‘Fluidity of language profiles in a follow-up study: Is early gesture predictive of later language profile?’

A research paper submitted in partial fulfilment of the requirements for the award of

MSc (Professional Qualification) in Speech and Language Therapy

University of Limerick, May 2015

Submitted by: Caitriona MacNally (13012096)

Supervisors: Dr. Carol-Anne Murphy and Hilary O’Neill

Final word count: 7,709

Proposed journal of submission: Journal of Speech, Language, and Hearing Research
# Table of Contents

Acknowledgements ........................................................................................................ 02

Abstract .......................................................................................................................... 03

Introduction .................................................................................................................... 04

Methodology .................................................................................................................. 10

Results ........................................................................................................................... 15

Discussion ....................................................................................................................... 28

References ...................................................................................................................... 35
Acknowledgements

I would like to thank the children and parents who participated in this project for their time and enthusiasm.

To my friends and research partners, Charlotte, Keira and Roz, thank you for your energy, hard work and humour during the last few months.

To Hilary O’Neill, thank you for your support and guidance during the project.

To Dr. Carol-Anne Murphy, thank you for your continued knowledge, guidance and patience on this project and throughout the course. It is very greatly appreciated.

Finally, I would like to sincerely thank my family and friends for their support and love throughout this project and during the last two years.
Abstract

Many studies have investigated whether non-verbal skills such as gesture are reliable early markers of later language delay, for example, Rowe and Goldin-Meadow (2009); Chiat and Roy (2008).

Researchers such as Reilly et al (2009) have looked at the value of following children longitudinally to see how communicative and expressive language profiles develop and change over time and to investigate which clinical markers appear to predict language profiles at a later stage. However a lack of longitudinal research specifically related to early gesture as a clinical marker of later language profile exists in the literature. This Time 2 (T2) study evaluates the relationship between early gesture at Time 1 (T1) and language profiles two years further along the language development trajectory.

Early non-verbal skills including gesture use and the ability to process meaning from symbolic representation were assessed in a cohort of 22 children aged 2-3 years at T1. At T2 follow-up, receptive and expressive language skills, phonological skills, non-verbal cognitive skills and non-word and sentence repetition were assessed in a group of 15 children aged 48-70 months with a previously identified receptive and/or expressive language delay from the same original T1 cohort of 22 children.

Results showed that language development appears to be fluid. Language profile at T2 changed for 86.6% of the T1 sample. Children with mixed receptive and expressive language difficulties at age 2-3 were more likely to have persisting language problems at age 4-5. Gesture did not appear to predict language profile or persistence of language impairment at age 4-5. Symbolic comprehension and morphosyntactic elements of language showed moderate correlations with expressive language. This small sample is very limited in terms of clinical significance and further research with a larger sample is necessary.

Keywords: Gesture, symbolic comprehension, language development trajectory, fluidity of profile, clinical markers of language impairment
Introduction
Speech and Language therapists are keen to identify early predictors or clinical markers of language which may be used at an early age to detect children at risk of language impairment (Reilly et al 2009). This would allow these children to be prioritised for earlier intervention than those who present with clinical markers suggestive of a language profile more likely to resolve with time. Early identification of speech and language requirements is crucial so that intervention resources are directed at those children who need them most and whose language impairments are least likely to resolve with age. Although we know that language develops at different rates for different children and that the language development trajectory is fluid, we do not currently know enough about clinical markers which predict long-term language outcomes for a young child. Children who present with a receptive language difficulty at age 2 may present with an expressive language difficulty at age 4. Others may persist in having a receptive language difficulty at age 4. Others still may resolve and may not present with any language difficulty at age 4. Research on reliable clinical markers that may be used to predict later language profile is therefore very important so that early intervention services can be better targeted towards the children who will benefit most from them.

Transient language difficulties versus persistent language difficulties
Brown (1973) found that onset and rate of language development varies widely in children. Bishop and Edmundson (1987) found that a large proportion of children with Expressive Language Delay had resolved and reached typical development within 1-2 years of initial referral. Building on this, Dale et al (2003) highlighted the importance of differentiating children with transient language difficulties from those with persistent language difficulties. Chiat and Roy (2013) highlight the clinical importance of distinguishing children who are ‘late bloomers’ from pre-schoolers at risk of language impairment in terms of both early identification and gaining more insight into the heterogeneous language profiles evident.

Provision of Speech and Language Therapy
Paul (2001) stated that it is inefficient if not unethical to provide speech and language therapy to children whose difficulties are likely to resolve. In addition, the Irish
Association of Speech and Language Therapists (IASLT) Code of Ethics states that it is the clinician’s responsibility to evaluate the effectiveness of intervention and to recommend intervention only when benefit can reasonably be expected (IASLT 2006). According to HSE figures from December 2013, 16,000 children were on waiting lists for an initial speech and language assessment in Ireland that month and a further 16,423 were waiting for therapy with 18 percent of those waiting for one to two years for intervention (Irish Times, March 2014).

The need for better research on clinical markers of later language profile

In light of these findings, more research on reliable clinical markers of persistent language impairment and predictors of specific later language profiles is indicated so that tight resources can be better targeted where needed. Previous triage criteria for language intervention used theories such as the ‘watch and see’ multi-modal assessment approach advocated by Paul (1996). This approach involves acknowledging the fluid nature and individual rates of language development, observing changes in a child’s language profile over time and adapting the intervention programme as necessary. Reilly et al (2009) also advocate the need to allow for fluidity in language profile and the need to watch and see how children’s language development develops when recommending intervention. The ‘watch and see’ approach is not practical however in a system where resources are tight and where a child may receive one eight-week block of therapy in a year and then not be seen again for a year afterwards. If we could clinically predict which children were more likely to resolve and which children showed signs of certain language profiles that would really benefit from speech and language therapy, we could provide targeted intervention for receptive or expressive language needs at an earlier age.

Gesture as a clinical marker

Various research has shown gesture to be a clinical marker of language development. Children first use deictic gestures, such as pointing, showing and referring to an external object, typically at 8-10 months according to Bates and Dick (2002), who found that children who used more gesture had greater comprehension of vocabulary than children who used less gesture. Acredolo and Goodywn (1988) noted that infants’ use of gesture indicates that an underlying linguistic relational structure already exists,
that is, the recognition that objects have names. Hulit and Howard (2002) found that phonological development begins with crying, vocalisations, babble and gesture. Iverson and Goldin-Meadow (2005) found gesture to be a predictor of later vocabulary development, where specific words that will become part of a child’s lexicon can be predicted by the child’s use of gesture. They found that, on average, children used a gesture to represent a word around 3 months before producing a verbal word for the object. They also established that increased use of gesture at 1-2 years indicated increased vocabulary at 3-4 years. Various research has shown gesture to be a precursor of both language development and phonological development. Gesture has also been shown to be a significant indicator of language delay in children (Chiat and Roy 2008; Rowe and Goldin-Meadow 2009). Gesture is significantly associated with receptive language abilities (Thal et al 1991).

The need to sub-group children according to language profile in order to prioritise for early intervention

Desmarais et al (2008, p.385) recommended that future studies of ‘late talkers’ should ‘include children from a clinical sample and document the various communication variables involved in language delay that have to do with both expressive and receptive skills’ in order to identify subgroups of children with distinct language impairment profiles.

Desmarais et al (2008) highlight the tendency in previous literature to give ‘late talkers’ an overarching diagnosis of expressive vocabulary delay at age 2 and they argue that future research should allow for the lack of homogeneity in a typical sample of 2 year old children with vocabulary delay. Dale et al (2003) looked at sub-grouping children according to language profile for more effective screening and early intervention, given that it is not efficient or cost-effective to provide early intervention across the board for children presenting with Expressive Language Impairment if we can find a way to clinically differentiate between those children whose language difficulty is likely to resolve and those whose language difficulty is likely to persist. Desmarais et al (2008) also highlight the potential value of using clinical markers to predict later language impairment. Early intervention for language difficulty is very important for several reasons. Early identification and intervention helps children who are at risk of
developmental delay to be more prepared for the linguistic and academic demands of school (Shonkoff 2003). In addition, children at risk of developmental issues and language delay who receive early intervention services are more socially and emotionally competent in early academic settings and mix better with peers than their counterparts who did not receive early intervention (Brophy-Herb et al 2007). The positive impact of early intervention on a child’s development and subsequent academic achievement presents a compelling argument for targeted screening using reliable clinical markers.

The issue of fluidity in the language development trajectory
A few longitudinal studies have demonstrated that although subgroups appear to exist in language profiles throughout school years, children who make up these subgroups move fluidly between differing types of language difficulties over time (Conti-Ramsden et al 1999). According to Chiat and Roy (2013), with each year from age 2 to age 5, roughly half of the children with delay will resolve and move into the normal range of language development. Reilly et al (2009) have investigated the fluid trajectory of language development, whereby children can present with very different language profiles at age 2 and at age 4. However, unlike Chiat and Roy (2008), Reilly et al (2009) have not examined fluidity using gesture as a specific clinical marker. As such, there is a gap in longitudinal research on the potential value of gesture as a reliable clinical marker which may point to those children who are more likely to resolve and those who are more likely to have persisting language difficulties at age 4. If more reliable clinical markers could be established, 2 year old children with profiles which are more likely to persist could be prioritised for early intervention. This would lead to more efficient and cost effective therapy and better outcomes.

Basis for current study
Reilly et al (2009) state that language impairment can be better defined at age 4 when language profiles develop further and when children whose language is likely to resolve have had time to do so. Longitudinal research following the same cohort of children studied at age 2 is therefore important.

A T1 study by O’Neill and Chiat (2012) found a link between gesture use and language profiles in a cohort of 22 typically developing participants age 2-3. The T1 study looked
at gesture use and symbolic skills in two subgroups of children with language delay. The two subgroups were RELD (Receptive Expressive Language Delay) and ELD (Expressive Language Delay). In the T1 study, gesture use and symbolic comprehension were assessed using the MacArthur Bates Communicative Development Inventories (MDI; Fenson et al 2007). The Pre-School Language Scale 3 (PLS-3 UK; Zimmerman et al 1997) was used to measure receptive and expressive language skills. Feldman (2000) highlights the need to use measures of word comprehension and word production in association with measuring gesture, that is, measuring gesture in isolation as a predictor of language delay is not sufficient. Chiat and Roy (2013) found the Pre-School Language Scale to be the most comprehensive measure of language and communication skills in pre-school children.

Results of the T1 study showed that gesture use and symbolic comprehension were significantly correlated with receptive language. Gesture and symbolic comprehension were not as significantly correlated with expressive language. This contradicts the findings of Dale et al (2003), who found correlations between early gesture and expressive language. This disparity indicates the need for further research on gesture as a clinical marker of later language profiles. In regard to symbolic comprehension, Chiat and Roy (2008) hypothesised that children with poor symbolic comprehension, that is, children who have difficulty deciphering meaning from other people’s use of symbols will have difficulties understanding the purpose of communication and deciphering the meaning behind words which will affect their language development.

The Early Sociocognitive Battery (Roy and Chiat 2006) administered at Time 1 looked at early responsiveness to symbolic representation.

This T2 longitudinal study investigates long term outcomes for the same group of children to see if gesture has predictive value for profiles at age 4-5. We aim to examine if gesture and symbolic comprehension scores at T1 are related to different language profiles at T2, and if gesture at age 2 is significantly related to expressive, receptive, mixed or resolved profiles at age 4. Our aim is to establish if gesture and symbolic comprehension can reliably be used as clinical markers of later language profile. Measures of receptive and expressive language, phonological and non-verbal
cognitive ability will be used in the T2 study in addition to the measures of gesture and symbolic comprehension taken at T1.
Methodology

Ethical approval
Ethical approval for the study was obtained from the Health Service Executive Regional Ethics Committee and City University London Ethics Committee, to which the T1 study was submitted.

Design
This study was designed as a T2 longitudinal cohort study, which aimed to follow the original T1 sample of 22 participants. Mean age in months at T1 was 27.3, SD=2.26. Only 15 of the T1 cohort chose to participate in the T2 follow up study. Mean age in months at T2 was 54.43, SD=3.6.

Participants
Referral criteria for participation in this study was for children to have participated in the original Time 1 (T1) study in 2012, be from monolingual English speaking families and have had an identified receptive and/or expressive language delay. Exclusionary criteria included any identified learning disability, behavioural disturbance, neurological and/or social emotional impairment.

At T1, the children were recruited through Speech and Language Therapists and Public Health Nurses working in a community based clinic service in the East of Ireland, whom the researcher met with to present the criteria for participation. The Speech and Language Therapists and Public Health Nurses then identified children who met the criteria from their existing caseloads, and provided their parents with an information sheet on the study. Parents contacted the researcher if they wanted their children to take part in the study.

At T2, one of the two principal investigators (O’Neill) contacted the parents of the original 22 participants by mail with information on the follow up study. Parents of participants had a period of time between receiving the information about the study and giving consent to participate. Again, parents contacted the researcher if they wished to participate in the T2 follow up study. Fifteen of the original 22 children were recruited and assessed. Consent forms were signed by a parent in the first session. Participants’ anonymity was protected. When assessments had been completed,
participants were number coded and were identified in the SPSS data file by their number ID and chronological age only. Sessions were recorded with parental consent for improved scoring accuracy and inter-rater reliability. Audio files were stored on a password protected computer.

While no strict nonverbal IQ was required to participate in the T1 study, a measure of nonverbal cognitive ability was introduced at T2 to rule out cognitive issues. However, the stability of IQ measures within this age group of children is problematic (Gilliam et al 2005) and it is widely recognised that cognitive abilities change rapidly in the early years.

Aims and Objectives
The primary aims of this follow up study were:

1. To investigate if early gesture and symbolic comprehension are predictive of persisting language impairment.
2. To investigate if early gesture and symbolic comprehension are reliable clinical markers of later language profile.
3. To add to the body of existing research on early clinical markers of later language profile.

These aims were addressed through following up a cohort of children approximately two years after they had initially been assessed.

Procedure
Before the language assessments were administered, the lead researcher conducted telephone interviews with a parent of each participant. Information gathered included language progress and SLT input since the T1 assessments.

Testing took place in February-March 2015. Each child was tested individually over two 45-75 minute sessions in a quiet unfamiliar clinic room with a parent present. Where the child did not complete the PLS-4 UK in one session, an additional third session was required. Each child sat at a small table directly across from the researcher, with parents seated next to the child. Assessments in each session were administered in a set order by one student researcher and double marked by another student researcher.
for inter-rater reliability. Inter-rater reliability is the degree to which multiple assessors make consistent judgements (Dane 2010).

The first session began with a warm-up activity to develop rapport between the child and the researcher. Games and stickers were provided at the end of sessions and during sessions where waning motivation levels were observed. In the first session, the PLS-4 UK Auditory Comprehension and Expressive Communication were administered.

During the second session, the DEAP, ERB and Raven’s Progressive Coloured Matrices were administered. The pace of each session was determined by the child and an activity was terminated if the child lost considerable interest or displayed signs of upset. Overall high cooperation was achieved and the participants were well motivated to complete assessments. At the end of the assessment stage, parents were provided with a Speech and Language Therapy report, giving details of assessment results and progress in the sessions.

**Measures**

**Figure 1: Table of assessments used at Time 1 and at Time 2**

<table>
<thead>
<tr>
<th>Assessments used at T1</th>
<th>Authors</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Assessments used at T2</th>
<th>Authors</th>
</tr>
</thead>
</table>

**Parental report measures of language ability**

Parental opinion of their child’s language was assessed prior to the first assessment session though a telephone interview. Parental opinion of language ability was further assessed using The Children’s Communication Checklist-2 (Bishop 2006), a
standardised measure designed to be completed by an adult who has regular contact with the child, typically a parent. The CCC-2 consists of 70 questions and aims to screen for children who are likely to have a language impairment, to identify pragmatic difficulties in children with communicative problems and to assist in identifying children who merit further assessment for an autistic spectrum disorder. The CCC-2 was completed by a parent of each participant during the first session. Gilliam et al (2005) note that whenever possible, any professionally generated screening data for young children should be combined with parental and teacher information, including first hand observations, to generate the most comprehensive and useful results.

**Direct measures of language**
Language was assessed using the Preschool Language Scale-Fourth Edition (PLS-4 UK) (Zimmerman et al 2007). This test is a standardized and norm-referenced evaluation tool for assessing receptive and expressive language skills in infants and young children (2 weeks to 6 years, 11 months old). The Auditory Comprehension component measures receptive language and the Expressive Communication component measures expressive language. Chiat and Roy (2013) found the Auditory Comprehension measure of the PLS 3-UK to be the best all round predictor of language and communication outcomes.

**Direct measures of speech and phonological processing skills**
Phonological ability was assessed using the phonological subtest of the Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al 2002). This is an assessment which investigates phonological skills. The mean standard score is 10 with the average range from 7-13. A score of 6 or less was used to classify phonological impairment. The Irish standardisation of the DEAP was used, which allows for acceptable Irish dialectal speech sound substitutions.

The Early Repetition Battery (ERB; Seeff-Gabriel et al 2008) (M=100, 1.25 SD=3.75) was administered to further assess underlying phonological processing skills. The ERB is a UK-developed and standardised test battery, which assesses phonological and morphosyntactic processing abilities in children aged 2 to 6 years. Non-word and sentence repetition tasks have been identified as clinical markers of Specific Language Impairments (SLI), (Coady and Evans 2008). The test consists of two expressive tasks,
PSRep (Preschool Repetition Test) which assesses repetition of single words and an SIT (Sentence Imitation Test), which assesses repetition of sentences.

**Measure of non-verbal cognitive ability**
A Ravens Coloured Progressive Matrices (CPM; Raven 1998) was used as a measure of non-verbal cognitive ability. This test is designed to provide a brief non-verbal measure of general ability and can be administered by Speech and Language Therapists. The test is made up of diagrammatic puzzles that are designed to assess intellectual processes of young children from 4 to 11 years. The mean standard score is 100, with any score between 85 and 115 considered to be within the average range of ability. As the participants of this study were at the very lower recommended age for this assessment and as some of them became distracted and did not complete the assessment, raw scores were used.

**Analysis**
Due to the small sample size and small number of measures taken, only simple correlational analysis could be undertaken. To better determine which clinical markers might predict later outcomes, a larger sample size and more in-depth analyses using regression analyses would be required.

**Sub-grouping at T1**
In the T1 study, participants were grouped into two subgroups: ELI (Expressive Language Delay) and RELD (Receptive and Expressive Language Delay), with a cut-off criterion for delay of at least -1.0 standard deviations below the mean of the reference population (Mean=100, SD=15; Zimmerman et al 1997).

**Sub-grouping at T2**
At T2, participants were allocated to one of four subgroups: Receptive impairment only, Receptive Expressive Language Impairment (RELI), Expressive Language Impairment (ELI) and No impairment. For classification purposes, participants were classified as impaired in one or more modality (receptive or expressive language) using a cut-off score of -1.25 standard deviations below the mean on the relevant domain of the PLS-4 UK (M=100, SD=18.75) (see Figure 2). The literature on what standard deviation below the mean needs to be present to diagnose impairment varies, with a current recommendation of -1.25 suggested in Reilly et al (2014), who discusses the
recommendations by the Iowa cohort of using -1.25 SD to diagnose language impairment within ‘socially defined functional expectations’ (Tomblin 2008, p. 95). For most studies the cut-off is between 1 and 2 standard deviations below the mean (Tomblin et al 2008). Lower prevalence of language problems, for example SLI, are found in studies with more stringent cut-off criteria, such as language skills at least 2 standard deviations below the mean (Williams 2002).

**Figure 2: Criteria for classification of subgroups at T2**

<table>
<thead>
<tr>
<th>RELI</th>
<th>Score of 81 or less on PLS 4-UK Receptive and Expressive</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELI</td>
<td>Score of 81 or less on PLS 4-UK Receptive; Score of 82 or more on PLS 4-UK Expressive</td>
</tr>
<tr>
<td>Receptive only</td>
<td>Score of 81 or less on PLS 4-UK Expressive; Score of 82 or more on PLS 4-UK Receptive</td>
</tr>
<tr>
<td>No impairment</td>
<td>Score of 82 or more on PLS 4-UK Receptive and Expressive</td>
</tr>
</tbody>
</table>

**Data analysis**

Data was analysed using the Statistical Package for the Social Sciences (SPSS), Version 22 (IBM 2013). Data was tested for normality. Descriptive statistics were obtained for all assessments used at Time 2. Spearman’s Rho correlations were run (a) between Time 1 and Time 2 variables and (b) between Time 2 variables to ascertain if statistically significant correlations existed between variables.

**Results**

This study was designed to longitudinally follow up the language profiles at T2 of a cohort of children studied at T1 and to explore the potential relationships between gesture, symbolic comprehension and later language profiles. Descriptive statistics for all assessment measures used at T2 are presented in Figure 3 below.
Figure 3: Mean, standard deviation (SD), minimum and maximum scores according to T2 group

<table>
<thead>
<tr>
<th>Classification T2</th>
<th>PLS Receptive T2</th>
<th>PLS Expressive T2</th>
<th>DEAP scaled score</th>
<th>ERB Word and Non-word Standard Score</th>
<th>ERB Sentence Processing Standard Score</th>
<th>Raven's CPM Raw Score</th>
<th>CCC-2 score</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Impairment N</td>
<td>92.86</td>
<td>90.57</td>
<td>5.00</td>
<td>91.17</td>
<td>85.17</td>
<td>100.00</td>
<td>70.67</td>
</tr>
<tr>
<td>SD</td>
<td>5.014</td>
<td>5.192</td>
<td>2.769</td>
<td>17.634</td>
<td>15.224</td>
<td>21.448</td>
<td>14.949</td>
</tr>
<tr>
<td>Min.</td>
<td>84</td>
<td>83</td>
<td>3</td>
<td>75</td>
<td>66</td>
<td>75</td>
<td>49</td>
</tr>
<tr>
<td>Max.</td>
<td>98</td>
<td>96</td>
<td>10</td>
<td>124</td>
<td>108</td>
<td>135</td>
<td>92</td>
</tr>
<tr>
<td>RELI N</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>72.50</td>
<td>64.50</td>
<td>6.50</td>
<td>93.00</td>
<td>68.00</td>
<td>80.00</td>
<td>68.00</td>
</tr>
<tr>
<td>Min.</td>
<td>66</td>
<td>54</td>
<td>4</td>
<td>91</td>
<td>66</td>
<td>80</td>
<td>68</td>
</tr>
<tr>
<td>Max.</td>
<td>79</td>
<td>75</td>
<td>9</td>
<td>95</td>
<td>70</td>
<td>80</td>
<td>68</td>
</tr>
<tr>
<td>ELI N</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>95.50</td>
<td>68.50</td>
<td>3.00</td>
<td>63.50</td>
<td>65.00</td>
<td>105.50</td>
<td>42.00</td>
</tr>
<tr>
<td>Min.</td>
<td>95</td>
<td>60</td>
<td>3</td>
<td>52</td>
<td>63</td>
<td>105</td>
<td>42</td>
</tr>
<tr>
<td>Max.</td>
<td>96</td>
<td>77</td>
<td>3</td>
<td>75</td>
<td>67</td>
<td>106</td>
<td>42</td>
</tr>
<tr>
<td>Receptive only N</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Mean</td>
<td>74.75</td>
<td>86.50</td>
<td>4.00</td>
<td>79.00</td>
<td>79.50</td>
<td>101.33</td>
<td>56.50</td>
</tr>
<tr>
<td>SD</td>
<td>3.304</td>
<td>2.646</td>
<td>2.000</td>
<td>5.657</td>
<td>9.192</td>
<td>2.309</td>
<td>7.853</td>
</tr>
<tr>
<td>Min.</td>
<td>71</td>
<td>83</td>
<td>3</td>
<td>75</td>
<td>73</td>
<td>100</td>
<td>47</td>
</tr>
<tr>
<td>Max.</td>
<td>79</td>
<td>89</td>
<td>7</td>
<td>83</td>
<td>86</td>
<td>104</td>
<td>66</td>
</tr>
<tr>
<td>Total N</td>
<td>15</td>
<td>15</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Mean</td>
<td>85.67</td>
<td>83.07</td>
<td>4.79</td>
<td>84.83</td>
<td>78.00</td>
<td>99.58</td>
<td>63.33</td>
</tr>
<tr>
<td>Min.</td>
<td>66</td>
<td>54</td>
<td>3</td>
<td>52</td>
<td>63</td>
<td>75</td>
<td>42</td>
</tr>
<tr>
<td>Max.</td>
<td>98</td>
<td>96</td>
<td>10</td>
<td>124</td>
<td>108</td>
<td>135</td>
<td>92</td>
</tr>
</tbody>
</table>
Figure 4: Histogram showing negatively skewed non-parametric data (PLS Receptive T2)

Figure 5: Histogram showing non-parametric data (DEAP PCC T2)
Testing for normality
Data was tested for normality using a Shapiro Wilk test due to the small sample size of less than fifty participants (Pallant 2013). The p-value for most data tested was <0.05, indicating that data was not normally distributed. For data with a p-value >0.05, the Q-Q plot which verifies the Shapiro Wilk test further was viewed and showed that although p-value was >0.05, the data was not normally distributed. Histograms (see Figures 4 and 5 above) for most variables showed negative skew, indicating that variables were not normally distributed and therefore indicating that non-parametric testing should be used.

More measured change in expressive language abilities than in receptive language abilities over time in this group of children
A Wilcoxon signed rank test was used to test whether there was a statistically significant difference or not between the results of assessments administered at T1 and again at T2, that is, PLS-4 UK Receptive and Expressive measures. This indicated no statistically significant difference between Receptive Language standard scores at T1 and Receptive Language standard scores at T2, $z = -.22$, $p > .05$, with a small effect size ($r = -.06$). The median score on the Receptive Language standard score decreased from T1 (Md= 93) to T2 (Md=89).

The Wilcoxon Signed Rank Test revealed statistical significance between Expressive Language standard scores at T1 and T2, $z = -2.814$, with a large effect size ($r = -0.73$). The median score on the Expressive Language measure increased from T1 (Md = 72) to T2 (Md = 86). Based on within group comparisons over time, there was more measured change in expressive language abilities over time in this group of children than in receptive abilities. Whilst the group at T2 differed in age from 51 to 70 months tests used allowed for age in scoring procedure so age is not a confounding variable.

Fluidity of profile
Results outlined in Figure 6 below show that the language development trajectory appears to be fluid for this cohort of children at T2. The nature of language impairment at T2 changed for 86.6% of the 15 T2 participants. Some of the T1 ELI group presented with receptive language difficulties at T2 and some of the T1 RELI group presented
with expressive only or receptive only language difficulties at T2, further demonstrating that children can present with a very different language profile at age 2 and at age 4 and move fluidly between different types of language impairment.

Results of this study show (in line with previous research) that receptive language impairment is more likely to persist than expressive language impairment. Of the 6 Time 1 RELI group, all 6 still had some kind of persistent language impairment at T2. 4 of these had a receptive element in their impairment at T2. Of the 9 Time 1 ELI group, 7 showed no persistent impairment at Time 2, that is, their language impairment had resolved. Unusually, 2 of 9 now had a Receptive Only impairment. This could be due to a number of factors such as poor test compliance at T1 or T2 or perhaps due to the increased demands of testing at an older age.

Figure 6: Fluidity of language profile subgroups from T1 to T2

Key:

<table>
<thead>
<tr>
<th>RELI</th>
<th>Receptive Expressive Language Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELI</td>
<td>Expressive Only Language Impairment</td>
</tr>
<tr>
<td>ROLI</td>
<td>Receptive Only Language Impairment</td>
</tr>
<tr>
<td>No imp</td>
<td>No language impairment</td>
</tr>
<tr>
<td>SP</td>
<td>Speech (Deap PCCSS &lt; 7)</td>
</tr>
</tbody>
</table>

Relationship between gesture at Time 1 and language profile at T2

Spearman’s rho correlations were run in SPSS to look at the relationship between non-parametric variables. Scatter plots were also used to examine correlations.
Correlations between T1 and T2 variables
A moderate correlation was found between Symbolic Comprehension T1 and ERB content word score T2 \((r=0.586, p=0.00)\) (see figure 7 above).

Moderate correlations were found between PLS Receptive T1 and ERB sentence repetition T2 \((r=0.643, p=0.24)\) and PLS Expressive T1 and ERB sentence repetition T2 \((r=0.696, p=0.012)\). Moderate correlations were also found between Distal Gesture T1 and ERB sentence repetition T2 \((r=0.635, p=0.026)\) and Conventional Gesture T1 and ERB sentence repetition T2 \((r=0.593, p=0.042)\). A moderate correlation was also found between Symbolic Comprehension T1 and PLS Receptive T2 \((r=0.412, p=0.127)\).

Correlations between T2 variables
A strong correlation was found between PLS Expressive T2 and ERB Content word score \((r=0.874, p=0.001)\). Moderate correlations were found between PLS Expressive score at T2 and ERB Sentence Repetition \((r=0.629, p=0.028)\), PLS Expressive T2 and ERB function word score \((r=0.691, p=0.27)\) and PLS Expressive T2 and ERB inflection word score \((r=0.439, p=0.204)\).
Figure 8: Scatter plot showing no correlation between conventional gesture at T1 and PLS Receptive Language score at T2
Figure 9: Scatter plot showing no correlation between Symbolic Comprehension at T1 and PLS Expressive Language score at T2.

Boxplots
Boxplots indicated higher Symbolic Comprehension scores at T1 for the T2 No impairment group. No similar effect was found for Conventional or Distal gesture at T1.
Figure 10: Box plots showing scores for symbolic comprehension at T1 in T2 subgroups
Figure 11: Box plots showing scores for distal gesture at T1 in T2 subgroups
Phonology scores at T2 for T1 RELI and ELI

Three of the T1 Receptive Expressive group and seven of the T1 Expressive group had phonological difficulties at T2.

The T1 RELD group had a slightly higher mean DEAP score (5.2) than the T1 ELI group (4.56) at T2 (see Figure 13).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEAP PCC scaled score</td>
<td>5</td>
<td>3</td>
<td>9</td>
<td>5.20</td>
<td>2.683</td>
</tr>
<tr>
<td>ERB word+nonword standard score</td>
<td>4</td>
<td>52</td>
<td>95</td>
<td>78.25</td>
<td>19.517</td>
</tr>
<tr>
<td>PLS Expressive Language Time 1</td>
<td>6</td>
<td>61</td>
<td>72</td>
<td>69.00</td>
<td>4.336</td>
</tr>
</tbody>
</table>
**Figure 14: Descriptive statistics for T1 ELI group**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEAP PCC scaled score</td>
<td>9</td>
<td>3</td>
<td>10</td>
<td>4.56</td>
<td>2.555</td>
</tr>
<tr>
<td>ERB word+nonword</td>
<td>8</td>
<td>75</td>
<td>124</td>
<td>88.13</td>
<td>16.075</td>
</tr>
<tr>
<td>standard score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLS Expressive Language</td>
<td>9</td>
<td>67</td>
<td>82</td>
<td>73.89</td>
<td>4.936</td>
</tr>
<tr>
<td>Time 1</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEAP PCC scaled score</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No correlations found between cognitive and phonological processing measures

Studies such as Crosbie et al (2009) have linked underlying cognitive linguistic skills such as non-verbal cognitive skills (measured by Raven’s Progressive Coloured Matrices) to phonological development but this study did not find any strong correlations between non-verbal ability and phonological ability (see Figures 15 and 16; r=0.406, r=0.281).
Figure 15: Moderate correlation found between Raven’s CPM and DEAP PCC at Time 2, $r=0.406$
Discussion
This longitudinal study is useful in that it looks at the relationship of gesture and symbolic comprehension to receptive and expressive language at Time 1 and looks at the relationship of these variables to language profiles at Time 2. This is the first study to look longitudinally at this particular range of assessments as predictors of later language profile.

The aims of this research project were to investigate if early gesture and symbolic comprehension predict persistence of certain types of language impairment and if early gesture and symbolic comprehension can be used as reliable clinical markers of later language profile. This discussion will look at the findings of the study,
comparisons to other studies, limitations that affected this study, clinical implications and recommendations.

Findings

More measured change in expressive language abilities over time
This study found more measured change in expressive language abilities over time in this group of children than in receptive abilities. This corresponds to existing literature which cites poorer prognosis for early receptive difficulties and the fact that receptive language difficulties are more likely to persist (Thal et al 1991; Chiat and Roy 2008).

Language development appears to be fluid in the early years
This study did not find a significant link between early gesture at 2-3 years and language profile at 4-5 years. However it did reveal the very fluid nature of the language development trajectory. This study therefore contributes to existing literature such as Dale et al (2003) regarding the fluid nature of language development.

Clinical markers
1). Symbolic comprehension: The ‘No Impairment’ group at Time 2 had significantly higher symbolic comprehension scores at Time 1 (See Figure 10), suggesting that higher symbolic comprehension at age 2-3 may indicate a greater likelihood that language difficulties will resolve by age 4-5. This suggests that early symbolic comprehension could be a clinical marker of an early language profile which may be likely to resolve. The very small sample of participants used in this study is too small to state this conclusively and the test would need to be repeated in a larger sample. This finding correlates with that of Chiat and Roy (2008), however, who found symbolic comprehension measured on the Early Sociocognitive Battery (ESB; Roy and Chiat 2006) (the same test used at our T1 to measure symbolic comprehension) to be the strongest predictor of social communication outcome and a relatively strong predictor of language outcomes.

2). Morphosyntactic clinical markers (Content word, Function word and Inflection word) and Sentence Repetition: Within Time 2 variables, the morphosyntactic elements of the Early Repetition Battery (ERB; Seeff-Gabriel et al 2008) were found to
be positively correlated with expressive language measures. A strong correlation was found between PLS Expressive T2 and ERB Content word score \( (r=0.874, p=0.001) \). Moderate correlations were found between PLS Expressive score at T2 and ERB Sentence Repetition \( (r=0.629, p=0.028) \), PLS Expressive T2 and ERB function word score \( (r=0.691, p=0.027) \) and PLS Expressive T2 and ERB inflection word score \( (r=0.439, p=0.204) \). This suggests that Sentence Repetition and morphosyntactic elements such as Content, Function and Inflection words may be clinical markers of a strong expressive language profile. These findings link to recent research on the ‘phonological hypothesis’ regarding Specific Language Impairment (Chiat and Roy 2008). Children with Specific Language Impairment (SLI) have been found to have difficulty with aspects of morphosyntax that are phonologically challenging, especially function words and inflection words (Leonard 1998; Chiat 2001). Children with SLI also have difficulty with sentence repetition tasks (Devescovi and Caselli 2007), with unstressed function words being particularly difficult for them (Seeff-Gabriel et al 2005). These findings led to the development of the Early Repetition Battery (ERB; 2008) which we used in this study to measure morphosyntactic elements. In addition to findings that phonological factors influence morphosyntactic difficulties, several studies have shown that children with SLI have difficulty with non-word repetition due to problems with phonological processing and memory (Gathercole 2006; Chiat 2006).

**Poor expressive language at T1 linked to weak phonology at T2**
Reviewing the outcomes for the T1 ELI group indicates that children with expressive only language difficulties at 2-3 years who have not resolved may present with phonological difficulties at 4-5 years (See Figures 13 and 14). There is a very small difference in the mean scores for both groups and a very small sample of 14 children completed DEAP assessments in this T2 study. This would need to be further investigated in a larger cohort.

**Comparisons with other studies**
The findings of this study suggest that gesture alone is not a sensitive enough measure to reliably predict later language impairment. This is similar to the findings of Nelson et al (2006), Zubrick et al (2007) and Reilly et al (2009), who found that the clinical predictors they investigated, such as socio-economic disadvantage, non-English
speaking background, gender, twining, gestation and birth order should not reliably be used alone to screen for language delay. Clinical markers need to be investigated in conjunction with standardised measures of receptive and expressive language.

The findings of this study are also in line with those of Chiat and Roy (2008) who found that language measures which draw on a wide range of skills were the strongest overall predictors of general language outcomes. Chiat and Roy (2008) found that receptive and expressive language skills measured on the PLS 3-UK were a stronger predictor of language outcome than early processing skills (VEPS) such as social responsiveness, joint attention and symbolic comprehension. This fits with the findings of our T2 study, which found the results of the PLS 4-UK Auditory Comprehension and Expressive Communication tests to be more reliable predictors of language profile than early gesture and symbolic comprehension measures, such that children who achieved low receptive language scores at T1 were more likely to have persistent language impairment at T2. Chiat and Roy (2008) concluded that a general language test such as the PLS, in particular the Auditory Comprehension scale, is a good starting-point for assessment, as performance on this test is broadly predictive of language profile.

**Recommendations**

Further research is indicated in line with Law et al (2008)'s assertion that a minimum of three observations of each participant are needed at regularly spaced time intervals for valid longitudinal profiling. That would mean that this T2 cohort of 15 children should be assessed again two years from now when they are aged 6-7 years.

This would allow researchers to further investigate fluidity of the language development trajectory, that is, to see if the four T2 subgroups have remained the same or if profiles have changed again. It would also be recommended to reassess phonological skills in two years at T3 when phonological profiles become more stable and when more delayed speech processes may be evident.

It would be important to establish at T3 if any diagnoses such as Specific Language Impairment (SLI) or Autistic Spectrum Disorder (ASD) have been made, in which case
the relationship of early gesture to these diagnoses could be investigated and more clinical usefulness of gesture as a clinical marker may be established.

**Clinical implications**
This study could be repeated in a larger cohort for more predictive outcomes. Given the -1.25 standard deviation cut-off in this study, it is worth considering that informal clinician based observations should be taken into account as well as what the results of standardised assessments indicate. Pool and Hourcade (2011 p. 273) note that ‘there may be legitimate concerns for those children who score slightly above identified cut-off scores (such as between one and -1.25 standard deviations below the mean as in the current study), as such scores may be indicative of limited skill repertoires’. Pool and Hourcade (2011) suggest a monitoring process for children whose scores are just above cut-off criteria and not low enough to indicate immediate intervention or referral. A monitoring policy would of course depend on the service delivery pathway in which a clinician works, but a monitoring period such as this where the child is scheduled for review in six months time for example would help avoid overlooking a developmental delay which may not have fully established at the time of assessment.

**Limitations**

**Attrition**
Attrition is one of the main limitations of a longitudinal study as the participants who drop out may differ from those who remain in the study. This reduces generalisability of findings and leads to biased estimates of means and biased associations between variables (Gustavson et al 2002; Farrington 1991). This T2 study was affected by attrition as only 15 of the 22 T1 participants took part in the T2 study.

**Small sample size**
The very small sample of 15 children in this Time 2 study is extremely limited in terms of predictive outcome. Further research with a larger sample is necessary. 7 of the original cohort did not return which led to an unbalanced sample. There is a possibility that parents of participants who returned felt that their children were less likely to have resolved since T1 and may have been influenced by the possibility of further intervention as an outcome of the T2 study, even though it was clearly noted that the study was solely for assessment and research purposes. There were limited repeated
observations of each child. For valid longitudinal research and to obtain effective trajectories using growth curve analysis, Law et al (2008) state that a minimum of three observations of each participant are needed at regularly spaced time intervals. Only 15 of the original 22 participants took part in the T2 study, which reduces the validity of the longitudinal findings. These 15 children would need to be followed up a third time two years after the T2 study to obtain an effective trajectory.

**Unequal levels of intervention from T1 to T2**
The T2 cohort had had varied amounts of therapy input since T1 and therefore they did not represent an evenly matched sample. Glascoe (2003) notes that child development changes rapidly during the birth to five year period, with this development greatly impacted by environmental factors and learning experiences, including language intervention.

**Observation points in the longitudinal study are predetermined**
This can be a disadvantage of a longitudinal study in that what happens between T1 and T2 observation points is not observed. This study has tried to offset this factor by conducting parent interviews at T2 on developmental progress, language progress, and Speech and Language input since T1.

**Compliance with assessments**
Some of the participants did not complete the full test battery. This was because the assessment battery was quite long and some participants lost attention or became distressed so assessments were discontinued which led to missing data. In general, testing is often not a natural experience for young children (Gilliam et al 2005). Traditional assessments may be administered by unfamiliar people and may require a very young child to sustain attention to adult-directed instructions and tasks (McLean and Crais 2004) Only 14 of the T2 participants completed the Diagnostic Evaluation of Articulation and Phonology (DEAP) assessment and only 12 of them completed the Early Repetition Battery (ERB) and Raven’s Coloured Progressive Matrices (CPM). This missing data with the small sample size, limits the potential of the study to find significant effects if there are any to be found.
The Sentence Repetition section of the Early Repetition Battery in particular relies on willingness to imitate, which is not always present in young children. It also relies on speech production (Gathercole 2006). Some of the participants had significant phonological difficulties which negatively impacted their ability to perform on the ERB word and non-word and sentence repetition tasks.

Conclusions
This study has demonstrated the value of studying the same cohort of children longitudinally as they get older to examine the relationship of early clinical markers to later language profiles. Despite its limitations, this study has pointed up the increased significance of morphosyntactic clinical markers of language profile as children get older and as their language develops and becomes more multifactorial, a point also highlighted by Chiat and Roy (2013). This study has corroborated the findings of Chiat and Roy (2008) that multifactorial language assessment which draws on a wide range of skills and language measures provides the strongest overall predictor of general language outcomes. It has also found early symbolic comprehension to be predictive of later language profile, similar to Roy and Chiat (2008), who found Very Early Processing Skills (VEPS) assessments to be strong predictors of language outcome. In terms of clinical markers, the potential value of symbolic comprehension at age 2-3 and morphosyntactic elements such as content, function and inflection words and sentence repetition at age 4-5 as clinical markers of language profile has been indicated in this very small sample and warrants further research in a larger sample.
References


