

Newcastle University

**Creating a Story-Based Dynamic Assessment to Identify Developmental
Language Disorder in Children Learning English as an Additional
Language**

Teresa Garrido-Tamayo

This thesis is submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Speech and Language Sciences
School of Education, Communication and Language Sciences
Newcastle University

Newcastle upon Tyne

March 2024

Abstract

Background: Developmental Language Disorder (DLD) hinders the language acquisition of around 8% of UK children at school entry, including those with English as an Additional Language (EAL). Early diagnosis is essential for mitigating its impact on everyday learning and communication. However, detecting DLD in multilingual children is difficult due to the absence of suitable assessment tools and practitioners fluent in the children's home languages. This exacerbates the challenge of distinguishing between language issues stemming from DLD and from a limited familiarity with English. In this context, Dynamic Assessment (DA) emerges as a solution that integrates teaching with assessment to uncover a child's capacity to learn language, rather than their current language skills, reducing linguistic and cultural bias and catering to children with variable English abilities.

This work focuses on creating a DA designed to assist DLD detection in UK school-aged children with EAL from diverse home language backgrounds. Using storytelling in English, this DA targets learning potential across three DLD-vulnerable areas: narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and receptive affective prosody.

Methods: Following pilot studies, the DA was trialled with 14 children with EAL aged 4;06–8;11 years from Northeast England. Its effectiveness was tested by comparing children's DA performances (scores in receptive affective prosody and story generation tasks, and modifiability) relative to outcomes in measures relevant for DLD diagnosis: the *Crosslinguistic Nonword Repetition Test* (CL-NWRT), *New Reynell Developmental Language Scales* (NRDLS) in English, and variables concerning English experience, proficiency, and presence of DLD risk factors.

Results: In children's narratives, story grammar usage significantly improved over the DA's teaching phase, but episodic structure complexity and emotional vocabulary did not, nor did affective prosody understanding.

Correlational and predictive relationships between DA performances and NRDLS scores, along with English proficiency, highlight the need to refine the DA to assess learning potential irrespective of existing English language skills. Such relationships with the CL-NWRT and DLD risk factor scores support the DA's sensitivity to DLD-related vulnerabilities and value in identifying DLD among children with EAL.

Conclusions: Additional testing with a larger sample is essential to strengthening the findings, which call for refinement of the DA to better evaluate DLD risk across UK multilingual children with varied English language experience and proficiency.

Acknowledgements

Completing this PhD has been a journey of profound learning and growth, made possible by the UK's Economic and Social Research Council, the Northern Ireland and North East Doctoral Training Partnership, and countless individuals who have supported me through the years, a few of whom I owe special thanks and dedicate this thesis to:

The adults and children, along with their families and teachers, who took part in the studies. I deeply appreciate your trust and generosity with your time during such challenging years.

Carolyn Letts and Laurence White, my PhD supervisors. It has been a privilege to work under your mentorship. Thank you for your wisdom, patience, and kindness, and for encouraging me to think critically, stay curious, and not get too tied up by the old impostor syndrome feelings! Thank you also to Cristina McKean and Faye Smith for your perceptive advice and kind encouragement during the annual progress reviews, which were pivotal in evolving my research and lifting my confidence as a beginning researcher.

Antonio Bernal Benítez, a wonderful, caring mentor whose passion for science and research ignited my own; Isabel Gómez Veiga, whose insightful teaching and kind, attentive guidance played a crucial role in shaping my research path; and Judith Sullivan, who took my English skills to the next level and offered me the warmest friendship and support along the way.

The staff, colleagues, and friends in the HaSS Faculty and School of ECLS for bringing much-needed light and connection to my PhD experience, and Emily, Mary Pat, Louise, and Ciara for our treasured cross-continental chats on all things dynamic assessment and beyond.

Clara Garrido-Tamayo, Debbie Foster, Alyona Chufistova, Sarah Barnett, Eva Gilmour, Olivia Dearden, and Charlotte Dawson, for your valuable assistance with material development and reliability work for this thesis, and Sean Pert for your insights on engaging multilingual and multicultural families in research.

My beautiful (and understanding!) friends and family, whose love and support I simply could not have done this PhD without. Special thanks to the NUSU PGR Community and writing groups (especially Emnani, Anastassia, and Francesca within the latter), Lisa, Mily, Nan, John, and my wee buddies Agua, Kairi, and Polly. *Gracias infinitas* to my beloved parents, Manuel and Rosa, my sister Clara, my grandparents, Rita and Pepe, and my late grandmother, Estrella. Finally, Garry, there are just not enough words to express my gratitude! You are the coolest, most wonderful life companion anyone could ever wish for.

Table of Contents

Chapter 1. Introduction.....	1
1.1 Chapter Outline.....	1
1.2 Motivation Behind this Thesis: Assessment of Developmental Language Disorder in Multilingual Children	1
1.3 The Present Thesis: Aims and Rationale.....	3
1.4 Research Methodology Overview.....	4
1.5 Main Findings	6
1.6 Structure of the Thesis	7
Chapter 2. Setting the Stage: A Literature Review on English as an Additional Language, Developmental Language Disorder, and the Value of Dynamic Assessment.....	9
2.1 Introduction.....	9
2.2 English as an Additional Language in the United Kingdom.....	9
2.3 Developmental Language Disorder: Prevalence and Diagnostic Criteria	14
2.4 Impact of Developmental Language Disorder in Language and Beyond.....	17
2.4.1 <i>Narrative macrostructure skills: Story grammar and episodic structure complexity</i>	17
2.4.2 <i>Emotional vocabulary knowledge</i>	19
2.4.3 <i>Receptive affective prosody skills</i>	21
2.4.4 <i>Impact of developmental language disorder in other areas of life</i>	22
2.5 Language Assessment for Multilingual Children	23
2.5.1 <i>Barriers and potential solutions in the assessment of children with EAL</i>	26
2.5.2 <i>Dynamic assessment of language learning</i>	29
2.6 Dynamic Assessment in the Present Study: Narrative Macrostructure, Emotional Vocabulary, and Receptive Affective Prosody	39
Chapter 3. Development and Piloting of the Dynamic Assessment Methods.....	41
3.1 Introduction.....	41
3.2 Development of the Receptive Affective Prosody Task	44
3.2.1 <i>Pilot testing 1: Receptive Affective Prosody Task (Congruent vs Incongruent Prosody)</i>	45
3.2.2 <i>Pilot testing 2: Receptive Affective Prosody Task (Congruent vs Incongruent Prosody, Congruent vs Neutral Prosody, and Neutral Content)</i>	52
3.2.3 <i>Pilot testing 3: Receptive Affective Prosody Task (Neutral Content)</i>	58
3.2.4 <i>Conclusions</i>	59
3.3 Development of the Storytelling Activity	60

3.3.1 Materials.....	60
3.3.2 Instructions	67
3.3.3 Pilot testing 1: Storytelling activity	81
3.3.4 Pilot testing 2: Storytelling activity	87
3.3.5 Conclusions.....	89
3.4 Summary of Piloting Phase and Design Decisions for the Main Study	91
Chapter 4. Methods: Main Study of the Dynamic Assessment	93
4.1 Introduction	93
4.2 Recruitment Procedure	95
4.2.1 <i>Participant eligibility criteria</i>	97
4.3 Data Collection Procedure.....	98
4.3.1 <i>Comparative measures</i>	99
4.3.2 <i>Formulation of the English Language Experience Score and DLD Risk Factors Score</i>	102
4.4 Participants' Profiles.....	105
4.4.1 <i>Demographic characteristics</i>	106
4.4.2 <i>Language experience profiles</i>	108
4.4.3 <i>General and language development profiles</i>	111
4.5 Interrater Reliability for the Dynamic Assessment's Storytelling Activity	113
4.6 Implementation Fidelity for the Dynamic Assessment's Storytelling Activity.....	114
4.7 Data Analysis Procedure	117
Chapter 5. Results: Main Study of the Dynamic Assessment.....	119
5.1 Introduction	119
5.2 Performance Across the Sample in the Comparative Measures and the Dynamic Assessment	121
5.2.1 <i>NRDLS and CL-NWRT: Descriptives</i>	121
5.2.2 <i>DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage: Descriptives</i>	122
5.2.3 <i>Dynamic assessment's Receptive Affective Prosody Task: Descriptives and repeated measures t-test results</i>	123
5.2.4 <i>Dynamic assessment's Story Generation Task: Descriptives, repeated measures t-test, and Wilcoxon signed-rank test results</i>	124
5.2.5 <i>Dynamic assessment's Modifiability Rating Scale: Descriptives</i>	126
5.2.6 <i>Summary of the sample performance in the comparative measures and the dynamic assessment</i>	126
5.2.7 <i>Individual results in the dynamic assessment</i>	127

5.3 Correlational Analyses between the Dynamic Assessment Measures and the Comparative Measures	129
5.3.1 <i>Correlations between the Receptive Affective Prosody Task and the NRDLS and CL-NWRT</i>	130
5.3.2 <i>Correlations between the Story Generation Task and the NRDLS and CL-NWRT</i>	133
5.3.3 <i>Correlations between the Modifiability Rating Scale and the NRDLS and CL-NWRT</i>	137
5.3.4 <i>Correlations between the Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage</i>	140
5.3.5 <i>Correlations between the Story Generation Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage</i>	140
5.3.6 <i>Correlations between the Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage</i>	143
5.4 Multiple Regression Analyses between the Dynamic Assessment Measures and the Comparative Measures	143
5.4.1 <i>Receptive Affective Prosody Task and the NRDLS and CL-NWRT</i>	144
5.4.2 <i>Story Generation Task and the NRDLS and CL-NWRT</i>	145
5.4.3 <i>Modifiability Rating Scale and the NRDLS and CL-NWRT</i>	146
5.4.4 <i>Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage</i>	146
5.4.5 <i>Story Generation Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage</i>	146
5.4.6 <i>Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage</i>	148
5.5 Conclusions.....	148
Chapter 6. Discussion: Main Study of the Dynamic Assessment	153
6.1 Introduction.....	153
6.2 Teaching Phase Impact on Dynamic Assessment Performance: A Pretest–Post-test Comparison	154
6.3 Relationship between the Dynamic Assessment Measures and the Comparative Measures	157
6.3.1 <i>Dynamic assessment measures and NRDLS</i>	158
6.3.2 <i>Dynamic assessment measures and English Language Proficiency Stage</i>	160
6.3.3 <i>Dynamic assessment measures and English Language Experience Score</i>	161
6.3.4 <i>Dynamic assessment measures and CL-NWRT</i>	161
6.3.5 <i>Dynamic assessment measures and DLD Risk Factors Score</i>	162

6.3.6 <i>Conclusions on the dynamic assessment and comparative measures relationships</i>	163
6.4 Limitations and Future Research Directions	164
Chapter 7. Conclusions	169
References	173
Appendices	199
Appendix 1. Star Points Chart Used During the Dynamic Assessment	201
Appendix 2. Advertising Flyer Used in the Child Pilot Study 1	202
Appendix 3. Parental Information Sheet Used in the Child Pilot Study 1	203
Appendix 4. Parental Consent Form Used in the Child Pilot Study 1	206
Appendix 5. Child-Friendly Information Video Used in the Child Pilot Study 1	207
Appendix 6. Facilitator Checklist Used in the Child Pilot Study 1	208
Appendix 7. Extract from the Dynamic Assessment Protocol Used in the Child Pilot Study 1: Receptive Affective Prosody Task Script	210
Appendix 8. Scoring Sheet for the Receptive Affective Prosody Task Used in the Child Pilot Study 1	213
Appendix 9. Information Sheet Used in the Adult Pilot Study of the Receptive Affective Prosody Task	214
Appendix 10. Consent Form Used in the Adult Pilot Study of the Receptive Affective Prosody Task	216
Appendix 11. Adult Trial of the Audio Stimuli in the Congruent vs Neutral Prosody Version of the Receptive Affective Prosody Task	217
Appendix 12. Response Tables Used in the Adult Pilot Study of the Receptive Affective Prosody Task	219
Appendix 13. Receptive Affective Prosody Task Script Used in the Child Pilot Study 2 and Main Study	222
13A. Pretest Phase	222
13B. Post-test Phase	226
Appendix 14. Scoring Sheets for the Receptive Affective Prosody Task Used in the Child Pilot Study 2 and Main Study	229
Appendix 15. Paper Copies of the Graphic Organisers and Pictograms Used as Part of the Storytelling Activity in the Main Study	230
Appendix 16. Script for the Storytelling Activity Used in the Main Study	232
16A. Pretest Phase	232
16B. Teaching Phase	234
16C. Post-test Phase	257

Appendix 17. Version 1 of Scoring Sheet 1: Story Grammar Units for Stories A, B and C	258
Appendix 18. Version 2 of Scoring Sheet 1: Story Grammar Units for Stories A and B ...	261
Appendix 19. Version 1 of Scoring Sheet 3: Emotional Vocabulary for Stories A, B and C	265
Appendix 20. COVID-19 Risk Assessment for Working in Schools	266
Appendix 21. Information Sheet for Schools Used in the Main Study	271
Appendix 22. Parental Information Sheet Used in the Main Study.....	274
Appendix 23. Parental Consent Form Used in the Main Study.....	278
Appendix 24. Child-Friendly Information Sheet Used in the Main Study	279
Appendix 25. Advertising Flyers Used in the Main Study	281
Appendix 26. Lone Working Risk Assessment Used for the Main Study	282
Appendix 27. Certificate of Research Participation Used in the Main Study	285
Appendix 28. Parental Questionnaire: Development & Language Experience.....	286
Appendix 29. Teacher Questionnaire: Language Development & Experience in the School Context.....	305
Appendix 30. Observation Form Used in the Main Study	314
Appendix 31. Participation Guidance for Parents Provided in the Main Study	316
Appendix 32. Participation Guidance for Teachers Provided in the Main Study	317
Appendix 33. Pearson's and Spearman's Correlation Matrices between Scores in the Receptive Affective Prosody Task and the CL-NWRT and NRDLS	318
Appendix 34. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 1: Story Grammar Units and the CL-NWRT and NRDLS	320
Appendix 35. Spearman's Correlation Matrix between Scores in Scoring Sheet 2: Episodic Structure Complexity and the CL-NWRT and NRDLS	322
Appendix 36. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 3: Emotional Vocabulary and the CL-NWRT and NRDLS	323
Appendix 37. Pearson's and Spearman's Correlation Matrices between Scores in Modifiability Rating Scale and the CL-NWRT and NRDLS	325
Appendix 38. Pearson's and Spearman's Correlation Matrices between Scores in the Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage	326
Appendix 39. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 1: Story Grammar Units and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage	328
Appendix 40. Spearman's Correlation Matrix between Scores in Scoring Sheet 2: Episodic Structure Complexity and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage	330

Appendix 41. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 3: Emotional Vocabulary and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage 331

Appendix 42. Pearson's and Spearman's Correlation Matrices between Scores in Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage 333

List of Tables

Table 1. <i>Department for Education (DfE)’s Five-Point Scale of Reading, Writing, and Spoken Language Proficiency in English</i>	11
Table 2. <i>Chapter 3 Overview</i>	44
Table 3. <i>Stimuli Sentences for the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task</i>	48
Table 4. <i>Stimuli Sentences for the Neutral Content Version of the Receptive Affective Prosody Task</i>	54
Table 5. <i>Stimuli Sentences for the Congruent vs Neutral Prosody Version of the Receptive Affective Prosody Task</i>	55
Table 6. <i>Descriptive Statistics for the Performance in the Receptive Affective Prosody Tasks in the Adult Pilot Study</i>	57
Table 7. <i>Story A: Scoring Sheet 1. Story Grammar Units – Version 3</i>	72
Table 8. <i>Story B: Scoring Sheet 1. Story Grammar Units – Version 3</i>	74
Table 9. <i>Story A: Scoring Sheet 2. Episodic Structure Complexity</i>	76
Table 10. <i>Story B: Scoring Sheet 2. Episodic Structure Complexity</i>	77
Table 11. <i>Story A: Scoring Sheet 3. Emotional Vocabulary – Version 2</i>	77
Table 12. <i>Story B: Scoring Sheet 3. Emotional Vocabulary – Version 2</i>	78
Table 13. <i>Modifiability Rating Scale</i>	81
Table 14. <i>Descriptive Statistics for the Children’s Performance in Scoring Sheet 1: Story Grammar Units (Story Generation Task) in the Online Pilot Study</i>	84
Table 15. <i>Descriptive Statistics for the Children’s Performance in Scoring Sheet 2: Episodic Structure Complexity (Story Generation Task) in the Online Pilot Study</i>	85
Table 16. <i>Descriptive Statistics for the Children’s Performance in Scoring Sheet 3: Emotional Vocabulary (Story Generation Task) in the Online Pilot Study</i>	85
Table 17. <i>Descriptive Statistics for the Children’s Performance in the Modifiability Rating Scale in the Online Pilot Study</i>	86
Table 18. <i>Overview of Key Demographic, Language Experience, and Developmental Characteristics for Each Child</i>	105
Table 19. <i>Storytelling Activity Implementation Fidelity Monitoring Scale</i>	115
Table 20. <i>Rating Form for the Assessor’s Use of Mediated Learning Experience Strategies</i>	116
Table 21. <i>Descriptive Statistics for the Children’s Performance in the NRDLS and CL-NWRT</i>	121
Table 22. <i>Descriptive Statistics for the Children’s English Language Experience Score and DLD Risk Factors Score</i>	122
Table 23. <i>Descriptive Statistics for the Children’s English Language Proficiency Stage</i>	123
Table 24. <i>Descriptive Statistics for the Children’s Performance in the Receptive Affective Prosody Task</i>	124
Table 25. <i>Descriptive Statistics for the Children’s Performance in Scoring Sheet 1: Story Grammar Units (Story Generation Task)</i>	124
Table 26. <i>Descriptive Statistics for the Children’s Performance in Scoring Sheet 2: Episodic Structure Complexity (Story Generation Task)</i>	125

Table 27. <i>Descriptive Statistics for the Children's Performance in Scoring Sheet 3: Emotional Vocabulary (Story Generation Task)</i>	126
Table 28. <i>Descriptive Statistics for the Children's Performance in the Modifiability Rating Scale</i>	126
Table 29. <i>Overview of Individual Results in the Comparative Measures and the Dynamic Assessment</i>	128
Table 30. <i>Overview of Significant Correlations between the Children's Performance in the Dynamic Assessment Measures and Comparative Measures</i>	151
Table 31. <i>Overview of Comparative Measures Showing as Predictors for the Dynamic Assessment Measures in the Regression Analyses</i>	152

List of Figures

Figure 1. Classification System for the 'Speech, Language and Communication Needs' (SLCN) Label	16
Figure 2. Illustration of Vygotsky's Zone of Proximal Development (ZPD) Concept.....	31
Figure 3. Overview of the Dynamic Assessment Developed in the PhD Research Project	42
Figure 4. Timeline of the Pilot and Main Studies of the PhD Research Project	43
Figure 5. Examples from the Training Phase in the Receptive Affective Prosody Task.....	47
Figure 6. Examples from the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task	49
Figure 7. Total Score by Age in the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task in the Online Child Pilot Study	50
Figure 8. Total Score by Age in the Congruent Trials of the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task in the Online Child Pilot Study.....	50
Figure 9. Total Score by Age in the Incongruent Trials of the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task in the Online Child Pilot Study.....	51
Figure 10. Example of Scenario Illustration in the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task Before and After Removing Potentially Distracting Elements	51
Figure 11. Example from the Neutral Content Version of the Receptive Affective Prosody Task	54
Figure 12. Example from the Congruent vs Neutral Prosody Version of the Receptive Affective Prosody Task	55
Figure 13. Stimulus Picture Story A	62
Figure 14. Stimulus Picture Story B	63
Figure 15. Stimulus Picture Story C	64
Figure 16. Example of Picture 3 from Stimulus Story A Before Speech Bubble Was Removed	65
Figure 17. Jigsaw Puzzles Used During the Teaching Phase of the Storytelling Activity	67
Figure 18. Feelings Wheel Used During the Teaching Phase of the Storytelling Activity.....	67
Figure 19. Clustered Bar Chart of Frequencies of English Language Interactions Between Children and their Parents	109
Figure 20. Pie Chart of Frequencies of English Language Interactions between Children and their Siblings.....	110
Figure 21. Scatterplot of the Correlation between the Pretest Scores in the Receptive Affective Prosody Task and the NRDLS Comprehension Scale Scores	131
Figure 22. Scatterplot of the Correlation between the Pretest Scores in the Receptive Affective Prosody Task and the NRDLS Production Scale Scores	131
Figure 23. Scatterplot of the Correlation between the Pretest Scores in the Receptive Affective Prosody Task and the CL-NWRT Scores	132
Figure 24. Scatterplot of the Correlation between the Post-test Scores in the Receptive Affective Prosody Task and the NRDLS Production Scale Scores	132
Figure 25. Scatterplot of the Correlation between the Pretest Scores in Scoring Sheet 1: Story Grammar Units and the NRDLS Comprehension Scale Scores	134

Figure 26. Scatterplot of the Correlation between the Pretest Scores in Scoring Sheet 1: Story Grammar Units and the NRDLS Production Scale Scores	134
Figure 27. Scatterplot of the Correlation between the Post-test Scores in Scoring Sheet 1: Story Grammar Units and the NRDLS Production Scale Scores	135
Figure 28. Scatterplot of the Correlation between the Pretest Ratings in Scoring Sheet 2: Episodic Structure Complexity and the NRDLS Production Scale Score.....	136
Figure 29. Scatterplot of the Correlation between the Post-test Scores in Scoring Sheet 3: Emotional Vocabulary and the NRDLS Production Scale Scores	137
Figure 30. Scatterplot of the Correlation between the Modifiability Rating Scale Scores and the NRDLS Comprehension Scale Scores	138
Figure 31. Scatterplot of the Correlation between the Modifiability Rating Scale Scores and the NRDLS Production Scale Scores	139
Figure 32. Scatterplot of the Correlation between the Modifiability Rating Scale Scores and the CL-NWRT Scores	139
Figure 33. Scatterplot of the Correlation between the Change Scores in Scoring Sheet 1: Story Grammar Units and the English Language Proficiency Stage.....	141

Chapter 1. Introduction

1.1 Chapter Outline

This PhD thesis is introduced in this opening chapter by first presenting the background that inspired the research, its objectives and the rationale underpinning it (see Sections 1.2 and 1.3). This is followed by an overview of the methodology (Section 1.4) and key findings (Section 1.5) from the main study, along with a detailed outline of the structure of the thesis, covering each chapter (Section 1.6).

1.2 Motivation Behind this Thesis: Assessment of Developmental Language Disorder in Multilingual Children

Latest estimates indicate that over a million immigrants arrived in the United Kingdom (UK) in 2023 (ONS, 2023), continuing the country's long-standing history of immigration from around the globe (ONS, 2016). For the children of these families, whether they have recently arrived or settled in the country generations ago, multilingualism becomes a necessity, far beyond a mere educational choice: in addition to their home language(s), learning English as an Additional Language (EAL) is an essential aspect of their educational success and social integration. However, the prevalence of developmental language difficulties among these children is believed to be on par with that of monolingual children (Mennen & Stansfield, 2006), with approximately 7.6% of all children starting school in the UK likely to encounter persistent language learning difficulties that cannot be explained by other identifiable causes (e.g., hearing loss, intellectual disabilities, limited experience with English as the language used in school) (Bishop, Snowling, Thompson, Greenhalgh & the CATALISE-2 Consortium, 2017; Norbury et al., 2016). This condition, referred to as 'Developmental Language Disorder' (DLD), not only poses significant challenges to social and educational growth but, given that language is embedded in most facets of our daily functioning, it is also linked with difficulties in other domains, such as employment, and social and emotional wellbeing throughout the lifespan (Orrego, McGregor & Reyes, 2023). As a result, early and accurate detection of DLD is a vital first step to lessen its extensive effects and improve the prospects for children's success in school and life in general.

Assessment of DLD in children with EAL is conducted in a similar manner to that of monolingual children but with a notable addition: the need to account for the greater variability inherent to learning more than one language. Although multilingual language acquisition does not cause nor aggravate DLD, this variability affects each language's

developmental rate. As such, it is advised that assessments cover all of a child's languages (Pert & Bradley, 2018), acknowledging that even typically developing children with EAL may take time to catch up to their monolingual native peers regarding their English language skills (e.g., Paradis & Jia, 2017), and show overlap with monolinguals with DLD during the early stages of acquiring this language (e.g., Paradis, Rice, Crago & Marquis, 2008).

The vast array of home languages spoken in the UK, however, complicates assessments in languages other than English, as there is a shortage of appropriate multilingual testing resources for both English and the children's home language(s), as well as practitioners skilled in these languages, resulting in misallocated and delayed support services (Letts, 2012; Mennen & Stansfield, 2006; Oxley, Cattani, Chondrogianni, White & De Cat, 2019). Among the solutions proposed to navigate these challenges is the emphasis on alternative assessment methods that do not require the assessor to be proficient in the child's home language(s). Dynamic Assessment (DA) is one such approach, prioritising the language learning process over existing language knowledge by integrating intervention into the evaluation process. DAs have been recognised as effective for identifying DLD risk in multilingual children (Hunt, Nang, Meldrum & Armstrong, 2022; Orellana, Wada & Gillam, 2019), especially because they can be conducted in English, provided the child understands the language well enough to follow task instructions (Hunt et al., 2022). Moreover, DA mitigates the cultural, educational and life experiential biases that often disadvantage children in conventional static assessments, which only consider current language skills at the time of testing (Camilleri & Law, 2007).

A fundamental method of DA is 'pretest–teach–post-test', which begins by evaluating a child's ability in a particular language area, proceeds with child-centred teaching on that skill, and concludes with a reassessment of the targeted skill to gauge progress (e.g., Kapantzoglou, Restrepo & Thompson, 2012; Peña, Gillam & Bedore, 2014b; Petersen, Chanthongthip, Ukrainetz, Spencer & Steeve, 2017). This method measures the child's performance gains from pretest to post-test and their *modifiability* (response to teaching) to help clarify the root causes of any pretest difficulties. Whereas small post-test improvements and poor learning behaviours signal a risk of DLD, marked post-test gains and positive learning behaviours suggest that initial challenges stem from insufficient English language exposure, or cultural and life experiential differences affecting performance (Peña et al., 2014b).

1.3 The Present Thesis: Aims and Rationale

The objective of this PhD research was to design and evaluate a DA aimed at identifying DLD risk among primary school-aged children with EAL. Using a pretest–teach–post-test sequence centred around storytelling in English, such an aim would be achieved by observing how well children respond to the learning opportunities provided in narrative macrostructure (including story grammar and episodic structure complexity), emotional vocabulary, and receptive affective prosody—three domains which are susceptible to DLD. These domains were chosen for their diagnostic relevance and compatibility with a storytelling framework that supports accurate, meaningful testing.

Focusing on narrative macrostructure first, the DA explores children’s capacity to learn to structure stories in a coherent and complete manner, recognising storytelling as a natural activity worldwide that is helpful for assessing language development in linguistically and culturally diverse populations. This can be especially valuable when examining skills related to organising a story—such as story grammar, which refers to the essential units of information constituting a narrative (e.g., characters, problem), and episodic structure complexity, referring to the production of complete stories—, as these skills transcend language-specific knowledge, like grammar or vocabulary, and are more easily shared across languages. Consequently, typically developing children with EAL reach age-appropriate norms in narrative macrostructure skills faster than in areas reliant on English syntax and lexical knowledge (Govindarajan & Paradis, 2019). Such findings suggest that testing narrative macrostructure offers a fairer evaluation for children with EAL still developing their English language skills. This advantage, however, does not extend to multilingual children with DLD, who generally exhibit poorer performance in narrative macrostructure than their typically developing peers with EAL (e.g., Govindarajan & Paradis, 2019; Squires et al., 2014).

The DA also evaluates children’s ability to enhance their vocabulary usage, focusing on terms that describe the story characters’ emotional states, as this is an area of difficulty associated with DLD (e.g., Boerma, Leseman, Timmermeister, Wijnen & Blom, 2016), in line with more general emotion recognition deficits observed in this disorder (e.g., Löytölä, Ohtonen, Laakso & Huttunen, 2020; Taylor, Maybery, Grayndler & Whitehouse, 2015). By targeting emotion-related vocabulary, the DA also taps into word learning skills often impaired in children with DLD (e.g., Hasson, Camilleri, Jones, Smith & Dodd, 2012a; Jackson, Leitão, Claessen & Boyes, 2019).

Further, the DA addresses children's ability to process affective prosody, a key element in enhancing the emotional content of stories (Reilly, 2001). While the understanding of affective prosody improves with age (Ma, Zhou & Thompson, 2022), research indicates that children with DLD struggle to identify and interpret emotional states in voices (e.g., Fujiki, Spackman, Brinton & Illig, 2008; Löytömäki et al., 2020; Taylor et al., 2015).

Supporting multilingual children at DLD risk to maximise their communication skills, wellbeing, and social and educational participation begins with efficient and effective assessment. The DA developed in this study addresses the need for appropriate tools to recognise and equitably meet these children's unique language needs within the UK's resource-constrained health, educational and social care system. Despite extensive evidence of the validity of DA approaches for informing the diagnosis of DLD in multilingual children, their adoption by practitioners remains limited (Oxley et al., 2019; Newbury, Bartoszewicz Poole & Theys, 2020), and there is a pronounced need for further research into DA that enhances the sparse DA resources available, especially for older, school-aged children with EAL (Oxley et al., 2019) and those from a wide array of linguistic and cultural backgrounds (Hunt et al., 2022). The DA in this study is tailored specifically for school-aged multilingual English-speaking children who speak a variety of home languages, mirroring the diverse language learning backgrounds and English proficiency levels observed in educational and speech and language therapy contexts in the UK (Mennen & Stansfield, 2006; Oxley et al., 2019). Overall, this research adheres to Newcastle University's School of Education, Communication and Language Sciences' aim to improve children's life chances, its wider commitment to social justice, and the UK's Economic and Social Research Council health, wellbeing, and social care priorities.

1.4 Research Methodology Overview

After receiving ethical approval from the Faculty of Humanities and Social Sciences Ethics Committee at Newcastle University, the practicability of the DA was initially explored through a series of pilot studies, culminating in a main study with 14 multilingual children aged 4;06–8;11, recruited from local families and schools.

The recruitment process faced significant obstacles related to engaging a population often difficult to access, compounded by complications from the COVID-19 pandemic, which led to a shift from the originally envisioned two-group design —aimed at comparing DA results between children with and without possible DLD— to a single-group approach, wherein children displayed a range of English proficiencies. Instead, this revised approach examined

the relationships between the children's performance in the DA and various comparative measures intended for distinguishing DLD from the need for increased English language exposure. These measures included the *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015; Chiat, Polišenská, Yanushevskaya & Antonijevic, 2020), and the *New Reynell Developmental Language Scales* (NRDLS; Edwards, Letts & Sinka, 2011) in English, combined with scores of the children's *English Language Proficiency Stage*, *English Language Experience*, and *DLD Risk Factors* obtained from parental and teacher reports, and observations of their communicative interactions within home and school settings. While the analyses of the DA's relationships with the *English Language Experience Score*, *English Language Proficiency Stage*, and NRDLS aimed to illustrate the possible influence of children's current English language experience and skills on the resource's effectiveness, the DA's links with the CL-NWRT and *DLD Risk Factors Score* investigated whether it is capable of identifying DLD risk among multilingual children. The study was guided by the two Research Questions (RQs) below, each accompanied by a sub-question.

- RQ1: What is the relationship between the participating children's performance in the DA and their scores in the NRDLS and the CL-NWRT?
 - Sub-RQ1: Which elements of the DA show the strongest relationship with the NRDLS and CL-NWRT?
- RQ2: What is the relationship between the participating children's performance in the DA and their *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*?
 - Sub-RQ2: Which elements of the DA show the strongest relationship with the children's *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*?

The data collection procedure was organised into two individual sessions lasting 45 to 60 minutes each, held in quiet spaces within the children's schools, with two exceptions made for home settings when school access was not possible. The first session involved the administration of the NRDLS and CL-NWRT. Around this same period, parental and teacher questionnaires were administered, and observations of the children's communicative interactions in their familiar environments were conducted. The second session introduced the DA, which was built around the child becoming a 'storytelling superstar'. This was divided into three phases:

- 1) A pretest phase, during which children's performance was tested through a *Receptive Affective Prosody Task (Feelings Game 1)* and a *Story Generation Task*, marking the beginning of the storytelling activity (*Storytime*).
- 2) A teaching phase within the storytelling activity, where visual aids (e.g., pictograms) and the pretest story were applied to instruct children on narrative macrostructure, emotional vocabulary, and affective prosody, integrating these aspects into the story context.
- 3) A post-test phase, where parallel forms of the pretest tasks were carried out to reassess children's performance, including a second *Story Generation Task* using a new story and another *Receptive Affective Prosody Task (Feelings Game 2)*.

The following DA measures were obtained: pretest, post-test, and change scores in the *Receptive Affective Prosody Task* and *Story Generation Task*, along with ratings on an ad hoc *Modifiability Rating Scale*, which assessed children's capacity to respond to instruction during the teaching phase.

For reliability and fidelity purposes, a second evaluator, blind to the story presentation order and the children's developmental profiles, scored a subset of participants' performances in the *Story Generation Tasks*. They also rated children's modifiability and the researcher's adherence to the storytelling activity guidelines.

1.5 Main Findings

The main study suggested that participation in the DA's teaching phase led to enhancements in children's story grammar usage without similar substantial gains in episodic structure complexity, emotional vocabulary, or understanding of affective prosody. Furthermore, the potential capacity of the DA to identify DLD risk among multilingual children was suggested by the correlational and predictive relationships between the children's DA results and their outcomes in the CL-NWRT and *DLD Risk Factors Score*. Nonetheless, this value of the DA is compromised by its reliance on children's English proficiency, as indicated by its relationships with the NRDLS and *English Language Proficiency Stage*, despite no evident links with the *English Language Experience Score*.

These findings point to the need to test the DA with a larger sample and further refine the resource's stimuli materials, scoring systems, and teaching strategies. Such efforts would aim to surpass the limitations found in conventional static tests, facilitating a more accurate

understanding of multilingual children's learning capabilities without bias towards their existing English language skills.

1.6 Structure of the Thesis

The thesis is organised into seven chapters. Following *Chapter 1*, which introduces the thesis, *Chapter 2* contextualises the creation and testing of the DA developed in this study within previous research, delving into key aspects of the EAL child population in the UK and discussing the prevalence, diagnostic criteria, and impact of DLD on language and other domains, with an emphasis on the target areas of the DA: narrative macrostructure, emotional vocabulary, and receptive affective prosody. This chapter also addresses the challenges of assessing multilingual children, presenting DA approaches as one viable solution.

Chapters 3 and 4 are dedicated to the development and testing of the DA in this research. Specifically, *Chapter 3* describes this resource's creation and pilot testing, considering the *Receptive Affective Prosody Task* and the storytelling activity, which includes the *Story Generation Task*. *Chapter 4* details the participant recruitment process and eligibility criteria for the main study, as well as the participating children's demographic, language experience, and general and language development characteristics. It also presents the data collection methods, including the comparative measures used, the formulation of the *English Language Experience Score* and *DLD Risk Factors Score*, and the study's interrater reliability, implementation fidelity, and data analysis procedures.

Chapter 5 presents the findings from the main study, starting with descriptive statistics to summarise the sample's performance across the DA and comparative measures. It also shows the results on the effectiveness of the DA's teaching phase for boosting pretest-to-post-test performance across all participants, and the correlational and predictive relationships between the DA outcomes and comparative measures, addressing the RQs and Sub-RQs.

Chapter 6 interprets the findings from the main study, focusing on the utility of the DA's teaching phase and the relationships observed between the DA and comparative measures. The implications of these findings are discussed for the DA's purpose of assessing DLD risk among the UK's multilingual children, reflecting on the study's limitations and suggesting future research directions.

Finally, *Chapter 7* concludes the thesis, reiterating its contribution to the research field of DA with DLD diagnostic purposes in multilingual child populations and its particular potential within the UK context.

Chapter 2. Setting the Stage: A Literature Review on English as an Additional Language, Developmental Language Disorder, and the Value of Dynamic Assessment

2.1 Introduction

The present chapter offers a literature review to contextualise the creation and testing of a Dynamic Assessment (DA) in this PhD study. This resource was designed to identify possible Developmental Language Disorder (DLD) risk in school-aged children who have English as an Additional Language (EAL) in the UK. Recognising these children's unique language needs and DLD when it is present is crucial for providing the necessary support, enabling them to reach their full potential and improve their future outcomes.

The chapter begins with an exploration of key aspects concerning the EAL child population in the UK context (Section 2.2), followed by an overview of DLD, detailing its prevalence and diagnostic criteria (Section 2.3). Subsequent sections examine the impact of DLD on language skills and other domains (Section 2.4), emphasising the specific areas targeted by the proposed DA resource: narrative macrostructure, emotional vocabulary, and receptive affective prosody skills (Sections 2.4.1 to 2.4.3). The challenges inherent in assessing multilingual children and potential solutions are then addressed (Section 2.5), setting the stage for the introduction of DA as a favourable and reliable approach for this demographic. The chapter concludes by integrating these themes with the present study, focusing on the application of DA through the context of storytelling to estimate DLD risk based on multilingual children's learning potential (Section 2.6).

2.2 English as an Additional Language in the United Kingdom

In our increasingly interconnected world, multilingualism (including bilingualism) has emerged as the most common approach to language learning, with over half the world's population estimated to speak two or more languages in their everyday lives (Grosjean, 2013). This global trend is notably evident in the UK, where over 1.7 million pupils in England and Wales alone are recorded as having 'English as an Additional Language' (EAL), including 22% of children in state-funded primary schools (DfE, 2023). Defined by the DfE (2023), EAL includes any child 'exposed to a language at home that is known or believed to be other than English', which could be any of the more than 88 home languages spoken in the UK (ONS, 2013), including Polish, Romanian, Panjabi, Urdu, Portuguese, Spanish, Arabic, Bengali, Gujarati, and Italian in the latest top ten (ONS, 2022).

While the EAL label might help in flagging children with potential needs for extra language support, it falls short in capturing the actual diverse experiences and developmental trajectories in English language acquisition that characterise the path to multilingualism for these children. For example, some children experience simultaneous multilingualism as they grow up exposed to multiple languages from birth or shortly after (Kohnert, 2010). This is common among second or third-generation ethnic minority children in the UK, who are fluent in English and engage with other languages at home or within their community as part of their cultural heritage (Strand, Malmberg & Hall, 2015). In such cases, families may embrace a multilingual strategy in the home environment, as well as spaces such as churches and complementary language schools (e.g., Kallis & Yarwood, 2022; Lam, Chaudry, Pinder & Sura, 2020), where both English and their home language(s) are incorporated, including code-switched varieties (Pert & Letts, 2006). This approach contributes to the formation of children's ethnic and cultural identities and their sense of belonging. It also fosters communication among family members, enabling intergenerational interaction (e.g., Kallis & Yarwood, 2022; Lam et al., 2020). Conversely, sequential multilingualism occurs when children learn an additional language(s) after making substantial progress in a first language, usually after age 3, through exposure in educational or community settings (Kohnert, 2010). This pattern is observed in recent migrant children in the UK who arrive with varying literacy skills in their first language(s) and are exposed to English in school for the first time (Strand et al., 2015). It also applies to children of recent migrants who are born in the UK and receive limited exposure to English before entering nursery.

As a result of these varied experiences, moving beyond the basic EAL label becomes essential to understanding children's unique English language needs and backgrounds and identifying the support they require across the curriculum (Hessel & Strand, 2021; Strand & Hessel, 2018). To facilitate such assessment within a common framework, between 2016–2017, the English Schools Census asked schools to make a 'best fit' judgement and categorise children with EAL by their English proficiency level on a 5-point scale, from 'New to English' (A) to 'Fluent' (E) (Table 1) (DfE, 2017). However, this requirement was withdrawn in 2018, a move that has been widely criticised and led to calls for its reintroduction (Hessel & Strand, 2021; NALDIC, 2018; Scott, 2021; Strand & Hessel, 2018).

Table 1. Department for Education (DfE)'s Five-Point Scale of Reading, Writing, and Spoken Language Proficiency in English

Proficiency stage	Description
A. New to English	<ul style="list-style-type: none"> May use first language for learning and other purposes. May remain completely silent in the classroom. May be copying/repeating some words or phrases. May understand some everyday expressions in English but may have minimal or no literacy in English. Needs a considerable amount of EAL support.
B. Early Acquisition	<ul style="list-style-type: none"> May follow day-to-day social communication in English and participate in learning activities with support. Beginning to use spoken English for social purposes. May understand simple instructions and can follow narrative/accounts with visual support. May have developed some skills in reading and writing. May have become familiar with some subject specific vocabulary. Still needs a significant amount of EAL support to access the curriculum.
C. Developing Competence	<ul style="list-style-type: none"> May participate in learning activities with increasing independence. Able to express self orally in English, but structural inaccuracies are still apparent. Literacy will require ongoing support, particularly for understanding text and writing. May be able to follow abstract concepts and more complex written English. Requires ongoing EAL support to access the curriculum fully.
D. Competent	<ul style="list-style-type: none"> Oral English will be developing well, enabling successful engagement in activities across the curriculum. Can read and understand a wide variety of texts. Written English may lack complexity and contain occasional evidence of errors in structure. Needs some support to access subtle nuances of meaning, to refine English usage, and to develop abstract vocabulary. Needs some/occasional EAL support to access complex curriculum material and tasks.
E. Fluent	<ul style="list-style-type: none"> Can operate across the curriculum to a level of competence equivalent to that of a pupil who uses English as his/her first language. Operates without EAL support across the curriculum.

Note. Alongside the above scale, 'Not Yet Assessed' (Code 'N') was available for use where the school had not yet had time to assess a child's proficiency in English.

Source: DfE (2017).

Research linking English proficiency to school attainment further reinforces the need for assessing English language ability rather than just EAL status (Demie, 2018; Hessel & Strand, 2021; Strand et al., 2015; Strand & Hessel, 2018; Strand & Lindorff, 2020; Whiteside, Gooch & Norbury, 2016). Studies show that children with EAL at the lower levels —A ('New to English') and B ('Early Acquisition')— achieve below the national average. In contrast, those at an intermediate level C ('Developing Competence') perform close to the national average, while those at the top end of the scale (levels D, 'Competent', and E, 'Fluent') attain above the national average, outperforming their monolingual English-speaking peers (Strand & Hessel, 2018). This disparity indicates that while multilingualism can enhance educational outcomes, insufficient proficiency in English (the primary language of instruction in UK schools) may pose a challenge to the likelihood of succeeding academically, with the timing of assessment playing a decisive role here; the earlier the child receives targeted support to gain fluency in English, the greater their potential to catch up (Hessel & Strand, 2021; Strand & Hessel, 2018; Whiteside et al., 2016). Accordingly, relying solely on binary EAL/non-EAL labelling is inadequate for determining individual children's attainment in school and the type of support they require, as the high performance of learners who are already fluent in English could mask the struggles of less proficient ones (Demie, 2018).

Aside from English language proficiency, risk factors for low academic results in children with EAL should also be considered, such as arriving in the English school system during a key stage, belonging to certain ethnic groups (e.g., Black-African, White-Other), and pupil mobility between schools (Strand et al., 2015). Factors shared with monolingual English-speaking peers, like experiencing family and neighbourhood socioeconomic deprivation, being male, being young for the year group, and having an identified Special Educational Need (SEN), should also be accounted for (Strand et al., 2015). In particular, the presence of a SEN has been identified as the most substantial risk factor for educational attainment among children with EAL (Strand et al., 2015). While this emphasises the need for awareness that having EAL is not a SEN, just as English-speaking monolingual children, children with EAL may also have additional educational needs and, as such, require both EAL and SEN support. This highlights the importance of differentiating between multilingual children who experience transient language deficits in English due to limited exposure to this language and those with more persistent issues due to genuine language impairment, such as DLD, which would require a referral to speech and language therapy services for further assessment and diagnosing (for more details, see Section 2.5).

Alongside these complexities in assessing English language proficiency, identifying multilingual pupils in need of language support has become particularly demanding given the adverse effects of the COVID-19 pandemic. Significant challenges were amplified in this area, with the impact being especially severe for children in the early stages of English language learning (Demie, Hay, Bellsham-Revell & Gay, 2022). In addition to children performing behind normal expectations, a clear pattern of English language loss has been observed, stemming from limited opportunities for these children to hear, speak, read, and write in this language during school closures (Demie et al., 2022; Scott, 2021).

Concerns about the effect of the pandemic on multilingual children's social and emotional wellbeing have also been raised. This includes issues like reduced motivation to learn due to language barriers, a lack of confidence in their ability to interact in English with their peers or in class, and feelings of loneliness and isolation (Demie et al., 2022). In line with this, evaluating, and consequently supporting, English language proficiency is crucial not only for boosting the academic performance of multilingual children, but also for improving their social, emotional, and behavioural functioning. Research indicates that greater English language skills predict fewer difficulties in these areas, with children who have EAL often showing better outcomes compared to their monolingual peers with similar English proficiency levels (Halle, Hair, Wandner, McNamara & Chien, 2012; Whiteside et al., 2016). These advantages associated with having experience in more than one language could be related to the enhanced cognitive functioning sometimes observed in multilingual children, including those with DLD (e.g., nonverbal processing speed: Ebert, 2021; visuospatial working memory, selective attention, and interference suppression: Engel de Abreu, Cruz-Santos & Puglisi, 2014; orienting attentional network: Park et al., 2019).

Ultimately, multilingual children in the UK represent a highly heterogeneous group with a wide range of English proficiency levels, from new to the language to fluent. The limitations of the EAL classification system highlight the need for school staff to have a more detailed assessment of English language skills that allows for effective, tailored support, benefitting both children's educational results and their social, emotional, and behavioural development. This will also be key for pinpointing cases of possible DLD versus temporary language difficulties in pupils who are not progressing at the expected rate. The following sections, 2.3, 2.4, and 2.5, delve deeper into DLD and its assessment in the context of multilingual children, respectively.

2.3 Developmental Language Disorder: Prevalence and Diagnostic Criteria

For the vast majority of children, the acquisition of language is an essential part of development that unfolds with relative ease, regardless of whether they are monolingual or multilingual. Still, about 7.6% of children in the UK experience issues with this on school entry, equivalent to two in every Year 1 classroom of 30 (Norbury et al., 2016). These language learning difficulties interfere with their daily functioning and cannot be explained by another condition, like hearing loss, or circumstances, such as reduced experience with the ambient language. While children with EAL were excluded from Norbury et al. (2016)'s population study due to the unfeasibility of assessing language ability in English and the 64 home languages represented in the cohort, prevalence rates are not expected to differ for them (Mennen & Stansfield, 2006). Inevitably, multilingual children are just as likely to have significant and unexplained language difficulties as their monolingual English-speaking peers.

Despite their widespread presence, the varied language learning problems children may face in the absence of a known cause have led to the use of inconsistent diagnostic terminology. Throughout time, this inconsistency has caused confusion in research and clinical practice, inequity over access to services, and limited recognition and understanding by scientists, practitioners, and the general public (Bishop et al., 2017; Ebbels, 2014). Recognising the need for consensus in diagnosing such language difficulties, the CATALISE panel —a multidisciplinary consortium of experts from English-speaking countries with backgrounds in speech and language therapy, education, psychology, paediatrics, and child psychiatry— came together in 2015–16 to conduct two studies using the Delphi method, in which they anonymously rated and discussed statements about the criteria and terminology for children's language difficulties (Bishop, Snowling, Thompson, Greenhalgh & the CATALISE Consortium, 2016; Bishop et al., 2017). The first study involved 59 experts, though two declined to participate in the second, leaving 57 for that phase (Bishop et al., 2016, 2017).

The CATALISE panel's recommendations on terminology proposed 'Language Disorder' as a term for children whose language difficulties significantly hinder everyday communication and learning and persist beyond the age of 5 years into adolescence and adulthood (Bishop et al., 2017). If such language difficulties cannot be explained by the presence of a primary biomedical condition, in which they would be part of a complex pattern of impairments (e.g., brain injury, acquired epileptic aphasia in childhood, certain neurodegenerative conditions, cerebral palsy, sensorineural hearing loss, genetic conditions like Down syndrome, autism or intellectual disability), the term 'Developmental Language Disorder' (DLD) should be used

instead. ‘Language Disorder associated with X’, where X is the name of one of the above differentiating biomedical conditions, is recommended otherwise (Bishop et al., 2017). While the exact cause of DLD remains unclear, rather than having a single identifiable origin, it is thought to be a multifactorial neurodevelopmental disorder resulting from the complex interaction of multiple biological, genetic, and environmental risk factors (Bishop, 2017; Mountford, Braden, Newbury & Morgan, 2022; Reilly et al., 2010).

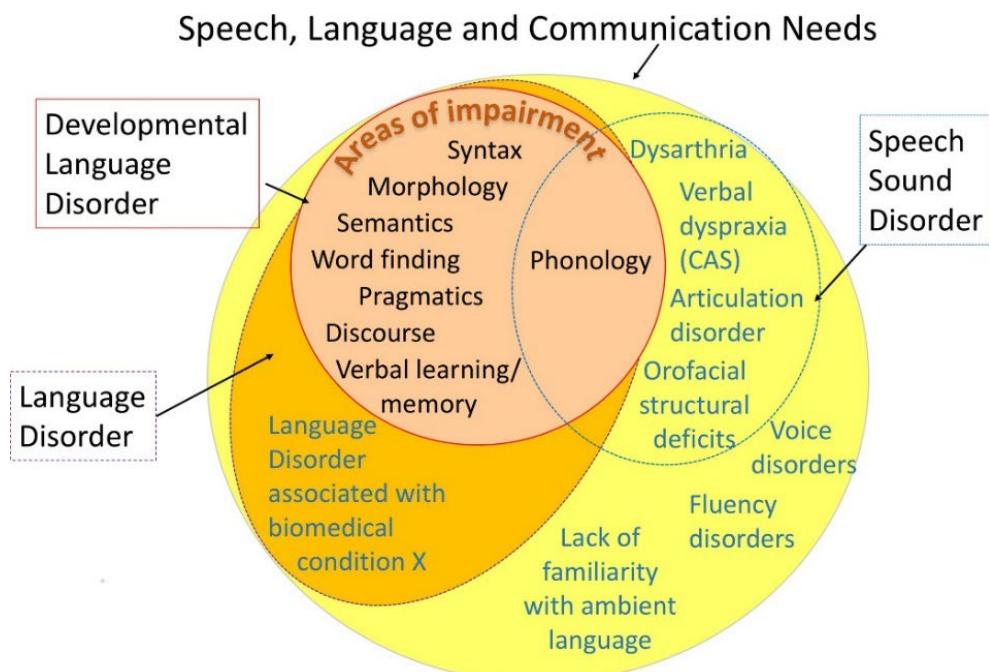
Additional recommendations from the CATALISE panel (Bishop et al., 2016, 2017) involved defining DLD more precisely, and the presence of the following does not preclude a diagnosis of DLD:

- Biological or environmental risk factors that have a partial or unclear causal relationship to language problems, but which are statistically more common in children with language disorders than typically developing children. These include lower 5-minute Apgar scores at birth, lower maternal education levels, lower socioeconomic status, a family history of speech, language, learning or literacy difficulties, being male, or being a younger sibling in a large family (Bishop et al., 2017; Reilly et al., 2010; Rudolph, 2017).
- Co-occurring neurodevelopmental disorders involving impairments in cognitive, sensorimotor, or behavioural domains which do not have a clear aetiology nor a causal relation to the language problems, such as attention deficit hyperactivity disorder, developmental co-ordination disorder, developmental dyslexia, speech difficulties, limitations of adaptive behaviour, and behavioural and emotional disorders (Bishop et al., 2017).
- A lack of discrepancy between nonverbal and verbal ability. Children with low nonverbal ability who do not meet the criteria for intellectual disability (i.e., a non-verbal IQ score below 70, limited adaptive functioning; APA, 2013; RCSLT, 2020a) can be diagnosed with DLD. Otherwise, the diagnosis would be ‘Language Disorder associated with Intellectual Disability’ (Bishop et al., 2017).

In light of these clarifications, the CATALISE panel also considered integrating both ‘Language Disorder’ and its subtypes —‘Developmental Language Disorder’ and ‘Language Disorder associated with X (biomedical condition)’— under the broad category ‘Speech, Language and Communication Needs’ (SLCN). This term is adopted by educational services in the UK and, as illustrated in Figure 1, covers a wide range of conditions affecting speech, language and communication, irrespective of their specific nature or cause, including children

with needs due to unfamiliarity with English, the language used in school (Bishop et al., 2017). The ‘SLCN’ categorisation aids in informing policy and service planning, as well as in highlighting issues in cases of preschool children for whom a DLD diagnosis may not be immediately feasible due to unclear prognoses before the age of 5 (Bishop, 2017; Bishop et al., 2017). Regardless of a formal diagnosis, there is a recognised need for extra language support at this stage, which must be addressed. While the accuracy of prognosis improves with age, it is essential to identify young children whose language development lags behind that of their peers, whether they eventually catch up or not. Special attention should be paid to any risk factors predictive of longer-term, persisting language difficulties, such as poor language comprehension, poor use of gestures to communicate, socio-economic disadvantage and/or a family history of language impairment (Bishop et al., 2016; RCSLT, 2020a).

Figure 1. Classification System for the ‘Speech, Language and Communication Needs’ (SLCN) Label



Note. Source: Bishop et al. (2017).

Following the suggested terminology outlined above, this PhD thesis uses ‘Developmental Language Disorder’ (DLD) to refer to both the participants of this research and those in earlier studies, including instances where older terms were employed to describe the same group of children with primary language difficulties (e.g., specific language impairment, language learning impairment, developmental dysphasia). This approach aligns with the shift towards a unified DLD terminology that enhances clarity and awareness of this condition.

2.4 Impact of Developmental Language Disorder in Language and Beyond

DLD interferes with learning, understanding, and using language, presenting challenges for classification into predictable subtypes (e.g., Conti-Ramsden & Botting, 1999) due to the numerous ways in which it can present itself. Acknowledging this heterogeneity, the CATALISE panel (Bishop et al., 2017) suggested focusing instead on identifying the main areas of functioning where a child shows impairment. Broadly speaking, these may include phonology (e.g., Aguilar-Mediavilla, Buil-Legaz & Sánchez-Azanza, 2020), syntax and morphology (e.g., Leonard, 2014), semantics (e.g., Jackson et al., 2019), pragmatics (e.g., Andreou, Lymeropoulou & Aslanoglou, 2022), discourse/narrative (e.g., Govindarajan & Paradis, 2019), verbal learning and memory difficulties (e.g., Leonard, 2014).

Three DLD-affected skills are of particular relevance to the dynamic assessment resource developed in this PhD study for detecting DLD risk in school-aged children with EAL, each aligning with the general areas outlined by the CATALISE panel. These skills are narrative macrostructure, with a focus on story grammar and episodic structure complexity (Section 2.4.1), which falls under discourse; emotional vocabulary (Section 2.4.2), pertaining to semantics; and receptive affective prosody (Section 2.4.3), associated with pragmatics. Further explanations of these in the context of this study will be provided in *Chapter 3*.

The selection of these three areas of focus was driven by both theoretical rationales based on their prospective diagnostic relevance (see Sections 2.4.1, 2.4.2, and 2.4.3), and practical factors based on their ease of integration into a storytelling framework. Storytelling underpins the dynamic assessment resource due to its engaging, inclusive quality across all ages, languages, and cultures (Gagarina et al., 2012; Spencer & Petersen, 2020), facilitating an enjoyable, natural, and meaningful testing experience that could promote optimal performance and diagnostic accuracy. Measuring children's grammar learning abilities was also considered—given the prevalence of deficits in this area associated with DLD (Leonard, 2014)—but ultimately avoided. This was due to the difficulty of targeting and capturing learning potential in this highly language-specific area within the brief, storytelling-centred teaching phase in English as part of the dynamic assessment.

2.4.1 Narrative macrostructure skills: Story grammar and episodic structure complexity

While microstructure refers to a story's word- and sentence-level components, such as lexical diversity and syntax complexity, macrostructure denotes the global, underlying organisation of a narrative, which is similar across languages and essential for constructing a logically

coherent plot. In both mono- and multilingual contexts, children with DLD often struggle with microstructure elements (e.g., monolingual: Norbury, Gemmell & Paul, 2014; multilingual: Altman, Armon-Lotem, Fichman & Walters, 2016; Rezzonico et al., 2015; Squires et al., 2014; Tsimpli, Peristeri & Andreou, 2016) and are generally less adept in macrostructure skills than their typically developing peers (e.g., monolingual: Norbury et al., 2014; multilingual: Boerma et al., 2016; Govindarajan & Paradis, 2019; Rezzonico et al., 2015; Squires et al., 2014).

Some evidence, however, suggests that macrostructure abilities may not consistently distinguish between typically developing children and those with DLD (Altman et al., 2016; Tsimpli et al., 2016). Govindarajan and Paradis (2019) attribute these conflicting findings to methodological differences, including the types of narrative tasks employed (story generation versus story retelling). Story generation tasks, which require children to produce their own stories, often with pictorial support, may differentiate between typical and atypical language development better, as they are more demanding than story retelling tasks, where children simply retell a story they have just heard. Additionally, these mixed results could be influenced by multilingual proficiency. Like children with DLD, typically developing children may also find microstructure components challenging in the early stages of learning English as an additional language, as these depend on language-specific skills like syntax and lexical knowledge of English. In contrast, their macrostructure skills are less affected because they reside at the cognitive-linguistic interface and can be shared between languages. This makes the ability to organise a narrative a less biased and more reliable indicator of narrative development at this stage (Govindarajan & Paradis, 2019; Squires et al., 2014). For this reason, the present study, which involves the participation of UK-based multilingual children with varying English language skills, targets macrostructure using a story generation task, as explained in *Chapter 3*.

In the dynamic assessment resource developed in this study, narrative macrostructure is assessed through the presence of story grammar units and the complexity of the episodic structure in the children's stories. Relevant to these measurements are Stein and Glenn (1979)'s *Story Grammar Model* and Westby (2005)'s *Story Grammar Decision Tree*, consistent with the approaches of two extensively used tools for studying narrative macrostructure in mono- and multilingual child populations: the *Edmonton Narrative Norms Instrument* (ENNI; Schneider, Dubé & Hayward, 2005) and the *Multilingual Assessment Instrument for Narratives* (MAIN, Gagarina et al., 2012, 2019).

Stein and Glenn (1979)'s *Story Grammar Model*, commonly employed for evaluating children's narrative macrostructure (e.g., Gagarina et al., 2012, 2019; Schneider et al., 2005), outlines that stories typically start with a setting introducing characters, place, and time, followed by at least one episode. This episode focuses on an initiating event or problem, and the characters' feelings, plans, and attempts to resolve this problem, culminating in the resolution of these efforts. Children's use of story grammar units increases from ages 3 to 9, reflecting their growing awareness of story organisation and coherent plot formation as a function of cognitive maturity (Khan et al., 2016; Schneider, Hayward & Dubé, 2006).

Westby (2005)'s *Story Grammar Decision Tree*, based on Stein and Glenn's *Story Grammar Model*, is applied to categorise stories into various levels depending on the complexity of the episode's structure. A story is an 'Isolated Description' if it lacks a temporally related sequence of events. Instead, it is considered a sequence when it has a temporally related sequence of events without causal links ('Action Sequence') or a causally related sequence of events but no goal-directed behaviour ('Reactive Sequence'). If goal-directed behaviour is implied but the planning of this behaviour is not clear through an attempt to achieve the goal or an outcome of such attempt, the story is an 'Abbreviated Episode'. The story becomes a 'Complete Episode' when it includes a plan, an attempt, and an outcome that makes goal-directed behaviour explicit. Finally, a story can be 'Elaborated' if it includes several attempts or outcomes, multiple sequential episodes or embedded episodes, or is told from various character perspectives. As children age, their narratives evolve from simple descriptions and series of events to more detailed episodic organisations involving goals, attempts to reach these goals, and outcomes from these attempts (Westby, 2005).

2.4.2 Emotional vocabulary knowledge

Story grammar units that relate to the characters' internal feelings in response to events are among the latest elements to emerge in childhood (Brinton, Fujiki & Asai, 2019; Govindarajan & Paradis, 2022; Khan et al., 2016) and can be particularly challenging for children with DLD when compared to their typically developing peers (monolingual: Brinton et al., 2019; Norbury et al., 2014; multilingual (Dutch–minority home language): Boerma et al., 2016).

Interestingly, the difficulties that children with DLD have in this area could also be greater than those experienced by autistic children, despite the expectation that the latter would particularly struggle with this aspect, in line with the autism-related vulnerabilities in Theory of Mind (e.g., Siller, Swanson, Serlin & Teachworth, 2014). Indeed, when Norbury et al.

(2014) contrasted the narratives of monolingual English-speaking children aged 6;06–15;00 with DLD (n = 23), autism (n = 25), and typical development (n = 27) using the wordless picture book *A Boy, a Dog and a Frog* (Mayer, 1967), they found that those with DLD produced fewer instances of language describing the characters' internal states, including emotions (e.g., 'The frog was lonely'), than both their typically developing and autistic peers. This remained the case despite the DLD and autism participant groups being selected to be as contrastive as possible, acknowledging the debated phenotypic overlap between DLD and autism (i.e., the children with DLD did not present clinically significant social-pragmatic impairments, and the autistic children did not have substantial structural language deficits on standardised tests). According to Norbury et al. (2014), such findings suggest that children with DLD lack the words to discuss the characters' thoughts and feelings.

Importantly, however, a more recent study by Govindarajan and Paradis (2022), published after the main data collection phase of this PhD research had started, presented somewhat contrasting evidence to previous studies considered (e.g., Boerma et al., 2016; Norbury et al., 2014). In their study, 29 multilingual English-speaking children aged 5;04–9;01 from diverse home language backgrounds participated. These children were either typically developing (n = 10) or clinically diagnosed with autism (n = 9) or DLD (n = 10), with diagnoses corroborated by the *Alberta Language Development Questionnaire* (ALDeQ; Paradis, Emmerzael & Sorenson Duncan, 2010). When comparing the narratives of the three groups using the six stories from the *Edmonton Narrative Norms Instrument* (ENNI; Schneider et al., 2005), they found that children with DLD used more internal state terms to describe characters' reactions to story outcomes (e.g., 'happy') than autistic children, with their performance in this aspect similar to that of their typically developing peers. At the same time, no group differences emerged in the production of characters' internal responses to the initiating events that set off the stories (e.g., 'mad'), although all groups used very few of these, reflecting developmental trends where these become more common in older children's narratives. Overall, these findings highlighted a need for studies with larger samples of multilingual participants with DLD to better understand the value of story grammar units related to internal states, including those pertinent to emotions, in distinguishing DLD from typical development and autism in multilingual populations.

So far, acknowledging the mixed evidence and the requirement for further research, aside from the limited vocabulary and linguistic capacity for articulating one's own and others' emotional states (Brinton et al., 2019; Norbury et al., 2014), any such difficulties would be consistent with broader DLD-related deficits in emotion processing and recognition (Bahn,

Vesker, Schwarzer & Kauschke, 2021; Griffiths, Goh, Norbury & the SCALES team, 2020; Löytömäki et al., 2020; Taylor et al., 2015), as well as in areas central to effective emotion perception, including Theory of Mind (Löytömäki et al., 2020; Nilsson & Jensen de López, 2016) and working memory (Löytömäki et al., 2020).

With the above in consideration, the dynamic assessment resource created in this study targets emotional vocabulary, which is integral to the emotional content of story plots and linked to successful emotion recognition (Brinton et al., 2019; Streubel, Gunzenhauser, Grosse & Saalbach, 2020). This is done by assessing children's capacity to make use of basic emotional terms (e.g., 'happy', 'angry') relevant to the protagonists' emotions in their stories. By evaluating and teaching these words describing emotions, the dynamic assessment also taps into children's ability to learn new words, as those with DLD often experience difficulty acquiring vocabulary, regardless of whether they are monolingual or multilingual (Camilleri & Law, 2007; Hasson et al., 2012a; Jackson et al., 2019; Kapantzoglou et al., 2012).

2.4.3 Receptive affective prosody skills

Affective prosody refers to the modulation of speech intonation, rhythm, and timing to express emotional states, which is similar across languages (e.g., sadness is associated with a slow speaking rate and low pitch, while happiness tends to be conveyed with a faster rate and higher pitch; Pell, Paulmann, Dara, Alasseri & Kotz, 2009a). This plays a key role in narratives, enriching the emotional content beyond the literal meaning of words (Reilly, 2001), and in social-emotional competence (Griffiths et al., 2020; Löytömäki, Laakso & Huttunen, 2023).

The recognition of emotions, an area affected by DLD as noted in Section 2.4.2, extends to challenges in this area. A growing body of research in monolingual populations indicates that children with DLD have difficulty interpreting emotional prosody in voices (Boucher, Lewis & Collis, 2000; Courtright & Courtright, 1983; Creusere, Alt & Plante, 2004; Fujiki et al., 2008; Griffiths et al., 2020; Löytömäki et al., 2020; Taylor et al., 2015; Trauner, Ballantyne, Chase & Tallal, 1993). This is similar in extent to autistic children (Boucher et al., 2000; Löytömäki et al., 2020; Taylor et al., 2015). Notably, even though three out of the 13 children with DLD in Löytömäki et al. (2020)'s study had a comorbid diagnosis of autism, and 30 out of the 97 children who met the criteria for DLD in Griffiths et al. (2020)'s study presented additional diagnoses, including an unspecified number with autism, both studies provide evidence that children with DLD alone have substantial difficulties in recognising emotions in the voice.

In contrast, typically developing children, whether monolingual or multilingual, demonstrate a steady improvement in their development of emotional prosody comprehension with age. This begins early in childhood, as evidenced by Ma et al. (2022)'s recent study, where they examined how 3–5-year-old American and Chinese children identified four emotions (happiness, sadness, surprise, anger) that were expressed through speech prosody in English, Chinese, French, and Spanish. Children's ability to interpret emotional prosody was consistent across both native and unfamiliar languages, suggesting the presence of cross-linguistic prosodic decoding ability, and that the enhanced sensitivity to emotional prosody in one's native language seen in adults (Pell et al., 2009a) may not be as evident at these ages.

Building on this understanding, and acknowledging that the discriminative value of receptive affective prosody in detecting DLD is still under exploration, especially with research needed involving multilingual populations, this study's dynamic assessment resource also targets receptive affective prosody by exploring children's capacity to interpret basic emotional states (e.g., sad, scared) conveyed through prosody.

2.4.4 Impact of developmental language disorder in other areas of life

Children with DLD face challenges that extend beyond language difficulties, significantly impacting other areas of their lives. For instance, the emotion recognition difficulties noted in Sections 2.4.2 and 2.4.3 place these children at long-term heightened risk for poor social, emotional, and behavioural functioning (Conti-Ramsden, Mok, Pickles & Durkin, 2013; Creusere et al., 2004; Löytömäki et al., 2020; Löytömäki et al., 2023; Norbury et al., 2016; Taylor et al., 2015). Such struggles contribute to a lower quality of life, regardless of the severity of the language difficulties (Eadie et al., 2018), and are linked to an increased vulnerability to experiencing peer problems —such as rejection and victimisation (Van den Bedem, Dockrell, van Alphen, Kalicharan & Rieffe, 2016), limited friendships and social networks (Chen, Justice, Rhoad-Drogalis, Lin & Sawyer, 2020; Löytömäki et al., 2023)—, as well as mental health issues like anxiety, depression, and low self-esteem (Botting, Toseeb, Pickles, Durkin & Conti-Ramsden, 2016; van den Bedem et al., 2018).

The repercussions of DLD are also present in the educational and employment spheres. DLD results in lower literacy levels and failure to meet academic expectations (Aguilar-Mediavilla, Buil-Legaz, López-Penadés, Sánchez-Azanza & Adrover-Roig, 2019; Conti-Ramsden, Durkin, Toseeb, Botting & Pickles, 2018; Norbury et al., 2016), along with leaving education substantially earlier than typically developing individuals (Conti-Ramsden et al., 2018). Later

in life, adults with DLD are more likely to be in non-professional occupations and to be unemployed for longer (Conti-Ramsden et al., 2018).

These wide-ranging impacts highlight the necessity for early identification of children at risk for DLD to ensure they receive timely and appropriate intervention. Acknowledging this priority, the dynamic assessment resource developed through this study is intended to complement the existing tools available for DLD diagnosis. Accurate detection of DLD is essential in providing children with the support they require for their language needs, helping to lessen the profound and multifaceted effects of DLD across their lifespan (Orrego et al., 2023). In this endeavour, speech and language therapists, working alongside families and other health and education professionals, play a vital role in enhancing a child's communication abilities, social participation, wellbeing, and educational access (Ebbels, McCartney, Slonims, Dockrell & Norbury, 2019; Orrego et al., 2023).

2.5 Language Assessment for Multilingual Children

Early language development rates are highly variable across children, owing to the interplay of factors specific to the child (i.e., child-internal factors) and features of the environment in which they are developing (i.e., child-external factors) (Reilly et al., 2010; Ukoumunne et al., 2012). This can complicate the assessment and diagnosis of language disorders, such as DLD, with an additional layer of complexity arising when children grow up exposed to more than one language, whether simultaneously or sequentially. The language learning context of these children is more intricate than that of monolingual children, with differences observed in the quantity and quality of language input and output for each language over time. Thus, while the assessment of DLD in children with EAL resembles that for monolingual children, there is a further need to account for the greater variability experienced in multiple language learning (Boerma & Blom, 2017; Paradis & Jia, 2017).

Importantly, multilingualism neither causes nor exacerbates DLD. In fact, contrary to common, unfounded concerns that children exposed to more than one language may become confused, it is recognised as offering linguistic, social, and possible cognitive advantages, regardless of the presence of DLD (Pert & Bradley, 2018). Multilingualism also plays a critical role in children's wellbeing, sense of identity, and relationships with their families and communities (Kohnert, 2010; Müller, Howard, Wilson, Gibson & Katsos, 2020). However, the diverse learning experiences inherent to a multilingual upbringing do impact children's skills and developmental rates in each language, and even those whose language is developing as expected often show higher dominance in at least one language (Kohnert, 2010). It has

been observed that children with EAL take 4–7 years of English exposure to converge with monolingual norms for their English language abilities. This timespan varies depending on the linguistic subdomain considered, the task difficulty, and individual factors, both child-internal (e.g., verbal short-term memory, vocabulary size) and external (e.g., amount and richness of language exposure) (Demie, 2013; Paradis & Jia, 2017; Paradis & Kirova, 2014; Paradis, Tulpar & Arppe 2016).

Furthermore, at the early stages of acquiring English, certain language profile features of typically developing multilingual children can overlap with those of monolingual children with DLD of the same age, such as difficulties with grammatical morphology (Paradis, 2005; Paradis et al., 2008). This overlap, coupled with the fact that DLD results in expressive and/or receptive difficulties in the home language(s) too, not just English, highlights the importance of gathering information on the child's development in all their languages for an evidence-based diagnosis (Pert & Bradley, 2018). A multilingual child is less likely to have DLD if their difficulties appear only in English (the new language being learned) and they show expected progress in their home language(s). This would indicate that the necessary mechanisms for language acquisition are in place, and they will readily acquire the new language with increased exposure in school. In contrast, a child with DLD would require specialist intervention, as their skills in both English and the home language(s) fall considerably below those of their typically developing peers from similar cultural and linguistic backgrounds (Pert & Bradley, 2018; Roseberry-McKibbin, 2021).

Determining whether a child has DLD or a temporary language difficulty due to language experience factors is relatively straightforward in more homogenous multilingual populations. For example, the large minority of Spanish-English speakers in the United States (US) provides a context where suitable assessment tools are available (e.g., Peña, Gutiérrez-Clellen, Iglesias, Goldstein & Bedore, 2014a). The UK, however, presents a unique challenge with over 88 home languages spoken other than English (ONS, 2013). The scarcity of tests with multilingual norms for every language combination limits the provision of assessment in languages other than English, complicating the task of discerning whether a child's slow progress in English is a result of limited exposure or is symptomatic of DLD (Oxley et al., 2019).

In this context, a survey study by Oxley et al. (2019) investigated the assessment practices of children with EAL in the UK. A total of 140 practitioners involved in language disorder identification and decision-making about children with EAL, mainly speech and language

therapists and teachers, were asked about the tools and procedures that they use with these children, their opinions and satisfaction levels regarding these tools, as well as the barriers they face (Oxley et al., 2019). The study revealed that, in line with official recommendations for when standardised assessments are unavailable (Pert & Bradley, 2018), most practitioners primarily rely on informal measures, such as observations of the child's communicative interactions and information gathered from parents and other professionals (Oxley et al., 2019).

Indeed, indirect assessment of the child's language development through teacher (Bedore, Peña, Joyner & Macken, 2011; Pua, Lee & Rickard Liow, 2017) and parental questionnaires or interviews (Bedore et al., 2011; Boerma & Blom, 2017; Li'el, Williams & Kane, 2019; Paradis et al., 2010) is vital for informing of DLD presence. In particular, parents and key caregivers such as grandparents can provide valuable insights into children's general and language development. This includes identifying indicators of DLD relevant across languages and cultures in both mono- and multilingual learning contexts, such as the child presenting difficulties in their home language(s) and experiencing late-onset of early language milestones like the first word and multi-word utterances (Boerma & Blom, 2017; Grimm & Schulz, 2014; Roseberry-McKibbin, 2021; Tuller, 2015). It is also essential to gather data about the quantity and quality of the child's language usage and exposure in their environment. Factors such as how long they have been exposed to a language, how much they hear and use it every day, and how often they engage in language-rich activities involving this language, such as reading, telling stories or watching films, are central to their developmental rates and proficiency in both English and their home language(s) (Paradis, 2011).

Notably, the use of formal tests standardised on monolingual children is discouraged. This is because monolingual norms in either English or the home language(s) are not applicable to children with EAL, especially since they may be less dominant in the language of testing due to reduced language-specific experience and thus risk being over-diagnosed with DLD (De Lambo White & Jin, 2011; Letts, 2012). Despite this, nearly half of the speech and language therapists in Oxley et al.'s (2019) study indicated using standardised tests in English to assess spoken language. Popular choices included the *Clinical Evaluation of Language Fundamentals* (CELF-IV; Semel, Wiig & Secord, 2003), *British Picture Vocabulary Scale* (BPVS; Dunn, Dunn, Whetton & Burley, 1997; Dunn, Dunn, Styles & Sewell, 2009), *Test of Abstract Language Comprehension* (TALC; Elks, McLachlan & Blank, 2012), *Test for Reception of Grammar* (TROG-2; Bishop, 2003), *Children's Communication Checklist* (CCC; Bishop, 1998, 2003), and *Renfrew Action Picture Test* (RAPT; Renfrew, 2003). With

the exception of the BPVS II (Dunn et al., 1997), which includes multilingual norms, the latest edition (BPVS III; Dunn et al., 2009), along with all other mentioned tests, does not provide norms for children with EAL.

This suggests a gap between practitioners' positive attitudes toward multilingualism and their actual practices. Although they believe in children's potential to learn multiple languages and support assessment in both English and the home language(s), as recommended by official guidelines and research, the reality often diverges due to restricted access to the training and resources (e.g., funding and time for adequate assessment, employment of interpreters) required for implementing best practice (Oxley et al., 2019), in line with other studies in largely English-speaking countries (Marinova-Todd et al., 2016; Newbury et al., 2020; Williams & McLeod, 2012). These challenges associated with disentangling a language difference from a language impairment place culturally and linguistically diverse children at high risk for misdiagnosis. Children may either be over-diagnosed with DLD or under-diagnosed due to their DLD-derived difficulties being wrongly attributed to limited language exposure, leading to inefficient resource use or the delayed provision of vital intervention (Grimm & Schulz, 2014; Oxley et al., 2019; Mennen & Stansfield, 2006). Possible detrimental effects on children's self-worth could also be a consequence of over-diagnosis, stemming from the improper association of multilingualism with a SEN (Winter, 1999).

2.5.1 Barriers and potential solutions in the assessment of children with EAL

Among the major barriers to the accurate assessment of DLD in children with EAL, professionals report a lack of practitioners competent in the children's home languages who can incorporate these languages into the assessment, the scarcity of multilingual assessments standardised for both English and the children's home languages, limited access to interpreters—especially those trained in speech and language therapy practices—, and little knowledge of developmental norms in the children's home languages (Letts, 2012; Mennen & Stansfield, 2006; Oxley et al., 2019). Other obstacles include practitioners' lack of cultural knowledge and information about multilingual development (Mennen & Stansfield, 2006; Oxley et al., 2019).

To address these barriers and the general lack of confidence that practitioners feel in their ability to assess multilingual children (Newbury et al., 2020; Oxley et al., 2019; Williams & McLeod, 2012), the provision of targeted guidance to support children with EAL is stressed, as well as additional, specialised training along with initial and continued professional development focused on typical multilingual language development (Newbury et al., 2020;

Oxley et al., 2019; Parry, 2020). For instance, a clear understanding of common processes and phenomena associated with multilingual language use is critical to avoid misjudging diversity for a disorder (De Lambo White & Jin, 2011). These may include interference or transfer between languages, codeswitching, and language loss or attrition. Sometimes, there is also a ‘silent period’ in the early stages of acquiring a new language, where the focus is more on listening and understanding than speaking (Bligh, 2014; Roberts, 2014; Siraj-Blatchford & Clarke, 2000).

Other suggestions for overcoming assessment barriers concern the assessment methods used, with enhancing the availability of standardised tests in the appropriate languages as the most commonly suggested solution (Newbury et al., 2020; Oxley et al., 2019). A few assessments standardised on multilingual children for whom one of the languages is English already exist in the UK. These include the *Sandwell Bilingual Screening Assessment Scales for Expressive Punjabi and English* for Punjabi-English-speaking children (Duncan, Gibbs, Noor & Whittaker, 1988), the vocabulary test *Prawf Geirfa Cymraeg* for Welsh-English-speaking children (Gathercole, Thomas & Hughes, 2008), the *Rochdale Assessment of Mirpuri Phonology* (RAMP; Stow & Pert, 1998), and the *Bilingual Assessment of Simple Sentences* (BASS; Pert & Stow, 2019) for children who speak a Pakistani heritage language (Mirpuri, Punjabi or Urdu) in combination with English. Similarly, in the US, several tests have been published for Spanish-English-speaking children, such as the *Bilingual English-Spanish Assessment* (BESA; Peña et al., 2014a), the Spanish version of the *Clinical Evaluation of Language Fundamentals* (CELF-IV Spanish; Semel, Wiig & Secord, 2006), and the bilingual versions of the *Expressive and Receptive One-Word Picture Vocabulary Tests* (EOWPVT-SBE and ROWPVT-SBE; Martin & Brownell, 2012a, 2012b). Furthermore, some multilingual standardised assessments extend beyond specific language combinations to include a broader variety of home languages. These are the *Bilingual Verbal Ability Tests* for children learning American English and one of 17 minority home languages (BVAT; Muñoz-Sandoval, Cummins, Alvarado & Ruef, 2005) and the *UK Bilingual Toddlers Assessment Tool* (UKBTAT) for 2-year-old toddlers learning British English along with any other additional language (Floccia et al., 2018).

When standardised multilingual tests for the required languages and age groups are unavailable, the informal, supplementary use of tests normed for monolingual children may be helpful, as long as the influence of the child’s formal education experience, cultural background, and familiarity with the test content and procedures is considered (De Lambo White & Jin, 2011; Peña & Quinn, 1997). These tests can provide directional, qualitative

insights, enabling comparison of a child's performance with that of other child speakers of the language for whom there are no concerns, as well as with broader knowledge about language acquisition stages (Letts, 2012). In this context, the *New Reynell Developmental Language Scales* (NRDLS; Edwards et al., 2011) include a *Multilingual Toolkit* (Letts & Sinka, 2011) with guidance on how to adapt and apply the scales to the home languages of children with EAL, taking into account their linguistic and cultural context.

Another solution for assessment barriers is collecting and analysing language samples as an additional source of information. In this case, the assistance from interpreters becomes key to correctly gauge the child's home language(s) skills (Ebert, 2020; Pieretti & Roseberry McKibbin, 2016). Narrative elicitation tools can be valuable for eliciting narrative samples, such as the *Multilingual Assessment Instrument for Narratives* (MAIN; Gagarina et al., 2012, 2019), a resource that allows for comparisons in over 90 languages for different aspects of narrative ability. The MAIN is part of a battery of assessment tools titled *Language Impairment Testing in Multilingual Settings* (LITMUS), which was developed under COST Action IS0804 (2009–2013) 'Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment', a European initiative focused on advancing the identification of multilingual children with DLD.

Additionally, given the unfeasibility of establishing norms for every language combination, alternative approaches for determining the presence of DLD in multilingual children have also been recommended as part of a comprehensive assessment repertoire. These methods, suitable for administration by English-speaking practitioners, are intended to rely more on underlying language-learning mechanisms and less on children's language-specific knowledge and experience. Among these, language-processing measures have been explored that show potential utility as clinical markers of DLD in both monolingual and multilingual children, including non-word repetition (e.g., Antonijevic-Elliott et al., 2019; Boerma & Blom, 2017; Chiat, 2015; Li'el et al., 2019; Ortiz, 2021; Schwob et al., 2021) and sentence repetition tasks (e.g., Li'el et al., 2019; Marinis & Armon-Lotem, 2015).

Of particular interest for multilingual populations is the *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015) within the LITMUS test battery mentioned above. Research has consistently shown the value of non-word repetition tasks for distinguishing DLD from typical development in monolingual populations across many languages, showing significant typically developing/DLD group differences in performance (e.g., English: Graf Estes, Evans & Else-Quest, 2007; Italian: Dispaldro, Leonard & Deevy, 2013; Vietnamese: Pham & Ebert,

2020; Arabic: Taha, Stojanovik & Pagnamenta, 2021). However, the evidence for these tasks' discriminatory potential is less consistent for multilingual children, with the language specificity of the tasks introducing bias in this regard (Ortiz, 2021; Schwob et al., 2021). Non-word repetition tasks where items have language-specific features make children's scores more dependent on their experience with the language in which the task was constructed, with children finding it easier to repeat non-words that were created based on their own language's phonological characteristics. Bearing in mind that this type of task disadvantages multilingual children whose exposure to the task language is limited (Farabolini, Taboh, Ceravolo & Guerra, 2023; Ortiz, 2021; Schwob et al., 2021), the CL-NWRT was designed to be compatible with diverse phonological systems. This aimed to allow children to use knowledge from any language learned and thereby boost diagnostic accuracy by maximising the gap between the performance of those with and without DLD. The CL-NWRT has been proposed as a suitable tool for distinguishing DLD from typical development in multilingual children (e.g., English–minority home language: Antonijevic-Elliott et al., 2019; Dutch–minority home language: Boerma et al., 2015, Boerma & Blom, 2017), although further research is needed given that evidence from recent studies questions its diagnostic utility, as the performance of typically developing children can still overlap with that of those with DLD despite the task being language-independent (Swedish–Arabic: Öberg & Bohnacker, 2022; Swedish–Turkish: Öberg & Bohnacker, 2024).

Similar to language-processing tasks, as discussed next in Section 2.5.2, dynamic assessments that measure a child's ability to learn language and are less dependent on pre-existing knowledge and experience have also been suggested as a valuable approach to assisting in the diagnosis of DLD in children with EAL (De Lamo White & Jin, 2011; Hasson & Joffe, 2007).

2.5.2 Dynamic assessment of language learning

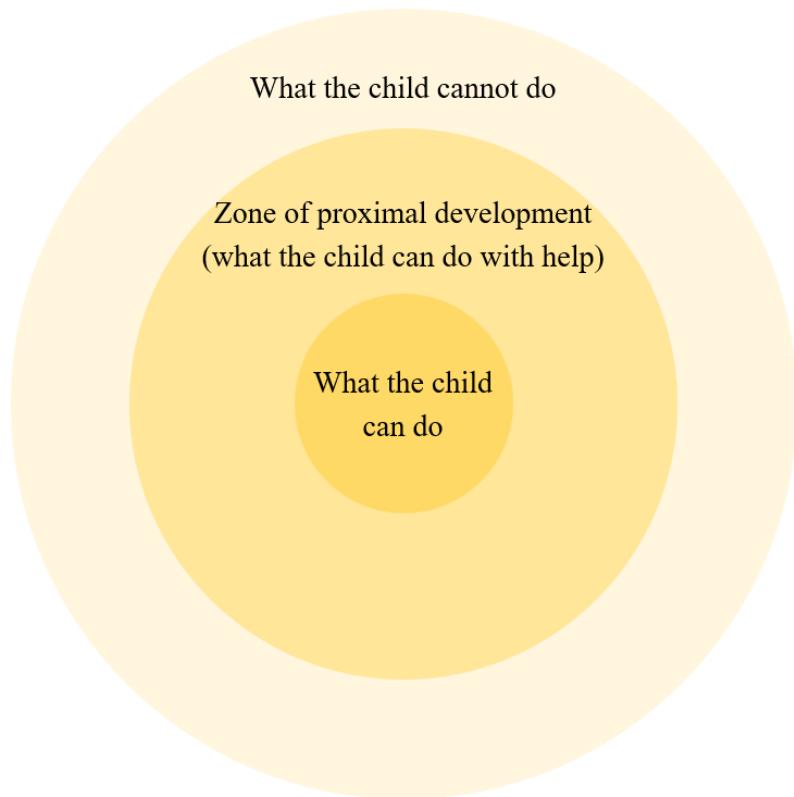
Introduction: The dynamic assessment approach

Dynamic Assessment (DA), also known as *interactive assessment* or *learning potential assessment*, is an umbrella term for a range of methods that integrate intervention into the assessment process (De Lamo White & Jin, 2011; Hasson & Joffe, 2007). Unlike conventional, so-called 'static' assessments —whether standardised or informal procedures— which provide a one-time 'snapshot' of a child's performance on a specific task, DA evaluates both the child's existing skills and, most crucially, their ability to learn with support, including the nature of the support required (De Lamo White & Jin, 2011; Hasson & Joffe, 2007).

The emphasis of DA on learning potential implies a focus on the process of learning language itself, rather than on the specific language being learned, such as English. This aspect makes DA particularly suitable for assessing children with EAL from diverse home language backgrounds for two main reasons. Firstly, DA can be administered in the practitioner's language (e.g., English), assuming the child has enough proficiency in that language to understand the task instructions (Hunt et al., 2022; Petersen et al., 2017). Secondly, DA helps to reduce the cultural and linguistic biases often present in static tests, where poor performance might be influenced by factors other than or in addition to DLD, like educational and cultural background (Camilleri & Law, 2007).

The theoretical foundations of DA trace back to Vygotsky's sociocultural theory and the *Zone of Proximal Development* (ZPD) concept (Vygotsky, 1978), as well as Feuerstein's theory of *Mediated Learning Experience* (Feuerstein, 1980). According to Vygotsky, a child's cognitive development occurs due to social interactions with more capable individuals, with learning taking place within the ZPD. The ZPD lies between what the child can achieve independently (actual developmental level) and what they can attain when assisted by an adult or a more experienced peer (potential development level) (Vygotsky, 1978) (see Figure 2). DA approaches aim to assess the 'size' of the ZPD; that is, to determine the extent of change or growth in a child's performance that can be induced through interactions with the assessor during the DA process (Peña, Iglesias & Lidz, 2001).

Figure 2. Illustration of Vygotsky's Zone of Proximal Development (ZPD) Concept



Feuerstein referred to this adult-supported learning within the ZPD as 'mediation' or a 'mediated learning experience' (Peña et al., 2001). In such experiences, the 'mediator', a more experienced person, interposes themselves between the stimuli in the environment and the child, using several mediating strategies (Lidz, 1991; Peña et al., 2001), including mediation of:

- *Intentionality*, to convey the goal and purpose of the learning activity for active and reciprocal engagement by the child;
- *Meaning*, to highlight the relevance of the activity and its benefits;
- *Transcendence*, to relate the activity to experiences in the child's life beyond the DA session; and
- *Competence*, to direct the child's attention to their acquired knowledge and how it can be applied, and to highlight their accomplishments with the aim of fostering self-confidence.

Using dynamic assessment of language learning with diagnostic aims

Research on the use of DA in the evaluation of child language disorders applies the concepts described above in ‘Introduction: The dynamic assessment approach’ (Section 2.5.2) to identify DLD in culturally and linguistically diverse children on the basis of their ability to benefit from intervention provided as part of the assessment. Various DA methods addressing learning potential across different language domains have been adopted, often in combination, such as *graduated prompting* (Camilleri & Law, 2007; Petersen, Tonn, Spencer & Foster, 2020), *testing the limits/clinical interview* (Gutiérrez-Clellen & Peña, 2001; Hasson, Dodd & Botting, 2012b; Peña, 2001) and, predominantly, *pretest–teach–post-test* (e.g., Peña et al., 2014b; Petersen et al., 2017).

In the US, extensive research has been conducted by Peña and colleagues using pretest–teach–post-test DA procedures to examine the ability of culturally and linguistically diverse children, mainly African-American and Latino-American, to acquire vocabulary (Gutiérrez-Clellen & Peña, 2001; Lidz & Peña, 1996; Peña et al., 2001; Peña, Quinn & Iglesias, 1992) and narrative skills (Peña et al., 2006; Peña et al., 2014b; Peña, Resendiz & Gillam, 2007). These DA procedures involved three phases: initial testing of a specific language skill, child-centred mediated teaching of that skill, and a concluding targeted skill test. Assessing the child’s learning capacity through their *modifiability*, or response to instruction, and their performance changes from pretest to post-test, contributes to understanding the nature of any initial language struggles. Specifically, when positive learning behaviours and noticeable post-test improvements were shown, pretest difficulties could be explained by limited English exposure and/or life experience and cultural differences leading to poor performance in static tests. In contrast, less optimal learning behaviours, minimal post-test improvements, and the need for more intensive assessor effort and input could be signs of DLD (Peña et al., 2014b).

For instance, Peña et al. (2014b) assessed the accuracy of a DA focused on narrative ability in English in identifying DLD among 54 Spanish-speaking preschoolers with EAL. This sample included 18 children with DLD, 18 typically developing control children matched for age, sex, IQ, and language experience, and 18 typically developing children matched only for age and language experience. The DA was conducted in three sessions, totalling about 70 minutes over a 7- to 14-day period. The first session consisted of the pretest and a first 30-minute intervention, while the second and third sessions included a second 30-minute intervention and the post-test, respectively. Following the DA procedure from Peña et al. (2006)’s study, the scripted intervention phases used a mediated learning approach to increase the length and

complexity of the children's narratives. During the pre- and post-tests, the children told stories based on two wordless picture books. The results showed that a combination of modifiability ratings (specifically compliance, metacognition, and task orientation), post-test narrative scores (setting, knowledge of dialogue, and complexity of vocabulary), and proportion of ungrammatical utterances derived from the post-test narrative sample could classify children with 80.6% accuracy when the results of the discriminant analysis were cross-validated with the age and language experience-matched typically developing group. These could also classify children with 97.2% accuracy when doing so with the age-, sex-, language experience-, and IQ- matched typically developing group. There were no substantial differences in the narrative ability and modifiability of the two typically developing groups (Peña et al., 2014b).

Other researchers have also found evidence supporting the suitability of DA of narrative ability as a diagnostic approach with multilingual children (Henderson, Restrepo & Aiken, 2018; Kramer, Mallett, Schneider & Hayward, 2009; Petersen et al., 2017). For example, to address some of the practical challenges preventing practitioners from using DA (e.g., excessive length of training, administration and scoring time, lack of validated cut-off points to indicate typical development versus DLD), Petersen et al. (2017) investigated the classification accuracy of a narrative DA measure for detecting DLD in 6;04–9;06-year-old Latino-American children, incorporating a few modifications which made it more practical to implement.

In Petersen et al. (2017)'s study, a narrative retell sample in English and Spanish was collected on the first day to determine language dominance and DLD status. This was followed by the administration of the DA over the next two days, involving a 25-minute pretest–teach–post-test session on each day. Each DA session was conducted in English and included a pretest narrative retell, a narrative retell teaching phase in which assessors cycled one to four times through a set of structured steps targeting story grammar units and adverbial subordinate clauses, and a post-test narrative retell. The pre- and post-test narrative retells were scored in real-time during the sessions, and a modifiability rating form was completed immediately after the teaching phases of each session. Four classification predictors were analysed: post-test scores, pretest–to–post-test change scores, modifiability ratings, and teaching duration. An overall modifiability rating was found to be the best single predictor of DLD, with 100% sensitivity and 88% specificity after one DA session and 100% sensitivity and specificity after two DA sessions (Petersen et al., 2017). This was in line with previous studies where modifiability was more consistently predictive than post-test scores or change

scores (Peña et al., 2006, 2014b; Ukrainetz, Harpell, Walsh & Coyle, 2000). Additionally, any two combinations of narrative post-test scores, modifiability ratings, and teaching duration for just one DA session resulted in sensitivity and specificity rates over 90%. As in Peña et al. (2014b)'s study, narrative change scores did not distinguish between the typically developing children and those with DLD, with the second gaining as much as typical learners on average (Petersen et al., 2017). This might be explained by the relatively easy learning content and the extra supportive procedure used in the teaching phase (Petersen et al., 2017), unlike in Peña et al.'s (2014b) study, where the story generation task proved too demanding for typically developing children to make larger gains than the DLD group.

Moreover, although the DA in Petersen et al. (2017)'s study was already considerably shorter than what has been used in previous research, it may be possible to abbreviate the process even further since post-hoc exploration of the first teaching cycle indicated that similar classification accuracy could be obtained after a 10-minute session. To further expand the DA's practicality, Petersen et al. (2017) also identified the cut-off points on each predictive indicator that best separated typical versus impaired performance.

DA procedures focusing on vocabulary acquisition and categorisation abilities have also proven effective for distinguishing between stronger and weaker language skills in children from Native-American (Ukrainetz et al., 2000) and Latino-American backgrounds (Kapantzoglou et al., 2012; Petersen et al., 2020). A recent study by Petersen et al. (2020) demonstrated that a 15-minute-long DA of inferential word learning, employing a hybrid pretest–teaching–post-test and graduated prompting approach, was more accurate in identifying DLD among 31 Hispanic children aged 5;09–9;07 years than two bilingual English/Spanish static tests of receptive and expressive vocabulary: the EOWPVT-SBE and the ROWPVT-SBE (Martin & Brownell, 2012a, 2012b). The combination of post-test scores and modifiability ratings from the DA yielded 90% sensitivity and 90.5% specificity, surpassing the classification results obtained from the static vocabulary tests, even when these were administered in Spanish and English.

An exception is Lazewnik et al. (2019)'s study with 30 Mexican American Spanish–English speaking children aged 4–5 years. This study aimed to determine whether the BESA Morphosyntax and Semantics subtests (Peña et al., 2014a) were more effective than the English and Spanish versions of the *Clinical Evaluation of Language Fundamentals Preschool–Second Edition* (CELF-P2; Semel et al., 2006) at distinguishing children with DLD from those without. It also evaluated whether adding informal language assessment

measures (DA task, parent interview, and mean length of utterance in words from a narrative language sample) improved DLD classification accuracy beyond the BESA alone. For the DA, a pretest–teach–post-test task based on Peña et al. (2001)’s protocol targeted vocabulary and was delivered in English or Spanish, depending on the child’s language use. Children completed the *EOWPVT–Fourth Edition: Spanish Bilingual Edition* (EOWPVT-4: SBE; Brownell, 2012) as both pre- and post-tests, with two 20-minute mediated sessions over three weeks in between them. Differing from Peña et al. (2001), Lazewnik et al. (2019) did not consider whether the children already knew the targeted words, hindering the DA task’s ability to distinguish between those with and without DLD, as 11 of 15 typically developing children and 10 of 15 language-impaired children initially scored within normal limits. Still, 10 typically developing children showed pretest–to–post-test gains of more than 3 points compared to only six children in the DLD group, suggesting a trend favouring the former.

In the UK, Camilleri and Law (2007) developed a 45-minute-long pretest–teach–post-test DA of word learning to compare the performance of 14 typically developing preschoolers with 40 preschoolers who had been referred to speech and language therapy, including 12 children with EAL and their 28 monolingual English-speaking peers within the latter group. The process began with a static vocabulary test (the BPVS), followed by an interactive teaching phase where children were prompted to match the targeted words with corresponding referents within picture cards. Subsequently, the children’s ability to retain those words for expressive and/or receptive purposes was assessed (Camilleri & Law, 2007). The DA was found to differentiate between typically developing children and those referred to speech and language therapy. Importantly, referred children with EAL performed comparably to their referred monolingual peers on the DA despite scoring significantly lower on the static vocabulary test, thus adding evidence to the unsuitability of static tests standardised in monolingual population for multilingual children (Camilleri & Law, 2007). It was further observed that the dynamic vocabulary scores were highly correlated with the static vocabulary scores when the whole group of referred children was considered, but not when focusing on the group with the lower static vocabulary scores separately. Among these lower-scoring children, the variability in the dynamic vocabulary scores could not be predicted by their static vocabulary scores (Camilleri & Law, 2007).

In a follow-up study, Camilleri and Law (2014) examined 37 preschoolers of the original 40 from their 2007 study. They explored the differences between the lower-scoring children and the referred group as a whole on static and dynamic tests across time. The findings showed that the DA substantially increased the predictive capacity of the static receptive vocabulary

measure, especially for children whose static vocabulary scores were below the 25th percentile. In a related effort, Camilleri and Botting (2013) employed a revised version of the DA introduced by Camilleri and Law (2007, 2014) to devise and test the *Dynamic Assessment of Word Learning* (DAWL). This study included 15 preschoolers, 10 referred for speech and language therapy, and among them, seven with EAL (five within the referred group and two in the typically developing group). Their findings provided further evidence of this DA procedure's reliability and validity in enhancing static assessment and in offering valuable insights into the likelihood of a child having DLD.

Furthermore, Hasson et al. (2012a) also incorporated a version of Camilleri and Law (2007)'s DA protocol into the *Dynamic Assessment of Preschoolers' Proficiency in Learning English* (DAPPLE). This DA battery uses brief pretest–teach–post-test procedures to evaluate the ability of preschoolers with EAL to learn vocabulary, phonology, and sentence structure within a single session lasting 30–40 minutes. In a trial with 12 children referred to speech and language therapy and 14 typically developing controls matched for age and socioeconomic status, the DAPPLE successfully discriminated between the two groups (Hasson et al., 2012a). The referred children required more prompting to identify targeted words in the receptive vocabulary assessment and retained fewer words in the post-test expressive component. These children also needed more assistance to acquire the targeted clause elements in the teaching phase and produced shorter clauses in the post-test. Similarly, they produced fewer words accurately in the pre- and post-tests and showed less consistency across trials (Hasson et al., 2012a). Notably, Hasson et al. (2012a) were the first researchers to explore the value of DAs focused on syntactic and phonological abilities to identify DLD in multilingual children (see also Hasson et al., 2012b).

Research on DA in countries where English is not the majority language has further corroborated its efficacy in detecting DLD among multilingual children. For instance, in Switzerland, Maragkaki and Hessel (2016) adapted the vocabulary subtest of the DAPPLE to German and trialled it with 12 preschoolers, including six typically developing and six language-impaired, who had German as an additional language and spoke a variety of other languages (French, Serbian, Albanian, Portuguese, and Spanish). Their findings supported the DA's advantage over the standardised static German vocabulary test in classifying these groups correctly, consistent with the results from Hasson et al. (2012a). Furthermore, a study in France by Hadjadj, Kehoe and Delage (2022) demonstrated the utility of a DA designed to evaluate morphosyntactic skills in distinguishing between typical development and DLD in a

sample of 79 French-speaking children aged 5;01–11;09 years, regardless of their mono- or multilingual background.

Ultimately, the expectation in all diagnostic DA studies was that typically developing children who may score low at the pretest would show sizable improvements following a short-term intervention. In contrast, children at risk for DLD who also score low at the pretest—often indistinguishable from their typically developing peers on static tests—may be less able to benefit from such interventions, showing limited or no progress afterwards (e.g., Gutiérrez-Clellen & Peña, 2001; Peña et al., 2014b). Overall, research has found that post-test scores and practitioner judgements of the degree of assessor effort and the child's modifiability are more effective in predicting language impairment status than (static) pretest scores and change scores that reflect the magnitude of change between pre- and post-tests (e.g., Gutiérrez-Clellen & Peña, 2001; Peña et al., 2006, 2014b; Petersen et al., 2017, 2020; Ukrainetz et al., 2000).

These findings on the utility of DA-derived post-test and modifiability scores are echoed in two recent reviews examining the evidence of applying DA for identifying DLD in multilingual children (Hunt et al., 2022; Orellana et al., 2019). In line with this, Hunt et al. (2022) suggested incorporating a delayed post-test in the DA procedure to assess skill retention more effectively, as immediate post-tests may not always capture significant differences between the results of children with and without DLD (Hasson et al., 2012a; Peña et al., 2014b; Petersen et al., 2017). Such a delay, whether brief—lasting only a few minutes within the same session, similar to the method used by Hasson et al. (2012a) and Maragkaki and Hessels (2016) in their evaluation of vocabulary learning abilities—or occurring in a different session (Peña et al., 2014b; Petersen et al., 2017), is particularly pertinent for challenging skills commonly affected by DLD, like mapping, phonological short-term memory, and working memory.

Moreover, there is an emphasis on carefully selecting modifiability measures and acknowledging possible cultural biases stemming from practitioners' expectations of typical learner behaviour within their own culture (Hunt et al., 2022; Orellana et al., 2019). A combined approach of language skill measures and modifiability scores (e.g., Peña et al., 2014b) can offer a more complete and unbiased view of a child's language acquisition abilities, boosting the DA's classification accuracy.

Further reflections from Hunt et al. (2022)'s review concern the appropriateness of DAs for children as young as 3 years old (e.g., Camilleri & Law, 2007; Hasson et al., 2012a), although

the accuracy of the results may be higher as they grow older, depending on the targeted language skill. Hunt et al. (2022) also recommend assessing language skills that the child has not yet mastered to prevent ceiling effects, citing Lazewnik et al. (2019)'s study, where the DA provided limited room to demonstrate learning potential, as many children had already acquired the targeted skill. They also suggest comparing the outcomes across different language skills (e.g., Hasson et al., 2012a) to achieve a more comprehensive assessment rather than focusing on a single skill.

In addressing the common criticisms regarding the practicality and efficiency of DA, particularly its perceived time-consuming and individualised nature, it is noteworthy that the duration of DA approaches is generally on par with that of standard language assessments (Hunt et al., 2022; Orellana et al., 2019). Moreover, the reliability and validity of DA outcomes can be enhanced by incorporating prespecified prompts (e.g., Camilleri & Law, 2007; Hasson et al., 2012a; Maragkaki & Hessels, 2016), scripts (e.g., Kapantzoglou et al., 2012; Peña et al., 2014b; Petersen et al., 2017), and practitioner training (e.g., Peña et al., 2014b; Petersen et al., 2017). These strategies promote greater consistency in assessor-child interactions and help generalise results to actual practice (Hunt et al., 2022). Equally important are the aspects of interrater reliability for the various DA measures, achieved through the assistance of second raters (e.g., Kapantzoglou et al., 2012; Petersen et al., 2020), and the use of implementation fidelity checklists and rating forms to ensure that assessors adhere to the DA administration procedures (e.g., Peña et al., 2014b; Petersen et al., 2017).

In conclusion, the collective evidence highlights the potential of DA as a promising complementary method for diagnosing DLD in multilingual children. This is especially relevant in contexts like the UK, where many practitioners do not share the languages of the children they evaluate. Even so, it is important to recognise the emergent and novel nature of this research area (Hunt et al., 2022; Orellana et al., 2019). Existing studies often have limitations, such as small sample sizes and case-control designs, which can affect the robustness of the results. There is also a clear need for further research involving multilingual children from diverse linguistic backgrounds beyond the US Spanish-English speaking population. Such research is critical to ensure that the positive outcomes concerning the diagnostic and practical effectiveness of DA procedures are generalisable, even when practitioners and children do not share the same culture or language (Hasson & Joffe, 2007; Hunt et al., 2022). Some European studies mentioned in this section (e.g., Hasson et al., 2012a; Maragkaki & Hessels, 2016), begin to address this gap, contributing to a more

inclusive and comprehensive understanding of the role of DA in DLD identification in multilingual settings.

2.6 Dynamic Assessment in the Present Study: Narrative Macrostructure, Emotional Vocabulary, and Receptive Affective Prosody

This chapter has explored the intricacies of multilingual language development, stressing the central role of early DLD detection in ensuring effective support for this demographic to mitigate the disorder's impact on language development and broader life areas. Within this scope, DA is introduced as a viable approach to shed light on whether children's language difficulties stem from DLD or the process of acquiring a new language. With its focus on a child's learning ability through integrated intervention and evaluation, DA is notable for its ability to adapt to different language backgrounds. This flexibility is crucial in the accurate diagnosis of DLD among multilingual children in the UK, a task often complicated by practitioners' inability to speak the children's home language(s) and the limited availability of adequate assessment tools.

In response to these identified challenges, this PhD research project is dedicated to the development of a DA tool for primary school-aged children in the UK who have EAL and may be at risk for DLD. This resource is intended to accommodate the varied English proficiency levels within this population, employing storytelling in English as an engaging and comprehensive medium to assess learning potential through changes in performance before and after instruction. The areas of focus —narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and affective prosody processing— were chosen for their diagnostic value and relevance to the challenges children with DLD face, as identified in Sections 2.4.1, 2.4.2, and 2.4.3. The subsequent chapters, *Chapters 3 and 4*, detail the development, implementation, and evaluation of the DA tool.

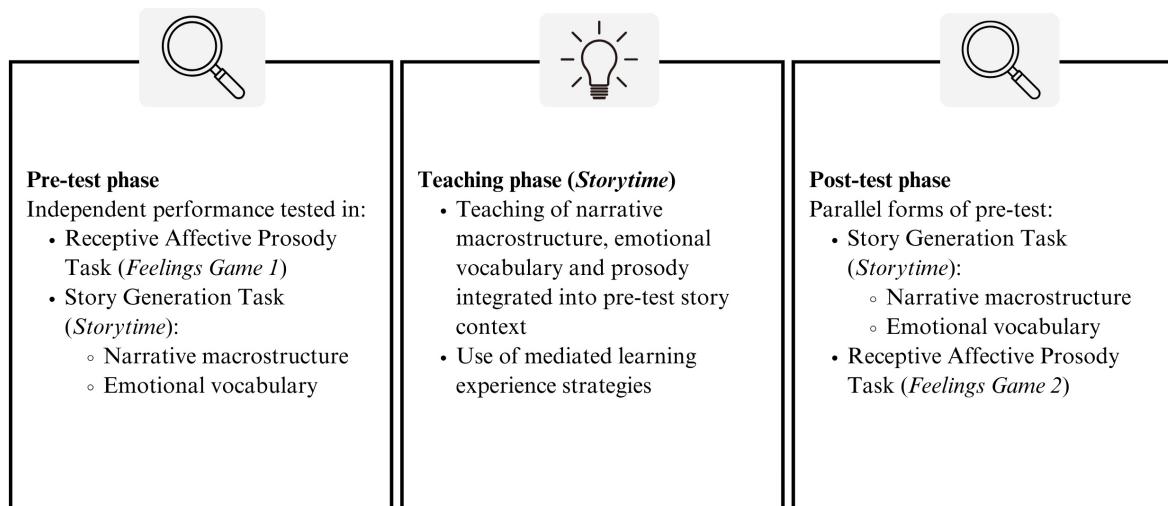
Chapter 3. Development and Piloting of the Dynamic Assessment Methods

3.1 Introduction

In the UK's diverse multilingual landscape, considerable challenges are posed in differentiating Developmental Language Disorder (DLD) from typical language variation in children with English as an Additional Language (EAL), as explained in Section 2.5.1. The purpose of this PhD research was to create and test a Dynamic Assessment (DA) for UK-based, primary school-aged children with EAL at risk for DLD. This resource employs storytelling and other activities in English to target learning potential across three diagnostically relevant areas for DLD: narrative macrostructure (story grammar and episodic structure complexity); emotional vocabulary; and, receptive affective prosody. The DA was designed to be a practical and accessible tool that serves as a source of convergent evidence for identifying DLD in this population, facilitating timely support.

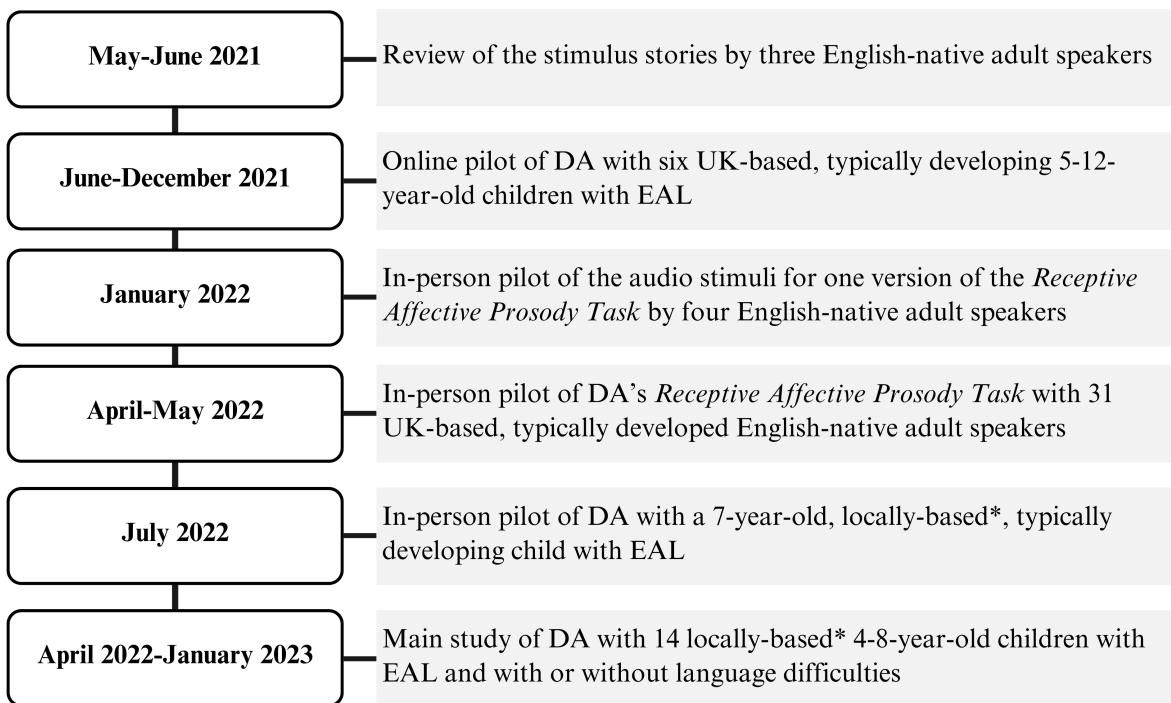
As summarised in Figure 3, the DA consists of a pretest phase, a teaching phase, and a post-test phase centred around becoming a 'storytelling superstar' during which three activities are undertaken: *Feelings Game 1*, *Storytime* and *Feelings Game 2*. In the pretest, children's independent performance (i.e., without adult guidance) is first tested in the *Receptive Affective Prosody Task (Feelings Game 1)*, followed by a *Story Generation Task* at the beginning of the storytelling activity (*Storytime*). Subsequently, a short teaching phase is conducted within the storytelling activity that incorporates the main strategies of a mediated learning experience. The story used in the pretest is employed here to teach narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and affective prosody integrated into the story context. Finally, in the post-test, parallel forms of the pretest tasks — identical in procedure but with different stimuli — are carried out to assess children's independent performance again: a second *Story Generation Task* using a new story and another *Receptive Affective Prosody Task (Feelings Game 2)*. Children receive a 'star' point for each task completed (see star points chart in Appendix 1) and a small prize at the end. The DA measures consist of the *Modifiability Rating Scale* scores, which reflect the children's capacity to respond to instruction during the teaching phase, and the pretest, post-test and (pretest-to-post-test) change scores in both the *Receptive Affective Prosody Task*, assessing emotional prosody processing, and the *Story Generation Task*, which targets narrative macrostructure (story grammar and episodic structure complexity) and emotional vocabulary.

Figure 3. Overview of the Dynamic Assessment Developed in the PhD Research Project



The research conducted to test this DA's methods encompassed several phases. As shown in Figure 4, three pilot studies were conducted between June 2021 and July 2022 to test the initial feasibility of the DA activities. These included an online pilot study with six typically developing primary school-aged children with EAL residing in the UK to test the initial versions of the pretest *Receptive Affective Prosody Task* and the storytelling activity (including the pre- and post-test *Story Generations Tasks*, and the teaching phase), as well as an in-person pilot study with 31 typically developed English-speaking adults to identify the most effective version of the three *Receptive Affective Prosody Tasks* created for the DA. Adults were selected due to concerns about the reliability of the original task and easier access to this demographic, facilitating the identification of potential issues before testing with children. Another in-person pilot study was conducted with a seven-year-old child with EAL to test the updated DA activities before proceeding to the main study. Additionally, to further refine the materials used in these activities, three English-native adult speakers revised the stimulus stories between May and June 2021, with another four English-native adult speakers reviewing the audio stimuli for one version of the *Receptive Affective Prosody Task* in January 2022.

Figure 4. Timeline of the Pilot and Main Studies of the PhD Research Project



Note. *Locally-based refers to Newcastle upon Tyne and surrounding areas. Abbreviations: EAL: English as an Additional Language; DA: Dynamic Assessment.

As outlined in Table 2, these pilot studies are the focus of the present chapter, starting with the DA's *Receptive Affective Prosody Task* in Section 3.2, which includes the design of its three versions across the pilot studies with children (Sections 3.2.1 and 3.2.3) and adults (Section 3.2.2). The development of the DA's storytelling activity is then described in Section 3.3, covering the materials (Section 3.3.1), instructions (Section 3.3.2), and pilot studies with children (Sections 3.3.3 and 3.3.4). This storytelling activity involves the pre- and post-test *Story Generation Tasks* targeting narrative macrostructure and emotional vocabulary, and a teaching phase in between focusing on these two domains plus affective prosody. The chapter concludes with a summary of the piloting phase and the key decisions that informed the design of the main study.

Table 2. Chapter 3 Overview

3.1	Introduction
3.2	Development of the Receptive Affective Prosody Task
3.2.1	Pilot testing 1: Receptive Affective Prosody Task (Congruent vs Incongruent Prosody)
3.2.2	Pilot testing 2: Receptive Affective Prosody Tasks (Congruent vs Incongruent Prosody, Congruent vs Neutral Prosody, and Neutral Content)
3.2.3	Pilot testing 3: Receptive Affective Prosody Task (Neutral Content)
3.2.4	Conclusions
3.3	Development of the Storytelling Activity
3.3.1	Materials
3.3.2	Instructions
3.3.3	Pilot testing 1: Storytelling activity
3.3.4	Pilot testing 2: Storytelling activity
3.3.5	Conclusions
3.4	Summary of Piloting Phase and Design Decisions for the Main Study

Prior to identifying the participants for all studies, the research was submitted for ethical approval by the Faculty of Humanities and Social Sciences Ethics Committee at Newcastle University. Approval was obtained on 6/4/2020, with subsequent amendments to the ethics application on 20/1/21, 8/12/21, 14/1/22, 17/3/22 and 14/12/22 due to research methods adjustments as a result of the COVID-19 pandemic.

3.2 Development of the Receptive Affective Prosody Task

During the pre- and post-test phases of the DA, children's receptive affective prosody is informally evaluated through the *Receptive Affective Prosody Task (Feelings Game)*, which measures their ability to interpret emotions in people's voices. Three versions were devised and pilot-tested to determine the most effective design for a task with such purpose: a *Congruent vs Incongruent Prosody* version, a *Congruent vs Neutral Prosody* version, and a *Neutral Content* version. In the *Congruent vs Incongruent Prosody* version, the child listens to utterances where the affective prosody either aligns or does not align with the emotional lexical content, and indicates if the voice heard reflects the intended emotion. In the *Congruent vs Neutral Prosody* version, the child listens to pairs of utterances —one with affective prosody matching their emotional lexical content, and one with neutral prosody— and determines, for each pair, which voice better conveys the intended emotion. In the *Neutral Content* version, the child listens to utterances with emotionally neutral lexical content and indicates which of two emotions the voice expresses.

This part of the chapter describes the development and piloting of these three task designs. The *Congruent vs Incongruent Prosody* task was initially trialled with children (Section 3.2.1). It was then further trialled with adults, along with the *Congruent vs Neutral Prosody* task and the *Neutral Content* task (Section 3.2.2). Finally, the *Neutral Content* task was trialled again with a child (Section 3.2.3).

3.2.1 Pilot testing 1: Receptive Affective Prosody Task (Congruent vs Incongruent Prosody)

Pilot testing 1: Participants' characteristics and recruitment procedure

The *Congruent vs Incongruent Prosody* version of the *Receptive Affective Prosody Task* was the first to be developed. This was tested in the initial pilot study of the DA—which included this task's pretest (*Feelings Game 1*), alongside the storytelling activity (*Storytime*) comprising the pre- and post-test *Story Generation Tasks* and teaching phase—with six typically developing, UK-based children with EAL. They were aged 5;06 to 12;11 years ($M = 8;05$, $SD = 2;10$) and consisted of four girls and two boys. Three children had Arabic as their home language, with Spanish being the home language for the other three children.

The initial recruitment through the researcher's contacts was supplemented with word-of-mouth advertising and the distribution of the flyer in Appendix 2 across the researcher's networks to boost participation, as the pandemic severely slowed participant response and recruitment.

The children's parents received the participant information sheets and consent forms in Appendices 3 and 4 ahead of the session, and they were given the opportunity to have these documents translated into their most proficient language if needed. The option to review these documents with the researcher via phone or video call was also offered, and when required, an interpreter could be arranged. In addition, children were shown a child-friendly information video sent to them in advance (see Appendix 5), and their verbal consent was sought at the beginning of the session.

The facilitator checklist in Appendix 6 was also shared with the primary caregiver who would be located with the child during the session. This person was asked to support the processes with the child during the video call and assist with technology on the other end. To adequately prepare them, the checklist covered how to get ready for the session and what participation would involve. The researcher was also available before the session to provide further clarity.

Pilot testing 1: Data collection procedure

The pilot of the initial *Congruent vs Incongruent Prosody* version of the *Receptive Affective Prosody Task* took place between June and December 2021 as part of a 45-60-minute Zoom session where the rest of the DA was also piloted, except for the post-test *Receptive Affective Prosody Task*, which had yet to be developed. Before the session, the parental questionnaire from Section 4.3.1 was used to collect data on children's language use and exposure at home, as well as general and language development.

The challenges posed by the pandemic for in-person data collection made the use of remote methods necessary in this pilot study. This approach was informed by the increasing body of evidence supporting the efficacy of conducting language assessments via telepractice, which produce outcomes comparable to those of in-person assessments (e.g., Raman et al., 2019; Waite, Theodoros, Russell & Cahill, 2010). It also took into account guidelines for remote delivery of speech and language therapy from Newcastle University (Moxam, 2021) and official organisations (ASHA, 2020; RCSLT, 2020b). Accordingly, Zoom was chosen as the most suitable video-conferencing platform for this pilot study due to its functionality and security characteristics (i.e., user interface, GDPR-compliant, and end-to-end encryption; RCSLT, 2020b; Zoom, 2020), and a parent or another primary caregiver was asked to act as the facilitator and be in the same room as the child. The session was conducted in a quiet, private space at the researcher's home, who used headphones to ensure calls were not overheard and that only she could hear the child and the facilitator speaking. Likewise, efforts were made to ensure that the child and the facilitator were in a confidential space.

The session adopted a game format divided into two parts —pretest *Receptive Affective Prosody Task (Feelings Game 1)* and storytelling activity (*Storytime*), including the pre- and post-test *Story Generation Tasks*— and focused on becoming a 'storytelling superstar' to motivate the children. They received a 'star' point for each task completed, and a final small prize of stickers and a bookmark that was posted to them. For the *Receptive Affective Prosody Task*, the script in Appendix 7 was employed for consistency of procedure across participants, with items displayed via PowerPoint using Zoom's screen-share function. Before beginning the task, a training phase was included to ensure children's understanding of the targeted emotions. Children were first shown the pictograms with the happy, sad, angry, and scared faces one by one, using an illustration to exemplify each emotion (see Figure 5). They were then presented with all pictograms at once and instructed to point at the face that looked happy, sad, angry, or scared. Following this, the researcher introduced the task by explaining

how feelings can also be expressed through our voices. As an example, the audio clip of the sentence 'My friends are great', spoken with prosody conveying happiness, was played and related to the happy pictogram.

Figure 5. Examples from the Training Phase in the Receptive Affective Prosody Task



What about the boy in this picture?

How does he feel? What does he have on his face? [A tear]

You can see other boys playing football who seem to be having fun, but this boy looks very sad, perhaps because they don't want to play with him.

In this picture we see a woman knocking on the door of a house where children are playing inside. The children are making lots of noise, jumping around on a chair and playing music with pot and pans.

The woman is going to tell the children to stop the noise and she looks very angry.

After the training phase, the task began with children listening to two pre-recorded practice items, repeated as necessary to familiarise them with what they were required to do. This was followed by 16 pre-recorded stimuli sentences with a happy, angry, sad, or scared lexical content (see Table 3). The affective prosody with which the practice and stimuli utterances were spoken was either congruent or incongruent with the emotional lexical content. For each sentence, the child was asked to indicate whether the voice they heard matched how the person felt or not (Congruent and Incongruent conditions, respectively). Responses were recorded on the scoring sheet in Appendix 8. This design is in line with previous studies that used stimuli sentences with both consistent and discrepant lexical content and affective prosody to investigate children's ability to interpret speakers' emotional states based on the prosodic cues (e.g., English: Friend, 2000, Morton & Trehub, 2001; Japanese: Ikeda, 2021; Swedish as first language: Champoux-Larsson & Dylman, 2018).

Table 3. Stimuli Sentences for the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task

Lexical content	Congruent affective prosody	Incongruent affective prosody
Happy	These books are fun This dog is adorable	The butterflies are pretty (sad prosody) The drawing is lovely (sad prosody)
Sad	The sandwich is wet (P1) These balloons aren't green She is being mean	The bus is gone (P2; happy prosody) The shop is closed (happy prosody) My team is losing (happy prosody)
Angry	The floor is dirty The ground is disgusting	The wall is filthy (scared prosody) The children are noisy (scared prosody)
Scared	The eagle is big The water is cold	The wolf is scary (angry prosody) The cheetah is fast (angry prosody)

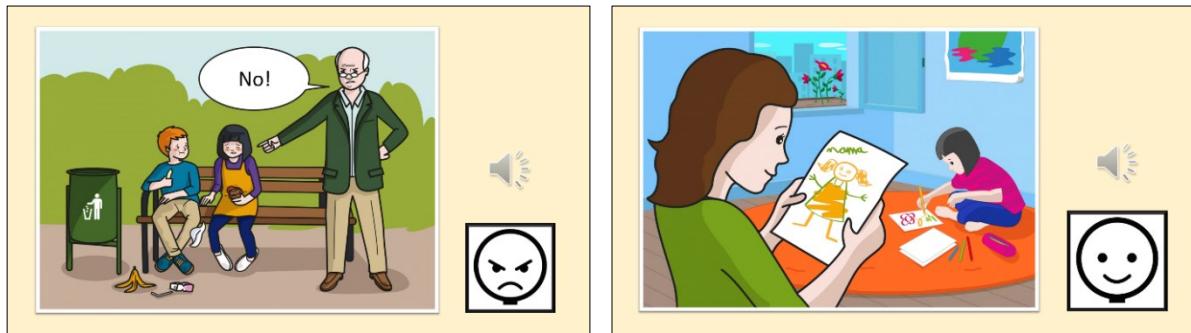
Note. Abbreviations: P1: Practice item 1; P2: Practice item 2.

All stimuli sentences were audio recorded to convey one of four targeted emotions by a female British English-native speaker who is a speech and language therapist. Audacity audio editing software was used to reduce background noise, normalise file waveforms to the same decibel loudness, and to remove silence before or after each utterance. The researcher, a fluent second language English speaker, and her supervisors, both first language English speakers, evaluated all recordings to ensure they conveyed the intended basic emotional states based on their prosodic attributes. There is evidence that these states are recognised from consistent prosodic patterns regardless of language and are therefore helpful for assessing children's ability to identify emotions in the voice regardless of their English proficiency (Ma et al., 2022; Pell et al., 2009a).

The structure of the stimuli sentences was limited to simple S-V/S-V-Adj sentences of equal four-word length. The subject matter of these sentences and their contextual scenarios were meant to be cross-culturally suitable and mirror general knowledge that would normally be within the grasp of children in their early to mid-primary school years. Each stimulus sentence was embedded in a simple scenario beginning with a brief introductory passage where the central character experiences one of four basic emotions. This passage was accompanied by an illustration and a pictogram depicting a happy, sad, angry, or scared face to aid task comprehension and response (see examples in Figure 6). The illustrations are property of Fundación Orange and were created by Tropical Estudio, María de la Fuente, and Rubén Rodríguez for #Soyvisual (for further details, see <https://www.soyvisual.org>), while the pictograms are property of the Government of Aragón (Spain) and were created by Sergio

Palao for ARASAAC (see <https://arasaac.org>). Both are distributed under Creative Commons licenses (BY-NC-SA).

Figure 6. Examples from the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task



The children littered the floor. The man tells them to stop. He seems angry! Listen to him talk. Does he sound angry? [Play audio clip of the sentence 'The ground is disgusting' said with congruent angry prosody]

The mum looks at the drawing her daughter made. She is smiling and seems to like it. Listen to the mum talk. Does she sound happy? [Play audio clip of the sentence 'The drawing is lovely' said with incongruent sad prosody]

Pilot testing 1: Results and discussion

The children's performance was higher for the congruent stimuli than for the incongruent stimuli (see Figures 7, 8, and 9). Accurate judgments of congruent stimuli remained relatively stable across ages, indicating that the ability to match utterances with emotional content to their corresponding emotional vocal cues is established early in development. However, the children's ability to identify incongruent emotional cues correctly generally increased as a function of age. These findings suggest that the younger the child is, the more they struggle to grasp the concept that the voice does not always match how the person feels, which is consistent with studies showing that the younger children are, the more they tend to rely on lexical content rather than prosody to judge speakers' emotions (Friend, 2000; Ikeda, 2021; Morton & Trehub, 2001).

Figure 7. Total Score by Age in the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task in the Online Child Pilot Study

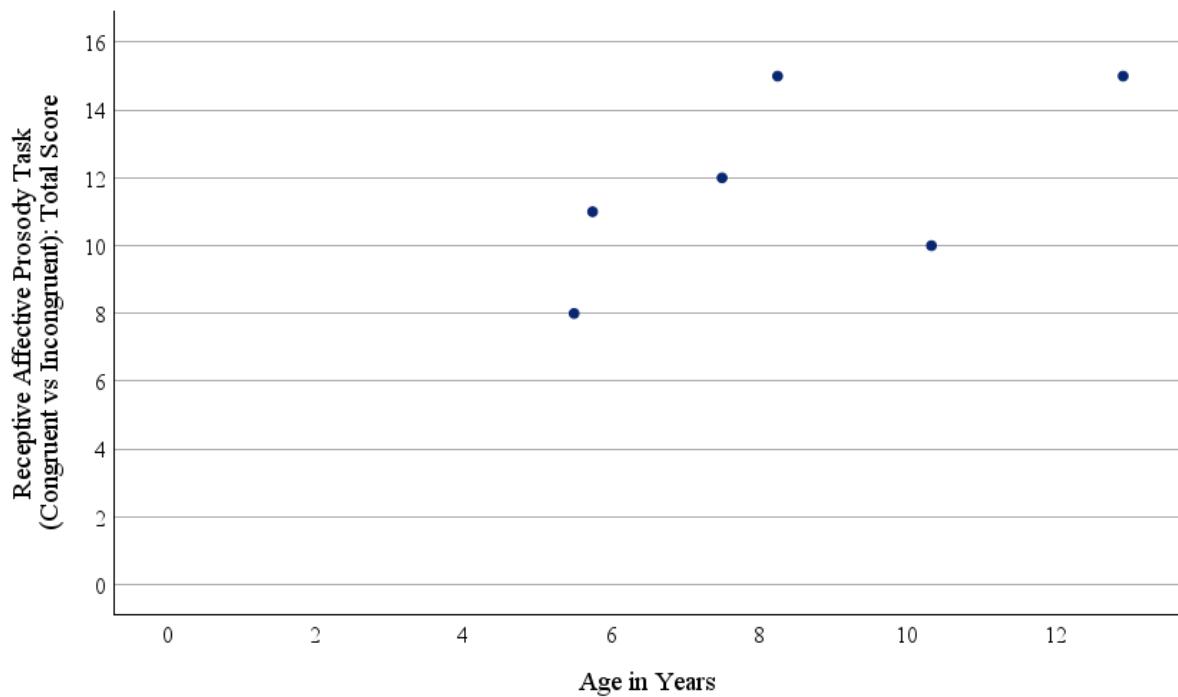


Figure 8. Total Score by Age in the Congruent Trials of the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task in the Online Child Pilot Study

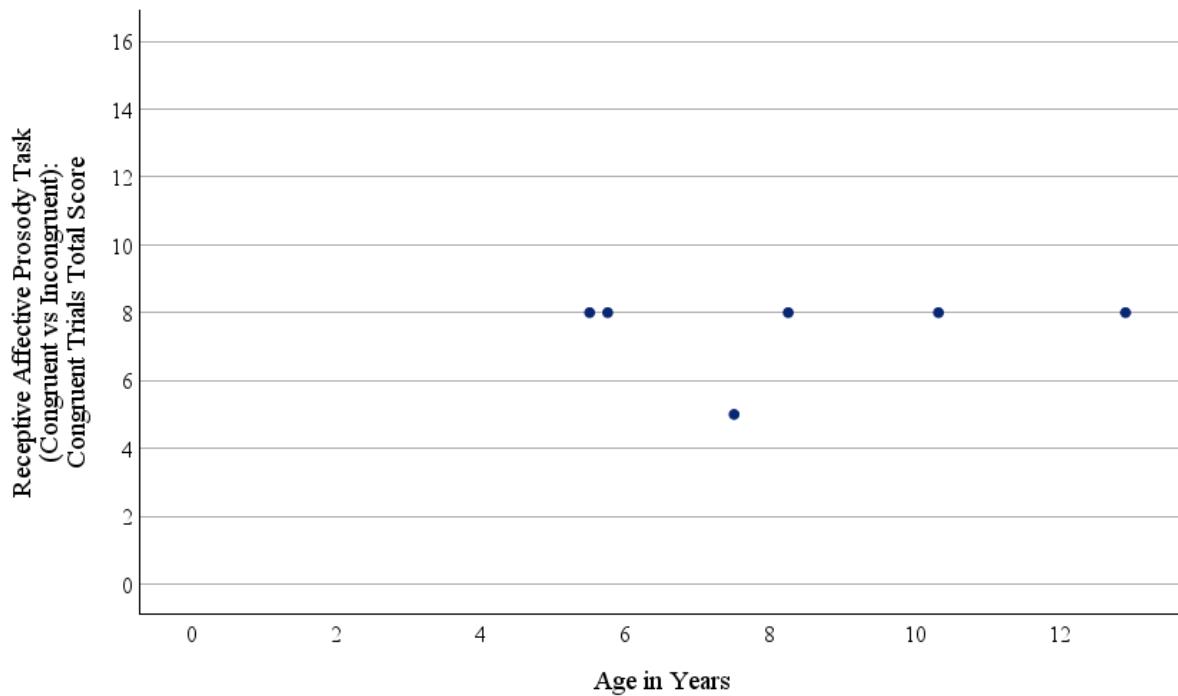
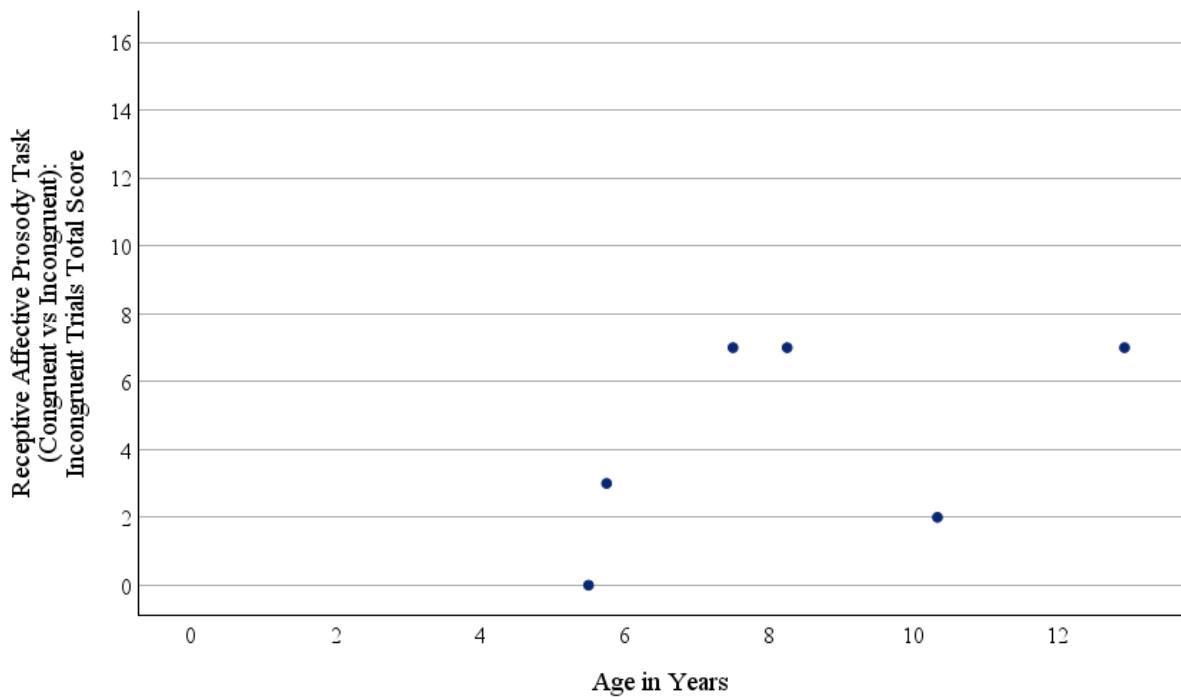
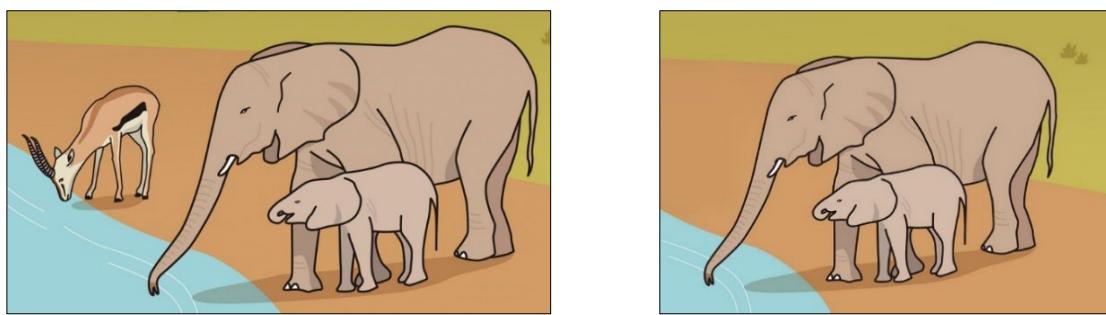


Figure 9. Total Score by Age in the Incongruent Trials of the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task in the Online Child Pilot Study



An important additional observation from this pilot study concerned the youngest participant getting distracted by the visual supports employed in the task. As a result, any potentially distracting elements of the scenario illustrations that were not essential to the situation presented to contextualise the stimuli utterances were removed following this piloting phase (see example in Figure 10).

Figure 10. Example of Scenario Illustration in the Congruent vs Incongruent Prosody Version of the Receptive Affective Prosody Task Before and After Removing Potentially Distracting Elements



Note. The gazelle was removed due to being irrelevant to the contextualising scenario for this stimulus utterance: *The baby elephant is a little scared of the water. His mum tells him that it's okay and completely safe. Listen to the baby elephant talk. Does he sound scared?* [Play audio clip of the sentence 'The water is cold' said in a congruent scared prosody]

Based on these findings from the pilot of the *Congruent vs Incongruent Prosody* task — particularly given younger children's challenges with mismatches between emotional prosodic cues and lexical content —, further piloting of this task and the creation of two new *Receptive Affective Prosody Tasks* (*Neutral Content* and *Congruent vs Neutral Prosody*) were deemed pertinent to determine the most effective task for the final DA resource. As detailed in Section 3.2.2, subsequent piloting with adults was conducted for all three tasks due to the more straightforward, quicker access to this demographic and the rationale that if adults found a task challenging, we could expect such a task to be problematic for children.

3.2.2 Pilot testing 2: Receptive Affective Prosody Task (Congruent vs Incongruent Prosody, Congruent vs Neutral Prosody, and Neutral Content)

Pilot testing 2: Participants' characteristics and recruitment procedure

The effectiveness of the three versions of the *Receptive Affective Prosody Task (Congruent vs Incongruent Prosody, Congruent vs Neutral Prosody, and Neutral Content)* was investigated in a second pilot study with typically developed English-native adult speakers. Thirty-six individuals took part, five of whom were excluded from the data analysis due to not being English-native speakers ($n = 2$) or having speech, language or hearing difficulties ($n = 3$). The 31 participants included in the analysis were 18 to 40 years old ($M = 25;00$, $SD = 7;05$), with a large majority being female ($n = 29$). Seven participants (22.6%) spoke other languages in addition to English, including Spanish ($n = 5$), French ($n = 3$), and Polish ($n = 1$).

The study was advertised through the researcher's professional networks and the SONA system for Speech and Language Sciences studies at Newcastle University. Before any participation, the individuals interested in taking part were provided with the research information sheet and consent form in Appendices 9 and 10, and the researcher was available to address any questions. Participants recruited through SONA were students and received course credit for taking part.

Pilot testing 2: Data collection procedure

The pilot study of the three versions of the *Receptive Affective Prosody Task (Congruent vs Incongruent Prosody, Congruent vs Neutral Prosody, and Neutral Content)* was conducted during 30-minute sessions in a quiet lab room inside the King George Building VI at Newcastle University in May 2022. After presenting participants with the information sheet and consent form, they were asked to provide basic personal information. This included their age, gender, whether they have any speech, language or hearing difficulties, whether they are

native English speakers, and other languages they may know. The three tasks were then shown on a computer via PowerPoint presentations, labelled as Tasks A, B, and C. These corresponded to the *Neutral Content*, *Congruent vs Neutral Prosody*, and *Congruent vs Incongruent Prosody* versions of the *Receptive Affective Prosody* task, respectively.

For the *Neutral Content* task, participants listened to two pre-recorded practice items, followed by 16 pre-recorded stimuli sentences with emotionally neutral lexical content, spoken with prosody expressing one of two emotions: happy/sad or angry/scared (see Table 4). For each sentence, they indicated which of two possible emotions the voice was expressing (please note: if running this task with children, they would either point at the pictogram of the two which best represents the emotion expressed prosodically, or respond verbally by naming the emotion). As in the other two versions of the *Receptive Affective Prosody Task*, contextualising scenarios introduced each sentence to support response and comprehension (see example in Figure 11). Previous studies exploring the understanding of emotional prosody in monolingual children with DLD have used stimuli containing emotionally neutral content, which ranged in length from single real words (Boucher et al., 2000) and nonsense words (Löytömäki et al., 2020), to sentences (Courtright & Courtright, 1983; Taylor et al., 2015; Trauner et al., 1993) and short narrative passages (Fujiki et al., 2008). The decision to use simple S-V/S-V-Adj, three-to-four-word sentences in the *Neutral Content* task was made on the basis that these would be easier for children with limited English language skills whilst still providing enough context for them to be able to perceive the affective prosody of the stimuli and perform to the highest possible level. The content of the sentences was deemed suitable for early to mid-primary school-aged children from diverse cultural backgrounds. These were audio recorded by a female British English-native speaker with acting experience, edited using Audacity audio software, and validated by the researcher and her supervisors.

Table 4. Stimuli Sentences for the Neutral Content Version of the Receptive Affective Prosody Task

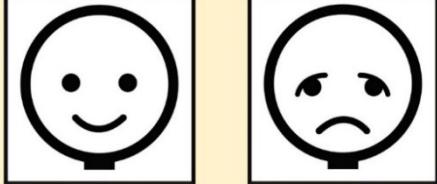
Happy prosody	Sad prosody
The baby is drinking (P1)	They are napping
The ball is red	My friends are running
The car is parked	I am waking up
The cat is brown	My dad is fishing
The moon is round	
Angry prosody	Scared prosody
My hair is short (P2)	The phone is ringing
We are reading	The animals are eating
I am cooking	The table is green
The flowers are growing	The dog is playing
The tap is running	

Note. Abbreviations: P1: Practice item 1; P2: Practice item 2.

Figure 11. Example from the Neutral Content Version of the Receptive Affective Prosody Task



The children do exercise. Listen to one of them talk. Does she sound happy or sad?



[Play audio clip of the sentence 'My friends are running' said with sad prosody]

In the *Congruent vs Neutral Prosody* task, participants listened to two pre-recorded practice items, followed by 16 pre-recorded stimuli sentences with emotional lexical content, each one read twice, one version with congruent affective prosody and the other version with neutral affective prosody (see Table 5). For each pair of utterances, they indicated in which version the voice sounded more like the intended emotion (either 1 or 2) (see example in Figure 12). As in the other two *Receptive Affective Prosody Task* versions, contextualising scenarios were included before each pair of stimuli utterances. Except for practice item 2, the stimuli utterances and contextualising scenarios from the *Congruent vs Incongruent Prosody* task were re-used here. Two female British English-native speakers initially recorded the stimuli sentences as part of their psychology student placement experience. However, only one set of

recordings was selected to keep the speaker's voice consistent throughout the task, both within and between pairs of sentences. For this reason, in addition to the researcher and her supervisors evaluating the two sets of recordings, the non-neutral recordings were further trialled with four British English-native adult speakers in January 2022. Details about this testing are provided in Appendix 11. Audio files were edited as required using Audacity audio software.

Table 5. Stimuli Sentences for the Congruent vs Neutral Prosody Version of the Receptive Affective Prosody Task

Happy lexical content	Sad lexical content
Today was great (P2)	The sandwich is wet (P1)
The butterflies are pretty	The shop is closed
The drawing is lovely	These balloons aren't green
These books are fun	My team is losing
This dog is adorable	She is being mean
Angry lexical content	Scared lexical content
The floor is dirty	The eagle is big
The wall is filthy	The water is cold
The ground is disgusting	The wolf is scary
The children are noisy	The cheetah is fast

Note. Abbreviations: P1: Practice item 1; P2: Practice item 2.

Figure 12. Example from the Congruent vs Neutral Prosody Version of the Receptive Affective Prosody Task



*The wolf is strong and powerful. The baby penguin fears him. Listen to the baby penguin talk. Which one sounds more scared?
Number 1 or number 2?*

[Play audio clip of the sentence 'The wolf is scary' twice, said with neutral prosody (1) and scared prosody (2)]

Finally, the method for the *Congruent vs Incongruent Prosody* task mirrored the one described for the initial pilot study with children in ‘Pilot testing 1: Data collection procedure’ (Section 3.2.1), with minor adjustments made to remove any potentially distracting elements from the contextual scenario illustrations as necessary.

Indeed, all three tasks featured scenarios with a brief introductory passage and an accompanying illustration depicting a character in a core emotional state to provide background for the stimuli utterances. However, to allow the study to be run with multiple participants simultaneously, the researcher recorded the audio for these contextual scenarios and inserted it into the slides, instead of being presented live as would be done for the children.

For each task, participants listened to the audio recordings of the contextual scenarios and the stimuli utterances through headphones and recorded their responses on paper scoring sheets (see Appendix 12). No feedback information was given about whether the response was correct or incorrect. Each participant was provided with instructions and tested individually. For all tasks, instructions included asking participants to listen to the contextualising audio recordings first. It was explained that this would give some background for the stimuli utterances, which were to be played only once. In addition, the following directions were provided for each task, with A corresponding to *Neutral Content*, B to *Congruent vs Neutral Prosody*, and C to *Congruent vs Incongruent Prosody*:

- *Task A*: You will listen to a series of sentences and have to indicate which of two possible emotions the voice was expressing. Please record your responses in the table below.
- *Task B*: You will listen to a series of pairs of sentences. For each pair, you have to indicate the one in which the voice better matches the intended emotion. Please record your responses in the table below.
- *Task C*: You will listen to a series of sentences and have to indicate whether or not the voice you hear matches how the person feels. Please record your responses in the table below.

The task order was counterbalanced considering every possible order so that 6 participants were presented with ABC, 5 with ACB, 5 with BAC, 5 with CAB, 5 with BCA, and 5 with CBA. In sessions where more than one participant was present, they were separated by sufficient space to prevent any potential audio or visual distractions between them.

Pilot testing 2: Results and discussion

The three receptive affective prosody tasks were generally performed by the adults with a high level of accuracy, as shown in Table 6.

Table 6. Descriptive Statistics for the Performance in the Receptive Affective Prosody Tasks in the Adult Pilot Study

Receptive Affective Prosody Task	N	Range	Minimum	Maximum	Median	M	SD
A: Neutral Content	31	1	15	16	16	16.0	0.2
B: Congruent vs Neutral Prosody	31	8	8	16	16	15.2	1.8
C: Congruent vs Incongruent Prosody	31	2	14	16	16	15.5	0.8

A repeated measures ANOVA was conducted to compare the effect of the type of task on participants' performance. Mauchly's test indicated that the assumption of sphericity could not be assumed, $\chi^2(2) = 35.71, p < 0.05$. Thus, the degrees of freedom were adjusted using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.58$). Although the slightly higher mean score in *Task A* makes it appear somewhat easier than the other two tasks, the ANOVA indicated that there was no significant effect of the task type on participants' scores, $F(1.17, 35.13) = 3.68, p = .057$.

The results of this pilot study show the general easiness of the *Receptive Affective Prosody Task* for the adult participants regardless of the version considered. The finding of this ceiling effect is reassuring as it provides an anchoring point for children's performance. However, it also made it challenging to select the most appropriate version of the task to be included in the DA under development. To fulfil this aim, careful consideration was given to all evidence gathered from this pilot and the previous pilot study conducted with children. The *Congruent vs Incongruent Prosody* task was ruled out to begin with because, as previously discussed, younger children would likely focus primarily on the lexical content (e.g., Morton & Trehub, 2001), making judgements related to whether the sentences' emotional meanings are congruent or incongruent with the affective prosody more susceptible to distraction effects. Following this, given that the *Congruent vs Neutral Prosody* task received the lowest mean score in the adult pilot, this task was eliminated. The *Neutral Content* task was consequently selected, as this appeared the most straightforward and suitable to complete for the target age group of the DA.

3.2.3 Pilot testing 3: Receptive Affective Prosody Task (Neutral Content)

Pilot testing 3: Participant's characteristics and recruitment procedure

An in-person pilot of the latest version of the *Neutral Content Receptive Affective Prosody Task* and the rest of the DA (see Section 3.3.4) was carried out with a 7-year-old Arabic-English-speaking child. Although input from the child's parents could not be obtained, the teacher expressed no concerns about their language development and marked their English proficiency level as 'Stage C. Developing Competence'. This child was initially recruited for the main study of the DA but selected for the in-person pilot instead due to their inability to participate in the main study.

Pilot testing 3: Data collection procedure

The in-person pilot study of the *Neutral Content Receptive Affective Prosody Task* took place in the child's school in July 2022. Similarly to the previous pilot study with adults (see 'Pilot testing 2: Data collection procedure' within Section 3.2.2), in this task, the child had to match emotionally neutral sentences, spoken with emotional prosody that conveyed one of two emotions —happiness/sadness or anger/fear—, to pictograms that represented the expressed emotion. However, as reflected in the script in Appendix 13 and the scoring sheets in Appendix 14, this time the task was divided into two to create the post-test *Receptive Affective Prosody Task*, meaning that the pre- and post-test tasks (*Feelings Game 1* and *2*) would each consist of eight trials, rather than 16, to avoid overburdening and demotivating the children with a potentially long and repetitive activity. Accordingly, a maximum score of 8 points could be given for both tasks. This would also help shorten the time required to administer the entire DA, intended to fall under an hour to make the resource more amenable for practitioners to use, as excessive administration time compared to traditional language assessments is one of the practical challenges often preventing practitioners from making use of DA (Hasson & Joffe, 2007; Petersen et al., 2017). Additionally, a training phase identical to the one included at the start of the *Congruent vs Incongruent Prosody* task in the initial pilot study with children was incorporated to ensure understanding of the targeted emotions, as described in 'Pilot testing 1: Data collection procedure' under Section 3.2.1.

Pilot testing 3: Results and discussion

The child performed at the ceiling level in the pretest (8/8) and marginally worse in the post-test (6/8). While they responded well to the task and were generally focused throughout the session, their performance in the final stage was likely affected by distractions and fatigue due

to a particularly busy school day with occasional interruptions, ambient noise, and hot weather. Moving forward, plans to liaise with school staff were made to improve the session environment and provide more breaks if required without affecting the completion of the DA. Regarding the administration rules, it was decided that each sentence would be played once. However, a replay would be allowed if the first recording is interrupted, if the child is distracted and does not respond, or if they state that they could not hear it well the first time.

3.2.4 Conclusions

In summary, the first online pilot study of the *Receptive Affective Prosody Task* with six typically developing children with EAL indicated better performance with congruent stimuli (e.g., ‘These books are fun’ said with a happy voice) and an improved ability to identify incongruent cues with age, suggesting that younger children could struggle with the latter because they prioritise emotional lexical content over prosodic cues. Consequently, two additional task designs were created, as the *Congruent vs Incongruent Prosody* version seemed less reliable for measuring receptive affective prosody skills as an indicator of DLD across the young school-aged children targeted by the DA.

A second in-person pilot study was conducted to test all three versions of the *Receptive Affective Prosody Task* (*Congruent vs Incongruent Prosody*, *Congruent vs Neutral Prosody*, and *Neutral Content*) with 31 typically developed English-speaking adults. This demographic was chosen for their accessibility and to identify tasks that might be problematic for children. While they found all versions of the task easy, their performance was slightly lower in the *Congruent vs Neutral Prosody* task, signalling a subtle difficulty that could be magnified in a child population.

These observations led to the in-person trial of the *Neutral Content* task with one typically developing child with EAL, whose performance was likely affected by environmental distractions and fatigue. This prompted the need to improve the conditions in which the task is administered to allow for optimal performance. Considering the outcomes from all three pilot studies, this *Neutral Content* version was selected as the definitive form of the *Receptive Affective Prosody Task* for the main study of the DA, as its more straightforward nature was expected to aid in accurately judging receptive affective prosody skills to help detect whether a child could have DLD without the interference of potential confounders.

3.3 Development of the Storytelling Activity

The storytelling activity (*Storytime*) of the DA involves three phases (pretest, teaching, and post-test) in which wordless picture stories in English are used to teach and evaluate children's potential to learn narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and receptive affective prosody. During the pre- and post-tests, narrative macrostructure and emotional vocabulary are assessed through two parallel *Story Generation Tasks*. In the teaching phase, both narrative macrostructure and emotional vocabulary are targeted, alongside affective prosody.

Following the account of the development and piloting of the *Receptive Affective Prosody Task* in Section 3.2, this part of the chapter describes the materials (Section 3.3.1) and instructions (Section 3.3.2) prepared for the storytelling activity, as well as the piloting with children in Sections 3.3.3 and 3.3.4.

3.3.1 Materials

The materials for the storytelling activity consist of text-free illustrated narratives and additional visual aids (e.g., pictograms representing each story grammar unit). Focusing first on the former, three stories with a six-picture sequence were constructed based on Stein and Glenn (1979)'s *Story Grammar Model: Story A* (see Figure 13), *Story B* (Figure 14), and *Story C* (Figure 15). These stories were designed to present clearly defined plots and be equivalent in length, emotional vocabulary, macrostructure, and cognitive and linguistic complexity. Each story was tailored to highlight the story grammar units as various human and animal characters encountered problems to resolve.

The content of the stories —characters, objects, actions, and settings— was intended to be suitable for primary school-aged children's linguistic and cognitive skills, as well as cross-culturally appropriate and representative of universal knowledge typical to this age group. Efforts were made to clearly and explicitly depict distinct elements of the story corresponding to story grammar units, with characters portrayed in a readily recognisable manner using clear facial expressions and body language to express their actions and emotions. To facilitate focus on the narratives, the background illustrations in the pictures were kept simple and uncluttered while still providing enough details for children to comprehend the contexts and gain a sense of place and time.

Several wordless picture books and narrative assessment instruments were examined for inspiration to create the stories, such as *Frog, Where Are You?* (Mayer, 1969), the platform *Storybooks UK*, the *Edmonton Narrative Norms Instrument* (ENNI; Schneider et al., 2005), and the *Multilingual Assessment Instrument for Narratives* (MAIN; Gagarina et al., 2012, 2019). The decision was made to develop original stories rather than use existing ones so that they would align with the research aims more effectively and ensure they would not be overly demanding for the children's abilities, which could adversely impact their performance. The latter aligns with Peña et al. (2014b)'s reflection on using simple, supported story sequences to better reveal pretest–to–post-test differences.

Figure 13. Stimulus Picture Story A



Setting & Characters

A girl called Rita (or any other name given) is watering a plant. There is also a cat.
It's night time, and they are in the living room at their home.



Problem

Her cat Gus (or any other name given) jumps and knocks over the plant.
The plant pot breaks.



Emotion 1

Rita feels sad about it.



Attempt

Rita tries to fix the broken plant pot with glue. Rita manages to fix the plant pot.
Rita feels happy.

Note. The story was presented as a series of images without text to the children, with the text provided in this figure only serving as an illustrative example to the reader, similar to the verbal guidance offered to the children before discussing the story grammar units.

Figure 14. Stimulus Picture Story B



Setting & Characters

A man called Ali (or any other name given) is eating a sandwich. There is also a bird named Bob (or any other name given). It's daytime, and they are outside in a park.



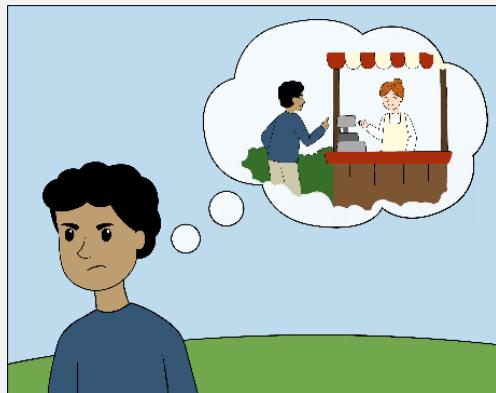
Problem

Bob flies by and steals the man's sandwich.



Emotion 1

Ali feels angry about it.



Plan

Ali then plans to go to the shop to buy another sandwich to replace the one that the bird stole.



Attempt

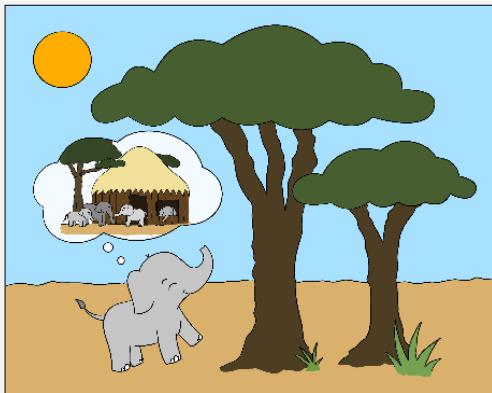
Ali goes to the shop to buy another sandwich.

Outcome & Emotion 2

Ali manages to buy a new sandwich. Ali feels happy.

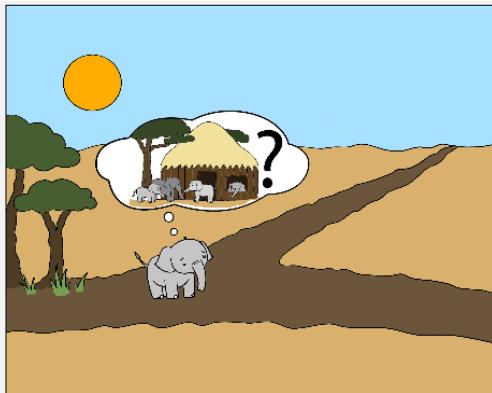
Note. The story was presented as a series of images without text to the children, with the text provided in this figure only serving as an illustrative example to the reader, similar to the verbal guidance offered to the children before discussing the story grammar units.

Figure 15. Stimulus Picture Story C



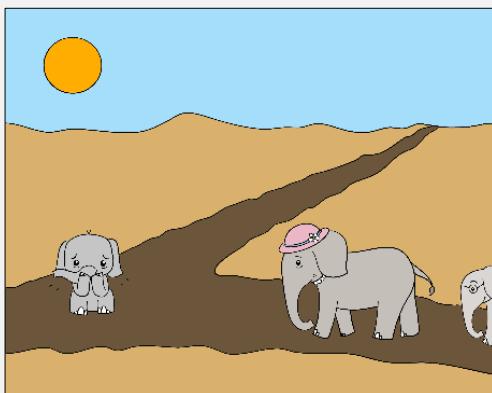
Setting & Characters

A little elephant called Momo (or any other name given) is walking home. It's daytime and she is in the savanna.



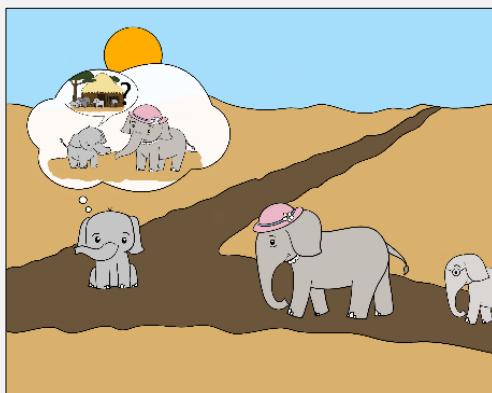
Problem

Momo realises that she is lost and doesn't know the way home.



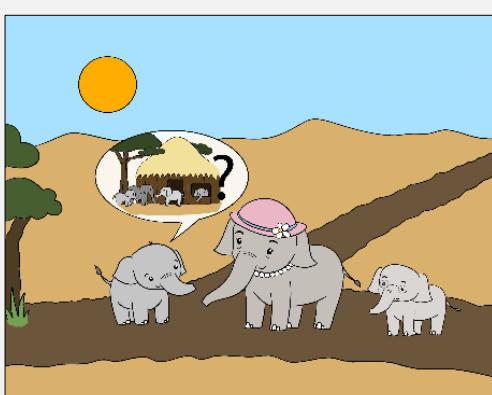
Emotion 1

Momo feels scared about it.



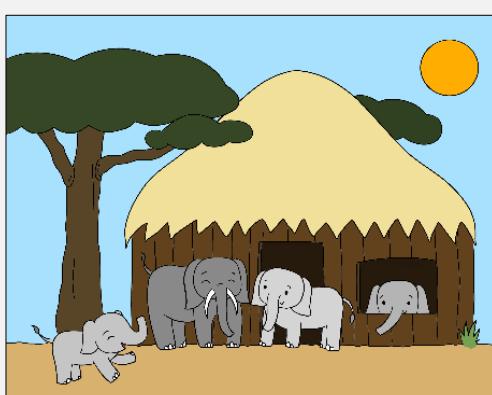
Plan

Momo sees an older elephant passing by and decides to ask her for directions to get back home.



Attempt

Momo asks the older elephant for directions back home.



Outcome

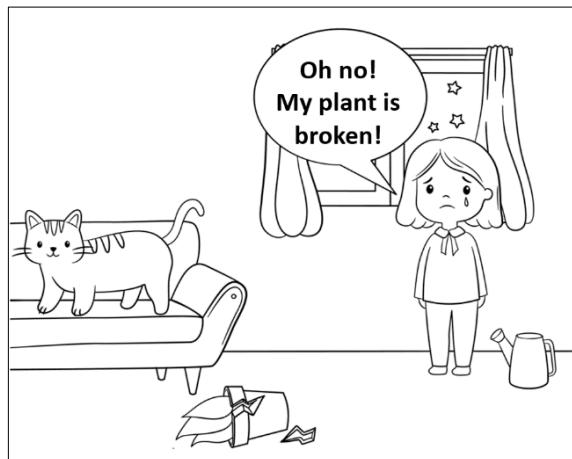
The older elephant tells Momo how to get home and she arrives home safely. Momo feels happy.

Note. The story was presented as a series of images without text to the children, with the text provided in this figure only serving as an illustrative example to the reader, similar to the verbal guidance offered to the children before discussing the story grammar units.

Three adult British English speakers reviewed the initial black-and-white pencil sketches that illustrated stories A, B, and C. The researcher invited them to contribute due to their understanding of children and refugee communities stemming from their personal and professional experiences, which included direct work and active involvement with these demographics. They commented on whether the stories were age and culturally appropriate, and easily understood. Based on this feedback and the observations from the online pilot study with children (see Section 3.3.3), various edits were made to improve the quality and clarity of these. During the processes of digitalising and colouring the drawings with an accessible palette, the assistance of experienced artists Alyona Chufistova and Clara Garrido-Tamayo was sought to achieve a higher-quality outcome. This would allow for the story elements to be portrayed in a more precise and visually appealing manner that would contribute to children's perception and enjoyment of the stories.

Other revisions included removing any speech bubbles with text (e.g., the girl in *Story A* saying, 'Oh no! My plant is broken!', as exemplified in Figure 16) to prevent children's varying reading skills from skewing the results and encourage them to make up the stories themselves.

Figure 16. Example of Picture 3 from Stimulus Story A Before Speech Bubble Was Removed



The more substantial alterations were made to *Story C* to enhance the clarity of the plot. Specifically, a thought bubble was included to show the little elephant thinking of home in picture 1, suggesting a journey back. Moving on from the static tree in the first drawings, a diverging path was added in pictures 2, 3, 4 and 5 to amplify the idea of her getting lost. Further, taking into account possible 'stranger danger'-related influences, the adult elephant in *Story C* was made to be accompanied by a younger elephant, fostering a sense of safety and

trust when the little elephant approaches them for directions. A family of elephants around a house was also incorporated to better represent the protagonist's home. This image first appears within thought bubbles in pictures 1, 2 and 4, where the little elephant ponders over her home and how to get back to this, having found herself lost. The image is then included in a speech bubble in picture 5, showing her asking for directions home. Picture 6 depicts the little elephant's home one last time as she reunites with her family. Efforts were made to make the thought bubbles in all stories as simple as possible.

In addition to the stimulus stories, the visual aids used during the storytelling activity involve a series of slides created through CANVA and PowerPoint. Along with illustrative stock images, these include graphic organisers and ARASAAC pictograms representing the overall story structure and each story grammar unit. As explained in Section 3.3.2, these visual aids and the story from the pretest are employed in the teaching phase to instruct on story grammar, as well as emotional vocabulary and affective prosody relevant to the story. After the online and in-person pilot studies (Sections 3.3.3 and 3.3.4), laminated paper copies of the graphic organisers and pictograms were produced for the main study (Appendix 15). This would reduce screen dependence, enhance the interactive, hands-on nature of the teaching activity, and strengthen children's comprehension and involvement in the learning process. These materials were included based on research supporting the use of physical referents for learning abstract concepts (Byrne et al., 2023) and pilot study observations of children's interest in interacting with toys like those from the NRDLS. Likewise, as described under 'B. Teaching Phase' in Section 3.3.2, a simple jigsaw puzzle is employed when encouraging children to reflect on what would happen if they told a story that was missing story grammar units, as this is compared with doing a jigsaw that is missing some pieces (Figure 17). When focusing on the story characters' emotional responses, the wheel in Figure 18 depicting the four targeted basic emotions is used (for further details, see 'B. Teaching phase', Section 3.3.2).

Figure 17. Jigsaw Puzzles Used During the Teaching Phase of the Storytelling Activity

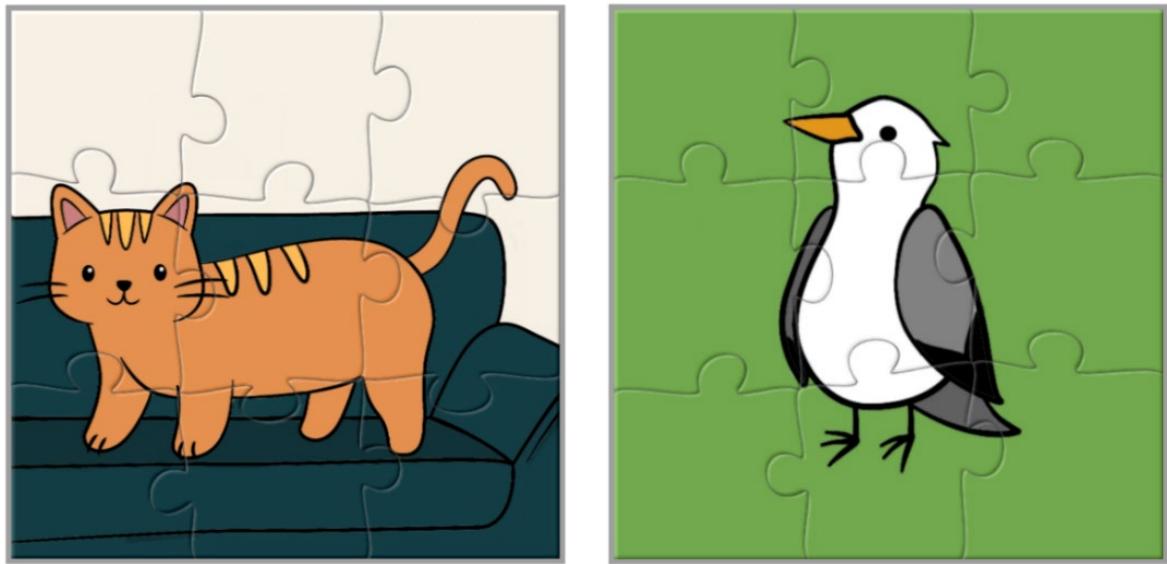
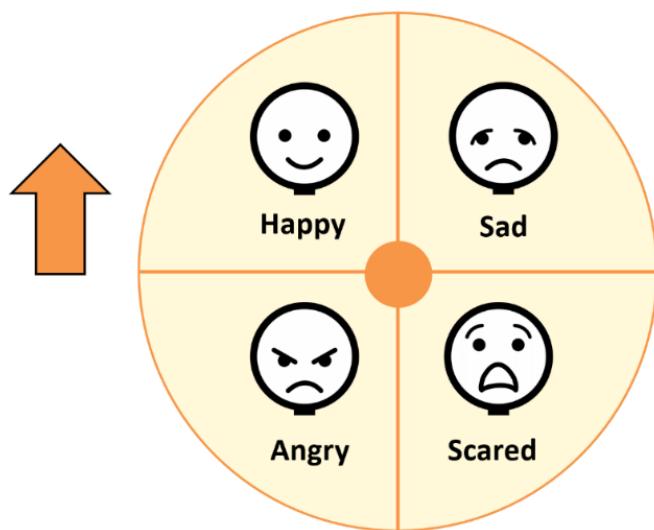


Figure 18. Feelings Wheel Used During the Teaching Phase of the Storytelling Activity



3.3.2 Instructions

The administration and scoring procedures for each phase of the storytelling activity (pretest, teaching, and post-test) are described in the following sections.

A. Pretest phase

The pretest phase of the storytelling activity involves assessing children's independent narrative creation skills through the *Story Generation Task*, which establishes a baseline measure of their unassisted competence in this domain before any teaching occurs.

As shown in the script in Appendix 16A, this phase begins with warm-up questions (e.g., ‘What is your favourite story?’). Following this, two mediating strategies are used in preparation for the teaching phase, during which the key strategies of a mediated learning experience—intentionality, meaning, transcendence, and competence (Lidz, 1991; Peña et al., 2001)—are incorporated (see Section 2.5.2 for further details). These include mediation of *intentionality* to state the purpose of the activity (‘Today we are going to learn how to tell good stories. You will become a storytelling superstar!’), *meaning* to state the relevance (‘A storytelling superstar can tell stories very well’), and *transcendence* to connect it with the child’s life experiences (‘Telling stories well is important because it helps us to communicate with friends and family, do great in school, and learn lots of things about life and the world’). These strategies are intended to stimulate children’s active engagement in the task and help them recognise its significance to their learning experience. Subsequently, children are asked to look through the pictures of *Story A* (Figure 13 in Section 3.3.1) or *Story B* (Figure 14 in Section 3.3.1), and then narrate without assistance based on those pictures while looking at them. The only cues used here are to prompt the children to continue if they fall silent (e.g., ‘Tell me more’).

Regarding the scoring of children’s performance, the audio or video recordings of the pre- and post-test *Story Generation Tasks* are used to transcribe their stories orthographically. Two scoring sheets are then completed to grade their narrative macrostructure abilities quantitatively. These focus on the presence of story grammar units and the complexity of the episodic structure in the children’s stories, in line with the ENNI (Schneider et al., 2005) and the MAIN (Gagarina et al., 2012, 2019), both of which are tools that have been used extensively to study narrative macrostructure in monolingual and multilingual child populations. A third scoring sheet is used to evaluate children’s emotional vocabulary knowledge pertinent to the protagonists’ emotional responses in their stories.

Scoring Sheet 1: Story Grammar Units

Three adaptations of *Scoring Sheet 1* were devised, and their usefulness was tested to gauge children’s inclusion of story grammar units in their narratives. Nine story grammar units are considered (characters, time, place, problem, emotion 1, plan, attempt, outcome, emotion 2) following Stein & Glenn (1979)’s classic model for analysing narrative macrostructure (see Section 2.4.1). This reflects the setting followed by an episode comprised of several components and centred around a goal to solve a problem.

As seen in Appendix 17, in the first version of *Scoring Sheet 1*, a point was granted for every story grammar unit that could be identified by a listener new to the story who could not see the pictures. The maximum score was 10 points. This first version of *Scoring Sheet 1* was used in both pilot studies and at the start of the data analysis for the main study. However, this seemingly straightforward binary 0 or 1-point approach proved ineffective in capturing children's varying English language abilities and overlooked the diverse ways they tried to express the story grammar units.

With the second version of *Scoring Sheet 1* (Appendix 18), a more nuanced approach involving a dual-scoring system was taken. The criterion for Scoring System 1 was whether a first-time listener unfamiliar with the story could understand the story grammar units based on the information the child provided overall, while the criterion for Scoring System 2 was whether the child had attempted to convey the story grammar unit. Within the first system, 0 points were given when the story grammar unit was not expressed or was so limited that a new listener would really struggle to understand; 1 point was given when the story grammar unit was partially expressed, making it somewhat difficult for a new listener to understand; and 2 points were given when the story grammar unit was fully expressed, allowing a new listener to understand clearly. Within the second scoring system, 0 points were given when the child did not make an evident attempt to convey the story grammar unit; 1 point was given when the child clearly tried to express the story grammar unit but was held back by their limited English language skills; and 2 points were given when the child accurately expressed the story grammar unit verbally in English. The maximum score was 40 points, with 20 points for each scoring system. This dual-scoring system offered a fairer, more comprehensive and precise evaluation of children's ability to structure a story, capturing a wider array of English language competencies. However, it proved to be overly intricate in practice.

The final *Scoring Sheet 1* shown in Tables 7 and 8 was created to preserve the thorough nature of the second scoring sheet while making it more user-friendly to ensure it could be applied consistently and efficiently. This would allow the number of story grammar units the child produces to be calculated using a single scoring system that still grants partial credit to recognise varying performance degrees and acknowledge efforts that demonstrate some achievement. With a maximum score of 30 points, the child can receive between 0 and 3 points for each story grammar unit according to the following scoring criteria:

- 0 points: The child does not express any information about the story grammar unit.
- 1 point: The child attempts to convey the story grammar unit but struggles to express it entirely due to English language proficiency limitations. These attempts include using pronouns when the lexical noun is unknown (e.g., ‘she’ instead of ‘girl’) or general all-purpose verbs rather than more specific or specialised verbs (e.g., ‘get’ in place of ‘steal’). Other common examples are the child asking for the word, gesturing or substituting words from their home languages.
- 2 points: The child refers to the story grammar unit verbally in a more elaborate manner but still does not express it fully.
- 3 points: The child expresses the story grammar unit verbally in a highly comprehensible and comprehensive manner.

Further information and scoring rules complementing the general scoring criteria above can be found below for each story grammar unit:

1. *Characters 1 and 2.* Both the protagonist and the other main character in the story are considered. Three points are given when the characters are introduced with an appropriate noun, regardless of where they are first mentioned, even if this is later in the story. If only pronouns are used to refer to the characters, 2 points are given for consistently accurate use of pronouns (e.g., ‘she’ to refer to the girl in *Story A*). In contrast, 1 point is awarded for occasional inaccuracies in pronoun usage (e.g., ‘she’ to refer to the man in *Story B*).
2. *Time.* Points are awarded based on the clarity and detail of the reference to the time when the story’s events occur. A highly explicit and complete reference (e.g., ‘It’s night time.’) receives 3 points. A less explicit and complete mention (e.g., ‘The moon is shining.’) obtains 2 points. An even more vague, less elaborate reference (e.g., ‘A star.’) is awarded 1 point.
3. *Place.* Points are given based on the clarity and detail of the reference to the place where the story’s events are set. A highly explicit and complete reference (e.g., ‘He’s in the park’) earns 3 points. A less explicit and complete mention (e.g., ‘The man sits on a bench.’) receives 2 points. An even more vague, less elaborate reference (e.g., ‘The bench is here.’) is granted 1 point.

4. *Problem*. Credit is awarded for the complication that drives the story's plot and triggers some response from the protagonist.
5. *Internal response: Emotion 1*. Credit is given for the protagonist's emotional response to the problem: 'sad' for *Story A*, 'angry' for *Story B*, or 'surprised' for both stories. References to emotional expressions (e.g., 'crying') are also credited. Likewise, points are given for emotions that do not reflect what the protagonists feel according to the pictures but still fit in the story's context (e.g., the girl being angry rather than sad when the cat breaks the plant pot in *Story A*).
6. *Internal response: Plan*. Credit is granted for the protagonist's idea targeted towards a goal to solve the problem. This includes goal-orientated language (e.g., 'plans to', 'decides to').
7. *Attempt*. Credit is awarded for the protagonist's action to achieve their goal of solving the problem. This includes action-orientated language (e.g., 'tries to', 'goes to').
8. *Outcome*. Credit is given for the result of the protagonist's attempt to solve the problem.
9. *Reaction: Emotion 2*. Credit is given for the protagonist's emotional response to the outcome, which is 'happy' for both stories. Emotional expressions (e.g., 'smiling') also receive points.

Table 7. Story A: Scoring Sheet 1. Story Grammar Units – Version 3

Story A: Scoring Sheet 1. Story Grammar Units		Score	Notes
Unit	Response	0 1 2 3 /3	
Character 1	<p>1 point: Exclusive and occasionally inaccurate use of pronouns to refer to the character</p> <p>2 points: Exclusive and consistently accurate use of pronouns to refer to the character</p> <p>3 points: Use of an appropriate noun to refer to the character at any point in the story, such as:</p> <ul style="list-style-type: none"> - Girl/Woman (or synonym) - Any proper name given - Also acceptable: Sister (or other family member name) 		
Character 2	<p>1 point: Exclusive and occasionally inaccurate use of pronouns to refer to the character</p> <p>2 points: Exclusive and consistently accurate use of pronouns to refer to the character</p> <p>3 points: Use of an appropriate noun to refer to the character at any point in the story, such as:</p> <ul style="list-style-type: none"> - Cat/Kitten (or synonym) - Any proper name given 		

Unit	Response	Score					Notes
Time	<p>1 point: Highly implicit and incomplete reference to the time, such as 'A star.'</p> <p>2 points: Implicit and incomplete reference to the time, such as 'The moon is shining.'</p> <p>3 points: Highly explicit and complete reference to the time, such as:</p> <ul style="list-style-type: none"> - Night time/Evening/Dark (or synonym) - Also acceptable: Bedtime/After dinner 	0	1	2	3	/3	
Place	<p>1 point: Highly implicit and incomplete reference to the place, such as 'The sofa is here.'</p> <p>2 points: Implicit and incomplete reference to the place, such as 'She's next to the sofa.'</p> <p>3 points: Highly explicit and complete reference to place, such as:</p> <ul style="list-style-type: none"> - Living room (or synonym) - Also acceptable: Home/House (or synonym) 	0	1	2	3	/3	
Problem	<p>Cat breaks/knocks over the plant/plant pot (or synonyms)</p> <p>The plant/plant pot breaks/falls (or synonyms)</p>	0	1	2	3	/3	
Emotion 1	<p>Sad (or synonym)</p> <p>Also acceptable: Crying/Surprised/Angry (or synonyms)</p>	0	1	2	3	/3	
Plan	<p>The girl/she plans* to fix/glue the plant/plant pot (or synonyms)</p> <p>*Examples of accepted goal-orientated language: decides/wants to/thinks she will</p>	0	1	2	3	/3	
Attempt	<p>The girl/she tries to* fix/glue the plant/plant pot with glue (or synonyms)</p> <p>*Examples of accepted action-orientated language: goes to/is going to/is glueing/glues</p>	0	1	2	3	/3	
Outcome	<p>The plant/plant pot is fixed/glued/fine (or synonyms)</p> <p>The girl/she fixed/has finished fixing the plant/plant pot (or synonyms)</p>	0	1	2	3	/3	
Emotion 2	<p>Happy (or synonym)</p> <p>Also acceptable: Smiling (or synonym)</p>	0	1	2	3	/3	
Story grammar units score							/30

Table 8. Story B: Scoring Sheet 1. Story Grammar Units – Version 3

Story B: Scoring Sheet 1. Story Grammar Units						
Unit	Response	Score			Notes	
Character 1	<p>1 point: Exclusive and occasionally inaccurate use of pronouns to refer to the character</p> <p>2 points: Exclusive and consistently accurate use of pronouns to refer to the character</p> <p>3 points: Use of an appropriate noun to refer to the character at any point in the story, such as:</p> <ul style="list-style-type: none"> - Boy/Man (or synonym) - Any proper name given - Also acceptable: Dad (or other family member name) 	0	1	2	3	/3
Character 2	<p>1 point: Exclusive and occasionally inaccurate use of pronouns to refer to the character</p> <p>2 points: Exclusive and consistently accurate use of pronouns to refer to the character</p> <p>3 points: Use of an appropriate noun to refer to the character at any point in the story, such as:</p> <ul style="list-style-type: none"> - Bird/Seagull (or synonym) - Any proper name given 	0	1	2	3	/3

Unit	Response	Score	Notes
Time	<p>1 point: Highly implicit and incomplete reference to the time, such as 'The sun.'</p> <p>2 points: Implicit and incomplete reference to the time, such as 'The sun is shining.'</p> <p>3 points: Highly explicit and complete reference to the time, such as:</p> <ul style="list-style-type: none"> - Daytime/Morning/Afternoon (or synonym) - Also acceptable: Lunchtime 	0 1 2 3	/3
Place	<p>1 point: Highly implicit and incomplete reference to the place, such as 'The bench is here.'</p> <p>2 points: Implicit and incomplete reference to the place, such as 'The man sits on a bench.'</p> <p>3 points: Highly explicit and complete reference to the place, such as:</p> <ul style="list-style-type: none"> - Park/Garden (or synonym) - Also acceptable: Outdoors/Outside (or synonym) 	0 1 2 3	/3
Problem	<p>Bird steals/takes the sandwich (or synonyms)</p> <p>The sandwich gets taken/stolen (or synonyms)</p>	0 1 2 3	/3
Emotion 1	<p>Angry (or synonym)</p> <p>Also acceptable: Shouting/Surprised/Sad (or synonyms)</p>	0 1 2 3	/3
Plan	<p>The man/he plans* to buy another sandwich (or synonyms)</p> <p>*Examples of accepted goal-orientated language: decides/wants to/thinks he will</p>	0 1 2 3	/3
Attempt	<p>The man/he goes to* the shop to buy another sandwich (or synonyms)</p> <p>*Examples of accepted action-orientated language: is going to/tries to/is buying/buys</p>	0 1 2 3	/3
Outcome	<p>The man/he bought/has another/a new sandwich (or synonyms)</p> <p>The man/he eats the new sandwich (or synonyms)</p>	0 1 2 3	/3
Emotion 2	<p>Happy (or synonym)</p> <p>Also acceptable: Smiling (or synonym)</p>	0 1 2 3	/3
Story grammar units score			/30

Scoring Sheet 2: Episodic Structure Complexity

The complexity of the episode's structure in children's stories was evaluated by prioritising the plan-attempt-outcome sequence in accordance with Westby (2005)'s *Story Grammar Decision Tree* (see Section 2.4.1). As shown in *Scoring Sheet 2* (Tables 9 and 10), children's stories are classified into one of four levels:

- Level 1. Description: The story consists of isolated descriptions of objects, characters, surroundings, or actions without a temporally related series of events. An attempt or the outcome may be mentioned, but they are presented in a descriptive manner, as events that happened without any explanation of connection to the protagonist's plan.
- Level 2. Sequence: The story includes a series of events that follow a sequence in time and may be linked by cause and effect but lacks goal-directed behaviour. While there are attempt and outcome statements, the protagonist's plan is not mentioned.
- Level 3. Incomplete episode: The story contains the protagonist's plan but is missing either an attempt to achieve it or the outcome.
- Level 4. Complete episode: The story includes a plan, an attempt, and an outcome that clearly shows the protagonist's goal-directed behaviour.

Table 9. Story A: Scoring Sheet 2. Episodic Structure Complexity

Story A: Scoring Sheet 2. Episodic Structure Complexity			
Structure complexity level	Description	✓	Notes
1. Description	Attempt, Outcome		
2. Sequence	Attempt + Outcome		
3. Incomplete episode	Plan, Plan + Attempt, Plan + Outcome		
4. Complete episode	Plan + Attempt + Outcome		
Episodic structure complexity level			/4

Table 10. Story B: Scoring Sheet 2. Episodic Structure Complexity

Story B: Scoring Sheet 2. Episodic Structure Complexity			
Structure complexity level	Description	✓	Notes
1. Description	Attempt, Outcome		
2. Sequence	Attempt + Outcome		
3. Incomplete episode	Plan, Plan + Attempt, Plan + Outcome		
4. Complete episode	Plan + Attempt + Outcome		
Episodic structure complexity level			/4

Scoring Sheet 3: Emotional Vocabulary

Scoring Sheet 3 (see Tables 11 and 12) is used to assess children's emotional vocabulary knowledge in the contexts of the pre- and post-test stories. More specifically, the focus is on the vocabulary pertinent to the protagonists' emotional responses to the problem and the outcome, with a maximum score of 4 points.

This way, children receive 2 points if they use specific terms or expressions indicating the protagonists' emotional responses to the problem (e.g., 'sad', 'heartbroken') and outcome (e.g., 'happy', 'over the moon'). One point is given for more general emotional expressions (e.g., crying) or emotions that do not precisely match the protagonists' feelings towards the problem as shown in the pictures, but which remain contextually fitting (e.g., the girl being angry when the plant pot breaks). Only emotions relevant to the problems and outcomes are considered, with feelings described elsewhere in the stories not given credit.

Table 11. Story A: Scoring Sheet 3. Emotional Vocabulary – Version 2

Story A: Scoring Sheet 3. Emotional Vocabulary			
Corresponding story grammar unit	Response	Score	Notes
Emotion 1	0 points: No emotion or irrelevant emotion 1 point: Crying (or another relevant emotional expression)/Angry (or synonym) 2 points: Sad/Surprised (or synonyms)	0 1 2	
Emotion 2	0 points: No emotion or irrelevant emotion 1 point: Smiling (or another relevant emotional expression) 2 points: Happy (or synonym)	0 1 2	
Emotional vocabulary score (sum of 1 to 2)		/4	

Table 12. Story B: Scoring Sheet 3. Emotional Vocabulary – Version 2

Story B: Scoring Sheet 3. Emotional Vocabulary			
Corresponding story grammar unit	Response	Score	Notes
Emotion 1	0 points: No emotion or irrelevant emotion 1 point: Shouting (or another relevant emotional expression)/Sad (or synonym) 2 points: Angry/Surprised (or synonyms)	0 1 2	
Emotion 2	0 points: No emotion or irrelevant emotion 1 point: Smiling (or another relevant emotional expression) 2 points: Happy (or synonym)	0 1 2	
Emotional vocabulary score (sum of 1 to 2)		/4	

In its initial version, *Scoring Sheet 3* allocated a maximum score of 2 points (see Appendix 19). One point was awarded for identifying ‘sad’ and ‘angry’ or synonyms as Emotion 1 in stories A and B, respectively, and for ‘happy’ or an equivalent as Emotion 2 in both stories. Moreover, for *Story C*, applied during the online pilot study, one point was assigned for ‘scared’ (Emotion 1) and another for ‘happy’ (Emotion 2). However, this binary scoring was revised to align it with the updated *Scoring Sheet 1* and reflect children’s diverse English language skills and response styles.

B. Teaching phase

Following the pretest phase, children participate in a short teaching session focusing on the story used in the pretest (*A* or *B*) that integrates the main strategies of a mediated learning experience (Lidz, 1991; Peña et al., 2001). Such strategies target the learning of story grammar, intending to help children learn to tell complete stories, as well as emotional vocabulary and affective prosody contextualised within these stories. The script in Appendix 16B was designed following Peña et al. (2006, 2014b)’s English narrative DA approach. Its application facilitates the effective administration of the teaching session. It also ensures adherence to the assessment protocol and consistency across participants while still allowing some flexibility for feedback and support to be adjusted to each child’s individual learning needs.

As recommended by Spencer & Petersen (2020), the structure of the teaching phase follows a whole–part–whole framework where the story is modelled for the child before exploring each story grammar unit one by one and reconstructing the story together in its entirety at the end.

Once the child finishes telling the pretest story, the researcher praises their efforts irrespective of their performance (*competence*) and informs them that they will work together to practice telling this story and learn about all the parts every story should include (*intentionality*). An overview of the story grammar units and their placement within the story's beginning, middle, and end is subsequently presented using pictograms and graphic organisers. At this point, the pretest story is modelled for the child, highlighting the importance of including all parts for the story to be complete and make sense (*meaning*). Each story grammar unit is then individually explored, beginning with the setting and the characters' information and continuing with the episode components. Before recapping what has been learned with the help of the visual aids previously employed (*competence*), the relevance of mentioning all story parts is reinstated. The jigsaw puzzle presented earlier in Figure 17 (Section 3.3.1) is used at this moment to stimulate children's reflection on the consequences of telling an incomplete story, likening it to a jigsaw lacking pieces (*transcendence*). To conclude the teaching phase, the researcher and the child reconstruct the story in its entirety, and the child is reminded of the need to practice telling stories at home and in school and to include the various elements when doing so (*transcendence* and *competence*). Before progressing to the post-test phase, the child is invited to colour in the second star on the star points chart (Appendix 1), representing the completion of the initial story, and a short break ensues.

Notably, the instruction of the protagonist's emotional reactions to the pretest story's problem and outcome serves as an opportunity to educate the child about the relevant emotional vocabulary. This is achieved through inferential comprehension questions about the characters' emotions and reflections that draw parallels with relatable instances from the child's own experiences. To support this process, the wheel previously shown in Figure 18 (Section 3.3.1) illustrating the four targeted basic emotions is employed as an additional visual aid, enabling the child to select the character's emotional state. Furthermore, this context allows for a connection to the learning acquired in the pretest *Receptive Affective Prosody Task*, reinforcing and directing the child's focus toward specific prosodic cues linked with each target emotion. A brief activity addressing receptive affective prosody is integrated at this point, where the child listens to two audio clips featuring the same sentence spoken with prosody that is either congruent or incongruent with the characters' emotions. Subsequently, the child is asked to identify which voice best represents the character in question and invited to imitate what the character says by putting on a voice that expresses the relevant emotion.

With regards to the prompting used, a two-step prompting procedure following Spencer & Petersen (2020) is applied when exploring each story grammar unit: the child is first asked a question to direct their attention to the specific story grammar unit that should be said (e.g., ‘What is the problem in this story? What happens to (Rita)’s plant?’). If the child responds correctly within a few seconds, their response is praised, recasted, and expanded (e.g., if they answer, ‘Broken plant’, we would say, ‘That’s right! The cat knocks over the plant pot and it breaks.’). On the contrary, if the child cannot tell the expected target in a timely manner, a model of what they should say is provided with a consequent opportunity for them to try again (e.g., ‘The cat knocks over the plant pot and it breaks. Now you say it.’). In addition, for the recap at the end of the teaching phase, increasingly supportive prompting questions are asked depending on the child’s difficulty in responding. For example, if they are unable to respond to ‘How does (Rita) feel about the problem?’, the next question would be ‘How does (Rita) feel when she sees that her plant pot is broken?’ (wait for response) followed by ‘Is she happy or sad?’ (wait for response), and, finally, ‘(Rita) feels sad because (Gus) the cat broke her plant pot.’.

Modifiability Rating Scale

The rating scale in Table 13 is used to record children’s modifiability or capacity to respond to instruction based on the assessor’s observations during the teaching phase. The scale features six items that capture three critical aspects to be assessed from 0 (‘Never’) to 2 (‘Often’) based on their frequency of occurrence: responsiveness (items 1–4), transfer of learning (item 5), and assessor effort (item 6). These items were judged essential based on modifiability measures from previous DA research with multilingual children who have DLD (Peña et al., 2014; Petersen et al., 2020).

In the *Responsiveness* category, items 3 and 4, which are associated with a child’s compliance and task orientation, were found to be significant factors in Peña et al. (2014)’s study for detecting DLD among English language learners. Metacognition was also important in this regard, which initially led to including an item related to error awareness. However, scoring this proved problematic as high-performing children who make fewer evident errors could unfairly receive lower scores due to limited chances to demonstrate error awareness. There was consideration to widen this item’s scope to other behaviours indicative of the child’s performance awareness. However, these might generally overlap with the *Transfer of learning* item, posing a risk of score inflation due to double-scoring.

Table 13. Modifiability Rating Scale

Modifiability Rating Scale	Never	Sometimes	Often
	0	1	2
Responsiveness			
1. The child acted on prompts and guidance provided.			
2. The child was attentive and focused on the task.			
3. The child cooperated and followed the task instructions.			
4. The child fully understood the task.			
Transfer of learning			
5. The child showed the ability to apply newly learned knowledge as the teaching session progressed.			
Assessor effort			
6. Minimal effort was required to facilitate the child's learning.			
Modifiability score (sum of 1 to 6):		/12	
Notes:			

C. Post-test phase

In the post-test phase of the storytelling activity, a parallel form of the pretest *Story Generation Task* is carried out using a new story (*Story A* or *B*) to assess children's independent performance again when telling a story without assistance following the teaching provided. The script and scoring sheets for this stage are near-identical to those from the pretest, with adjustments made to account for the different story used (see Appendix 16C).

3.3.3 Pilot testing 1: Storytelling activity

Pilot testing 1: Participants' characteristics and recruitment procedure

The storytelling activity—which integrated the pre- and post-test *Story Generations Tasks*, and the teaching phase targeting narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and affective prosody—was piloted with six children with EAL residing in the UK. These same children had also participated in the online pilot study of the pretest *Receptive Affective Prosody Task*. Their ages ranged from 5;06 to 12;11 years, and they exhibited typical language development, with further details about their characteristics and the participant recruitment procedure available in Section 3.2.1.

Pilot testing 1: Data collection procedure

The pilot study of the storytelling activity ran between June and December 2021 through the same 45-60-minute game-based Zoom session centred on becoming a ‘storytelling superstar’ described in ‘Pilot testing 1: Data collection procedure’ (Section 3.2.1), where a parent was present with the child as a facilitator. Along with the pretest *Receptive Affective Prosody Task*, the storytelling activity was delivered via PowerPoint using Zoom’s screen-share function.

The original black-and-white pencil sketches of stories A, B, and C were used for the first five children and their digitalised versions for the sixth child. These digitalised drawings were still uncoloured, and although they were more highly defined due to the transition from pencil to a digital medium and the speech bubbles had been removed, the content remained fundamentally the same as those used with the previous participants. As explained in Section 3.3.1, changes included removing speech bubbles containing text and refining the plot’s clarity in *Story C* about the elephant (see Figure 15 in Section 3.3.1) through some revisions.

For the storytelling activity procedure, a preliminary, extended version of the final one described in Section 3.3.2 was employed. This included a second teaching cycle during which the researcher worked with the child to co-construct *Story B*, a new narrative different from *Story A*, which was used in the pretest and first teaching cycle. *Story C* was used in the post-test. A similar procedure was applied during both teaching cycles to foster a mediated learning experience, with the second slightly briefer. At this stage, the whole–part–whole instructional sequence had yet to be introduced for the teaching phase, and a more flexible prompting approach was used that lacked the specificity of the final two-step method, particularly regarding how to react to children’s possible responses when working on each story grammar unit. Also, rather than asking increasingly supportive prompting questions during the recap at the end of the teaching cycles, open discussion about each story grammar unit was encouraged. This initial prompting method gave the assessor more room for individual interpretation, hence needing further refinement to improve the consistency of its application.

The teaching of the emotional vocabulary and affective prosody targets within the story contexts slightly differed from that of the in-person pilot and main study (see Section 3.3.2). Aside from inferential questions on the characters’ emotions and relating these to the children’s experiences, this involved revisiting the definitions of the emotions provided in the pretest *Receptive Affective Prosody Task*, and overstressing the prosodic cues associated with

each emotion to show its correct use and prompt children to imitate these (e.g., You could say, ‘Rita felt sad and said, “Oh no! My plant is broken!” [sad vocal expression] Now, you give it a go – what do you think Rita is saying here?).

For scoring purposes, the first versions of *Scoring Sheet 1: Story Grammar Units* (maximum score: 10 points) and *Scoring Sheet 3: Emotional Vocabulary* (maximum score: 2 points) were used to analyse how children incorporated the story grammar units and emotional vocabulary in their pre- and post-test stories, along with *Scoring Sheet 2: Episodic Structure Complexity* (4 possible levels) to evaluate the complexity of the episodic structures. The development of these scoring sheets is detailed in Section 3.3.2.

Pilot testing 1: Results and discussion

The following subsections in ‘Pilot testing 1: Results and discussion’ (Section 3.3.3) detail the children’s pre- and post-test performance in the *Story Generation Task* within the DA’s storytelling activity. Only four of the six children who participated in the online pilot study were included in this analysis. Their ages ranged from 5;09 to 10;04 years ($M = 8;00$, $SD = 1;11$) (for further details, see ‘Pilot Testing 1: Participants’ characteristics and recruitment procedure’ in Section 3.2.1). The scores of the youngest child (5;06) and the oldest child (12;11) were examined separately. The reason for this is that, for the youngest child, the session had to be concluded early due to their growing inattentiveness as the second teaching cycle neared its end, which aligns with their *Modifiability Rating Scale* of 6/12. Even so, in the pretest, this child showed emergent narrative macrostructure and emotional vocabulary skills, with scores of 5/10 in *Scoring Sheet 1*, 3/4 in *Scoring Sheet 2*, and 2/2 in *Scoring Sheet 3*.

On the other hand, the oldest child did not take part in the teaching phase as it was deemed excessively simple for them. Instead, they were invited to narrate the three stories and share any reflections that could inform the ongoing development of these. The child told the three stories as would be expected based on the illustrations. They stated that they enjoyed these and suggested the addition of colour to make them more compelling.

Scoring Sheet 1: Story Grammar Units

The first version of *Scoring Sheet 1* (see Section 3.3.2 and Appendix 17) was applied in this pilot study. This sheet assigned a point for each of nine story grammar units identified in a child’s narrative: characters 1 and 2, time, place, problem, emotion 1, plan, attempt, outcome, and emotion 2. The highest attainable score was 10 points at both the pre- and post-tests.

Table 14 shows that the mean pretest score was 8.8 ($SD = 0.5$, range = 1), indicating high initial performance with limited variability in story grammar unit usage among the four children. The mean post-test score was 6.8 ($SD = 2.2$, range = 5), reflecting decreased scores and increased variability after the teaching phase.

Table 14. Descriptive Statistics for the Children's Performance in Scoring Sheet 1: Story Grammar Units (Story Generation Task) in the Online Pilot Study

Scoring Sheet 1: Story Grammar Units	N	Range	Minimum	Maximum	M	SD
Pretest score	4	1	8	9	8.8	0.5
Post-test score	4	5	4	9	6.8	2.2
Change score	4	4	-4	0	-2.0	1.8

This performance decline reflects the limited potential for substantial post-teaching improvements due to the high baseline scores. In fact, post-test scores were lower for three children, and for the two youngest children, their unique interpretation of *Story C*'s plot played a role in this regard. Specifically, their view of the adult elephant as a member of the little elephant's family looking for her, rather than an unrelated passerby whom she asks for directions home, affected their ability to earn points for several story grammar units, such as the Plan (e.g., 'The elephant is thinking about her family finding her.' instead of 'The elephant is thinking of asking the older elephant for directions home'). This compromised the reliability of the pre- and post-test performance comparison in *Scoring Sheet 1*.

Scoring Sheet 2: Episodic Structure Complexity

Scoring Sheet 2 (see Section 3.3.2) was used in this pilot study to assess the complexity of the episodic structure within a child's narrative. This consisted of a 4-level ordinal scale based on the plan-attempt-outcome sequence: Level 1: 'Description', 'Level 2: Sequence', 'Level 3: Incomplete Episode', and 'Level 4: Complete Episode'.

Table 15 shows that the four children's pretest narratives uniformly reached the highest complexity level, corresponding to a complete episode (*Median* = 4, range = 0). However, after the teaching phase, there was a performance decline and the emergence of variability (*Median* = 3, range = 3). This was particularly evident for the two youngest children due to the interpretative issues concerning *Story C* mentioned in the previous subsection about the *Scoring Sheet 1* results.

Table 15. Descriptive Statistics for the Children’s Performance in Scoring Sheet 2: Episodic Structure Complexity (Story Generation Task) in the Online Pilot Study

Scoring Sheet 2: Episodic Structure Complexity	N	Range	Minimum	Maximum	Median	Mode	M	SD
Pretest score	4	0	4	4	4	4	4.0	0.0
Post-test score	4	3	1	4	3	4	2.8	1.5
Change score	4	3	-3	0	-1	0	-1.3	1.5

Scoring Sheet 3: Emotional Vocabulary

Scoring Sheet 3 (see Section 3.3.2 and Appendix 19) was employed in this pilot study to evaluate the use of emotional vocabulary in a child’s narrative. A maximum of 2 points was possible: one for identifying ‘sad’ and ‘scared’ (or synonyms) as Emotion 1 in stories A and C, respectively, and another for identifying ‘happy’ (or synonyms) as Emotion 2 in both stories.

Table 16 reports a mean pretest score of 1.5 ($SD = 0.6$, range = 1), reflecting high and consistent performance on emotional vocabulary usage among the four children. The post-test mean score was 1.0 ($SD = 0.8$, range = 2), signalling reduced performance and more variability. Notably, in three of the children’s post-test stories, Emotion 2 was omitted. In two cases, this omission appeared to be linked to parental interruptions, such as parents interjecting their own interpretation of the story, which likely disrupted the children’s thought processes.

Table 16. Descriptive Statistics for the Children’s Performance in Scoring Sheet 3: Emotional Vocabulary (Story Generation Task) in the Online Pilot Study

Scoring Sheet 3: Emotional Vocabulary	N	Range	Minimum	Maximum	M	SD
Pretest score	4	1	1	2	1.5	0.6
Post-test score	4	2	0	2	1.0	0.8
Change score	4	2	-1	1	-0.5	1.0

Modifiability Rating Scale

The *Modifiability Rating Scale* (see Section 3.3.2) was trialled in this pilot study to gauge how well each child responded during the DA’s teaching phase, as well as the level of support required. Six items were scored from 0 (‘Never’) to 2 (‘Often’) across three domains: responsiveness, transfer of learning, and assessor effort. The highest possible score was 12 points.

As shown in Table 17, the four children's mean score was 10.5 ($SD = 1.9$, range = 4), reflecting an overall high capacity to respond to instruction and integrate new knowledge during the teaching phase, despite some variability. On average, it also seems that a low amount of assessor effort was needed to facilitate learning for these children.

Table 17. Descriptive Statistics for the Children's Performance in the Modifiability Rating Scale in the Online Pilot Study

	N	Range	Minimum	Maximum	M	SD
Modifiability Rating Scale	4	4	8	12	10.5	1.9

Pilot testing 1: Reflections on online vs in-person data collection for the main study

Other issues impacted the children's performance in the *Story Generation Task* besides those mentioned above. As the storytelling activity progressed, their focus generally declined, with signs of restlessness becoming apparent. These included more frequent movement, fidgeting with objects, or turning away from the screen to glance at, or interact with, other things and people in the room. Breaks were encouraged as needed, and parents were asked in advance to ensure that the session took place in a distraction-free environment. Despite these measures, the researcher's limited control over disruptions in the children's surroundings and other noticeable factors, such as the late timing of some sessions due to families' limited availability and children being slightly unwell with an incipient cold or lacking computer or desk experience, contributed to their performance during the session.

In light of these observations from the online pilot study, the pros and cons of continuing with remote data collection versus transitioning to in-person data collection for the main study were weighed. The advantages of remote data collection included its expanded geographical reach, enabling the inclusion of a more extensive and diverse sample of children with EAL from across the UK. It also removed the need for COVID 19-related risk assessments and personal protective equipment, which could be costly and time-consuming.

However, despite these benefits, remote data collection posed increased challenges in maintaining children's focus and engagement during the sessions, especially with younger participants. It also made building rapport with them and handling disruptions in their environment more difficult. Possible technical difficulties, internet connection issues, and parental interference with children's performance, such as giving unsolicited prompts or answers, could also affect the reliability of the data collected. At the same time, the

requirement for a facilitator to be present during the session where the comparative measures (e.g., NRDLS) would be administered, in addition to the DA session, could overburden parents and educators already managing busy schedules. Gathering data in person would also ease the burden on parents still developing their English proficiency to serve as intermediaries, helping to circumvent language barrier challenges.

Additionally, the use of direct online equivalents of the comparative measures was planned, such as adjusting the NRDLS and CL-NWRT for remote delivery, subject to the authors' approval. In the case of the NRDLS, the *Language Intervention in the Early Years* (LIVELY) project research team at Newcastle University (for further details, see <https://research.ncl.ac.uk/lively>), who were also exploring the possibility of remote NRDLS administration, advised using software like PowerPoint and Miro for image screen-sharing on Zoom and pre-recording videos of the researcher doing the relevant actions with the objects. For parts involving the child moving toys, options such as posting materials to them beforehand or using online alternatives were considered. However, these online adaptations, being untested, raised reliability concerns and were time-consuming to prepare, adding to the drawbacks of remote data collection.

With these factors in mind, a decision was made to adjust the research methods further and adopt the in-person approach for the main study as it became viable within the fieldwork period. This adjustment followed UK Government guidelines and involved conducting the risk assessment in Appendix 20. The in-person approach, initially intended before the pandemic, was recognised as a method that would meet the needs of the studied population more successfully and allow the collection of higher-quality data. Accordingly, additional piloting was undertaken with a child at their school, as detailed in Sections 3.2.3 (*Receptive Affective Prosody Task*) and 3.3.4 (*Storytelling Activity*).

3.3.4 Pilot testing 2: Storytelling activity

Pilot testing 2: Participant's characteristics and recruitment procedure

An in-person pilot study of the storytelling activity and the rest of the DA was carried out with the same 7-year-old Arabic-English-speaking, typically developing child mentioned in Section 3.2.3.

Pilot testing 2: Data collection procedure

The in-person pilot study of the storytelling activity took place in the child’s school in July 2022. The coloured final versions of the stories were used, and the brief activity described in Section 3.3.2 was incorporated to make the teaching of affective prosody within the story context more engaging. A prompting procedure like that of the online pilot study was implemented throughout the teaching phase, along with identical versions of *Scoring Sheets 1, 2, and 3* for the pre- and post-test stories. However, in this instance, *Story B* was used for the post-test and *Story C* for the second teaching cycle, while *Story A* continued to be employed in the pretest. Noticeably, in *Story C*, this child also interpreted the elephants, whom the little elephant consults for directions, as representations of the little elephant’s family members.

Pilot testing 2: Results and discussion

The child scored high in both the pre- and post-test stories, slightly improving in the post-test as scores in *Scoring Sheet 1: Story Grammar Units* rose from 7/10 to 8/10. *Scoring Sheet 2: Episodic Structure Complexity* was rated at 4/4 for both stories, while scores in *Scoring Sheet 3: Emotional Vocabulary* improved from 1/2 to 2/2 from the pretest to the post-test. In line with their *Modifiability Rating Scale* score (11/12), the child displayed a positive disposition toward the activities and was eager to participate, despite a mild attention decline as the session progressed due to distraction and fatigue effects, as explained in Section 3.2.3.

Considering the insights from the online and in-person pilot studies, the second teaching cycle was eliminated from the teaching phase within the storytelling activity. This would help to fit the DA into a session of under an hour that is amenable to practitioner use and still captures a range of skill levels within the child population studied. The second teaching cycle was instead scheduled as a separate follow-up session for those children who struggled during the initial post-test. The same story as the original post-test (i.e., *Story A* or *B*) would be applied here, while *Story C* about the little elephant would be introduced in the new post-test. Any necessary adjustments would be made to the scoring sheets to allow for the alternative interpretation of *Story C*’s plot to be granted credit if deemed appropriate. This refers to the elephants whom the little elephant asks for help being identified as relatives rather than strangers (see ‘Pilot testing 1: Results and discussion’ in Section 3.3.3 and ‘Pilot testing 2: Data collection procedure’ in Section 3.3.4). However, a tight schedule prevented the completion of this extra session within the research project’s timeframe, leaving it as a potential area for future exploration of the DA resource.

Other refinements involved ensuring the uniformity of the storytelling activity script's content across all story grammar units to guarantee that equal importance and time are dedicated to learning each unit. Additionally, the teaching phase was redesigned to follow a whole–part–whole sequence, paired with a two-step prompting method, to improve its effectiveness and application consistency. The teaching phase activities were also made more interactive by incorporating story-relevant physical materials, as explained in Section 3.3.1.

Moreover, as mentioned in 'Pilot testing 3: Results and discussion' (Section 3.2.3), at least one short break would be scheduled within future DA sessions to enhance testing conditions and optimise children's performance. School staff would be consulted about extending session times slightly if needed to counteract potential fatigue without compromising session completion. At the same time, they would be requested to kindly ensure that the designated session space is as quiet and distraction-free as possible on the day.

3.3.5 Conclusions

In summary, the first online pilot study of the storytelling activity with six typically developing children with EAL showed high competence across *Scoring Sheet 1: Story Grammar Units*, *Scoring Sheet 2: Episodic Structure Complexity*, and *Scoring Sheet 3: Emotional Vocabulary* in the *Story Generation Task* pretest. However, the post-teaching results indicated a decline in performance and increased variability, despite generally good modifiability ratings reflecting receptiveness to the teaching activities. This decline was attributed to the children's high baseline abilities, which left limited room for improvement, specific interpretative issues with the plot of *Story C* among the youngest participants, and parental interferences with some children's performances. Additionally, focus and engagement weakened as the storytelling activity continued, a problem exacerbated by the remote data collection setting. These challenges, along with others expressed in 'Pilot testing 1: Reflections on online vs in-person data collection for the main study' (Section 3.3.3), prompted a transition to in-person data collection for the main study of the DA to enhance participation and data quality.

Accordingly, one last pilot study was conducted with a typically developing child with EAL at their school. This showed high performance across all three scoring sheets in the *Story Generation Task*, with some pretest–to–post-test gains in story grammar units and emotional vocabulary. However, no gains were observed in episodic structure complexity, where a ceiling score was reached on both occasions. Again, the child exhibited good modifiability

ratings and engagement, although attention mildly declined eventually due to distraction and fatigue.

Aside from the edits made to the stimulus picture stories for clarity and visual appeal following initial feedback from three adult British English speakers and the online pilot study with children, further adjustments to the storytelling activity included:

- Reducing the teaching phase to one cycle instead of two, using two stories rather than three. A possible second cycle with a third story is reserved for a follow-up session for children who struggle in the initial post-test. This helps to fit the DA session length to under an hour.
- Implementing a whole–part–whole sequence into the teaching phase’s structure, complemented by a two-step prompting method for better effectiveness and application consistency.
- Ensuring equal emphasis on learning each story grammar unit throughout the storytelling activity script.
- Boosting the teaching phase’s interactivity by introducing physical materials (e.g., laminated pictograms for the story grammar units, feelings wheel, jigsaw puzzles) and an activity addressing receptive affective prosody more engagingly.

Although made after both pilot studies and the start of the data analysis process for the main study, another key adjustment to the storytelling activity involved adjusting *Scoring Sheets 1* and *3* for the pre- and post-test *Story Generation Tasks* to more accurately evaluate children’s use of story grammar and emotional vocabulary, considering their varying English language skills.

Moreover, regarding the overall DA, to better manage fatigue and minimise distractions, sessions would incorporate at least one scheduled short break and be conducted in quiet, distraction-free spaces with the support of school staff.

3.4 Summary of Piloting Phase and Design Decisions for the Main Study

The present chapter has detailed the piloting stage of this PhD research, which evaluated the initial DA methods aimed at capturing the learning potential of school-aged, multilingual children in narrative macrostructure, emotional vocabulary, and receptive affective prosody.

Between May and June 2021, three English-native adult speakers with expertise in children and refugee communities reviewed the clarity, age, and cultural appropriateness of the initial sketches for the stimulus picture stories. Following this, from June 2021 to July 2022, three pilot studies tested the practicality of the early versions of the DA activities. An online pilot study was initially run with six typically developing UK-based multilingual children, aged 5;06–12;11 years, to evaluate the pretest *Receptive Affective Prosody Task* and the storytelling activity. This was followed by an in-person pilot study with 31 typically developed English-speaking adults to refine the *Receptive Affective Prosody Task*, and an in-person pilot study with a seven-year-old, typically developing multilingual child to finalise the DA activities before the main study.

For the *Receptive Affective Prosody Task*, three designs (*Congruent vs Incongruent Prosody*, *Congruent vs Neutral Prosody*, and *Neutral Content*) were tested, with contribution from four English-native adult speakers who reviewed the audio stimuli for the *Congruent vs Neutral Prosody* design in January 2022. The *Neutral Content* task was ultimately selected for the DA because it seemed better suited to capture receptive affective prosody skills in children within the DA's target age group. This design minimised the risk of confusion from incongruent emotional cues in the *Congruent vs Incongruent Prosody* task, as suggested by the findings from the pilot study with children, and avoided possible difficulties with the *Congruent vs Neutral Prosody* task, as implied by the pilot study with adults.

Refinements were also made within the storytelling activity based on the feedback from the three English-speaking adults and findings from the pilot studies with children. These included editing the stimulus picture stories for better clarity and visual appeal, condensing the teaching phase to one cycle with two stories, introducing a whole-part-whole sequence and two-step prompting method in the teaching phase, and revising the script for balanced emphasis on each story grammar unit. Additionally, the teaching phase was made more engaging and interactive with physical materials and an activity targeting receptive affective prosody. *Scoring Sheets 1 and 3* for the pre- and post-test *Story Generation Tasks* were also revised to enhance the accuracy of evaluating children's story grammar and emotional vocabulary according to their diverse English language abilities. Finally, the overall DA

session structure was adjusted to include at least one break and to be held in quiet, distraction-free settings.

Overall, the piloting phase provided critical insights that informed the design and implementation of the DA for the main study. The subsequent *Chapter 4* details the methods employed in this main study. This in-person research, involving 14 children, expanded upon the initial insights from the pilot studies to further evaluate the DA's utility in identifying DLD in multilingual children. This phase entailed comparing the children's performance in the DA with measures relevant to differentiating, as sources of difficulty with the English language, between DLD and limited familiarity with English among children with EAL. These included the *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015; Chiat et al., 2020), the *New Reynell Developmental Language Scales* (NRDLS; Edwards et al., 2011), observations of the children's home and school interactions, and parental and teacher reports on their language development and experience.

Chapter 4. Methods: Main Study of the Dynamic Assessment

4.1 Introduction

A main study was conducted, building on the results of the pilot studies presented in *Chapter 3*, to further trial the usefulness of the DA as a resource in aiding the recognition of DLD in multilingual children. The participating children's performances in the DA measures — including their capacity to respond to instruction in the teaching phase as per their *Modifiability Rating Scale* scores (see Section 3.3.2), as well as their pretest, post-test, and change scores in both the *Receptive Affective Prosody Task* ('Pilot testing 3: Data collection procedure' in Section 3.2.3) and the *Story Generation Task* (including *Scoring Sheet 1: Story Grammar Units*, *Scoring Sheet 2: Episodic Structure Complexity*, and *Scoring Sheet 3: Emotional Vocabulary*) (Section 3.3.2)— were compared with their performances in several measures appropriate for discerning DLD from English language difficulties attributable to limited experience with this language. These comparative measures were the language-neutral version of the *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015; Chiat et al., 2020), the *New Reynell Developmental Language Scales* in English (NRDLS; Edwards et al., 2011), observations of the children's interactions with peers and adults at school, home, or familiar public spaces, and parental and teacher questionnaires focused on the children's language development and experience. Three key variables were extracted from the questionnaire and observational sources: *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*. The details of these comparative measures are elaborated in Sections 4.3.1 and 4.3.2.

Two Research Questions (RQs) with their corresponding sub-questions were proposed:

- RQ1: What is the relationship between the participating children's performance in the DA and their scores in the NRDLS and the CL-NWRT?
 - Sub-RQ1: Which elements of the DA show the strongest relationship with the NRDLS and CL-NWRT?
- RQ2: What is the relationship between the participating children's performance in the DA and their *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*?

- o Sub-RQ2: Which elements of the DA show the strongest relationship with the children's *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*?

Specifically, the DA elements in both Sub-RQs 1 and 2 refer to the *Modifiability Rating Scale* scores, and the pretest, post-test, and change scores in the *Receptive Affective Prosody Task* and in *Scoring Sheet 1: Story Grammar Units*, *Scoring Sheet 2: Episodic Structure Complexity*, and *Scoring Sheet 3: Emotional Vocabulary* within the *Story Generation Task*.

These RQs were addressed through a mixed-quantitative methods design, which combined both correlational and regression analyses (multiple and ordinal) without predefined hypotheses. Analysing how children's DA performance relates to their NRDLS and CL-NWRT scores (for RQ1), and to their *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score* (for RQ2), aimed to shed light on the DA's utility for identifying DLD risk, as distinct from any effect of the level of English language exposure. This would be based on children's learning potentials in the targeted DLD-sensitive areas: narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and receptive affective prosody. Consistent with Sub-RQs 1 and 2, the analyses also sought to determine which DA measures were most strongly associated with the selected comparative measures, offering critical insights into the potential for detecting DLD risk. In particular, by investigating the DA's connections with the NRDLS, *English Language Experience Score*, and *English Language Proficiency Stage*, the study explored how a child's existing English language skills and experience could impact the DA's effectiveness.

Concurrently, the analyses of the DA's relationships with the CL-NWRT and *DLD Risk Factors Score* were intended to evaluate the DA's capacity to discern DLD risk in multilingual children.

This study's methodological approach is situated within the pre-accuracy phase of the diagnostic research framework (Dollaghan & Horner, 2011; Klee, 2008; Sackett & Haynes, 2002), which delineates four phases for the development of diagnostic tools, each addressing progressively detailed questions about their diagnostic accuracy. In the context of the target disorder in this research (i.e., DLD), these would read:

- Phase I: 'Do test results in children with DLD differ from those in typically developing children?'

- Phase II: ‘Are children with certain test results more likely to have DLD than those with other test results?’
- Phase III: ‘Does the test distinguish children with and without DLD among those in whom it is clinically reasonable to suspect that the disorder is present?’
- Phase IV: ‘Do children who undergo the test fare better (in their ultimate health outcomes) than similar children who are not tested?’

As reflected above, Phase I would compare the results in the DA (*index test*) between groups of children categorised as having DLD —according to a *reference* standard— and their typically developing peers. Since no single reference standard exists for diagnosing DLD in multilingual children, group assignment would need to rely on the use of various evidence-based, complementary measures suited for identifying DLD risk in these children (De Lamo White & Jin, 2011; Letts, 2012; Li’el et al., 2019; Roseberry-McKibbin, 2021). This study takes an earlier, pre-Phase I approach to test development, as it compares the DA performance of a single group of children who exhibit a range of English proficiencies with their results in measures relevant to diagnosing DLD. This lays the basis for subsequent research (e.g., Phases I to IV, as above) into the DA’s value to contribute to the detection of DLD risk among children with EAL.

This chapter begins by outlining the recruitment process and eligibility criteria for the participants in this study (refer to Section 4.2). The data collection methods are then presented in Section 4.3, which contains information on the comparative measures and the formulation of the *English Language Experience Score* and *DLD Risk Factors Score*. Following this, Section 4.4 describes the participants’ profiles, including their demographic, language experience, and both general and language development characteristics. Subsequent sections focus on the study’s interrater reliability (Section 4.5), implementation fidelity (Section 4.6), and data analysis procedures (Section 4.7).

4.2 Recruitment Procedure

Fourteen children with EAL took part in the main study. Details about their demographics, language experience, and general and language development characteristics are provided in Section 4.4. Five children were recruited from two primary schools within Newcastle upon Tyne and the surrounding regions, whereas the remaining nine children were recruited through families, with parents being the first to reach out to the researcher about their

children's participation in the study after hearing about it via word of mouth or through promotions on social media and Newcastle University's channels. Headteachers from potentially participating schools were first approached via email to request permission for their schools to take part in the study, at which point a detailed description of the research was provided (see Appendix 21). The teaching staff then distributed the research information sheets and consent forms in Appendices 22 and 23 to the parents of pupils who met the eligibility criteria outlined in Section 4.2.1, ranging from a handful to over 30 in each school. A website with details about the project, including an informational video, was also shared to further encourage participation among families and facilitate understanding (see <https://blogs.ncl.ac.uk/tgarridotamayo2>). As an alternative, when recruitment was initiated through families, parents were first provided with thorough information about the study, with contact with the children's schools being made once parental consent had been obtained. These schools were also encouraged to invite other multilingual pupils to participate.

In all cases, the option of translating the information and consent sheets into the parents' most proficient language was available, as well as the possibility of going through both documents with the researcher, whether in person or via phone or video call, with the presence of an interpreter if necessary. Moreover, participating children were guided through the child-friendly information sheet in Appendix 24 and asked for their verbal consent at the start of each assessment session. A £10 Amazon gift voucher was offered to their parents as a token of thanks for their cooperation. While this may have encouraged participation, the amount involved should not mean undue pressure was exerted.

Sixteen consent forms were returned within the specified timeframe, which was highly flexible to allow as many children as possible to participate. Due to factors such as illness, unresponsive parents, and changing schools, one child was excluded from the main study, while the other was selected for the in-person pilot of the DA instead (see Sections 3.2.3 and 3.3.4). Notably, reaching the target of 50 participants within the research project's timeframe became unattainable due to the recruitment complications posed by the COVID-19 pandemic. These difficulties were in addition to the already anticipated challenges given the low number of multilingual children with DLD relative to the general population. The final small sample size reflects this, even though extensive efforts were made towards recruiting as many participants as possible to enhance sample representability and study validity. To this end, multiple strategies were pursued: searching in the Department for Education's school database, filtering by education phase and various relevant local authorities; contacting key networks like *Engage with Developmental Language Disorder* (E-DLD), the *National*

Association for Language Development in the Curriculum (NALDIC), and an EAL network of teachers within the *School Effectiveness* team in Newcastle; liaising with Newcastle University's *Language Intervention in the Early Years* (LIVELY) project research team; reaching out to recommended contacts with links to schools; and advertising via Newcastle University's channels and social media (e.g., WhatsApp, Twitter, and Facebook). The promotional flyers for this main study are available in Appendix 25.

4.2.1 Participant eligibility criteria

Families and schools were asked to identify children who met the following criteria:

- Aged 4 (Reception Year) to 8 years old. A range of 5–7 years old was initially selected as later stages of narrative development are progressively reached at these ages; thus, children are expected to produce more structurally complex stories (Khan et al., 2016; Squires et al., 2014). However, as a result of recruitment complications, the age range was slightly extended to allow more children who had expressed interest to take part in the study, as their performance would still provide useful information on the suitability of the DA for different age groups.
- Use languages other than, or in addition to, English at home. This would be in line with the Department for Education's guidelines, which consider pupils to have EAL if they are exposed to a language at home other than English (DfE, 2023).

These children would either present language difficulties or not, based on whether their parents and/or teachers report concerns about the children's language development in English and/or their home language(s). Consistent with the CATALISE panel's recommendations for DLD diagnostic criteria (Bishop et al., 2017), should language difficulties be present, they would not occur in the context of a more complex pattern of impairments characteristic of a biomedical condition (e.g., autism, cerebral palsy, intellectual disability).

In light of the recruitment challenges encountered, the original plan of dividing participants into two groups matched by age and language experience was revised to account for the smaller number of participants. One group would have consisted of children with typically developing language. In contrast, the other group intended to include children whose language difficulties are a primary cause for concern (i.e., at risk of DLD), based on parents' and teachers' concerns about their language skills, along with considerations of the children's observed interactions and their performance in the NRDLS and CL-NWRT relative to other

children of the same age and language background. However, instead of these two distinct groups, the study incorporated a single group of children across a spectrum of language proficiency.

4.3 Data Collection Procedure

Between April 2022 and January 2023, children participated in two 45-60-minute individual assessment sessions in a quiet space within their school buildings. On two occasions, these occurred in a quiet room in the children's homes as school access could not be facilitated, subject to a risk assessment (Appendix 26).

The first session involved administering the NRDLS and CL-NWRT (see Section 4.3.1) to gain a deeper understanding of the children's language development and to provide a reference point for contrasting the results of the DA. Although the session was estimated to last approximately 45 minutes, additional sessions took place on the same day or a different close day as required, depending on the children's needs and factoring in both their age and attention span. Recreational activities such as drawing, colouring, and playing with toys of interest to each child were used to motivate and provide a break between tasks. At the end of the session, stickers were given as a little prize to acknowledge the children's efforts, in addition to verbal praise.

As these initial assessment sessions with the children were underway, parents and teachers were requested to complete the pertinent questionnaires (see Section 4.3.1) alone at their convenience or in an interview with the researcher. Concurrently, the researcher carried out non-intrusive observations of the children's interactions in their schools and either their homes or familiar public spaces (see Section 4.3.1). Out of the 14 participating children, school observations could not be facilitated for two children, whereas for the remaining 12 children, these were carried out in person. Eight of these children were also observed outside of the school context: four at home (one in person and three via video recordings) and four in familiar public spaces. Observations outside the school context were not facilitated for the remaining six children due to parental refusal or non-response.

The second session entailed the implementation of the DA and was audio or video recorded for subsequent analysis. At the beginning of the session, children were welcomed and introduced to the three main activities: the *Feelings Game*, *Storytime*, and the second part of the *Feelings Game*. It was explained that completing each activity would earn star points on a journey to becoming a 'storytelling superstar', and they could win a prize at the end. The

protocols for administering the pre- and post-test *Receptive Affective Prosody Tasks* (*Feelings Games 1* and *2*), as well as the storytelling activity (*Storytime*; including the pre- and post-test *Story Generation Tasks*, and the teaching phase) were adhered to, as outlined in Section 3.2.3 ('Pilot testing 3: Data collection procedure') and Section 3.3.2, respectively. The versions of the materials for both the *Receptive Affective Prosody Tasks* and the storytelling activity were counterbalanced to mitigate potential order effects, with half the children receiving version AB, which included *Story A* as the pretest and *Story B* as the post-test, and the other half BA, with *Story B* as the pretest and *Story A* as the post-test. Scoring of the children's pre- and post-test performance in the *Receptive Affective Prosody Task* and the *Story Generation Task*, along with their modifiability during the teaching phase, was completed according to the guidelines in Sections 3.2.3 ('Pilot testing 3: Data collection procedure') and 3.3.2. Upon conclusion of the session and completion of the stars point chart, children were awarded a certificate of participation in the study (Appendix 27), complemented with a bookmark and a pencil as final small prizes.

4.3.1 Comparative measures

To explore children's developmental status and establish a basis for judging the effectiveness of the DA, the direct and indirect assessment measures outlined below were designated as comparative measures in the main study based on previous research for diagnosing DLD in multilingual children (e.g., Boerma & Blom, 2017; Letts, 2012; Li'el et al., 2019).

Amongst the direct measures, the following were included:

- *New Reynell Developmental Language Scales* (NRDLS; Edwards et al., 2011). This comprehensive assessment tool evaluates the language development of 2;00–7;06-year-old children through play-based activities. It comprises the *Comprehension Scale*, which investigates children's understanding of certain vocabulary items and grammatical features, and the *Production Scale*, which focuses on their ability to produce these language features. Although the NRDLS norms are based on monolingual English-speaking children living in the UK, they were used in this study to gain insights into the participants' expressive and receptive English language abilities, considering only their raw scores. Due to the rise in difficulty level as the NRDLS progresses to reflect language development over the age range, both scales were administered from the first sections and continued until the child failed a whole section. At that point, the practice items and a few items from the subsequent section were tried. If the child answered any item correctly on that occasion, the assessment continued until they failed an entire section again. Testing

with the NRDLS was conducted exclusively in English due to the broad diversity of home languages spoken by the children in this study, which would make adapting and conducting the scales in each of these home languages highly impractical given the research project's time constraints.

- Language-neutral version of the *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015; Chiat et al., 2020). Nonword repetition tasks evaluate phonological short-term memory and processing and are considered a clinical marker for DLD in monolingual and multilingual learning contexts (Boerma & Blom, 2017; Chiat, 2015). Unlike other nonword repetition tests, the CL-NWRT is designed to reduce the influence of language-specific experience on performance by incorporating phonological properties common across languages (Chiat, 2015). In this study, the 5-minute test was presented as a PowerPoint-based game on a laptop computer. Children were told a scenario accompanied by pictures in which the necklace two children made for their mother broke, causing them sadness. They were then required to repeat 16 nonwords ('magic words') preceded by two practice items to create a new necklace. A bead appeared on the necklace with an animated effect after each repetition, allowing the children to see their progress. Each stimulus nonword was played once, with a replay permitted if the first recording was interrupted or the child was distracted. Children listened to the nonwords over child-sized headphones and were praised for their responses regardless of accuracy. The researcher annotated these in real-time and scored later, accounting for the number of whole items correct (Boerma et al., 2015; Chiat & Polišenská, 2016). To be considered correct, responses had to include all phonemes in the target nonword in the proper order, with additions, omissions, and substitutions scored as errors. Non-responses were scored as incorrect. The CL-NWRT materials and permission to use these were kindly received from the authors.

For the indirect measures, previous research demonstrates that parents (Bedore et al., 2011; Boerma & Blom, 2017; Paradis et al., 2010) and teachers (Bedore et al., 2011; Pua et al., 2017) can be reliable informants of multilingual children's language experience and development. With this in view, the two questionnaires explained below and found in Appendices 28 and 29 were developed via Online Surveys.

- *Parental questionnaire*. This questionnaire was designed to gather key information about the children's general and language development, as well as their language usage and exposure at home and outside of the school setting. To develop it, well-known DLD risk

factors across cultures and languages (e.g., absence of differentiating biomedical condition, delayed early linguistic milestones, poor home language(s) abilities, family history of language problems (Bishop et al., 2017; Paradis et al., 2010) were considered, and the *Alberta Language Development Questionnaire* (Paradis et al., 2010), the *Alberta Language Environment Questionnaire* (Paradis, 2011), and the *Questionnaire for Parents of Bilingual Children* (Tuller, 2015) were used as guides to ensure a robust approach.

The questionnaire consists of 58 questions distributed across five sections: *Introductory details*, *General information about the child*, *Languages used with and by the child*, *The child's general and language development*, and *Information about the child's family*. Parents were given the choice to complete it in their own time or to be interviewed by the researcher, either face-to-face or through a phone or video call with an interpreter present, if required, depending on their preference and English language proficiency and literacy levels.

- *Teacher questionnaire*. This questionnaire sought to collect essential data on the children's English language abilities and language experience within the school environment. It is divided into four sections, totalling 17 questions: *Introductory details*, *The child's language use and exposure in school*, *The child's English proficiency* (according to the DfE (2017)'s 5-point scale detailed in Section 2.2), and *Concerns about the child's language development*. Teachers could complete it in their own time or with the researcher in person or via phone or video call.

A further indirect measure included observations in everyday settings. These provide insights into children's communication abilities, needs and preferences beyond what parental and teacher reports can capture, helping to understand their engagement in educational and social activities and identify possible signs of DLD (Letts, 2012, 2013). For instance, limited verbal interactions in both English and the home language(s), perhaps replaced with non-verbal communication attempts, might signal DLD risk. Alternatively, a tendency to interact in the home language(s) and rely heavily on nonverbal communication in English-speaking situations may point to limited English proficiency due to fewer opportunities to learn this language. For this reason, the main study included observations of the children's natural interactions, as guided by the form in Appendix 30, focusing on their attention, social, and play behaviours alongside language and communication.

Accordingly, parents and teachers were asked to video record the child at least twice for 10-25 minutes each time while interacting with peers and adults in everyday settings like home or

the classroom. Video recordings included naturally occurring situations where each child's language was used, such as during playtime with classmates or while conversing with a parent. These videos were then shared with the researcher using Newcastle University's File Drop-Off Service. The relevant instructions for parents and teachers are in Appendices 31 and 32, respectively. When parents were unable to share videos of their children's home interactions, the researcher observed the children in familiar public spaces (e.g., playground, community centre) or at their homes, depending on the family's needs. These in-person observations were subject to the risk assessment in Appendix 26 to ensure the safety and wellbeing of the participants and the researcher, as required by Newcastle University's research policies and ethics standards. Similarly, observations of the children were carried out directly in their schools when teachers requested this option instead of sending video recordings.

4.3.2 Formulation of the English Language Experience Score and DLD Risk Factors Score

Two summary variables were formulated for use in the main study's data analyses, focusing on evaluating the usefulness of the DA (see *Chapter 5*). These included an *English Language Experience Score* reflecting English language usage and exposure, and a *DLD Risk Factors Score* indicating the potential for atypical language development. Both involved the use of data from the parental and teacher questionnaires and the observations of the children's interactions described in Section 4.3.1.

English Language Experience Score

A summary variable titled *English Language Experience Score* was devised to integrate the four components of the multilingual experience defined below, including three quantity-oriented factors (1–3) and one quality-oriented (4). Each factor is key in determining the rate at which children with EAL acquire English. This summary variable is computed by ranking children from 1 (highest value) to 14 (lowest value) on the four factors. The rankings for each child across these factors are then summed and averaged by dividing by 4. This produces a score where a lower value indicates a more extensive English language experience, whereas a higher value captures a less extensive experience. Z-scores were calculated to confirm the validity of this ranking process, ensuring an accurate representation of each child's language experience relative to the group.

- 1) *Length of exposure to English (in months)*: This is determined by subtracting the age at which the children were first exposed to English from their age at the time of testing.

Previous studies indicate that the overall time a child has been exposed to English predicts individual differences in their skills in this additional language, with more prolonged exposure linked to proficiency increases across diverse linguistic domains (e.g., narratives: Govindarajan & Paradis, 2019; vocabulary and verb morphology: Paradis, 2011; vocabulary, grammar, and global/discourse comprehension: Paradis & Jia, 2017).

- 2) *Average frequency of current English language exposure:* This is calculated by scoring the English language input children receive from various interlocutors in their lives (parents, grandparents, siblings, school staff, and peers) on a 5-point scale, from 'Never' (0) to 'Always' (4). The scores from relevant interlocutor categories are then summed and divided by the number of categories applicable to the child, providing a final score for the average English input. The highest possible score is 4.
- 3) *Average frequency of current English language usage:* This is determined by scoring children's use of English with the same interlocutors considered for their English language exposure. An identical 5-point scale is applied, with scores from pertinent interlocutor categories summed and then divided by the number of categories that apply to the child, yielding a final average English output score. The highest possible score is 4.

The amount of English that children encounter and use across daily settings influences their abilities in this language (e.g., semantics and morphosyntax: Bohman, Bedore, Peña, Mendez-Perez & Gillam, 2010; syntax and vocabulary: Sorenson Duncan & Paradis, 2020). While the input quality can impact this effect (e.g., English use at home does not lead to enhanced children's English language proficiency when parents are not fluent in this language; Govindarajan & Paradis, 2019; Hoff, Core & Shanks, 2020; Paradis, 2011; Paradis & Jia, 2017; Sorenson Duncan & Paradis, 2020), all regular communicators are incorporated into the *English Language Experience Score*, acknowledging their contribution as regular sources of input and output to the children's multilingual development.

- 4) *Richness of the English language environment in the home setting:* This is measured by evaluating how often children engage in five language-related activities in English outside the school context on a weekly basis. These activities include reading books or having books read to them, telling stories or listening to them, singing songs or listening to music, watching movies or videos, and playing with siblings or friends. Each activity is scored on a 3-point scale, ranging from 'Rarely/never' (0) to 'Almost every day/every day' (2), with a maximum possible total score of 10 points. The quality of the English language

environment also signals children's proficiency in this language, with those who more frequently engage in language-rich activities demonstrating superior language outcomes (e.g., narratives: Govindarajan & Paradis, 2019; vocabulary and verb morphology: Paradis, 2011; vocabulary and grammar: Paradis & Jia, 2017).

DLD Risk Factors Score

A second summary variable termed *DLD Risk Factors Score* was computed by adding up the scores assigned for the four DLD risk factors specified below, with a maximum of 4 points. A decision was made to exclude family history of speech, language, learning or literacy difficulties, a well-known DLD risk factor (Grimm & Schulz, 2014; Restrepo, 1998; Roseberry-McKibbin, 2021; Tuller, 2015), due to its rare occurrence in the study's sample, affecting only three of the 14 children and mostly relating to hearing difficulties. Given this and the complexities of accurately documenting this kind of family history in culturally diverse contexts (Boerma & Blom, 2017; Paradis et al., 2010), as noted in Section 4.4.1, this factor's inclusion was not deemed appropriate.

- 1) *Production of the first word (in any language)*: 0 points were allocated for 'On time' (≤ 18 months) and 1 point for 'Delayed' (> 18 months).
- 2) *Production of the first multi-word utterances (in any language)*: 0 points were assigned for 'on time' (≤ 24 months) and 1 point for 'delayed' (> 24 months).

The timing of these early milestones is a significant DLD indicator across monolingual and multilingual learning contexts (Boerma & Blom, 2017; Grimm & Schulz, 2014; Roseberry-McKibbin, 2021; Tuller, 2015), with multilingual children attaining them similarly to monolingual peers (Boerma & Blom, 2017; Paradis et al., 2010; Roseberry-McKibbin, 2021). Considering the potential bias observed in the parental reports in this study, such as unrealistic milestones like the first word at 6 months —likely reflecting parental satisfaction with the child's early language development rather than accurate timing (Tuller, 2015)— milestone timing ('on time' or 'delayed') was prioritised over specific ages of achievement.

- 3) *Parental concerns about the child's language abilities*: 0 points were granted for 'No', 1 point for 'Yes'.
- 4) *Teacher concerns about the child's language abilities*: 0 points were granted for 'No', 1 point for 'Yes'.

Insights and concerns reported by parents and teachers over children's language abilities serve as important contributors to the identification of DLD risk in multilingual populations (Bedore et al., 2011; Boerma & Blom, 2017; Li'el et al., 2019; Paradis et al., 2010; Pua et al., 2017; Restrepo, 1998; Tuller, 2015).

4.4 Participants' Profiles

This section provides details on the participants' profiles, focusing on their demographic background (Section 4.4.1), language experience (Section 4.4.2), and general and language development characteristics (Section 4.4.3). As an introductory overview, key attributes for each child are summarised in Table 18.

Table 18. Overview of Key Demographic, Language Experience, and Developmental Characteristics for Each Child

Child	Age	Gender	Sibs	Birth Order	Parental Edu	DLD Risk	Eng Exp	Eng Prof	Home Lang
1	4;06	F	1	1	U-U	3	6.3	5	Italian, Spanish
2	4;08	M	2	3	U-U	2	5.3	5	Finnish
3	4;11	M	1	2	U-U	3	7.5	3	Telugu
4	5;07	M	1	1	U-U	0	13.8	2	Turkish
5	6;02	M	1	2	U	2	8.5	4	Chinese
6	6;06	F	0	1	U-U	1	6.3	5	German, Spanish
7	6;06	M	1	1	S-S	0	9.5	3	Vietnamese
8	6;11	F	2	2	U-U	0	4.0	5	Finnish
9	7;00	F	1	1	U-U	2	7.5	5	Telugu
10	7;01	M	0	1	S-S	1	6.8	2	Russian
11	7;04	M	3	3	U-U	1	3.5	4	Arabic
12	7;05	M	3	3	S-S	2	3.5	5	Kurdish
13	7;10	F	3	4	U-U	1	6.5	5	Arabic
14	8;11	F	2	3	S-P	1	9.0	5	Chinese

Note. Abbreviations: F: Female; M: Male; Sibs: Number of siblings; Parental Edu: Combined parents' education level (proxy for socioeconomic background); P: Primary school; S: Secondary school; U: University; DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; Home Lang: Home language.

Demographic variables in Table 18 include age, gender, number of siblings, birth order, and parents' education (an indicator of socioeconomic status). Excluded are parents' birth countries, languages used at work, and English proficiencies, as they are not directly relevant

to the children. Family history of speech, language, or hearing issues is also omitted to protect confidentiality, given the small number of cases and potential sensitivity of these data in culturally diverse contexts (see Section 4.4.1).

Within the language experience variables, the children's home languages are listed, with other aspects summarised in the *English Language Experience Score*, integrating data on the length of English exposure, frequency of English input and output, and richness of the English language environment at home (see Sections 4.3.2 and 5.2.2). Code-switching practices are excluded due to their high prevalence among participants as a common feature of multilingual language use (see Section 4.4.2).

The children's English proficiency stages are also included. Additional characteristics related to language development are captured in the *DLD Risk Factors Score* (see Sections 4.3.2 and 5.2.2), accounting for parents' and teachers' concerns about the child's language abilities, as well as the timing of first words and word combinations. Further developmental factors discussed in Section 4.4.3 are omitted for clarity.

4.4.1 Demographic characteristics

Fourteen children with EAL recruited from families and schools in Newcastle and nearby areas participated in the main study. They were aged 4;06–8;11 years old ($M = 6;06$, $SD = 1;04$) and consisted of six girls and eight boys born in England, UK, except for one child born in Turkey.

Twelve of the 14 children had siblings. The distribution of birth order within the families was as follows: six children were first-borns (including those without siblings), three children were the second child, four children were the third child, and one child was the fourth-born child.

Regarding the parents, 10 of the 14 children had at least one parent who held a university-level education, possibly associated with a middle or high socioeconomic status among these families. Of these 10 children, nine had parents who both had university degrees. In one case, only the mother had a university degree, while the father's education level was unreported due to an unfilled section in the questionnaire for caregiver 2 (commonly observed to be the father when the mother is caregiver 1). This omission will be noted in each instance where data about the fathers are presented throughout the chapter. Additionally, three children had parents who both attained secondary education, whereas one child had a mother with a primary school education and a father with a secondary school education.

In examining the parents' countries of birth, the data indicate significant diversity. For seven out of the 14 children, both parents were born in the same country: Libya (two children), India (two children), Bulgaria (one child), China (one child), and Lithuania (one child). Conversely, the parents of five children were born in different countries: Finland-England (two children), Venezuela-Italy (one child), Iraq-Kurdistan (one child), and Mexico-Germany (one child). For the child whose father's data were unreported, the mother's country of birth was China. The country of birth data were also missing for one child's parents; however, their home language, Vietnamese, denotes a possible origin from Vietnam for both.

The languages spoken by the parents at their workplaces also varied among the 14 children. Eleven children had at least one parent using English in their professional settings. Specifically, eight had parents who both used English at work, while one child's mother employed English and Turkish, with data for the father marked as not applicable in this regard. For the remaining two children of the 11, one had a father who spoke English at work, with data for the mother also marked as not applicable. The other had a mother using English with no corresponding data reported for the father. Furthermore, for one child, both parents spoke Kurdish in their workplaces, while for another, both parents used Chinese. Data were not reported for the parents of one child.

For self-perceived language proficiency, parents rated their skills in English and their home languages on a 4-point scale: 'Not at all,' 'Basic understanding and speaking ability,' 'Good understanding and can express myself in many situations,' and 'Excellent understanding and can express myself in most situations.' Focusing on the English language, seven of the 14 children had both parents rate their proficiency as 'excellent'. For the remaining seven, the data varied. The parents of two children refrained from providing data on this aspect. Two other children had both parents rate their English language skills as 'basic', while in another case, the mother's proficiency was 'basic' and the father's 'good'. For another child, the mother reported 'good' proficiency, with no information available for the father. Finally, one child's mother rated her proficiency as 'basic' and the father's as 'not at all'.

Finally, the presence of speech, language, hearing problems or learning difficulties within the family was reported for three of the 14 children, with the father having hearing complications in all three cases and the mother experiencing speech difficulties in one case. Notably, factors related to culturally diverse contexts—such as reluctance to disclose these relatives' issues due to stigma, difficulties in tracking educational experiences in families affected by conflicts and displacements, and limited access to special education services in some societies—could

have affected parents' willingness and ability to report the presence of these problems in their families (Boerma & Blom, 2017; Paradis et al., 2010).

4.4.2 Language experience profiles

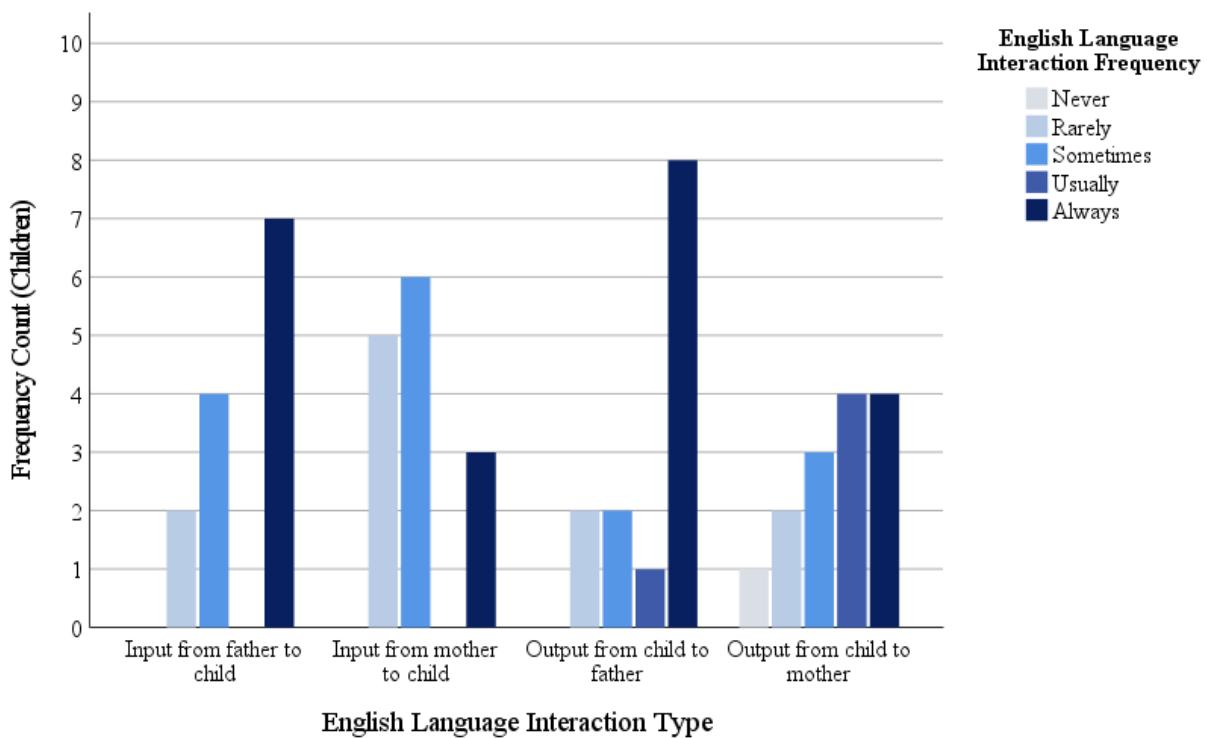
In addition to English, the languages spoken at home by the children included Arabic (two children), Chinese (two children), Finnish (two children), Kurdish (one child), Russian (one child), Telugu (two children), Turkish (one child), and Vietnamese (one child). Two children were from a trilingual household where English, Spanish, and either German or Italian were used.

Eight children were exposed to English from birth, while another six began their exposure between the ages of 3;00 and 4;06 years. The length of English exposure across the 14 children spanned from 1;06 to 7;05 years ($M = 5;01$, $SD = 2;00$). Complete data for two children were missing as obtaining clarifications from their parents was not possible regarding their responses to the question about the age of first exposure to English. However, further information from the children's parents and teachers allows for reasonable inference that they experienced some level of English exposure from birth.

Regarding code-switching practices, 12 children were reported to use more than one language concurrently. Amongst those, only one parent expressed concern, fearing that teachers and others might not understand the child when they spoke their home language, even though this practice was deemed appropriate at home.

The frequency of English input and output varied across different sources in the home context. As presented in Figure 19, when examining the children's English language interactions with their parents, mothers 'always' spoke English to three children, 'sometimes' to six, and 'rarely' to five. Four children each 'always' and 'usually' spoke English to their mother, while three 'sometimes' and two 'rarely' did so. One child 'never' spoke English to their mother. In contrast, among the 13 children with data, fathers 'always' spoke English to seven children, 'sometimes' to four and 'rarely' to two. Eight children 'always' spoke English to their fathers, one 'usually' did, two 'sometimes', and two 'rarely'. This shows that all children used and were exposed to some English from both parents, with distinct patterns in parent-child English language interactions.

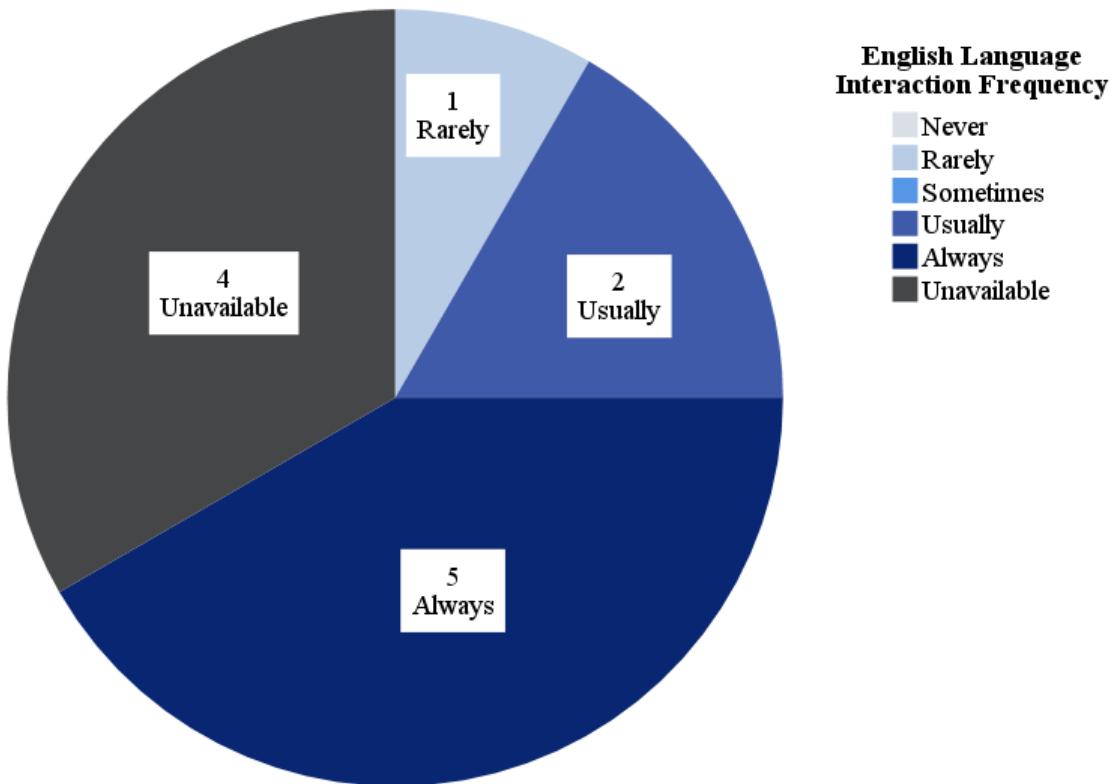
Figure 19. Clustered Bar Chart of Frequencies of English Language Interactions Between Children and their Parents



Additionally, grandparents were reported as regular caregivers for four children. They ‘always’ spoke to the children in their home languages, never in English. As for the children, three ‘always’ spoke in their home languages to their grandparents, with one of them ‘rarely’ using English. The fourth child exclusively spoke in English to them.

For the 12 children with siblings, English language input and output varied within sibling interactions. As shown in Figure 20, five children, each with one to three siblings, ‘always’ spoke and were spoken to in English by their siblings. Two children with two siblings each ‘usually’ used and were spoken to in English. One child with a single sibling ‘rarely’ used and was spoken to in English. Incomplete data were noted in several cases. One child with three siblings was ‘always’ spoken to in English by their eldest sibling, but data on input from the other siblings and all sibling-related output was unavailable. Another child with one sibling ‘usually’ spoke English, but there was no information about the input from the sibling. For two children with one sibling each, no input or output data were provided, with one case attributed to the sibling being deemed too young by the parent.

Figure 20. Pie Chart of Frequencies of English Language Interactions between Children and their Siblings



Note. The pie chart combines the frequencies of both English language input and output in the children's interactions with their siblings, since available data showed identical patterns.

'Unavailable' represents instances of incomplete or missing data.

In addition to these familial interactions, the richness of the English language environment at home, measured by children's weekly engagement in activities such as reading or watching films in this language as explained in Section 4.3.2, ranged from 3 to 10 points. Three children scored 10, three scored 9, two scored 8, and one scored 7, reflecting frequent engagement in these activities. Two children scored 6, one scored 4, and one scored 3, indicating less frequent engagement. Data were missing for one child.

In the school setting, three teachers —two from the same institution— reported the children using languages besides English with specialised teachers at their schools, specifically Spanish and French. In addition, languages such as Arabic, Bengali, Chinese, Czech, Kurdish, Persian, Polish, Romanian, Russian, Urdu, and Vietnamese were spoken informally and outside the curriculum. Regardless of this multilingual environment in two schools, all 14 children's teachers 'always' spoke English to the children, with 13 children 'always' speaking in English to them and one 'usually' doing so. Two children 'sometimes' or 'rarely' spoke in their home language.

As for interactions with other children, peers ‘always’ spoke in English to 13 of the children who also ‘always’ spoke in English to them. One child was ‘usually’ spoken to in English by peers and ‘sometimes’ in their home language. This child also ‘usually’ spoke in English to their peers but ‘sometimes’ used their home language.

In summary, across the diverse linguistic environments, the majority of children experienced substantial exposure to English from various sources both at home and in school. English usage is notably prominent in interactions with fathers and siblings, and remains consistent in communication with teachers and peers, even in the presence of several other languages. Additionally, while there is variation in the quality of the English language environment at home, most children frequently engage in language-rich activities in English, which could be beneficial to their proficiency in this language (e.g., Govindarajan & Paradis, 2019; Paradis, 2011; Paradis & Jia, 2017).

4.4.3 General and language development profiles

With regards to the 14 children’s general development, complications during pregnancy or birth were reported for two children, while five children were born outside the standard term. Twelve children started walking before 18 months, with only one child beginning after this period; data for one child were unavailable. Non-DLD-related, treated health conditions were reported in five children.

Other developmental aspects that could be informative to the risk of DLD were also examined. These included children’s excessive reliance on gestures over speech for communication, particularly when considering their dominant language, inappropriate social use of language, and difficulty paying attention (Letts, 2012; Roseberry-McKibbin, 2021). Play-related struggles, such as difficulties accessing play (Lloyd-Esenkaya, Russell & Clair, 2020), were also considered. However, across the sample, these characteristics appeared to fall within the range of typical developmental variability. Specifically, attention difficulties were observed in six children, but these instances were within the normal variability for their developmental stages, sometimes reflecting parental expectations of behaviour rather than the children’s actual capabilities. Additionally, while issues with social language use, such as turn-taking and staying on topic, were noted in the youngest children, these were considered typical for their age.

The parents of eight children expressed concerns about their speech and language abilities, with expressive difficulties noted for seven children and receptive difficulties for one child.

The children's expressive difficulties led to reduced understanding from others, particularly among friends and more distant family members. Despite this, children usually understood their parents and could follow multi-step instructions. When asked about the impact of these speech and language difficulties on the children's lives, parents noted people's need for their children to repeat what they said and their children's frustration when not being understood.

Teachers revealed concerns for three children consistent with the issues raised by the parents of these children, suggesting uniformity across home and school observations. Teachers did not voice any concerns for the four other children for whom their parents had expressed concerns, possibly supporting the observation that parents of multilingual children often express greater apprehension about their child's language development than parents of monolingual children (Boerma & Blom, 2017). Despite the concerns for three children, no teacher perceived any child as being slower in their English language acquisition than children of similar age, language, and cultural background.

In terms of the children's *English Language Proficiency Stage*, as per the DfE (2018)'s 5-point scale of reading, writing and spoken language, eight children were categorised as 'E. Fluent', while two children each fell into the 'D. Competent' and 'C. Developing Competence' categories. Finally, two were identified at the 'B. Early Acquisition' level. Furthermore, as far as the teachers were aware, only two experienced a silent period upon first exposure to English. The duration of this period was unknown for one child, while it lasted approximately 10 weeks for the other, falling within the typical range (Bligh, 2014; Siraj-Blatchford & Clarke, 2000). However, see Roberts (2014) for a review questioning the extent and quality of the evidence for a silent stage in early second language acquisition.

Children were also evaluated on whether they began uttering their first words around the age of 12 months and before turning 18 months, as well as their first word combinations around 18 months and before 24 months, as per typical developmental expectations (Roseberry-McKibbin, 2021; Tuller, 2015). These data pertain to only 12 children because it could not be recalled for the other two. The age at which these 12 children pronounced their first recognisable words (in any language) ranged from 0;06 to 3;00 years ($M = 1;06$, $SD = 0;10$). Eight children said their first word at or before 18 months, with the remaining four doing so after 18 months. Regarding the age at which they began joining words into short sequences (in any language), the range was 1;04 to 3;06 years ($M = 2;01$, $SD = 1;00$). Eight children started combining words at or before the first 24 months, and four began doing so after 24 months (two also had their first words past 18 months).

4.5 Interrater Reliability for the Dynamic Assessment's Storytelling Activity

For interrater reliability purposes, a psychology placement student was trained to act as a second rater and independently transcribe and score the stories of five randomly selected children, representing 36% of the sample. The student was blind to the children's developmental profiles and the order in which the pre- and- post-test stories were presented (AB or BA). Scoring was performed using *Scoring Sheet 1: Story Grammar Units*, *Scoring Sheet 2: Episodic Structure Complexity*, and *Scoring Sheet 3: Emotional Vocabulary* for both stories A and B.

In *Scoring Sheet 1*, the initial interrater agreement was moderate at 60% for *Story A* and low at 20% for *Story B*. These percentages were calculated by taking the total number of instances where the two raters agreed for the children's stories and dividing that number by 5, which was the total number of children. Discrepancies arose from differing interpretations of what three children said and the scoring criteria. After clarification and discussion among the raters, the agreement increased to 80% for both stories, with any remaining disagreements due to the student's difficulty in clearly hearing one child's utterances once for each story. In these instances, the researcher's ratings were accepted given her advantage of direct, in-person hearing of this child.

For *Scoring Sheet 2*, the initial interrater agreement was high at 80% for *Story A* and moderate at 60% for *Story B*, which improved to 80% following further discussion. Persisting discrepancies were again linked to the student's difficulties in hearing the same child. The researcher's ratings, supported by her in-person hearing, were again accepted.

Regarding *Scoring Sheet 3*, an initial agreement of 100% was achieved for both stories, demonstrating complete consistency between raters.

Additionally, the student examined the storytelling activity recordings of another three randomly selected children (21% of the sample) to determine their *Modifiability Rating Scale* score. The interrater agreement was 100%, indicating total consistency.

Importantly, the student did not rate children's performance in the *Receptive Affective Prosody Task* nor the researcher's adherence to this task's protocol, as its administration procedure and scoring criteria are straightforward and objective compared to the storytelling activity, including the *Story Generation Task*. The latter involves a more subjective and flexible approach to both administration and scoring, introducing more variability in protocol adherence and the interpretation of children's performance.

4.6 Implementation Fidelity for the Dynamic Assessment's Storytelling Activity

For implementation fidelity purposes, the psychology placement student also reviewed the researcher's adherence to the storytelling activity guidelines using the same three recordings of the three randomly selected children evaluated for their *Modifiability Rating Scale* score. The aim was to check the accuracy and consistency of application across participants, including the use of the relevant mediated learning experience strategies during the teaching phase. Accordingly, two tools were designed and implemented: the *Storytelling Activity Implementation Fidelity Monitoring Scale* (Table 19) and the *Rating Form for the Assessor's Use of Mediated Learning Experience Strategies* (Table 20), with the latter drawing inspiration from Lidz (1991)'s *Mediated Learning Experience Rating Scale*, which aids in examining an assessor's use of mediational behaviours and has been utilised in previous DA research (Peña et al., 2006; 2014b; Ukrainetz et al., 2000).

Focusing on the *Storytelling Activity Implementation Fidelity Monitoring Scale* in Table 19, it was found that the researcher scored 18/18 for two children, and 17/18 for the other child due to the occasional use of positive feedback during the post-test phase (e.g., 'That's excellent, well done!'). Although the script advises against feedback on performance at this stage, a subsequent review of the recording confirmed its appropriateness for this particular child. The feedback was not aimed at evaluating their performance, but rather at reinforcing their self-confidence during moments of difficulty in producing certain words due to limited English proficiency. This aided in fostering a positive testing environment and would be in line with the researcher's use of 'competence' as one of the mediated learning experience strategies in the teaching phase (see Section 3.3.2).

Table 19. Storytelling Activity Implementation Fidelity Monitoring Scale

Storytelling Activity Implementation Fidelity Monitoring Scale	Never 0	Sometimes 1	Often 2
Pretest phase			
1. Followed the script for the pretest phase			
2. Used the correct materials for the pretest phase			
3. Ensured an appropriate environment, considering unavoidable limitations (e.g., background noise)			
Teaching phase			
4. Followed the script for the teaching phase			
5. Used the correct materials for the teaching phase			
6. Ensured an appropriate environment, considering unavoidable limitations (e.g., background noise)			
Post-test phase			
7. Followed the script for the post-test phase			
8. Used the correct materials for the post-test phase			
9. Ensured an appropriate environment, considering unavoidable limitations (e.g., background noise)			
Implementation fidelity score (sum of 1 to 9):	/18		
Notes:			

In addition, when assessed with the *Rating Form for the Assessor's Use of Mediated Learning Experience Strategies* in Table 20, the researcher was awarded the full score of 4/4 for all three children, indicating a proper application of the strategies.

Table 20. Rating Form for the Assessor's Use of Mediated Learning Experience Strategies

Rating Form for the Assessor's Use of Mediated Learning Experience Strategies	
1. Mediation of intentionality	
Statements that convey the goal and purpose of the activity, such as:	
<ul style="list-style-type: none"> - <i>Today we are going to learn how to tell good stories. You will become a storytelling superstar!</i> - <i>We are going to learn all the parts that a story should have.</i> 	
Please mark: Observed/Not observed	
Notes:	
2. Mediation of meaning	
Statements of the relevance of the activity and what we are working on, such as:	
<ul style="list-style-type: none"> - <i>Telling stories well is important because it helps us to communicate with friends and family, do really well in school, and learn lots of things about life and the world.</i> - <i>Talking about how the girl feels is important because it helps us to understand what feelings or emotions she is experiencing and why she wants to fix the plant pot.</i> 	
Please mark: Observed/Not observed	
Notes:	
3. Mediation of transcendence	
Statements connecting the activity and what we are working on with relevant experiences in the child's life outside the context of the session, such as:	
<ul style="list-style-type: none"> - <i>This plant was important to the girl, possibly one of her favourite things! Do you have a favourite thing? Maybe a toy you love? How would you feel if this got broken?</i> - <i>What would happen if we told a story that was missing some parts? For example, if your friend or teacher asks you to tell them a story but you only say who the characters are...</i> 	
Please mark: Observed/Not observed	
Notes:	
4. Mediation of competence	
Statements directing the child's focus to their acquired knowledge and how this can be applied, or highlighting the child's accomplishments to foster self-confidence in their abilities. For example:	
<ul style="list-style-type: none"> - <i>To help you remember all the parts the story should have for it to be complete, you can use these pictures we have been working with. Let's use them while we recap what we have learned.</i> - <i>Well done! You are on your way to becoming a storytelling superstar!</i> 	
Please mark: Observed/Not observed	
Notes:	
Assessor's use of mediated learning experience strategies score (out of 4):	/4

4.7 Data Analysis Procedure

This study's data analysis procedure focused on further evaluating the DA's effectiveness for DLD risk detection among children with EAL. Initial steps involved summarising sample performance in the DA —*Receptive Affective Prosody Task, Story Generation Task, and Modifiability Rating Scale*— and the comparative measures —NRDLS, CL-NWRT, *DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage*— through descriptive statistical methods, as well as testing the impact of the DA's teaching phase on children's performance using repeated measures t-tests and a Wilcoxon signed-rank test. Consecutively, the research questions examining the relationships between children's outcomes in the DA and the comparative measures were addressed through a series of correlational and regression analyses. Findings from these analyses are presented in *Chapter 5*.

Chapter 5. Results: Main Study of the Dynamic Assessment

5.1 Introduction

Building on earlier pilot studies detailed in *Chapter 3*, this chapter presents the findings of the main study in this PhD research, which was designed to further trial the effectiveness of the Dynamic Assessment (DA) being developed. This pretest–teach–post-test DA resource aims to identify Developmental Language Disorder (DLD) risk in primary school-aged, multilingual children with English as an Additional Language (EAL) by exploring their learning potential across narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and receptive affective prosody. Examining the children’s learning capabilities across these three DLD-vulnerable areas can provide a sense of whether any difficulties in English result from DLD or limited familiarity with this language.

Fourteen children with EAL aged 4;06–8;11 years old ($M = 6;06$, $SD = 1;04$) from families and schools in Newcastle and nearby areas took part in this study, with details about the sample found in Section 4.4. These children’s outcomes in the DA measures—including their *Modifiability Rating Scale* scores, and their pretest, post-test, and change scores in the *Receptive Affective Prosody Task* and *Story Generation Task* (including *Scoring Sheet 1: Story Grammar Units*, *Scoring Sheet 2: Episodic Structure Complexity*, and *Scoring Sheet 3: Emotional Vocabulary*)—were evaluated against those of measures pertinent to identifying DLD risk among multilingual children, distinguishing this from their possible lack of experience with the English language. These included the language-neutral version of the *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015; Chiat et al., 2020) and the *New Reynell Developmental Language Scales* in English (NRDLS; Edwards et al., 2011) (for further details, refer to Section 4.3.1). Additionally, the children’s *English Language Proficiency Stage* was considered, along with two summary variables—a *DLD Risk Factors Score* and an *English Language Experience Score*—, which were devised using selected data from the parental and teacher questionnaires, as well as from observations of the children. Details about how these summary variables were calculated are covered in Section 4.3.2.

In this framework, two Research Questions (RQs) were investigated, along with their corresponding sub-questions:

- RQ1: What is the relationship between the participating children’s performance in the DA and their scores in the NRDLS and the CL-NWRT?

- Sub-RQ1: Which elements of the DA show the strongest relationship with the NRDLS and CL-NWRT?
- RQ2: What is the relationship between the participating children's performance in the DA and their *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*?
 - Sub-RQ2: Which elements of the DA show the strongest relationship with the children's *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*?

As specified in *Chapter 4*, the DA elements in Sub-RQs 1 and 2 encompass the *Modifiability Rating Scale* scores, as well as the pretest, post-test, and change scores in the *Receptive Affective Prosody Task* and in *Scoring Sheet 1: Story Grammar Units*, *Scoring Sheet 2: Episodic Structure Complexity*, and *Scoring Sheet 3: Emotional Vocabulary* within the *Story Generation Task*.

An open exploration of the relationships identified in the RQs, without predefined hypotheses, can provide evidence of the DA's utility in disentangling DLD risk from a need for additional English language exposure. While the DA's relationships with the NRDLS, *English Language Experience Score*, and *English Language Proficiency Stage* serve to examine whether the DA's usefulness is confounded by a child's current English language abilities and experience, the relationships with the CL-NWRT and *DLD Risk Factors Score* offer direct indications into its value to detect DLD risk in children with EAL.

Within this chapter, Sections 5.2.1 to 5.2.6 cover the results of the initial data analysis conducted to assess performance levels within the sample. This involved summarising children's outcomes in the NRDLS, CL-NWRT, *DLD Risk Factors Score*, *English Language Experience Score*, *English Language Proficiency Stage*, and DA using descriptive statistics. Additional insight into the usefulness of the DA's teaching phase was then gained through repeated measures t-tests and a Wilcoxon signed-rank test (see Sections 5.2.3 and 5.2.4). These tests were employed to examine the changes between the pre- and post-test scores across all participants. Furthermore, Section 5.2.7 explores individual outcomes in the DA.

Following this, to address RQs and Sub-RQs 1 and 2, Pearson and Spearman's rank-order correlational analyses were carried out, alongside stepwise (forward) multiple and ordinal regression analyses. The outcomes of these correlational analyses, which can be found in Sections 5.3.1 to 5.3.3, provided insights into the relationships between children's scores in

the DA measures (*Receptive Affective Prosody Task*, *Story Generation Task*, and *Modifiability Rating Scale*), and their respective scores in the NRDLS and CL-NWRT. These analyses also incorporated the children's *English Language Proficiency Stage*, as well as their *DLD Risk Factors* and *English Language Experience* scores (see Sections 5.3.4 to 5.3.6). Additionally, the regression analyses investigated the extent to which these variables predicted variance in the DA performance (Section 5.4).

The data were first inputted into Microsoft Excel and then, after cleaning and preparation, imported into IBM SPSS Statistics 27, where the statistical analyses were conducted.

5.2 Performance Across the Sample in the Comparative Measures and the Dynamic Assessment

5.2.1 NRDLS and CL-NWRT: Descriptives

The descriptive statistics for the 14 children's performance in the *New Reynell Developmental Language Scales* (NRDLS) and the *Crosslinguistic Nonword Repetition Test* (CL-NWRT) are presented in Table 21. These measures supplemented the evaluation of the children's development and served as a basis for appraising the DA's effectiveness (see Section 4.4.1).

For the NRDLS, the *Comprehension Scale* (maximum raw score: 72) reflects a relatively high and consistent average performance in comprehension skills within the children. In contrast, the *Production Scale* (maximum raw score: 64) shows more variability and potentially greater challenges in language production across the sample.

In the CL-NWRT (maximum score: 16), performance was reasonably high with modest variability among participants.

Table 21. Descriptive Statistics for the Children's Performance in the NRDLS and CL-NWRT

	N	Range	Minimum	Maximum	M	SD
NRDLS Comprehension Scale score	14	20	49	69	61.0	6.9
NRDLS Production Scale score	14	34	28	62	50.3	11.8
CL-NWRT score	14	7	8	15	12.1	2.2

Note. NRDLS: *New Reynell Developmental Language Scales*; CL-NWRT: *Crosslinguistic Nonword Repetition Test*.

5.2.2 DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage: Descriptives

Children's *English Language Proficiency Stage* and two summary variables —*English Language Experience Score* and *DLD Risk Factors Score*— were derived from parental and teacher questionnaires, along with observational data, to gauge the extent of their English proficiency, usage, and exposure, as well as the likelihood that they could have atypical language development. In conjunction with the NRDLS and CL-NWRT, they also served as a basis for evaluating the usefulness of the DA. Details about the summary variables and the *English Language Proficiency Stage* are specified in Sections 4.3.2 and 4.4.3, respectively.

Table 22 presents the descriptive statistics for the *DLD Risk Factors Score* and *English Language Experience Score* across the 14 participants. The *DLD Risk Factors Score* (maximum value: 4) considers the production of the first word and multi-word utterances, and parental and teacher concerns about the child's language abilities. As shown in Table 22, this score indicates a low DLD risk level among the children with some variability. Missing data for two children on the ages of their first word and multi-word utterances (see Section 4.4.3) were handled by scoring these factors as 0 ('On time'), hypothesising that the lack of recall implied no parental concerns regarding these milestones.

Table 22. Descriptive Statistics for the Children's English Language Experience Score and DLD Risk Factors Score

	N	Range	Minimum	Maximum	M	SD
DLD Risk Factors Score	14	3.0	0.0	3.0	1.4	1.0
English Language Experience Score	14	10.3	3.5	13.8	7.0	2.7

The *English Language Experience Score* operates on a ranking system where 1 represents the most extensive experience and 14 the least, based on combining scores for: the length of exposure to English; the average frequency of current English language exposure and usage; and the richness of the English language environment at home. As shown in Table 22, this score reflects varied but generally moderate experience with English across the sample. Data for one child were missing regarding the richness of the English language environment at home. In line with standard practices for handling missing data, where an average score is given if estimating missing continuous data, this factor was ranked as in the middle (7).

It could be argued that the factor concerning the richness of the English language environment at home is a qualitatively different and more subjective scale when compared to those of the three quantity-orientated factors. Therefore, a separate calculation of the *English Language Experience Score* was performed excluding this factor. This recalculated score was tried in all relevant correlational and regression analyses conducted with the DA measures. However, the results of these analyses are not reported in this chapter, as excluding this factor did not yield different outcomes from when it was included.

Additionally, as stated in Section 4.4.3, the children's teachers categorised their *English Language Proficiency Stage* using the Department for Education (2018)'s 5-point scale: 'A. New to English', 'B. Early Acquisition', 'C. Developing Competence', 'D. Competent', and 'E. Fluent'. For ease of analysis, these categories were recoded numerically from 1 (A) to 5 (E). Eight children were classified as 5 ('Fluent'), representing most of the sample (see Table 23).

Table 23. Descriptive Statistics for the Children's English Language Proficiency Stage

	N	Range	Minimum	Maximum	Median	Mode	M	SD
English Language Proficiency Stage	14	3	2	5	5.0	5	4.1	1.2

5.2.3 Dynamic assessment's Receptive Affective Prosody Task: Descriptives and repeated measures t-test results

Shifting the focus to the 14 children's performance in the Dynamic Assessment (DA) measures, the *Receptive Affective Prosody Task* probes their ability to decode emotional cues in voices by having them match utterances —emotionally neutral in lexical content but spoken with emotional prosody (happiness/sadness or anger/fear)— to pictograms depicting the conveyed emotion. The task consists of eight stimuli, with a maximum score of 8 points in both the pre- and post-tests (see 'Pilot testing 3: Data collection procedure' in Section 3.2.3).

The sample's performance data for the *Receptive Affective Prosody Task* did not significantly deviate from a normal distribution, as confirmed by a Shapiro-Wilk test ($p = .088$).

Accordingly, a repeated measures t-test was conducted to evaluate the impact of the teaching phase on performance in this task across participants. As shown in Table 24, while the mean score rose slightly from pretest to post-test, this improvement was not statistically significant, $t(13) = 1.71$, $p = .111$. The effect size was small (Cohen's $d = 0.457$, Hedges' $g = 0.444$).

Table 24. Descriptive Statistics for the Children’s Performance in the Receptive Affective Prosody Task

Receptive Affective Prosody Task	N	Range	Minimum	Maximum	M	SD
Pretest score	14	5	3	8	6.7	1.4
Post-test score	14	5	3	8	7.1	1.7
Change score	14	3	-1	2	0.4	0.9

5.2.4 Dynamic assessment’s Story Generation Task: Descriptives, repeated measures t-test, and Wilcoxon signed-rank test results

Story Grammar Units: Descriptives and repeated measures t-test results

During the pre- and post-tests of the DA’s storytelling activity, children participated in the *Story Generation Task*. Three measures were completed to evaluate their stories: *Scoring Sheet 1: Story Grammar Units*, *Scoring Sheet 2: Episodic Structure Complexity*, and *Scoring Sheet 3: Emotional Vocabulary*. From now on, these will be referred to as *Story Grammar Units*, *Episodic Structure Complexity*, and *Emotional Vocabulary* throughout this chapter.

For *Story Grammar Units*, each unit (characters, time, place, problem, emotion 1, plan, attempt, outcome, and emotion 2) is scored between 0 and 3 points, with a maximum possible score of 30 points in both the pre- and post-tests (see Section 3.3.2).

After verifying the normality of the data with a Shapiro-Wilk test ($p = .788$), a repeated measures t-test was performed to assess the impact of the teaching phase on *Story Grammar Units* scores across the sample. The results showed that the pretest-to-post-test increase in the mean score (see Table 25) was significant, $t(13) = 2.85$, $p = .014$, with a medium effect size (Cohen’s $d = 0.762$, Hedges’ $g = 0.740$), indicating a meaningful impact of the teaching phase on children’s capacity to include story grammar elements in their narratives.

Table 25. Descriptive Statistics for the Children’s Performance in Scoring Sheet 1: Story Grammar Units (Story Generation Task)

Scoring Sheet 1: Story Grammar Units	N	Range	Minimum	Maximum	M	SD
Pretest score	14	15	9	24	18.8	4.7
Post-test score	14	17	10	27	21.5	5.5
Change score	14	13	-4	9	2.7	3.6

Episodic Structure Complexity: Descriptives and Wilcoxon signed-rank test results

Episodic Structure Complexity consists of a 4-level scale to assess the intricacy of the episode's structure within children's narratives: 'Level 1: Description', 'Level 2: Sequence', 'Level 3: Incomplete Episode', and 'Level 4: Complete Episode' (see Section 3.3.2).

Due to the ordinal nature of the scores, a Wilcoxon signed-rank test was used to assess the effect of the teaching phase on the participants' scores. The complexity of the episodic structures in children's stories did not significantly improve from pretest to post-test, $z = -0.91$, $p = .366$, with a small effect size ($r = .171$). As reflected in Table 26, the mode remained at 4 for the pre- and post-tests, indicating that six children were already performing at ceiling level before any instruction was provided.

Table 26. Descriptive Statistics for the Children's Performance in Scoring Sheet 2: Episodic Structure Complexity (Story Generation Task)

Scoring Sheet 2: Episodic Structure Complexity	N	Range	Minimum	Maximum	Median	Mode	M	SD
Pretest score	14	2	2	4	3.0	4	3.1	0.9
Post-test score	14	3	1	4	4.0	4	3.4	0.9
Change score	14	3	-1	2	0.0	0	0.2	0.9

Emotional Vocabulary: Descriptives and repeated measures t-test results

Emotional Vocabulary targets the use of vocabulary related to the protagonists' emotional responses to the story challenges and outcomes, with a maximum achievable score of 4 points in the pre- and post-tests (see Section 3.3.2).

After establishing data normality with a Shapiro-Wilk test ($p = .177$), a repeated measures t-test showed that the small pretest-to-post-test increase in the mean score (see Table 27) was not statistically significant, $t(13) = 0.38$, $p = .710$. The effect size was small (Cohen's $d = 0.102$, Hedges' $g = 0.099$). Thus, across the sample, the teaching phase did not significantly boost children's awareness of emotional vocabulary in the story contexts.

Table 27. Descriptive Statistics for the Children’s Performance in Scoring Sheet 3: Emotional Vocabulary (Story Generation Task)

Scoring Sheet 3: Emotional Vocabulary	N	Range	Minimum	Maximum	M	SD
Pretest score	14	3	1	4	2.7	1.1
Post-test score	14	4	0	4	2.9	1.2
Change score	14	4	-2	2	0.1	1.4

5.2.5 Dynamic assessment’s Modifiability Rating Scale: Descriptives

The *Modifiability Rating Scale* evaluates children’s ability to respond to instruction during the DA’s teaching phase. It consists of six items, each scored from 0 (‘Never’) to 2 (‘Often’), across three domains: responsiveness, transfer of learning, and assessor effort, with a total possible score of 12 points (see Section 3.3.2).

As shown in Table 28, while there was notable variability among the children, the results indicate a generally high capacity to respond to instruction and apply new knowledge during the teaching phase. These findings also suggest that the assessor may typically need to invest less effort to facilitate learning within this particular group.

Table 28. Descriptive Statistics for the Children’s Performance in the Modifiability Rating Scale

	N	Range	Minimum	Maximum	M	SD
Modifiability Rating Scale	14	8	4	12	9.1	3.3

5.2.6 Summary of the sample performance in the comparative measures and the dynamic assessment

Across the sample, children showed reasonably strong CL-NWRT performance, alongside solid language comprehension and a range of language production skills in the NRDLS, which aligns with the tendency for comprehension to precede production in language acquisition (Edwards et al., 2011; Giguere & Hoff, 2022). Moreover, they generally presented a low risk of DLD, moderate English language experience, and proficient English language skills.

With regard to the DA, in the *Story Generation Task*, the usage of *Story Grammar Units* significantly improved in the children’s narratives after the teaching phase. However, within this same task, only small, non-significant improvements were observed in *Episodic Structure Complexity* and *Emotional Vocabulary*. Similar small, non-significant progress also occurred

in the *Receptive Affective Prosody Task*. Finally, the *Modifiability Rating Scale* indicated the children's relatively high capacity to respond to teaching.

5.2.7 Individual results in the dynamic assessment

Building on the summarised sample outcomes in Section 5.2.6, this section presents the individual results across the DA measures for the 14 children, as shown in Table 29, which also includes comparative measure results for reference. These individual-level data reflect the general trends discussed in Sections 5.2.1 to 5.2.6.

A closer examination of the children's DA results identifies Children 2 and 10 as cases where higher *Modifiability Rating Scale* scores do not necessarily correspond to improved post-test outcomes. Child 2 (age 4;08) showed the largest gains in *Story Grammar Units* and *Episodic Structure Complexity* across the sample, with a 9-point gain in the former and a 2-point gain in the latter, along with a 1-point gain in *Emotional Vocabulary*. This progress occurred despite a low *Modifiability Rating Scale* score of 4, reflective of age-related attention difficulties and the need for greater assessor effort to facilitate learning. Child 2's low score of 3 in both the *Receptive Affective Prosody Task* pre- and post-tests was also influenced by these attention challenges. In contrast, Child 10 (age 7;01) had a *Modifiability Rating Scale* score of 10, indicating good engagement with the learning process, yet showed gains only in the *Receptive Affective Prosody Task* (2 points). Child 10's score in *Story Grammar Units* declined by 3 points, and in *Emotional Vocabulary* by 1 point, primarily due to missing Emotion 2, while their *Episodic Structure Complexity* score remained stable.

These findings suggest that pre-existing English skills may impact children's ability to benefit from instruction, regardless of responsiveness, since Child 2 was classified as 'Fluent' in English by their teacher, with a *NRDLS Comprehension Scale* score of 60 and a *Production Scale* score of 38, and Child 10 was at the 'Early Acquisition' stage of English proficiency, with a *NRDLS Comprehension Scale* score of 50 and a *Production Scale* score of 28. Therefore, as discussed in *Chapter 6*, the DA should be refined to more accurately capture teaching-induced performance changes in the targeted areas across various English proficiency stages, enhancing its capacity for distinguishing DLD risk from limited English skills.

Table 29. Overview of Individual Results in the Comparative Measures and the Dynamic Assessment

Child	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Age	4;06	4;08	4;11	5;07	6;02	6;06	6;06	6;11	7;00	7;01	7;04	7;05	7;10	8;11
Eng Prof	5	5	3	2	4	5	3	5	5	2	4	5	5	5
Eng Exp	6.3	5.3	7.5	13.8	8.5	6.3	9.5	4.0	7.5	6.8	3.5	3.5	6.5	9.0
DLD Risk	3	2	3	0	2	1	0	0	2	1	1	2	1	1
CL-NWRT	12	8	13	10	11	15	9	14	14	11	13	15	14	11
NRDLS Comp.	55	60	64	52	69	65	49	59	67	50	66	66	66	66
NRDLS Prod.	48	38	46	28	56	58	45	58	58	28	58	60	62	61
RAPT	Pretest	7	3	6	5	8	8	6	7	6	8	8	7	8
	Post-test	8	3	7	4	8	8	8	6	8	8	8	8	8
	Change	1	0	1	-1	0	0	2	-1	1	2	0	0	0
SS1	Pretest	18	11	18	9	24	20	23	18	23	13	23	21	21
	Post-test	20	20	14	13	24	26	25	24	24	10	27	24	23
	Change	2	9	-4	4	0	6	2	6	1	-3	4	3	2
SS2	Pretest	3	2	2	2	4	4	4	2	3	3	4	4	3
	Post-test	4	4	2	1	4	3	4	3	4	3	3	4	4
	Change	1	2	0	-1	0	-1	0	1	1	0	-1	0	1
SS3	Pretest	4	1	1	4	2	2	3	2	4	3	4	2	4
	Post-test	2	2	0	3	4	4	3	3	2	2	4	3	4
	Change	-2	1	-1	-1	2	2	0	1	-2	-1	0	1	0
Modifiability	4	4	5	5	9	12	8	11	12	10	12	12	12	12

Note. Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; Modifiability: *Modifiability Rating Scale* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; RAPT (pretest, post-test and change): *Receptive Affective Prosody Task* scores; SS1 (pretest, post-test and change): *Scoring Sheet 1: Story Grammar Units* scores; SS2 (pretest, post-test and change): *Scoring Sheet 2: Episodic Structure Complexity* scores; SS3 (pretest, post-test and change): *Scoring Sheet 3: Emotional Vocabulary* scores.

5.3 Correlational Analyses between the Dynamic Assessment Measures and the Comparative Measures

This section presents the results of the bivariate correlational analyses conducted to investigate possible relationships between the 14 children's scores in the Dynamic Assessment (DA) measures and their performance in comparative measures. These analyses address Research Question 1 (RQ1), concerning the relationship between the children's DA performance and their scores in the NRDLS and CL-NWRT, and Research Question 2 (RQ2), examining the relationship between the children's DA performance and their *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*. Furthermore, the analyses contribute to exploring Sub-RQ1, which seeks to identify which elements of the DA are most strongly related to outcomes in the NRDLS and CL-NWRT, and Sub-RQ2, which focuses on the DA elements that relate most strongly with the children's *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*.

The DA measures included the *Receptive Affective Prosody Task*, *Story Generation Task* (*Story Grammar Units*, *Episodic Structure Complexity*, and *Emotional Vocabulary*), and *Modifiability Rating Scale*. Comparative measures comprised the NRDLS and CL-NWRT, pertinent to RQ1 and Sub-RQ1, along with the children's *English Language Proficiency Stage*, *DLD Risk Factors Score*, and *English Language Experience Score* for RQ2 and Sub-RQ2. These correlational analyses were exploratory and conducted without predefined hypotheses, in line with the novel nature of the DA resource being trialled and the study's position within the pre-accuracy phase of the diagnostic research framework (for further details, see Section 4.1).

For most analyses, both Pearson's product-moment and Spearman's rank-order correlation coefficients were conducted for a comprehensive understanding of the data. This dual approach was adopted to address the challenges posed by the small sample size and the non-normal distribution of some scores, as indicated by Shapiro-Wilk tests. While Pearson's correlations are appropriate for data with symmetrical distributions, Spearman's correlations are preferred for data that deviate from normal distribution patterns. On this basis, both types of analyses were included to obtain a more robust insight into the relationships being explored. The only exceptions were the analyses involving data from *Episodic Structure Complexity*, where only Spearman's correlations were run due to its ordinal nature.

In the sub-sections that follow, Spearman's correlations are primarily reported, with Pearson's correlations only referenced when they show a different outcome from Spearman's, as the results from Shapiro-Wilk tests indicated deviations from the normal distribution in the following measures (p -values $< .05$):

- NRDLS scores (*Comprehension*: $p = .027$, *Production*: $p = .013$)
- *Receptive Affective Prosody Task*'s pretest ($p = .012$) and post-test ($p < .001$) scores
- *Story Grammar Units*' pretest ($p = .041$) and post-test ($p = .016$) scores
- *Emotional Vocabulary*'s pretest ($p = .017$) and post-test ($p = .018$) scores
- *Modifiability Rating Scale* scores ($p = .004$)
- *English Language Proficiency Stage* ratings ($p = .001$)

5.3.1 Correlations between the Receptive Affective Prosody Task and the NRDLS and CL-NWRT

The relationships between the children's pretest, post-test, and change scores in the *Receptive Affective Prosody Task* and their scores in the NRDLS and CL-NWRT were examined using Pearson's and Spearman's correlations (see Tables 33A and 33B in Appendix 33).

Spearman's correlations are prioritised below due to normality deviations in the *Receptive Affective Prosody Task* and NRDLS data (see Section 5.3).

The pretest scores presented highly significant positive correlations with the NRDLS scores (*Comprehension*: $r_s(12) = .69$, $p = .006$, Figure 21; *Production*: $r_s(12) = .76$, $p = .001$, Figure 22), and a significant positive correlation with the CL-NWRT scores, $r_s(12) = .59$, $p = .027$ (Figure 23), suggesting that the *Receptive Affective Prosody Task* pretest scores largely align with the NRDLS and CL-NWRT scores.

The post-test scores showed a trend towards a significant positive correlation with the *NRDLS Production Scale* scores, $r_s(12) = .47$, $p = .091$ (Figure 24), but no correlations with the scores in the *NRDLS Comprehension Scale*, $r_s(12) = .37$, $p = .200$, and CL-NWRT, $r_s(12) = .34$, $p = .237$. Pearson's correlations supported these results and further revealed a trend towards a significant positive correlation with the CL-NWRT scores (see Table 33A). These findings suggest some correlation between the *Receptive Affective Prosody Task* post-test scores and performance in the *NRDLS Production Scale*, as well as some evidence of a possible link with the CL-NWRT scores, but no clear relationships with the *NRDLS Comprehension Scale* scores.

The change scores showed no correlations with the scores in the NRDLS (*Comprehension*: $r_s(12) = -.21, p = .463$; *Production*: $r_s(12) = -.18, p = .547$ and CL-NWRT, $r_s(12) = -.11, p = .719$, indicating a lack of relationships between the *Receptive Affective Prosody Task* change scores and the NRDLS and CL-NWRT scores.

Figure 21. Scatterplot of the Correlation between the Pretest Scores in the Receptive Affective Prosody Task and the NRDLS Comprehension Scale Scores

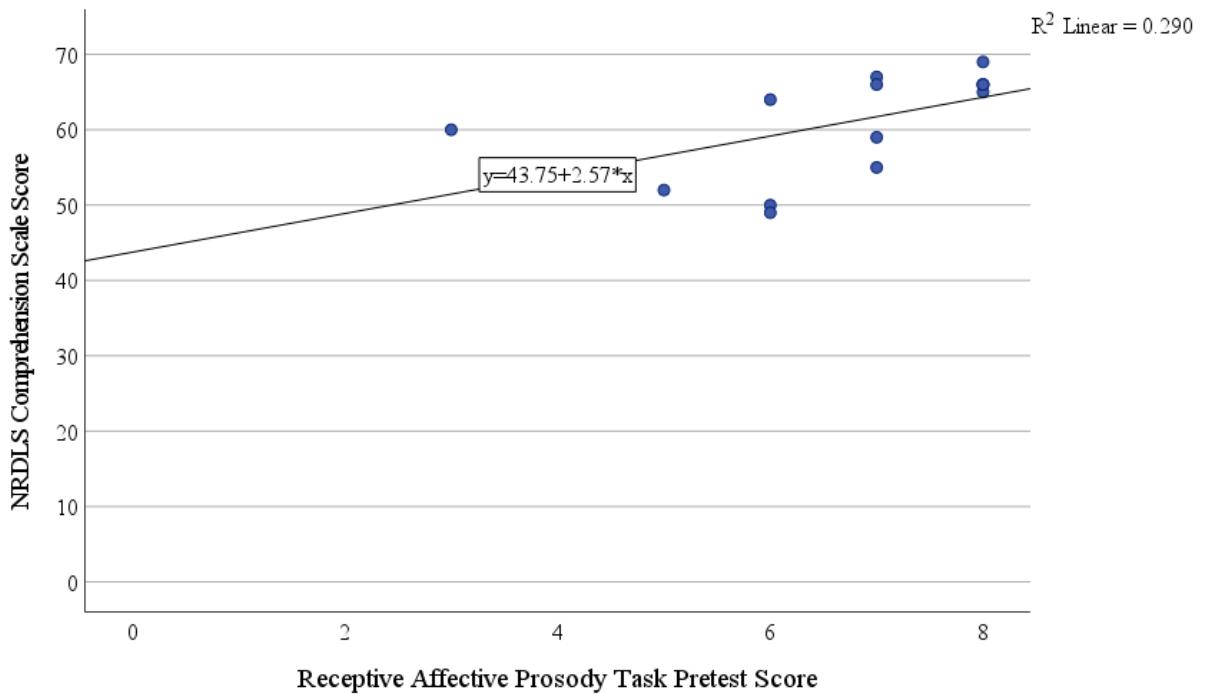


Figure 22. Scatterplot of the Correlation between the Pretest Scores in the Receptive Affective Prosody Task and the NRDLS Production Scale Scores

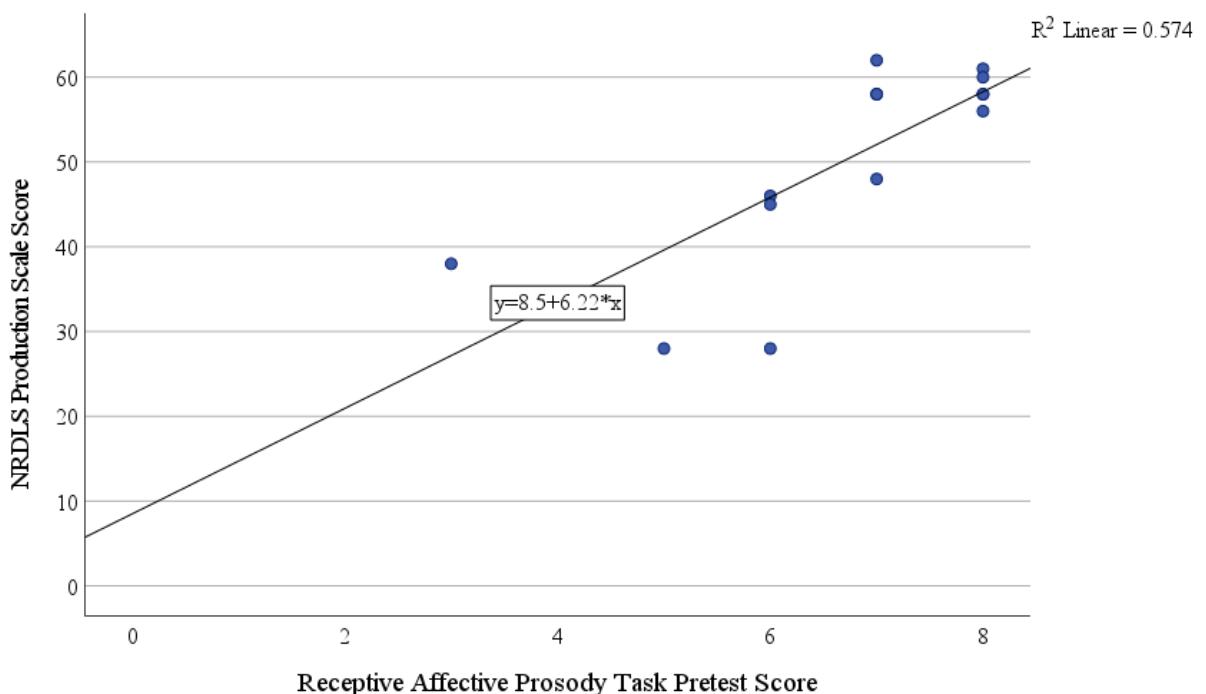


Figure 23. Scatterplot of the Correlation between the Pretest Scores in the Receptive Affective Prosody Task and the CL-NWRT Scores

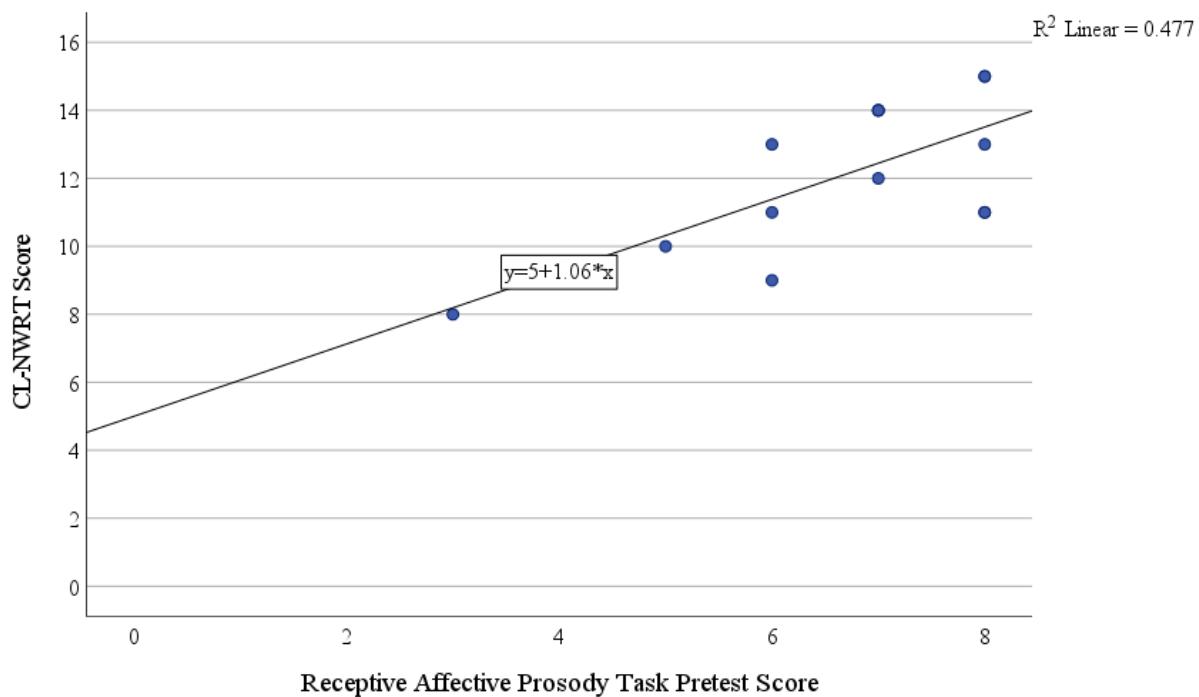
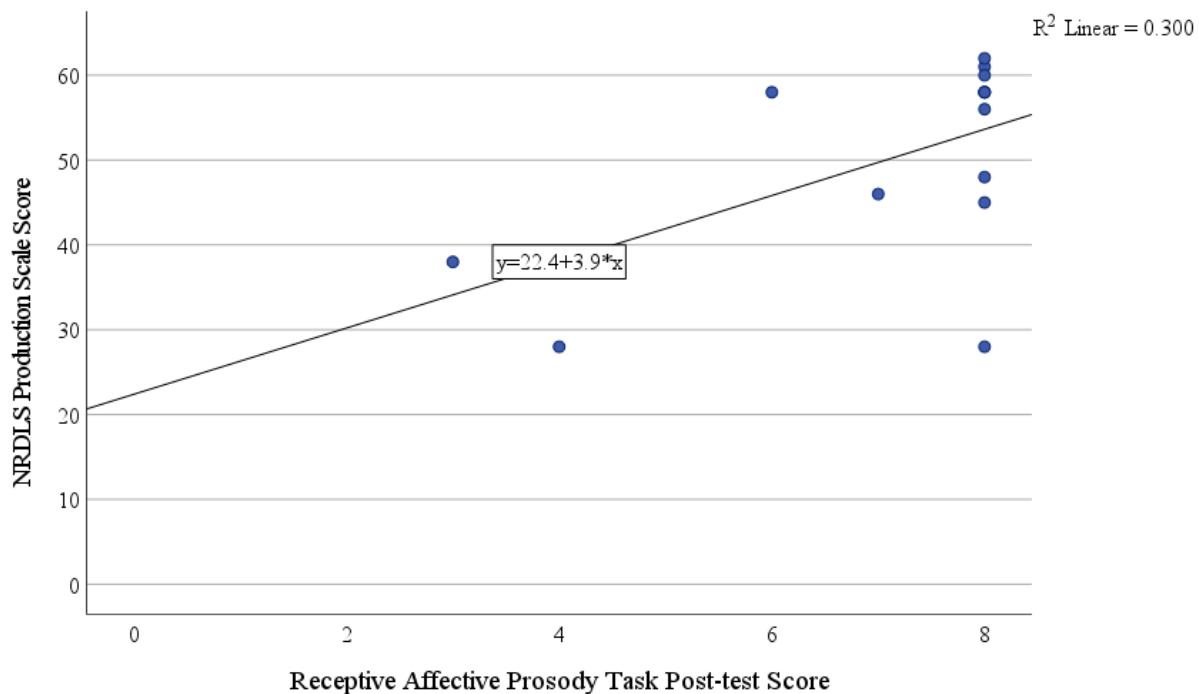


Figure 24. Scatterplot of the Correlation between the Post-test Scores in the Receptive Affective Prosody Task and the NRDLS Production Scale Scores



5.3.2 Correlations between the Story Generation Task and the NRDLS and CL-NWRT

Story Grammar Units and the NRDLS and CL-NWRT

The relationships between the children's pretest, post-test, and change scores in *Story Grammar Units* (which pertains to the DA's *Story Generation Task*) and their NRDLS and CL-NWRT scores were studied through Pearson's and Spearman's correlations (see Tables 34A and 34B in Appendix 34). Given normality deviations in the *Story Grammar Units* and NRDLS data (see Section 5.3), Spearman's correlations are prioritised.

The pretest scores showed significant positive correlations with the NRDLS scores (*Comprehension*: $r_s(12) = .64, p = .013$, Figure 25; *Production*: $r_s(12) = .53, p = .049$, Figure 26), and no relationship with the CL-NWRT scores, $r_s(12) = .24, p = .411$, suggesting that the *Story Grammar Units* pretest scores consistently correlate with the NRDLS scores, but not with the CL-NWRT scores.

The post-test scores displayed a significant positive relationship with the *NRDLS Production Scale* scores, $r_s(12) = .64, p = .014$ (Figure 27), but not with the scores in the *NRDLS Comprehension Scale*, $r_s(12) = .46, p = .101$, and CL-NWRT, $r_s(12) = .28, p = .338$. Pearson's correlations supported these findings, except for a significant positive correlation with the *NRDLS Comprehension Scale* scores (see Table 34A). Overall, the *Story Grammar Units* post-test scores correlate with the *NRDLS Production Scale* performance, whereas this alignment appears less pronounced with the *NRDLS Comprehension Scale* and is absent with the CL-NWRT.

The change scores showed no correlations with the scores in the NRDLS (*Comprehension*: $r_s(12) = -.05, p = .864$; *Production*: $r_s(12) = .23, p = .440$) and CL-NWRT, $r_s(12) = -.02, p = .942$, implying that the *Story Grammar Units* change scores do not reflect outcomes in the NRDLS and CL-NWRT.

Figure 25. Scatterplot of the Correlation between the Pretest Scores in Scoring Sheet 1: Story Grammar Units and the NRDLS Comprehension Scale Scores

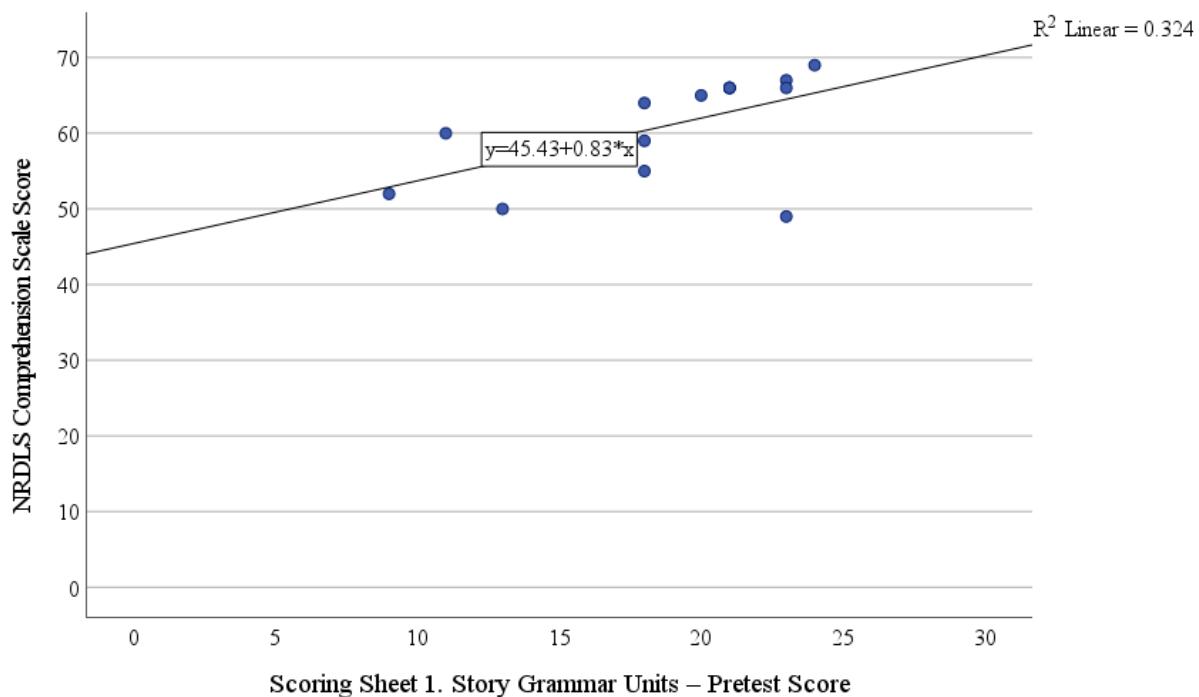


Figure 26. Scatterplot of the Correlation between the Pretest Scores in Scoring Sheet 1: Story Grammar Units and the NRDLS Production Scale Scores

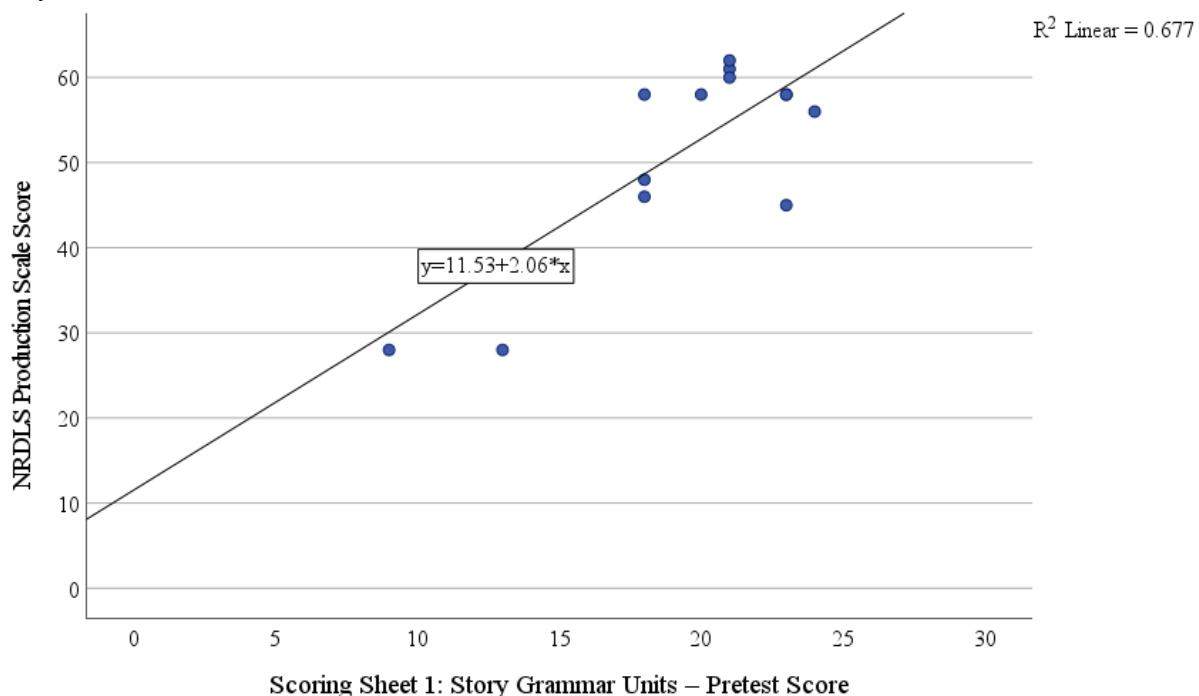
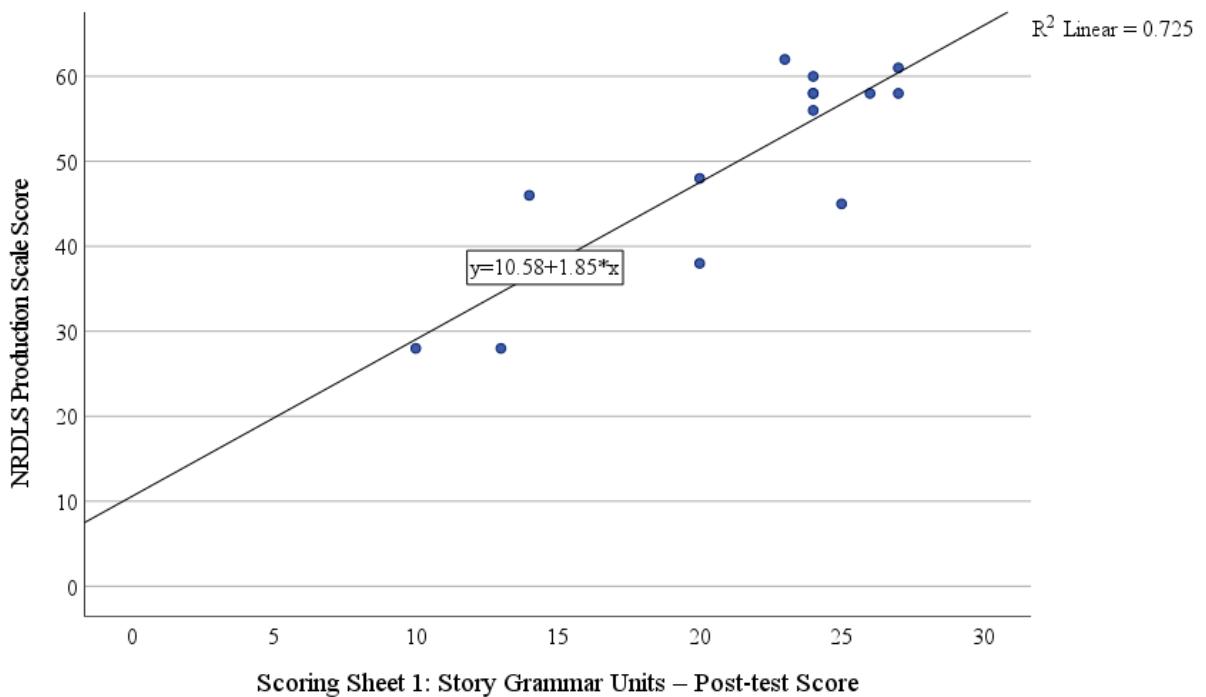


Figure 27. Scatterplot of the Correlation between the Post-test Scores in Scoring Sheet 1: Story Grammar Units and the NRDLS Production Scale Scores



Episodic Structure Complexity and the NRDLS and CL-NWRT

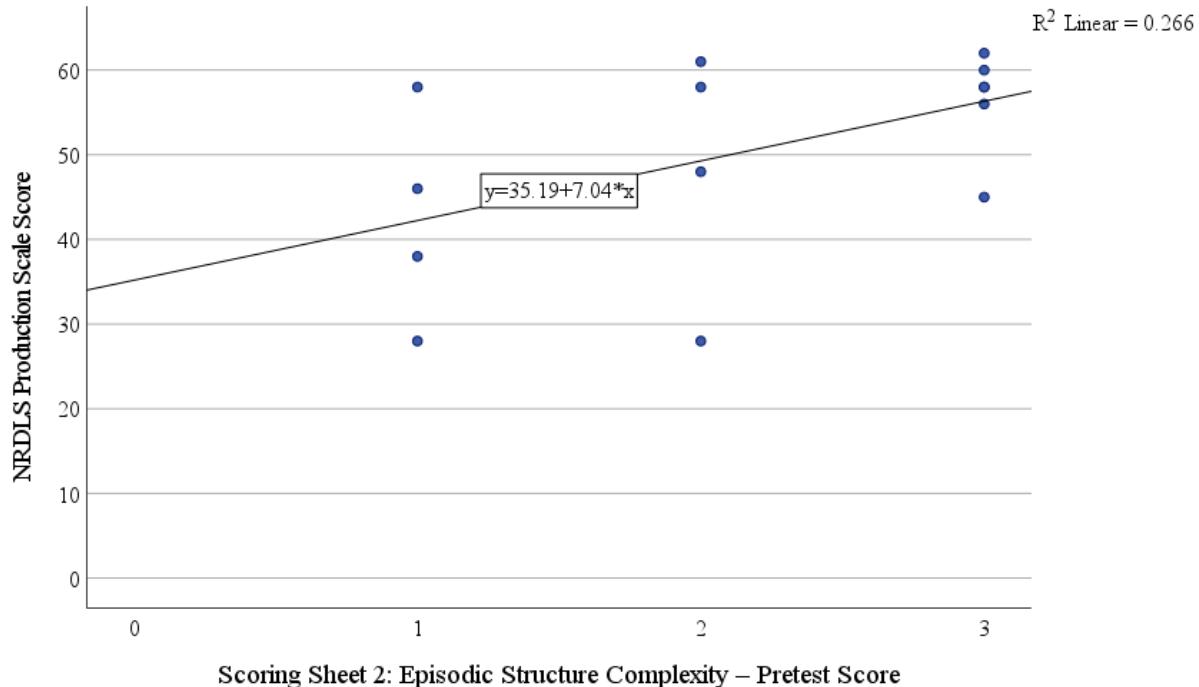
Considering the ordinal nature of *Episodic Structure Complexity* (DA's *Story Generation Task*), only Spearman's correlations were performed to explore the relationships between children's pretest, post-test, and change scores in *Episodic Structure Complexity* and their scores in the NRDLS and CL-NWRT (see Table 35A in Appendix 35). In this case, Pearson's correlations were unsuitable (see Section 5.3).

Pretest ratings showed no correlations with the scores in the *NRDLS Comprehension Scale*, $r_s(12) = .40, p = .161$, and *CL-NWRT*, $r_s(12) = .33, p = .257$. However, a trend emerged towards a significant positive correlation with the *NRDLS Production Scale* scores, $r_s(12) = .47, p = .090$ (Figure 28). Post-test ratings also did not correlate with the scores in the NRDLS (*Comprehension*: $r_s(12) = .37, p = .194$; *Production*: $r_s(12) = .39, p = .167$) and CL-NWRT, $r_s(12) = -.07, p = .816$. Change scores further resulted in a lack of correlations with the scores in the NRDLS (*Comprehension*: $r_s(12) = .03, p = .927$; *Production*: $r_s(12) = .07, p = .807$) and CL-NWRT, $r_s(12) = -.19, p = .519$.

Given these findings, while there is indication of a possible relationship between the *Episodic Structure Complexity* pretest ratings and the *NRDLS Production Scale* outcomes, overall,

performance in the NRDLS and CL-NWRT does not align with the pretest, post-test, and change scores in *Episodic Structure Complexity*.

Figure 28. Scatterplot of the Correlation between the Pretest Ratings in Scoring Sheet 2: Episodic Structure Complexity and the NRDLS Production Scale Score



Emotional Vocabulary and the NRDLS and CL-NWRT

The correlations between the children’s pretest, post-test, and change scores in *Emotional Vocabulary* (DA’s *Story Generation Task*) and their NRDLS and CL-NWRT scores were evaluated using Pearson’s and Spearman’s correlations (see Tables 36A and 36B in Appendix 36). Given normality deviations in the *Emotional Vocabulary* and NRDLS data (see Section 5.3), Spearman’s correlations are prioritised.

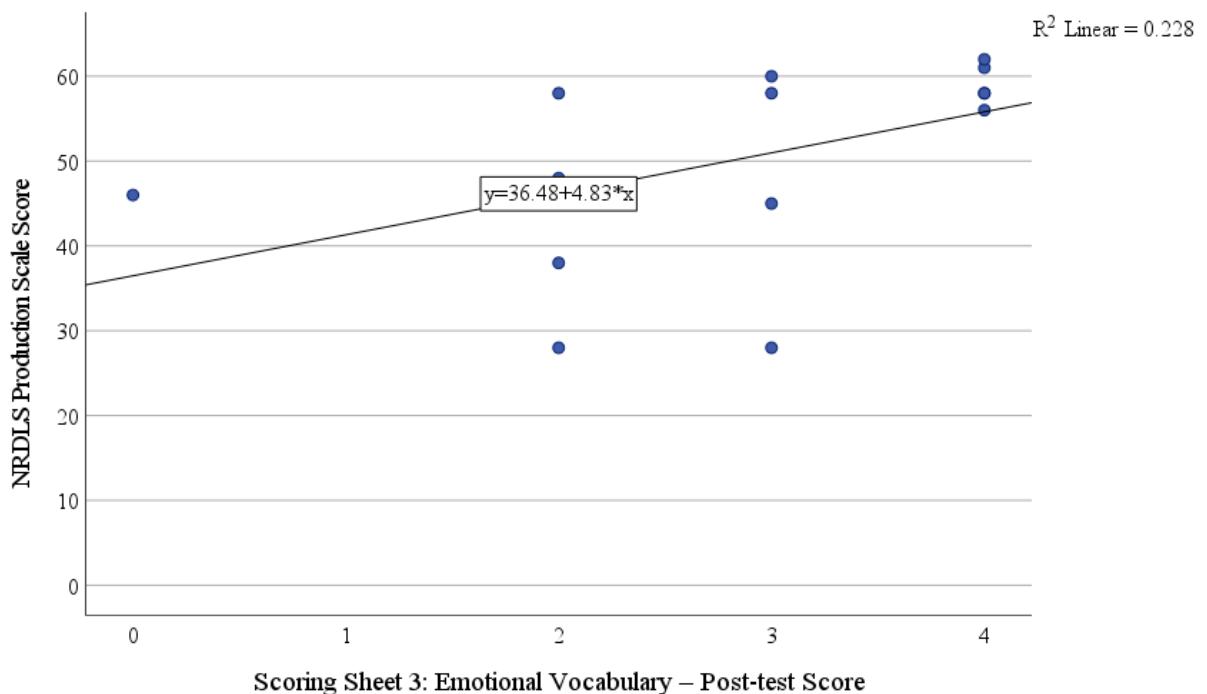
The pretest scores showed no relationships with the scores in the NRDLS (*Comprehension*: $r_s(12) = -.05, p = .868$; *Production*: $r_s(12) = .025, p = .934$) and CL-NWRT, $r_s(12) = -.14, p = .636$, suggesting no correlations between the *Emotional Vocabulary* pretest scores and performance in the NRDLS and CL-NWRT.

The post-test scores showed no correlations with the scores in the *NRDLS Comprehension Scale*, $r_s(12) = .44, p = .112$, and CL-NWRT, $r_s(12) = .20, p = .500$. However, a significant positive correlation was observed with the *NRDLS Production Scale* scores, $r_s(12) = .59, p = .027$ (Figure 29). Pearson’s correlations mirrored these findings, including a trend towards a

significant positive correlation with the *NRDLS Production Scale* scores (see Table 36A). Overall, while the *Emotional Vocabulary* post-test scores did not align with performance in the *NRDLS Comprehension Scale* and CL-NWRT, they seem to be correlated with the *NRDLS Production Scale* results.

The change scores displayed no correlations with the scores in the NRDLS (*Comprehension*: $r_s(12) = .35, p = .218$; *Production*: $r_s(12) = .43, p = .125$) and CL-NWRT, $r_s(12) = .24, p = .413$, indicating no clear relationship between the *Emotional Vocabulary* change scores and the NRDLS and CL-NWRT outcomes.

Figure 29. Scatterplot of the Correlation between the Post-test Scores in Scoring Sheet 3: Emotional Vocabulary and the NRDLS Production Scale Scores



5.3.3 Correlations between the Modifiability Rating Scale and the NRDLS and CL-NWRT

The relationships between the children's *Modifiability Rating Scale* scores and their NRDLS and CL-NWRT scores were tested via Pearson's and Spearman's correlations (see Tables 37A and 37B in Appendix 37). With normality deviations in the *Modifiability Rating Scale* and NRDLS data (see Section 5.3), Spearman's correlations are prioritised.

A significant positive correlation was found with the *NRDLS Comprehension Scale* scores, $r_s(12) = .60, p = .022$ (Figure 30), and highly significant positive correlations with the scores in the *NRDLS Production Scale*, $r_s(12) = .81, p = .001$ (Figure 31), and CL-NWRT, $r_s(12) =$

.68, $p = .008$ (Figure 32). Pearson's correlations supported these results; however, the positive correlation with the *NRDLS Comprehension Scale* scores approached significance (see Table 37A). Hence, in general, children's ability to respond to the teaching provided during the DA correlates with their performance in the NRDLS and CL-NWRT.

Figure 30. Scatterplot of the Correlation between the Modifiability Rating Scale Scores and the NRDLS Comprehension Scale Scores

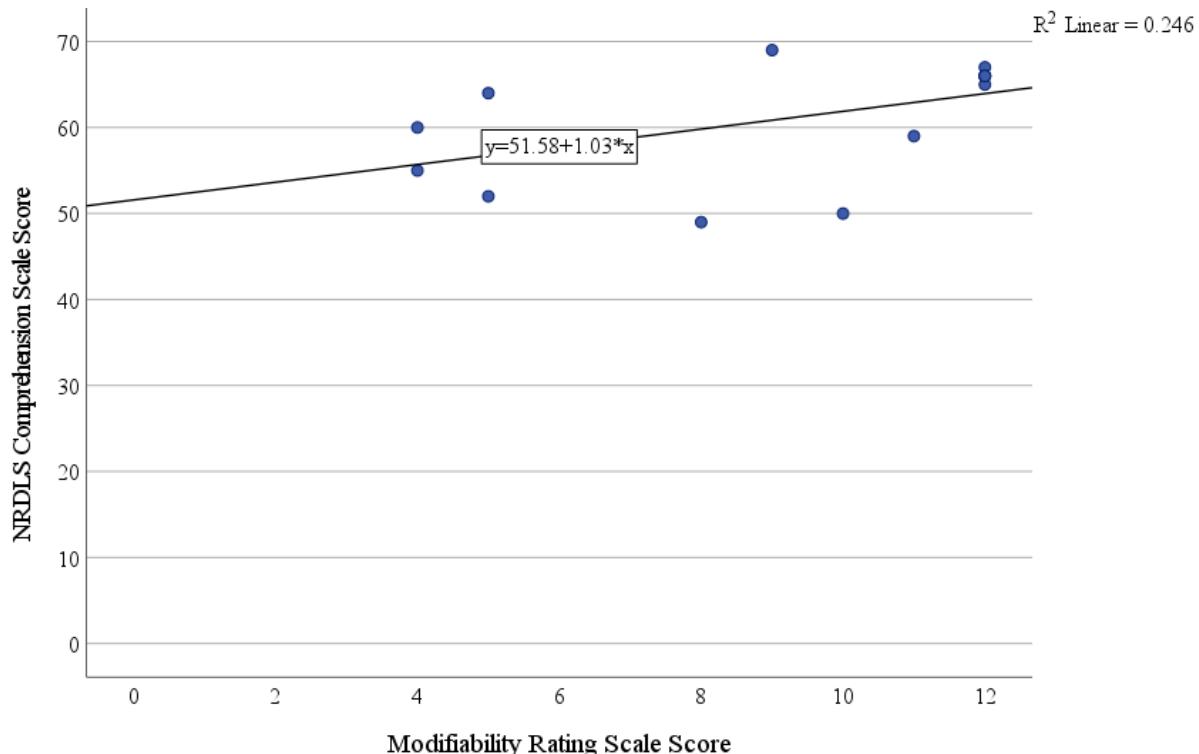


Figure 31. Scatterplot of the Correlation between the Modifiability Rating Scale Scores and the NRDLS Production Scale Scores

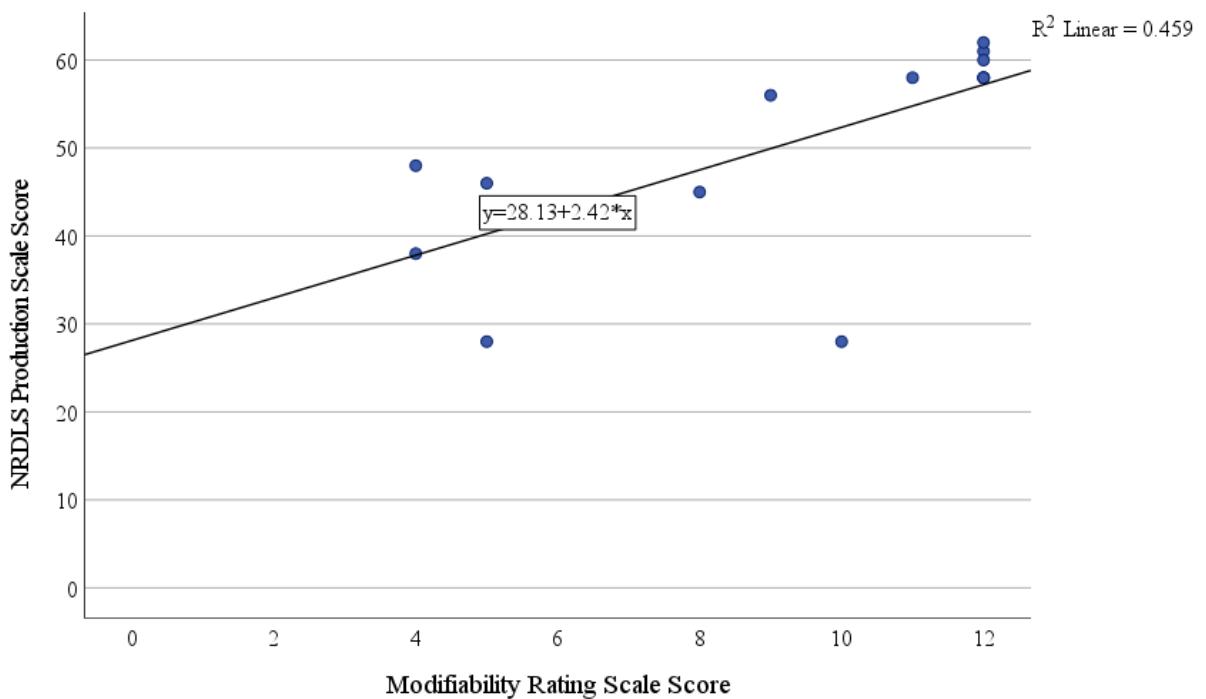
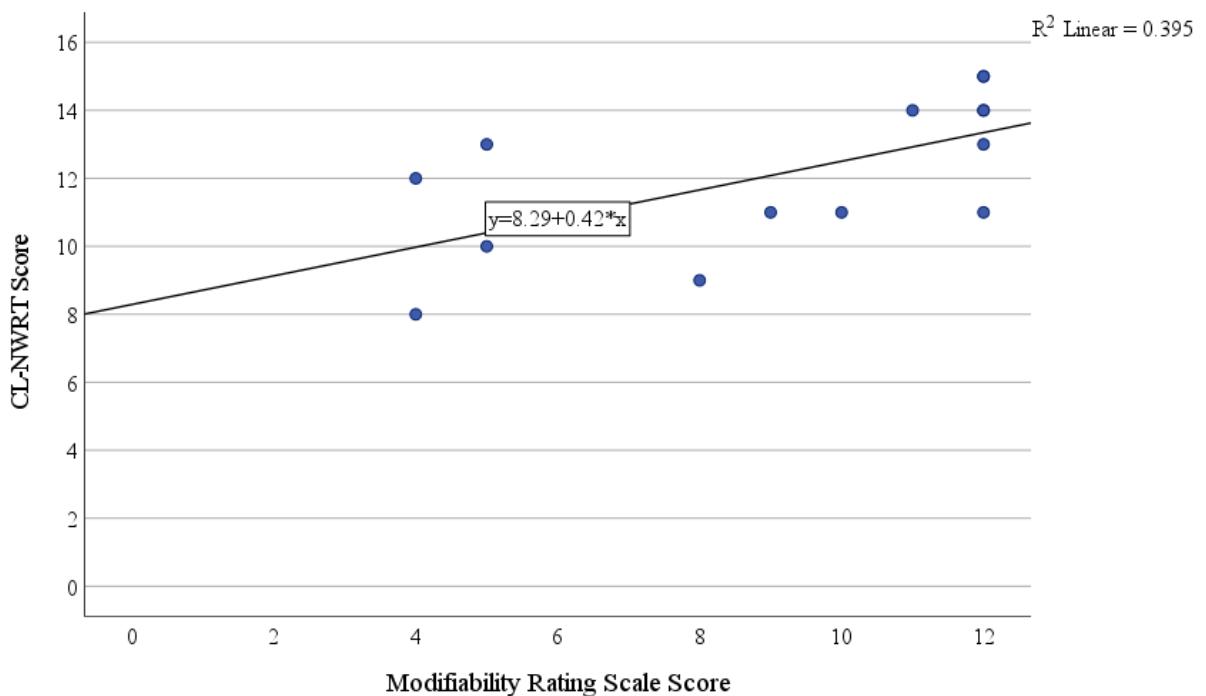


Figure 32. Scatterplot of the Correlation between the Modifiability Rating Scale Scores and the CL-NWRT Scores



5.3.4 Correlations between the Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Pearson's and Spearman's correlations were run to assess the relationships between the children's pretest, post-test, and change scores in the *Receptive Affective Prosody Task* and their *DLD Risk Factors Score*, *English Language Experience Score*, and *English Language Proficiency Stage* (see Tables 38A and 38B in Appendix 38). Spearman's correlations are prioritised due to normality deviations in the *Receptive Affective Prosody Task* and *English Language Proficiency Stage* data (see Section 5.3).

There were no correlations between the pretest scores and the *DLD Risk Factors Score*, $r_s(12) = .10, p = .734$; *English Language Experience Score*, $r_s(12) = -.31, p = .278$; and *English Language Proficiency Stage*, $r_s(12) = .43, p = .123$. Similarly, the post-test scores showed no correlations with the *DLD Risk Factors Score*, $r_s(12) = .11, p = .718$; *English Language Experience Score*, $r_s(12) = -.00, p = .989$; and *English Language Proficiency Stage*, $r_s(12) = .14, p = .646$. The change scores also did not correlate with the *DLD Risk Factors Score*, $r_s(12) = .27, p = .354$; *English Language Experience Score*, $r_s(12) = .22, p = .447$; and *English Language Proficiency Stage*, $r_s(12) = -.22, p = .446$.

Overall, there appears to be a lack of correlations between the children's pretest, post-test, and change scores in the *Receptive Affective Prosody Task* and their chance of presenting atypical language development, as well as their level of English language experience and competence.

5.3.5 Correlations between the Story Generation Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Story Grammar Units and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

The correlations between the children's pretest, post-test, and change scores in *Story Grammar Units* (DA's *Story Generation Task*) and their *DLD Risk Factors Score*, *English Language Experience Score*, and *English Language Proficiency Stage* were analysed using Pearson's and Spearman's correlations (see Tables 39A and 39B in Appendix 39). Given normality deviations in the *Story Grammar Units* and *English Language Proficiency Stage* data (see Section 5.3), Spearman's correlations are prioritised.

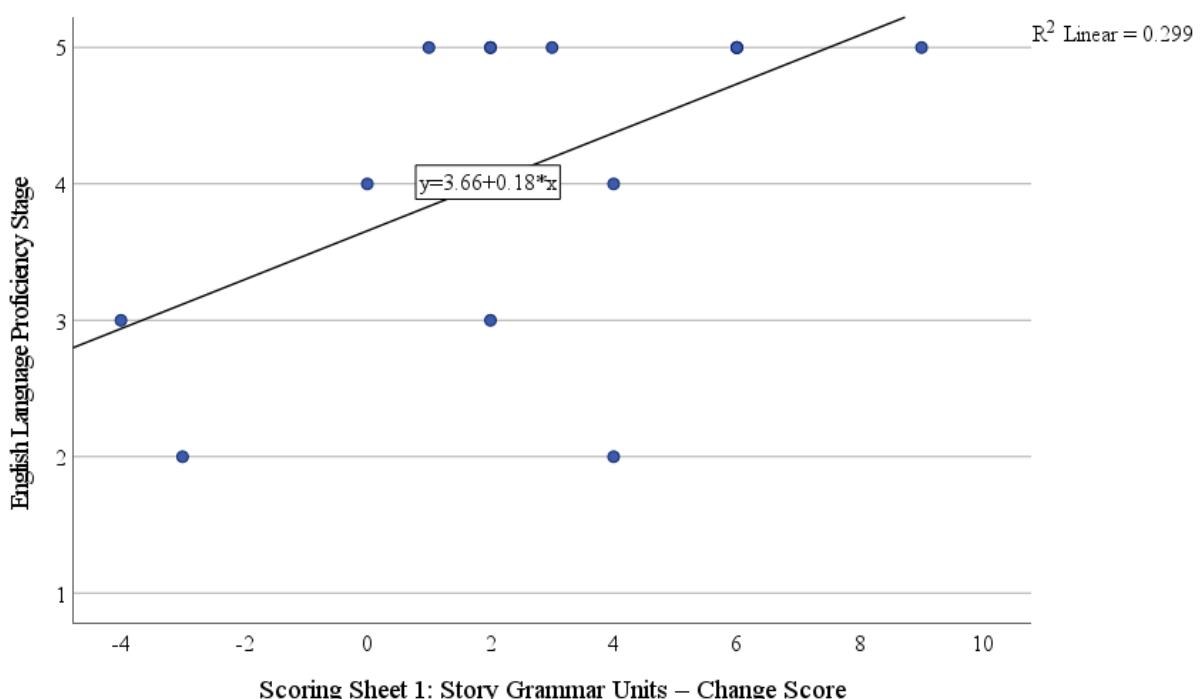
The pretest scores did not correlate with the *DLD Risk Factors Score*, $r_s(12) = .06, p = .850$; *English Language Experience Score*, $r_s(12) = .06, p = .838$; and *English Language*

Proficiency Stage, $r_s(12) = .15$, $p = .612$, suggesting that the *Story Grammar Units* pretest scores are unrelated to the risk of having DLD, and the degree of experience and proficiency in English.

The post-test scores did not correlate with the *DLD Risk Factors Score*, $r_s(12) = -.24$, $p = .404$; *English Language Experience Score*, $r_s(12) = -.15$, $p = .599$; and *English Language Proficiency Stage*, $r_s(12) = .40$, $p = .161$. Pearson's correlations supported these findings, except for a highly significant positive correlation with the *English Language Proficiency Stage* (see Table 39A). However, this relationship remains tentative due to the ordinal nature of the *English Language Proficiency Stage* variable and normality deviations in both this variable and *Story Grammar Units* data. No correlations emerged with English experience nor the risk of having DLD.

The change scores did not correlate with the *DLD Risk Factors Score*, $r_s(12) = -.39$, $p = .167$, and *English Language Experience Score*, $r_s(12) = -.34$, $p = .232$, but there was a trend towards a significant positive correlation with the *English Language Proficiency Stage*, $r_s(12) = .49$, $p = .073$ (Figure 33). Pearson's results were consistent, though they indicated that the positive correlation with the *English Language Proficiency Stage* was significant (see Table 39A). Therefore, the *Story Grammar Units* change scores seem somewhat correlated with English proficiency, but not with English experience or the risk of having DLD.

Figure 33. Scatterplot of the Correlation between the Change Scores in Scoring Sheet 1: Story Grammar Units and the English Language Proficiency Stage



Episodic Structure Complexity and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Only Spearman's correlations were run to study the relationships between the children's pretest, post-test, and change scores in *Episodic Structure Complexity* (DA's *Story Generation Task*) and their *DLD Risk Factors Score*, *English Language Experience Score*, and *English Language Proficiency Stage* (see Table 40A in Appendix 40). This was due to the ordinal nature of *Episodic Structure Complexity*.

Pretest ratings showed no correlations with the *DLD Risk Factors Score*, $r_s(12) = -.06, p = .842$; *English Language Experience Score*, $r_s(12) = -.14, p = .636$; and *English Language Proficiency Stage*, $r_s(12) = .11, p = .709$. Post-test ratings also did not correlate with the *DLD Risk Factors Score*, $r_s(12) = .32, p = .273$, and *English Language Experience Score*, $r_s(12) = -.03, p = .910$. However, a significant positive correlation was observed with the *English Language Proficiency Stage*, $r_s(12) = .57, p = .035$. Change scores were not correlated with the *DLD Risk Factors Score*, $r_s(12) = .34, p = .230$, and *English Language Experience Score*, $r_s(12) = -.09, p = .757$, but a significant positive correlation was found with the *English Language Proficiency Stage*, $r_s(12) = .54, p = .045$.

Consequently, while the *Episodic Structure Complexity* post-test and change scores appear to be associated with English proficiency, the *Episodic Structure Complexity* pretest, post-test, and change scores do not align with the prospect of having atypical language development, nor with English experience.

Emotional Vocabulary and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Pearson's and Spearman's correlations were run to examine how the children's pretest, post-test, and change scores in *Emotional Vocabulary* (DA's *Story Generation Task*) correlate with their *DLD Risk Factors Score*, *English Language Experience Score*, and *English Language Proficiency Stage* (see Tables 41A and 41B in Appendix 41). Due to normality deviations in the *Emotional Vocabulary* and *English Language Proficiency Stage* data (see Section 5.3), Spearman's correlations are prioritised.

The pretest scores showed no correlations with the *DLD Risk Factors Score*, $r_s(12) = -.25, p = .385$; *English Language Experience Score*, $r_s(12) = .28, p = .326$; and *English Language Proficiency Stage*, $r_s(12) = -.13, p = .669$, indicating that the *Emotional Vocabulary* pretest

scores are not associated with the possibility of having DLD, nor with English experience or proficiency.

The post-test scores showed no correlations with the *DLD Risk Factors Score*, $r_s(12) = -.45, p = .104$; *English Language Experience Score*, $r_s(12) = -.02, p = .959$; and *English Language Proficiency Stage*, $r_s(12) = .19, p = .523$. Pearson's results were consistent, except for a significant negative relationship with the *DLD Risk Factors Score* (see Table 41A). Thus, aside from a possible relationship with the risk of experiencing DLD, the *Emotional Vocabulary* post-test scores did not correlate with the English experience and proficiency levels.

The change scores showed no correlations with the *DLD Risk Factors Score*, $r_s(12) = -.21, p = .472$; *English Language Experience Score*, $r_s(12) = -.28, p = .327$; and *English Language Proficiency Stage*, $r_s(12) = .33, p = .251$, implying no consistent relationships between the *Emotional Vocabulary* change scores and the risk of atypical language development, or English experience or proficiency.

5.3.6 Correlations between the Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Pearson's and Spearman's correlations were conducted to evaluate how children's *Modifiability Rating Scale* scores relate to their *DLD Risk Factors Score*, *English Language Experience Score*, and *English Language Proficiency Stage* (see Tables 42A and 42B in Appendix 42). Spearman's correlations are prioritised due to normality deviations in the *Modifiability Rating Scale* and *English Language Proficiency Stage* data (see Section 5.3).

There were no correlations with the *DLD Risk Factors Score*, $r_s(12) = -.23, p = .423$; *English Language Experience Score*, $r_s(12) = -.25, p = .387$; and *English Language Proficiency Stage*, $r_s(12) = .38, p = .187$, suggesting that children's ability to respond to the teaching provided during the DA does not correlate with their risk of DLD, nor with their English experience or proficiency.

5.4 Multiple Regression Analyses between the Dynamic Assessment Measures and the Comparative Measures

To complement the correlational analyses discussed in Section 5.3, this section presents the results of the stepwise multiple regression analyses conducted using the forward selection method (entry criterion: $p \leq .05$). These analyses delve deeper into the relationships

highlighted in RQs and Sub-RQs 1 and 2 to identify which independent measure scores were significant predictors of DA performance. Independent variables included the NRDLS and CL-NWRT on one hand, and the *English Language Proficiency Stage* (treated as a continuous variable in line with previous research, e.g., Hessel & Strand, 2021), *DLD Risk Factors Score*, and *English Language Experience Score* on the other, whereas dependent variables were the DA tasks. Independent variables were grouped in this manner consistent with the RQs, rather than being included all at once, to mitigate potential issues such as overfitting and to enhance the robustness of the findings. Consequently, the NRDLS and CL-NWRT measures were analysed separately from variables related to English exposure, proficiency, and DLD risk factors. The forward selection method was chosen for its suitability in exploratory analyses, particularly relevant due to the DA's novelty and the study's pre-accuracy stage in the diagnostic research framework (see Section 4.1).

When considering the ordinal ratings from *Episodic Structure Complexity (Story Generation Task)* as the dependent variables, ordinal logistic regression analyses were conducted instead. The same independent variables specified above were included here, with the *English Language Proficiency Stage* treated as a categorical independent variable in this analysis.

The necessary assumptions for the multiple regression analyses were checked. There were no notable concerns regarding the independence and normal distribution of the values of the residuals, homoscedasticity, or influential cases biasing the model. While a Pearson's correlation suggested potential multicollinearity between the *NRDLS Comprehension* and *Production* scales ($r = .787$), this was not substantiated by tolerance values above 0.1 and VIF values below 10. The forward selection method further aided in addressing any potential multicollinearity issues. Additional assumptions for the ordinal regression analyses are specified in the relevant sections below.

5.4.1 Receptive Affective Prosody Task and the NRDLS and CL-NWRT

Forward selection multiple regression analyses were conducted for the *Receptive Affective Prosody Task* scores with the NRDLS and CL-NWRT as independent variables.

For the pretest scores, the model was significant, $F(1, 12) = 16.155, p = .002$, adj. $R^2 = .538$. The *NRDLS Production Scale* scores emerged as the only predictor ($\beta = .757, p = .002$), while the *NRDLS Comprehension Scale* and CL-NWRT scores did not contribute significantly to the model.

For the post-test scores, the model was also significant, $F(1, 12) = 5.149, p = .042$, adj. $R^2 = .242$. The *NRDLS Production Scale* scores were again the only predictor ($\beta = .548, p = .042$), with no significant predictive value from the *NRDLS Comprehension Scale* and CL-NWRT scores.

For the change scores, neither the NRDLS nor the CL-NWRT scores emerged as predictors.

5.4.2 Story Generation Task and the NRDLS and CL-NWRT

Story Grammar Units and the CL-NWRT and NRDLS

Forward selection multiple regression analyses were conducted for the *Story Grammar Units* scores using the NRDLS and CL-NWRT scores as independent variables.

The model for the pretest scores was significant, $F(1, 12) = 25.141, p < .001$, adj. $R^2 = .650$, with only the *NRDLS Production Scale* scores as a predictor ($\beta = .823, p < .001$). The *NRDLS Comprehension Scale* and CL-NWRT scores did not contribute significantly to the model.

For the post-test scores, two models emerged. In the first model, $F(1, 12) = 31.698, p < .001$, adj. $R^2 = .703$, the *NRDLS Production Scale* scores were the only predictor ($\beta = .852, p < .001$). In the second model, $F(1, 12) = 29.950, p < .001$, adj. $R^2 = .817$, both the scores in the *NRDLS Production Scale* ($\beta = 1.158, p < .001$) and CL-NWRT ($\beta = -.462, p = .014$) were found to be predictors. The *NRDLS Comprehension Scale* scores were not a predictor in either model.

Neither the NRDLS nor the CL-NWRT scores were predictors for the change scores.

Episodic Structure Complexity and the CL-NWRT and NRDLS

Ordinal logistic regression analyses were conducted for the *Episodic Structure Complexity* ratings with the NRDLS and CL-NWRT scores as independent variables. These variables were not predictors for the *Episodic Structure Complexity* pretest and change scores.

However, for the post-test ratings, the model significantly predicted the dependent variable over and above the intercept-only model, $\chi^2(3) = 8.780, p = .032$. An increase in the *NRDLS Production Scale* scores was associated with an increase in the odds of obtaining higher ratings in *Episodic Structure Complexity*, with an odds ratio of 1.349, 95% CI [1.000, 1.819], $\chi^2(1) = 3.838, p = .050$. Additionally, the CL-NWRT scores were marginally predictive (odds ratio = .339, 95% CI [.112, 1.025], $\chi^2(1) = 3.673, p = .055$). The assumption of proportional

odds was met, as assessed by a full likelihood ratio test comparing the fitted model to a model with varying location parameters, $\chi^2(6) = 9.363, p = .154$. The deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi^2(36) = 20.753, p = .980$, despite most cells being sparse with zero frequencies in 75.0% of cells, as expected given the continuous nature of the independent variables.

Emotional Vocabulary and the CL-NWRT and NRDLS

Forward selection multiple regression analyses were conducted to predict the *Emotional Vocabulary* scores using the NRDLS and CL-NWRT scores as independent variables. In these analyses, neither the NRDLS nor the CL-NWRT scores were predictors for any of the *Emotional Vocabulary* scores.

5.4.3 Modifiability Rating Scale and the NRDLS and CL-NWRT

Forward selection multiple regression analyses were run for the *Modifiability Rating Scale* scores with the NRDLS and CL-NWRT scores as independent variables. The model was significant, $F(1, 12) = 10.185, p = .008$, adj. $R^2 = .414$, with the *NRDLS Production Scale* scores as the only predictor ($\beta = .678, p = .008$). The *NRDLS Comprehension Scale* and CL-NWRT scores did not contribute significantly to the model.

5.4.4 Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Forward selection multiple regression analyses were run for the *Receptive Affective Prosody Task* pretest, post-test, and change scores with the *English Language Experience Score*, *DLD Risk Factors Score*, and *English Language Proficiency Stage* as independent variables. None of these variables emerged as predictors for any of the receptive affective prosody scores.

5.4.5 Story Generation Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Story Grammar Units and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Forward selection multiple regression analyses were conducted for the scores in *Story Grammar Units* using the *English Language Experience Score*, *DLD Risk Factors Score*, and *English Language Proficiency Stage* as independent variables.

Neither of these variables was a predictor for the pretest scores. However, for the post-test scores, the model was significant, $F(1, 12) = 13.379, p = .003$, adj. $R^2 = .488$, with the *English Language Proficiency Stage* as a predictor ($\beta = .726, p = .003$) and the *DLD Risk Factors Score* as a marginal predictor ($\beta = -.365, p = .075$). The *English Language Experience Score* did not add significantly to the model.

Two models emerged for the change scores. In the first model, $F(1, 12) = 5.129, p = .043$, adj. $R^2 = .241$, the *English Language Proficiency Stage* was the only predictor ($\beta = .547, p = .043$). In the second model, $F(1, 12) = 7.590, p = .008$, adj. $R^2 = .503$, both the *English Language Proficiency Stage* ($\beta = .702, p = .005$) and *DLD Risk Factors Score* ($\beta = -.552, p = .020$) were predictors. The *English Language Experience Score* was not a predictor in either model.

Episodic Structure Complexity and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

For the *Episodic Structure Complexity* ratings, ordinal logistic regression analyses were run considering the *English Language Experience Score*, *DLD Risk Factors Score*, and *English Language Proficiency Stage* as independent variables. These variables were not predictors for the *Episodic Structure Complexity* pretest and change scores.

For the post-test ratings, the model did not significantly predict the dependent variable over the intercept-only model, $\chi^2(5) = 6.270, p = .281$. However, level '2. Early Acquisition' within the *English Language Proficiency Stage* approached significance for the post-test ratings (odds ratio = .018, 95% CI [0.000, 1.561], $\chi^2(1) = 3.115, p = .078$). The assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the fitted model to a model with varying location parameters, $\chi^2(10) = 9.381, p = .496$. The deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi^2(34) = 23.262, p = .918$, despite most cells being sparse with zero frequencies in 75% of cells due to the continuous nature of two of the three independent variables: *English Language Experience Score* and *DLD Risk Factors Score*.

Emotional Vocabulary and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Forward selection multiple regression analyses were conducted for the scores in *Emotional Vocabulary* with the *English Language Experience Score*, *DLD Risk Factors Score*, and *English Language Proficiency Stage* as independent variables.

Neither the *English Language Experience Score*, *DLD Risk Factors Score*, nor *English Language Proficiency Stage* emerged as predictors for the pretest or change scores. However, for the post-test scores, two models were significant. In the first model, $F(1, 12) = 4.980, p = .045$, adj. $R^2 = .234$, the *DLD Risk Factors* was the only predictor ($\beta = -.542, p = .045$). In the second model, $F(1, 12) = 5.796, p = .019$, adj. $R^2 = .425$, the *English Language Proficiency Stage* ($\beta = .488, p = .048$) and *DLD Risk Factors Score* ($\beta = -.678, p = .010$) were predictors. The *English Language Experience Score* was not a predictor in either model.

5.4.6 Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Forward selection multiple regression analyses were run for the *Modifiability Rating Scale* scores using the *English Language Experience Score*, *DLD Risk Factors Score*, and *English Language Proficiency Stage* as independent variables. None of these variables were identified as predictors for any of the *Modifiability Rating Scale* scores.

5.5 Conclusions

This study continued to evaluate the effectiveness of the newly developed DA to detect DLD risk among the UK's multilingual child population. Fourteen 4;06–8;11-year-old children with EAL from Newcastle and its vicinity took part. In line with the proposed RQs, correlational and regression analyses were employed to explore the relationships between these children's performance in the DA and comparative measures relevant to identifying DLD. Uncovering these was key for drawing conclusions about the DA's current utility to help understand the nature of a child's English language struggles (i.e., language disorder versus temporary difficulty due to limited experience with the language).

In addressing RQ and Sub-RQ 1, various significant relationships were observed between the children's outcomes in the DA and the NRDLS, suggestive of English proficiency influences on DA outcomes, and CL-NWRT, which reflected the DA's capability to reveal underlying language learning ability.

On the one hand, within the DA's *Story Generation Task* measures, the *Story Grammar Units* pre- and post-test scores significantly and positively correlated with the NRDLS results, especially with the *NRDLS Production Scale*, which positively predicted such scores. However, the *Story Grammar Units* change scores did not correlate with the NRDLS scores, nor were they predicted by them. The *Episodic Structure Complexity* pretest scores trended towards a significant positive correlation with the *NRDLS Production Scale* scores, but this

pattern did not extend to post-test or change scores. Still, the *NRDLS Production Scale* scores did positively predict the *Episodic Structure Complexity* post-test scores. The *Emotional Vocabulary* post-test scores showed a possible significant and positive correlation with the *NRDLS Production Scale*, with no correlations observed between its pretest and change scores with the NRDLS scores. The NRDLS scores were not predictors for the *Emotional Vocabulary* scores either. Additionally, the *Receptive Affective Prosody* pre- and post-test scores, along with the *Modifiability Rating Scale* scores, all significantly and positively correlated with the NRDLS scores, particularly with the *NRDLS Production Scale*, which also positively predicted these scores.

On the other hand, the CL-NWRT scores did not correlate with any of the *Story Grammar Units* scores but negatively predicted the *Story Grammar Units* post-test scores. Similarly, no correlations emerged between the CL-NWRT scores and the *Episodic Structure Complexity* scores, but the CL-NWRT scores were a marginal negative predictor for the *Episodic Structure Complexity* post-test scores. The CL-NWRT scores did not correlate with any of the *Emotional Vocabulary* scores and did not predict any of these either. However, they did present significant positive correlations with the *Receptive Affective Prosody Task* pretest scores and a possible trend towards a positive significant correlation with the post-test scores. The *Receptive Affective Prosody* change scores, however, were not correlated with the CL-NWRT scores, and these did not predict any *Receptive Affective Prosody* scores. Finally, the *Modifiability Rating Scale* scores also correlated significantly and positively with the CL-NWRT scores, but these did not predict the former.

Exploring RQ and Sub-RQ 2 uncovered various links between children's DA performance and their *English Language Proficiency Stage* and *English Language Experience Score* — indicating whether the DA outcomes are affected by exposure to and proficiency in English —, and *DLD Risk Factors Score*, suggesting that the DA results can somewhat reflect the likelihood of having DLD (i.e., underlying language learning ability).

Children's *DLD Risk Factors Score* showed a possible significant negative correlation with their *Emotional Vocabulary* post-test scores, negatively predicting them as well. They also emerged as a marginal negative predictor for the *Story Grammar Units* post-test scores, and a negative predictor for the *Story Grammar Units* change scores. Yet, no further correlations were found between the *DLD Risk Factors Score* and the other DA outcomes, nor was it a predictor for any of these. On a different note, the *English Language Experience Score* did not correlate with nor predict the results in any of the DA measures, whereas the *English*

Language Proficiency Stage did show some trend towards a significant and positive correlation with the *Story Grammar Units* change scores and a tentative positive correlation with the *Story Grammar Units* post-test scores. The *English Language Proficiency Stage* also positively predicted these *Story Grammar Units* scores. There was also a significant positive relationship between the *English Language Proficiency Stage* and the *Episodic Structure Complexity* post-test and change scores. Children's English proficiency further positively predicted their *Emotional Vocabulary* post-test scores, but no additional correlations were identified with the other DA outcomes, and this did not predict any of these.

These findings, summarised in Tables 30 and 31, are further elaborated in *Chapter 6*, along with those related to the overall impact of the DA's teaching phase on enhancing performance across the participating children. The implications of what these findings mean in terms of the DA's preliminary capacity for evaluating DLD risk in the UK's multilingual children with diverse English language abilities are discussed.

Table 30. Overview of Significant Correlations between the Children's Performance in the Dynamic Assessment Measures and Comparative Measures

		NRDLS Comp.		NRDLS Prod.		CL-NWRT		DLD Risk		Eng Exp		Eng Prof	
		P	S	P	S	P	S	P	S	P	S	P	S
RAP	Pretest	**+	***+	**+	**+	**+	**+	*					
T	Post-test			*+	Trend		Trend						
	Change												
SS1	Pretest	*+	*+	**+	*+								
	Post-test	*+		**+	*+							***+	
	Change											*+	Trend
SS2	Pretest	.	.	.	Trend	
	Post-test	*+
	Change	*+
SS3	Pretest												
	Post-test			Trend	*+	*+				*-			
	Change												
Modifiability	Trend	*+	*+	***+	**+	*+	**+						

Note. All correlations are 2-tailed. Abbreviations: ..: Not applicable; *-: Negative correlation significant at the .05 level; *+: Positive correlation significant at the .05 level; **+: Positive correlation highly significant at the .01 level; Trend *+: Positive correlation showing a trend toward significance with $.1 > p > .05$; P: Pearson; S: Spearman; CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; Modifiability: *Modifiability Rating Scale* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; RAPT (pretest, post-test and change): *Receptive Affective Prosody Task* scores; SS1 (pretest, post-test and change): *Scoring Sheet 1: Story Grammar Units* scores; SS2 (pretest, post-test and change): *Scoring Sheet 2: Episodic Structure Complexity* scores; SS3 (pretest, post-test and change): *Scoring Sheet 3: Emotional Vocabulary* scores.

Table 31. Overview of Comparative Measures Showing as Predictors for the Dynamic Assessment Measures in the Regression Analyses

		NRDLS Comp.	NRDLS Prod.	CL-NWRT	DLD Risk	Eng Exp	Eng Prof
RAPT	Pretest		Predict +				
	Post-test		Predict +				
	Change						
SS1	Pretest		Predict +				
	Post-test		Predict +	Predict –	Margin –	Predict +	
	Change				Predict –		Predict +
SS2	Pretest						
	Post-test		Predict +	Margin –			
	Change						
SS3	Pretest						
	Post-test				Predict –		Predict +
	Change						
Modifiability		Predict +					

Note. Abbreviations: +: Positive direction; -: Negative direction; Margin: Marginal predictor ($.1 > p > .05$); Predict: Predictor ($p \leq .05$); CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; Modifiability: *Modifiability Rating Scale* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; RAPT (pretest, post-test and change): *Receptive Affective Prosody Task* scores; SS1 (pretest, post-test and change): *Scoring Sheet 1: Story Grammar Units* scores; SS2 (pretest, post-test and change): *Scoring Sheet 2: Episodic Structure Complexity* scores; SS3 (pretest, post-test and change): *Scoring Sheet 3: Emotional Vocabulary* scores.

Chapter 6. Discussion: Main Study of the Dynamic Assessment

6.1 Introduction

The UK's diverse multilingual environment can make it difficult to differentiate Developmental Language Disorder (DLD) from typical multiple language acquisition among children with English as an Additional Language (EAL). As explained in Section 2.5.1, some major obstacles include the absence of appropriate multilingual assessments covering English and the children's home languages, and a scarcity of practitioners proficient in these languages (Oxley et al., 2019). This often leads to misallocated speech and language therapy provision and delayed interventions (Mennen & Stansfield, 2006; Oxley et al., 2019), highlighting an urgent need for more effective assessment within the UK's educational and healthcare systems, which face limited resources, particularly for multilingual children (Flynn & Curdt-Christiansen, 2018; Oxley et al., 2019). The COVID-19 pandemic may have added further strain in this sense (Clegg, O'Flynn & Just, 2021). In response, there is an increasing interest in Dynamic Assessment (DA), grounded in the works of Vygotsky (1978) and Feuerstein (1980), as a complementary method for diagnosing DLD in English-speaking multilingual children based on their underlying language learning potential rather than a snapshot of their language skills (Hasson & Joffe, 2007; Hunt et al., 2022; Orellana et al., 2019).

This chapter interprets the findings of the main study within this PhD research, previously presented in *Chapter 5*. The study investigates the preliminary usefulness of a novel DA in assessing DLD risk in children with EAL, contributing to the emerging body of DA research with DLD diagnostic purposes in multilingual child populations. Following a series of pilot studies detailed in *Chapter 3*, the sample in this study consisted of 14 children aged 4;06–8;11 years from Newcastle upon Tyne and nearby areas, all of whom have EAL and speak a range of home languages. The DA, which uses a pretest–teach–post-test format, investigates the potential to learn through storytelling in English in three DLD-sensitive areas: narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and receptive affective prosody. This aims to help determine DLD risk on the basis that while a typically developing child would show notable performance gains with little assessor effort, indicating high modifiability, a child with DLD would exhibit only small improvements, suggesting low modifiability and greater assessor support needs. In this study, the DA's effectiveness in detecting DLD risk was evaluated in relation to several comparative measures relevant to differentiating DLD from lack of opportunity to learn English: the language-

neutral *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015; Chiat et al., 2020), the *New Reynell Developmental Language Scales* (NRDLS; Edwards et al., 2011) in English, and children's *English Language Proficiency Stage*, *English Language Experience Score*, and *DLD Risk Factors Score*.

The present chapter begins by discussing the findings concerning the impact of the DA's teaching phase on the sample's performance (see Section 6.2). Next, aligning with the research questions, the relationships observed between the children's performance in the DA and the comparative measures are interpreted in Section 6.3, explaining their significance for the DA's initial capacity for evaluating DLD risk in UK multilingual children with varied English language abilities. Following that, Section 6.4 addresses the study's implications and limitations, reflecting on the challenges encountered and proposing directions for future research on the DA resource.

6.2 Teaching Phase Impact on Dynamic Assessment Performance: A Pretest–Post-test Comparison

The effectiveness of the DA in measuring teaching-induced changes in ability among children with EAL was first evaluated by analysing the impact of the teaching phase on the 14 participants' performance. The results of a Wilcoxon signed-rank test and repeated measures t-tests revealed varying levels of impact across different skill domains from pretest to post-test. In the *Receptive Affective Prosody Task*, a slight, non-significant improvement was noted following the teaching phase, suggesting a limited effect of this phase on enhancing children's abilities to interpret emotional cues in voices. In contrast, there was a moderate significant improvement in children's ability to include story grammar units in their narratives, as assessed by *Scoring Sheet 1: Story Grammar Units* as part of the *Story Generation Task*. However, within the same task, only small, non-significant changes were observed in the complexity of the stories' episodic structures and emotional vocabulary usage, as measured by *Scoring Sheet 2: Episodic Structure Complexity* and *Scoring Sheet 3: Emotional Vocabulary*, respectively.

These mixed outcomes, particularly the limited improvements in some areas, could be partially attributed to a ceiling effect. Several children had already attained the highest possible scores in the pretest for *Receptive Affective Prosody* ($n = 5$), *Episodic Structure Complexity* ($n = 6$), and *Emotional Vocabulary* ($n = 5$), which restricted the scope for observing marked improvement post-teaching. The small sample size may have exacerbated this ceiling effect, potentially overstating the perceived performance level across participants.

relative to what might be seen in a larger group. Importantly, however, considering the relatively wide age range of the sample (4;06–8;11 years), a consequence of accommodating a greater amount of interested participants amid recruitment difficulties—and understanding that their performance would offer valuable data into the DA’s suitability for different ages—it was observed that achieving high scores was not confined to older children, as might have been expected. This suggests a need to further adjust the difficulty levels of measures related to *Receptive Affective Prosody*, *Episodic Structure Complexity*, and *Emotional Vocabulary*, as well as the associated stimuli and teaching strategies for these areas, to better match each child’s starting abilities, aiming for a more precise representation of the spectrum of skills across these areas at various developmental stages.

This need for adjustment leads to the consideration of limited scoring ranges within these three DA measures that may have facilitated the higher scores. In contrast to *Scoring Sheet 1: Story Grammar Units*, which allows for a broader range of scores with a maximum of 30 points, the scoring capacity for the *Receptive Affective Prosody Task* (maximum of 8 points), *Scoring Sheet 2: Episodic Structure Complexity* (4 levels), and *Scoring Sheet 3: Emotional Vocabulary* (4 points) appears constrained. As described in Section 3.3.2, revisions to *Scoring Sheets 1* and *3* were implemented following the initial piloting to improve their precision in capturing children’s story grammar and emotional vocabulary skills. Furthermore, to avoid overburdening the children, the original scoring sheet for the *Receptive Affective Prosody Task* was divided into two, each allocated 8 points for the pre- and post-tests, as opposed to the initial 16 points (see ‘Pilot testing 3: Data collection procedure’ in Section 3.2.3). *Scoring Sheet 2: Episodic Structure Complexity*, however, was not changed.

Increasing the scoring scale for *Scoring Sheet 2: Episodic Structure Complexity*, for instance, could be achieved by introducing additional levels beyond the current 4-point scale (i.e., ‘Level 1: Description’, ‘Level 2: Sequence’, ‘Level 3: Incomplete Episode’, and ‘Level 4: Complete Episode’ based on the plan-attempt-outcome sequence; Westby, 2005). This could be realised by expanding the single-episode stimuli stories used in the DA’s storytelling activity. As exemplified in the MAIN (Gagarina et al., 2012, 2019), where each story depicts three short episodes, and the ENNI (Schneider et al., 2005), which progresses from simple stories with a single episode to complex stories with three episodes, this adaptation would afford children multiple opportunities to demonstrate each targeted story grammar unit, including plans, attempts and outcomes. Conducting this more detailed evaluation of children’s abilities to produce complete episodes in their narratives could be beneficial, especially for those at more advanced stages in developing these story-structuring skills.

Having more than one episode per stimulus story would also allow for a more thorough exploration of children's learning capabilities across both story grammar and emotional vocabulary in *Scoring Sheets 1* and *3*, thereby aiding in a more effective detection of DLD risk through the DA. In addition to featuring several episodes, the integration of subplots within the stimuli stories could also be explored as a means to broaden the scope for evaluating children's *Story Generation Task* performance. However, a careful balance should be maintained to avoid excessively lengthening the DA session, ideally restricting this to under an hour so that it is comparable to traditional language assessments, as prolonged administration time could challenge the DA's applicability in real practice (Hasson & Joffe, 2007; Hunt et al., 2022).

The study by Lazewnik et al. (2019) with Spanish–English speaking 4–5-year-olds reflects the importance of tailoring the DA to each child's learning stage to show their growth potential, setting appropriately challenging skill targets that extend beyond their current abilities. As discussed under 'Using dynamic assessment of language learning with diagnostic aims' (Section 2.5.2), Lazewnik et al.'s (2019) results did not distinguish between children with DLD and those without, with both groups initially scoring within normal limits (11/15 in the typically developing group and 10/15 in the language-impaired group). Lazewnik et al. applied Peña et al. (2001)'s pretest–teach–post-test DA protocol to target vocabulary, which had proved successful in identifying low-language ability groups. However, unlike in that study, they did not differentiate between known and unknown words to the children in their pre- and post-test evaluations, contributing to the DA's inability to discern between groups, which resulted in it not being useful in identifying DLD in these children.

In line with these considerations, another proposal to adapt the DA to children's individual skill stages would involve conducting a follow-up DA session for those who encountered difficulties during the initial post-test and who may benefit from a second teaching phase and post-test to demonstrate their abilities, as explained in 'Pilot testing 2: Results and discussion' (Section 3.3.4). The same story from the original *Story Generation Task* post-test —either *Story A*, which involves a girl and her cat, or *Story B*, about a man and a bird — in the first DA session would be reused here for instruction to maintain continuity, while *Story C*, featuring elephants, would be introduced in the new post-test.

Shifting the focus to the DA's teaching methods, the effectiveness of the teaching phase might have been influenced by the inherent complexity of the targeted skills. For example, fostering social communication-related behaviours, such as emotion awareness and understanding,

requires considerable time and effort (Fujiki & Brinton, 2017). This suggests that teaching receptive affective prosody, which entails interpreting emotional information conveyed through vocal cues, may necessitate more prolonged, focused practice to achieve more meaningful gains than what was provided in the DA's current teaching phase. As elaborated in Section 3.3.2, this phase addressed affective prosody on two different occasions through an activity that involved listening to two audio clips of the same sentence, each articulated with congruent or incongruent emotional prosody. Children were asked to discern which voice best represented the character's emotions in response to the problem and outcome of the story and to imitate what the character said using an emotionally congruent voice. Incorporating a slightly longer, more in-depth activity that provides additional opportunities to pay attention to emotional prosodic cues and practice distinguishing between various emotional tones could more successfully reveal children's abilities to enhance their vocal emotion comprehension from pretest to post-test, thus enabling a more accurate assessment of DLD risk via the DA.

In summary, this study revealed notable improvements in children's ability to include story grammar units into their narratives over the DA's teaching phase, though only small advances in producing complete episodes and employing emotional vocabulary within the stories, as well as processing affective prosody. The constraints of a small sample size, alongside the presence of a ceiling effect and the potential inadequacies of the DA's scoring and teaching strategies, emphasise the need for additional testing of this resource with a larger sample and its further refinement to cater to the varied learning potentials of children at different developmental stages.

6.3 Relationship between the Dynamic Assessment Measures and the Comparative Measures

Building on the review of the DA's teaching phase impact on the 14 children's performance in Section 6.2, this section discusses findings from correlational and regression analyses that test the DA's effectiveness as a tool for identifying DLD risk in children with EAL by exploring the relationships between their outcomes in the DA (*Modifiability Rating Scale, Receptive Affective Prosody Task, and Story Generation Task*) and the NRDLS and CL-NWRT to address Research Question (RQ) 1. These analyses also focused on the DA's relationships with the *DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage* for RQ2. In each instance, the DA elements showing the strongest relationships with these comparative measures were identified (Sub-RQs 1 and 2).

These exploratory analyses shed light on the DA's current strengths and limitations in revealing learning capacity to differentiate DLD-related English language struggles from those due to limited experience with this language. Insights into how a child's existing English language skills and experience may affect the DA's diagnostic accuracy for assessing learning potential were gained from its relationships with the NRDLS, *English Language Experience Score*, and *English Language Proficiency Stage*. Meanwhile, analysing the DA's relationships with the CL-NWRT and *DLD Risk Factors Score* was key to determining its utility in detecting the likelihood of DLD among children with EAL.

6.3.1 Dynamic assessment measures and NRDLS

The correlational and predictive relationships observed between the NRDLS scores and specific DA outcomes highlight the importance of further adjusting the DA's methods to measure a child's language learning potential more clearly without excessive interference from their English language abilities.

Focusing first on the DA's *Story Generation Task*, while no significant relationships emerged for the *Story Grammar Units* and *Episodic Structure Complexity* change scores, the *Story Grammar Units* results, both in the pre- and post-tests, were positively associated with the NRDLS scores, especially with the *NRDLS Production Scale*, which positively predicted these scores. Additionally, a trend towards a positive link was noted between the *Episodic Structure Complexity* pretest scores and the *NRDLS Production Scale* scores, also in the direction of positively predicting the post-test scores for this measure. These relationships may be capturing the intertwined nature of narrative macrostructure development and linguistic skills. As evidenced in both mono- and multilingual children, producing a coherent, well-structured narrative is, to some extent, contingent upon the integration of lexical and grammatical skills, with studies showing that children with advanced syntactic competency (monolingual: Gardner-Neblett, 2022; multilingual: Bitetti, Hammer & López, 2020; Chan, Chen, Hamdani, Tse & Cheng, 2023) and higher vocabulary knowledge (monolingual: Khan, Logan, Justice, Bowles & Piasta, 2021, multilingual: Bitetti et al., 2020; Chan et al., 2023; Fiani, Henry & Prévost, 2022; Lindgren & Bohnacker, 2022) are more adept at organising their narratives, as they can draw upon a broader lexicon for story construction and make better use of syntax to link different story parts.

Indeed, while macrostructure constitutes the overarching organisation of a story shared across languages, this is not entirely language-independent and relies on foundational language skills at the microstructural level. Therefore, it is essential to refine the DA to account for the

interplay between a child’s English language skills and their capacity to develop story grammar and episodic structure complexity skills. Practical steps in this direction could be introducing more episodes into the stimuli stories of the DA’s storytelling activity, as well as additional teaching and post-test phases in a second DA session, as proposed in Section 6.2. Adjusting this would allow the assessment to be more attuned to individual English proficiency levels by offering children increased opportunities to show their narrative macrostructure learning skills without heavily relying on their current English language abilities.

Expanding the stimuli stories to include more episodes, and/or incorporating more teaching and post-test phases in the DA’s storytelling activity, could also lead to a more reliable understanding of children’s abilities to use basic emotional terms (e.g., ‘happy’, ‘angry’) within their narratives, moving beyond the constraints set by their existing English language skills. This amendment would be motivated by the possible positive association detected between the *NRDLS Production Scale* scores and the *Emotional Vocabulary* post-test scores. While the NRDLS scores did not predict any of this measure’s outcomes and no correlations were found with the NRDLS for the pretest and change scores, the potential relationship after the teaching phase hints at a parallel between children’s language production skills and their ability to refer to story characters’ emotional states, a key aspect of narrative macrostructure (monolingual: Khan et al., 2021; multilingual: Chan et al., 2023; Fiani et al., 2022). Recognising this insight is important for improving the DA’s capacity to accurately assess children’s use of emotional vocabulary while minimising the influence of their English language skills.

The positive correlations between the *Receptive Affective Prosody* pre- and post-test scores and performance in the NRDLS —more so in the *NRDLS Production Scale*, which also positively predicted these scores— suggest that children’s ability to interpret prosodic cues pertinent to core emotional states (e.g., sadness, fear) is linked, to some extent, with their English language competence. This echoes findings from studies showing that linguistic abilities and sociocultural factors —such as preferences in how emotions should be vocally expressed to conform to social display rules— enhance sensitivity to emotional prosody in one’s native language (Pell, Monetta, Paulmann & Kotz, 2009b), a development which emerges between the ages of 5 and 8 (Chronaki, Wigelsworth, Pell & Kotz, 2018; Ma et al., 2022). The observation of this native-language advantage in processing vocally conveyed emotions co-occurs with evidence on the universality of affective prosody recognition from early childhood, independent of word meaning (Chronaki et al., 2018; Ma et al., 2022). While

emotional states are displayed similarly across cultures, with vocal emotions containing pan-cultural perceptual properties that enable the perception of essential emotions across languages, there is some cultural variation in how emotions are vocally expressed (van Rijn & Larrouy-Maestri, 2023), and cultural and linguistic familiarity can influence this recognition process (Chronaki et al., 2018; Ma et al., 2022; Pell et al., 2009b). To enhance the assessment of children's capability to recognise targeted emotions through prosody, it may be beneficial to transition from using semantically neutral sentences in English to pseudo-utterances without meaningful linguistic content as the *Receptive Affective Prosody Task* stimuli (Chronaki et al., 2018; Pell et al., 2009b). Such a shift would aim to further minimise the effect of English proficiency, centring the evaluation on prosodic interpretation abilities.

Lastly, the positive link between the *Modifiability Rating Scale* scores and NRDLS results—particularly pronounced with the *NRDLS Production Scale*, which positively predicted these scores—indicates that children with higher English language skills are more likely to respond successfully to the DA's teaching phase. This observation concurs with the findings discussed above, reiterating the need for a more balanced DA approach that accommodates the range of English language abilities among the children, lessening the impact of these on their capacity to engage with instruction and enabling more precise detection of DLD risk in children with EAL.

6.3.2 Dynamic assessment measures and English Language Proficiency Stage

Correlational and predictive relationships were identified between the *English Language Proficiency Stage* and certain *Story Generation Task* outcomes. Specifically, a trend towards a positive correlation was noted between the *English Language Proficiency Stage* and *Story Grammar Units* post-test and change scores, with the former also positively predicting these scores. Moreover, a positive association was found with the *Episodic Structure Complexity* post-test and change scores, and the *English Language Proficiency Stage* also positively predicted the *Emotional Vocabulary* post-test scores.

Mirroring the findings related to the NRDLS in Section 6.3.1, these reinforce the importance of adjusting the DA's methods further to gauge children's potential for expanding narrative macrostructure and emotional vocabulary skills through DA instruction, distinctly from their English proficiency. Adding more episodes in the stimuli stories could assist with this process, along with supplementary teaching and post-test phases (see Section 6.2).

6.3.3 Dynamic assessment measures and English Language Experience Score

Contrary to expectations derived from research with multilingual children —where increased rich English language input and prolonged exposure are linked to both higher narrative macrostructure (Govindarajan & Paradis, 2019) and vocabulary skills in this language (Paradis, 2011; Paradis & Jia, 2017; Sorenson Duncan & Paradis, 2020)— no correlational or predictive relationships were found between the *English Language Experience Score* and the *Story Generation Task* outcomes.

In fact, this lack of relationships concerning the *English Language Experience Score* extended to all other DA measures. While this outcome supports the DA's value in assessing language learning potential independently of a child's accumulated English language experience, it also conflicts with observations connecting children's English language skills with their performance in certain DA measures, as captured by the NRDLS (Section 6.3.1) and *English Language Proficiency Stage* (6.3.2). Discrepancies like this may suggest that these proficiency measures tap into aspects of language ability that are more closely aligned with the DA's targets, such as the NRDLS' assessment of vocabulary and grammatical aspects that may contribute to the production of narrative macrostructure features (e.g., Bitetti et al., 2020), as opposed to the *English Language Experience Score*, which reflects broader features of language usage and exposure. While we would expect these language experience factors to contribute to language proficiency, they may not be as directly relevant to the specific skills the DA seeks to measure.

6.3.4 Dynamic assessment measures and CL-NWRT

The correlational and predictive relationships between certain DA outcomes and the CL-NWRT scores —a potential diagnostic marker for DLD focusing on phonological short-term memory and processing, less reliant on language-specific knowledge (Boerma & Blom, 2017; Chiat, 2015)— offer insights into the DA's sensitivity to detect DLD-related vulnerabilities and its preliminary value for identifying DLD risk among children with EAL.

In the *Story Generation Task*, although no correlations were observed between the CL-NWRT scores and the *Story Grammar Units* scores, the former negatively predicted this measure's post-test scores. A similar pattern was observed with the *Episodic Structure Complexity* scores, where, despite the absence of correlations with CL-NWRT scores, a marginal negative prediction was noted for this measure's post-test scores, hinting at a complex relationship between nonword repetition skills and narrative macrostructure abilities. Furthermore, no

significant relationships were observed with the *Emotional Vocabulary* scores, despite known associations between vocabulary acquisition and nonword repetition skills, which involve the temporary storage and retrieval of novel strings of phonemes, akin to the process of learning new words (Gathercole, 2006; Melby-Lervåg et al., 2012).

For the *Receptive Affective Prosody Task*, positive correlations were noted with the CL-NWRT scores in the pretest, with a possible near-significant positive correlation in the post-test. In this manner, robust nonword repetition skills may correspond with a better understanding of emotional prosody, potentially signalling a lower risk of DLD on the basis that this disorder can affect both nonword repetition (Boerma & Blom, 2017; Chiat, 2015) and receptive affective prosody (Griffiths et al., 2020; Löytömäki et al., 2020; Taylor et al., 2015). However, neither correlational nor predictive relationships were observed between the *Receptive Affective Prosody* change scores and CL-NWRT scores. This, along with the less robust correlation in the post-test compared to the pretest, may be related to the DA's teaching phase having a limited impact on improving affective prosody comprehension, which in turn also affected the initial link between these skills and nonword repetition abilities.

To conclude, a positive correlation was noted between the *Modifiability Rating Scale* scores and CL-NWRT scores, although the latter did not predict the former. Stronger nonword repetition skills could, therefore, be consistent with increased responsiveness to the DA's teaching phase, suggesting the receptiveness of the *Modifiability Rating Scale* to identifying DLD risk in multilingual children, mirroring previous research (Peña et al., 2006, 2014; Petersen et al., 2017; Ukrainetz et al., 2000).

6.3.5 Dynamic assessment measures and DLD Risk Factors Score

The relationships between the *DLD Risk Factors Score* and specific DA measures, alongside those observed in Section 6.3.4 regarding the CL-NWRT scores, illuminate the DA's capacity for determining the likelihood of a multilingual child experiencing DLD risk.

Focusing on the DA's *Story Generation Task*, the *DLD Risk Factors Score* marginally negatively predicted the *Story Grammar Units* post-test scores. It also negatively predicted this measure's change scores. Similarly, the *DLD Risk Factors Score* possibly negatively correlated with, and also negatively predicted, the *Emotional Vocabulary* post-test scores. These relationships imply that an increased DLD risk in children is linked to reduced abilities in integrating story grammar units and emotional vocabulary into narratives, consistent with DLD's recognised detrimental effects in both domains (Boerma et al., 2016; Govindarajan &

Paradis, 2019). Despite the lack of further predictive or correlational relationships between the *DLD Risk Factors Score* and other DA measures, including *Episodic Structure Complexity*, *Receptive Affective Prosody*, and *Modifiability*, which was possibly attributed to the constrained statistical power from the small sample size, the results underscore the DA's potential value in signalling DLD in children with EAL.

6.3.6 Conclusions on the dynamic assessment and comparative measures relationships

The correlational and predictive relationships observed between the children's outcomes in the DA and the CL-NWRT and *DLD Risk Factors Scores* indicate that the DA holds some value for detecting DLD in multilingual children as a resource sensitive to DLD-related vulnerabilities. However, the DA's current effectiveness is constrained by its dependency on the children's English language skills and exposure, as evidenced by the relationships between the performances in the DA and the NRDLS and *English Language Proficiency Stage*, despite a lack of associations involving the *English Language Experience Score*. The influence of pre-existing English language skills on DA outcomes is further exemplified by the cases of the two children discussed in Section 5.2.7.

Further testing of the DA with a larger sample is essential to validate and expand upon these preliminary findings. Overall, these advocate for further refinement of the DA to enhance its diagnostic independence from pre-existing English language abilities, aiming to reliably evaluate learning potential in UK multilingual children with varied degrees of experience and proficiency in English. Given the intricacies of diagnosing DLD in multilingual language learning settings —where language disorder and natural variation in multiple language development may intersect— this adjustment would be crucial for the DA to fulfil its intended role, which is to prioritise the child's ability to learn with support, as opposed to merely assessing skills at the time of testing, a limitation of conventional static tests.

A notable point from the correlational and regression analyses concerns the role of the DA change scores. Significant relationships were identified between DA measures' pretest scores and both the NRDLS and CL-NWRT, as well as between post-test scores and these measures, in addition to the *DLD Risk Factors Score* and *English Language Proficiency Stage*. Additionally, the *Modifiability Rating Scale* scores were related to the NRDLS and CL-NWRT. However, significant relationships for DA change scores were only found with *Story Grammar Units* in relation to the *English Language Proficiency Stage* and *DLD Risk Factors Score*. At first sight, this pattern seems consistent with findings from previous DA studies reviewed in Section 2.5.2 where, together with the pretest scores, the change scores proved

less helpful in differentiating between typically developing children and those at risk for DLD, with post-test scores and modifiability ratings being more indicative of language impairment (e.g., Peña et al., 2006, 2014; Petersen et al., 2017, 2020; Ukrainetz et al., 2000). However, the one-group design in this research contrasts with the between-group comparisons commonly used in these prior studies, and the absence of significant relationships involving change scores here is likely related to the fact that several higher-achieving children's pretest scores reached ceiling in *Receptive Affective Prosody*, *Episodic Structure Complexity*, and *Emotional Vocabulary*, as detailed in Section 6.2. This ceiling effect likely restricted the observable improvements in these areas over the DA's teaching phase, stressing the importance of fine-tuning this resource's assessment and teaching strategies and related stimuli, as proposed in Section 6.2, and conducting more testing with a larger sample to enhance and more precisely evaluate the DA's diagnostic capabilities. Aligning with this, as suggested by Hunt et al. (2022), the possibility of incorporating a shortly delayed post-test within the same or a second DA session to take into account DLD-impaired skill retention could make it easier to differentiate between the results of children with and without DLD. Specifically, incorporating an additional DA session with a second teaching phase and post-test phase (see Section 6.2) would be consistent with this approach.

6.4 Limitations and Future Research Directions

In conducting this study, several limitations were encountered that warrant consideration. As first stated in Section 4.2, the recruitment process proved to be particularly challenging and time-consuming, mirroring the difficulties faced in earlier pilot studies of the DA. The anticipated relatively low prevalence of multilingual children with DLD compared to the general population was further complicated by the challenges of accessing, engaging, and retaining children and families from culturally and linguistically diverse backgrounds.

Significant extra time and effort were required to ensure effective communication, build rapport, and overcome potential barriers related to language, literacy, and scepticism and lack of awareness around research (Bonevski et al., 2014). In an attempt to improve inclusivity, extensive efforts were made to simplify interactions with parents/caregivers and the research information materials (e.g., Appendices 22, 31; research project website: <https://blogs.ncl.ac.uk/tgarridotamayo2>). Consistent with these efforts, the possibility of translating the materials into the participants' home languages and receiving interpreter assistance was offered to all parents/caregivers of the children involved. This initiative aimed to further facilitate their understanding of the study, including aspects of participation and consent, as well as the developmental and language experience questionnaire. However,

despite repeated offers, there were no requests for interpreters or translations, suggesting that, even though only the parents of seven children rated their English proficiency as ‘excellent’, the remaining parents might have also believed their level of English was sufficient for the study’s requirements (Pert, 2022b).

Looking ahead to future studies of the DA, proactive collaboration with professional interpreters and translators could be beneficial, ensuring parents/caregivers perceive this use of language support services as a measure to guarantee accessibility and full understanding, rather than a critique of their English proficiency (Pert, 2022b). Such an approach could encourage more families to take part in the study, facilitating clearer communication, bridging cultural barriers, and fostering trust (Pert, 2022b; Pert & Bradley, 2018). Focusing on specific sectors of the UK’s multilingual population who speak common languages, such as Polish, Romanian, and Panjabi (ONS, 2022), could simplify collaboration with interpreters and translators, making it a more feasible approach within the research’s budget and time constraints, in contrast with the broader challenge of accommodating the sheer number of home languages spoken across the UK all at once. Likewise, the recently developed *Quantifying Bilingual Experience* (Q-Bex) questionnaire (De Cat et al., 2022), which is available in numerous languages, could also be explored as an option for gathering parental reports, complementing it with study-specific questions as necessary.

Additional hurdles to this study’s recruitment and participation efforts were introduced by the COVID-19 pandemic, which profoundly disrupted research activities across the UK (UKRI, 2021). The pandemic’s impact, including poor health, along with the introduction of social distancing and safety measures, made many potential participants unavailable, even those who were initially eager to take part in the study. In this context, while online, remote methods appeared to be an obvious viable solution to motivate participation during a period fraught with hardship and competing priorities for most families, their limitations, identified in the pilot studies of the DA, ultimately led to a preference for conducting this main study in person, as detailed in ‘Pilot testing 1: Reflections on online vs in-person data collection for the main study’ (Section 3.3.3).

Overall, these difficulties in securing participants, indicative of the complexities of conducting field research with *hard-to-reach* populations during a global health crisis, resulted in the DA being tested with a small cohort of 14 multilingual children with and without reported language difficulties. This aligns with previous studies of DA with DLD diagnostic aims in multilingual children, exposing the emerging nature of this research area

(Hunt et al., 2022; Orellana et al., 2019). Although the limited sample size restricts the statistical robustness and generalisability of the findings, which require cautious interpretation, the sample remains valuable, as it reflects the reality of educational and speech and language therapy practice in the Northeast of England and across the UK, where multilingual children originate from a range of language learning backgrounds and exhibit diverse English proficiency levels (Mennen & Stansfield, 2006; Oxley et al., 2019). While this supports the application of the results to a linguistically and culturally diverse caseload, in line with studies such as those by Hasson et al. (2012a) and Camilleri and Law (2007), future research of the DA should aim for a higher number of participants and broaden socioeconomic representation, as 10 of the 14 children had at least one university-educated parent, suggesting a predominance of middle to high socioeconomic statuses. The use of the growing database of participants created as part of the *Engage with Developmental Language Disorder* (E-DLD) project, which consists of families and individuals affected by DLD, could be useful for this purpose (for further details, see <https://www.engage-dld.com>).

Looking forward, having a larger sample would allow research on the DA to evolve beyond the diagnostic research framework's pre-accuracy phase (Dollaghan & Horner, 2011; Klee, 2008; Sackett & Haynes, 2002)—focused on appraising the DA's initial utility for identifying children with EAL at risk for DLD—and advance into Phase I. As first mentioned in Section 4.1, this next phase would explore group-level differences in the performance of children with and without possible DLD in the DA (*index test*). Given the absence of a single *gold* standard assessment for DLD diagnosis in multilingual children, assignment to groups would rely on their potential developmental status as per a *reference* standard, which could be based on either registration on a speech and language therapist's caseload or, for a more consistent approach across the sample, a variety of evidence-based, complementary measures relevant for detecting DLD in this demographic (De Lam White & Jin, 2011; Letts, 2012; Li'el et al., 2019; Roseberry-McKibbin, 2021), similar to those used in the current study. Crucially, these findings would provide suggestive, rather than definitive, evidence of DLD presence, requiring further comprehensive evaluation in English and the children's home language(s) (Pert & Bradley, 2018).

Finally, refining the DA's scoring systems, teaching strategies, and associated stimuli, as proposed in Sections 6.2 and 6.3, could improve its effectiveness over a static assessment approach, offering a more reliable understanding of children's language learning potential and the specific challenges they face, whether arising from DLD or the typical process of acquiring an additional language. To enhance the reliability and validity of the DA outcomes,

this study adopted specific strategies, including the use of scripts, assessor training, and the support of a second assessor for inter-rater reliability and implementation fidelity purposes, which future DA research should continue to employ to promote consistency in assessor-child interactions and the generalisability of findings, often criticised due to the individualised nature of DA approaches (Hasson & Joffe, 2007; Hunt et al., 2022).

In conclusion, the present study advances the development of the DA created in this thesis for diagnosing DLD risk among multilingual children in the UK. It also contributes to the broader field of DA research in this area, where the need has been highlighted for further investigation that adds to the very few DA resources already available, especially for older, school-aged children with EAL (Oxley et al., 2019) and those from a wider array of linguistic and cultural backgrounds to validate the DA's adaptability across varied settings where more than one language is used for interaction on a regular basis (Hunt et al., 2022; Orellana et al., 2019). Moreover, the study tackles the gap in tools appropriate for recognising and addressing the unique language needs of multilingual children (Newbury et al., 2020; Oxley et al., 2019), ensuring they are provided with equitable educational and healthcare support within the UK. Future studies should also account for the adverse pandemic effects on children's communication and language development (Clegg et al., 2021; Jeffreys, 2021), including those with EAL (Demie et al., 2022). These effects, which could derive from factors such as reduced social interaction, schooling disruptions, and poor mental health, may add another layer of complexity in differentiating DLD-related English language struggles from those due to limited experience with this language.

Chapter 7. Conclusions

This thesis has focused on the challenge of diagnosing Developmental Language Disorder (DLD) in the multilingual child population of the UK, a task made complex by the country's extensive linguistic diversity and the limited availability of suitable assessment tools for this demographic (Oxley et al., 2019). At its core, the research aimed to devise and evaluate a Dynamic Assessment (DA) resource for primary school-aged children with English as an Additional Language (EAL) who have multiple language learning backgrounds and English proficiency levels, just as practitioners and educators typically encounter in the UK (Mennen & Stansfield, 2006; Oxley et al., 2019).

This approach was motivated by a comprehensive review of the literature, which identified important gaps in the assessment of DLD among multilingual children; a group with naturally widely variable language development due to factors such as differing degrees of exposure to each of their languages, including English. The research responded to this by introducing a storytelling-centred DA characterised by its pretest–teach–post-test structure, which moved the focus away from static language proficiency to learning potential in areas sensitive to DLD—narrative macrostructure (story grammar and episodic structure complexity), emotional vocabulary, and receptive affective prosody—aiming to help disentangle DLD from typical multilingual language acquisition. Making such a distinction is key for providing timely and appropriate support to these children, whether through increased exposure to English in school—if their English language issues are attributable to insufficient language input—or through direct, specialist intervention if they have a true language impairment (Pert & Bradley, 2018; Roseberry-McKibbin, 2021).

The primary data collection of this thesis involved the participation of 14 multilingual children aged 4;06–8;11 years to explore the DA's preliminary effectiveness in identifying DLD risk. Specifically, the DA's usefulness in measuring teaching-induced changes in the children's performance was evidenced by their improved story grammar usage, suggesting the resource's early potential for revealing learning capabilities in this area in a manner that conventional, static assessments may not capture. However, such capability did not currently extend to the other targeted areas: episodic structure complexity, emotional vocabulary, and affective prosody processing. A small sample size along with factors including a ceiling effect and limitations of the DA's methods may have influenced these outcomes.

Furthermore, when exploring the DA's diagnostic effectiveness for DLD based on its relationships with comparative measures to address this study's research questions and sub-

questions, the findings indicated a potential usefulness in recognising DLD risk in multilingual children. This was drawn from correlational and predictive relationships noted between DA outcomes and the language-neutral *Crosslinguistic Nonword Repetition Test* (CL-NWRT; Chiat, 2015; Chiat et al., 2020) and *DLD Risk Factors Score*. However, the DA's initial utility is limited by its current dependence on the children's English proficiency, as shown by its relationships with the *New Reynell Developmental Language Scales* (NRDLS; Edwards et al., 2011) in English and *English Language Proficiency Stage*, despite no evident connections with the *English Language Experience Score*.

As explained in Section 6.4, these results pave the way for future work, including the need for larger-scale studies to validate the findings and delve deeper into the DA's effectiveness, as well as for refining the DA. Below is an overview of the main suggestions explored in *Chapter 6* for further development of the DA to better accommodate individual skill levels and enhance its sensitivity to DLD risk compared to a static assessment approach:

- Expanding the scoring range of the DA measures: This adjustment would enable the incorporation of more difficulty levels, especially pertinent to the *Story Generation Task* measures, allowing for a finer, more detailed assessment.
- Incorporating more episodes into the single-episode stimuli stories in the DA's storytelling activity: Following the previous suggestion, introducing more than one episode in the stimuli stories would give children additional opportunities to produce each targeted story grammar unit. This expansion would facilitate a more thorough evaluation of their learning abilities across story grammar (*Scoring Sheet 1*), episodic structure complexity (*Scoring Sheet 2*), and emotional vocabulary (*Scoring Sheet 3*). The possibility of incorporating subplots within the stimuli stories could also be explored.
- Adding an extra DA cycle in a follow-up session: For children who struggled in the initial post-test, a second DA session could offer further opportunities for teaching and testing to better gauge their learning abilities. This would involve reusing the story from the first *Story Generation Task* post-test to ensure continuity in teaching, and adding a new story in the post-test.
- Enhancing the teaching methods: Certain targeted skills, such as children's processing of affective prosody, may benefit from the inclusion of activities within the teaching phase that allow for more in-depth, prolonged practice of these skills, taking into account their inherent complexity.

In summary, this thesis represents a step forward in the ongoing efforts to develop effective DA resources that help to identify DLD within multilingual child populations, particularly in contexts such as the UK, where many minority languages other than English, the societal language, are spoken. Specifically, the research contributes to the sparse array of available DA tools for children with EAL, especially those that seek to identify the language learning skills of school-aged children, as well as children with varied linguistic and cultural upbringings (Hunt et al., 2022; Orellana et al., 2021; Oxley et al., 2019).

DA procedures like the one developed in this study can prove very useful as a source of meaningful, reliable evidence when used alongside other evaluation methods, within a comprehensive assessment framework that accounts for areas of strength and weaknesses in all of a child's languages and across different language domains (De Lambo White & Jin, 2011; Hasson et al., 2012a). DAs allow us to look beyond a multilingual child's current deficits in English, as static forms of testing reveal, by providing crucial insights into why they may be struggling with this language and the nature of the support that may benefit them. This approach can play an important role in improving diagnostic processes to accurately distinguish DLD from language differences, which are largely dependent on the opportunities the child has had so far to learn English as an additional language. Fundamentally, securing a valid diagnosis is the basis for preventing misidentification and ensuring that children receive the appropriate kind of support through school or, in the case of DLD, specialist treatment, helping them fulfil their communication potential and ultimately enhancing their quality of life outcomes.

Overall, these assessment practices are also vital for dispelling the widespread misconception, prevalent within society and in educational and clinical settings, that multilingualism complicates a child's language acquisition —rooted in misguided beliefs that view learning multiple languages as unusual and divergent from the monolingual norm (Pert, 2022a). Indeed, these practices can contribute to reinforcing the understanding that language acquisition is enriched, rather than hindered, by multilingualism, even in cases where DLD may be present (Pert & Bradley, 2018). It is critical to raise awareness that multilingualism neither causes nor contributes to speech, language, or communication impairments, including DLD, and that the real challenge faced by education and health professionals lies in the correct detection of these difficulties when they lack assessment resources and language competence adequate for this purpose, among other barriers (Newbury et al., 2020; Oxley et al., 2019). Integrating DAs with other culturally appropriate and linguistically sensitive assessment tools can enable teachers and practitioners in the UK to base their decisions

regarding the absence or presence of DLD in children with EAL on robust evidence. By doing so, these professionals are equipped to avoid baseless myths and to promote non-discriminatory, equitable care and support for all children regardless of their linguistic and cultural backgrounds, positively framing multilingualism as an asset rather than a challenge.

References

Aguilar-Mediavilla, E., Buil-Legaz, L., & Sánchez-Azanza, V. A. (2020). Speech profiles of Spanish-Catalan children with developmental language disorder. *Clinical Linguistics & Phonetics*, 34(1–2), 110–130. <https://doi.org/10.1080/02699206.2019.1619096>

Aguilar-Mediavilla, E., Buil-Legaz, L., López-Penadés, R., Sánchez-Azanza, V. A., & Adrover-Roig, D. (2019). Academic outcomes in bilingual children with developmental language disorder: A longitudinal study. *Frontiers in Psychology*, 10(531), 1–17. <https://doi.org/10.3389%2Ffpsyg.2019.00531>

Altman, C., Armon-Lotem, S., Fichman, S., & Walters, J. (2016). Macrostructure, microstructure, and mental state terms in the narratives of English–Hebrew bilingual preschool children with and without specific language impairment. *Applied Psycholinguistics*, 37(1), 165–193. <https://doi.org/10.1017/S0142716415000466>

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorder* (5th ed.). American Psychiatric Association.

American Speech-Language-Hearing Association (ASHA). (2020). *Telepractice (Practice Portal)*. <https://www.asha.org/practice-portal/professional-issues/telepractice/>

Andreou, G., Lymeropoulou, V., & Aslanoglou, V. (2022). Developmental Language Disorder (DLD) and Autism Spectrum Disorder (ASD): Similarities in pragmatic language abilities. A systematic review. *International Journal of Developmental Disabilities*, 0(0), 1–15. <https://doi.org/10.1080/20473869.2022.2132669>

Antonijevic-Elliott, S., Lyons, R., O’Malley, M. P., Meir, N., Haman, E., Banasik, N., Carroll, C., McMenamin, R., Rodden, M., & Fitzmaurice, Y. (2019). Language assessment of monolingual and multilingual children using non-word and sentence repetition tasks. *Clinical Linguistics & Phonetics*, 34(4), 293–311. <https://doi.org/10.1080/02699206.2019.1637458>

Bahn, D., Vesker, M., Schwarzer, G., & Kauschke, C. (2021). A multimodal comparison of emotion categorization abilities in children with developmental language disorder. *Journal of Speech, Language, and Hearing Research*, 64(3), 993–1007. https://doi.org/10.1044/2020_JSLHR-20-00413

Bedore, L., Peña, E., Joyner, D., & Macken, C. (2011). Parent and teacher rating of bilingual language proficiency and language development concerns. *International Journal of Bilingual Education and Bilingualism, 14*(5), 489–511.
<https://doi.org/10.1080/13670050.2010.529102>

Bishop, D. V. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 39*(6), 879–891. <https://doi.org/10.1017/S0021963098002832>

Bishop, D. V. (2003). *The Children's Communication Checklist—2*. Psychological Corporation.

Bishop, D. V. (2017). Why is it so hard to reach agreement on terminology? The case of Developmental Language Disorder (DLD). *International Journal of Language & Communication Disorders, 52*(6), 671–680. <https://doi.org/10.1111/1460-6984.12335>

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & the CATALISE-2 Consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry, 58*(10), 1068–1080.
<https://doi.org/10.1111/jcpp.12721>

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & the CATALISE Consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS ONE, 11*(7), Article e0158753. <https://doi.org/10.1371/journal.pone.0168066>

Bitetti, D., Hammer, C. S., & López, L. M. (2020). The narrative macrostructure production of Spanish–English bilingual preschoolers: Within-and cross-language relations. *Applied Psycholinguistics, 41*(1), 79–106.
<https://doi.org/10.1017/S0142716419000419>

Bligh, C. (2014). *The silent experiences of young bilingual learners: A sociocultural study into the silent period*. SensePublishers. <https://doi.org/10.1007/978-94-6209-797-1>

Boerma, T., & Blom, E. (2017). Assessment of bilingual children: What if testing both languages is not possible? *Journal of Communication Disorders*, 66, 65–76. <https://doi.org/10.1016/j.jcomdis.2017.04.001>

Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *Journal of Speech, Language, and Hearing Research*, 58(6), 1747–1760. https://doi.org/10.1044/2015_JSLHR-L-15-0058

Boerma, T., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2016). Narrative abilities of monolingual and bilingual children with and without language impairment: Implications for clinical practice. *International Journal of Language and Communication Disorders*, 51(6), 626–638. <https://doi.org/10.1111/1460-6984.12234>

Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish–English bilinguals. *International Journal of Bilingual Education and Bilingualism*, 13(3), 325–344. <https://doi.org/10.1080/13670050903342019>

Bonevski, B., Randell, M., Paul, C., Chapman, K., Twyman, L., Bryant, J., Brozek, I., & Hughes, C. (2014). Reaching the hard-to-reach: A systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Medical Research Methodology*, 14(42), 1–29. <https://doi.org/10.1186/1471-2288-14-42>

Botting, N., Toseeb, U., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2016). Depression and anxiety change from adolescence to adulthood in individuals with and without language impairment. *PLoS One*, 11(7), Article e0156678. <https://doi.org/10.1371%2Fjournal.pone.0156678>

Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology and Psychiatry*, 41(7), 847–857. <https://doi.org/10.1111/1469-7610.00672>

Brinton, B., Fujiki, M., & Asai, N. (2019). The ability of five children with developmental language disorder to describe mental states in stories. *Communication Disorders Quarterly*, 40(2), 109–116. <https://doi.org/10.1177/1525740118779767>

Brownell, R. (2012). *Expressive One-Word Picture Vocabulary Test—Fourth Edition: Spanish-Bilingual Edition*. Academic Therapy Publications.

Byrne, E. M., Jensen, H., Thomsen, B. S., & Ramchandani, P. G. (2023). Educational interventions involving physical manipulatives for improving children's learning and development: A scoping review. *Review of Education*, 11(2), Article e3400. <https://doi.org/10.1002/rev3.3400>

Camilleri, B., & Botting, N. (2013). Beyond static assessment of children's receptive vocabulary: the dynamic assessment of word learning (DAWL). *International Journal of Language & Communication Disorders*, 48(5), 565–581.
<https://doi.org/10.1111/1460-6984.12033>

Camilleri, B., & Law, J. (2007). Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary. *International Journal of Speech-Language Pathology*, 9(4), 312–322.
<https://doi.org/10.1080/14417040701624474>

Camilleri, B., & Law, J. (2014). Dynamic assessment of word learning skills of pre-school children with primary language impairment. *International Journal of Speech-Language Pathology*, 16(5), 507–516. <https://doi.org/10.3109/17549507.2013.847497>

Champoux-Larsson, M. F., & Dylman, A. S. (2018). A prosodic bias, not an advantage, in bilinguals' interpretation of emotional prosody. *Bilingualism: Language and Cognition*, 1–9. <https://doi.org/10.1017/S1366728918000640>

Chan, A., Chen, S., Hamdani, S., Tse, B., & Cheng, K. (2023). Story telling in bilingual Urdu–Cantonese ethnic minority children: Macrostructure and its relation to microstructural linguistic skills. *Frontiers in Psychology*, 14, Article 924056.
<https://doi.org/10.3389/fpsyg.2023.924056>

Chen, J., Justice, L.M., Rhoad-Drogalis, A., Lin, T.J. & Sawyer, B. (2020). Social networks of children with developmental language disorder in inclusive preschool programs. *Child Development*, 91(2), 471–487. <https://doi.org/10.1111/cdev.13183>

Chiat, S. (2015). Nonword repetition. In S. Armon-Lotem, J. de Jong & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 125–150). Multilingual Matters.

Chiat, S., & Polišenská, K. (2016). A framework for crosslinguistic nonword repetition tests: Effects of bilingualism and socioeconomic status on children's performance. *Journal of Speech, Language, and Hearing Research, 59*(5), 1179–1189.
https://doi.org/10.1044/2016_JSLHR-L-15-0293

Chiat, S., Polišenská, K., Yanushevskaya, I., & Antonijevic, S. (2020). *Crosslinguistic nonword repetition test: language-neutral version* [Measurement instrument]. Unpublished instrument.

Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports, 8*(1), Article 8659. <https://doi.org/10.1038/s41598-018-26889-1>

Clegg, J., O'Flynn, P., & Just, P. (2021). *Speech and language therapy during and beyond COVID-19: Building back better with people who have communication and swallowing needs*. Royal College of Speech and Language Therapists.
www.rcslt.org/wp-content/uploads/2021/03/Building-back-better-March2021.pdf

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42*(5), 1195–1204. <https://doi.org/10.1044/jslhr.4205.1195>

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorders. *International Journal of Language and Communication Disorders, 53*(2), 237–255. <https://doi.org/10.1111%2F1460-6984.12338>

Conti-Ramsden, G., Mok, P. L., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities, 34*(11), 4161–4169. <https://doi.org/10.1016%2Fj.ridd.2013.08.043>

Courtright, J. A., & Courtright, I. C. (1983). The perception of nonverbal vocal cues of emotional meaning by language-disordered and normal children. *Journal of Speech, Language, and Hearing Research, 26*(3), 412–417.
<https://doi.org/10.1044/jshr.2603.412>

Creusere, M., Alt, M., & Plante, E. (2004). Recognition of vocal and facial cues to affect in language-impaired and normally-developing preschoolers. *Journal of Communication Disorders*, 37(1), 5–20. [https://doi.org/10.1016/S0021-9924\(03\)00036-4](https://doi.org/10.1016/S0021-9924(03)00036-4)

De Cat, C., Kašćelan, D., Prévost, P., Serratrice, L., Tuller, L., & Unsworth, S. (2022). *Quantifying Bilingual EXperience (Q-BEx): Questionnaire manual and documentation*. <https://doi.org/10.17605/OSF.IO/GB9M7>

De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language & Communication Disorders*, 46(6), 613–627. <https://doi.org/10.1111/j.1460-6984.2011.00049.x>

Demie, F. (2013). English as an additional language pupils: How long does it take to acquire English fluency? *Language and Education*, 27(1), 59–69. <https://doi.org/10.1080/09500782.2012.682580>

Demie, F. (2018). English language proficiency and attainment of EAL (English as second language) pupils in England. *Journal of Multilingual and Multicultural Development*, 39(7), 641–653. <https://doi.org/10.1080/01434632.2017.1420658>

Demie, F., Hay, A., Bellsham-Revell, A., & Gay, A. (2022). *The impact of school closures on pupils with English as an additional language: Evidence from teaching staff, school leaders, pupils and parents*. Schools' Research and Statistics Service Education and Learning, London Borough of Lambeth. https://beta.lambeth.gov.uk/sites/default/files/2022-02/The_Impact_of_School_Closures_on_Pupils_with_English_as_an_Additional_Language_2022.pdf

Department for Education (DfE). (2018). *DfE School census 2017 to 2018: Guide for schools and LAs* (DFE-00171-2017). UK Government. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/708550/2017_to_2018_School_Census_Guide_V1_8.pdf

Department for Education (DfE). (2023). *Academic year 2022/2023: Schools, pupils and their characteristics*. UK Government. <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics/2022-23>

Department for Education (DfE). (2017). *Collection of data on pupil nationality, country of birth and proficiency in English. Summary Report.*
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/665127/Data_on_pupil_nationality__country_of_birth_and_proficiency.pdf

Dispaldro, M., Leonard, L. B., & Deevy, P. (2013). Real-word and nonword repetition in Italian-speaking children with specific language impairment: A study of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research, 56*(1), 323–336. [https://doi.org/10.1044/1092-4388\(2012/11-0304\)](https://doi.org/10.1044/1092-4388(2012/11-0304))

Dollaghan, C. A., & Horner, E. A. (2011). Bilingual language assessment: A meta-analysis of diagnostic accuracy. *Journal of Speech, Language and Hearing Research, 54*(4), 1077–1088. [https://doi.org/10.1044/1092-4388\(2010/10-0093\)](https://doi.org/10.1044/1092-4388(2010/10-0093))

Duncan, D., Gibbs, D., Noor, N. S., & Whittaker, H. M. (1988). *Sandwell bilingual screening assessment scales for expressive Panjabi and English*. NFER-Nelson.

Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *British Picture Vocabulary Scale* (3rd Ed.). GL Assessment.

Dunn, L. M., Dunn, D. M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale* (2nd Ed.). NFER-Nelson.

Eadie, P., Conway, L., Hallenstein, B., Mensah, F., McKean, C., & Reilly, S. (2018). Quality of life in children with developmental language disorder. *International Journal of Language & Communication Disorders, 53*(4), 799–810.
<https://doi.org/10.1111/1460-6984.12385>

Ebbels, S. (2014). Introducing the SLI debate. *International Journal of Language & Communication Disorders, 49*(4), 377–380. <https://doi.org/10.1111/1460-6984.12119>

Ebbels, S. H., McCartney, E., Slonims, V., Dockrell, J. E., & Norbury, C. F. (2019). Evidence-based pathways to intervention for children with language disorders. *International Journal of Language & Communication Disorders, 54*(1), 3–19.
<https://doi.org/10.1111/1460-6984.12387>

Ebert, K. D. (2020). Language sample analysis with bilingual children: translating research to practice. *Topics in Language Disorders, 40*(2), 182–201.
<https://doi.org/10.1097/TLD.0000000000000209>

Ebert, K. D. (2021). Revisiting the influences of bilingualism and developmental language disorder on children's nonverbal processing speed. *Journal of Speech, Language, and Hearing Research*, 64(9), 3564–3570. https://doi.org/10.1044/2021_JSLHR-21-00156

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

Elks, L., McLachlan, H., & Blank, M. (2012). *TALC: Test of Abstract Language Comprehension* (2nd ed.). Elklan.

Engel de Abreu, P. M., Cruz-Santos, A., & Puglisi, M. L. (2014). Specific language impairment in language-minority children from low-income families. *International Journal of Language & Communication Disorders*, 49(6), 736–747. <https://doi.org/10.1111/1460-6984.12107>

Farabolini, G., Taboh, A.R., Ceravolo, M.G., & Guerra, F. (2023). The association between language exposure and nonword repetition performance in bilingual children: A systematic review and meta-analysis. *Bilingualism: Language and Cognition*, 26(3), 621–635. <https://doi.org/10.1017/S1366728922000906>

Feuerstein, R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability*. University Park Press.

Fiani, R., Henry, G., & Prévost, P. (2022). Macrostructure in narratives produced by Lebanese Arabic-French bilingual children: Developmental trends and links with language dominance, exposure to narratives and lexical skills. *Linguistic Approaches to Bilingualism*, 12(4), 446–478. <https://doi.org/10.1075/lab.20035.fia>

Floccia, C., Sambrook, T., Delle Luche, C., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J., & Plunkett, K. (2018). Vocabulary of 2-year-olds learning English and an additional language: norms and effects of linguistic distance. *Monographs of the Society for Research in Child Development*, 83(1), 1–135. <https://doi.org/10.1111/mono.v83.1>

Flynn, N., & Curdt-Christiansen, X.L. (2018). Intentions versus enactment: Making sense of policy and practice for teaching English as an additional language. *Language and Education*, 32(5), 410–427. <https://doi.org/10.1080/09500782.2018.1475484>

Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage. *Developmental Science*, 3(2), 148–162. <https://doi.org/10.1111/1467-7687.00108>

Fujiki, M., & Brinton, B. (2017). Pragmatics and social communication in child language disorders. In R. G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 441–460). Psychology Press.

Fujiki, M., Spackman, M. P., Brinton, B., & Illig, T. (2008). Ability of children with language impairment to understand emotion conveyed by prosody in a narrative passage. *International Journal of Language & Communication Disorders*, 43(3), 330–345. <https://doi.org/10.1080/13682820701507377>

Gagarina N., Klop D., Kunnari S., Tantale K., Välimaa T., Balčiūnienė I., Bohnacker U., & Walters J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, 56, 1–140. <https://doi.org/10.21248/zaspil.56.2019.414>

Gagarina N., Klop D., Kunnari S., Tantale K., Välimaa T., Bohnacker U., & Walters J. (2019). MAIN: Multilingual Assessment Instrument for Narratives – Revised. *ZAS Papers in Linguistics*, 63, 1–36. <https://doi.org/10.21248/zaspil.56.2019.414>

Gardner-Neblett, N. (2022). What predicts oral narrative competence among African American children? Exploring the role of linguistic and cognitive skills. *Journal of Speech, Language, and Hearing Research*, 65(8), 2931–2947. https://doi.org/10.1044/2022_jslhr-22-00002

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(4), 513–543. <https://doi.org/10.1017/S0142716406060383>

Gathercole, V. C. M., Thomas, E. M., & Hughes, E. (2008). Designing a normed receptive vocabulary test for bilingual populations: A model from Welsh. *International Journal of Bilingual Education and Bilingualism*, 11(6), 678–720. <https://doi.org/10.1080/13670050802149283>

Giguere, D. & Hoff, E. (2022). Bilingual development in the receptive and expressive domains: They differ. *International Journal of Bilingual Education and Bilingualism*, 25(10), 3849–3858. <https://doi.org/10.1080/13670050.2022.2087039>

Govindarajan, K., & Paradis, J. (2019). Narrative abilities of bilingual children with and without developmental language disorder (SLI): Differentiation and the role of age and input factors. *Journal of Communication Disorders*, 77, 1–16.
<https://doi.org/10.1016/j.jcomdis.2018.10.001>

Govindarajan, K., & Paradis, J. (2022). Narrative macrostructure and microstructure profiles of bilingual children with autism spectrum disorder: differentiation from bilingual–1390. <https://doi.org/10.1017/S0142716422000406>

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 50(1), 177–195. [https://doi.org/10.1044/1092-4388\(2007/015\)](https://doi.org/10.1044/1092-4388(2007/015))

Griffiths, S., Goh, S. K. Y., Norbury, C. F., & the SCALES team (2020). Early language competence, but not general cognitive ability, predicts children's recognition of emotion from facial and vocal cues. *PeerJ*, 8, Article e9118.
<https://doi.org/10.7717/peerj.9118>

Grimm, A., and Schulz, P. (2014). Specific language impairment and early second language acquisition: the risk of over-and underdiagnosis. *Child Indicators Research*, 7(4), 821–841. <https://doi.org/10.1007/s12187-013-9230-6>

Grosjean, F. (2013). Bilingualism: A short introduction. In Grosjean, F. & Li, P. (Eds.), *The psycholinguistics of bilingualism* (pp. 5–25). Wiley-Blackwell.
https://www.francoisgrosjean.ch/bilin_bicult/25%20Grosjean.pdf

Gutiérrez-Clellen, V. F., & Peña, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language, Speech and Hearing Services in Schools*, 32(4), 212–224.
[https://doi.org/10.1044/0161-1461\(2001/019\)](https://doi.org/10.1044/0161-1461(2001/019))

Hadjadj, O., Kehoe, M., & Delage, H. (2022). Dynamic assessment identifies morphosyntactic deficits in mono-and bilingual children with developmental language disorder. *Languages*, 7(4), 295. <https://doi.org/10.3390/languages7040295>

Halle, T., Hair, E., Wandner, L., McNamara, M., & Chien, N. (2012). Predictors and outcomes of early versus later English language proficiency among English language

learners. *Early Childhood Research Quarterly*, 27(1), 1–20. <https://doi.org/10.1016/j.ecresq.2011.07.004>

Hasson, N., & Joffe, V. (2007). The case for dynamic assessment in speech and language therapy. *Child Language Teaching and Therapy*, 23(1), 9–25. <https://doi.org/10.1177%2F0265659007072142>

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2012a). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy*, 29(1), 57–75. <https://doi.org/10.1177%2F0265659012459526>

Hasson, N., Dodd, B. & Botting, N. (2012b). Dynamic Assessment of Sentence Structure (DASS): Design and evaluation of a novel procedure for the assessment of syntax in children with language impairments. *International Journal of Language & Communication Disorders*, 47(3), 285–299. <https://doi.org/10.1111/j.1460-6984.2011.00108.x>

Henderson, D. E., Restrepo, M. A., & Aiken, L. S. (2018). Dynamic assessment of narratives among Navajo preschoolers. *Journal of Speech, Language, and Hearing Research*, 61(10), 2547–2560. https://doi.org/10.1044/2018_JSLHR-L-17-0313

Hessel, A. K., & Strand, S. (2021). Proficiency in English is a better predictor of educational achievement than English as an Additional Language (EAL). *Educational Review*, 75(4), 763–786. <https://doi.org/10.1080/00131911.2021.1949266>

Hoff, E., Core, C., & Shanks, K. (2020). The quality of child-directed speech depends on the speaker's language proficiency. *Journal of Child Language*, 47(1), 132-145. <https://doi.org/10.1017/S030500091900028X>

Hunt, E., Nang, C., Meldrum, S., & Armstrong, E. (2022). Can dynamic assessment identify language disorder in multilingual children? Clinical applications from a systematic review. *Language, Speech & Hearing Services in Schools*, 53(2), 598–625. https://doi.org/10.1044/2021_LSHSS-21-00094

Ikeda, S. (2021). Dual development of affective-speech-based emotion perception. *The Journal of Genetic Psychology*, 182(6), 462–470. <https://doi.org/10.1080/00221325.2021.1967270>

Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2019). The evaluation of word-learning

abilities in people with developmental language disorder: A scoping review. *International Journal of Language & Communication Disorders*, 54(5), 742–755. <https://doi.org/10.1111/1460-6984.12490>

Jeffreys, B. (2021, April 27). Lockdowns hurt child speech and language skills – report. *BBC News*. <https://www.bbc.co.uk/news/education-56889035>

Kallis, G., & Yarwood, R. (2022). Intergenerational language practices, linguistic capital and place: The case of Greek-Cypriot migrant families in the UK. *Children's Geographies*, 20(6), 931–943. <https://doi.org/10.1080/14733285.2021.2003302>

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech, and Hearing Services in Schools*, 43(1), 81–96. [https://doi.org/10.1044/0161-1461\(2011/10-0095\)](https://doi.org/10.1044/0161-1461(2011/10-0095))

Khan, K. S., Gugiu, M. R., Justice, L. M., Bowles, R. P., Skibbe, L. E., & Piasta, S. B. (2016). Age-related progressions in story structure in young children's narratives. *Journal of Speech, Language, and Hearing Research*, 59(6), 1395–1408. https://doi.org/10.1044/2016_jslhr-1-15-0275

Khan, K. S., Logan, J., Justice, L. M., Bowles, R. P., & Piasta, S. B. (2021). The contribution of vocabulary, grammar, and phonological awareness across a continuum of narrative ability levels in young children. *Journal of Speech, Language, and Hearing Research*, 64(9), 3489–3503. https://doi.org/10.1044/2021_JSLHR-20-00403

Klee, T. (2008). Considerations for appraising diagnostic studies of communication disorders. *Evidence-Based Communication Assessment and Intervention*, 2(1), 34–45. <https://doi.org/10.1080/17489530801927757>

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*, 43(6), 456–473. <https://doi.org/10.1016/j.jcomdis.2010.02.002>

Kramer, K., Mallett, P., Schneider, P., & Hayward, D. (2009). Dynamic Assessment of Narratives with Grade 3 Children in a First Nations Community. *Canadian Journal of*

Speech-Language Pathology & Audiology, 33(3).
https://cjslpa.ca/files/2009_CJSLPA_Vol_33/CJSLPA_2009_Vol_33_No_03_Fall.pdf

Lam, V. L., Chaudry, F. R., Pinder, M., & Sura, T. (2020). British Sikhs in complementary schooling: The role of heritage language proficiency and ‘culture learning’ in ethnic identity and bicultural adaptation. *Language and Education*, 34(1), 81-96.
<https://doi.org/10.1080/09500782.2019.1634095>

Lazewnik, R., Creaghead, N. A., Smith, A. B., Prendeville, J.-A., Raisor-Becker, L., & Silbert, N. (2019). Identifiers of language impairment for Spanish–English dual language learners. *Language, Speech & Hearing Services in Schools*, 50(1), 126–137.
https://doi.org/10.1044/2018_LSHSS-17-0046

Leonard, L. B. (2014). Specific language impairment across languages. *Child Development Perspectives*, 8(1), 1–5. <https://doi.org/10.1111/cdep.12053>

Letts, C. & Sinka, I. (2011). *Multilingual Toolkit*. GL Assessment.

Letts, C. (2012). Multicultural issues in assessment and management. In M. Kersner, J. A. Wright (Eds.), *Speech and language therapy: The decision making process when working with children* (pp. 135–142). Routledge.

Letts, C. (2013). 3. What are the building blocks for language acquisition? Underlying principles of assessment for language impairment in the bilingual context. In V. Gathercole (Ed.), *Solutions for the Assessment of Bilinguals* (pp. 36–56). Multilingual Matters. <https://doi.org/10.21832/9781783090150-005>

Li’el, N., Williams, C., & Kane, R. (2019). Identifying developmental language disorder in bilingual children from diverse linguistic backgrounds. *International Journal of Speech-Language Pathology*, 21(6), 613–622.
<https://doi.org/10.1080/17549507.2018.1513073>

Lidz, C. S. (1991). *Practitioner’s guide to dynamic assessment*. Guilford Press.

Lidz, C. S., & Peña, E. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children. *Language, Speech, and Hearing Services in Schools*, 27(4), 367–372. <https://doi.org/10.1044/0161-1461.2704.367>

Lindgren, J., & Bohnacker, U. (2022). How do age, language, narrative task, language proficiency and exposure affect narrative macrostructure in German-Swedish bilingual children aged 4 to 6? *Linguistic Approaches to Bilingualism*, 12(4), 479–508. <https://doi.org/10.1075/lab.20020.lin>

Lloyd-Esenkaya, V., Russell, A. J., & Clair, M. C. S. (2020). What are the peer interaction strengths and difficulties in children with developmental language disorder? A systematic review. *International Journal of Environmental Research and Public Health*, 17(9), 3140. <https://doi.org/10.3390%2Fijerph17093140>

Löytömäki, J., Laakso, M.L. & Huttunen, K. (2023). Social-emotional and behavioural difficulties in children with neurodevelopmental disorders: Emotion perception in daily life and in a formal assessment context. *Journal of Autism and Developmental Disorders*, 53, 4744–4758. <https://doi.org/10.1007/s10803-022-05768-9>

Löytömäki, J., Ohtonen, P., Laakso, M. L., & Huttunen, K. (2020). The role of linguistic and cognitive factors in emotion recognition difficulties in children with ASD, ADHD or DLD. *International Journal of Language & Communication Disorders*, 55(2), 231–242. <https://doi.org/10.1111/1460-6984.12514>

Ma, W., Zhou, P., & Thompson, W. F. (2022). Children's decoding of emotional prosody in four languages. *Emotion*, 22(1), 198–212. <https://doi.org/10.1037/emo0001054>

Maragkaki, I., & Hessel, M. (2016). A pilot study of dynamic assessment of vocabulary in German for bilingual preschoolers in Switzerland. *Journal of Studies in Education*, 7(1), 32–49. <https://doi.org/10.5296/jse.v7i1.10392>

Marinis, T., & Armon-Lotem, S. (2015). Sentence repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 95–125). Multilingual Matters.

Marinova-Todd, S. H., Colozzo, P., Mirenda, P., Stahl, H., Kay-Raining Bird, E., Parkington, K., Genesee, F. (2016). Professional practices and opinions about services available to bilingual children with developmental disabilities: An international study. *Journal of Communication Disorders*, 63, 47–62. <https://doi.org/10.1016/j.jcomdis.2016.05.004>

Martin, N., & Brownell, R. (2012a). *Expressive One-Word Picture Vocabulary Test: Spanish-Bilingual Edition (4th Ed.)*. Pro-Ed.

Martin, N., & Brownell, R. (2012b). *Receptive One-Word Picture Vocabulary Test: Spanish-Bilingual Edition (4th Ed.)*. Pro-Ed.

Mayer, M. (1967). *A boy, a dog and a frog*. Dial Books for Young Readers.

Mayer, M. (1969). *Frog, where are you?* Dial Books for Young Readers.

Melby-Lervåg, M., Lervåg, A., Lyster, S. A. H., Klem, M., Hagtvæt, B., & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science*, 23(10), 1092–1098.
<https://doi.org/10.1177/0956797612443833>

Mennen, I., & Stansfield, J. (2006). Speech and language therapy service delivery for bilingual children: A survey of three cities in Great Britain. *International Journal of Language & Communication Disorders*, 41(6), 635–652.
<https://doi.org/10.1080/13682820600623911>

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development*, 72(3), 834–843. <https://doi.org/10.1111/1467-8624.00318>

Mountford, H.S., Braden, R., Newbury, D.F. & Morgan, A.T. (2022). The genetic and molecular basis of developmental language disorder: A review. *Children (Basel)*, 9(5), 586. <https://doi.org/10.3390/children9050586>

Moxam, C. (2021). *Telehealth: A toolkit for practice (Version 3)* [Unpublished; Newcastle University].

Müller, L. M., Howard, K., Wilson, E., Gibson, J., & Katsos, N. (2020). Bilingualism in the family and child well-being: A scoping review. *International Journal of Bilingualism*, 24(5-6), 1049–1070. <https://doi.org/10.1177/1367006920920939>

Muñoz-Sandoval, A. F., Cummins, J., Alvarado, C. G., & Ruef, M. L. (2005). *Bilingual verbal ability tests*. Riverside Publishing.

National Association for Language Development in the Curriculum (NALDIC). (2018). *NALDIC's position statement: Withdrawal of English as an Additional Language (EAL) proficiency data from the School Census returns*. <https://naldic.org.uk/about-naldic/activism/position-statements/dfe-proficiency-data>

Newbury, J., Bartoszewicz Poole, A., & Theys, C. (2020). Current practices of New Zealand speech-language pathologists working with multilingual children. *International Journal of Speech-Language Pathology*, 22(5), 1–12.
<https://doi.org/10.1080/17549507.2020.1712476>

Nilsson, K. K., & Jensen de López, K. (2016). Theory of mind in children with specific language impairment: A systematic review and meta-analysis. *Child Development*, 87(1), 143–153. <https://doi.org/10.1111/cdev.12462>

Norbury, C. F., Gemmell, T., & Paul, R. (2014). Pragmatics abilities in narrative production: A cross-disorder comparison. *Journal of Child Language*, 41(3), 485–510.
<https://doi.org/10.1017/S030500091300007X>

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. *Journal of Child Psychology and Psychiatry*, 57(11), 1247–1257. <https://doi.org/10.1111/jcpp.12573>

Öberg, L., & Bohnacker, U. (2022). Non-word repetition and vocabulary in Arabic-Swedish-speaking 4–7-year-olds with and without developmental language disorder. *Languages*, 7(3), 204. <https://doi.org/10.3390/languages7030204>

Öberg, L., & Bohnacker, U. (2024). Beyond language scores: How language exposure informs assessment of nonword repetition, vocabulary and narrative macrostructure in bilingual Turkish/Swedish children with and without developmental language disorder. *Children*, 11(6), 704. <https://doi.org/10.3390/children11060704>

Office for National Statistics (ONS). (2013). *2011 Census: Detailed analysis – English language proficiency in England and Wales, main language and general health characteristics*. UK Government.
<https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/articles/detailedanalysisenglishlanguageproficiencyinenglandandwales/2013-08-30>

Office for National Statistics (ONS). (2016). *Explore 50 years of international migration to and from the UK*. UK Government.
<https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/explore50yearsofinternationalmigrationtoandfromtheuk/2016-12-01>

Office for National Statistics (ONS). (2022). *2021 Census: Main language, English language proficiency, and household language in England and Wales*. UK Government.
<https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/bulletins/languageenglandandwales/census2021>

Office for National Statistics (ONS). (2023). *Long-term international migration, provisional: year ending June 2023*. UK Government.

<https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/longterminternationalmigrationprovisional/yearendingjune2023>

Orellana, C. I., Wada, R., & Gillam, R. B. (2019). The use of dynamic assessment for the diagnosis of language disorders in bilingual children: A meta-analysis. *American Journal of Speech-Language Pathology*, 28(3), 1298–1317.
https://doi.org/10.1044/2019_ajslp-18-0202

Orrego, P. M., McGregor, K. K., & Reyes, S. M. (2023). A first-person account of developmental language disorder. *American Journal of Speech-Language Pathology*, 32(4), 1383–1396. https://doi.org/10.1044/2023_AJSLP-22-00247

Ortiz, J. A. (2021). Using nonword repetition to identify language impairment in bilingual children: A meta-analysis of diagnostic accuracy. *American Journal of Speech-Language Pathology*, 30(5), 2275–2295.
https://doi.org/10.1044/2021_AJSLP-20-00237

Oxley, E., Cattani, A., Chondrogianni, V., Floccia, C., White, L., & De Cat, C. (2019). *Assessing bilingual children: A survey of professional practices in the UK*.
<https://doi.org/10.31219/osf.io/2vjwe>

Paradis, J. (2005). Grammatical morphology in children learning English as a second language. *Language, Speech, and Hearing Services in Schools*, 36(3), 172–187. [https://doi.org/10.1044/0161-1461\(2005/019\)](https://doi.org/10.1044/0161-1461(2005/019))

Paradis, J. (2011). Individual differences in child English second language acquisition: comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1(3), 213–237. <https://doi.org/10.1075/lab.1.3.01par>

Paradis, J., & Jia, R. (2017). Bilingual children's long-term outcomes in English as a second language: Language environment factors shape individual differences in catching up with monolinguals. *Developmental Science*, 20(1), Article e12433.
<https://doi.org/10.1111/desc.12433>

Paradis, J., & Kirova, A. (2014). English second-language learners in preschool: Profile effects in their English abilities and the role of home language environment. *International Journal of Behavioral Development*, 38(4), 342–49.
<https://doi.org/10.1177%2F0165025414530630>

Paradis, J., Emmerzael, K., & Sorenson Duncan, T. (2010). Assessment of English language learners: using parent report on first language development. *Journal of Communication Disorders*, 43(6), 474–497.
<https://doi.org/10.1016/j.jcomdis.2010.01.002>

Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics*, 29(4), 689–722.
<https://doi.org/10.1017/S0142716408080296>

Paradis, J., Tulpar, Y., & Arppe, A. (2016). Chinese L1 children's English L2 verb morphology over time: Individual variation in long-term outcomes. *Journal of Child Language*, 43(3), 553–580. <https://doi.org/10.1017/S0305000915000562>

Park, J., Miller, C. A., Sanjeevan, T., van Hell, J. G., Weiss, D. J., & Mainela-Arnold, E. (2019). Bilingualism and attention in typically developing children and children with developmental language disorder. *Journal of Speech, Language and Hearing Research*, 62(11), 4105–4118. https://doi.org/10.1044/2019_JSLHR-L-18-0341

Parry, F. (2020, May 11). Teachers' perceptions and recognition of EAL learners with additional needs. *EAL Journal Blog*. <https://ealjournal.org/2020/05/11/teachers-perceptions-and-recognition-of-eal-learners-with-additional-needs>

Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., & Kotz, S. A. (2009a). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics*, 37(4), 417–435. <https://doi.org/10.1016/j.wocn.2009.07.005>

Pell, M., Monetta, L., Paulmann, S., & Kotz, S. (2009b). Recognising emotions in a foreign language. *Journal of Nonverbal Behavior*, 33(2), 107–120.

<https://doi.org/10.1007/s10919-008-0065-7>

Peña E. D., Gutiérrez-Clellen, V. F., Iglesias A., Goldstein B. A., & Bedore L. M.

(2014a). *BESA: Bilingual English Spanish Assessment manual*. AR-Clinical Publications.

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014b). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language

learners. *Journal of Speech, Language, and Hearing Research*, 57(6), 2208–2220.

https://doi.org/10.1044/2014_JSLHR-L-13-0151

Peña, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T.

(2006). Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research*, 49(5), 1037–1057. [https://doi.org/10.1044/1092-4388\(2006/074\)](https://doi.org/10.1044/1092-4388(2006/074))

Peña, E. D., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic

assessment of children's word learning ability. *American Journal of Speech-*

Language Pathology, 10(2), 138–154. [https://doi.org/10.1044/1058-0360\(2001/014\)](https://doi.org/10.1044/1058-0360(2001/014))

Peña, E., & Quinn, R. (1997). Task familiarity: Effects on the test performance of Puerto

Rican and African American children. *Language, Speech, and Hearing Services in*

Schools, 28(4), 323–332. <https://doi.org/10.1044/0161-1461.2804.323>

Peña, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language

assessment: A nonbiased procedure. *The Journal of Special Education*, 26(3), 269–

280. <https://doi.org/10.1177%2F002246699202600304>

Peña, E., Resendiz, M., & Gillam, R. B. (2007). The role of clinical judgement of

modifiability in the diagnosis of language impairment. *International Journal of*

Speech-Language Pathology, 9(4), 332–345. <https://doi.org/10.1080/14417040701413738>

Peña, E.D. (2001). Assessment of semantic knowledge: Use of feedback and clinical interviewing. *Seminars in Speech and Language*, 22(1), 51–63.
<https://doi.org/10.1055/s-2001-13865>

Pert, S. & Bradley, B. (2018). *Clinical guidelines for speech and language therapists: Bilingualism: Working with bilingual clients/patients with speech, language and communication needs*. Royal College of Speech and Language Therapists. https://research.manchester.ac.uk/files/69559895/mainbilingualism_1_.pdf

Pert, S. & Letts, C. (2006) Codeswitching in Mirpuri-speaking Pakistani heritage preschool children: Bilingual language acquisition. *International Journal of Bilingualism*, 10(3), 349–374. <https://doi.org/10.1177/13670069060100030501>

Pert, S. & Stow, C. (2019). *Bilingual Assessment of Simple Sentences (BASS): An expressive language assessment for children of Pakistani heritage who speak Mirpuri, Punjabi or Urdu*. Royal College of Speech and Language Therapists.
<https://research.manchester.ac.uk/en/publications/bilingual-assesment-of-simple-sentences-bass-an-expressive-langua>

Pert, S. (2022a). Introduction. In S. Pert (Ed.), *Working with children experiencing speech and language disorders in a bilingual context: A home language approach* (pp. 1–16). Routledge Taylor and Francis Group.

Pert, S. (2022b). Working alongside interpreters. In S. Pert (Ed.), *Working with children experiencing speech and language disorders in a bilingual context: A home language approach* (pp. 67–89). Routledge Taylor and Francis Group.

Petersen, D. B., Chanthongthip, H., Ukrainetz, T. A., Spencer, T. D., & Steeve, R. W. (2017). Dynamic assessment of narratives: Efficient, accurate identification of language impairment in bilingual students. *Journal of Speech, Language, and Hearing Research*, 60(4), 983–998. https://doi.org/10.1044/2016_JSLHR-L-15-0426

Petersen, D. B., Tonn, P., Spencer, T. D., & Foster, M. E. (2020). The classification accuracy of a dynamic assessment of inferential word learning for bilingual English/Spanish-speaking school-age children. *Language, Speech, and Hearing Services in Schools*, 51(1), 144–164. https://doi.org/10.1044/2019_LSHSS-18-0129

Pham, G., & Ebert, K. D. (2020). Diagnostic accuracy of sentence repetition and nonword repetition for developmental language disorder in Vietnamese. *Journal of Speech, Language, and Hearing Research, 63*(5), 1521–1536.
https://doi.org/10.1044/2020_JSLHR-19-00366

Pieretti, R. A., & Roseberry-McKibbin, C. (2016). Assessment and intervention for English language learners with primary language impairment: Research-based best practices. *Communication Disorders Quarterly, 37*(2), 117–128.
<https://doi.org/10.1177/1525740114566652>

Pua, E. P. K., Lee, M. L. C., & Rickard Liow, S. J. (2017). Screening bilingual preschoolers for language difficulties: Utility of teacher and parent reports. *Journal of Speech, Language, and Hearing Research, 60*(4), 950–968.
https://doi.org/10.1044/2016_jslhr-l-16-0122

Raman, N., Nagarajan, R., Venkatesh, L., Monica, D. S., Ramkumar, V., & Krumm, M. (2019). School-based language screening among primary school children using telepractice: A feasibility study from India. *International Journal of Speech Language Pathology, 21*(4), 425–434. <https://doi.org/10.1080/17549507.2018.1493142>

Reilly, J. (2001). From affect to language: development of evaluation in narratives in spoken English and American Sign Language. In L. Verhoeven, S. Strömqvist (Eds.), *Narrative Development in a Multilingual Context* (pp. 399–417). John Benjamins.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., Conway, L., Eadie, P., & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from early language in Victoria study. *Pediatrics, 126*(6), e1530-e1537.
<https://doi.org/10.1542/peds.2010-0254>

Renfrew, C. (2003). *Renfrew Action Picture Test*. Speechmark Publishing.

Restrepo, A. (1998). Identifiers of predominately Spanish-speaking children with language impairment. *Journal of Speech. Language and Hearing Research, 41*(6), 1398–1411.
<https://doi.org/10.1044/jslhr.4106.1398>

Rezzonico, S., Chen, X., Cleave, P. L., Greenberg, J., Hipfner-Boucher, K., Johnson, C. J., Milburn, T., Pelletier, J., Weitzman, E., & Girolametto, L. (2015). Oral narratives in monolingual and bilingual preschoolers with SLI. *International Journal of Language*

& Communication Disorders, 50(6), 830–841. <https://doi.org/10.1111/1460-6984.12179>

Roberts, T. A. (2014). Not so silent after all: Examination and analysis of the silent stage in childhood second language acquisition. *Early Childhood Research Quarterly*, 29(1), 22–40. <https://doi.org/10.1016/j.ecresq.2013.09.001>

Roseberry-McKibbin, C. (2021). Utilising comprehensive preassessment procedures for differentiating language difference from language impairment in English learners. *Communication Disorders Quarterly*, 42(2), 93–99. <https://doi.org/10.1177/1525740119890314>

Royal College of Speech and Language Therapists (RCSLT). (2020a). *RCSLT briefing paper on Language Disorder with a specific focus on Developmental Language Disorder*. <https://www.rcslt.org/wp-content/uploads/media/docs/Covid/language-disorder-briefing-paper-with-edit.pdf>

Royal College of Speech and Language Therapists (RCSLT). (2020b). *Telehealth guidance*. <https://www.rcslt.org/members/delivering-quality-services/telehealth/telehealth-guidance/>

Rudolph, J. M. (2017). Case history risk factors for specific language impairment: A systematic review and meta-analysis. *American Journal of Speech-Language Pathology*, 26(3), 991–1010. https://doi.org/10.1044/2016_AJSLP-15-0181

Sackett, D. L., & Haynes, R. B. (2002). The architecture of diagnostic research. *British Medical Journal*, 324(7336), 539–541. <https://www.jstor.org/stable/25227608>

Schneider, P., Dubé, R. V., & Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*. <https://www.rehabresearch.ualberta.ca/enni>

Schneider, P., Hayward, D., & Dubé, R. V. (2006). Storytelling from pictures using the Edmonton narrative norms instrument. *Journal of Speech-Language Pathology and Audiology*, 30(4), 224–238. <http://www.rehabmed.ualberta.ca/spa/enni/Schneider%20Hayward%20Dube%202006.pdf>

Schwob, S., Eddé, L., Jacquin, L., Leboulanger, M., Picard, M., Oliveira, P. R., & Skoruppa, K. (2021). Using nonword repetition to identify developmental language disorder in

monolingual and bilingual children: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, 64(9), 3578–3593.

https://doi.org/10.1044/2021_JSLHR-20-00552

Scott, R. (2021). *Language and learning loss: The evidence on children who use English as an Additional Language*. The Bell Foundation. <https://www.bell-foundation.org.uk/app/uploads/2021/06/Language-learning-loss.pdf>

Semel, E. M., Wiig, E. H., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals: Spanish* (4th ed.). The Psychological Corporation.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals* (4th ed.). The Psychological Corporation.

Semel, E., Wiig, E. H., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals: Preschool* (2nd ed.). The Psychological Corporation.

Siller, M., Swanson, M. R., Serlin, G., & Teachworth, A. G. (2014). Internal state language in the storybook narratives of children with and without autism spectrum disorder: Investigating relations to theory of mind abilities. *Research in Autism Spectrum Disorders*, 8(5), 589–596. <https://doi.org/10.1016/j.rasd.2014.02.002>

Siraj-Blatchford, I., & Clarke, P. (2000). *Supporting identity, diversity, and language in the early years*. Open University Press.

Sorenson Duncan, T. & Paradis, J. (2020) How does maternal education influence the linguistic environment supporting bilingual language development in child second language learners of English?. *International Journal of Bilingualism* 24(1), 46–61.
<https://doi.org/10.1177/1367006918768366>

Spencer, T. D., & Petersen, D. B. (2020). Narrative intervention: Principles to practice. *Language, Speech, and Hearing Services in Schools*, 51(4), 1081–1096.
https://doi.org/10.1044/2020_LSHSS-20-00015

Squires, K. E., Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., Bohman, T. M., & Gillam, R. B. (2014). Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language & Communication Disorders*, 49(1), 60–74. <https://doi.org/10.1111/1460-6984.12044>

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.). *New Directions in Discourse Processing* (pp. 53–102). Ablex.

Stow, C., & Pert, S. (1998). The development of a bilingual phonology assessment. *International Journal of Language & Communication Disorders*, 33(S1), 338–342. <https://doi.org/10.3109/13682829809179447>

Strand, S. & Hessel, A. (2018). *English as an additional language, proficiency in English and pupils' educational achievement: An analysis of Local Authority data*. University of Oxford, Unbound Philanthropy and The Bell Foundation. <https://www.bell-foundation.org.uk/app/uploads/2018/10/EAL-PIE-and-Educational-Achievement-Report-2018-FV.pdf>

Strand, S. & Lindorff, A. (2020). *English as an Additional Language: Proficiency in English, educational achievement and rate of progression in English language learning*. University of Oxford, Unbound Philanthropy and The Bell Foundation. <https://www.bell-foundation.org.uk/app/uploads/2020/02/University-of-Oxford-Report-Feb-2020-web.pdf>

Strand, S., Malmberg, L., & Hall, J. (2015). *English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database*. University of Oxford, Education Endowment Foundation, Unbound Philanthropy and The Bell Foundation. https://ore.exeter.ac.uk/repository/bitstream/handle/10871/23323/EAL_and_education al_achievement2.pdf

Streubel, B., Gunzenhauser, C., Grosse, G., & Saalbach, H. (2020). Emotion-specific vocabulary and its contribution to emotion understanding in 4- to 9-year-old children. *Journal of Experimental Child Psychology*, 193, Article 104790. <https://doi.org/10.1016/j.jecp.2019.104790>

Taha, J., Stojanovik, V., & Pagnamenta, E. (2021). Nonword repetition performance of Arabic-speaking children with and without developmental language disorder: A study on diagnostic accuracy. *Journal of Speech, Language, and Hearing Research*, 64(7), 2750–2765. https://doi.org/10.1044/2021_jslhr-20-00556

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. (2015). Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. *International Journal of Language & Communication Disorders*, 50(4), 452–466. <https://doi.org/10.1111/1460-6984.12146>

Trauner, D. A., Ballantyne, A., Chase, C., & Tallal, P. (1993). Comprehension and expression of affect in language-impaired children. *Journal of Psycholinguistic Research*, 22(4), 445–452. <https://doi.org/10.1007/BF01074346>

Tsimpli, I. M., Peristeri, E., & Andreou, M. (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguistics*, 37(1), 195–216. <https://doi.org/10.1017/S0142716415000478>

Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong & N. Meir (Eds.), *Assessing multilingual children: Disentangling Bilingualism from Language Impairment*. (pp. 301–330). Multilingual Matters.

UK Research and Innovation (UKRI). (2021). *Survey findings of the impact of COVID-19 on researchers*. <https://www.ukri.org/news/survey-findings-of-the-impact-of-covid-19-on-researchers>

Ukoumunne, O. C., Wake, M., Carlin, J., Bavin, E. L., Lum, J., Skeat, J., Williams, J., Conway, L., Cini, E., & Reilly, S. (2012). Profiles of language development in pre-school children: A longitudinal latent class analysis of data from the Early Language in Victoria Study. *Child: Care, Health and Development*, 38(3), 341–349. <https://doi.org/10.1111/j.1365-2214.2011.01234.x>

Ukrainetz, T. A., Harpell, S., Walsh, C., & Coyle, C. (2000). A preliminary investigation of dynamic assessment with Native American kindergartners. *Language, Speech, and Hearing Services in Schools*, 31(2), 142–154. <https://doi.org/10.1044/0161-1461.3102.142>

Van den Bedem, N. P., Dockrell, J. E., van Alphen, P. M., Kalicharan, S. V., & Rieffe, C. (2016). Victimization, bullying, and emotional competence: Longitudinal associations in (pre)adolescents with and without Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, 61(8), 2028–2044. https://doi.org/10.1044/2018_JSLHR-L-17-0429

Van den Bedem, N.P., Dockrell, J.E., van Alphen, P.M., de Rooij, M., Samson, A.C., Harjunen, E.L. and Rieffe, C. (2018). Depressive symptoms and emotion regulation strategies in children with and without developmental language disorder: a longitudinal study. *International Journal of Language & Communication Disorders*, 53(6), 1110–1123. <https://doi.org/10.1111%2F1460-6984.12423>

Van Rijn, P., & Larrouy-Maestri, P. (2023). Modelling individual and cross-cultural variation in the mapping of emotions to speech prosody. *Nature Human Behaviour*, 7, 386–396. <https://doi.org/10.1038/s41562-022-01505-5>

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010). Internet-based telehealth assessment of language using the CELF-4. *Language, Speech, and Hearing Services in Schools*, 41(4), 445–458. [https://doi.org/10.1044/0161-1461\(2009/08-0131\)](https://doi.org/10.1044/0161-1461(2009/08-0131))

Westby, C. E. (2005). Assessing and remediating text comprehension problems. In H. W. Catts & A. G. Kamhi (Eds.), *Language and reading disabilities* (pp. 157–232). Allyn & Bacon.

Whiteside, K. E., Gooch, D., & Norbury, C. F. (2016). English language proficiency and early school attainment among children learning English as an additional language. *Child Development*, 88(3), 812–827. <https://doi.org/10.1111/cdev.12615>

Williams, C.J., & McLeod, S. (2012). Speech-language pathologists' assessment and intervention practices with multilingual children. *International Journal of Speech-Language Pathology*, 14(3), 292–305. <https://doi.org/10.3109/17549507.2011.636071>

Winter, K. (1999). Speech and language therapy provision for bilingual children: Aspects of the current service. *International Journal of Language & Communication Disorders*, 34(1), 85–98. <https://doi.org/10.1080/136828299247658>

Zoom. (2020). *Zoom Video Communications GDPR Compliance*. <https://zoom.us/gdpr>

Appendices

Appendix 1. Star Points Chart Used During the Dynamic Assessment

STEPS TO BECOMING A STORYTELLING S U P E R STAR	1. Play the Feelings Game 1	★
	2. Complete Story A	★
	3. Complete Story B	★
	4. Play the Feelings Game 2	★

BILINGUAL CHILDREN NEEDED!

Help us try out a resource that will assist in detecting Developmental Language Disorder in bilingual children

We are looking for 5-12 year old children living in the UK who:

- Speak languages other than or in addition to English at home.
- Have **no** speech, language, or communication difficulties.

Participation will involve playing fun language games in English over Zoom!



For more details, please contact Teresa:
t.garrido-tamayo2@newcastle.ac.uk
07565 527743



Appendix 3. Parental Information Sheet Used in the Child Pilot Study 1

Using a story-based dynamic assessment to identify developmental language disorder in children learning English as an additional language

Piloting phase stage 1 (May 2021)



Speech & Language Sciences

Parent/Guardian Information Sheet

Dear Parent/Guardian,

I am a PhD student at the School of Education, Communication and Language Sciences at Newcastle University carrying out research on the assessment of language difficulties in children learning English as an additional language. This project is under the supervision of Drs Carolyn Letts and Laurence White. I am contacting to ask whether your child can be involved in this research, and whether you can inform us about your child's general and language development and the language(s) they use in the home and outside of school.

What is the study about?

Many children have difficulties understanding and/or using language regardless of whether English is their first language or not. These children are identified as having a Developmental Language Disorder when these difficulties are significant, persistent and cannot be explained by an obvious cause. For children learning English as an additional language, identifying a Developmental Language Disorder is difficult because there is a lack of tools we can use with children growing up in environments where more than language is spoken.

I want to find out how useful a tool I have created is in detecting Developmental Language Disorder in children learning English as an additional language, so that their needs can be addressed as quickly as possible. For the first phase of the study, I am looking to recruit children of various ages who have English as an additional language and do not present language difficulties.

Why was your child chosen?

For this initial phase of the study, I am recruiting children through personal contacts, such as yourself. These are children of different ages who have English as an additional language and do not present language difficulties.

What am I asking parents and children to do?

If you agree, you will be asked to provide information on your child's general and language development, as well as the language(s) you use at home and outside of the school environment. This will be in the form of an online questionnaire you can complete in your own time, or I may meet with you in a Zoom video call where we can go through the questionnaire together – whichever works best for you. I can also supply an interpreter to help if necessary.

Following this, I will meet with your child once through a video call which will last around 30-45 minutes. For the session, I will require you to be present with your child. We will arrange for this video call to take place at the most convenient time for you and your child. During the video call, I will conduct some activities that involve listening to and telling stories, as well as playing a game where your child will learn about how we can express our feelings with our voices.

The session will be carried out in a quiet, private room in my home, and I will ask that you and your child are also in a confidential space that is free from distractions and interruptions. I will not do any of these activities with your child unless I have your full consent. All activities are designed to be enjoyable and fun for children. However, I will stop at once if your child indicates in any way that they do not wish to participate or continue with any part of the session. I have full Disclosure and Barring clearance, which means that I am authorised and qualified to work safely with children.

Ethics

The study has been approved by the Faculty of Humanities & Social Sciences Research Ethics Committee at Newcastle University.

Risks and benefits

There are no known risks to participation in this study. Your child would benefit in that their language development would be studied closely and, therefore, any extra support they may need will be recognised.

Does your child have to do this?

No. If you do not want your child to participate, this is, of course, absolutely understandable. If you change your mind, you can withdraw your child from the research at any time. If your child does not want to join in with the activities or becomes upset, we will stop at once. If you would like a brief report on what happened in the activities with your child, this can be provided.

Who will know about this?

Solely myself and my two supervisors. We will not pass any personal information about you or your child to anyone else, nor store this information in a location where others may see it. If your child chats to me about topics unrelated to the project, this will be fully confidential, except in the unlikely case that they disclose something that suggests their safety is at risk.

The activities I will undertake with your child will be recorded on paper and electronic documents, as well as video recordings. I will give a unique participant code to all the information I collect from you and your child, and neither your name nor your child's name will appear on any documents or files. Papers will be stored in a locked filing cabinet in the university, and video recordings and any other electronic files will be stored on secure, password-protected University servers. All of this will be destroyed within 5 years of the project. When I write about this study in my thesis, or when I talk about it to others, for example at a conference or a seminar, no information will be included that will make it possible to identify any individual children.

Consent

If you are willing for your child to participate in this study, please sign the attached consent form.

Thank you for taking the time to read this. If you have any questions, please do not hesitate to contact me.

Best wishes,

Teresa Garrido-Tamayo
PhD Speech & Language Sciences
Email: t.garrido-tamayo2@newcastle.ac.uk

Appendix 4. Parental Consent Form Used in the Child Pilot Study 1

Using a story-based dynamic assessment to identify developmental language disorder in children learning English as an additional language

Piloting phase stage 1 (May 2021)



Speech & Language Sciences

Parent/Guardian Consent Form

		Please initial box
1.	I have read the information sheet dated May 2021 (piloting phase stage 1) for the study above. I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.	
2.	I understand that my child will carry out activities through Zoom with a researcher, some of which will be video recorded.	
3.	I understand that I will need to be present with my child during the Zoom video session.	
4.	I am responsible for arranging a confidential space within my home that is free from distractions and disruptions for the Zoom video session.	
5.	I agree to provide information about the language(s) the child speaks at home and when outside of school, as well as their general and language development.	
6.	I understand that it will not be possible for anyone other than the researcher to identify my child.	
7.	I understand that my participation and my child's participation are voluntary. I can withdraw at any time without giving reason and without detriment to any care or services my child or I may be receiving or may receive in the future. Any activity will stop immediately if my child does not want to join in or becomes upset.	
I agree for my child (name) _____ to participate in this study.		
Parent/guardian: Signature (please provide your digital signature, or a scanned or photographed copy): _____		
Date: _____		
Print full name: _____		
Relationship to child: _____		

Teresa Garrido-Tamayo
PhD Speech & Language Sciences
Email: t.garrido-tamayo2@newcastle.ac.uk

Appendix 5. Child-Friendly Information Video Used in the Child Pilot Study 1

The video file can be found in the following link:

https://drive.google.com/file/d/1Zb989zcZPpAfsPUnXgwZuQ67UAQwcHM2/view?usp=drive_link

Appendix 6. Facilitator Checklist Used in the Child Pilot Study 1

Using a story-based dynamic assessment to identify developmental language disorder in children learning English as an additional language



Piloting phase stage 1 (May 2021)

Facilitator checklist for parents/guardians

Thank you for agreeing to you and your child's participation in this research study. Please carefully read the information on the checklist below to understand how you should prepare yourself and your child ahead of the session, as well as what you and your child will be required to do during the session. If you have any questions, please do not hesitate to get in touch.

Things I need to do <u>before</u> the session	Done
Location	
Choose a quiet, private room in my home	<input type="checkbox"/>
Remove toys and any other distractions from the room, such as pets	<input type="checkbox"/>
Put my mobile phone on silent mode	<input type="checkbox"/>
Make sure the room is well lit	<input type="checkbox"/>
If possible, close the door to avoid interruptions	<input type="checkbox"/>
Seating	
Find a comfortable place for my child and I to sit	<input type="checkbox"/>
Make sure that my child and I can be seen clearly on the computer camera	<input type="checkbox"/>
Computer	
Ensure my computer is fully charged or plugged into a charging outlet	<input type="checkbox"/>
Check that the Wi-Fi/internet connection is good	<input type="checkbox"/>
Check that the computer camera and microphone work as intended (a testing session with the researcher will be held to ensure everything works correctly)	<input type="checkbox"/>
Be prepared	
Have any necessary materials available for the session	<input type="checkbox"/>
Have a drink of water available nearby	<input type="checkbox"/>

1

Teresa Garrido-Tamayo
PhD Speech & Language Sciences
Email: t.garrido-tamayo2@newcastle.ac.uk

Things I need to do during the session

1. Be present in the same room with my child for the duration of the session.
2. Assist with technology on my child's end by accessing Zoom and making sure that the computer camera and microphone work correctly throughout the session.
3. Assist with sharing with my child any necessary materials previously provided, as indicated by the researcher.
4. Encourage my child to listen and participate in the different tasks, especially if they get distracted.

The session will consist of four parts:

1. Introduction

I will welcome you and your child and explain the session plan using a visual schedule. Please ensure that you are both listening and ask any questions you might have at this point.

2. Feelings Game

Your child will play a game where they will learn about how we can express our feelings by using our voices. During this game, your child will be asked to guess if people's voices match how they feel, whether this is happy, sad, angry, or scared.

3. Break Time

At this point, we will have a short break. This is a good time for you and your child to have a quick stretch or a little jump, and to go to the bathroom if needed.

4. Storytime

We will cover three stories. First, your child will be asked to tell a story, and then we will practice telling this story and a new one together. Finally, your child will be asked to tell a third story. Please encourage your child to listen and participate in the activities.

Appendix 7. Extract from the Dynamic Assessment Protocol Used in the Child Pilot
Study 1: Receptive Affective Prosody Task Script

Slide	Pre-test phase script for <i>Receptive Affective Prosody Task</i>
1-2	<p>To start with, we are going to play the <i>Feelings Game</i>. In this game, we're going to talk and learn about our feelings. We all have different feelings that we experience for different reasons. In our game, we're going to focus on four feelings: happy, sad, angry, and scared. For example, we can feel...</p> <ul style="list-style-type: none"> - Happy is when something good has happened and we smile a lot. This picture means happy because the person is smiling. - Sad is when we feel like crying because something bad or upsetting has happened. This picture means sad because the person has a sad, unhappy face. - Angry is when we feel very annoyed and want to shout or break something because something we don't like has happened. This picture means angry. - We can also feel scared if we are filled with fear or frightened by something scary or worrying. This picture means scared.
4-7	<ul style="list-style-type: none"> - For example, in this picture we can see a girl meeting her friend at the airport. The girls haven't seen each other for such a long time and they look very happy! They're both smiling and just about to hug. - What about the boy in this picture? How does he feel? What does he have on his face? <i>[A tear]</i> You can see other boys playing football who seem to be having fun, but this boy looks very sad, perhaps because they don't want to play with him. - In this picture we can see a little boy who is watching a scary film on TV. How does he feel? He looks very scared and his teddy bear looks scared too! - In this picture we see a woman knocking on the door of a house where children are playing inside. The children are making lots of noise, jumping around on a chair and playing music with pot and pans. The woman is going to tell the children to stop the noise and she looks very angry. <p><i>[Show four pictograms at once]</i> Now that you know how to recognise feelings on the face, point at the picture that looks... happy, sad, angry, scared.</p>
9	<p>Up until now we have learned that we can show how we feel by using our faces (like by making a happy or a sad face), but there are other things we can use to show how we feel to others. For example, we can show our feelings by using our voices, changing the way we talk.</p> <p><i>[Show happy pictogram and play audio clip of sentence 'My friends are great' said with prosody conveying happiness]</i> Listen to how this person sounds. They have a happy face, and they sound happy.</p>
10	<p>Now let's do something a little different! First, you will see a picture of someone who is doing something. They might be happy or sad, angry or scared. Then you</p>

	will hear that person talking. You will have to tell me if their voice matches how they feel. Sometimes the voice matches how they feel, but other times it doesn't. This is just part of the game. Let's try!
11-12	<p><i>Practice item 1:</i> Here we see a boy who offers to share his sandwich with the girl. The girl seems sad because her sandwich has fallen into a puddle. Listen to the girl talk. Does she sound sad? <i>[Play audio clip of the sentence 'The sandwich is wet' said in a congruent sad prosody. If child is correct say, 'Yes, well done! Her voice sounds like she is sad']</i></p> <p><i>Practice item 2:</i> Here we see a girl who has finished her day at school and is going to take the bus home. However, when she gets outside, she sees the bus is gone. She missed the bus. Listen to the girl talk. Does she sound sad because she missed the bus? <i>[Play audio clip of the sentence 'The bus is gone' said in an incongruent happy prosody. If the child is correct, say, 'Yes, well done! Her voice sounds like she is happy, which is not right because she feels sad. Her voice doesn't match how she feels', and begin the prosody task by presenting the 16 audio clips consecutively. Otherwise, repeat practice items until the child responds correctly]</i></p>
13	Here we see a little mouse who is scared of the big eagle flying towards him. Listen to the mouse talk. Does he sound scared? <i>[Play audio clip of the sentence 'The eagle is big' said in a congruent scared prosody]</i>
14	The little boy and the cat look at the butterflies. They look happy and peaceful. Listen to the little boy talk. Does he sound happy? <i>[Play audio clip of the sentence 'The butterflies are pretty' said in an incongruent sad prosody]</i>
15	The dog and the cat have muddy paws after playing outside in the rain. The woman looks angry. Listen to her talk. Does she sound angry? <i>[Play audio clip of the sentence 'The floor is dirty' said in a congruent angry prosody]</i>
16	The girl asks her mum to buy her a dinosaur toy for her birthday. However, it's Sunday and the shop is closed. Listen to the girl talk. Does she sound sad? <i>[Play audio clip of the sentence 'The shop is closed' said in an incongruent happy prosody]</i>
17	The baby elephant is a little scared of the water. His mum tells him that it's okay and completely safe. Listen to the baby elephant talk. Does he sound scared? <i>[Play audio clip of the sentence 'The water is cold' said in a congruent scared prosody]</i>
18	The mum looks at the drawing her daughter made. She is smiling and seems to like it. Listen to the mum talk. Does she sound happy? <i>[Play audio clip of the sentence 'The drawing is lovely' said in an incongruent sad prosody]</i>
19	The little boy is making a mess on the wall with paint. His dad is not happy and tells him to stop. Listen to the dad talk. Does he sound angry? <i>[Play audio clip of the sentence 'The wall is filthy' said in an incongruent scared prosody]</i>
20	The little boy asks his dad for a green balloon, but the balloon seller says he doesn't have that colour. The little boy is sad. Listen to him talk. Does he sound sad? <i>[Play</i>

	<i>audio clip of the sentence 'These balloons aren't green' said in a congruent sad prosody]</i>
21	The wolf is strong and powerful. The baby penguin fears him. Listen to the baby penguin talk. Does she sound scared? <i>[Play audio clip of the sentence 'The wolf is scary' said in an incongruent angry prosody]</i>
22	The boy buys a comic book. He loves comic books and seems happy. Listen to him talk. Does he sound happy? <i>[Play audio clip of the sentence 'These books are fun' said in a congruent happy prosody]</i>
23	The children littered the floor. The man tells them to stop. He seems angry! Listen to him talk. Does he sound angry? <i>[Play audio clip of the sentence 'The ground is disgusting' said in a congruent angry prosody]</i>
24	The children are playing basketball. The boy with the glasses is sad because his team is losing. Listen to the boy talk. Does he sound sad? <i>[Play audio clip of the sentence 'My team is losing' said in an incongruent happy prosody]</i>
25	The cheetah is chasing the zebra. The zebra is scared and is running as fast as she can. Listen to the zebra talk. Does she sound scared? <i>[Play audio clip of the sentence 'The cheetah is fast' said in an incongruent angry prosody]</i>
26	The man hugs his dog. They both look very happy! Listen to the man talk. Does he sound happy? <i>[Play audio clip of the sentence 'This dog is adorable' said in a congruent happy prosody]</i>
27	Two children are making a lot of noise in the school corridor. The teacher is a little angry and tells them to be quiet. Listen to the teacher talk. Does she sound angry? <i>[Play audio clip of the sentence 'The children are noisy' said in an incongruent scared prosody]</i>
28	The little boy is crying because his friend has taken the toy car from him. His mum suggests that they could play together. Listen to the little boy talk. Does he sound sad? <i>[Play audio clip of the sentence 'She is being mean' said in a congruent sad prosody]</i>
29	Good job! You've earned your first star!

Appendix 8. Scoring Sheet for the Receptive Affective Prosody Task Used in the Child Pilot Study 1

Receptive affective prosody task: Pretest phase		Score	Notes
1	The eagle is big	0	1
2	The butterflies are pretty	0	1
3	The floor is dirty	0	1
4	The shop is closed	0	1
5	The water is cold	0	1
6	The drawing is lovely	0	1
7	The wall is filthy	0	1
8	These balloons aren't green	0	1
9	The wolf is scary	0	1
10	These books are fun	0	1
11	The ground is disgusting	0	1
12	My team is losing	0	1
13	The cheetah is fast	0	1
14	This dog is adorable	0	1
15	The children are noisy	0	1
16	She is being mean	0	1
Total score (sum of 1 to 16)			

Appendix 9. Information Sheet Used in the Adult Pilot Study of the Receptive Affective Prosody Task

Evaluating the effectiveness of a receptive affective prosody assessment task

Participant Information Sheet

Dear Participant,

You are being invited to take part in a research study entitled *Evaluating the effectiveness of a receptive affective prosody assessment task*. I am conducting this study as part of my PhD project on the assessment of language difficulties in children learning English as an additional language. This project is under the supervision of Dr Carolyn Letts and Dr Laurence White from the School of Education, Communication & Language Sciences at Newcastle University.

Before you decide whether you wish to take part, it is essential that you understand why the research is being done and what it will involve. Please read this form carefully and ask any questions you may have. If you do decide to take part, you will be asked to sign a consent form. However, you are free to withdraw at any time, without giving any reason and without any negative consequences for you.

What is the purpose of the research?

The purpose of this study is to pilot a receptive affective prosody task with English-native adult speakers, though non-native speakers can also take part. This task will ultimately be used to assess children's ability to interpret emotions in people's voices. This is an important ability that some children can struggle with, such as those with autism and developmental language disorder.

What does taking part involve?

Participation in this study will involve providing some basic personal information, including your age, gender, whether you have any speech, language, or hearing difficulties, whether you are a native English speaker, and other languages you may know. Following this, you will be asked to take part in three similar activities:

1. You will listen to a series of sentences and have to indicate whether or not the voice you hear matches how the person feels.
2. You will listen to a series of pairs of sentences. For each pair, you have to indicate the one in which the voice better matches the intended emotion.
3. You will listen to a series of sentences and have to indicate which of two possible emotions the voice was expressing.

In all three activities, you will be asked to record your responses on a paper response sheet.

Your participation in this study will take approximately 40 minutes.



Evaluating the effectiveness of a receptive affective prosody assessment task

Confidentiality and data management

All data collected will be anonymised and stored securely on password-protected University servers and in a locked filing cabinet at Newcastle University. This data will only be held for the purpose of this research and will be destroyed within 5 years. It will not be possible to identify individual participants from the resulting PhD thesis and/or other research outputs, such as journal articles and conference or seminar presentations.

Ethical approval

The study has been approved by the Faculty of Humanities & Social Sciences Research Ethics Committee at Newcastle University.

Risks and benefits

There are no foreseeable risks associated with this study. Your valuable participation would contribute to the development of a receptive affective prosody task to assess children's ability to understand emotions in people's voices. This task will be used within a broader assessment resource to help identify children with English as an additional language at risk of having developmental language disorder and/or who need extra help with their English.

Many thanks for taking the time to read this. If you have any questions or would like more information, please do not hesitate to contact me at t.garrido-tamayo2@newcastle.ac.uk. You can also email my supervisors Dr Carolyn Letts at carolyn.letts@newcastle.ac.uk and Dr Laurence White at laurence.white@newcastle.ac.uk.

Best wishes,

Teresa Garrido-Tamayo



Appendix 10. Consent Form Used in the Adult Pilot Study of the Receptive Affective Prosody Task

Evaluating the effectiveness of a receptive affective prosody assessment task

Consent Form

Thank you for your interest in taking part in this research. Please complete this form after you have read the Participant Information Sheet.

Please put your initials in the box to indicate you have read and agree with the following statements:

	Initials
1. I have read the information sheet for the study entitled <i>Evaluating the effectiveness of a receptive affective prosody assessment task</i> . I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.	
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. I understand that if I decide to withdraw, any data that I have provided up to that point will be omitted from the study.	
3. I understand that all data collected will be anonymised and confidentiality will be maintained in all outputs relating to this study. This data will only be held for the purpose of this research and will be destroyed within 5 years.	
4. I agree to take part in this research project.	

Participant

Signature:

Date:

Print full name:



Appendix 11. Adult Trial of the Audio Stimuli in the Congruent vs Neutral Prosody Version of the Receptive Affective Prosody Task

The stimuli sentences for the *Congruent vs Neutral Prosody* version of the *Receptive Affective Prosody Task* were initially audio recorded by two female British English-native speakers as part of their psychology student placement experience. However, only one set of recordings were to be used, in order to keep the speaker's voice consistent throughout the task, both within and between pairs of sentences. For this reason, in addition to the researcher and her supervisors evaluating the two sets of recordings, the non-neutral recordings were further trialled with four British English-native adult speakers (two males aged 31 and 63 and two females aged 35 and 63) in January 2022.

Sentences were played over loudspeakers in random order. After listening to each sentence twice—once per speaker—they were asked which emotion they thought the voice evoked in each case (i.e., happiness, sadness, anger, fear or other). A score of 1 was awarded each time a listener identified the recording in accordance with the intended affective prosody, and a score of 0 was given if the listener's response did not align with the type of prosody the stimulus was supposed to represent. Total scores obtained for every sentence in each emotion category (happy, sad, angry, and scared) were calculated for both speakers (see Table 11A). The four listeners recognised the intended emotions more frequently in Speaker 1's audio recordings, except for the angry utterances, for which Speaker 2's recordings were preferred on the whole. Therefore, Speaker 1's recordings were selected, with the angry utterances and any others being re-recorded by this speaker as necessary. Audio files were edited as required using Audacity audio software.

Table 11A. Scores for the Adult Trial of the Audio Stimuli in the Congruent vs Neutral Prosody Version of the Receptive Affective Prosody Task

Sentences		Scores	
Happy sentences		Speaker 1	Speaker 2
Today was great		3	1
The butterflies are pretty		4	0
The drawing is lovely		3	1
These books are fun		4	0
This dog is adorable		3	3
Total score		17	5
Sad sentences		Speaker 1	Speaker 2
The sandwich is wet		1	2
The shop is closed		1	2
These balloons aren't green		4	0
My team is losing		2	2
She is being mean		3	1
Total score		11	7
Angry sentences		Speaker 1	Speaker 2
The floor is dirty		0	4
The wall is filthy		2	3
The ground is disgusting		1	3
The children are noisy		1	4
Total score		4	14
Scared sentences		Speaker 1	Speaker 2
The eagle is big		4	0
The water is cold		2	0
The wolf is scary		2	2
The cheetah is fast		1	0
Total score		9	2

Appendix 12. Response Tables Used in the Adult Pilot Study of the Receptive Affective Prosody Task

Activity A		Response	
Practice item		Happy	Sad
Practice item	My hair is short	Angry	Scared
1	The phone is ringing	Angry	Scared
2	The ball is red	Happy	Sad
3	We are reading	Angry	Scared
4	They are napping	Happy	Sad
5	I am cooking	Angry	Scared
6	I am waking up	Happy	Sad
7	The animals are eating	Angry	Scared
8	The car is parked	Happy	Sad
9	The cat is brown	Happy	Sad
10	The flowers are growing	Angry	Scared
11	My friends are running	Happy	Sad
12	The moon is round	Happy	Sad
13	The table is green	Angry	Scared
14	The tap is running	Angry	Scared
15	The dog is playing	Angry	Scared
16	My dad is fishing	Happy	Sad

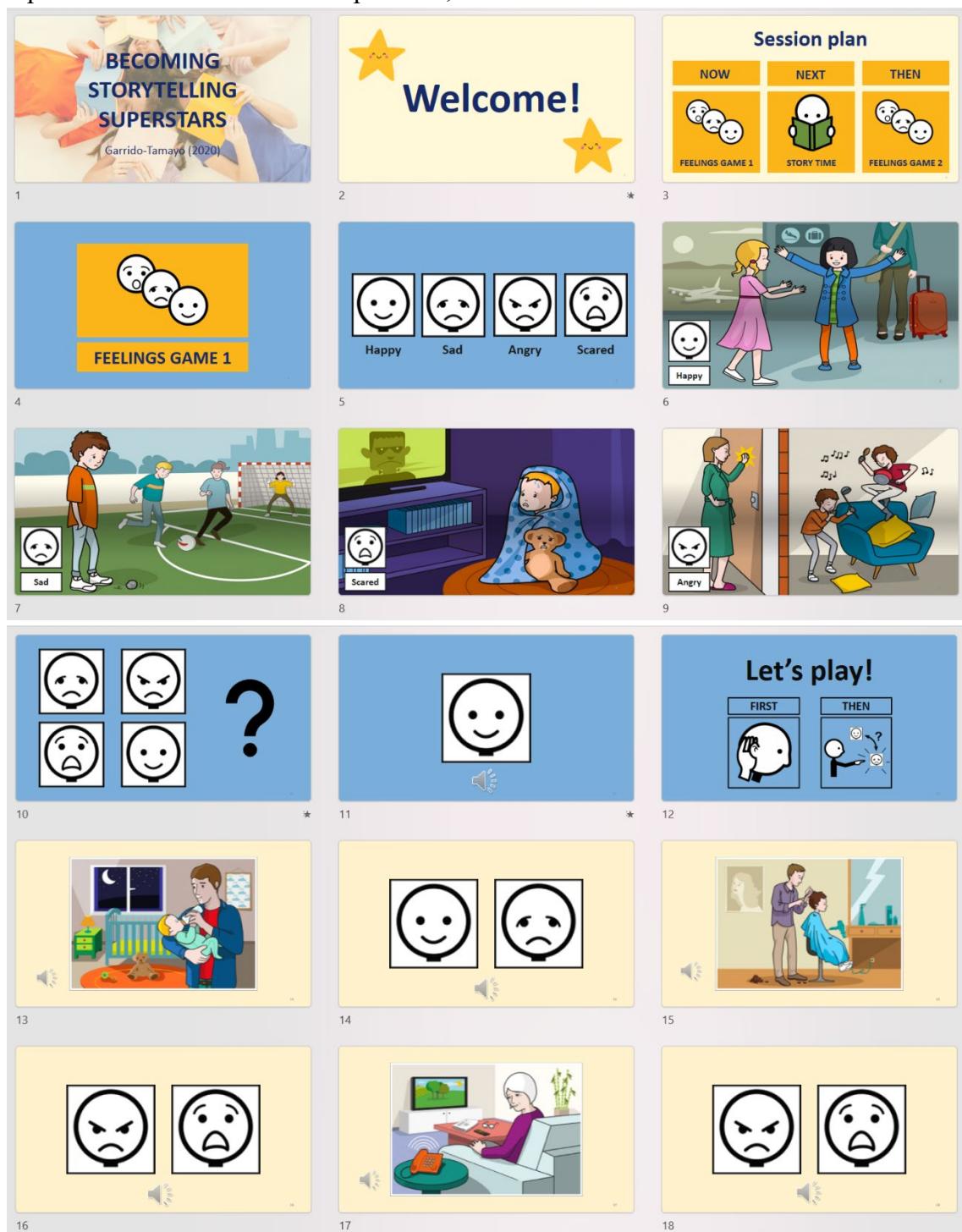
Activity B		Response	
Practice item	The sandwich is wet	1	2
Practice item	Today was great	1	2
1	The eagle is big	1	2
2	The butterflies are pretty	1	2
3	The floor is dirty	1	2
4	The shop is closed	1	2
5	The water is cold	1	2
6	The drawing is lovely	1	2
7	The wall is filthy	1	2
8	These balloons aren't green	1	2
9	The wolf is scary	1	2
10	These books are fun	1	2
11	The ground is disgusting	1	2
12	My team is losing	1	2
13	The cheetah is fast	1	2
14	This dog is adorable	1	2
15	The children are noisy	1	2
16	She is being mean	1	2

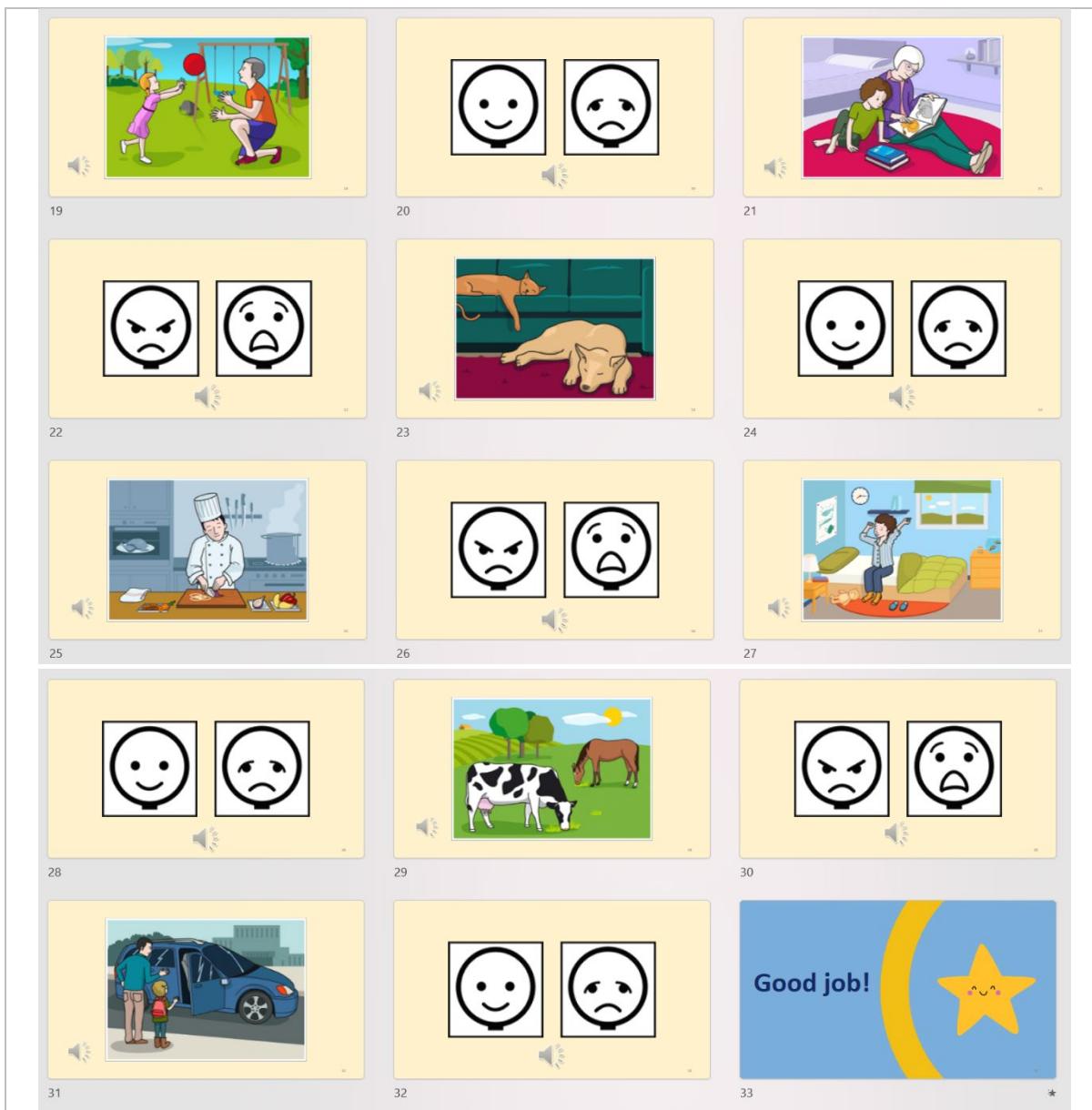
Activity C		Response	
Practice item	The sandwich is wet	Yes	No
Practice item	The bus is gone	Yes	No
1	The eagle is big	Yes	No
2	The butterflies are pretty	Yes	No
3	The floor is dirty	Yes	No
4	The shop is closed	Yes	No
5	The water is cold	Yes	No
6	The drawing is lovely	Yes	No
7	The wall is filthy	Yes	No
8	These balloons aren't green	Yes	No
9	The wolf is scary	Yes	No
10	These books are fun	Yes	No
11	The ground is disgusting	Yes	No
12	My team is losing	Yes	No
13	The cheetah is fast	Yes	No
14	This dog is adorable	Yes	No
15	The children are noisy	Yes	No
16	She is being mean	Yes	No

Appendix 13. Receptive Affective Prosody Task Script Used in the Child Pilot Study 2 and Main Study

13A. Pretest Phase

To clarify the context of the script that follows, an overview of the slides used during the pretest *Receptive Affective Prosody Task* is provided below. This overview pertains to children who followed the AB sequence (*Story A* in the pretest and *Story B* in the post-test). For children in the BA sequence (*Story B* in the pretest and *Story A* in the post-test), the slides were identical, except that the *Receptive Affective Prosody Task* stimuli used here for the pretest were then used for the post-test, and vice versa.





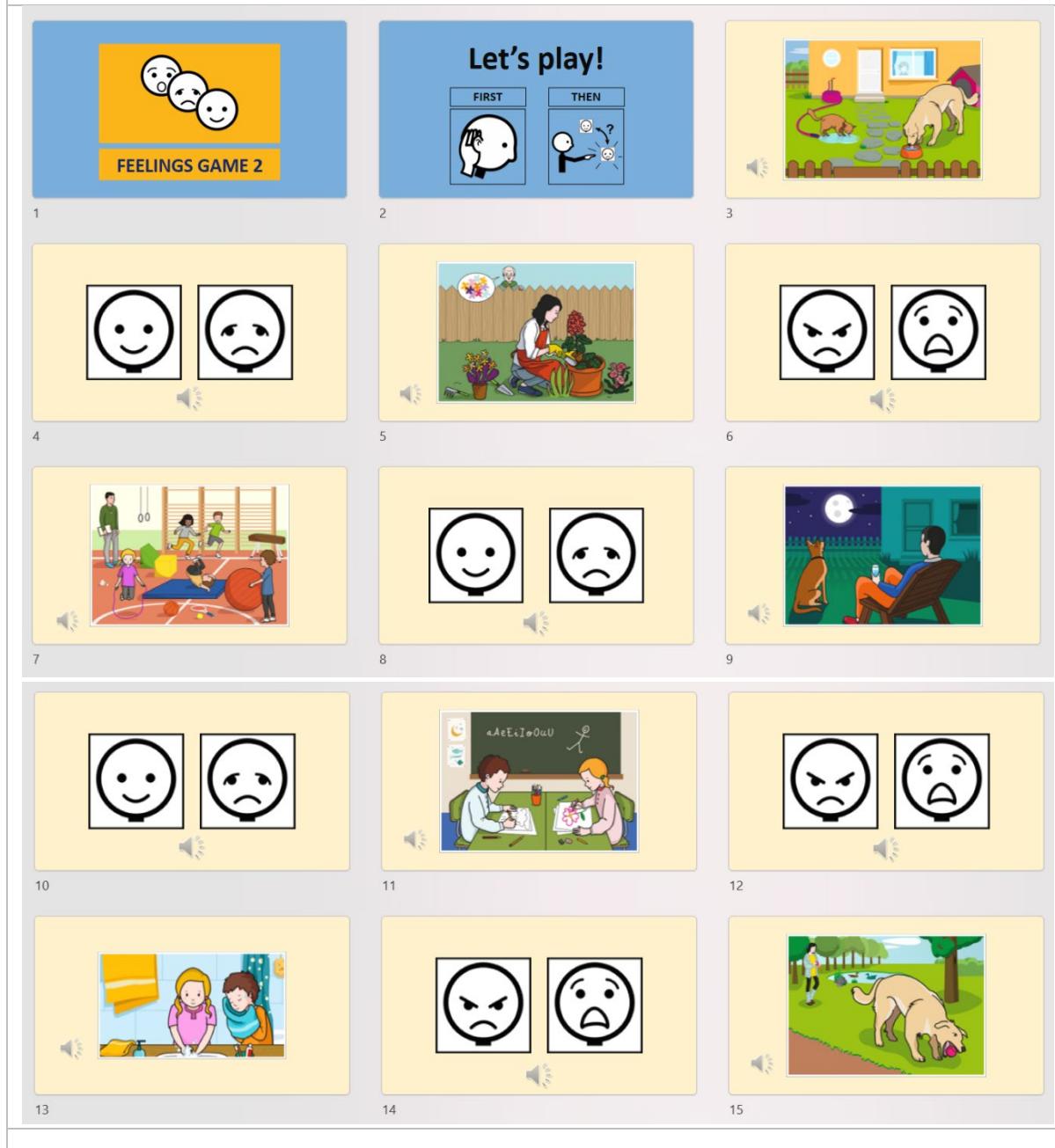
Slide	Pre-test phase script for Receptive Affective Prosody Task (Task order AB)
4	<p>To start with, we are going to play the <i>Feelings Game</i>. In this game, we're going to talk and learn about our feelings. We all have different feelings that we experience for different reasons.</p>
5	<p>In our game, we're going to focus on four feelings: happy, sad, angry, and scared. For example, we can feel...</p> <ul style="list-style-type: none"> - Happy is when something good has happened and we smile a lot. This picture means happy because the person is smiling. - Sad is when we feel like crying because something bad or upsetting has happened. This picture means sad because the person has a sad, unhappy face. - Angry is when we feel very annoyed and want to shout or break something because something we don't like has happened. This picture means angry.

	<ul style="list-style-type: none"> - We can also feel scared if we are filled with fear or frightened by something scary or worrying. This picture means scared.
6-9	<ul style="list-style-type: none"> - For example, in this picture we can see a girl meeting her friend at the airport. The girls haven't seen each other for such a long time and they look very happy! They're both smiling and just about to hug. - What about the boy in this picture? How does he feel? What does he have on his face? <i>[A tear]</i> You can see other boys playing football who seem to be having fun, but this boy looks very sad, perhaps because they don't want to play with him. - In this picture we can see a little boy who is watching a scary film on TV. How does he feel? He looks very scared and his teddy bear looks scared too! - In this picture we see a woman knocking on the door of a house where children are playing inside. The children are making lots of noise, jumping around on a chair and playing music with pot and pans. The woman is going to tell the children to stop the noise and she looks very angry.
10	<p><i>[Show four pictograms at once]</i> Now that you know how to recognise feelings on the face, point at the picture that looks... happy, sad, angry, scared.</p>
11	<p>Up until now we have learned that we can show how we feel by using our faces (like by making a happy or a sad face), but we can also show our feelings by using our voices, changing the way we talk.</p> <p>Listen to how this person sounds. They have a happy face, and they sound happy.</p> <p><i>[Show happy pictogram and play audio clip of sentence 'My friends are great' said with prosody conveying happiness]</i></p>
12	<p>Now let's do something a little different! You have to guess how people are feeling based on how their voice sounds. Sometimes the voice will be happy. And sometimes the voice will be sad, or angry, or scared. Each time, first listen and then point to the picture that shows how the person is feeling. Let's try!</p> <p><i>[Verbal responses will be accepted, i.e., if the child says the name of the emotion instead of pointing at the picture]</i></p>
13-14	<p><i>[Begin with Practice item 1]</i> Here the dad gives the baby milk. Listen to the dad talk and point at the picture. Does he sound happy or sad?</p> <p><i>[Play audio clip of the sentence 'The baby is drinking' said with prosody expressing happiness. If child is correct say, 'Yes, well done! His voice sounds like he is happy'.]</i></p>
15-16	<p><i>Practice item 2:</i> Here we can see a boy getting a haircut. Listen to him talk and point at the picture. Does he sound angry or scared?</p> <p><i>[Play audio clip of the sentence 'My hair is short' said with prosody expressing anger. If child is correct say, 'Yes, well done! His voice sounds like he is angry', and begin the prosody task by presenting the 16 audio clips consecutively. Otherwise, repeat training items until child responds correctly.]</i></p>

17- 18	The woman hears the phone ring. Listen to her talk and point at the picture. Does she sound angry or scared? [Play audio clip of the sentence 'The phone is ringing' said in a scared prosody]
19- 20	The girl and her dad play with the ball. Listen to the girl talk and point at the picture. Does she sound happy or sad? [Play audio clip of the sentence 'The ball is red' said in a happy prosody]
21- 22	The girl and her grandmother read together. Listen to the girl talk and point at the picture. Does she sound angry or scared? [Play audio clip of the sentence 'We are reading' said in an angry prosody]
23- 24	The girl sees the cat and the dog sleeping. Listen to her talk and point at the picture. Does she sound happy or sad? [Play audio clip of the sentence 'They are napping' said in a sad prosody]
25- 26	The chef is cooking dinner. Listen to him talk and point at the picture. Does he sound angry or scared? [Play audio clip of the sentence 'I am cooking' said in an angry prosody]
27- 28	The boy is getting up to go to school. Listen to him talk and point at the picture. Does he sound happy or sad? [Play audio clip of the sentence 'I am waking up' said in a sad prosody]
29- 30	The cow and the horse eat grass. Listen to the farmer talk and point at the picture. Does he sound angry or scared? [Play audio clip of the sentence 'The animals are eating' said in a scared prosody]
31- 32	The girl and her dad get out of the car and wave goodbye. Listen to the girl talk. Does she sound happy or sad? [Play audio clip of the sentence 'The car is parked' said in a happy prosody]
33	Good job! You've earned your first star! You can colour in the first star on the chart.

13B. Post-test Phase

To clarify the context of the script that follows, an overview of the slides used during the post-test *Receptive Affective Prosody Task* is provided below. This overview pertains to children who followed the AB sequence (*Story A* in the pretest and *Story B* in the post-test). For children in the BA sequence (*Story B* in the pretest and *Story A* in the post-test), the slides were identical, except that the *Receptive Affective Prosody Task* stimuli used here for the pretest were then used for the post-test, and vice versa.





16



17



18

Amazing! You are now a...

Storytelling



19

Slide	Post-test phase script for Receptive Affective Prosody Task (Task order AB)
1-2	To finish with, we are going to play the second part of the Feeling Game. Just like before, you have to guess how people feel based on how their voice sounds. Sometimes the voice will be happy, and sometimes the voice will be sad, angry or scared. Each time, first listen and then point at the picture that shows how the person is feeling.
3-4	The girl looks out the window and sees the cat and the dog. Listen to the girl talk. Does she sound happy or sad? <i>[Play audio clip of the sentence 'The cat is brown' said in a happy prosody]</i>
5-6	The man sees the gardener planting flowers. Listen to the man talk. Does he sound angry or scared? <i>[Play audio clip of the sentence 'The flowers are growing' said in an angry prosody]</i>
7-8	The children do exercise. Listen to one of them talk. Does she sound happy or sad? <i>[Play audio clip of the sentence 'My friends are running' said in a sad prosody]</i>
9-10	The man and his dog look at the moon. Listen to the man talk. Does he sound happy or sad? <i>[Play audio clip of the sentence 'The moon is round' said in a happy prosody]</i>
11-12	The children draw at the table. Listen to the girl and point at the picture. Does she sound angry or scared? <i>[Play audio clip of the sentence 'The table is green' said in a scared prosody]</i>
13-14	The children are in the bathroom. Listen to the boy talk. Does he sound angry or scared? <i>[Play audio clip of the sentence 'The tap is running' said in an angry prosody]</i>
15-16	The dog picks up the ball. Listen to the woman talk. Does she sound angry or scared?

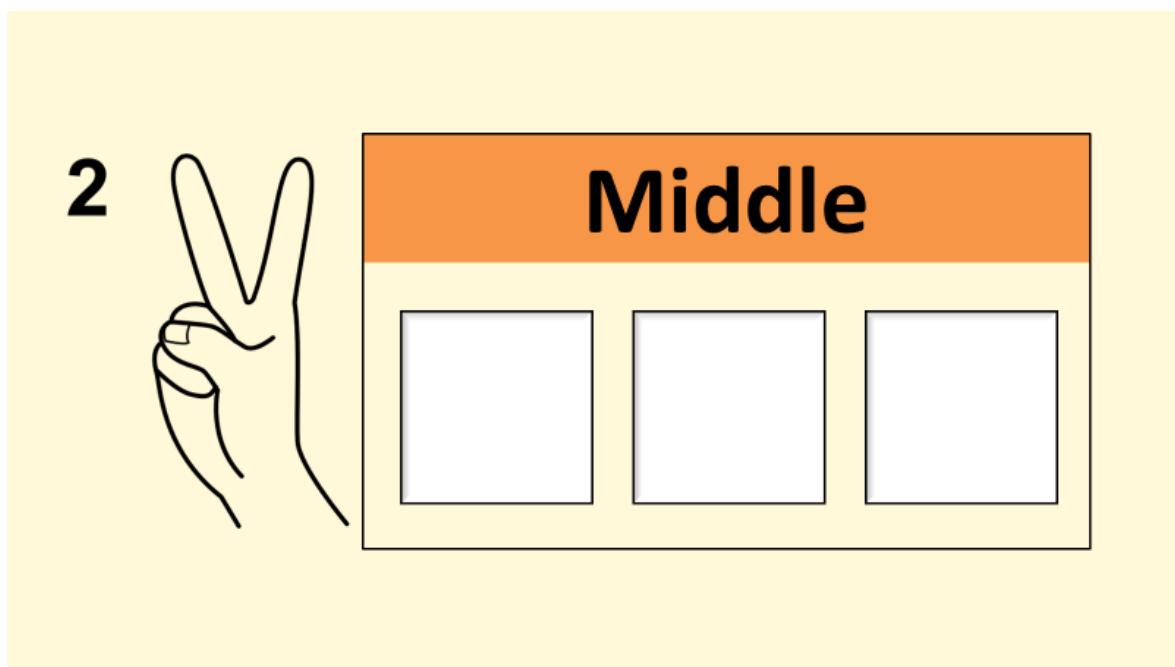
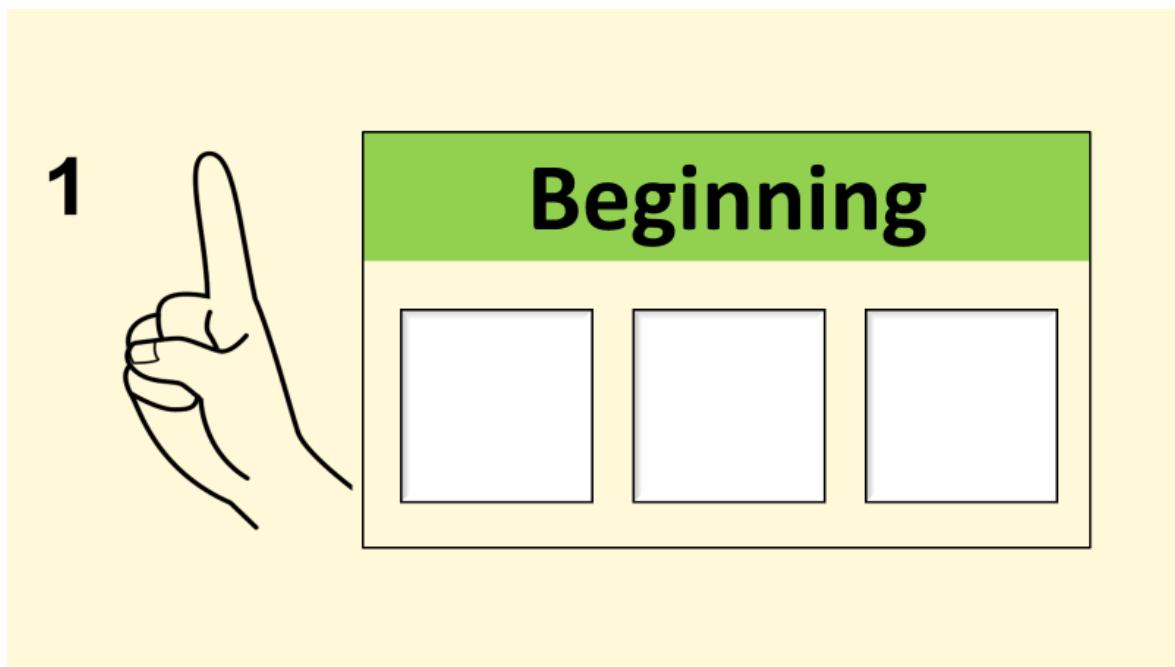
	<i>[Play audio clip of the sentence 'The dog is playing' said in a scared prosody]</i>
17- 18	The boy and his dad are fishing on the ice. Listen to the boy talk. Does he sound happy or sad? <i>[Play audio clip of the sentence 'My dad is fishing' said in a sad prosody]</i>
19	Amazing! You have done a fantastic job completing all the activities and are now a storytelling superstar! We have learned lots about expressing our feelings using our voices and telling really good stories that include all the different parts.

Appendix 14. Scoring Sheets for the Receptive Affective Prosody Task Used in the Child Pilot Study 2 and Main Study

Please select as appropriate: Pretest/Post-test				
Practice items (if pretest)		Response		Notes
P1	The baby is drinking	Happy	Sad	
P2	My hair is short	Angry	Scared	
Test items		Response		Score
1	The cat is brown	Happy	Sad	
2	The flowers are growing	Angry	Scared	
3	My friends are running	Happy	Sad	
4	The moon is round	Happy	Sad	
5	The table is green	Angry	Scared	
6	The tap is running	Angry	Scared	
7	The dog is playing	Angry	Scared	
8	My dad is fishing	Happy	Sad	
Total score			/8	

Please select as appropriate: Pretest/Post-test				
Practice items (if pretest)		Response		Notes
P1	The baby is drinking	Happy	Sad	
P2	My hair is short	Angry	Scared	
Test items		Response		Score
1	The phone is ringing	Angry	Scared	
2	The ball is red	Happy	Sad	
3	We are reading	Angry	Scared	
4	They are napping	Happy	Sad	
5	I am cooking	Angry	Scared	
6	I am waking up	Happy	Sad	
7	The animals are eating	Angry	Scared	
8	The car is parked	Happy	Sad	
Total score			/8	

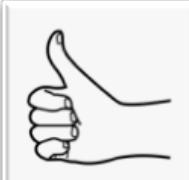
Appendix 15. Paper Copies of the Graphic Organisers and Pictograms Used as Part of the Storytelling Activity in the Main Study



3



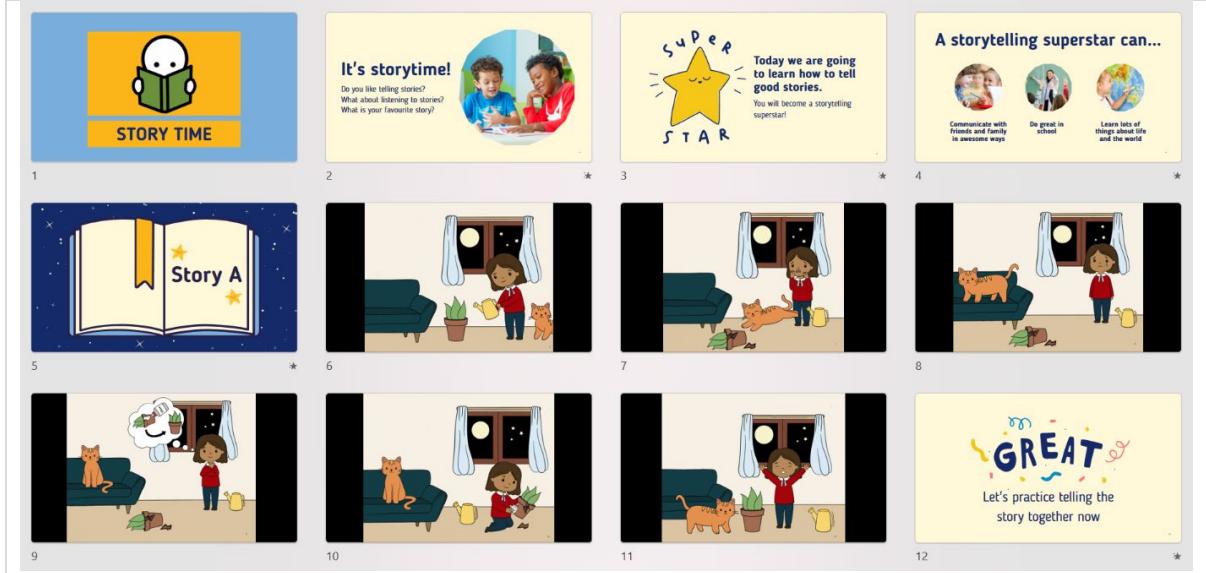
End



Appendix 16. Script for the Storytelling Activity Used in the Main Study

16A. Pretest Phase

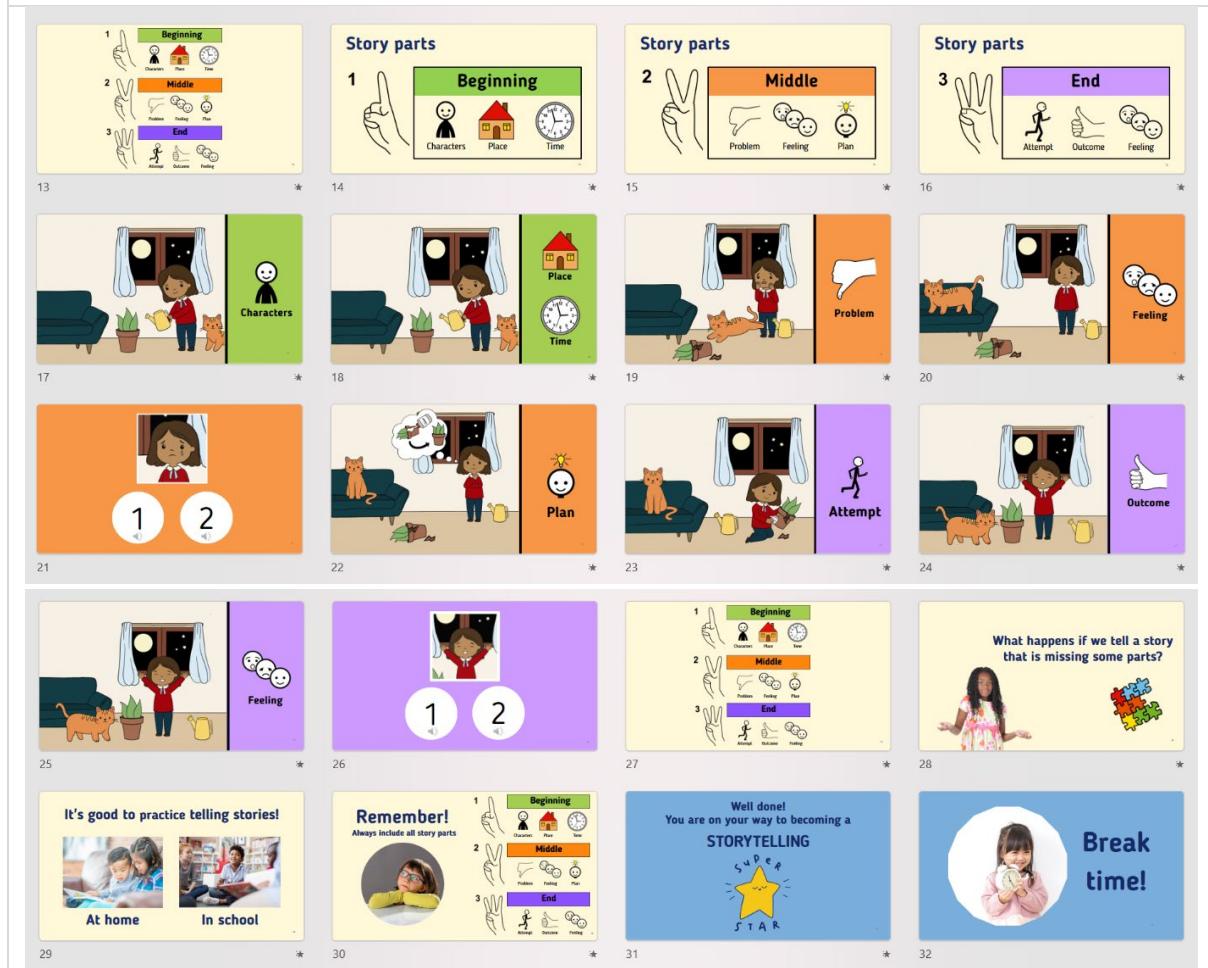
To clarify the context of the script that follows, an overview of the slides used during the pretest phase of the storytelling activity is provided below. This overview pertains to children who followed the AB sequence, experiencing *Story A* as the pretest and *Story B* as the post-test. For children in the BA sequence, the slides were identical, except that *Story B* was used in the pretest and *Story A* in the post-test.



Slide	Pre-test phase script for stories A/B	MLE strategies
1	<ul style="list-style-type: none"> Now it's story time! 	(Warm-up time)
2	<ul style="list-style-type: none"> Do you like telling stories? What about listening to stories? What is your favourite story? <i>[Possible warm-up questions]</i> <i>[If the child is hesitant to respond or indicates they do not like stories/books, say:]</i> Stories can also come in the form of films/movies, video games... Not only books! Do you like any of those? 	
3	<ul style="list-style-type: none"> Today we are going to learn how to tell good stories. You will become a storytelling superstar! 	Intentionality
4	<ul style="list-style-type: none"> A storytelling superstar can tell stories very well. Telling stories well is important because it helps us to communicate with friends and family, do really well in school, and learn lots of things about life and the world. 	Meaning, transcendence
5	<ul style="list-style-type: none"> We are going to start with this story. First, let's have a look at all the pictures here. <i>[Wait for the child to look at the pictures; ask them to indicate when they want to see the next picture.]</i> <i>[Once you reach Story A/B's picture 6 say:]</i> Okay, let's go back to the start. <i>[and then say:]</i> Now look at the pictures again and tell me the story that you see. I will be quiet and listen to your story. 	N/A
6-11	<ul style="list-style-type: none"> Tell me the story that you see, starting with this picture. <i>[Wait for the child to tell the story; prompt them to continue if they fall silent, e.g., 'Tell me more.' Do not give feedback based on the child's performance at this stage.]</i> 	

16B. Teaching Phase

To clarify the context of the first script that follows, an overview of the slides used during the pretest phase of the storytelling activity is provided below. This overview pertains to children who followed the AB sequence, experiencing *Story A* as the pretest and *Story B* as the post-test. For children in the BA sequence, the slides were identical, except that *Story B* was used in the pretest and *Story A* in the post-test.



Slide	Teaching phase script for <i>Story A</i> (Task Order AB)	MLE strategies
12	<ul style="list-style-type: none"> <i>[Praise the child's performance when they finish telling Story A regardless of how they did.]</i> Great! Well done! Now we are going to practice telling this story together. We are going to learn all the parts that a story should have. 	Competence, intentionality
13	<ul style="list-style-type: none"> Good stories have a beginning, a middle and an end. 	(Overview of the story parts)
14	<ul style="list-style-type: none"> <i>[As you introduce each story grammar unit, stick the pictograms representing these on the graphic organiser.]</i> At the beginning of the story, we talk about: <ul style="list-style-type: none"> The characters. These are the people or the animals in the story, and Where and when the story takes place; that is, place and time. 	
15	<ul style="list-style-type: none"> Next, in the middle of the story, we talk about: <ul style="list-style-type: none"> The problem. This is something that happens that needs to be fixed or sorted out. How the characters feel about the problem, and What they plan to do to fix the problem. 	
16	<ul style="list-style-type: none"> Finally, at the end of the story, we talk about: <ul style="list-style-type: none"> The attempt. This is what the characters do to fix the problem. The outcome. This is what happens when they try to fix it, and How they feel about the outcome. 	
6-11	<ul style="list-style-type: none"> <i>[Provide the child with a model of the story before discussing each story grammar unit one by one. Point at the relevant story pictures and printed story grammar pictograms as you do so.]</i> In the story you have just told, a girl is watering a plant. There is also a cat. It's night time, and they are in the living room at their home. The cat jumps and knocks over the plant. The plant pot breaks, and the girl feels sad about it. She then plans to fix the plant pot so that the plant will be okay. The girl tries to fix the broken plant pot with glue. She manages to fix it and feels happy because her plant looks healthy and good again. 	(Model of the story)
13	<ul style="list-style-type: none"> It's important that we include all these parts so that the story is complete and people can understand what happens in it. 	Meaning, intentionality

	<ul style="list-style-type: none"> Now we are going to learn a little more about every part that needs to be in the story you have just told. 	
17	<ul style="list-style-type: none"> First, let's talk about the characters. These are the people and the animals in the story. <i>[Click for the Characters pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick the Characters pictogram on the graphic organiser; e.g., 'Put the Characters picture here to remember to include them in your story'. Do this successively for all story grammar units. When possible, provide encouraging feedback to acknowledge any story grammar units that the child already included in the pretest, e.g., 'I know you are good at this bit because you said it earlier when you were telling the story!']</i> Talking about the characters is important because it's their story that we are going to follow, so we will want to know about them. For example, we could say what their names are, what they are doing, and what they look like. Name: <i>[Refer to picture 1]</i> In this story: Does the girl have a name? <i>[Wait for response]</i> What about the cat? <i>[Wait for response]</i> <i>[Use the names the child gives you from now on; if the child does not respond within a few seconds, follow up with a model sentence: 'You could say, the girl is called (Rita), and her cat is called (Gus). Now you say it.']</i> Activity: We also need to say what they are doing. What is (Rita) doing here? <i>[Let the child describe the activity; encourage/praise and recast/expand as required; if the child cannot tell the expected target within a few seconds, follow up with a model sentence: '(Rita) is watering the plant. Now you say it.']</i> What they look like: We can also describe what they look like. What does (Gus) look like? <i>[Wait for response]</i> What about (Rita)? What does she look like? <i>[Wait for response]</i> <i>[Let the child describe each character, recasting/expanding and encouraging/praising as required (e.g., 'Good job! You have used lots of great words to describe the cat!'). If the child cannot tell the targets within a few seconds, follow up with the model sentences: 'The cat (Gus) is big and orange with stripes on his head and back.', 'The girl (Rita) has medium short brown hair and is wearing a red top. Now you say it.']</i> 	Intentionality, meaning, competence

18	<ul style="list-style-type: none"> • We should also talk about where and when the story happens. That is, place and time. <i>[Click for the Place and Time pictograms to show on the slide and draw the child's attention to it. Encourage the child to stick these pictograms on the graphic organiser.]</i> • Talking about the story's place and time helps us learn about where the characters are and at what time the story happens. • So, in this story: Where do you think they are? In which place? <i>[Wait for response]</i> What time do you think it is? Is it daytime or night time? <i>[Wait for response]</i> <p><i>[Encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentences: '(Rita) and (Gus) are in the living room at home.', 'It's night time because we can see the moon and the stars in the dark sky. Now you say it.]</i></p>	Intentionality, meaning, competence
19	<ul style="list-style-type: none"> • Then in the middle of the story, something happens that needs to be fixed or sorted out. Something goes wrong, and there is a problem. <i>[Click for the Problem pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • We must always say what the problem is to know what goes wrong in the story. • What is the problem in this story? What goes wrong? What happens to (Rita)'s plant? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Gus) the cat knocks over the plant pot and it breaks. Now you say it.]</i> 	Intentionality, meaning, competence
20	<ul style="list-style-type: none"> • After we say what the problem is, we talk about how the character –(Rita), the girl in this story– feels about it. <i>[Click for the Feeling pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • Talking about how (Rita) feels is important because it helps us to understand what feelings or emotions she is experiencing and why she wants to fix the plant pot. • As we saw in the Feelings Game, to express how (Rita) feels, we should use special words like 'happy', 'sad', 'angry' or 'scared', and change our voice to make it sound 'happy', 'sad', 'angry' or 'scared'. To help us remember we have this wheel with pictures. <i>[Illustrate each emotion prosodically and</i> 	Intentionality, meaning, transcendence, competence

	<p><i>through exaggerated facial expressions. Point at the relevant pictograms on the Feelings wheel.]</i></p> <ul style="list-style-type: none"> How does (Rita) feel when she sees her plant pot is broken? Tell me and pick how she feels on the wheel. <i>[Wait for response]</i> <ul style="list-style-type: none"> <i>[If the child responds correctly, say, ‘That’s right, (Rita) feels sad! Well done!’ and move on to the next slide.]</i> <i>[If the child responds incorrectly, say, ‘This plant was important to (Rita), possibly one of her favourite things! Do you have a favourite thing? Maybe a toy you love? How would you feel if this got broken?’ Use the child’s favourite thing as an example to help them understand and relate to the situation. ‘That’s right; you would feel sad! (Rita) is sad because her plant is broken.’ Move on to the next slide when pertinent.]</i> 	
21	<ul style="list-style-type: none"> Listen to (Rita) talk – which one sounds more like her? 1 or 2? <i>[Play the two audio clips of the sentence ‘My plant is broken!’ spoken with happy and sad prosody.]</i> <ul style="list-style-type: none"> <i>[If the child selects the correct audio (2), say, ‘Well done! (Rita) is sad and she sounds sad.’, expressing the sentence with a sad-sounding voice yourself and encouraging the child to do the same by saying, ‘Can you repeat? Make your voice sound sad.’]</i> <i>[If the child selects the incorrect audio (1), play the audio clips again and explain why the sad-sounding audio is the correct one by saying, ‘(Rita) is sad and she is crying!’ Using a sad vocal expression, say, ‘Oh no! I’m sad because my plant is broken! Can you repeat? Make your voice sound sad.’]</i> <i>[If you haven’t said this in the previous slide, say, ‘This plant was important to (Rita), possibly one of her favourite things! Do you have a favourite thing? Maybe a toy you love? How would you feel if this got broken?’ Use the child’s favourite thing as an example to help them understand and relate to the situation. ‘That’s right; you would feel sad! Well done!]’</i> 	Transcendence, competence
22	<ul style="list-style-type: none"> In our story, we should also talk about what the character – (Rita), the girl – plans to do to fix the problem. <i>[Click for the Plan pictogram to show on the slide and draw the child’s attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> 	Intentionality, meaning, competence

	<ul style="list-style-type: none"> • We should talk about the plan to know what (Rita) is thinking of doing to fix the problem. • What does (Rita) plan to do to fix the plant pot? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Rita) plans to fix the plant pot using glue. Now you say it.]</i> 	
23	<ul style="list-style-type: none"> • At the end of the story, we say what the character –(Rita)– does to try to fix the problem. This is the attempt. <i>[Click for the Attempt pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • Talking about the attempt helps us to understand how (Rita) follows her plan to fix the problem. • What does (Rita) try to do to fix the plant pot? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Rita) tries to fix the plant pot with glue. Now you say it.]</i> 	Intentionality, meaning, competence
24	<ul style="list-style-type: none"> • We then have to say what happens when the character –(Rita)– tries to fix the problem. This is what we call the outcome of the story, where the problem gets fixed (or not!) <i>[Click for the Outcome pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • We must always include the outcome to understand how the story ends. • Is the problem in this story fixed: what happens when (Rita) uses glue to fix the plant pot? Does she manage to fix it? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Rita) fixes the plant pot with glue. Now you say it.]</i> 	Intentionality, meaning, competence
25	<ul style="list-style-type: none"> • Finally, we should talk about how the character –(Rita)– feels about the outcome. <i>[Click for the Feeling pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • How does (Rita) feel now that her plant pot is fixed? Tell me and pick how she feels on the wheel. <i>[Wait for response]</i> 	Intentionality, transcendence, competence

	<ul style="list-style-type: none"> ○ <i>[If the child responds correctly, say, 'That's right, (Rita) feels happy! Well done!' and move on to the next slide.]</i> ○ <i>[If the child responds incorrectly, say, '(Rita) is happy because she fixed the plant pot and her plant looks great, just like before it broke'. Once again, refer to the child's favourite thing. It's just like if... was fixed after it got broken. How would you feel if this happened? That's right; you would feel happy! Just like here (Rita) is happy because her plant is fixed.] Move on to the next slide when pertinent.]</i> 	
26	<ul style="list-style-type: none"> • Listen to (Rita) talk – which one sounds more like her? 1 or 2? <i>[Play the two audio clips of the sentence 'My plant looks great!' spoken with happy and sad prosody.]</i> ○ <i>[If the child selects the correct audio (1), say, 'Well done! (Rita) is happy and she sounds happy.', expressing the sentence with a happy-sounding voice yourself and encouraging the child to do the same by saying, 'Can you repeat? Make your voice sound happy.'].</i> ○ <i>[If the child selects the incorrect audio (2), play the audio clips again and explain why the happy-sounding audio is the correct one by saying, '(Rita) is happy, and she is smiling!' Smile and using a happy vocal expression say, 'Yay! I'm happy because my plant is fixed! Can you repeat? Make your voice sound happy.'].</i> • <i>[If you haven't said this in the previous slide, say, 'It's just like if... was fixed after it got broken. How would you feel if this happened? That's right; you would feel happy! Just like here (Rita) is happy because her plant is fixed. Well done!']</i> 	Transcendence, competence
27-28	<ul style="list-style-type: none"> • So, when telling stories, we should include all these parts that we have talked about. <i>[Point at the completed graphic organiser as you say this.]</i> • What would happen if we told a story that was missing some parts? For example, if your friend or teacher asks you to tell them a story, but you only say who the characters are and nothing else. You don't say where they are, what the problem is... Then your story wouldn't be... <i>[Wait for response; if necessary, say, 'finished! The story wouldn't be complete']</i> • It would be like doing a jigsaw puzzle, and you don't have all the pieces. We wouldn't be able to complete it. Look, let's try 	Transcendence, competence

	<p>it with this. <i>[Encourage the child to help you do the simple cat jigsaw puzzle but hide the piece corresponding to the cat's face; emphatically highlight that the piece is missing and we can't see the full picture of the cat. Then, add the missing piece and say, 'Just like we need all the pieces of the jigsaw for it to be complete, we need all the parts of a story for it to make sense'.]</i></p> <ul style="list-style-type: none"> • To help you remember all the parts the story should have for it to be complete, you can use these pictures we have been working with <i>[Point at the story grammar pictograms as you say this.]</i> Let's use them while we recap what we have learned. 	
6	<ul style="list-style-type: none"> • <i>[Go back to Story A's pictures, starting on slide 6.]</i> So, we have learned that at the beginning of the story, we talk about... <i>[Begin by sticking the Characters pictogram in its place on the empty story grammar graphic organiser and wait for response; if necessary, say, 'The characters!', then ask, 'What else?' and wait for response, encouraging the child to stick the consecutive corresponding pictograms on the graphic organiser.]</i> <ul style="list-style-type: none"> ○ The characters. These are the people or the animals in the story. ○ Where and when the story takes place. That is, place and time. • In this story: <i>[Point at Story A's picture 1 and wait for the child to tell you each story grammar unit; encourage/praise and recast/expand as required. Offer prompts with increasing support depending on the child's difficulty to respond.]</i> <ul style="list-style-type: none"> ○ Who are the characters? <i>[Wait for response. The next prompts in order of least to most amount of assistance would be:]</i> <ol style="list-style-type: none"> 1. What person and animal is the story about? <i>[Wait for response]</i> 2. What can we say about the girl and the cat? For example, their names, what they are doing, what they look like... <i>[Wait for response]</i> 3. A girl called (Rita) is watering her plant in the living room. Her cat (Gus) is also there. (Rita)'s hair is brown. (Gus)'s hair is orange. 	(Recap)

	<ul style="list-style-type: none"> ○ Where does the story happen? In which place? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. Where/In which place are the girl and the cat? <i>[Wait for response]</i> 2. Are they at home or in school? In the kitchen or the living room? <i>[Wait for response]</i> 3. (Rita) and (Gus) are in the living room at home. ○ When does the story take place?/At what time? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. Is it night time or daytime? <i>[wait for response]</i> 2. Is it night time with the moon and the stars in the sky, or is it daytime with the sun in the sky? <i>[Wait for response]</i> 3. It's night time because the moon and the stars are in the dark sky. 	
7-9	<ul style="list-style-type: none"> • Following this, in the middle of the story, we talk about... <i>[Wait for the child to respond and stick the Problem pictogram on the story grammar graphic organiser. If they don't respond within a few seconds, stick the pictogram yourself and say, 'The problem!', then ask, 'What else?' and wait for the child to respond, encouraging them to stick the consecutive corresponding pictograms on the graphic organiser.]</i> <ul style="list-style-type: none"> ○ The problem –what goes wrong–. This is something that happens that needs to be fixed or sorted out. ○ How the characters feel about the problem, and ○ What they plan to do to fix the problem. • In this story: <i>[Point at Story A's picture 2,3,4 and wait for the child to tell you each story grammar unit; encourage/praise and recast/expand as appropriate. Offer prompts with increasing support as required.]</i> <ul style="list-style-type: none"> ○ What is the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. What goes wrong that needs to be fixed? <i>[Wait for response]</i> 	

	<ol style="list-style-type: none"> 2. What does the cat do with the girl's plant? <i>[Wait for response]</i> 3. A girl called (Rita) is watering her plant in the living room. Her cat (Gus) jumps and knocks the plant pot over. The plant pot breaks. <ul style="list-style-type: none"> ○ How does (Rita) feel about the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. How does (Rita) feel when she sees that her plant pot is broken? <i>[Wait for response]</i> 2. Is she happy or sad? <i>[Wait for response]</i> 3. (Rita) feels sad because (Gus) broke her plant pot. ○ What is (Rita)'s plan to fix the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. What does (Rita) plan to do to fix the broken plant pot? <i>[Wait for response]</i> 2. Does she think of using tape or glue to fix it? <i>[Wait for response]</i> 3. (Rita) plans to fix the plant pot with glue. 	
10-11	<ul style="list-style-type: none"> • Finally, at the end of the story, we talk about: <i>[Wait for the child to respond and stick the Attempt pictogram on the story grammar graphic organiser. If they don't respond within a few seconds, stick the pictogram yourself and say, 'The attempt!', then ask, 'What else?' and wait for the child to respond, encouraging them to stick the consecutive corresponding pictograms on the graphic organiser.]</i> <ul style="list-style-type: none"> ○ The attempt. This is what the characters do to fix the problem. ○ The outcome. This is what happens when they try to fix it, and ○ How they feel about the outcome. • In this story: <i>[Point at Story A's picture 5,6 and wait for the child to tell you each story grammar unit; encourage/praise and recast/expand as appropriate. Offer prompts with increasing support as required.]</i> 	

	<ul style="list-style-type: none"> ○ What does the character in this story attempt to fix the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. What does (Rita) try to do to fix the plant pot? <i>[Wait for response]</i> 2. Does she use tape or glue to fix it? <i>[Wait for response]</i> 3. (Rita) tries to fix the plant pot with glue. ○ What is the outcome of the story? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. Is the problem in this story fixed: what happens when (Rita) uses glue to fix the plant pot? <i>[Wait for response]</i> 2. Does (Rita) manage to fix the plant pot with glue? <i>[Wait for response]</i> 3. (Rita) fixes the plant pot. ○ How does (Rita) feel about the outcome? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. How does (Rita) feel when she fixes her plant pot? <i>[Wait for response]</i> 2. Is she happy or sad? <i>[Wait for response]</i> 3. (Rita) feels happy that her plant looks great, just like before. 	
6-11	<ul style="list-style-type: none"> • <i>[Finish by encouraging the child to retell the story with you in its entirety. Provide the model below and point at the relevant story pictures and printed story grammar pictograms as necessary.]</i> So, in this story, a girl called (Rita) is watering a plant. Her cat (Gus) is also there. (Rita)'s hair is brown. (Gus)'s hair is orange. It's night time and they are in the living room at their home. The cat (Gus) jumps and knocks over the plant pot. The plant pot breaks, and this makes (Rita) feel sad. (Rita) then plans to fix the plant pot with glue so that the plant will be okay. (Rita) tries to fix the broken plant pot with glue. She manages to fix it and feels happy because her plant looks healthy and good again. Yay! 	(Model of the story)
29	<ul style="list-style-type: none"> • It's important to practice telling stories as often as possible at home and in school! 	Competence, transcendence

30	<ul style="list-style-type: none"> Also, always include all the different story parts when doing so. 	
31	<ul style="list-style-type: none"> Well done! You are on your way to becoming a storytelling superstar! You've earned your second star! <i>[Encourage the child to colour in the second star on the star points chart.]</i> 	Competence
32	<ul style="list-style-type: none"> <i>[5-minute break, e.g., little stretch and activity of choice]</i> 	N/A

Slide	Teaching phase script for <i>Story B</i> (Task Order BA)	MLE strategies
12	<ul style="list-style-type: none"> [Praise the child's performance when they finish telling <i>Story A</i> regardless of how they did.] Great! Well done! Now we are going to practice telling this story together. We are going to learn all the parts that a story should have. 	Competence, intentionality
13	<ul style="list-style-type: none"> Good stories have a beginning, a middle and an end. 	(Overview of the story parts)
14	<ul style="list-style-type: none"> [As you introduce each story grammar unit, stick the pictograms representing these on the graphic organiser.] At the beginning of the story, we talk about: <ul style="list-style-type: none"> The characters. These are the people or the animals in the story, and Where and when the story takes place; that is, place and time. 	
15	<ul style="list-style-type: none"> Next, in the middle of the story, we talk about: <ul style="list-style-type: none"> The problem. This is something that happens that needs to be fixed or sorted out. How the characters feel about the problem, and What they plan to do to fix the problem. 	
16	<ul style="list-style-type: none"> Finally, at the end of the story, we talk about: <ul style="list-style-type: none"> The attempt. This is what the characters do to fix the problem. The outcome. This is what happens when they try to fix it, and How they feel about the outcome. 	
6-11	<ul style="list-style-type: none"> [Provide the child with a model of the story before discussing each story grammar unit one by one. Point at the relevant story pictures and printed story grammar pictograms as you do so.] In the story you have just told, a man is eating a sandwich. There is also a bird. It's daytime and they are outside in a park. The bird flies by and steals the man's sandwich, and the man feels angry about it. He then plans to go to the shop to buy another sandwich to replace the one that the bird stole. The man goes to the shop to buy another sandwich. He manages to buy it and feels happy because he can now continue eating. 	(Model of the story)
13	<ul style="list-style-type: none"> It's important that we include all these parts so that the story is complete and people can understand what happens in it. 	Meaning, intentionality

	<ul style="list-style-type: none"> Now we are going to learn a little more about every part that needs to be in the story you have just told. 	
17	<ul style="list-style-type: none"> First, let's talk about the characters. These are the people and the animals in the story. <i>[Click for the Characters pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick the Characters pictogram on the graphic organiser; e.g., 'Put the Characters picture here to remember to include them in your story'. Do this successively for all story grammar units. When possible, provide encouraging feedback to acknowledge any story grammar units that the child already included in the pretest, e.g., 'I know you are good at this bit because you said it earlier when you were telling the story!']</i> Talking about the characters is important because it's their story that we are going to follow, so we will want to know about them. For example, we could say what their names are, what they are doing, and what they look like. Name: <i>[Refer to picture 1]</i> In this story: Does the man have a name? <i>[Wait for response]</i> What about the bird? <i>[Wait for response]</i> <i>[Use the names the child gives you from now on; if the child does not respond within a few seconds, follow up with a model sentence: 'You could say, the man is called (Ali), and the bird is called (Bob). Now you say it. ']</i> Activity: We also need to say what they are doing. What is (Ali) doing here? <i>[Let the child describe the activity; encourage/praise and recast/expand as required; if the child cannot tell the expected target within a few seconds, follow up with a model sentence: '(Ali) is eating a sandwich. Now you say it. ']</i> What they look like: We can also describe what they look like. What does (Bob) look like? <i>[Wait for response]</i> What about (Ali)? What does he look like? <i>[Wait for response]</i> <i>[Let the child describe each character, recasting/expanding and encouraging/praising as required (e.g., 'Good job! You have used lots of great words to describe the bird!'). If the child cannot tell the targets within a few seconds, follow up with the model sentences: 'The bird (Bob) is white with grey wings and tail.', 'The man (Ali) has short black hair and is wearing a blue top. Now you say it. ']</i> 	Intentionality, meaning, competence

18	<ul style="list-style-type: none"> • We should also talk about where and when the story happens. That is, place and time. <i>[Click for the Place and Time pictograms to show on the slide and draw the child's attention to it. Encourage the child to stick these pictograms on the graphic organiser.]</i> • Talking about the story's place and time helps us learn about where the characters are and at what time the story happens. • So, in this story: Where do you think they are? In which place? <i>[Wait for response]</i> What time do you think it is? Is it daytime or night time? <i>[Wait for response]</i> <p><i>[Encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentences: '(Ali) and (Bob) are outside in a park. Now you say it.', 'It's daytime because we can see the sky is light blue. Now you say it.]</i></p>	Intentionality, meaning, competence
19	<ul style="list-style-type: none"> • Then in the middle of the story, something happens that needs to be fixed or sorted out. Something goes wrong, and there is a problem. <i>[Click for the Problem pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • We must always say what the problem is to know what goes wrong in the story. • What is the problem in this story? What goes wrong? What happens to (Ali)'s sandwich? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Bob) the bird flies by and steals (Ali)'s sandwich. Now you say it.]</i> 	Intentionality, meaning, competence
20	<ul style="list-style-type: none"> • After we say what the problem is, we talk about how the character –(Ali), the man in this story– feels about it. <i>[Click for the Feeling pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • Talking about how (Ali) feels is important because it helps us to understand what feelings or emotions he is experiencing and why he wants to get another sandwich. • As we saw in the Feelings Game, to express how (Ali) feels, we should use special words like 'happy', 'sad', 'angry' or 'scared', and change our voice to make it sound 'happy', 'sad', 'angry' or 'scared'. To help us remember we have this wheel with pictures. <i>[Illustrate each emotion prosodically and</i> 	Intentionality, meaning, transcendence, competence

	<p><i>through exaggerated facial expressions. Point at the relevant pictograms on the Feelings wheel.]</i></p> <ul style="list-style-type: none"> How does (Ali) feel when he sees his sandwich is stolen? Tell me and pick how he feels on the wheel. <i>[Wait for response]</i> <ul style="list-style-type: none"> <i>[If the child responds correctly, say, ‘That’s right, (Ali) feels angry! Well done!’ and move on to the next slide.]</i> <i>[If the child responds incorrectly, say, ‘This sandwich was important to (Ali) because he was hungry. How would you feel if you were eating a sandwich (or some other food that you like) and an animal took it?’ Use this hypothetical scenario as an example to help them understand and relate to the situation. ‘That’s right; you would feel angry! (Ali) is angry because the bird stole his sandwich.’ Move on to the next slide when pertinent.]</i> 	
21	<ul style="list-style-type: none"> Listen to (Ali) talk – which one sounds more like him? 1 or 2? <i>[Play the two audio clips of the sentence ‘That bird took my sandwich!’ spoken with scared and angry prosody.]</i> <ul style="list-style-type: none"> <i>[If the child selects the correct audio (2), say, ‘Well done! (Ali) is angry and he sounds angry.’, expressing the sentence with an angry-sounding voice yourself and encouraging the child to do the same by saying, ‘Can you repeat? Make your voice sound angry.’]</i> <i>[If the child selects the incorrect audio (1), play the audio clips again and explain why the angry-sounding audio is the correct one by saying, ‘(Ali) is angry, and he is shouting!’ Using an angry vocal expression, say, ‘Oh no! I’m angry because that bird took my sandwich! Can you repeat? Make your voice sound angry.’]</i> <i>[If you haven’t said this in the previous slide, say, ‘This sandwich was important to (Ali) because he was hungry. How would you feel if you were eating a sandwich (or some other food that you like) and an animal took it?’ Use this hypothetical scenario as an example to help them understand and relate to the situation. ‘That’s right; you would feel angry! Well done!]’</i> 	Transcendence, competence
22	<ul style="list-style-type: none"> In our story, we should also talk about what the character – (Ali), the man – plans to do to fix the problem. <i>[Click for the Plan pictogram to show on the slide and draw the child’s</i> 	Intentionality, meaning, competence

	<p><i>attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i></p> <ul style="list-style-type: none"> • We should talk about the plan to know what (Ali) is thinking of doing to fix the problem. • What does (Ali) plan to do to get another sandwich? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Ali) plans to go to the shop to buy another sandwich. Now you say it.]</i> 	
23	<ul style="list-style-type: none"> • At the end of the story, we say what the character –(Ali)– does to try to fix the problem. This is the attempt. <i>[Click for the Attempt pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • Talking about the attempt helps us to understand how (Ali) follows his plan to fix the problem. • What does (Ali) try to do to replace the sandwich the bird stole? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Ali) goes to the shop to try to buy another sandwich. Now you say it.]</i> 	Intentionality, meaning, competence
24	<ul style="list-style-type: none"> • We then have to say what happens when the character –(Ali)– tries to fix the problem. This is what we call the outcome of the story, where the problem gets fixed (or not!) <i>[Click for the Outcome pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> • We must always include the outcome to understand how the story ends. • Is the problem in this story fixed: what happens when (Ali) goes to the shop to buy another sandwich? Does he manage to buy another sandwich? <i>[Wait for response; encourage/praise and recast/expand as required. If the child cannot tell the target within a few seconds, follow up with the model sentence: '(Ali) goes to the shop to buy another sandwich. Now you say it.]</i> 	Intentionality, meaning, competence
25	<ul style="list-style-type: none"> • Finally, we should talk about how the character –(Ali)– feels about the outcome. <i>[Click for the Feeling pictogram to show on the slide and draw the child's attention to it. Encourage the child to stick this pictogram on the graphic organiser.]</i> 	Intentionality, transcendence, competence

	<ul style="list-style-type: none"> How does (Ali) feel now that he has a new sandwich? Tell me and pick how he feels on the wheel. <i>[Wait for response]</i> <ul style="list-style-type: none"> <i>[If the child responds correctly, say, 'That's right, (Ali) feels happy! Well done!' and move on to the next slide.]</i> <i>[If the child responds incorrectly, say, '(Ali) is happy because he has bought another sandwich to replace the one the bird stole'. Once again, refer to the hypothetical scenario. It's just like if you got another sandwich (or food) after yours got stolen/taken/you lost yours. How would you feel if this happened? That's right; you would feel happy! Just like here, (Ali) is happy because he has a new sandwich.' Move on to the next slide when pertinent.]</i> 	
26	<ul style="list-style-type: none"> Listen to (Ali) talk – which one sounds more like him? 1 or 2? <i>[Play the two audio clips of the sentence 'This sandwich is delicious!' spoken with happy and sad prosody.]</i> <ul style="list-style-type: none"> <i>[If the child selects the correct audio (1), say, 'Well done! (Ali) is happy and he sounds happy.', expressing the sentence with a happy-sounding voice yourself and encouraging the child to do the same by saying, 'Can you repeat? Make your voice sound happy.']}</i> <i>[If the child selects the incorrect audio (2), play the audio clips again and explain why the happy-sounding audio is the correct one by saying, '(Ali) is happy, and he is smiling!' Smile and using a happy vocal expression say, 'Yay! I'm happy because I have a new sandwich! Can you repeat? Make your voice sound happy.']}</i> <i>[If you haven't said this in the previous slide, say, 'It's just like if you got another sandwich (or food) after yours got stolen/taken/you lost yours. How would you feel if this happened? That's right; you would feel happy! Just like here, (Ali) is happy because he has a new sandwich. Well done!']</i> 	Transcendence, competence
27-28	<ul style="list-style-type: none"> So, when telling stories, we should include all these parts that we have talked about. <i>[Point at the completed graphic organiser as you say this.]</i> What would happen if we told a story that was missing some parts? For example, if your friend or teacher asks you to tell them a story, but you only say who the characters are and nothing else. You don't say where they are, what the problem 	Transcendence, competence

	<p>is... Then your story wouldn't be... <i>[Wait for response; if necessary, say, 'finished! The story wouldn't be complete']</i></p> <ul style="list-style-type: none"> • It would be like doing a jigsaw puzzle, and you don't have all the pieces. We wouldn't be able to complete it. Look, let's try it with this. <i>[Encourage the child to help you do the simple bird jigsaw puzzle but hide the piece corresponding to the bird's face; emphatically highlight that the piece is missing and we can't see the full picture of the bird. Then, add the missing piece and say, 'Just like we need all the pieces of the jigsaw for it to be complete, we need all the parts of a story for it to make sense']</i> • To help you remember all the parts the story should have for it to be complete, you can use these pictures we have been working with <i>[Point at the story grammar pictograms as you say this]</i>. Let's use them while we recap what we have learned. 	
6	<p>• <i>[Go back to Story B's pictures, starting on slide 6.]</i> So, we have learned that at the beginning of the story, we talk about... <i>[Begin by sticking the Characters pictogram in its place on the empty story grammar graphic organiser and wait for response; if necessary, say, 'The characters!', then ask, 'What else?' and wait for response, encouraging the child to stick the consecutive corresponding pictograms on the graphic organiser.]</i></p> <ul style="list-style-type: none"> ○ The characters. These are the people or the animals in the story. ○ Where and when the story takes place. That is, place and time. <p>• In this story: <i>[Point at Story B's picture 1 and wait for the child to tell you each story grammar unit; encourage/praise and recast/expand as required. Offer prompts with increasing support depending on the child's difficulty to respond.]</i></p> <ul style="list-style-type: none"> ○ Who are the characters? <i>[Wait for response. The next prompts in order of least to most amount of assistance would be:]</i> <ol style="list-style-type: none"> 1. What person and animal is the story about? <i>[Wait for response]</i> 2. What can we say about the man and the bird? For example, their names, what they are doing, what they look like... <i>[Wait for response]</i> 	(Recap)

	<p>3. A man called (Ali) is eating a sandwich in the park. A bird (Bob) is also there. (Ali)'s hair is black. (Bob)'s feathers are white and grey.</p> <ul style="list-style-type: none"> ○ Where does the story happen? In which place? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. Where/In which place are the man and the bird? <i>[Wait for response]</i> 2. Are they at home or outside? In the supermarket or the park? <i>[Wait for response]</i> 3. (Ali) and (Bob) are outside in a park. ○ When does the story take place?/At what time? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. Is it night time or daytime? <i>[wait for response]</i> 2. Is it night time with the moon and the stars in the sky, or is it daytime with light blue sky? <i>[Wait for response]</i> 3. It's day-time because the sun is shining and the sky is light blue.
7-9	<ul style="list-style-type: none"> • Following this, in the middle of the story, we talk about... <i>[Wait for the child to respond and stick the Problem pictogram on the story grammar graphic organiser. If they don't respond within a few seconds, stick the pictogram yourself and say, 'The problem!', then ask, 'What else?' and wait for the child to respond, encouraging them to stick the consecutive corresponding pictograms on the graphic organiser.]</i> <ul style="list-style-type: none"> ○ The problem –what goes wrong–. This is something that happens that needs to be fixed or sorted out. ○ How the characters feel about the problem, and ○ What they plan to do to fix the problem. • In this story: <i>[Point at Story B's picture 2,3,4 and wait for the child to tell you each story grammar unit; encourage/praise and recast/expand as appropriate. Offer prompts with increasing support as required.]</i> <ul style="list-style-type: none"> ○ What is the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i>

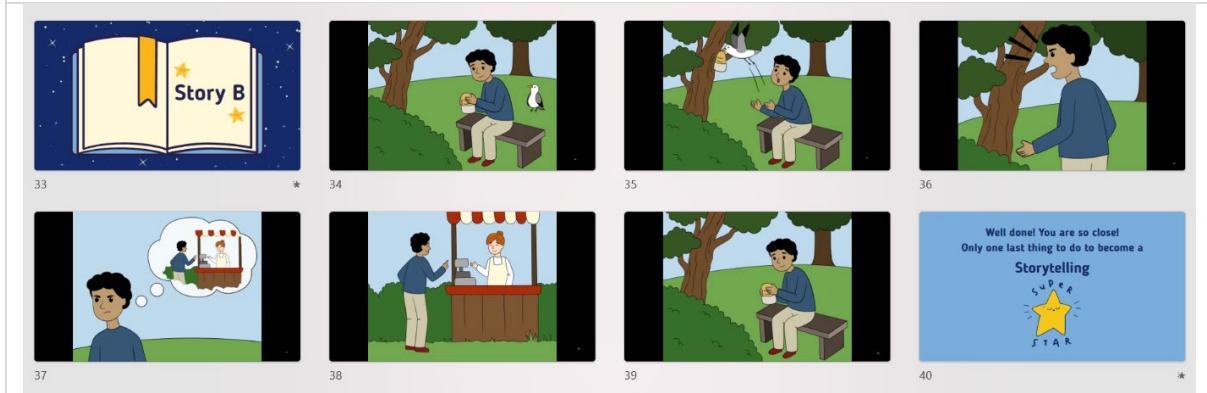
	<p>4. What goes wrong that needs to be fixed? <i>[Wait for response]</i></p> <p>5. What does the bird do with the man's sandwich? <i>[Wait for response]</i></p> <p>6. A man called (Ali) is eating a sandwich in a park. A bird (Bob) flies by and steals his sandwich.</p> <ul style="list-style-type: none"> ○ How does (Ali) feel about the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 4. How does (Ali) feel when he sees that his sandwich is stolen/taken? <i>[Wait for response]</i> 5. Is he angry or scared? <i>[Wait for response]</i> 6. (Ali) feels angry because (Bob) stole/took his sandwich. ○ What is (Ali)'s plan to fix the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. What does (Ali) plan to do to replace the sandwich? <i>[Wait for response]</i> 2. Does he think of going home to make a new sandwich or going to the shop to buy a new sandwich? <i>[Wait for response]</i> 3. (Ali) plans to go to the shop to buy a new sandwich. 	
10-11	<ul style="list-style-type: none"> • Finally, at the end of the story, we talk about: <i>[Wait for the child to respond and stick the Attempt pictogram on the story grammar graphic organiser. If they don't respond within a few seconds, stick the pictogram yourself and say, 'The attempt!', then ask, 'What else?' and wait for the child to respond, encouraging them to stick the consecutive corresponding pictograms on the graphic organiser.]</i> <ul style="list-style-type: none"> ○ The attempt. This is what the characters do to fix the problem. ○ The outcome. This is what happens when they try to fix it, and ○ How they feel about the outcome. • In this story: <i>[Point at Story B's picture 5,6 and wait for the child to tell you each story grammar unit; encourage/praise</i> 	

	<p><i>and recast/expand as appropriate. Offer prompts with increasing support as required.]</i></p> <ul style="list-style-type: none"> ○ What does the character in this story attempt to fix the problem? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. What does (Ali) try to do to replace the sandwich? <i>[Wait for response]</i> 2. Does he go home to make a new sandwich or go to the shop to buy a new sandwich? <i>[Wait for response]</i> 3. (Ali) goes to the shop to try to buy a new sandwich. ○ What is the outcome of the story? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. Is the problem in this story fixed: what happens when (Ali) goes to the shop to buy a new sandwich? <i>[Wait for response]</i> 2. Does (Ali) manage to buy a new sandwich? <i>[Wait for response]</i> 3. (Ali) buys a new sandwich. ○ How does (Ali) feel about the outcome? <i>[Wait for response. Next prompts in order of least to most amount of assistance:]</i> <ol style="list-style-type: none"> 1. How does (Ali) feel when he buys a new sandwich? <i>[Wait for response]</i> 2. Is he happy or sad? <i>[Wait for response]</i> 3. (Ali) feels happy that he has a new sandwich. 	
6-11	<ul style="list-style-type: none"> • <i>[Finish by encouraging the child to retell the story with you in its entirety. Provide the model below and point at the relevant story pictures and printed story grammar pictograms as necessary.]</i> So, in this story, a man called (Ali) is eating a sandwich. A bird (Bob) is also there. (Ali)'s hair is black. (Bob)'s feathers are white & grey. It's daytime, and they are outside in a park. The bird (Bob) flies by, and steals (Ali)'s sandwich, and this makes (Ali) feel angry. (Ali) then plans to go to the shop to buy another sandwich to replace the one the bird stole. (Ali) goes to the shop to try to buy another sandwich. He manages to buy it and feels happy because he can now continue eating. Yay! 	(Model of the story)

29	<ul style="list-style-type: none"> It's important to practice telling stories as often as possible at home and in school! 	Competence, transcendence
30	<ul style="list-style-type: none"> Also, always include all the different story parts when doing so. 	
31	<ul style="list-style-type: none"> Well done! You are on your way to becoming a storytelling superstar! You've earned your second star! <i>[Encourage the child to colour in the second star on the star points chart.]</i> 	Competence
32	<ul style="list-style-type: none"> <i>[5-minute break, e.g., little stretch and activity of choice]</i> 	N/A

16C. Post-test Phase

To clarify the context of the script that follows, an overview of the slides used during the pretest phase of the storytelling activity is provided below. This overview pertains to children who followed the AB sequence, experiencing *Story A* as the pretest and *Story B* as the post-test. For children in the BA sequence, the slides were identical, except that *Story B* was used in the pretest and *Story A* in the post-test.



Slide	Post-test phase script for stories A/B	MLE strategies
33	<ul style="list-style-type: none"> Let's tell one last story. You'll tell me the story that you see in the pictures, and I'll listen quietly. First, let's have a look at all the pictures here. <i>[Wait for the child to look at the pictures; ask them to indicate when they want to see the next picture.]</i> <i>[Once you reach Story A/B's picture 6 say:]</i> Okay, let's go back to the start. <i>[and then say:]</i> Now look at the pictures again and tell me the story that you see. I will be quiet and listen to your story. 	N/A
34-39	<ul style="list-style-type: none"> Tell me the story that you see, starting with this picture. <i>[Wait for the child to tell the story; prompt them to continue if they fall silent. Do not give feedback based on the child's performance at this stage.]</i> 	
40	<ul style="list-style-type: none"> Well done! You are so close! There's just one last thing to do. <i>[Encourage the child to colour in the third star on the star points chart.]</i> 	

Appendix 17. Version 1 of Scoring Sheet 1: Story Grammar Units for Stories A, B and C

Story A: Scoring Sheet 1. Story Grammar Units (Version 1)			
Unit	Response	Score	Notes
1. Character 1	Girl/Woman (or synonym) Any proper name given <i>Also acceptable:</i> Sister (or other family member name) <i>Not acceptable:</i> Pronoun		
2. Character 2	Cat/Kitten (or synonym) Any proper name given <i>Not acceptable:</i> Pronoun		
3. Time	Night time/Evening/Dark (or synonym) <i>Also acceptable:</i> Bedtime/After dinner		
4. Place	Living room (or synonym) <i>Also acceptable:</i> Home/House (or synonym)		
5. Problem	Cat breaks/knocks over the plant/plant pot (or synonyms) The plant/plant pot breaks/falls (or synonyms)		
6. Emotion 1	Sad (or synonym) <i>Also acceptable:</i> Crying (or synonym)		
7. Plan	The girl/she plans* to fix/glue the plant/plant pot (or synonyms) *Examples of accepted goal-orientated language: decides/wants to/thinks she will		
8. Attempt	The girl/she tries to* fix/glue the plant/plant pot with glue (or synonyms) *Examples of accepted action-orientated language: goes to/is going to/is gluing/glues		
9. Outcome	The plant/plant pot is fixed/glued/fine (or synonyms) The girl/she fixed/has finished fixing the plant/plant pot (or synonyms)		
10. Emotion 2	Happy (or synonym) <i>Also acceptable:</i> Smiling (or synonym)		
Story grammar units score (sum of 1 to 10)			/10

Story B: Scoring Sheet 1. Story Grammar Units (Version 1)

Unit	Response	Score	Notes
1. Character 1	Boy/Man (or synonym) Any proper name given <i>Also acceptable:</i> Dad (or other family member name) <i>Not acceptable:</i> Pronoun		
2. Character 2	Bird/Seagull (or synonym) Any proper name given <i>Not acceptable:</i> Pronoun		
3. Time	Day/Day-time/Morning/Afternoon (or synonym) <i>Also acceptable:</i> Lunch time		
4. Place	Park/Garden (or synonym) <i>Also acceptable:</i> Outdoors/Outside (or synonym)		
5. Problem	Bird steals/takes the sandwich (or synonyms) The sandwich gets taken/stolen (or synonyms)		
6. Emotion 1	Angry (or synonym) <i>Also acceptable:</i> Shouting (or synonym)		
7. Plan	The man/he plans* to buy another sandwich (or synonyms) <i>*Examples of accepted goal-orientated language:</i> decides/wants to/thinks he will		
8. Attempt	The man/he goes to* the shop to buy another sandwich (or synonyms) <i>*Examples of accepted action-orientated language:</i> is going to/tries to/is buying/buys		
9. Outcome	The man/he bought/has another/a new sandwich (or synonyms) The man/he eats the new sandwich (or synonyms)		
10. Emotion 2	Happy (or synonym) <i>Also acceptable:</i> Smiling (or synonym)		
Story grammar units score (sum of 1 to 10)			/10

Story C: Scoring Sheet 1. Story Grammar Units (Version 1)

Unit	Response	Score	Notes
1. Character 1	Elephant (or synonym) Any proper name given <i>Not acceptable:</i> Pronoun		
2. Character 2	Elephant (or synonym) Any proper name given <i>Not acceptable:</i> Pronoun		
3. Time	Day time/Morning/Afternoon (or synonym) <i>Also acceptable:</i> Lunch time		
4. Place	Savanna (or synonym) <i>Also acceptable:</i> Outdoors/Outside (or synonym)		
5. Problem	The elephant gets lost and can't find her way home (or synonyms)		
6. Emotion 1	Scared (or synonym) <i>Also acceptable:</i> Shaky/Crying (or synonyms)		
7. Plan	The elephant/she plans* to ask the older elephant for directions to get back home (or synonyms) *Examples of accepted goal-orientated language: decides/wants to/thinks she will		
8. Attempt	The elephant/she asks* the older elephant for directions to get back home (or synonyms) *Examples of accepted action-orientated language: tries to/goes to/is going to/is asking		
9. Outcome	The elephant/she arrives home safely (or synonyms) The older elephant tells the elephant/her how to get home, and she arrives home safely (or synonyms)		
10. Emotion 2	Happy (or synonym) <i>Also acceptable:</i> Smiling (or synonym)		
Story grammar units score (sum of 1 to 10)			/10

Appendix 18. Version 2 of Scoring Sheet 1: Story Grammar Units for Stories A and B

Story A: Scoring Sheet 1. Story Grammar Units (Version 2)								
Unit	Response	Scoring 1		Scoring 2		Total points	Notes	
Character 1	Girl/Woman (or synonym) Any proper name given <i>Also acceptable:</i> Sister (or other family member name) <i>Partial credit (1 point max for each system):</i> Exclusive use of pronouns to refer to character	0	1	2	0	1	2	/4
Character 2	Cat/Kitten (or synonym) Any proper name given <i>Partial credit (1 point max for each system):</i> Exclusive use of pronouns to refer to character	0	1	2	0	1	2	/4
Time	Night/Night-time/Evening/Dark (or synonym) <i>Also acceptable:</i> Bedtime/After dinner <i>Partial credit (1 point max for each system):</i> Descriptions of the story setting that suggest the time implicitly, e.g., ‘The moon is shining.’	0	1	2	0	1	2	/4
Place	Living/sitting room (or synonym) <i>Also acceptable:</i> Home/House (or synonym) <i>Partial credit (1 point max for each system):</i> Details about the characters’ location that suggest the surrounding place implicitly, e.g., ‘She is next to the sofa.’	0	1	2	0	1	2	/4
Problem	Cat breaks/knocks over the plant/plant pot (or synonyms) The plant/plant pot breaks/falls (or synonyms)	0	1	2	0	1	2	/4
Emotion 1	Sad (or synonym) <i>Also acceptable:</i> Crying/Surprised/Angry (or synonyms)	0	1	2	0	1	2	/4
Plan	The girl/she plans* to fix/glue the plant/plant pot (or synonyms) *Examples of accepted goal-orientated language: decides/wants to/thinks she will	0	1	2	0	1	2	/4

Unit	Response	Scoring 1	Scoring 2	Total points	Notes
Attempt	The girl/she tries to* fix/glue the plant/plant pot with glue (or synonyms) *Examples of accepted action-orientated language: goes to/is going to/is gluing/glues	0 1 2	0 1 2	/4	
Outcome	The plant/plant pot is fixed/glued/fine (or synonyms) The girl/she fixed/has finished fixing the plant/plant pot (or synonyms)	0 1 2	0 1 2	/4	
Emotion 2	Happy (or synonym) <i>Also acceptable:</i> Smiling (or synonym)	0 1 2	0 1 2	/4	
Story grammar units score		/20	/20	/40	

Story B: Scoring Sheet 1. Story Grammar Units (Version 2)						
Unit	Response	Scoring 1	Scoring 2	Total points	Notes	
Character 1	Boy/Man (or synonym) Any proper name given <i>Also acceptable:</i> Dad (or other family member name) <i>Partial credit (1 point max for each system):</i> Exclusive use of pronouns to refer to character	0 1 2	0 1 2	/4		
Character 2	Bird/Seagull (or synonym) Any proper name given <i>Partial credit (1 point max for each system):</i> Exclusive use of pronouns to refer to character	0 1 2	0 1 2	/4		
Time	Day/Day-time/Morning/Afternoon (or synonym) <i>Also acceptable:</i> Lunch time <i>Partial credit (1 point max for each system):</i> Descriptions of the story setting that suggest the time implicitly, e.g., 'The sun is shining.'	0 1 2	0 1 2	/4		
Place	Park/Garden (or synonym) <i>Also acceptable:</i> Outdoors/Outside (or synonym) <i>Partial credit (1 point max for each system):</i> Details about the characters' location that suggest the surrounding place implicitly, e.g., 'He is sitting on a bench.'	0 1 2	0 1 2	/4		
Problem	Bird steals/takes the sandwich (or synonyms) The sandwich gets taken/stolen (or synonyms)	0 1 2	0 1 2	/4		
Emotion 1	Angry (or synonym) <i>Also acceptable:</i> Shouting/Surprised/Sad (or synonyms)	0 1 2	0 1 2	/4		
Plan	The man/he plans* to buy another sandwich (or synonyms) <i>*Examples of accepted goal-orientated language: decides/wants to/thinks he will</i>	0 1 2	0 1 2	/4		

Unit	Response	Scoring 1	Scoring 2	Total points	Notes
Attempt	The man/he goes to* the shop to buy another sandwich (or synonyms) *Examples of accepted action-orientated language: is going to/tries to/is buying/buys	0 1 2	0 1 2	/4	
Outcome	The man/he bought/has another/a new sandwich (or synonyms) The man/he eats the new sandwich (or synonyms)	0 1 2	0 1 2	/4	
Emotion 2	Happy (or synonym) <i>Also acceptable: Smiling (or synonym)</i>	0 1 2	0 1 2	/4	
Story grammar units score		/20	/20	/40	

Appendix 19. Version 1 of Scoring Sheet 3: Emotional Vocabulary for Stories A, B and C

Story A: Scoring Sheet 3. Emotional Vocabulary (Version 1)			
Corresponding story grammar unit	Response	Score	Notes
Emotion 1	Sad (or synonym)		
Emotion 2	Happy (or synonym)		
Emotional vocabulary score (sum of 1 to 2)		/2	

Story B: Scoring Sheet 3. Emotional Vocabulary (Version 1)			
Corresponding story grammar unit	Response	Score	Notes
Emotion 1	Angry (or synonym)		
Emotion 2	Happy (or synonym)		
Emotional vocabulary score (sum of 1 to 2)		/2	

Story C: Scoring Sheet 3. Emotional Vocabulary (Version 1)			
Corresponding story grammar unit	Response	Score	Notes
Emotion 1	Scared (or synonym)		
Emotion 2	Happy (or synonym)		
Emotional vocabulary score (sum of 1 to 2)		/2	

Appendix 20. COVID-19 Risk Assessment for Working in Schools

The risk assessment provided below was subject to monthly reviews. The version presented here is the final and most recently updated.

School of Education, Communication and Language Sciences
King George VI Building
Queen Victoria Road
Newcastle upon Tyne
NE1 7RU
United Kingdom



Risk Assessment for Working in Schools

Project title	Creating a story-based dynamic assessment to detect developmental language disorder in children learning English as an additional language
Description of project activity	<p>During the course of this project, the researcher (Teresa Garrido-Tamayo) will be working in primary schools directly with children who have English as an additional language. It is expected that she will meet each child twice for two individual 45-minute sessions. The project aims to evaluate the effectiveness of an assessment resource designed to help identify developmental language disorder in children with English as an additional language.</p> <p>The risk assessment below outlines the project approach to managing risk when working in schools. The aim is to keep school staff, pupils, and the researcher safe and to avoid spreading the COVID-19 virus. Newcastle University's guidance has been followed in developing this document, which states that research should be undertaken following the latest government guidance applicable in the locality where the research is conducted (in this case, schools in the Newcastle upon Tyne area). Relevant Government advice can be found in the following links:</p> <ul style="list-style-type: none">- https://www.gov.uk/coronavirus- https://www.gov.uk/guidance/covid-19-coronavirus-restrictions-what-you-can-and-cannot-do- https://www.gov.uk/government/publications/coronavirus-covid-19-local-restrictions-in-education-and-childcare-settings- https://www.gov.uk/government/publications/actions-for-schools-during-the-coronavirus-outbreak/schools-covid-19-operational-guidance <p>The project has been approved by the Faculty of Humanities and Social Sciences Research Ethics Committee at Newcastle University.</p>
Assessor	Teresa Garrido-Tamayo
Manager/Responsible person	Drs Carolyn Letts and Laurence White
Unit name	Speech and Language Sciences Department, School of Education, Communication and Language Sciences, Newcastle University
Date of assessment	17/1/23
Review date	Monthly
Location of project activity	Schools in the Newcastle upon Tyne area

Details of hazards, risks and implemented controls	
Hazard 1	Each school may have differing policies and protocols
Risks	Each school would be expected to adhere at a minimum to national (Gov.uk) guidance. All schools would be expected to be doing all they can to keep children and staff safe within their own circumstances.
Controls	<ul style="list-style-type: none"> Schools will be contacted before the researcher's first visit to discuss this risk assessment and each school's own risk assessment. The researcher will request a copy of the school's risk assessment and visitor protocol. She will ask schools to inform her of changes to their risk assessments, visitor protocols and other relevant guidance. It is expected that most schools' risk assessments will be similar and based on Local Authority and national guidance. However, some schools may have specific needs. The researcher will ensure she is aware of schools' individual requirements. National guidance for schools will be the minimum standards the researcher will adhere to. She will follow the procedures in place within each school unless they fall short of national guidance, in which case this would be escalated to her project supervisors. The researcher will monitor University, national (Gov.uk), regional (Local Authority) and local (school) guidance regularly and will update this risk assessment as necessary. If significant changes are made, schools will be informed. Information about the risk assessment will be provided in the information sheets shared with parents and school staff. Consent forms will verify whether it has been read.
Hazard 2	Causing disruptions to the school day
Risks	Schools need to ensure that their plans are not disrupted to guarantee staff and children are as safe as possible.
Controls	<ul style="list-style-type: none"> The researcher will consult visitor protocols before entering a school. The researcher will consult with schools about any changes to the school day (e.g., staggered drop off and collection times, lunch and break times). All visits will be arranged in advance, and the researcher will only visit a school when it is essential. Visits will not last longer than necessary.
Hazard 3	Social distancing
Risks	The researcher not maintaining the appropriate distance from staff and children resulting in transmission of the COVID-19 virus.
Controls	<ul style="list-style-type: none"> The researcher will comply with appropriate social distancing requirements when navigating the school before, during and after the sessions. The researcher will comply with the school's guidance on distancing. If the school is working in bubbles, the researcher will ask key school contacts whether working with children from different bubbles is possible. If this were to be approved, she would pay particular attention to infection prevention/control measures.

Hazard 4	Collecting and returning child to class
Risks	Child is not monitored while moving around the school.
Controls	<ul style="list-style-type: none"> The researcher will ensure that the child she is working with on each occasion is not left unattended at any time. The researcher will follow school advice to ensure she minimises the number of people and rooms she has contact with.
Hazard 5	Cleaning and hygiene
Risks	Poor cleaning and hygiene practices resulting in transmission of the COVID-19 virus.
Controls	<ul style="list-style-type: none"> Government guidance states that face coverings are no longer required for staff and visitors in classrooms or communal areas in primary schools. PPE is no longer required to be worn either in response to COVID-19 beyond what staff would normally need for their work. Before her first visit, the researcher will ask schools about their individual requirements and wear face masks and PPE if necessary. A leaflet will be shared that can be shown to children with information on what to expect during the sessions and details about PPE/face coverings. The researcher will talk to schools about an appropriate space for conducting the sessions. This should be well ventilated. Regarding hand washing and general cleanliness, the researcher will: <ul style="list-style-type: none"> Always carry hand sanitiser and use it as appropriate. Hand sanitiser will always be used before entering and after leaving the school building. Always carry face masks and PPE (gloves, apron, etc.) for use as outlined above. Always carry anti-bacterial wipes and use them to wipe down resources and surfaces as necessary. Anti-bacterial wipes, tissues, and hand sanitiser will be used during the sessions to help children if coughing/sneezing and clean toys/wash hands as appropriate. Support children to wash their hands or use hand sanitiser before and after each session. School staff will be consulted to identify any child for whom medical or skin conditions/allergies would affect the hygiene policy and require extra care to be taken with the type of disinfectant products used. Wash her hands with soap or hand sanitiser before and after each session. Ensure the area she is assigned to is clean before and after use. Clean her laptop computer screen before and after each session and discourage children from touching the screen during the sessions.
Hazard 6	Managing resources
Risks	Ensure resources are clean to avoid transmission of the COVID-19 virus.
Controls	<ul style="list-style-type: none"> The activities will be conducted using the researcher's laptop computer. The child will not be required to touch the screen. The screen will be cleaned between sessions.

	<ul style="list-style-type: none"> The activities will involve the use of toys and laminated resources such as pictures. Most toys will be made of non-porous materials. However, the use of soft toys in one activity will be expected. In this case, schools will be consulted on their approaches to using soft toys, and these will be replaced with alternatives if needed. All toys and resources will be wiped with a disposable disinfectant wipe before and after use by any child. When appropriate, toys will be washed in warm water with a suitable detergent. Damaged and/or heavily contaminated toys and resources will be discarded and replaced.
Hazard 7	Child becomes unwell
Risks	<p>Child shows COVID-19 symptoms whilst in school (cough, high temperature, loss or change to sense of smell or taste, etc.).</p> <p>Child requires first aid or is unwell but not showing COVID-19 symptoms.</p>
Controls	<ul style="list-style-type: none"> The researcher will be aware of the relevant protocol in each school. Any unwell child will be taken to an identified school staff member who will ensure they are isolated. Depending on symptoms, the researcher will use PPE. Schools will be asked to notify the researcher if they have confirmed cases of COVID-19 among children or adults she is working with so that she can act accordingly.
Hazard 8	Researcher becomes unwell
Risks	Researcher shows COVID-19 symptoms whilst in school (cough, high temperature, loss or change to sense of smell or taste, etc.).
Controls	<ul style="list-style-type: none"> The researcher is fully vaccinated. She will monitor her health and take preventative measures. The researcher will follow the University and NHS processes relating to reporting and testing (https://www.ncl.ac.uk/wellbeing/). The researcher will not attend any school if she has COVID-19 symptoms. In this case, she will inform schools as soon as possible and cancel any arranged visits. If the LFD test is positive, she will self-isolate for 5 days or more if necessary until she tests negative and is asymptomatic. She will only return to schools once she is asymptomatic and has tested negative. During her visits to the schools, the researcher will carry identification and keep records of which staff members and children she is in close contact with. If she develops symptoms while in a school, she will inform key school contacts and notify them where she has been and which surfaces she has touched. She will leave the school and arrange an LFD test as soon as possible. If she is confirmed to have COVID-19, she will provide schools with details of anyone she has been in close contact with. If the person the researcher lives with has symptoms of COVID-19 or has tested positive for COVID-19, she will take a daily LFD test before leaving home and, if the result is positive, will immediately self-isolate and follow the guidelines above for being positive.

Hazard 9	If child needs the toilet during a session
Risks	Child needs to move around the school during the session.
Controls	<ul style="list-style-type: none"> The researcher will check the school's relevant protocol. Ideally, children should visit the toilet beforehand and be ready to attend the session.
Hazard 10	Emotional distress
Risks	Child is distressed during the session.
Controls	<ul style="list-style-type: none"> The researcher has experience working with children, and her DBS clearance will be made available to schools on request. The researcher can recognise distress signs and has strategies for reassuring children. Warm-up time will be provided at the beginning of the sessions to build rapport with the child. A relaxed and safe atmosphere will be created during the session so that children feel comfortable and confident. If possible, the researcher will spend additional warm-up time with the children before the session. If the child does not wish to participate in the activities at any point, the session will stop, and they will be immediately returned to their teacher.
Hazard 11	Researcher wellbeing
Risks	The researcher is stressed because of work pressures and responsibilities.
Controls	<ul style="list-style-type: none"> The researcher will manage her own timetabling and ensure time for breaks, cleaning, and meals. She will have regular opportunities to reflect and discuss with her project supervisors. The researcher has read and understood the national and local guidance. She has developed this risk assessment and can identify risks and control measures. National, regional, and local guidance will continue to be monitored throughout the project.
Hazard 12	Travel
Risks	The researcher is exposed to the COVID-19 virus while travelling to and between schools.
Controls	<ul style="list-style-type: none"> The researcher will follow national guidelines for travelling. Where possible, travel will be done via walking or by private transport. If this is not possible, taxis will be used instead of public transport to avoid excessive contact with members of the public.

Appendix 21. Information Sheet for Schools Used in the Main Study

Information for Schools

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

I am a PhD student based in the School of Education, Communication and Language Sciences at Newcastle University, supervised by Dr Carolyn Letts and Dr Laurence White. I am writing to request permission to carry out a study with 5- to 7-year-old children at your school who have English as an Additional Language (EAL).

This will involve me visiting your school at mutually convenient times to carry out activities with children identified as having EAL and whose parents consent to their participation. The study will also require these children's parents and teachers to share information on their language development and use as explained below.

Why am I carrying out this study?

Developmental Language Disorder (DLD) is identified when a child has significant, persistent difficulties producing and/or understanding language. DLD affects 2 children in the average classroom regardless of whether they grow up monolingually or bilingually, and can significantly impact most aspects of a child's life, including their learning and school achievement. Detecting DLD in children with EAL is currently challenging because there is a severe lack of tools that we can use with children growing up in environments where more than one language is spoken.

To address this, I am devising a dynamic assessment in English that aims to distinguish children with EAL who are struggling with language generally and who may be at risk for DLD, from their typically developing peers with EAL. This will be achieved by incorporating an element of teaching into the assessment through the use of stories, allowing for evaluation of the child's potential to learn language, rather than only their current language abilities.

By testing whether this dynamic assessment accurately identifies possible DLD in children with EAL, this study will contribute to the development of more efficient and reliable assessment resources that recognise the language needs of children with EAL as early as possible. Ultimately, it has the potential to lead to better quality and more equitable services to address these needs.



1

Information for Schools

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

What will the study involve?

With the help of school staff, I will recruit pupils in Key Stage 1 who are reported to have EAL. The study will be explained to parents and informed consent will be obtained for their children to take part. Parents will also be asked to complete a questionnaire on their child's general development, including with regard to language development and their use of language at home and outside the school environment. Additionally, they will be asked to provide video recordings of their child's interactions at home, following instructions on how to do so. As parents' English fluency is likely to vary, different approaches will be used to ensure they understand what they are agreeing to and what the questionnaire is asking. For those whose English is at a higher level, information about the project, the consent forms, the video recordings and the questionnaire will be given to complete in their own time. In other cases, I will interview parents via Zoom and work with an interpreter if necessary. Any further support the school staff can offer in liaising with parents to encourage their participation will be warmly welcomed. To show gratitude for their valuable cooperation, £10 Amazon vouchers will be provided to parents and children for participation.

Teachers will be asked to complete a short questionnaire about each child's development and use of English in school. If possible (pandemic permitting), they may also share a video recording of the child's interaction with others in the classroom that allows me to gain a better understanding of their language development. Written permission will be sought from the parents of all children potentially included on video recordings.

Following this, the children will attend two 30–45-minute individual sessions. I will ask school staff for a quiet, confidential space inside the school to conduct these. All activities involved are designed to be enjoyable and fun for children of this age. In the first session, the child's language development will be assessed directly using activities that involve talking and interacting with a range of resources such as toys and pictures. In the second session, the dynamic assessment will take place. This will involve telling stories in English and playing a game about expressing feelings with our voices. This latter session will be video recorded. I have current DBS clearance and will cease activities immediately if the child becomes distressed or signals that they do not wish to continue.

Information for Schools

Confidentiality and data management

All data collected will be anonymised and stored securely on password-protected University servers and in a locked filing cabinet at Newcastle University. This data will only be held for the purpose of this research and will be destroyed within 5 years. It will not be possible to identify individual participants from the resulting PhD thesis and/or other research outputs, such as journal articles and conference or seminar presentations.

Ethical approval

The study has been approved by the Faculty of Humanities & Social Sciences Research Ethics Committee at Newcastle University.

Risks and benefits

The outcomes of this research should provide valuable information that will assist in identifying children with EAL at risk of having DLD and/or who need extra help with their English. Schools taking part will also directly benefit from the close study of the participating children's language development, which could help to recognise any additional support they require.

There are no foreseeable risks associated with this study. Regarding COVID-19 risk, a risk assessment for working with children in schools has been developed following the latest Newcastle University and Government guidance. I aim to keep school staff, pupils and myself safe and to avoid spreading the virus. I will ensure that I am aware of your school's requirements and adhere to your risk assessment, visitor protocols and other relevant guidance.

Many thanks for taking the time to read this. If you have any questions or would like more information, please do not hesitate to contact me at t.garrido-tamayo2@newcastle.ac.uk. You can also email my supervisors Dr Carolyn Letts at carolyn.letts@newcastle.ac.uk and Dr Laurence White at laurence.white@newcastle.ac.uk.

Best wishes,
Teresa Garrido-Tamayo



3

Appendix 22. Parental Information Sheet Used in the Main Study

Information for Parents/Guardians

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

September 2022

Dear Parent/Guardian,

I am a PhD student at the School of Education, Communication and Language Sciences at Newcastle University carrying out research on the assessment of language difficulties in children learning English as an additional language. This project is under the supervision of Drs Carolyn Letts and Laurence White. I am contacting to ask whether your child can be involved in this research, and whether you can inform us about your child's language development and the language(s) they use in the home and outside of school. You will receive a £10 Amazon voucher to show gratitude for your and your child's valuable cooperation in this study.

What is the study about?

Many children have difficulties understanding and/or using language regardless of whether English is their first language or not. These children are identified as having a Developmental Language Disorder (DLD) when these difficulties are significant, persistent and cannot be explained by an obvious cause. Importantly, DLD can affect most aspects of a child's life, including their learning and school achievement.

Currently, detecting DLD in children learning English as an additional language is tricky because there is a lack of tools that we can use with children growing up in environments where more than one language is spoken. I want to find out how useful a tool I have created is in detecting DLD in children learning English as an additional language, so that their needs can be addressed as quickly as possible. I am therefore looking to recruit children between the ages of 4 and 8 years who have English as an additional language, and who either **present or do not present** language difficulties.

Why was your child chosen?

I contacted your school and they suggested your child would be suitable for this study as they are aged between 4 (reception year) and 8 and have English as an additional language.



1

Information for Parents/Guardians

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

September 2022

What am I asking parents and children to do?

If you agree to take part, you will be asked to provide information on your child's general and language development, as well as the language(s) you use at home and outside of the school environment. This will be in the form of a questionnaire you can complete in your own time, or I may meet with you in a Zoom video call where we can go through the questionnaire together – whichever works best for you. I can also supply an interpreter to help if necessary.

Additionally, I will need to observe your child's interactions to gain a better understanding of their language development. Normally I would do this in person myself, but the risks associated with the COVID-19 virus currently prevent me from doing so. For this reason, you will be asked to record your child's interactions during everyday situations (e.g., play time with siblings, or while looking at a book or holding a conversation with you) and share these video recordings with me. Instructions on the best way to do this will be provided. Please note that the child's teacher will be also asked to provide additional information on their language development and usage in school through a questionnaire and classroom video recordings.

Following this, I will come into the school twice and spend around 45 minutes with your child each time. During the first time, I will do some activities with your child that involve talking and interacting with toys and pictures. The second time, I will conduct some activities with your child that involve telling stories and playing a game where your child will learn about how we can express our feelings with our voices. This second session will be video recorded. All activities will take place in a quiet area in your child's school separated from the main classroom to avoid disruptions and interruptions, but we will always be within sight of their teachers. I will not do any of these activities with your child unless I have your full consent. All activities are designed to be enjoyable and fun for children of this age. However, I will stop at once if your child indicates in any way that they do not wish to participate or continue with any part of the session. I have full Disclosure and Barring clearance, which means that I am authorised and qualified to work safely with children.

Information for Parents/Guardians

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

September 2022

Ethics

The study has been approved by the Faculty of Humanities & Social Sciences Research Ethics Committee at Newcastle University.

Risks and benefits

There are no known risks to participation in this study. Your child would benefit in that their language development would be studied closely and, therefore, any extra support they may need will be recognised. In the broader context, the outcomes of this research should provide valuable information that will assist in identifying children with English as an additional language at risk of having DLD and/or who need extra help with their English.

Regarding COVID-19 risk, a risk assessment for working with children in schools has been developed following the latest Newcastle University and Government guidance. I aim to keep school staff, pupils and myself safe and to avoid spreading the virus. I will ensure that I am aware of the individual requirements at your child's school and that I adhere to the school's own risk assessment, visitor protocols and other relevant guidance.

Does your child have to do this?

No. If you do not want your child to participate, this will not affect anything that they do in school. If you change your mind, you can withdraw your child from the research at any time. If your child does not want to join in with the activities or becomes upset, we will stop at once. If you would like a brief report on what happened in the activities with your child, this can be provided.

Information for Parents/Guardians

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

September 2022

Who will know about this?

Solely your child's teachers, myself and the researchers I work with. We will not pass any personal information about you or your child to anyone else, nor store this information in a location where others may see it. If your child chats to me about topics unrelated to the project, this will be fully confidential, except in the unlikely case that they disclose something that suggests their safety is at risk. If so, I will follow appropriate school procedures to ensure your child stays safe.

The activities I will undertake with your child will be recorded on paper and electronic documents, as well as video recordings. I will give a unique participant code to all the information I collect from you and your child, and neither your name nor your child's name will appear on any documents or files. Papers will be stored in a locked filing cabinet in the university, and video recordings and any other electronic files will be stored on secure, password-protected University servers. All of this will be destroyed within 5 years of the project. When I write about this study in my thesis or talk about it to others, for example, at a conference or a seminar, no information will be included that will make it possible to identify any individual children.

Consent

If you are willing for your child to participate in this study, please sign the attached consent form.

Many thanks for taking the time to read this. If you have any questions or would like more information, please do not hesitate to contact me at t.garrido-tamayo2@newcastle.ac.uk. You can also email my supervisors Dr Carolyn Letts at carolyn.letts@newcastle.ac.uk and Dr Laurence White at laurence.white@newcastle.ac.uk.

Best wishes,
Teresa Garrido-Tamayo



4

Appendix 23. Parental Consent Form Used in the Main Study

Creating a story-based dynamic assessment to identify developmental language disorder in children learning English as an additional language | September 2022

Consent form for Parents/Guardians

Please put your initials in the box to indicate you have read and agree with the following statements:

	Initials
1. I have read the information sheet dated September 2022 for the study above. I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.	
2. I have had the opportunity to review the project risk assessment.	
3. I understand that my child will carry out activities with a researcher, some of which will be video recorded. These video recordings will be stored securely and confidentially, and will be destroyed within 5 years of the project.	
4. I agree for the researcher to share the video recordings of the sessions with the researchers working within this project.	
5. I agree to provide information about the language(s) the child speaks at home and when outside of school, as well as their general and language development. This will be in the form of a questionnaire that I can complete in my own time, or I may choose to meet with the researcher via Zoom to go through the questionnaire together.	
6. I agree to video record my child's interactions during everyday situations and share these video recordings with the researcher. These video recordings will be stored securely and confidentially, and will be destroyed within 5 years of the project.	
7. I agree for my child's teacher to provide additional information on their language development and usage in school through a questionnaire and classroom video recordings.	
8. I understand that all information collected will be anonymised and confidentiality will be maintained in all outputs relating to this study. It will not be possible for anyone other than the researcher to identify my child.	
9. I agree that the researcher can contact me regarding the questionnaire and the video recordings I provide if necessary.	
10. I understand that the participation of myself and my child is voluntary. I can withdraw at any time without giving reason and without detriment to any care or services my child or I may be receiving or may receive in the future. Any activity will stop immediately if my child does not want to join in or becomes upset.	

I agree for my child (name) _____ to participate in this study.

Parent/Guardian

Signature:
Date:
Print full name:
Relationship to child:
Email address and/or telephone number:

Please return the form to your child's teacher



Appendix 24. Child-Friendly Information Sheet Used in the Main Study

Information for Children



Hello! My name is Teresa.

I'm a researcher at Newcastle University.

I work with children like you who are learning English and speak other languages at home, which is awesome!

WHY ARE YOU HERE?

I'm coming to visit you at your school twice.

The first time we'll play fun games using pictures and toys!



The second time we'll play a game about feelings and then tell some stories.



Both times, we'll wash our hands before and after playing to keep the germs away. Sometimes we might use hand gel too.

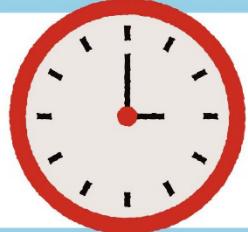


You might also see me wearing a face mask or visor.

Information for Children

HOW LONG WILL WE PLAY FOR?

Not long! Both times will take about the same as lunch break.



DO I HAVE TO SAY YES?

Absolutely not. If you don't want to play, just say no.
It's your choice.



Even if you say yes now, you can still change your mind later.

WHAT'S GOOD ABOUT PLAYING WITH YOU?



By playing these games, you will be helping me to make a tool that finds which children need extra help for learning language.

Language is very important! We use it when we speak and listen to others. And also when we write and read.

WHO WILL KNOW ABOUT THIS?

The information I collect about you (like your age and answers to questions) will be stored in a very secure place, and only myself and the researchers I work with will be able to see it.



It will have a special code on it, rather than your name.

Appendix 25. Advertising Flyers Used in the Main Study



BILINGUAL & MULTILINGUAL CHILDREN NEEDED!

Please help us try out a resource that will assist in detecting developmental language disorder in children with English as an additional language.

We are looking for 5-7-year-old children from Gateshead and Newcastle schools who:

- Use languages other than or in addition to English at home.
- Either present or do not present language difficulties.

Participation will involve engaging in fun language activities! Parents and teachers will also be asked to share details on children's language development and use. £10 Amazon vouchers will be provided to thank families for their cooperation.

Project website:
<https://blogs.ncl.ac.uk/tgarridotamayo2>

 **For more details, please contact Teresa (researcher) at:**
t.garrido-tamayo2@newcastle.ac.uk



LAST CALL: BILINGUAL & MULTILINGUAL CHILDREN NEEDED!

Please help us try out a resource that will assist in detecting developmental language disorder in children with English as an additional language.

We are looking for children from Gateshead/Newcastle/surrounding areas who:

- Are aged 4 (reception year) to 8.
- Use languages **other than** English at home, or **in addition to** English at home.
- Either present language difficulties or do not present language difficulties.

Participation will involve engaging in fun language activities! Parents and teachers will also be asked to share details on children's language development and use. £10 Amazon vouchers will be provided to thank families for their cooperation.

Project website:
<https://blogs.ncl.ac.uk/tgarridotamayo2>

 **For more details, please contact Teresa (researcher) at:**
t.garrido-tamayo2@newcastle.ac.uk

Appendix 26. Lone Working Risk Assessment Used for the Main Study

Newcastle University - Risk Assessment			
Project title	Creating a story-based dynamic assessment to detect developmental language disorder in children learning English as an additional language		
Description of work activity	<p>During the course of this project, the researcher (Teresa Garrido-Tamayo) will be visiting homes to observe or work with primary school-aged children who have English as an additional language. The researcher typically works with each child in their school during two individual sessions. Additionally, observations of the children's interactions in daily familiar settings outside the school context are carried out through video recordings shared by the parents using Newcastle University's File Drop-Off Service. Therefore, home visits will only take place as required in the following scenarios:</p> <ol style="list-style-type: none"> 1. When access to the child's school cannot be facilitated to carry out the sessions and parents request for the researcher to do home visits instead of bringing their child to the Speech & Language Clinic or a suitable alternative space in the University campus. 2. When parents are unable to share video recordings of their child's interactions and have requested for the researcher to observe the child at their home instead. <p>Additionally, given scenario 2, the researcher will visit children in public places when parents request this option instead of at home, depending on their needs.</p> <p>The project has been approved by the Faculty of Humanities and Social Sciences Research Ethics Committee at Newcastle University.</p>		
Unit name	Speech and Language Sciences Department, School of Education, Communication and Language Sciences, Newcastle University	Location	Newcastle upon Tyne, Gateshead and surrounding areas
Assessor	Teresa Garrido-Tamayo	Approver (Manager / Responsible person)	Drs Carolyn Letts and Laurence White
Date of assessment	14/12/2022	Review Date (2 years)	14/12/2024
Version: 6 Owner: DM Approved by: OHSS Date of creation: 08/2019 Review Date: 08/2021			
Page 1 of 5			

	Hazards	Risks (Who might be harmed & how?)	Controls
1.	<p>Observations of participants in their homes or public places (e.g. playground, community centre)</p> <p>Sessions with the participants in their homes</p>	<p>The researcher is at risk of physical or verbal abuse and harassment from participants, their relatives or members of the public.</p>	<ul style="list-style-type: none"> • The researcher will assess whether visiting the participant's home or the public place requested by their parents is safe and appropriate on an individual basis. Visits will only be carried out if they are deemed low risk after she has established sufficient contact with the participant's parents and is well-informed about the area/location where the visit will occur, who will be there at the time, whether these people are agreeable with the visit taking place, and any relevant background information on the participant/family. • The researcher will: <ul style="list-style-type: none"> ○ aim to carry out visits in well-lit and populated areas where risk is minimal (e.g. regarding crime activity, weak phone reception), ○ plan travel in advance and use a map as appropriate, ○ be informed of routes in and out of the site and ensure exits are always easily accessible, ○ avoid carrying large amounts of cash, valuables and jewellery, ○ minimise the presence of potential weapons (e.g. sharp/heavy objects) or clothing hazards (e.g. scarves) that could lead to her being harmed, ○ ensure that her attire is professional and does not cause offence, ○ ensure professional boundaries at all times, including discouraging inappropriate behaviour and requests for personal information,
Version: 6 Owner: DM Approved by: OHSS Date of creation: 08/2019 Review Date: 08/2021			
Page 2 of 5			

	Hazards	Risks (Who might be harmed & how?)	Controls
			<ul style="list-style-type: none"> ○ behave professionally at all times and avoid potentially offensive language or behaviour, ○ avoid approaching members of the public who appear agitated or under the influence of alcohol or drugs, ○ call the police or emergency services if anyone is at immediate risk or a crime is taking place, ○ always carry photo ID (University Smartcard), ○ ensure that her mobile phone is charged and switched on at all times to call for help if required, ○ leave her locations, travel plan and contact details with her supervisors and partner, ○ inform her supervisors and partner about the times that she intends to be doing the home visit and any schedule changes, ○ let her supervisors and partner know when the home visit has been completed and she has safely left the place. If the researcher does not make contact when expected, they will be prepared to contact police or emergency services as necessary.
2.	Illness or distress	Participant becomes unwell or shows signs of distress during the researcher's visit	<ul style="list-style-type: none"> ● The researcher has experience working with children and her DBS clearance will be made available to parents upon request. She can recognise distress signs and has strategies for reassuring children.

Version: 6

Owner: DM

Approved by: OHSS

Date of creation: 08/2019

Review Date: 08/2021

Page 3 of 5

	Hazards	Risks (Who might be harmed & how?)	Controls
		Researcher becomes unwell before or during her visit	<ul style="list-style-type: none"> ● Warm-up time will be provided at the start of the visits to build rapport with the child. When working directly with the child, a relaxed atmosphere will be established during the sessions so that the child feels comfortable and safe. ● During the sessions, the researcher will work with the child in a quiet room in their homes to avoid disruptions or interruptions, but the door will be left open and parents will be within close proximity so they can supervise the child. ● During the observations, whether these are in the child's home or a public place, parents will be present with their child or within close proximity so they can supervise them at all times. ● If the child does not wish to be observed or participate in the activities at any point, the researcher will immediately stop the observation/session and inform parents about this. ● If the child becomes unwell during the session, the researcher will immediately alert the parents so they can act accordingly. ● If the researcher becomes unwell before or during her visit, she will notify the child's parents and reschedule the visit as soon as possible.
3.	Travel via walking or by private transport/taxi	Slips, trips and falls Hazardous weather Collisions with vehicles, pedestrians or other objects causing injury or death to the researcher	<ul style="list-style-type: none"> ● As stated above, the researcher will inform her supervisors and partner of travel plans and the dates and times of the visits. ● The researcher will: <ul style="list-style-type: none"> ○ monitor traffic and surroundings hazards and determine the safest place to walk around obstacles on the footpath,

Version: 6

Owner: DM

Approved by: OHSS

Date of creation: 08/2019

Review Date: 08/2021

Page 4 of 5

	Hazards	Risks (Who might be harmed & how?)	Controls
			<ul style="list-style-type: none"> ○ wear appropriate clothing and footwear to suit weather conditions, ○ check the weather forecast and reschedule her visit if necessary, ○ seek medical attention where appropriate for injuries. <ul style="list-style-type: none"> ● Travel via private transport will only be done if the driver is fit to drive and the vehicle is in a roadworthy condition. In these cases, the researcher will wear her seat belt before departure and ensure any heavy equipment items are held in the boot.

Additional Controls (is there anything you need to plan for?)	Who	Target Date	Completion Date
N/A	[Insert Name]	Click or tap to enter a date.	Click or tap to enter a date.
Emergency procedures			

Signature of Responsible Person (Double click on the signature box below)

 Recoverable Signature

X Teresa Garrido-Tamayo

Signed by: 0cd5fe38-671e-477e-97e3-307767a8/a84

Version: 6

Owner: DM

Approved by: OHSS

Date of creation: 08/2019

Review Date: 08/2021

Page 5 of 5

Appendix 27. Certificate of Research Participation Used in the Main Study



Appendix 28. Parental Questionnaire: Development & Language Experience

Please note that the parental questionnaire presented below is an exported version of the original online questionnaire. Some formatting discrepancies, such as tables splitting across pages, may be observed due to the exportation process. These issues were not present in the questionnaire provided to the parents.

Parental Questionnaire: Development & Language Experience

Page 1: Welcome

Your answers to this questionnaire will help us get a sense of your child's general development and how their language is developing, as well as their language use and exposure at home and outside of the school environment.

Please fill in the questions to the best of your knowledge and remember that there are no right or wrong answers. You do not have to answer every question if you do not feel comfortable with one or more of them. However, please try to fill in as much as possible because your answers are very important to us. As stated in the information sheet, your responses will be kept confidential and anonymous.

Many thanks for your time and co-operation.

If you have any questions about this questionnaire, please contact Teresa Garrido-Tamayo at t.garrido-tamayo2@newcastle.ac.uk or 07565 527743.

Page 2: Introductory details

1. Child's participant code:

2. This questionnaire is to be completed by the child's main caregiver, that is, the person who spends the most time with the child. Please specify your relationship to the child (for example, mother, father, grandmother, etc.):

3. Today's date:

Dates need to be in the format 'DD/MM/YYYY', for example 27/03/1980.



(dd/mm/yyyy)

Page 3: General information about the child

4. What is your child's gender?

- Male
- Female
- Other

4.a. If you selected Other, please specify:

5. Please enter your child's date of birth:

Dates need to be in the format 'DD/MM/YYYY', for example 27/03/1980.

(dd/mm/yyyy)

6. Please enter your child's place of birth:

7. If the place of birth is not in the United Kingdom, date of arrival in the United Kingdom:

Dates need to be in the format 'DD/MM/YYYY', for example 27/03/1980.

(dd/mm/yyyy)

8. Does your child have any brothers or sisters?

- Yes
- No

8.a. If yes, please enter all your children's ages below in years and months.

1st born child:	<input type="text"/>
2nd born child:	<input type="text"/>
3rd born child:	<input type="text"/>

4th born child:	
5th born child:	
6th born child:	
Other:	

Page 4: Languages used with and by the child

9. At the present time, what language(s) are spoken at home?

Language 1:	
Language 2:	
Other language (please specify):	

Please take note of the numbers associated with the language(s) above, as these will be referred to throughout the questionnaire as Language 1, Language 2 and Other Language.

10. At what age was your child first in contact with each language? Please specify in years and months. If contact was from birth, say 0 years.

Language 1:	
Language 2:	
Other language:	

Main Caregiver 1

The following questions are for Main Caregiver 1. This is you, the person completing the questionnaire.

11. What language(s) do you speak to the child? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

12. What language(s) does the child speak to you? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

Main Caregiver 2 (if applicable)

If your child has a second main caregiver, this section applies to them. This would be someone who spends an equal amount of

time with the child as you, the main caregiver. Otherwise, please skip these questions.

13. Please specify the relationship of Main Caregiver 2 to the child (for example, mother, father, grandmother, etc.):

14. What language(s) does Main Caregiver 2 speak to the child? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

15. What language(s) does the child speak to Main Caregiver 2? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

Other adults (if applicable)

If there is any other adult that regularly takes care of your child (for example, grandparent, babysitter, etc.), this section applies to them. Otherwise, please skip these questions.

16. Please specify the relationship of this adult to the child (for example, grandparent, babysitter, etc.):

17. What language(s) does this person speak to the child? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

18. What language(s) does the child speak to this person? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

Siblings (if applicable)

If your child has brothers or sisters, this section applies to them. Otherwise, please skip these questions.

19. Language(s) used by oldest brother/sister with the child (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

20. Language(s) used by the child with oldest brother/sister (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

21. Language(s) used by second oldest brother/sister with the child (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

22. Language(s) used by the child with second oldest brother/sister (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				

Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

23. Language(s) used by third oldest brother/sister with the child (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

24. Language(s) used by the child with third oldest brother/sister (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

25. Language(s) used by fourth oldest brother/sister with the child (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

26. Language(s) used by the child with fourth oldest brother/sister (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

27. Language(s) used by fifth oldest brother/sister with the child (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

28. Language(s) used by the child with fifth oldest brother/sister (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

29. Language(s) used by sixth oldest brother/sister with the child (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

30. Language(s) used by the child with sixth oldest brother/sister (please tick):

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Other language	<input type="checkbox"/>				

The child's interests and activities

31. What are your child's main interests and favourite things to do?

32. Focusing on **Language 1**, which of the following activities involving language does your child do at home each week? Please tick:

Please don't select more than 1 answer(s) per row.

	Rarely/never	At least once a week	Almost everyday/everyday
Reading or looking at books and/or having books read to them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Telling stories and/or having stories told to them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Listening to music and/or singing songs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Watching movies or videos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Playing with siblings and/or friends	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

33. Focusing on **Language 2**, which of the following activities involving language does your child do at home each week? Please tick:

Please don't select more than 1 answer(s) per row.

	Rarely/never	At least once a week	Almost everyday/everyday
Reading or looking at books and/or having books read to them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Telling stories and/or having stories told to them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Listening to music and/or singing songs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Watching movies or videos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Playing with siblings and/or friends	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

34. Focusing on the **Other Language**, which of the following activities involving language does your child do at home each week? Please tick:

Please don't select more than 1 answer(s) per row.

	Rarely/never	At least once a week	Almost everyday/everyday
Reading or looking at books and/or having books read to them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Telling stories and/or having stories told to them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Listening to music and/or singing songs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Watching movies or videos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Playing with siblings and/or friends	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

35. List any other activities that the child does outside of school:

Activities:	Language(s) used:			Frequency (days per week):
	Language 1	Language 2	Other language	
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

10 / 19

5	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="text"/>
6	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="text"/>

36. Briefly describe a typical day in your child's life from the time they wake up in the morning until they go to bed at night. You do not need to give details of school activities, just indicate when your child is at school.

Page 5: The child's general and language development

37. Were there any complications during pregnancy or birth?

- Yes
- No

37.a. If yes, please describe:

38. Was your child born early/on time/late?

- Early
- On time
- Late

38.a. By how much were they born early or late? Please specify weeks and days if possible.

39. At what age did your child begin to walk? Please specify age in years and months if possible.

40. Does your child have any medical condition(s)?

- Yes
- No

40.a. If yes, please describe:

41. Has your child experienced any of the following?

- Frequent colds
- Ear infections
- Vision problems
- Hearing problems
- Other

41.a. If you selected Other, please specify:

41.b. If you ticked any of the above, please give details:

Language development

42. At what age did your child say their first recognisable words (any language)? Please specify age in years and months if possible.

43. At what age did your child begin joining words together into a short sequence? Please specify age in years and months if possible.

44. How easy is it for you to understand what your child is trying to say?

45. How easy is it for friends and family to understand what your child is trying to say?

46. Does your child understand what you say to them most of the time?

Yes
 No

46.a. If no, please give one or two examples of when your child appears to not understand.

47. Is your child able to follow instructions with more than one step most of the time?

Yes
 No

47.a. If no, please explain:

48. Mixing languages is normal for bilingual children. Does your child ever use more than one language at the same time?

Yes
 No

48.a. If yes, does this worry you?

49. Compared with other children of the same age, do you have any concerns about your child's speech and language abilities?

Yes
 No

49.a. If yes, please describe your concerns:

49.b. If you have indicated above that your child has speech and/or language difficulties, how do these difficulties affect their daily life?

Page 6: Information about the child's family

50. Has anyone in the family had, or have, speech, language and/or hearing problems or learning difficulties?

- Yes
- No

50.a. If yes, who and what kind of difficulties? Please describe:

Information about Main Caregiver 1

51. In which country were you born?

52. If you are currently working, what is the language you use at your workplace?

53. What is the highest level of education you have completed?

- Primary school
- Secondary school
- University
- Other

53.a. If you selected Other, please specify:

54. In your opinion, how well can you speak and understand the following languages?

Please don't select more than 1 answer(s) per row.

	Not at all	Basic understanding and speaking ability	Good understanding and can express myself in many situations	Excellent understanding and can express myself in most situations
Language 1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Language 2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Information about Main Caregiver 2 (if applicable)

55. In which country was Main Caregiver 2 born?

56. If Main Caregiver 2 is currently working, what is the language they use at their workplace?

57. What is the highest level of education Main Caregiver 2 has completed?

- Primary school
- Secondary school
- University
- Other

57.a. If you selected Other, please specify:

58. In your opinion, how well can Main Caregiver 2 speak and understand the following languages?

Please don't select more than 1 answer(s) per row.

	Not at all	Basic understanding and speaking ability	Good understanding and can express myself in many situations	Excellent understanding and can express myself in most situations
Language 1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Language 2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Page 7: End

Thank you for completing the questionnaire. Your answers have been submitted and you can now close this page.

If you have any questions about this questionnaire, please contact Teresa Garrido-Tamayo at t.garrido-tamayo2@newcastle.ac.uk or 07565 527743.

Appendix 29. Teacher Questionnaire: Language Development & Experience in the School Context

Please note that the teacher questionnaire presented below is an exported version of the original online questionnaire. Some formatting discrepancies, such as tables splitting across pages, may be observed due to the exportation process. These issues were not present in the questionnaire provided to the teachers.

Teacher Questionnaire: Language Development & Experience in the School Context

Page 1: Welcome

Your answers to this questionnaire will help us get a sense of the child's English language development and their language use and exposure in school. Please fill in the questions to the best of your knowledge and remember that there are no right or wrong answers. As stated in the information sheet for the school, your responses will be kept confidential and anonymous.

Many thanks for your time and co-operation.

If you have any questions about this questionnaire, please contact Teresa Garrido-Tamayo at t.garrido-tamayo2@newcastle.ac.uk or 07565 527743.

Page 2: Introductory details

1. Child's participant code: * *Required*

2. School:

3. School year:

4. Please specify your role as a school staff member:

5. Today's date:

Dates need to be in the format 'DD/MM/YYYY', for example 27/03/1980.



(dd/mm/yyyy)

Page 3: Child's language use and exposure in school

6. Date of entry into the school:

Dates need to be in the format 'DD/MM/YYYY', for example 27/03/1980.



(dd/mm/yyyy)

7. Prior to this, has the child attended any other schools in the United Kingdom?

- Yes
- No

8. In addition to English, are there any other languages used in the school?

- Yes
- No

8.a. If yes, please specify.

Language 1:	<input type="text"/>
Language 2:	<input type="text"/>
Language 3:	<input type="text"/>

Please take note of the numbers associated with the language(s) above, as these will be referred to throughout the questionnaire as Language 1, Language 2 and Language 3.

School staff

9. What language(s) do the school staff speak to the child? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
English	<input type="checkbox"/>				
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Language 3	<input type="checkbox"/>				

10. What language(s) does the child speak to the school staff? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
English	<input type="checkbox"/>				
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Language 3	<input type="checkbox"/>				

11. Please indicate the roles of the staff members who use languages other than English with the child:

Language 1:	<input type="text"/>
Language 2:	<input type="text"/>
Language 3:	<input type="text"/>

Peers

12. What language(s) do the child's peers speak to the child at school? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
English	<input type="checkbox"/>				
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Language 3	<input type="checkbox"/>				

13. What language(s) does the child speak to their peers at school? Please tick:

Please don't select more than 1 answer(s) per row.

	Never	Rarely	Sometimes	Usually	Always
English	<input type="checkbox"/>				
Language 1	<input type="checkbox"/>				
Language 2	<input type="checkbox"/>				
Language 3	<input type="checkbox"/>				

Silent period

14. During the earliest stages of acquiring English as an additional language, children sometimes go through a silent period as they adapt to the new language environment. Is this something the child has experienced or is currently experiencing?

Yes
 No

14.a. If yes, please indicate how long the silent period has lasted or is lasting and elaborate.

Page 4: Child's English proficiency

15. According to the Department for Education's Proficiency in English Scale below, which proficiency stage does the child correspond most closely to?

- Stage A. New to English. May use first/home language for learning and other purposes. May remain completely silent in the classroom. May be copying/repeating some words or phrases. May understand some everyday expressions in English but may have minimal or no literacy in English. Needs a considerable amount of English as an Additional Language (EAL) support to access curriculum content.
- Stage B. Early Acquisition. May follow day-to-day social communication in English and participate in learning activities with support. Beginning to use spoken English for social purposes. May understand simple instructions and can follow narrative/accounts with visual support. May have developed some skills in reading and writing. May have become familiar with some subject-specific vocabulary. Still needs a significant amount of EAL support to access the curriculum.
- Stage C. Developing Competence. May participate in learning activities with increasing independence. Able to express self orally in English, but structural inaccuracies are still apparent. Literacy will require ongoing support, particularly for understanding text and writing. May be able to follow abstract concepts and more complex written English. Requires ongoing EAL support to access the curriculum fully.
- Stage D. Competent. Oral English will be developing well, enabling successful engagement in activities across the curriculum. Can read and understand a wide variety of texts. Written English may lack complexity and contain occasional evidence of errors in structure. Needs some support to access subtle nuances of meaning, to refine English usage, and to develop abstract vocabulary. Needs some/occasional EAL support to access complex curriculum material and tasks.
- Stage E. Fluent. Can operate across the curriculum to a level of competence equivalent to that of a pupil who uses English as their first language. Operates without EAL support across the curriculum.

Page 5: Concerns about the child's language development

16. Do you have any concerns about the child's speech, language, communication, or social skills?

- Yes
- No

16.a. If yes, please describe your concerns.

17. Compared with children of similar age, language, and cultural background (if present in the school), is the child slower to learn English?

- Yes
- No

17.a. If yes, please explain.

Page 6: End

Thank you for completing the questionnaire. Your answers have been submitted and you can now close this page.

If you have any questions about this questionnaire, please contact Teresa Garrido-Tamayo at t.garrido-tamayo2@newcastle.ac.uk or 07565 527743.

Appendix 30. Observation Form Used in the Main Study

Participant's code:		
OBSERVATION NO.:		
Date and time:		
Setting:		
Main activity:		
Speech intelligibility		
Familiar listeners understand the child's speech	Seen	Notes
Unfamiliar listeners understand the child's speech		
Speaks differently compared to their peers (e.g., too loudly/softly, unusual intonation)		
Expressive communication		
Initiates interactions with adults/other children		
Maintains interactions with adults/other children		
Adapts communication style to meet listener needs		
Shows topic relevancy		
Can communicate wants and needs by speaking		
Can communicate wants and needs non-verbally		
Uses non-verbal communication appropriately (e.g., eye contact, proximity, touch, gesture, facial expressions)		
Further notes (e.g., main means of expression (words, phrases, gestures), language preference):		
Receptive communication		
Responds timely and appropriately to others		
Can follow instructions given to a whole group		
Can follow instructions just given to them		
Requires repetition or simplification of what was said		
Looks around for clues before carrying out instructions		
Watches and copies others' responses to instructions		
Further notes:		
Attention and listening		
Stays on task		
Sits still without fidgeting/becoming restless		
Is easily distracted		
‘Switches off’/daydreams		
Completes work		
Further notes (e.g., age-appropriateness of attention span):		

Play and social interaction	Seen	Notes
Shows age-appropriate turn-taking skills		
Engages in age-appropriate joint attention		
Shares toys and objects		
<i>Play – Cognitive perspective:</i>		
- Exploratory		
- Symbolic		
- Constructive		
- Rule-governed		
<i>Play – Social perspective:</i>		
- Unoccupied		
- Solitary		
- Onlooker		
- Parallel		
- Associative		
- Rough and tumble		
- Cooperative		
Further notes (e.g., Does the child show any difficulties in accessing play? Do they initiate play interactions or wait to be approached? What type of play interactions do they prefer?):		
Communicative functions	Seen	Notes
Label things/actions		
Request things/actions		
Describe things/actions		
Direct attention to self/things/events		
Ask for information		
Give information		
Agree		
Disagree, reject		
Protest, argue		
Show humour, tease		
Other functions (e.g., greet, ask permission, apologise, warn, promise, show off):		

Appendix 31. Participation Guidance for Parents Provided in the Main Study

Instructions for Parents/Guardians

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

Dear Parent/Guardian,

Many thanks for signing the consent form agreeing to your and your child's participation in my study. Please carefully read the information below to understand what you are required to do next. If you have any questions, please do not hesitate to get in touch with me (Teresa) at t.garrido-tamayo2@newcastle.ac.uk or 07565 527743 (I'm also available through WhatsApp).

1. Complete the questionnaire about your child's development and language experience.

You can do this in your own time using the online form, which can be accessed here: <https://newcastle.onlinesurveys.ac.uk/parental-questionnaire>

Alternatively, you can fill in the attached paper copy and return this to the school, or we may meet to go through the questionnaire together – whichever is easier for you. I can also supply an interpreter to help if necessary.



Your child's participant code is:

2. Video-record your child's interactions.

Use a smartphone or a camera to video-record your child interacting in any language with you and/or other close family members or friends during everyday situations at home. Examples of these situations may include:



- Playtime with other children such as siblings, cousins or family friends.
- Reading a book with you or a close family member.
- Holding a conversation with you or a close family member.

Please try to record the child at least twice for around 10-25 minutes each time. Following this, upload the video files to Newcastle University's File Drop-Off Service: <https://dropoff.ncl.ac.uk> Click on *drop off* and follow the steps. Please send the files to t.garrido-tamayo2@newcastle.ac.uk and indicate your child's participant code in the notes.

Appendix 32. Participation Guidance for Teachers Provided in the Main Study

Instructions for School Staff

CREATING A STORY-BASED DYNAMIC ASSESSMENT TO IDENTIFY DEVELOPMENTAL LANGUAGE DISORDER IN CHILDREN LEARNING ENGLISH AS AN ADDITIONAL LANGUAGE

Dear School Staff,

Many thanks for your help with my PhD project titled above. Please read the information below to understand what you are required to do for the children taking part. If you have any questions, please do not hesitate to contact me (Teresa) at t.garrido-tamayo2@newcastle.ac.uk or 07565 527743.

1. Complete the questionnaire about the child's English language development and language use and exposure in school.

You can do this in your own time using the online form, which can be accessed here:
<https://newcastle.onlinesurveys.ac.uk/teacher-questionnaire>

Alternatively, we may meet to go through the questionnaire together – whichever is easier for you.

The child's participant code is:

2. Video-record the child's interactions in the classroom.

Please share the provided video recordings consent forms with the parents/guardians of the participating children's classmates. Following this, video-record the child interacting with other children in the classroom for whom parents have consented to appear in these recordings. If the school does not have a suitable device for this, a camera will be provided.

For example, you could record the children engaging in free conversation and/or carrying out an activity together that encourages them to communicate and share their ideas. Please choose the times/activities you believe show the child "coming out of their shell" best.

Please try to record the child at least twice for around 10-25 minutes each time. Following this, upload the video files to Newcastle University's File Drop-Off Service: <https://dropoff.ncl.ac.uk> Click on *drop off* and follow the steps. Please send the files to t.garrido-tamayo2@newcastle.ac.uk and indicate the child's participant code in the notes.



Project website:
<https://blogs.ncl.ac.uk/tgarridotamayo2>

Appendix 33. Pearson's and Spearman's Correlation Matrices between Scores in the Receptive Affective Prosody Task and the CL-NWRT and NRDLS

Table 33A. Pearson's Correlations between Scores in the Receptive Affective Prosody Task and the CL-NWRT and NRDLS

		1	2	3	4	5	6
1. NRDLS Comp.	Pearson Correlation	1.00	.79**	.55*	.54*	.28	-.32
	Sig. (2-tailed)	.	.001	.041	.047	.325	.260
	N	14	14	14	14	14	14
2. NRDLS Prod.	Pearson Correlation	.79**	1.00	.66**	.76**	.55*	-.19
	Sig. (2-tailed)	.001	.	.010	.002	.042	.510
	N	14	14	14	14	14	14
3. CL-NWRT	Pearson Correlation	.55*	.66**	1.00	.69**	.52	-.14
	Sig. (2-tailed)	.041	.010	.	.006	.058	.626
	N	14	14	14	14	14	14
4. RAPT pretest	Pearson Correlation	.54*	.76**	.69**	1.00	.83**	-.07
	Sig. (2-tailed)	.047	.002	.006	.	<.001	.803
	N	14	14	14	14	14	14
5. RAPT post-test	Pearson Correlation	.28	.55*	.52	.83**	1.00	.50
	Sig. (2-tailed)	.325	.042	.058	<.001	.	.067
	N	14	14	14	14	14	14
6. RAPT change	Pearson Correlation	-.32	-.19	-.14	-.07	.50	1.00
	Sig. (2-tailed)	.260	.510	.626	.803	.067	.
	N	14	14	14	14	14	14

Note. **. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; RAPT pretest: *Receptive Affective Prosody Task* pretest score; RAPT post-test: *Receptive Affective Prosody Task* post-test score; RAPT change: *Receptive Affective Prosody Task* change score.

Table 33B. Spearman's Correlations between Scores in the Receptive Affective Prosody Task and the CL-NWRT and NRDLS

			1	2	3	4	5	6
Spearman's rho	1. NRDLS Comp.	Correlation Coefficient	1.00	.70**	.46	.69**	.37	-.21
		Sig. (2-tailed)	.	.005	.099	.006	.200	.463
		N	14	14	14	14	14	14
	2. NRDLS Prod.	Correlation Coefficient	.70**	1.00	.69**	.76**	.47	-.18
		Sig. (2-tailed)	.005	.	.006	.001	.091	.547
		N	14	14	14	14	14	14
	3. CL-NWRT	Correlation Coefficient	.46	.69**	1.00	.59*	.34	-.11
		Sig. (2-tailed)	.099	.006	.	.027	.237	.719
		N	14	14	14	14	14	14
	4. RAPT pretest	Correlation Coefficient	.69**	.76**	.59*	1.00	.66**	-.25
		Sig. (2-tailed)	.006	.001	.027	.	.010	.392
		N	14	14	14	14	14	14
	5. RAPT post-test	Correlation Coefficient	.37	.47	.34	.66**	1.00	.49
		Sig. (2-tailed)	.200	.091	.237	.010	.	.074
		N	14	14	14	14	14	14
	6. RAPT change	Correlation Coefficient	-.21	-.18	-.11	-.25	.49	1.00
		Sig. (2-tailed)	.463	.547	.719	.392	.074	.
		N	14	14	14	14	14	14

Note. **. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; RAPT pretest: *Receptive Affective Prosody Task* pretest score; RAPT post-test: *Receptive Affective Prosody Task* post-test score; RAPT change: *Receptive Affective Prosody Task* change score.

Appendix 34. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 1: Story Grammar Units and the CL-NWRT and NRDLS

Table 34A. Pearson's Correlations between Scores in Scoring Sheet 1: Story Grammar Units and the CL-NWRT and NRDLS

		1	2	3	4	5	6
1. NRDLS Comp.	Pearson Correlation	1.00	.79**	.55*	.57*	.56*	.10
	Sig. (2-tailed)	.	.001	.041	.034	.037	.724
	N	14	14	14	14	14	14
2. NRDLS Prod.	Pearson Correlation	.79**	1.00	.66**	.82**	.85**	.21
	Sig. (2-tailed)	.001	.	.010	<.001	<.001	.462
	N	14	14	14	14	14	14
3. CL-NWRT	Pearson Correlation	.55*	.66**	1.00	.45	.31	-.13
	Sig. (2-tailed)	.041	.010		.104	.287	.655
	N	14	14	14	14	14	14
4. SS1 pretest	Pearson Correlation	.57*	.82**	.45	1.00	.76**	-.16
	Sig. (2-tailed)	.034	<.001	.104	.	.001	.596
	N	14	14	14	14	14	14
5. SS1 post-test	Pearson Correlation	.56*	.85**	.31	.76**	1.00	.52
	Sig. (2-tailed)	.037	<.001	.287	.001	.	.057
	N	14	14	14	14	14	14
6. SS1 change	Pearson Correlation	.10	.21	-.13	-.16	.52	1.00
	Sig. (2-tailed)	.724	.462	.655	.596	.057	.
	N	14	14	14	14	14	14

Note. **. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; SS1 pretest: *Scoring Sheet 1: Story Grammar Units* pretest score; SS1 post-test: *Scoring Sheet 1: Story Grammar Units* post-test score; SS1 change: *Scoring Sheet 1: Story Grammar Units* change score.

Table 34B. Spearman's Correlations between Scores in Scoring Sheet 1: Story Grammar Units and the CL-NWRT and NRDLS

			1	2	3	4	5	6
Spearman's rho	1. NRDLS Comp.	Correlation Coefficient	1.00	.70**	.46	.64*	.46	-.05
		Sig. (2-tailed)	.	.005	.099	.013	.101	.864
		N	14	14	14	14	14	14
	2. NRDLS Prod.	Correlation Coefficient	.70**	1.00	.69**	.53*	.64*	.23
		Sig. (2-tailed)	.005	.	.006	.049	.014	.440
		N	14	14	14	14	14	14
	3. CL-NWRT	Correlation Coefficient	.46	.69**	1.00	.24	.28	-.02
		Sig. (2-tailed)	.099	.006	.	.411	.338	.942
		N	14	14	14	14	14	14
	4. SS1 pretest	Correlation Coefficient	.64*	.53*	.24	1.00	.71**	-.25
		Sig. (2-tailed)	.013	.049	.411	.	.004	.394
		N	14	14	14	14	14	14
	5. SS1 post-test	Correlation Coefficient	.46	.64*	.28	.71**	1.00	.42
		Sig. (2-tailed)	.101	.014	.338	.004	.	.140
		N	14	14	14	14	14	14
	6. SS1 change	Correlation Coefficient	-.05	.23	-.02	-.25	.42	1.00
		Sig. (2-tailed)	.864	.440	.942	.394	.140	.
		N	14	14	14	14	14	14

Note. **. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* scores; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; SS1 pretest: *Scoring Sheet 1: Story Grammar Units* pretest score; SS1 post-test: *Scoring Sheet 1: Story Grammar Units* post-test score; SS1 change: *Scoring Sheet 1: Story Grammar Units* change score.

Appendix 35. Spearman's Correlation Matrix between Scores in Scoring Sheet 2: Episodic Structure Complexity and the CL-NWRT and NRDLS

Table 35A. Spearman's Correlations between Scores in Scoring Sheet 2: Episodic Structure Complexity and the CL-NWRT and NRDLS

			1	2	3	4	5	6
Spearman's rho	1. NRDLS Comp.	Correlation Coefficient	1.00	.70**	.46	.40	.37	.03
		Sig. (2-tailed)	.	.005	.099	.161	.194	.927
		N	14	14	14	14	14	14
	2. NRDLS Prod.	Correlation Coefficient	.70**	1.00	.69**	.47	.39	.07
		Sig. (2-tailed)	.005	.	.006	.090	.167	.807
		N	14	14	14	14	14	14
	3. CL-NWRT	Correlation Coefficient	.46	.69**	1.00	.33	-.07	-.19
		Sig. (2-tailed)	.099	.006	.	.257	.816	.519
		N	14	14	14	14	14	14
	4. SS2 pretest	Correlation Coefficient	.40	.47	.33	1.00	.41	-.45
		Sig. (2-tailed)	.161	.090	.257	.	.145	.105
		N	14	14	14	14	14	14
	5. SS2 post-test	Correlation Coefficient	.37	.39	-.07	.41	1.00	.57*
		Sig. (2-tailed)	.194	.167	.816	.145	.	.034
		N	14	14	14	14	14	14
	6. SS2 change	Correlation Coefficient	.03	.07	-.19	-.45	.57*	1.00
		Sig. (2-tailed)	.927	.807	.519	.105	.034	.
		N	14	14	14	14	14	14

Note. **. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; SS2 pretest: *Scoring Sheet 2: Episodic Structure Complexity* pretest rating; SS2 post-test: *Scoring Sheet 2: Episodic Structure Complexity* post-test rating; SS2 change: *Scoring Sheet 2: Episodic Structure Complexity* change score.

Appendix 36. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 3: Emotional Vocabulary and the CL-NWRT and NRDLS

Table 36A. Pearson's Correlations between Scores in Scoring Sheet 3: Emotional Vocabulary and the CL-NWRT and NRDLS

		1	2	3	4	5	6
1. NRDLS Comp.	Pearson Correlation	1.00	.79**	.55*	-.23	.29	.42
	Sig. (2-tailed)	.	.001	.041	.436	.318	.132
	N	14	14	14	14	14	14
2. NRDLS Prod.	Pearson Correlation	.79**	1.00	.66**	-.05	.48	.43
	Sig. (2-tailed)	.001	.	.010	.879	.084	.122
	N	14	14	14	14	14	14
3. CL-NWRT	Pearson Correlation	.55*	.66**	1.00	-.07	.13	.17
	Sig. (2-tailed)	.041	.010	.	.801	.664	.571
	N	14	14	14	14	14	14
4. SS3 pretest	Pearson Correlation	-.23	-.05	-.07	1.00	.26	-.60*
	Sig. (2-tailed)	.436	.879	.801	.	.376	.024
	N	14	14	14	14	14	14
5. SS3 post-test	Pearson Correlation	.29	.48	.13	.26	1.00	.62*
	Sig. (2-tailed)	.318	.084	.664	.376	.	.017
	N	14	14	14	14	14	14
6. SS3 change	Pearson Correlation	.42	.43	.17	-.60*	.62*	1.00
	Sig. (2-tailed)	.132	.122	.571	.024	.017	.
	N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; SS3 pretest: *Scoring Sheet 3: Emotional Vocabulary* pretest score; SS3 post-test: *Scoring Sheet 3: Emotional Vocabulary* post-test score; SS3 change: *Scoring Sheet 3: Emotional Vocabulary* change score.

Table 36B. Spearman's Correlations between Scores in Scoring Sheet 3: Emotional Vocabulary and the CL-NWRT and NRDLS

			1	2	3	4	5	6
Spearman's rho	1. NRDLS Comp.	Correlation Coefficient	1.00	.70**	.46	-.05	.44	.35
		Sig. (2-tailed)	.	.005	.099	.868	.112	.218
		N	14	14	14	14	14	14
	2. NRDLS Prod.	Correlation Coefficient	.70**	1.00	.69**	.03	.59*	.43
		Sig. (2-tailed)	.005	.	.006	.934	.027	.125
		N	14	14	14	14	14	14
	3. CL-NWRT	Correlation Coefficient	.46	.69**	1.00	-.14	.20	.24
		Sig. (2-tailed)	.099	.006	.	.636	.500	.413
		N	14	14	14	14	14	14
	4. SS3 pretest	Correlation Coefficient	-.05	.03	-.14	1.00	.14	-.59*
		Sig. (2-tailed)	.868	.934	.636	.	.624	.026
		N	14	14	14	14	14	14
	5. SS3 post-test	Correlation Coefficient	.44	.59*	.20	.14	1.00	.68**
		Sig. (2-tailed)	.112	.027	.500	.624	.	.008
		N	14	14	14	14	14	14
	6. SS3 change	Correlation Coefficient	.35	.43	.24	-.59*	.68**	1.00
		Sig. (2-tailed)	.218	.125	.413	.026	.008	.
		N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; SS3 pretest: *Scoring Sheet 3: Emotional Vocabulary* pretest score; SS3 post-test: *Scoring Sheet 3: Emotional Vocabulary* post-test score; SS3 change: *Scoring Sheet 3: Emotional Vocabulary* change score.

Appendix 37. Pearson's and Spearman's Correlation Matrices between Scores in Modifiability Rating Scale and the CL-NWRT and NRDLS

Table 37A. Pearson's Correlations between Scores in Modifiability Rating Scale and the CL-NWRT and NRDLS

		1	2	3	4
1. Modifiability	Pearson Correlation	1.00	.50	.68**	.63*
	Sig. (2-tailed)	.	.071	.008	.016
	N	14	14	14	14
2. NRDLS Comp.	Pearson Correlation	.50	1.00	.79**	.55*
	Sig. (2-tailed)	.071	.	.001	.041
	N	14	14	14	14
3. NRDLS Prod.	Pearson Correlation	.68**	.79**	1.00	.66**
	Sig. (2-tailed)	.008	.001	.	.010
	N	14	14	14	14
4. CL-NWRT	Pearson Correlation	.63*	.55*	.66**	1.00
	Sig. (2-tailed)	.016	.041	.010	.
	N	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; Modifiability: *Modifiability Rating Scale* score.

Table 37B. Spearman's Correlations between Scores in Modifiability Rating Scale and the CL-NWRT and NRDLS

		1	2	3	4
Spearman's rho	1. Modifiability	Correlation Coefficient	1.00	.60*	.81**
		Sig. (2-tailed)	.	.022	.001
		N	14	14	14
	2. NRDLS Comp.	Correlation Coefficient	.60*	1.00	.70**
		Sig. (2-tailed)	.022	.	.005
		N	14	14	14
	3. NRDLS Prod.	Correlation Coefficient	.81**	.70**	1.00
		Sig. (2-tailed)	.001	.005	.
		N	14	14	14
	4. CL-NWRT	Correlation Coefficient	.68**	.46	.69**
		Sig. (2-tailed)	.008	.099	.006
		N	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: CL-NWRT: *Crosslinguistic Nonword Repetition Test* score; NRDLS Comp.: *NRDLS Comprehension Scale* score; NRDLS Prod.: *NRDLS Production Scale* score; Modifiability: *Modifiability Rating Scale* score.

Appendix 38. Pearson's and Spearman's Correlation Matrices between Scores in the Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Table 38A. Pearson's Correlations between Scores in the Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

		1	2	3	4	5	6
1. DLD Risk	Pearson Correlation	1.00	-.28	.28	.02	.15	.23
	Sig. (2-tailed)	.	.337	.332	.938	.606	.424
	N	14	14	14	14	14	14
2. Eng Exp	Pearson Correlation	-.28	1.00	-.57*	-.27	-.22	.02
	Sig. (2-tailed)	.337	.	.032	.357	.448	.951
	N	14	14	14	14	14	14
3. Eng Prof	Pearson Correlation	.28	-.57*	1.00	.35	.15	-.27
	Sig. (2-tailed)	.332	.032	.	.224	.614	.349
	N	14	14	14	14	14	14
4. RAPT pretest	Pearson Correlation	.02	-.27	.35	1.00	.83**	-.07
	Sig. (2-tailed)	.938	.357	.224	.	<.001	.803
	N	14	14	14	14	14	14
5. RAPT post-test	Pearson Correlation	.15	-.22	.15	.83**	1.00	.50
	Sig. (2-tailed)	.606	.448	.614	<.001	.	.067
	N	14	14	14	14	14	14
6. RAPT change	Pearson Correlation	.23	.02	-.27	-.07	.50	1.00
	Sig. (2-tailed)	.424	.951	.349	.803	.067	.
	N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; RAPT pretest: *Receptive Affective Prosody Task* pretest score; RAPT post-test: *Receptive Affective Prosody Task* post-test score; RAPT change: *Receptive Affective Prosody Task* change score.

Table 38B. Spearman's Correlations between Scores in the Receptive Affective Prosody Task and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

			1	2	3	4	5	6
Spearman's rho	1. DLD Risk	Correlation Coefficient	1.00	-.22	.27	.10	.11	.27
		Sig. (2-tailed)	.	.449	.361	.734	.718	.354
		N	14	14	14	14	14	14
	2. Eng Exp	Correlation Coefficient	-.22	1.00	-.51	-.31	-.00	.22
		Sig. (2-tailed)	.449	.	.065	.278	.989	.447
		N	14	14	14	14	14	14
	3. Eng Prof	Correlation Coefficient	.27	-.51	1.00	.43	.14	-.22
		Sig. (2-tailed)	.361	.065	.	.123	.646	.446
		N	14	14	14	14	14	14
	4. RAPT pretest	Correlation Coefficient	.10	-.31	.43	1.00	.66**	-.25
		Sig. (2-tailed)	.734	.278	.123	.	.010	.392
		N	14	14	14	14	14	14
	5. RAPT post-test	Correlation Coefficient	.11	-.00	.14	.66**	1.00	.49
		Sig. (2-tailed)	.718	.989	.646	.010	.	.074
		N	14	14	14	14	14	14
	6. RAPT change	Correlation Coefficient	.27	.22	-.22	-.25	.49	1.00
		Sig. (2-tailed)	.354	.447	.446	.392	.074	.
		N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; RAPT pretest: *Receptive Affective Prosody Task* pretest score; RAPT post-test: *Receptive Affective Prosody Task* post-test score; RAPT change: *Receptive Affective Prosody Task* change score.

Appendix 39. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 1: Story Grammar Units and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Table 39A. Pearson's Correlations between Scores in Scoring Sheet 1: Story Grammar Units and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

		1	2	3	4	5	6
1. DLD Risk	Pearson Correlation	1.00	-.28	.28	.12	-.13	-.36
	Sig. (2-tailed)	.	.337	.332	.697	.650	.213
	N	14	14	14	14	14	14
2. Eng Exp	Pearson Correlation	-.28	1.00	-.57*	-.34	-.01	-.29
	Sig. (2-tailed)	.337	.	.032	.236	.965	.323
	N	14	14	14	14	14	14
3. Eng Prof	Pearson Correlation	.28	-.57*	1.00	.43	.73**	.55*
	Sig. (2-tailed)	.332	.032	.	.129	.003	.043
	N	14	14	14	14	14	14
4. SS1 pretest	Pearson Correlation	.12	-.34	.43	1.00	.76**	-.16
	Sig. (2-tailed)	.697	.236	.129	.	.001	.596
	N	14	14	14	14	14	14
5. SS1 post-test	Pearson Correlation	-.13	-.01	.73**	.76**	1.00	.52
	Sig. (2-tailed)	.650	.965	.003	.001	.	.057
	N	14	14	14	14	14	14
6. SS1 change	Pearson Correlation	-.36	-.29	.55*	-.16	.52	1.00
	Sig. (2-tailed)	.213	.323	.043	.596	.057	.
	N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; SS1 pretest: *Scoring Sheet 1: Story Grammar Units* pretest score; SS1 post-test: *Scoring Sheet 1: Story Grammar Units* post-test score; SS1 change: *Scoring Sheet 1: Story Grammar Units* change score.

Table 39B. Spearman's Correlations between Scores in Scoring Sheet 1: Story Grammar Units and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

			1	2	3	4	5	6
Spearman's rho	1. DLD Risk	Correlation Coefficient	1.00	-.22	.27	.06	-.24	-.39
		Sig. (2-tailed)	.	.449	.361	.850	.404	.167
		N	14	14	14	14	14	14
	2. Eng Exp	Correlation Coefficient	-.22	1.00	-.51	.06	-.15	-.34
		Sig. (2-tailed)	.449	.	.065	.838	.599	.232
		N	14	14	14	14	14	14
	3. Eng Prof	Correlation Coefficient	.27	-.51	1.00	.15	.40	.49
		Sig. (2-tailed)	.361	.065	.	.612	.161	.073
		N	14	14	14	14	14	14
	4. SS1 pretest	Correlation Coefficient	.06	.06	.15	1.00	.71**	-.25
		Sig. (2-tailed)	.850	.838	.612	.	.004	.394
		N	14	14	14	14	14	14
	5. SS1 post-test	Correlation Coefficient	-.24	-.15	.40	.71**	1.00	.42
		Sig. (2-tailed)	.404	.599	.161	.004	.	.140
		N	14	14	14	14	14	14
	6. SS1 change	Correlation Coefficient	-.39	-.34	.49	-.25	.42	1.00
		Sig. (2-tailed)	.167	.232	.073	.394	.140	.
		N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; SS1 pretest: *Scoring Sheet 1: Story Grammar Units* pretest score; SS1 post-test: *Scoring Sheet 1: Story Grammar Units* post-test score; SS1 change: *Scoring Sheet 1: Story Grammar Units* change score.

Appendix 40. Spearman's Correlation Matrix between Scores in Scoring Sheet 2: Episodic Structure Complexity and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Table 40A. Spearman's Correlations between Scores in Scoring Sheet 2: Episodic Structure Complexity and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

		1	2	3	4	5	6	
Spearman's rho	1. DLD Risk	Correlation Coefficient	1.00	-.22	.27	-.06	.32	.34
		Sig. (2-tailed)	.	.449	.361	.842	.273	.230
		N	14	14	14	14	14	14
	2. Eng Exp	Correlation Coefficient	-.22	1.00	-.51	-.14	-.03	-.09
		Sig. (2-tailed)	.449	.	.065	.636	.910	.757
		N	14	14	14	14	14	14
3. Eng Prof	Correlation Coefficient	.27	-.51	1.00	.11	.57*	.54*	
		Sig. (2-tailed)	.361	.065	.	.709	.035	.045
		N	14	14	14	14	14	14
	4. SS2 pretest	Correlation Coefficient	-.06	-.14	.11	1.00	.41	-.45
		Sig. (2-tailed)	.842	.636	.709	.	.145	.105
		N	14	14	14	14	14	14
5. SS2 post-test	Correlation Coefficient	.32	-.03	.57*	.41	1.00	.57*	
		Sig. (2-tailed)	.273	.910	.035	.145	.	.034
		N	14	14	14	14	14	14
	6. SS2 change	Correlation Coefficient	.34	-.09	.54*	-.45	.57*	1.00
		Sig. (2-tailed)	.230	.757	.045	.105	.034	.
		N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; SS2 pretest: *Scoring Sheet 2: Episodic Structure Complexity* pretest rating; SS2 post-test: *Scoring Sheet 2: Episodic Structure Complexity* post-test rating; SS2 change: *Scoring Sheet 2: Episodic Structure Complexity* change score.

Appendix 41. Pearson's and Spearman's Correlation Matrices between Scores in Scoring Sheet 3: Emotional Vocabulary and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Table 41A. Pearson's Correlations between Scores in Scoring Sheet 3: Emotional Vocabulary and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

		1	2	3	4	5	6
1. DLD Risk	Pearson Correlation	1.00	-.28	.28	-.24	-.54*	-.26
	Sig. (2-tailed)	.	.337	.332	.410	.045	.378
	N	14	14	14	14	14	14
2. Eng Exp	Pearson Correlation	-.28	1.00	-.57*	.34	-.01	-.29
	Sig. (2-tailed)	.337	.	.032	.236	.965	.323
	N	14	14	14	14	14	14
3. Eng Prof	Pearson Correlation	.28	-.57*	1.00	-.14	.30	.36
	Sig. (2-tailed)	.332	.032	.	.632	.300	.204
	N	14	14	14	14	14	14
4. SS3 pretest	Pearson Correlation	-.24	.34	-.14	1.00	.26	-.60*
	Sig. (2-tailed)	.410	.236	.632	.	.376	.024
	N	14	14	14	14	14	14
5. SS3 post-test	Pearson Correlation	-.54*	-.01	.30	.26	1.00	.62*
	Sig. (2-tailed)	.045	.965	.300	.376	.	.017
	N	14	14	14	14	14	14
6. SS3 change	Pearson Correlation	-.26	-.29	.36	-.60*	.62*	1.00
	Sig. (2-tailed)	.378	.323	.204	.024	.017	.
	N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; SS3 pretest: *Scoring Sheet 3: Emotional Vocabulary* pretest score; SS3 post-test: *Scoring Sheet 3: Emotional Vocabulary* post-test score; SS3 change: *Scoring Sheet 3: Emotional Vocabulary* change score.

Table 41B. Spearman's Correlations between Scores in Scoring Sheet 3: Emotional Vocabulary and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

			1	2	3	4	5	6
Spearman's rho	1. DLD Risk	Correlation Coefficient	1.00	-.22	.27	-.25	-.45	-.21
		Sig. (2-tailed)	.	.449	.361	.385	.104	.472
		N	14	14	14	14	14	14
	2. Eng Exp	Correlation Coefficient	-.22	1.00	-.51	.28	-.02	-.28
		Sig. (2-tailed)	.449	.	.065	.326	.959	.327
		N	14	14	14	14	14	14
	3. Eng Prof	Correlation Coefficient	.27	-.51	1.00	-.13	.19	.33
		Sig. (2-tailed)	.361	.065	.	.669	.523	.251
		N	14	14	14	14	14	14
	4. SS3 pretest	Correlation Coefficient	-.25	.28	-.13	1.00	.14	-.59*
		Sig. (2-tailed)	.385	.326	.669	.	.624	.026
		N	14	14	14	14	14	14
	5. SS3 post-test	Correlation Coefficient	-.45	-.02	.19	.14	1.00	.68**
		Sig. (2-tailed)	.104	.959	.523	.624	.	.008
		N	14	14	14	14	14	14
	6. SS3 change	Correlation Coefficient	-.21	-.28	.33	-.59*	.68**	1.00
		Sig. (2-tailed)	.472	.327	.251	.026	.008	.
		N	14	14	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; SS3 pretest: *Scoring Sheet 3: Emotional Vocabulary* pretest score; SS3 post-test: *Scoring Sheet 3: Emotional Vocabulary* post-test score; SS3 change: *Scoring Sheet 3: Emotional Vocabulary* change score.

Appendix 42. Pearson's and Spearman's Correlation Matrices between Scores in Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

Table 42A. Pearson's Correlations between Scores in the Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

		1	2	3	4
1. DLD Risk	Pearson Correlation	1.00	-.28	.28	-.34
	Sig. (2-tailed)		.337	.332	.234
	N	14	14	14	14
2. Eng Exp	Pearson Correlation	-.28	1.00	-.57*	-.36
	Sig. (2-tailed)	.337		.032	.202
	N	14	14	14	14
3. Eng Prof	Pearson Correlation	.28	-.57*	1.00	.35
	Sig. (2-tailed)	.332	.032		.215
	N	14	14	14	14
4. Modifiability	Pearson Correlation	-.34	-.36	.35	1.00
	Sig. (2-tailed)	.234	.202	.215	
	N	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; Modifiability: *Modifiability Rating Scale* score.

Table 42B. Spearman's Correlations between Scores in Modifiability Rating Scale and the DLD Risk Factors Score, English Language Experience Score, and English Language Proficiency Stage

			1	2	3	4
Spearman's rho	1. DLD Risk	Correlation Coefficient	1.00	-.22	.27	-.23
		Sig. (2-tailed)	.	.449	.361	.423
		N	14	14	14	14
		Correlation Coefficient	-.22	1.00	-.51	-.25
	2. Eng Exp	Sig. (2-tailed)	.449	.	.065	.387
		N	14	14	14	14
		Correlation Coefficient	.27	-.51	1.00	.38
	3. Eng Prof	Sig. (2-tailed)	.361	.065	.	.187
		N	14	14	14	14
		Correlation Coefficient	-.23	-.25	.38	1.00
	4. Modifiability	Sig. (2-tailed)	.423	.387	.187	.
		N	14	14	14	14

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). Abbreviations: DLD Risk: *DLD Risk Factors Score*; Eng Exp: *English Language Experience Score*; Eng Prof: *English Language Proficiency Stage*; Modifiability: *Modifiability Rating Scale* score.

