Measuring Change in Speech over Time:
Words or Segments?

A research project submitted to the Department of Clinical Therapies, University of Limerick, as part of the requirement for the Master of Science in Speech and Language Therapy

May 2015

Intended Journal for Submission:

International Journal of Language and Communication Disorders

Name: Bernadette Anne Ryan
Student Number: 11010894
Supervisor: Aileen Wright
Word Count: 8,248
Acknowledgements

I would like to thank my family for their constant love and support, not just throughout the past two years, but always.

A special and heartfelt thank you to friends, both near and far, for your patience, encouragement and love at all times on this journey.

To my fellow colleagues who were a pleasure to work with, in particular Jessica, thank you for the laughs along the way.

I would also like to extend my gratitude to the parents and children who participated in this study. Without their involvement, this research would not have been possible.

Finally, I would like to convey my sincerest appreciation to my supervisor, Aileen Wright, for her guidance, support and encouragement at all times. Thank you.
Abstract

Background: Measuring change is an integral part of the therapy process; what is measured and how it is measured validates intervention. Commonly in the literature, phonological analysis has been carried out using measures that focus on segments, e.g. Percent Consonants Correct (PCC) rather than whole words. However, more recent studies suggest that analysis of children’s phonological development should include measures that track change towards correct word forms, rather than just correctness of individual consonants.

Objective: To determine whether a measure of percentage of whole word proximity is a more sensitive measure of change in children’s developing phonological systems than PCC.

Method: Twelve children aged between 3;02 and 4;09 with SSDs participated in this study. At initial assessment speech accuracy in single words was assessed using a standardised test of phonology, the Diagnostic Evaluation of Articulation and Phonology (DEAP) (Dodd et al 2002). Speech accuracy in connected speech was assessed using a constrained speech sample elicited using pictures from the Renfrew Action Picture Test (RAPT) (Renfrew 2003). Both assessments were re-administered 8 weeks later. PCC and PWP were calculated for each assessment at initial and follow up assessment, and results compared to establish whether one measure was more sensitive to change in the samples than the other.

Results: No one measure was more sensitive than the other for all cases. Individual analyses exemplify how change is better captured by one method or the other in different cases.

Conclusions: The effectiveness of this study is limited by smaller than expected amounts of change over the two time points. However, analysis of individual cases where some change occurred showed that PCC was more accurate for some and PWP was more accurate for others. Clinical implications are discussed.

Keywords: Measuring change, speech accuracy, segments, whole words, Percentage Consonants Correct, Proportion of Whole-Word Proximity.
**Abbreviations:** SSD = Speech Sound Disorder; PCC = Percentage Consonants Correct; PWP = Proportion of Whole-Word Proximity; DEAP = Diagnostic Evaluation of Articulation and Phonology; RAPT = Renfrew Action Picture Test; SLT = Speech and Language Therapist; sSLT = Student Speech and Language Therapist.
## Contents

Acknoweldgements .................................................................................................................................................. 3
Abstract ............................................................................................................................................................... 4
List of Tables ......................................................................................................................................................... 8
List of Figures ....................................................................................................................................................... 8
Introduction .......................................................................................................................................................... 9
  Factors involved in Severity Determination: Subjective Measures ............................................................... 9
  Objective Measures: Segmental Measures ....................................................................................................... 10
  Alternative Segmental Measures ..................................................................................................................... 11
  Objective Measures: Whole Word Measures ................................................................................................ 12
  Alternative Whole Word Measures ............................................................................................................... 14
Conclusion ............................................................................................................................................................ 15
Aims of Investigation ......................................................................................................................................... 15
Research Hypothesis .......................................................................................................................................... 16
Methodology ......................................................................................................................................................... 17
  Ethical Approval and Consent .......................................................................................................................... 17
  Experimental Design .................................................................................................................................... 17
  Sourcing Participants ................................................................................................................................... 17
Pre-Treatment Assessment .................................................................................................................................. 18
Materials and Measures Used .......................................................................................................................... 18
  1. Diagnostic Evaluation of Articulation and Phonology: (DEAP; Dodd et al 2002) .............................. 18
  2. Constrained Connected Speech Sample .................................................................................................. 18
  3. Measures Used: ......................................................................................................................................... 19
Procedure .......................................................................................................................................................... 19
Post Intervention Assessment ............................................................................................................................ 19
Reliability .............................................................................................................................................................. 20
Results ................................................................................................................................................................... 22
Test of Normality .................................................................................................................................................. 22
Amount of change by the Group .......................................................................................................................... 23
Individual Analyses .......................................................................................................................................... 25
Children who made significant change in measures of Single Word accuracy ............................................ 26
  Child 1 (Time 1 Information: Male, CA: 3; 6, Phonological Delay) ......................................................... 26
  Child 3 (Time 1 Information: Female, CA: 4; 5, Consistent Phonological Disorder) ............................ 28
  Child 5 (Time 1 Information: Male, CA: 3; 2, Consistent Phonological Disorder) .................................. 30
List of Tables

Table 1: Characteristics and results at initial assessment and following home programme

Table 2: Correlation between PCC and PWP

Table 3: Difference between PCC and PWP in single words and connected speech

Table 4: Mean increase in PCC and PWP between T1 and T2

Table 5: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 1

Table 6: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 3

Table 7: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 5

Table 8: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 6

Table 9: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 8

Table 10: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 13

List of Figures

Figure 1: Amount of change in single words between T1 and T2

Figure 2: Amount of change in connected speech between T1 and T2

Figure 3: Change in PCC and PWP at single word level for Child 1

Figure 4: Change in PCC and PWP at single word level for Child 3

Figure 5: Change in PCC and PWP at single word level for Child 5

Figure 6: Change in PCC and PWP in connected speech for Child 6

Figure 7: Change in PCC and PWP in connected speech for Child 8

Figure 8: Change in PCC and PWP in connected speech for Child 9

Figure 9: Change in PCC and PWP in connected speech for Child 13
Introduction
Measuring change is a vital part of the therapy process; what is measured and how it is measured validates intervention. Therefore, a method sensitive enough to show even small amounts of change needs to be applied. Commonly in the literature, phonological analysis has been carried out using measures that focus on segments, e.g. Percent Consonants Correct (PCC) rather than whole words. However, more recent studies suggest that analysis of children’s phonological development should include measures that track change towards correct word forms, rather than just correctness of individual consonants.

Factors involved in Severity Determination: Subjective Measures
Identification of a speech sound disorder is usually followed by the determination of severity (Flipsen et al 2005). Traditionally, clinicians have used subjective measures to determine these levels. One way of categorising children with speech disorder is by labelling the severity of their disorder: ‘Mild’, ‘Moderate’, or ‘Severe’. However, the reliability of this method is dependent on clinical experience (Dodd 2005). Flipsen et al (2005) examined the use of subjective measures in clinical practice and found only ‘poor’ to ‘fair’ listener agreement between ten experienced SLT’s evaluating 17 speech samples, using a Mild/Moderate/Severe rating scale, highlighting the use of impressionistic judgements as problematic and among the lowest levels of evidence to be used in clinical decision making.

A further approach to determining severity is the use of ordinal rating scales, such as the Intelligibility in Context Scale (ICS) (McLeod et al 2012). The ICS is a parent-completed questionnaire, developed to supplement clinical measures of intelligibility. It provides a more comprehensive picture of the impact of SSD on a child’s intelligibility in daily life. Use of these scales provides information for clinicians about a child’s intelligibility to the average listener. Since intelligibility across contexts is the ultimate goal, this is important to measure. However, use of tools based on impressionistic judgements are not sensitive enough for clinical evaluation. They offer little information about the underlying disorder and raise questions around reliability, standardisation and validity (Flipsen et al 2005). Subsequently, objective tools of measurement may provide more robust methods of analysis for clinicians.
Objective Measures: Segmental Measures

Objective measures of children’s phonological development have commonly been aimed at the level of the individual speech sound by using various methods of segmental analysis. One of the most significant contributions came from Shriberg and Kwiatkowski (1982) with their introduction of the now widely used Percentage of Consonants Correct (PCC). PCC compares the child’s production to the adult target on a phoneme by phoneme basis. The number of correct consonants produced is divided by the total number of target consonants and then multiplied by 100 to give an overall percentage of consonants correct (Stoel-Gammon 2010) in a given speech sample. According to Burrows and Goldstein (2010), PCC calculates accuracy on all individual phonemes and is the most widely used segmental measure across the literature. It is also frequently cited as an index of severity in textbooks (e.g., Bauman-Waengler, 2004; Bernthal and Bankson, 2004) and in standardised assessments such as the Diagnostic Evaluation of Articulation and Phonology (DEAP) (Dodd et al 2002). Historically, Shriberg et al (1997) examined PCC in children of various ages, including: typically developing children, children with mild-to-severe SSD, and those with residual errors. The results presented significant differences in PCC between sub-sets differentiated by disorder. Authors therefore concluded that PCC was valid for measuring the accuracy of productions by children aged 3 to 6 years with SSDs (Burrows and Goldstein 2010).

The literature offers many examples of research citing this method of segmental analysis in support of Shriberg’s work. Most significantly, a randomised control trial carried out by Almost and Rosenbaum (1998) measured PCC in 30 children aged between 33 and 61 months with severe phonological disorders. The results presented significant differences in PCC between treated and untreated groups following therapy over a 4-month period, showing PCC to be a robust method of measurement at the phoneme level. These findings are substantiated in an efficacy study of two different types of therapy by Dodd et al (2005). More recently, the merits of PCC as a reliable indicator of change is supported in the literature by Newbold et al (2013).
Alternative Segmental Measures

In addition to PCC, a number of alternative segmental measures focused at the level of individual speech sound have been proposed. The majority of these were presented by Shriberg et al (1997) as a series of extensions to PCC, and have the potential to serve as indexes of severity (Flipsen et al 2005). These include Percentage of Vowels Correct (PVC) and Percentage of Phonemes Correct (PPC) - calculated in the same way as PCC. PVC and PPC have also been frequently cited in the literature (Dodd et al 2003) and in standardised assessments (Dodd et al 2002). However, inter-rater reliability for PCC and PPC tends to be higher than PVC, which may be a result of wider variation across transcription of vowels (Powell 2001). Furthermore, it is widely agreed that impairments at consonant level are more common that vowel errors, and measurement of same provides more valuable information for intervention.

More recently, Preston et al (2011) investigated the psychometric properties of a new measure entitled the ‘Weighted Speech Sound Accuracy’ (WSSA) with speech samples from children across a wide range of ages and populations. Similarly to PCC, sounds are vertically aligned to compare the child’s production with the adult target. However, authors claim that the new measure is more sensitive than PCC in that different types of errors are weighted differently. For example, phoneme omissions and unusual errors are weighted more heavily, but errors involving common substitutions are given smaller weights (Preston et al 2011). PCC was used as a standard for comparison and results showed that WSSA and PCC scores were highly correlated across 3 groups. Conversely, WSSA scores in Group 4 'Spontaneous speech of typically developing toddlers', showed only moderate correlations with PCC. This suggests that WSSA scores account for improvements in how well segments in the word are represented, whereas binary judgments of accuracy (PCC) penalise for even minor errors and may be less sensitive to small changes in phonetic accuracy. Results provide preliminary support for the WSSA as a valid and reliable measure of phonetic accuracy in children’s speech. However, further research to show clinical effectiveness is needed. Additionally, computer competency is required, data input can be time-consuming and software may be expensive and not accessible to all clinicians.
In contrast, PCC offers simplicity of calculation, ease of interpretation and psychometric properties (Preston et al. 2011). However, it is also not without limitations, and concerns about the sensitivity of PCC have been raised. Firstly, errors such as dentalisations or lateralisations are scored in the same way as any other error type, such as deletions, though it is less likely to have the same impact on intelligibility (Flipsen et al. 2005). Furthermore, lack of sensitivity for different error types as the child’s production gets closer to the adult target, poses a problem (Hall et al. 1998). For example, a child may omit a phoneme from the adult target (e.g. [kæb] for /kræb/) and then progress to substituting the omission with an incorrect phoneme (e.g. [kwæb] for /kræb/), thus bringing it closer to the adult target but still awarding the same score. PCC provides us with important information and is a reasonable starting point for discussing numeric quantification of children’s speech. However, an argument can be made for measures which allow tracking of change in the direction of better proximity to the whole word, rather than a pure binary distinction of accurate or not accurate. This is supported in more recent literature and discussed in the subsequent section.

**Objective Measures: Whole Word Measures**
Although most investigations of phonological development in children have focused on varying methods of segmental analysis, a whole word perspective has been identified in the literature as an important aspect to be considered. Velleman and Vihman (2002) propose that early word productions are lexically driven, with little awareness of individual speech segments and an attempt to replicate models in the environment. As children are exposed to more language they begin to notice regularities and patterns that they integrate into their own productions. This need to approximate whole words will continue to influence phonological development (Bunta et al. 2009). Ingram and Ingram (2001) discuss a number of issues around this topic, including the need to determine when significant changes in whole word correctness occur, and how these changes should be measured.

To this end, Ingram (2002) proposed a variety of related measures to calculate the phonological complexity of words and the child’s ability to match them accurately. These measures include: the Proportion of Whole-Word Correctness (PWC), Phonological Mean...
Length of Utterance (PMLU) and the Proportion of Whole-Word Proximity (PWP). According to Ingram (2002), PWC determines what proportions of the child’s words are produced correctly out of the entire vocabulary. Schmitt et al (1983 cited in Ingram 2002) studied PWC on conversational samples from children between three and seven years, and found PWC increased significantly between the ages of three and three-and-a-half years, going from 68% to 76%. However, while PWC contributes to intelligibility, it does not tell us anything about movement towards accuracy of the target as it counts whole words as simply either right or wrong. Therefore, sensitivity to progress is likely to be even less sensitive than PCC.

Proposed by Ingram and Ingram (2001) and further detailed in Ingram (2002), PMLU measures the length of a child’s words and the number of correct consonants produced. Firstly, the PMLU of the adult target is calculated, one point is received for each speech sound, and one point for each consonant. For example, in the adult target ‘crab’ or /kræb/, there are four segments /k/, /r/, /æ/ and b/, plus three consonants /k/, /r/ and /b/, giving a total target score of 7. To calculate the child’s PMLU, one point is received for each speech sound produced (regardless of accuracy), and one point for each correct consonant when compared to the adult target. In a child’s realisation of ‘crab’ as [kwæb], there are four segments [k], [w], [æ], and [b], plus two correct consonants, [k] and [b], giving a total score of 6 out of a possible 7. The PMLU is calculated by adding the total number of points assigned for a selected set of words and dividing that number by the number of words (Ingram and Ingram 2001). The calculation of the PMLU for the child’s target word provides us with an idea of the complexity of the words that the child is attempting as well as providing some idea of the relation between the child’s forms and their targets.

Ingram and Ingram (2001) argue that a mediating factor in phonological acquisition is the child’s need to approximate the adult target. Their measure, PWP, is a whole-word measure directly related to PMLU that focuses on what the child is capable of producing, in terms of proximity to the adult target. PWP is calculated by dividing the PMLU of the child’s words by the PMLU of the corresponding target words (Arias and Lleó 2014). Using
the example outlined above, the PMLU of the adult target /kræb/ gave a total target score of 7 and the PMLU of the child’s realisation of [kwæb] gave a total of 6 out of a possible 7. To calculate PWP the PMLU of the child’s word [kwæb] (6) would be divided by the PMLU of the adult target /kræb/ (7), resulting in a PWP of 6/7 or 86%. As with the PMLU, the overall PWP equals the average across the sample (Flipsen et al 2005). PMLU and PWP have emerged as the most commonly used metrics of whole word analysis in the literature and Ingram (2002) demonstrates its usefulness by applying it in a wide range of contexts, including: a comparison of monolingual children, a comparison across languages, and the diagnosis of an impairment or delay.

Data has been provided by other researchers using these whole word measures offered by Ingram and Ingram (2001) and Ingram (2002). Most recently, Newbold et al (2013) compared a range of commonly used measures of speech output including PCC and PWP on speech samples from children with significant speech sound difficulties at 4 years and again at 6 years. Results showed PCC and PWP to be the most accurate indicators of change. Furthermore, both PCC and PWP captured change over time and differentiated between children who made different amounts of progress. However, while PCC showed maximal change in outcomes, PWP was found to be more sensitive to individual differences such as structural changes (Newbold et al 2013). PCC and PWP are thus both potentially useful tools for evaluating speech outcomes.

Alternative Whole Word Measures
An independent measure proposed by Stoel-Gammon (2010) offers an alternative approach to assessing phonology by examining the phonetic complexity of children’s productions. In contrast to most current measures of phonological development, this approach focuses on the complexity of segments, syllables, and words (Stoel-Gammon 2010). The Word Complexity Measure (WCM), provides quantitative and qualitative information about productions and provides an alternative approach to the works of both Ingram and Shriberg. Each word is awarded a complexity score based on an approach that assesses different levels within the phonological system: word patterns, syllable structures, and sound classes (Stoel-Gammon 2010). Higher scores indicate the presence of more difficult or later acquired phonological parameters. Examples illustrating the
scoring procedures include the following: [bɛd] (target: bed): WCM = 1 point (final consonant) and [ɡwin] (target: green): WCM = 3 points (cluster + final consonant + velar) (Stoel-Gammon 2010). The WCM is a sensitive measure of change that may offer an appropriate assessment for younger children with small vocabularies or restricted inventories. However, authors acknowledge issues such as the parameters of measurement and optimal sample size of words need to be resolved. Further research showing clinical effectiveness and ease of use would provide beneficial information into the reliability of this new measure.

**Conclusion**
Objective methods of analysis provide clinicians with more valid and reliable means of measuring change in children’s speech accuracy. Across the literature, PCC and PWP have been cited as being potentially the most useful for evaluating outcomes. PCC has been shown to be a robust indicator of change, being less affected by the choice of stimuli. However, PMLU and PWP have been shown to capture change across time and tasks, while appearing to be more sensitive than PCC to small differences such as structural changes (Newbold et al 2013). Few studies, have compared PCC and PWP as methods of measuring change in children with speech sound disorders (Burrows and Goldstein 2010; Newbold et al 2013). Therefore, research is warranted to determine which method provides a more sensitive measurement of change.

**Aims of Investigation**
The purpose of the intended study is to analyse the speech samples of children with speech sound disorders at two time points using the metrics of PCC and PWP and compare these results to conclude which is the more sensitive measure of change.

Therefore, the following research question has been proposed:

Is PWP a more sensitive method of measuring change than PPC in children with speech sound disorders?
Research Hypothesis
The following is hypothesised:

Results from this study will demonstrate that PWP is a more sensitive method of measuring change than PCC as a measurement of speech accuracy in children with speech sound disorders.
Methodology

Ethical Approval and Consent
Ethical approval for this study was granted by the Research Ethics Committee University Hospital Limerick. Informed written consent was obtained from the parents prior to their child’s participation.

Experimental Design
Both a longitudinal and pre and post-intervention test study design was used, as this study was part of a larger one testing efficacy of an intervention. Participants were randomly allocated to a ‘treatment now’ or ‘treatment later’ group at the initial assessment stage, and all were reassessed following the 8 week intervention period. It was expected that the intervention programme would effect significant changes in speech accuracy in the ‘treatment now’ group. However, analysis of the data shows that significant change by the ‘treatment now’ group over the ‘treatment later’ group did not occur. In fact, change occurred for individuals in both groups. Therefore data from all participants were analysed together as one group for this study.

Sourcing Participants
Fourteen children participated in the study, however two were ultimately excluded. One was excluded as they did not meet the eligibility criteria and the other was excluded at as they did not comply with the assessment procedure. Therefore, twelve children (five girls and seven boys) were examined in the study.

The principal investigator contacted two local Speech and Language Therapy service managers who acted as gatekeepers and sourced children from their waiting lists. The SLT’s contacted the parents of children and sent them the information and consent forms. Parents who were interested in participating in the study returned the forms to the SLT, who subsequently forwarded these to the principal investigator.

Inclusionary criteria: monolingual English speaking children aged between 3;0 and 4;11 years, who were referred to local SLT services owing to difficulties with speech sounds and who were awaiting initial assessment. Exclusionary criteria: bilingual children and children who were already receiving therapy. Also excluded were children with speech
deficits due to structural or organic causes and children with an additional diagnosis of
cognitive or sensory difficulties.

After a screening call, appointments were arranged if participants were deemed to be
eligible. At the assessment phase, the age range was between 3;02 and 4;09, with a mean
age of 3;11. Five of the children had previously received intervention for their speech
disorder.

**Pre-Treatment Assessment**
Assessment appointments were given for two Speech and Language Therapy clinics in the
mid-western region of Ireland. Two student Speech and Language Therapists (sSLT) were
present throughout each assessment session. One sSLT assessed the child in a quiet clinic
room. Each assessment lasted approximately 45 minutes and a parent was present
throughout. Following administration of the tests, the child’s case history was obtained by
the second sSLT.

**Materials and Measures Used**

1. **Diagnostic Evaluation of Articulation and Phonology: (DEAP; Dodd et al 2002)**
   - The diagnostic screener of the DEAP was carried out to establish the
     characteristics of the child’s speech errors.
   - The phonology subtest was administered to provide the single word sample for
     the baseline measurements of PCC and PWP for each participant.
   - The assessment consisted of picture naming where each child was instructed to
     name 50 items.

2. **Constrained Connected Speech Sample**
   - The Renfrew Action Picture Test (RAPT; Renfrew 2003) was utilised to provide a
     baseline sample of connected speech.
   - The assessment consisted of picture description where the child was instructed to
     describe what was happening in the picture.
   - This provided context for eliciting the same words at both time points to allow for
     change to be mapped against similar targets as well as interpret unintelligible
     utterances.
3. Measures Used:

- PCC was calculated using the guidelines outlined by Shriberg and Kwiatkowski (1982) and used in the DEAP manual (Dodd et al 2002) (Appendix A).
- PWP was calculated using the guidelines set outlined by (Ingram 2002) (Appendix B).

A separate research project carried out a full analysis of the results obtained from the tests of non-word and real word discrimination (see Barrett 2015 and Quinlan 2015) as well as the FOCUS tool and Likert questionnaire (see Kelleher 2015, Kerrigan 2015 and Martin 2015). Compliance was also measured using a diary completed by the parent.

**Procedure**

The diagnostic screener of the DEAP (Dodd et al 2002) was carried out in the first order to establish the characteristics of the child’s speech errors. Speech accuracy in single words was assessed using the standardised test of phonology, the Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al 2002). Speech accuracy in connected speech was assessed using a constrained speech sample, elicited using pictures from the Renfrew Action Picture Test (RAPT; Renfrew 2003). The pictures from the RAPT provided context for eliciting the same words at both time points to allow for change to be mapped against similar targets, as well as interpret unintelligible responses. Both are tools normally used by local SLT services. A test of non-word discrimination and real word discrimination of the type commonly used in speech intervention studies were also administered, but results are not reported in this study. A case history as normally used by SLT services was also completed. Parents perceptions of severity of speech impairment, pre and post intervention, was collected using the FOCUS tool (Thomas-Stonell et al 2010), and acceptability of the programme was measured using a Likert questionnaire.

**Post Intervention Assessment**

On completion of the home programme, each child was invited back for a re-assessment of their speech skills. The phonology subtest of the DEAP and the RAPT were re-administered. The parents were asked to complete the post intervention FOCUS tool and a Likert questionnaire. A diary completed by the parent detailing compliance was also returned.
Reliability
Of the eight sSLTs who were the investigators for this research project, two were present for each assessment. One sSLT administered the assessments while the other observed. Both sSLTs transcribed the responses. The second sSLT collected the signed information sheets and consent forms, carried out the case history, and answered any questions that arose. Parents also gave consent for the session to be recorded by signing the audio release form.

Phonetic Transcriptions
All transcriptions were recorded using an Olympus VN-5500PC. Following the assessment session, transcriptions were compared. In the event of inconsistencies, both sSLTs listened to the recording and agreed upon a transcription. The same transcribers were used pre-intervention and post-intervention to avoid issues of inter-rater reliability.

Scoring Assessments
Both sSLTs scored assessments and compared results. Scores were also checked by 6 other sSLTs working on the project until consistent agreement was reached.
### Results

<table>
<thead>
<tr>
<th>Child</th>
<th>Gender</th>
<th>Group</th>
<th>Age (months)</th>
<th>Diagnosis</th>
<th>Single Words</th>
<th>Connected Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PCC Time 1</td>
<td>PWP Time 1</td>
</tr>
<tr>
<td>1</td>
<td>Male</td>
<td>Now</td>
<td>42</td>
<td>Delay</td>
<td>75</td>
<td>65</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>Later</td>
<td>45</td>
<td>Disorder</td>
<td>88</td>
<td>81</td>
</tr>
<tr>
<td>3</td>
<td>Female</td>
<td>Later</td>
<td>53</td>
<td>Disorder</td>
<td>71</td>
<td>71</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>Now</td>
<td>47</td>
<td>Disorder</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>Male</td>
<td>Now</td>
<td>38</td>
<td>Disorder</td>
<td>29</td>
<td>54</td>
</tr>
<tr>
<td>6</td>
<td>Female</td>
<td>Later</td>
<td>40</td>
<td>Disorder</td>
<td>55</td>
<td>77</td>
</tr>
<tr>
<td>7</td>
<td>Female</td>
<td>Later</td>
<td>47</td>
<td>Disorder</td>
<td>50</td>
<td>67</td>
</tr>
<tr>
<td>8</td>
<td>Female</td>
<td>Now</td>
<td>42</td>
<td>Delay</td>
<td>63</td>
<td>81</td>
</tr>
<tr>
<td>9</td>
<td>Female</td>
<td>Now</td>
<td>50</td>
<td>Disorder</td>
<td>63</td>
<td>79</td>
</tr>
<tr>
<td>10</td>
<td>Male</td>
<td>Later</td>
<td>49</td>
<td>Delay</td>
<td>64</td>
<td>81</td>
</tr>
<tr>
<td>11</td>
<td>Male</td>
<td>Now</td>
<td>44</td>
<td>Delay</td>
<td>80</td>
<td>88</td>
</tr>
<tr>
<td>13</td>
<td>Male</td>
<td>Later</td>
<td>53</td>
<td>Disorder</td>
<td>59</td>
<td>73</td>
</tr>
</tbody>
</table>

*Table 1: Characteristics and results at initial assessment and following the home programme for each participant*
Results
The results of the study are presented through descriptive analysis and both quantitative and qualitative measures. Table 1 (on the previous page) reports characteristics and results at initial assessment and following the home programme for each participant.

Test of Normality
Although a Shapiro Wilk test showed that the data was normally distributed, owing to the small sample size non-parametric tests were used for analysis.

Spearman’s Rank Order Correlation (rho) explores the direction and strength of the relationship between two continuous variables. This test was used to investigate whether there was an association between PCC and PWP measures in single word and connected speech samples. Results showed a strong correlation between measures which suggests they may be evaluating the same entity, i.e. both measuring speech accuracy. See Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Correlation Coefficient (r)</th>
<th>Significance Value (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Words Time 1</td>
<td>0.973</td>
<td>0</td>
</tr>
<tr>
<td>Single Words Time 2</td>
<td>0.970</td>
<td>0</td>
</tr>
<tr>
<td>Connected Speech Time 1</td>
<td>0.993</td>
<td>0</td>
</tr>
<tr>
<td>Connected Speech Time 2</td>
<td>0.935</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2: Correlation between PCC and PWP (Large correlation (r) = .50 to 1.0, Cohen 1988).

Conversely, considerable difference was seen between PCC and PWP in single words and connected speech at both time points. PWP scores were consistently higher than PCC scores, which indicated that it awards greater accuracy to productions. See Table 3 (on the next page).
Table 3: Difference between PCC and PWP in single words and connected speech T1 and T2.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean increase in PCC score between T1 and T2</th>
<th>Mean increase in PWP score between T1 and T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single word sample (DEAP)</td>
<td>2.08</td>
<td>2.00</td>
</tr>
<tr>
<td>Connected speech sample</td>
<td>0.75</td>
<td>1.50</td>
</tr>
</tbody>
</table>

Table 4: Mean increase in PCC and PWP between Time 1 and Time 2

Further investigation using a Wilcoxon Signed Rank Test examined whether the amount of change measured by PCC was significantly different from the amount measured by PWP. Results from this test demonstrated no significant difference in the amount of change measured in PCC in single words compared to PWP in single words (Z=-.205, \( p=.838 \)). Similarly, no significant difference was evident in the amount of change measured in PCC in connected speech when compared to PWP in connected speech (Z= -.472, \( p=.637 \)).
While examination of the means had shown that both PCC and PWP showed small amounts of change between samples at T1 and T2 (see table 3 above), neither measure showed significant change. Therefore either of the following conclusions could be reached:

i. PWP is no more sensitive than PCC as a measure of change in speech accuracy over time.

ii. There was not sufficient change demonstrated by the group to show differences in sensitivity between measures.

If no change can be seen we would not expect one measure to look more sensitive than the other. We can therefore conclude that the amount of change made at group level was not sufficient to distinguish the relative sensitivity of the two measures.

As no change was seen at group level, individual cases were investigated to find any that showed significant change, which would allow comparison of the sensitivity of the different measures.
Individual Analyses
A statistical analysis of significant change over time in each child’s speech accuracy was carried out using a two-tailed Fishers exact test. This confirmed that significant change could be seen at single word level on the DEAP assessment for Child 1, 3 and 5. See Figure 1.

![Figure 1: Change made by individual participants between T1 and T2 in single words](image1)

Additionally, significant change over time could be seen in connected speech using a constrained speech sample elicited from the RAPT assessment for Child 6, 8, 9 and 13. See Figure 2.

![Figure 2: Change made by individual participants between T1 and T2 in connected speech](image2)
Children who made significant change in measures of Single Word accuracy

Child 1 (Time 1 Information: Male, CA: 3; 6, Phonological Delay)
A two tailed Fishers exact test confirmed a significant increase in PWP at single word level ($p = 0.01093$). Despite the fact that the increase in PCC was bigger, this was not statistically significant. PCC increased by 9% from 75% to 84% and PWP increased by 5% from 88% to 93%. Therefore, PCC was more sensitive than PWP in capturing change over time for this child. See Figure 3.

![Figure 3: Change in PCC and PWP at single word level for Child 1](image)

The sensitivity of PCC to capture greater change than PWP is illustrated by the following example. At Time 1 (T1) the adult target 'elephant' or $/ɛləfənt/$ was produced as $[ɛləbənt]$. This rendered a score of 75% accuracy in PCC as there are three correct consonants [$l$, $n$] and [$t$], giving a total score of 3 out of a possible 4. 90% was awarded in PWP as there are 7 segments, plus 3 correct consonants, giving a total score of 10 out of a possible 11. At Time 2 (T2) the child's production matched the adult target. Therefore PCC changed from 75% to 100%, increasing by 25% as all 4 consonants were correct, giving a total of 4 out of 4. Conversely, PWP changed from 90% to 100%, increasing by only 10% in accuracy as there are 7 segments [$ɛ$, [$l$], [$ə$, $f$, [$a$], [$n$] and [$t$], plus 4 correct consonants [$l$], $[f]$, [$n$] and [$t$], giving a total of 11 out of 11. As PWP had already awarded partial credit for there being a phoneme in the correct position
(albeit the wrong phoneme) at T1, replacement of the incorrect phoneme [b] with the correct phoneme [t] at T2, could only warrant additional credit of 10% at T2. See Table 5.

<table>
<thead>
<tr>
<th>Productions</th>
<th>PCC</th>
<th>PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: [ɛləbənt]</td>
<td>3/4 = 75%</td>
<td>10/11 = 90%</td>
</tr>
<tr>
<td>T2: [ɛləfənt]</td>
<td>4/4 = 100%</td>
<td>11/11 = 100%</td>
</tr>
<tr>
<td>Change</td>
<td>25% ↑</td>
<td>10% ↑</td>
</tr>
</tbody>
</table>

*Table 5: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 1*

Additional examples of these types of changes are demonstrated in the child's production of the adult target /gλnɔz/ as [gλbð] at T1 and [gλbz] at T2, and again in the adult target /брɛd/ as [bwɛd] at T1 and [брɛd] at T2. As PWP had already awarded maximum marks for structure, correction of a substitution error afforded a smaller increase in score. This child corrected 11 substitution errors and 3 structural errors between Time 1 and Time 2 (see Appendix C for examples of substitution and structural errors). As PCC gives more credit for corrections of substitution errors than structural errors, it appears to be more sensitive to change for this child.
Child 3 (Time 1 Information: Female, CA: 4; 5, Consistent Phonological Disorder)
Child 3’s PCC increased by 11% from 71% to 82% and PWP increased by 5% from 87% to 93%. Statistical analysis using a two-tailed Fisher’s exact test confirmed a significant increase in PCC (p = 0.014291) and PWP (p = 0.03674) at single word level. Therefore, both PCC and PWP were shown to be sensitive to change over time. However, PCC was shown to be more sensitive at capturing change in this case. See Figure 4.

Figure 4: Change in PCC and PWP at single word level for Child 3

In the same way as Child 1, Child 3 was shown to reduce the presence of substitution errors from T1 to T2. This is exemplified by the following production. The adult target ‘train’ or /treɪn/ was produced as [tʃweɪn] at Time 1 (T1). This resulted in a score of 33% accuracy in PCC as there is one correct consonants [n], giving a total score of 1 out of 3. 71% was awarded in PWP as there are 4 segments, plus 1 correct consonant, giving a total score of 5 out of 7. At Time 2 (T2) /treɪn/ was produced as [tʃreɪn] which was closer to the adult target. This showed PCC changing from 33% to 67% increasing by 34% as there were 2 correct consonants, giving a total of 2 out of 3. On the other hand, PWP changed from 71% to 86%, increasing by only 15% in accuracy as there are 4 segments [t], [r], [ei] (diphthongs are counted as 1) and [n], plus 2 correct consonants [r] and [n], giving a total of 6 out of 7. Again, as PWP had already awarded partial credit for 2 phonemes in the correct position (although incorrect) at T1, replacement of the
incorrect phoneme [w] with the correct one [r] at T2, could only award additional credit of 15% at T2. See Table 6.

<table>
<thead>
<tr>
<th>Productions</th>
<th>PCC</th>
<th>PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: [tfwein]</td>
<td>1/3 = 33%</td>
<td>5/7 = 71%</td>
</tr>
<tr>
<td>T2 : [tfrein]</td>
<td>2/3 = 67%</td>
<td>6/7 = 86%</td>
</tr>
<tr>
<td>Change</td>
<td>34% ↑</td>
<td>15% ↑</td>
</tr>
</tbody>
</table>

*Table 6: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 3*

Supplementary examples of these types of changes are illustrated in the child's production of /tuθbrʌʃ/ as [tuθbwʌs] at T1 and [tuθbrʌs] at T2, and again in /skweə/ as [tʃkweə] at T1 and [skweə] at T2. Similarly to Child 1, this child corrected 12 substitution errors and 1 structural error from T1 to T2. Therefore, as PCC gives more credit for corrections of substitution errors than structural errors, it appears to be more sensitive to change in this case.
Child 5 (Time 1 Information: Male, CA: 3; 2, Consistent Phonological Disorder)
Child 5’s PCC increased by 4% from 29% to 33% and PWP increased by 8% from 54% to 62%. A two-tailed Fisher’s exact test confirmed a significant increase in PWP at single word level (p = 0.035611). Therefore, PWP was more sensitive than PCC in capturing change for this child. See Figure 5.

![Figure 5: Change in PCC and PWP at single word level for Child 5](image)

In contrast to Child 1 and Child 3, Child 5 was shown to reduce the presence of structural errors over time. This is demonstrated by the following example. At Time 1 (T1) the adult target ‘three’ or /θri/ was produced as [fi]. This offered a score of 0% accuracy in PCC as there were no correct consonants, giving a score of 0 out of 3. PWP awarded a score of 40% as there were 3 segments but no correct consonants, giving a total of 3 out of 5. At Time 2 (T2) the child’s production [fwi] was closer to the adult target. Again, this offered a score of 0% accuracy in PCC because although the child now marked 2 consonant places, neither was correct; consequently, PCC made no change. In contrast, PWP changed from 40% to 60%, increasing by 20% in accuracy as there are 3 segments [f], [w] and [i] but no correct consonants, giving a total of 3 out of 5.
This resulted in a greater change shown by PWP from T1 to T2, having awarded partial credit for there being one phoneme in the correct position [f] (although incorrect) at T1 and additional credit for 2 incorrect phonemes [f] and [w] at T2. PCC gained no credit at either time points. See Table 7.

<table>
<thead>
<tr>
<th>Productions</th>
<th>PCC</th>
<th>PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: [fi]</td>
<td>0/2 = 0%</td>
<td>2/5 = 40%</td>
</tr>
<tr>
<td>T2: [fwi]</td>
<td>0/2 = 0%</td>
<td>3/5 = 60%</td>
</tr>
<tr>
<td>Change</td>
<td>0%</td>
<td>20% ↑</td>
</tr>
</tbody>
</table>

*Table 7: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 5*

Further instances of decreasing structural errors include the child's production of the adult target /treɪn/ as [ti] at T1 and [tsein] at T2, and again in /fəɪv/ as [fai] at T1 and [faiz] at T2. Additionally, eight structural changes and one substitution are noted from T1 to T2. As a result, PWP was shown to be more sensitive to change in this case.
Children who made significant change in Connected Speech
Although connected speech samples were constrained, some children did not use the same vocabulary at both time points. This made finding illustrative examples for comparative data difficult in some instances.

Child 6 (Time 1 Information: Female, CA: 3; 4, Consistent Phonological Disorder)
A two tailed Fishers exact test confirmed a significant increase in PWP in connected speech (p=0.032913). Despite the fact that the increase in PCC was bigger, this was not statistically significant. PCC increased by 10% from 67% to 77% and PWP increased by 6% from 79% to 85%. Therefore, PCC was more sensitive than PWP in capturing change over time for this child. See Figure 6.

The sensitivity of PCC to capture greater change than PWP is illustrated by the following example. At Time 1 (T1) the adult target ‘cat’ or /kæt/ was produced as [kæx]. This rendered a score of 50% accuracy in PCC as there is one correct consonants [k], giving a total score of 1 out of a possible 2. 80% was awarded in PWP as there are 3 segments, plus 1 correct consonant, giving a total score of 4 out of a possible 5. At Time 2 (T2) the child's production matched the adult target. Therefore, PCC changed from 50% to 100% increasing by 50% as both consonants were correct, giving a total of 2 out of 2. Conversely, PWP changed from 80% to 100%,
increasing by only 20% in accuracy as there are 3 segments [k], [æ] and [t], plus 2 correct consonants [k] and [t], giving a total of 5 out of 5.

Exploration of assessment results showed specific improvements in reducing phoneme substitutions from T1 to T2. This resulted in a greater change captured by PCC from T1 to T2. As PWP had already awarded partial credit for a phoneme in the correct position (albeit the wrong phoneme) at T1, replacement of the incorrect phoneme [x] with the correct phoneme [t], could only warrant additional credit of 20% at T2. See Table 8.

<table>
<thead>
<tr>
<th>Productions</th>
<th>PCC</th>
<th>PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T1: [kæx]</strong></td>
<td>1/2 = 50%</td>
<td>4/5 = 80%</td>
</tr>
<tr>
<td><strong>T2: [kæt]</strong></td>
<td>2/2 = 100%</td>
<td>5/5 = 100%</td>
</tr>
<tr>
<td><strong>Change</strong></td>
<td>50% ↑</td>
<td>20% ↑</td>
</tr>
</tbody>
</table>

*Table 8: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 6*

Additional examples of this type of change is demonstrated in the child's production of the adult target ‘glasses’ or /glæsəz/ as [gwædəs] at T1 and [gwæsəz] at T2. Therefore, PCC appears to be more sensitive to change for this child.
Child 8 (Time 1 Information: Female, CA: 3; 6, Phonological Delay)
This child’s performance worsened between the two time points. PCC decreased by 9% from 82% to 73%, and PWP decreased by 8% from 90% to 82%. Statistical analysis using a two-tailed Fisher’s exact test confirmed a significant decrease in PWP in connected speech (p=0.006813). This was despite the fact that measures of change in PCC and PWP differed by only 1%. Therefore, both PCC and PWP are shown to be sensitive to change over time. However, PCC was seen to be more sensitive in capturing change by the smallest of margins in this case. See Figure 7.

![Figure 7: Change in PCC and PWP in connected speech for Child 8](image)

Illustrative examples that showed the difference in PCC and PWP were limited as this child did not use the same vocabulary at both time points in the connected speech sample. However, the increased sensitivity of PCC in capturing change is exemplified by the following production. At Time 1 (T1) the adult target ‘glasses’ or /glæsəz/ was produced as [glæsəs]. This rendered a score of 75% accuracy in PCC as there are 3 correct consonants [g], [l] and [s], giving a total score of 3 out of a possible 4. 90% was awarded in PWP as there are 6 segments, plus 3 correct consonants, giving a total of 9 out of a possible 10. At Time 2 (T2) the child’s performance got worse and the adult target /glæsəz/ was produced as [glæθə]. Therefore, PCC changed from 75% to 50%, decreasing by 25% as only 2 consonants were correct. In contrast, PWP changed...
from 90% to 70%, also decreasing, but by a smaller margin of 20% in accuracy as there are 5 segments [g], [l], [æ], [θ] and [ə], plus 2 correct consonants [g] and [l]. This smaller margin of change illustrates how PCC was seen to be somewhat more sensitive than PWP in this case. See Table 9.

<table>
<thead>
<tr>
<th>Productions</th>
<th>PCC</th>
<th>PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: [glaesə]</td>
<td>3/4 = 75%</td>
<td>9/10 = 90%</td>
</tr>
<tr>
<td>T2 : [glæθə]</td>
<td>2/4 = 50%</td>
<td>7/10 = 70%</td>
</tr>
</tbody>
</table>

*Table 9: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 8*

Consequently, PCC appears to be marginally more sensitive to capturing change in substitutions than structures from T1 to T2 for this child.
**Child 9 (Time 1 Information: Male, CA: 4; 2, Consistent Phonological Disorder)**

Child 9’s PCC increased by 1% from 72% to 73% and PWP increased by 7% from 80% to 87%. A two-tailed Fisher’s exact test confirmed a significant increase in PWP in connected speech ($p=0.04721$). Therefore, PWP was more sensitive than PCC in capturing change for this child. See Figure 8.

![Figure 8: Change in PCC and PWP in connected speech for Child 9](image)

This child did not use the same vocabulary at both time points, which makes a direct comparison of substitutions and structural errors impossible, for example, production of ‘fence’ or [fɛns] at T1 and ‘gate’ or [geɪt] at T2. The number of words elicited in the sample was also relatively small; 53 words at T1 and 47 words at T2. It may be possible that this child used words at T2 which were structurally simpler than at T1 or that he selected words that he knew he could say. This would account for why PWP was more sensitive in capturing change than PCC in this case.
**Child 13 (Time 1 Information: Male, CA: 4; 5, Consistent Phonological Disorder)**

A two-tailed Fisher’s exact test confirmed a significant change in PWP in connected speech ($p=0.021717$). Child 13’s PCC made no change with scores of 58% at both time points but PWP was shown to increase by 8% from 68% to 76%. Therefore, PWP was more sensitive than PCC in capturing in this case. See Figure 9.

![Figure 9: Change in PCC and PWP in connected speech for Child 13](image)

These findings are illustrated in the child’s production of the following word. At Time 1 (T1) the adult target ‘stairs’ or /steɪr/ was produced as [steə]. This offered a score of 50% accuracy in PCC as there were two correct consonants [s] and [t]. PWP awarded a score of 56% as there were 3 segments, plus 2 correct consonants. At Time 2 (T2) the child’s production [steɪrd] was closer to the adult target. Therefore PCC changed from 50% to 75% increasing by 25% as 3 consonants were correct [s], [t] and [r]. In contrast, PWP changed from 56% to 89%, increasing by 33% in accuracy as there are 5 segments [s], [t], [ei], [r] and [d], plus 4 correct consonants [s], [t], [r] and [d].

Examination of assessments showed an increase in the presence of substitutions and a decrease in structural errors from T1 to T2. This resulted in a greater change captured by PWP from T1 to T2; having not awarded any credit for the omission of 2 phonemes [r] and [z] at T1,
gained maximum credit for 1 correct phoneme [r] at T2 and partial credit for 1 phoneme in the correct position [d] (although incorrect) at T2. In contrast, PCC did not award any credit for the incorrect phoneme [d] at T2. See Table 10.

<table>
<thead>
<tr>
<th>Productions</th>
<th>PCC</th>
<th>PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: [stea]</td>
<td>2/4 = 50%</td>
<td>5/9 = 56%</td>
</tr>
<tr>
<td>T2 : [stelrd]</td>
<td>3/4 = 75%</td>
<td>8/9 = 89%</td>
</tr>
<tr>
<td>Change</td>
<td>25% ↑</td>
<td>44% ↑</td>
</tr>
</tbody>
</table>

*Table 10: Calculations of accuracy in PCC and PWP from T1 to T2 for Child 13*

Further examples of similar types of errors are illustrated in the child's production of the adult target /dʒʌmpt/ as [dʌmt] at T1 [dʌmpɛd] and at T2. As a result, PWP was shown to be more sensitive to change for this child.

To summarize, significant change over time was seen at single word level for Child 1, 3 and 5, and in connected speech for Child 6, 8, 9 and 13. Of these 7 children, 4 showed greater change in PCC than PWP whereas 3 showed greater change in PWP than PCC.

<table>
<thead>
<tr>
<th>Child</th>
<th>Number of words in sample at Time 1</th>
<th>Number of words in sample at Time 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>70</td>
<td>59</td>
</tr>
<tr>
<td>8</td>
<td>66</td>
<td>48</td>
</tr>
<tr>
<td>9</td>
<td>53</td>
<td>47</td>
</tr>
<tr>
<td>13</td>
<td>60</td>
<td>84</td>
</tr>
</tbody>
</table>

*Table 11: Number of words in connected speech sample at Time 1 and Time 2*
Discussion

The aim of this study was to analyse the speech samples of children with SSD at two time points and compare results to determine whether a measure of Proportion of Whole Word Proximity (PWP) is a more sensitive measure of change in children’s developing phonological systems than Percentage Consonants Correct (PCC). This was investigated by addressing the following question:

‘Is PWP a more sensitive method of measuring change than PCC in children with speech sound disorders?’

It was hypothesised that results from this study would demonstrate that PWP is a more sensitive method of measuring change than PCC in speech accuracy over time. However, the analysis does not support this hypothesis as no one measure was more sensitive than the other for all cases. Both PWP and PCC were shown to be differentially sensitive to different types of change in individual cases across the group. Results showed PCC to be more sensitive for capturing change in substitution errors, whereas PWP tended to be more sensitive in capturing change in structural errors.

As this sample showed no statistically significant change in PCC or PWP at group level, no conclusions can be reached about whether one measure is more sensitive than the other to change at group level. However, these results do suggest that group results will be affected by the characteristics of the children making up the group, or by the targets of therapy in an intervention study. For example, if the target of intervention is correcting structural errors, PWP would be a more sensitive measure for capturing change. In contrast, if the target of intervention is correcting substitution errors, PCC would be more sensitive for capturing change.
Current findings in relation to previous research
Results from a Test of Correlation showed a strong relationship between Percentage Consonants Correct (PCC) and Proportion of Whole Word Proximity (PWP) which suggests they are evaluating the same entity. This is consistent with findings of Burrows and Goldstein (2010) who found that that PCC, PMLU and PWP were all significantly and highly correlated to each other. Results from individual analysis illustrate how PCC was sensitive to change in some cases whereas PWP was sensitive to change in others. This is consistent with findings of Newbold, Stackhouse and Wells (2013) who found that both PCC and PWP were sensitive to capturing change over time. They concluded that while PCC showed maximal change in outcomes, PWP was shown to be more sensitive than PCC to small differences such as structural changes. Results of this study expand on these findings by showing that PCC was more sensitive for capturing change in substitutions while PWP was more sensitive to change in structural errors.

Limitations
Due to the small number of participants and the small amount of change made, results of the present study must be treated with caution. Speech accuracy in connected speech was assessed using a constrained speech sample elicited using pictures from the RAPT (Renfrew 2003). Shriberg and colleagues (1989) suggest a sample size of at least one hundred words in connected speech in order to provide a representative sample for phonological analysis. Samples analysed varied from 53 to 70 words at Time 1 and 47 to 84 words at Time 2. In particular, Child 9 did not use the same vocabulary at both time points meaning that variability in measures could be introduced depending on the child’s selection of vocabulary. As this sample was much smaller than the recommended amount, it is not certain that it reflected the same number of changes. This same limitation applies to all children in this sample as none of them reached the recommended number of words. See Table 11 at end of Results section.
**Indications for Further Research**

It is recommended that further studies investigating the sensitivity of Proportion of Whole Word Proximity (PWP) and Percentage Consonants Correct (PCC) be carried out on larger samples. Analysis of samples showing greater amounts of change across the group is also suggested. Lastly, connected speech samples of more than one hundred words would create a greater opportunity for comparative data over two time points.

**Conclusions and Clinical Implications**

Analysis of individual cases where significant change occurred showed that PCC was more sensitive for four of the participants, whereas PWP was more sensitive for three participants. PCC was more sensitive for capturing change in substitution errors, while PWP was more sensitive for capturing change in structural errors.

The clinical implications of these findings are significant as they may form the basis for clinical decision making when choosing outcome measures and different methods of measuring change specific to the child. To conclude, results indicate that PCC provides a more reliable measurement of change and more robust measure of outcomes for a child who produces words characterised by substitution errors. Conversely, for a child who produces words characterised by structural errors, PWP provides a more reliable measurement of change and more robust measure of outcomes.
Appendices

Appendix A: Rules for calculating Percent of Consonants Correct (PCC)

Appendix B: Rules for calculating Phonological Mean Length of Utterance (PMLU) and Proportion of Whole Word Proximity (PWP)

Appendix C: Examples of Substitution and Structural Errors
Appendix A

Rules for calculating Percent of Consonants Correct (PCC)

(Shriberg and Kwiatkowski 1982)

Determining which Consonants to Use

1. Use only the intended or target sounds in words.
2. When /r/ occurs after a vowel, it is considered to be a consonant. However, the stressed and unstressed ‘vocalic r’, are considered to be vowels.
3. If the child repeats a syllable, for example, ba-baby, score only the consonants in the first syllable.
4. Do not include consonants in words whose meaning is questionable or in words that are partly or completely unintelligible.
5. If a child repeats a word several times, score only the consonants in the first two, unless the articulation changes.

Scoring the consonants

1. Use the following response definition: ‘Score as incorrect unless clearly heard as correct’.
2. If the child speaks a dialect of English, that dialect serves as the standard in determining if the consonants are correct or not.
3. If the child’s production is comparable to what an adult would say in casual conversation, then the consonants should be scored against the standards for casual speech, for example the word and said as a syllabic /n/. Note that in adult casual speech, the /h/ in pronouns can be deleted unless the pronoun is stressed, for example, I called him as [alkoldIm] in contrast to the stressed version He called me as [hikaldmi]. Similarly, /ŋ/ may be said as /n in unstressed syllables as in singing [sɪŋn].
4. Incorrect productions errors include:
   - Deletion of a target consonant
• Substitutions
• Partial voicing of initial voiceless consonants
• Distortions of a target consonant, even if it is only a slight distortion

**Calculating PCC**

Count up the total number of consonant targets and the number that were correct. Divide the number correct by the total and multiple by 100 to determine the PCC.

**For example:**

In the adult target ‘crab’ or /kræb/, there are three consonants /k/, /r/, and b/, giving a target consonant score of 3.

If the child produces /kræb/ as [kwæb], there are two correct consonants [k] and [b], giving a total of 2 out of a possible 3.

To calculate PCC of the child’s word [kwæb] the number of correct consonants (2) would be divided by the number of target consonants for /kræb/ (3), and then multiplied by 100, resulting in a PCC of 2/3 or 67%.
Appendix B

Rules for calculating Phonological Mean Length of Utterance (PMLU) and Proportion of Whole Word Proximity (PWP)
(Ingram 2002)

PMLU measures the length of a child’s words and the number of correct consonants produced.

1. **Sample-Size Rule**: Select at least 25 words, and preferably 50 words for analysis, depending on sample size. If the sample is larger than 50 words, select a selection of words that cover the entire sample, e.g. every other word in a sample of 100 words.

2. **Lexical-Class Rule**: Count words (e.g. common nouns, verbs, adjectives, prepositions and adverbs) that are used in normal conversation between adults. This excludes child words, e.g. mommy, daddy, tata, etc. Counting child words can inflate the PMLU if a child is a reduplicator.

3. **Compound Rule**: Do not count compounds as a single word unless they are spelled as a single word, e.g. ‘cowboy’ but not ‘teddy bear’, i.e. ‘teddy bear’ would be excluded from the count. This rule simplifies decisions about what constitutes a word in the child’s sample.

4. **Variability Rule**: Only count a single production for each word. If more than one occurs, then count the most frequent one. If there is none, then count the last one produced. Counting variable productions may distort the count if there is a highly variable single word.

5. **Production Rule**: Count 1 point for each consonant and vowel that occurs in the child's production. Syllabic consonants receive one point, e.g. syllabic ‘l’, ‘r’, and ‘n’. (Some transcriptions may show these as two segments, i.e. a schwa plus consonant, e.g. ‘bottle’ [bædəl], but it should be counted as one consonantal segment). Do not count more segments than are in the adult word. For example, a child who says ‘foot’ as [hwut] has two
consonants counted, not three. Otherwise, children who add segments will get higher scores despite making errors.

6. **Consonants Correct Rule**: Assign 1 additional point for each correct consonant. Correctness in vowels is not counted since vowel transcriptions are typically of low reliability. Syllabic consonants receive an additional point in the same way as nonsyllabic consonants. A child who applies liquid simplification, for example, will get 1 point for producing a vowel, e.g. ‘bottle’ [bado], but 2 points if the syllabic consonant is correct.

7. **PWP**: is a whole-word measure directly related to PMLU that focuses on what the child is capable of producing, in terms of proximity to the adult target. PWP is calculated by dividing the PMLU of the child’s words by the PMLU of the corresponding target words.

For example:

Firstly, the PMLU of the adult target is calculated, one point is received for each speech sound, and one point for each consonant. In the adult target ‘crab’ or /kræb/, there are four segments /k/, /r/, /æ/ and b/, plus three consonants /k/, /r/ and /b/, giving a total target score of 7.

To calculate the child’s PMLU, one point is received for each speech sound produced (regardless of accuracy), and one point for each correct consonant when compared to the adult target. For example, in a child’s realisation of ‘crab’ as [kwæb], there are four segments [k], [w], [æ], and [b], plus two correct consonants, [k] and [b], giving a total of 6 out of a possible 7.

The PMLU of the adult target /kræb/ gave a total target score of 7 and the PMLU of the child’s realisation of [kwæb] gave a total of 6 out of a possible 7.

To calculate PWP the PMLU of the child’s word [kwæb], (6) would be divided by the PMLU of the adult target /kræb/ (7), resulting in a PWP of 6/7 or 86%.
Appendix C
Substitution and Structural Errors
(Dodd 2005)

Substitution Errors include the following:

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gliding</td>
<td>Replacement of liquids /l, r/ with glides [w, j]</td>
<td>[wæˈbit] rabbit, [ʃæm] lamb</td>
</tr>
<tr>
<td>Deaffrication</td>
<td>Deletion or replacement of affricates</td>
<td>[wʊʃ] watch, [brɪʤ] bridge</td>
</tr>
<tr>
<td>Fronting</td>
<td>Place of articulation is moved to a more anterior position</td>
<td>[ɛg] egg, [sɪp] ship</td>
</tr>
<tr>
<td>Stopping</td>
<td>Replacement of fricatives with stops</td>
<td>[bæn] van, [dɪbra] zebra</td>
</tr>
<tr>
<td>Voicing</td>
<td>Pre-vocalic voicing and post-vocalic voicing</td>
<td>[bæm] pram, [pɪk] pig</td>
</tr>
<tr>
<td>Assimilation</td>
<td>A substitution of one sound in a word to harmonize with another sound</td>
<td>[keɪki] Katie, [bɔbʌm] bottom</td>
</tr>
</tbody>
</table>

Structural Errors include the following:

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster reduction</td>
<td>Deletion of one consonant from the cluster</td>
<td>[paɪdɛ] spider, [bɛd] bread</td>
</tr>
<tr>
<td>Weak Syllable Deletion</td>
<td>Deletion of an unstressed syllable</td>
<td>[matɔʊ] tomatoes, [bɛlʌ] umbrella</td>
</tr>
<tr>
<td>Final-consonant deletion</td>
<td>Deletion or glottal replacement of some (but not all) word final consonants</td>
<td>[dɔ] doll, [bɔʔ] book</td>
</tr>
<tr>
<td>Initial-consonant deletion</td>
<td>The deletion of word- or syllable- initial consonants</td>
<td>[ɔl] ball</td>
</tr>
<tr>
<td>Medial-consonant deletion</td>
<td>The deletion or glottal replacement of intervocalic consonants</td>
<td>[peɪr] paper, [doʔi] dolly</td>
</tr>
</tbody>
</table>
Reference List


