Title:

Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties

Authors:

Belinda Seeff-Gabriel, Shula Chiat and Barbara Dodd

Department of Language and Communication Science, City University, London, UK

Address correspondence to: Belinda Seeff-Gabriel or Shula Chiat, Department of Language and Communication Science, City University, Northampton Square, London, EC1V 0HB, UK; email: Belinda.Seeff-Gabriel.1@city.ac.uk or shula.chiat.1@city.ac.uk

Keywords: sentence imitation, moprhosyntax, speech difficulties.

ABSTRACT

Background: Sentence imitation has been identified as a good indicator of children's language skills, with performance differentiating children with specific language impairment and showing relationships with other language measures. It has a number of advantages over other methods of assessment. The assessment of morphosyntax in children who have severe speech difficulties presents unique challenges which currently available sentence imitation assessments do not address.

Aims: This paper presents a new sentence imitation test (The Sentence Imitation Test (SIT^{61})) and reports on an investigation which sets out to determine whether this test (1) reveals differences in performance between a group of children diagnosed with specific language impairment and a group of typically developing children, (2) reveals distinct profiles of performance amongst children with different speech difficulties, and (3) provides information about morphosyntactic strengths and difficulties.

Methods and Procedures: SIT^{61} is a finely graded sentence imitation test in which the phonotactic structure, segmental phonology and length of words were kept as developmentally simple as possible. Responses are scored for number of content words, function words and inflections correct. A novel scoring system was devised to credit a child where there was evidence of targeting a morpheme even if it was mispronounced. The test was administered to four groups of

children between the ages of 4 and 6 years: 33 children with typical development, 13 children with known expressive morphosyntactic difficulties (specific language impairment), and two groups of 14 children with different types of speech disorder: a group with consistent phonological disorder, who used atypical phonological error patterns consistently; and a group with inconsistent phonological disorder, who produced atypical phonological errors inconsistently.

Outcomes and Results: SIT[61] found differences in performance between the group of typically developing participants and the three clinic groups. While the consistent phonological disorder group obtained extremely high scores for content and function words, they obtained lower inflection scores reflective of their speech difficulties. The scores of the specific language impairment and inconsistent phonological disorder groups were comparable for content and function words, but the groups were differentiated through an analysis of their errors. Further analyses confirmed that low scores obtained by some children in the inconsistent phonological disorder group were due to morphosyntactic difficulties and not speech difficulties.

Conclusions and Implications: A new sentence imitation test, the SIT[61], is shown to be a valuable tool for identifying expressive morphosyntactic difficulties in children. It is informative about the morphosyntactic abilities of children with speech disorders and raises questions as to the nature of their difficulties.

WHAT THIS PAPER ADDS

What is already known on this subject:

Research shows sentence imitation to be a useful method of assessing expressive morphosyntactic skills in children. However, the usefulness of this method when working with children who have speech disorders has not been explored. Currently available sentence imitation tests do not address the challenges associated with assessing this population group.

What the study adds:

A new sentence imitation, the SIT[61], is shown to be a valuable tool for identifying expressive morphosyntactic difficulties in children, and is informative about the morphosyntactic abilities of children whose speech is unintelligible.

BACKGROUND

Until recently sentence imitation has played a limited role in the assessment of children's expressive language skills. While included as a subtest in some language assessments, for example the Clinical Evaluation of Language Fundamentals (CELF) and the Test of Language Development (TOLD), sentence imitation is rarely used as an assessment tool in its own right. This may be because sentence imitation is assumed to test auditory memory rather than expressive language skills. However, research has begun to show that sentence imitation draws on and is informative about a range of language skills.

It is several decades since researchers first noticed that there was a close relationship between spontaneous production and sentence imitation in typically developing children. Brown and Fraser (1963) observed that children between 2 and 3 years of age reproduce adult models in a manner consistent with their own spontaneous, telegraphic productions. They often preserve lexical words, while omitting the functional words and morphemes. Others have found that children omit or inaccurately repeat structures not yet used spontaneously (Sturner, Kunze, Funk, & Green, 1993). In line with these earlier findings, Devescovi and Caselli (2007) found sentence repetition to be reliable and discriminating when tracking the morphosyntactic development of typically developing Italian preschoolers. Chiat and Roy (2008) administered a sentence repetition task to a sample of 187 clinically referred children aged 4 to 5 years, and found that

scores on the sentence imitation task correlated significantly with scores on the Preschool Language Scale-3(UK) (Boucher & Lewis, 1997), the Receptive and Expressive One Word Picture Vocabulary Tests (Brownell, 2000) and the Sentence Recall Subtest of the Clinical Evaluation of Language Fundamentals (CELF) -Preschool (Wiig, Secord, & Semel, 1992). Consistent with Brown and Fraser's observations of patterns of repetition in younger typically developing children, they found that the repetition performance of their clinic sample was influenced by morphosyntax, with children better able to repeat content words than function words. This selective difficulty with function words in sentence repetition is in line with the wide-ranging evidence that unstressed grammatical elements present a particular problem for children with specific language impairment (SLI) (Chiat, 2001; McGregor & Leonard, 1994). It confirms that the sentence repetition task tests more than auditory memory, revealing strengths and weaknesses in children's language. Bernstein Ratner (2000) pointed out the 'general agreement by researchers working over the past 30 years that sentences constructed at a level slightly above that observed in the child's spontaneous speech are regularized in ways that reflect both the child's extraction of form and meaning and the child's productive linguistic capacity' (p.293). This conclusion is reinforced by the more recent findings on sentence repetition in children with language impairments and typically developing children reported above.

In the light of observed relations between sentence imitation and other language measures, it is not surprising that sentence imitation has been put forward as a possible marker of SLI (Conti-Ramsden et al., 2001). A study conducted by Conti-Ramsden et al. (2001) evaluated sentence recall, along with non-word repetition and tense-marking, as potential psycholinguistic markers for SLI in children aged 11 who had been identified as having SLI at age 7. The Recalling Sentences Subtest of the CELF-R (Semel, Wiig, & Secord, 1994) was used to assess the participants' sentence repetition abilities. The study revealed sentence repetition to be the strongest of the four candidate markers of SLI. Recent cross-linguistic studies have found sentence repetition to be informative in diverse languages. For example, in a study of Cantonese-speaking children, Stokes et al. (2006) found that sentence repetition differentiated children with SLI from those who were typically developing (while nonword repetition did not).

These research findings point to the potential of sentence imitation as a clinical tool. In support of its use, research in the 1980s suggested that elicited speech production produces more robust effects than spontaneous production (Kahmi, Catts, & Davis, 1984; Paul & Shriberg, 1982), perhaps because children have greater control over linguistic encoding in spontaneous speech production and may selectively avoid complex structures (Panagos & Prelock, 1984). Facilitation tasks such as sentence imitation may provide a fuller picture of the child's linguistic skills and limitations.

Furthermore, sentence imitation has distinct practical advantages over spontaneous production as a means of assessing expressive language skills. First, sentence imitation enables the examiner to elicit a range of carefully selected targets in a more systematic way than is possible with spontaneous production. Allowing systematic manipulation of sentence length and grading of stimuli according to syntactic complexity, sentence imitation has the potential to reveal more comprehensive information about a child's sentence production abilities, as well as his/her threshold of performance. Second, targets can include morphosyntactic structures that are difficult to elicit in spontaneous production, for example negative and question structures. In addition, for Speech and Language Therapists under increasing pressure to determine a child's abilities in a short period of time, sentence imitation offers an efficient alternative to spontaneous production as it is less time-consuming.

A further advantage of sentence imitation that has not previously received attention is its potential for assessing expressive language in children whose spontaneous productions are unintelligible to the listener. The utterances of children with inconsistent speech errors may be particularly difficult to understand because the same item is produced differently on different occasions. Unlike children with consistent errors, where the listener may tune into the pattern or system of errors, children with inconsistent errors do not have a system to 'tune into'. Hence, without knowing what word or structure the child is intending to say in spontaneous production, it may be impossible to determine what has been

targeted. With naming tests, such as the Renfrew Word Finding Test (Renrew & Mitchell, 1997) and the Expressive One-Word Picture Vocabulary Test (Brownell, 2000), the targets are known for picturable concepts at a single word level. With sentence imitation, the targets are known not only for a variety of word types, but also a range of syntactic structures.

Existing sentence imitation tests do not exploit the full potential of this assessment method. As exemplified in studies mentioned above, the CELF Recall tests are commonly used in research, as well as clinically. The TOLD-P:3 (Test of Language Development, Primary, 3rd edition; Newcomer & Hammill, 1997) is another omnibus language assessment that includes a sentence imitation subtest. These assessments are norm referenced, and the sentence imitation subtests comprise a set of sentences ranging in length. Within both tests, scoring is focused on the production of the entire sentence. The Recalling Sentences in Context Subtest scores a child's production according to the number of errors made in each sentence. In neither test is provision made for evaluating the production of specific sentence structures or elements. These tests therefore yield a broad score and are used primarily to determine whether a child has difficulty compared to his/her typically developing peers. While this fulfills an important purpose, it does not take full advantage of the potential of sentence imitation to identify specific strengths and difficulties, and hence to inform targets for intervention.

Furthermore, these tests do not lend themselves to the assessment of children who have speech difficulties. The selection of words appears not to take phonological factors into account, so includes phonological targets that are late emerging and that some children with speech difficulties may avoid (Stoel-Gammon, 1998). More crucially, there are no guidelines for scoring words or morphemes produced incorrectly due to speech difficulties. Therefore, while these tests have the implicit advantage of providing a target against which to measure a child's production, neither addresses the specific challenges associated with assessing young children with speech difficulties. As a result, these children may obtain scores which underestimate their true expressive language potential: their low scores may be attributed to expressive language difficulties even where they are in fact due to speech difficulties.

This paper presents a new sentence imitation test (SIT[61]) which focuses on morphosyntactic structures in simple sentences and aims to address the shortfalls of currently available assessments as outlined above. The test is systematically graded in terms of length, and the scoring allows for the identification of specific morphosyntactic strengths and difficulties within a sentence. In addition, the challenges of assessing the expressive language of children with speech difficulties are addressed by careful selection of stimuli and a supplementary scoring system. The paper reports an investigation which set out to determine whether this test (i) differentiates between a group of children diagnosed with SLI and a group of typically developing children, (ii) reveals

distinct profiles of performance amongst children with different speech difficulties, and (iii) provides information about a child's morphosyntactic strengths and difficulties. It also explores the patterns of error in different groups, with implications for diagnosis and intervention.

METHODS & PROCEDURES

Participants

Participants were recruited through Speech and Language Therapists working in a specialist centre (The Nuffield Hearing and Speech Centre in the Royal Throat, Nose and Ear Hospital), as well as clinics and mainstream services in the areas of Camden, Redbridge and Newcastle. A child was included in the study if s/he was an English first-language speaker and scored within the normal range (> -1 SD) on:

(i) a measure of receptive language (Test of Auditory Comprehension of Language: TACL-3, Carrow-Woolfolk, 1999)

(ii) a non-verbal screening (The Picture Completion Subtest of the Weschler Preschool and Primary Scale of Intelligence-Revised: WPPSI-R^{uk}, Wechsler, 1990)

(iii) an oro-motor screening (Oro-motor Screening Subtest of the Diagnostic Evaluation of Articulation and Phonology (DEAP): isolated movements, sequential movements and DDK) (Dodd et al., 2002).

Children were assigned to the SLI group if they fulfilled the following criteria:

(i) expressive language difficulties according to SLT's referral

(ii) age-appropriate speech development or speech delay (Bradford and Dodd, 1994) as measured on the Phonology Subtest of the DEAP.

Children with speech difficulties were assigned to one of the following two groups using the Phonology and Inconsistency Subtests of the DEAP:

(i) CPD group: children with consistent speech/phonological disorder who use at least two atypical error patterns consistently as indicated by a score of less than 40% on the Inconsistency Subtest of the DEAP.

(ii) IPD group: children with inconsistent speech/phonological disorder as indicated by a score of more than 40% on the inconsistency subtest of the DEAP, production of atypical errors with no observable pattern, and no noted articulatory groping on volitional phoneme production (Bradford & Dodd, 1994).

Children with a profile of phonological delay, who use only phonological error patterns that occur in typical development, were not included since they present no particular challenge for assessment of expressive language abilities.

A sample of typically developing children was recruited from the same schools as the children attending clinic and from schools in close proximity. Invitations were given to the parents of an entire class of children. Children were included if they fulfilled the following criteria:

(i) age-appropriate receptive and expressive language abilities according to the teacher

(ii) no apparent communication or motor difficulties according to the teacher

(iii) first-language English

(iv) parental consent.

A total of 41 clinically referred children were recruited, with roughly equal numbers in the specific language impairment (SLI), consistent phonological disorder (CPD) and inconsistent phonological disorder (IPD) groups, along with 33 typically developing (TD) children. Table 1 presents background information on the four groups of participants. One-way analysis of variance tests found no significant differences between the groups on age, receptive language scores and non-verbal scores.

INSERT TABLE 1 ABOUT HERE

Sentence Imitation Test (SIT[61])

The Sentence Imitation Test (SIT[61]) consisted of sixty-one stimuli which were graded for syntactic complexity using the Language Assessment, Remediation

and Screening Procedures (LARSP; Crystal, Fletcher, & Garman, 1976). This framework represents a developmental hierarchy of syntactic structures, and SIT-61 stimuli reflected structures ranging from Stage II (1.6-2.0 years) to Stage VI (3.6-4.6 years). The stimuli consisted of simple sentences apart from the final ten which contained complex sentence structures, including embedded verb phrase, relative clause, and a passive, e.g. *I want you to kick the ball to me*; *He is the boy who fell*; *The apple was eaten by the bird*. Hence, the SIT[61] could be considered a test of basic structures that develop in the pre-school years.

As the primary objective of the SIT-61 was to assess expressive morphosyntax, the stimuli were carefully designed to minimise the effects that non-targeted variables, including semantic familiarity and word frequency, could have on performance. Words were taken from the vocabulary checklist section of the MacArthur Communicative Developmental Inventory (Fenson, 1993), and the 'early' and 'very early acquired' words from the age-of-acquisition rated nouns and verbs in 'An Object and Action Naming Battery' (Druks & Masterson, 2000). In addition, phonotactic structure, segmental phonology and length of words were kept as developmentally simple as possible to accommodate children with speech difficulties: the phonotactic structures used were VC / CV / CVCV / CVC / CVCVC; consonant clusters were avoided apart from four instances; later developing phonemes such as affricates were not included; the length of nouns, verbs and adjectives was kept to one syllable or two syllables with initial stress;

and all prepositions contained one syllable apart from *under*. Inflections were not included if they created a word-final cluster.

The 61 stimuli contain a total of 355 morphemes. These were broken down into content words (nouns, verbs, adjectives, adverbs), n=157; function words (prepositions, pronouns, auxiliaries/copulas, determiners, conjunctions, subordinators/complementisers), n=159; and inflections (plural, past tense, 3rd person singular agreement, contracted negative), n=39. The sampling of inflections was limited by the exclusion of inflections that create a word-final cluster.

Procedure

Stimuli were presented in a fixed order in blocks of increasing length and complexity. Stimuli were rehearsed and presented live by the researcher to maintain the child's motivation and maximise participation in the task. Sessions were recorded and each child's repetitions were transcribed and scored by the researcher. To evaluate inter-rater reliability, the data from five randomly selected participants (two from each speech disordered group and one from the group of children with SLI), were transcribed and scored independently by a final-year SLT student trained in phonetic transcription.

Scoring

Each response was scored for number of correct content words, function words and inflections, and scores were summed to give a total content word, function word and inflection score. Total scores were represented as a percentage of the total numbers of target morphemes in each of these categories.

A special scoring system for children with unintelligible speech was devised. This laid out minimum requirements for crediting a child with the production of a correct morpheme, the aim being to maximise the possibility of identifying morphosyntactic skills that might be masked by speech difficulties. For content and function words, these requirements were as follows:

(i) the presence of at least one correctly placed consonant and correct vowel, allowing for substitutions due to systematic error patterns, e.g. [peɪ] for *face*, where the child is stopping fricatives.

(ii) the presence of at least one correctly placed consonant allowing for substitutions due to systematic error patterns, and a vowel that matches the target vowel in length and is close in vowel space, e.g. [kɛk] for *cat*, where the child has either backed target /t/ or assimilated this to the initial /k/.

(iii) all correct consonants, allowing for substitutions due to systematic error patterns, but with an incorrect vowel, e.g. [mæt] for *make*.

(iv) the presence of the correct initial consonant with a vowel and/or other consonants showing interference from a word preceding or proceeding it, e.g. [huz Σuz] for *her shoes*.

The marking of inflections may be challenging for a child who has difficulty producing particular consonants, e.g. /s/. The child was credited with the production of the inflection if there was evidence of any consonant in the position of the inflection, e.g. [hƆσIδ] for *horses*.

Analysis

Performance of the four groups was compared in terms of sentence element type (content word, function word, inflection) and sentence length (short: 2-4 words, long: 6-9 words). Non-parametric statistics were used when analysis involved the TD group, as this group clearly violated the assumption of normal distribution necessary for parametric analysis. Although the assumptions for parametric analyses were not always met by the clinic groups, i.e. some of the distributions were found to be negatively skewed, a decision was made to use parametric analyses, based on the case put forward by Howell (1997): 'in practice, however, the analysis of variance is a very robust statistical procedure and the assumptions can be violated with relatively minor effects. This is especially true for the normality assumption… in general, if the populations can be assumed to be symmetrical or at least symmetrical in shape e.g. all negatively skewed, and if

the largest variance is no more than four times the smallest, the analysis of variance is most likely to be valid'.

OUTCOMES & RESULTS

Inter-rater reliability

Correlations were determined for the two raters' scoring of content words, function words and inflections on each of the 61 stimuli for five randomly selected participants: two from each speech disordered group and one from the SLI group. Mean agreement for content words was found to be $r(59)=.876$ ($p<.01$) with a range of .82 to .92; for function words, $r(59)=.77$ ($p<.01$) with a range of .74 to .81; and for inflections, $r(59)=.788$ ($p<.01$) with a range of .75 to .84.

Comparison of TD and SLI groups

In order to investigate whether the SIT[61] differentiated between the typically developing group and the group with known expressive morphosyntactic difficulties, the scores of the TD and SLI groups were compared. As shown in table 2 and figure 1, the TD Group's means were above 99% for all three syntactic categories, with extremely small standard deviations. Thus, the typically developing children completed the task with ease and their performance was not affected by length. Scores in the SLI group were lower, and overlap with the TD

group was negligible: only two children's content word scores, one child's function word score, and one child's inflection score fell within the range of the TD group. Mann-Whitney tests revealed significant differences between the TD group and the SLI group on content words ($U=8$, $p<.001$), function words ($U=1$, $p<.001$) and inflections ($U=8.5$, $p<.001$). Therefore SIT differentiated between the children with typical development and those diagnosed with expressive morphosyntactic difficulties.

INSERT TABLE 2 ABOUT HERE

INSERT FIGURE 1 ABOUT HERE

As the performance of the TD group was close to or at ceiling for all three categories, their performance was not analysed further. Paired-samples *t*-tests were used to investigate the effects of sentence element type and sentence length on the performance of the SLI group. These showed that the SLI group's mean content-word score was significantly higher than its mean function-word score ($t_{13}=4.53$, $p<.01$) and mean inflection score ($t_{13}=3.25$, $p<.01$), which did not differ significantly from each other. These results demonstrate that performance was influenced by morphosyntactic factors. It was also found that the group performed better on short stimuli than on long stimuli ($t_{11}=-2.26$, $p<.05$).

Comparison of CPD and IPD with TD and SLI groups

In order to investigate whether the performance of the two speech disordered groups was indicative of typical or impaired expressive morphosyntax, the performance of each of the speech disordered groups was compared to that of the SLI and TD groups.

Table 3 presents the means, standard deviations and ranges obtained by the CPD, IPD and SLI groups for content words, function words and inflections. Figure 2 presents error bars for each group's scores.

INSERT TABLE 3 ABOUT HERE

INSERT FIGURE 2 ABOUT HERE

As table 3 and figure 2 show, the CPD group obtained high mean scores for content and function words (content words: 96.5%; function words: 90%) and its means were higher than those obtained by both the SLI and IPD groups (IPD content words: 89.4%; function words: 71.9%; SLI content words: 90%; function words: 76.8%). All groups had the smallest range of scores for content words and all showed a wider range of scores for function words. However, the range for the CPD group is deceptive as the score of 49.7% was an outlier. Excluding this outlier, the range was 82.4%-100%. Means were lowest and range of scores were widest for inflections for the CPD and IPD groups (CPD mean=78.8%, range=35.9-100%; IPD mean=68.9%, range= 23.1-97.4%), while the SLI group

obtained a slightly higher mean score for inflections compared to function words (function words=76.8%, inflections=80.8%). The CPD group's scores on content and function words were closest to the corresponding scores for the TD group (see table 2). Nevertheless, comparisons using Mann-Whitney revealed significant differences (content words: U=54, p<0.001; function words: U=37, p<.001).

A three-factor mixed ANOVA compared the performance of the CPD, IPD and SLI groups. The main effect of sentence element was significant ($F_{2,74}$=61.37, p<.001, η=.624). The main effect of group was not significant ($F_{2,37}$=2.2, p=.13, η=.106). This could be due to the fact that the mean scores of the SLI group were between the CPD and IPD groups' content and function word scores, as well as the general variability in the IPD and SLI groups' scores. The interaction of sentence element and group was significant ($F_{4,74}$=35.6, p<.01, η=.12) The main effect of length was significant ($F_{1,37}$=26.75, p<.001, η=.42).

Post hoc analyses using Independent *t*-tests showed that it was only on function words that the CPD group performed significantly better than the IPD group (t_{26}=2.65, p<0.05), though the difference in content-word performance of the two groups approached significance (t_{26}=1.98, p=.059). The CPD group obtained significantly higher scores than the SLI group for content words (t_{25}=2.12, p<.05) and function words (t_{25}=2.27, p<.05). Although the SLI group obtained a slightly higher mean score for inflections than the CPD group, this was not significant.

Differences in performance between the IPD and SLI groups turned out not to be significant for any of the sentence element types.

Turning to the comparison of sentence element types within groups, paired-samples *t*-tests showed that the CPD group's mean content-word score was significantly higher than its mean function-word score (t_{13}=2.68, p<.05), which was significantly higher than its mean inflection score (t_{13}=4.22, p<.01). While the IPD and SLI groups also obtained significantly higher mean content-word compared to function-word scores (IPD: t_{13}=5.36, p<.01, SLI: t_{12}=4.53, p<.01) and inflections (IPD: t_{13}=6.04, p<.01, SLI: t_{12}=3.25, p<.01), the difference between their mean function-word and inflection scores was not significant.

In summary, all clinical groups performed significantly below the TD group, with the CPD group's scores being closest to those of the TD group. The CPD group performed significantly better than the IPD and SLI groups on function words, and outperformed both groups on content words though this difference fell just short of significance in the case of the IPD group. The performance of the IPD and SLI groups did not differ significantly for any of the sentence elements. All groups were affected by the length of the stimuli.

Further comparison of IPD and SLI groups

The finding that the scores of the IPD and SLI groups did not differ significantly suggests one of two things: either that the lower scores obtained by the IPD group reflect speech difficulties which the special scoring system may have failed to overcome; or that some of the children with IPD also had expressive morphosyntactic difficulties characteristic of SLI. We consider each of these possibilities in turn.

(i) It is possible that the IPD group had more severe phonological difficulties than the CPD group, and that these affected their repetition of sentences even when the allowances of the special scoring system were made. The IPD group may have produced fewer phonemes correctly, affecting their realisation of target morphemes, which then failed to meet the criteria for crediting a morpheme. In order to address this possibility, the phonological accuracy scores of the two groups was compared in terms of percentage consonants correct (PCC) and percentage vowels correct (PVC) on the Phonology Subtest of the DEAP. Figure 3 shows the PCC and PVC scores for the two groups.

INSERT FIGURE 3 ABOUT HERE

A two-factor mixed ANOVA found a main effect of consonant versus vowel production ($F_{1,26}$= 248.32 p<.001, effect size p=.905), but the main effect of group was not significant. There was no interaction effect. As expected, post hoc analysis using Paired samples *t*-tests showed that the PVC score was

significantly higher than the PCC score for both the CPD (t_{13}=-9.06, p=.001) and IPD (t_{13}=-15.96, p=.001) groups.

Hence, the differences found between the two groups on the SIT[61] could not be attributed to differences in phonological accuracy. In addition, many of the CPD participants obtained scores at or almost at ceiling on the SIT[61] despite low PCC scores, confirming that performance is not dependent on PCC.

While there were no group differences in speech accuracy, it is possible that, within the IPD group, there were children whose speech was so inconsistent that even with the special scoring system, they could not be credited with the production of morphemes they were in fact targeting, and hence they obtained poor scores. To address this possibility, Spearman's rank order correlations were used to determine the relationship between the degree of inconsistency and the scores obtained by these participants for content words, function words and inflections. No significant correlations were found. For example, one child with an inconsistency score of 56% obtained 96.8% for content words, 74.2% for function words and 82.1% for inflections. In contrast, another child with an inconsistency score of 44%, obtained 79% for content words, 64% for function words and 61% for inflections. Hence, the scores obtained by the participants in the IPD group were not associated with the degree of inconsistency.

(ii) In light of the above analyses showing that poor scores obtained by some of the participants in the IPD group could not be attributed to the severity of their speech difficulty in terms of phonological accuracy or inconsistency, it may be concluded that their poor scores were a reflection of their expressive morphosyntactic abilities as opposed to their speech. However, their difficulties with expressive morphosyntax are not necessarily the same as the difficulties of the children with SLI. An analysis of errors may throw more light on this.

Error analysis

The following error categories were identified:

(i) Whole-word substitution from the same grammatical category: target word substituted by a word of the same grammatical category, e.g. [gɜl] for *lady*.

(ii) Whole-word substitution from a different grammatical category: target word substituted by a word of a different grammatical category, e.g. *The water made black shoes dirty* where *black* has replaced *her*.

(iii) Sound substitution: number of target syllables maintained but sounds bear no relation to target morphemes and do not meet criteria for crediting a child with the morpheme, e.g. [ʊ hʌhʌ kəu] for *put on your coat*.

(iv) Unmatched syllables: number of target syllables not maintained and syllables cannot be matched to target morphemes, e.g. [jæb gæm] for *There is the man*. Here it is uncertain which target syllable [jæb] corresponds to.

(v) Omission: all syllables produced can be matched to target morphemes and it is clear that a target morpheme has been omitted, e.g. [kɪtɨ ʌgʌ bʌk] for *The cat was under the bus*.

(vi) Distortion: child's production is too distorted or unintelligible to be transcribed.

Note: /ə/ for /ðə/ substitutions were not marked as sound or whole-word substitution errors. This is because it is extremely difficult to determine whether /ə/ has been produced due to difficulty with the production of the more difficult phoneme /ð/, or whether it is in fact a whole-word substitution.

All content and function word errors made by the SLI and IPD groups were classified using the above categories. Figure 4 presents each group's mean for each error type as a percentage of the total content- and function-word errors made by that group.

INSERT FIGURE 4 ABOUT HERE

The results in figure 4 show that the main types of error made by both groups were omissions (SLI=54.9%; IPD=41.3%) and whole-word substitutions from the same grammatical category (SLI=36.8% IPD=24.4%), with no significant differences between them. While the SLI group made a significantly higher proportion of whole-word substitutions from a different grammatical category ($U=27.5$, $p<.002$; SLI=4.1%; IPD=.5%), the IPD group's errors contained a significantly higher proportion of sound substitutions ($U=5$, $p<.001$, SLI=1.9%; IPD=20.6%) and unmatched syllables ($U=45$, $p<.01$; SLI=.4%; IPD=10.1%). Hence, the IPD group's repetition performance was to some extent differentiated from the SLI group by their errors. The nature of these errors suggests that they may have arisen from the co-occurrence of difficulties with speech and morphosyntax.

DISCUSSION

The results of the study revealed that:

(i) SIT[61] differentiated between the children with and without expressive morphosyntactic difficulties. This was shown through the significant difference in performance of the typically developing children and those with SLI who were known to have expressive morphosyntactic difficulties, with very little overlap between these groups. However, these very clear findings must be interpreted with caution. The numbers in each group were small.

Furthermore, given the recruitment process for both groups, it cannot be assumed that they represent the full spectrum of performance in the populations from which they are drawn. This study therefore represents only the first step in evaluating the diagnostic accuracy of SIT[61]; a robust diagnostic tool would require evidence of reliable differentiation between fully representative samples (Klee, 2008; Sackett & Haynes, 2002)[1].

(ii) The test overcomes the challenges associated with assessing the expressive language abilities of children with speech difficulties. As pointed out above, some children with speech difficulties may avoid phonemes they perceive to be difficult, and in some instances refuse to attempt tasks. All the participants, irrespective of their degree of speech or sentence-level difficulties, attempted all sentences on the test. Furthermore, there were children in the CPD and IPD groups who obtained scores that were well within the range of the TD participants, and analyses revealed that speech severity and inconsistency were not significantly related to performance on the SIT[61].

(iii) SIT[61] identified specific morphosyntactic weaknesses in the clinic groups. All clinic groups performed better on content words than function words and inflections. The CPD group performed better on function words than inflections, while the IPD group, like the SLI group, performed similarly on function words and inflections. The finding that the children with known expressive morphosyntactic difficulties in the SLI group were significantly

[1] A shortened version of SIT[61] has now been standardised (Seeff-Gabriel, Chiat, & Roy, 2008). In addition, an adapted version (SIT-16), has been administered to a large sample of clinically referred children (Chiat & Roy, 2008) providing further evidence of its diagnostic potential.

weaker on function words and inflections compared to content words is in line with previous research showing children with SLI to have particular difficulties with function words and inflections (Chiat, 2001; McGregor & Leonard, 1994). Similar findings have emerged in other languages. For example, in a study of pre-school Italian children, Bortolini et al. (2006) found that direct-object clitics were particularly vulnerable, and proposed that these might serve as a clinical marker for SLI.

Having established the different groups' performance on the SIT[61], what inferences can we make about the expressive morphosyntactic abilities of children who have speech difficulties?

CPD Group

Overall, the children with CPD did not perform as well as children in the TD group. However, the majority of children with CPD scored at or almost at ceiling for content and function words despite their speech difficulty. This indicates that performance on the SIT[61] need not be negatively affected by speech difficulties alone.

Since inflections are mostly subsyllabic, e.g. plural /s/ and /z/, it is likely that lower scores confined to this category were due to speech difficulties. In support of this, an analysis revealed that the majority of errors made by the CPD group

(70.68%) involved (i) the production of the later developing phonemes /s/ and /z/ as plurals, for example 'We have cake at *parties*', and as third person singular agreement marker, for example 'Tim *goes* in the house'; and (ii) clusters such as the contracted negative, illustrated in the sentence '*Don't* feed the horses'. Inflections that were produced more successfully by these participants included (i) the progressive inflection, for example 'Mummy is *giving* Peter cake' where *ing* is a whole syllable; and (ii) irregular past tense verbs, e.g. *gave*, *sat*, and *ate* where the vowel indicates tense. This distribution of errors indicates that the lower inflection scores obtained by some children with speech difficulties may be primarily attributable to their speech difficulties, rather than their linguistic knowledge.

IPD Group

There was significant variability in the performance of children with IPD. The performance of eight out of the 14 participants was comparable with that of typically developing children and children with CPD. These findings indicate that, despite their scores being lower than the scores of children with CPD at a group level, children with IPD do not necessarily have expressive morphosyntactic difficulties. The remaining six children with IPD obtained low scores across all three morphosyntactic categories, and the group as a whole did not differ from the SLI group. The majority of their errors, as with the SLI group, involved omissions and whole-word substitutions, suggesting that they have the same

difficulties as children with SLI. However, their errors included unmatched syllables, as well as significantly more sound substitutions than the SLI group. These errors are most plausibly related to their speech difficulties. Based on this analysis of errors, it may be concluded that the children with IPD who obtained low scores on the SIT[61] had co-occurring speech and language difficulties.

However, there is an alternative interpretation of their errors that requires consideration. Perhaps some children with inconsistent speech have underlying difficulties with speech planning not only at the word level, but at the prosodic phrase level of production as well, and these difficulties lead to omissions, particularly of weak elements in the prosodic phrase (Levelt et al., 1999). In this case, the similarity of this group's profile with the profile of the SLI group would be superficial, arising from distinct underlying difficulties. In order to explore this possibility, further research is needed to investigate speech planning at the prosodic phrase level. In addition, further investigation is needed to determine whether these children show morphosyntactic difficulties in tasks that do not involve speech production (Seeff-Gabriel and Chiat, in preparation).

Co-occurrence of speech and language disorders

The present research revealed that the children with speech disorders were not a homogeneous group: the performance of the CPD group was better than that of the IPD group especially with respect to function words, and was close to the

performance of the TD group except on inflections. This suggests a low probability of co-occurrence of CPD and expressive morphosyntactic difficulties, probably no greater than the incidence in the general population. However, the presence of one outlier within the CPD group points to the need for further investigation before drawing firm conclusions about morphosyntactic abilities in this group. Although some children with IPD scored within the normal range, as a group the IPD children appeared to be at greater risk of morphosyntactic difficulties like those demonstrated by the SLI group in this study. Overall, the co-occurrence of IPD and morphosyntactic difficulties was 43% (as mentioned above, 6 out of 14 children with IPD obtained low scores across all three morphosyntactic categories). Though the groups in this study were small, tentative comparisons can be made with other studies of co-occurrence of speech and language disorders. Different profiles of performance across the children with speech difficulties may partially explain the variability in co-morbidity estimates reported in a large scale review carried out by Shriberg and Austin (1998): it could be that samples of children in the different studies were made up of different proportions of children with CPD and IPD, differentially influencing the outcomes of each study. The low incidence of co-occurring CPD and morphosyntactic difficulties is also incongruent with the result of 45% for co-occurrence of CPD and language difficulties obtained by Broomfield and Dodd (2004). Here, the difference may be accounted for by the differing ages, assessments and cut-off criteria used in their study, compared with those in this study. The children included in Broomfield and Dodd's study ranged between the

ages of 0 and 16 years. Their expressive language abilities were determined by a variety of assessments, depending on the age of the child, and a cut-off of 1SD was used to indicate the presence of a difficulty. In contrast to this, the children in this study represented a narrow age range (from 4 to 6 years). A consistent measure of expressive language assessment was used (SIT[61]), and a more extreme cut-off of 1.5 SD below the mean of the TD group indicated a difficulty.

Given the small groups in the present study, and the substantial differences in samples, methods and criteria for impairment, no firm conclusions can be drawn on rates of co-morbidity. However, the present study points to the need for further investigation of the nature of speech and language difficulties in children who show both speech and language disorders. This will clarify whether the disorders are co-morbid i.e. due to the co-occurrence of difficulties observed in a pure speech disorder and in SLI, or whether they are related, reflecting difficulties in speech planning and production that impact on the sentence level.

Conclusion

This is the first study to use sentence imitation with children who have speech difficulties. The design and scoring were shown to accommodate these children. The SIT[61] discriminated between groups of children with consistent and inconsistent speech difficulties. The CPD group obtained morphosyntactic scores comparable with typically developing children, except for inflections, where

production was affected by their speech difficulties. The performance of the IPD group, on the other hand, showed an overlap with the SLI group's performance, in terms of scores on morphosyntactic categories and types of error. However, some children with IPD attained normal scores, implying that low scores on the SIT[61] were not just due to speech difficulties but indicated co-occurring morphosyntactic difficulties. Further research is needed to verify whether these difficulties are indeed distinct and comorbid, or whether they are interrelated, for example stemming from difficulties in speech production at the prosodic phrase as well as at the word level. It is also acknowledged that sentence imitation does not provide a complete picture of a child's sentence level abilities, and further investigations are needed to explore the exact nature of the underlying difficulties using assessment methods and including tasks which do not require the production of speech.

ACKNOWLEDGEMENTS

The authors would like to thank all the Speech and Language Therapists and Teachers who supported this research by referring suitable candidates, and the children themselves who willingly participated in the study. Particular thanks to the Speech and Language Therapists at the Nuffield Hearing and Speech Centre whose co-operation resulted in many of the participants being seen at their Centre.

REFERENCES

BERNSTEIN RATNER, N., 2000, Elicited Imitation and other methods for the analysis of trade-offs between speech and language skills in children. In L. B. R. Menn, N (ed), *Methods for Studying Language Production* (London: Lawrence Erlbaum Associates), pp. 291-311.

BORTOLINI, U., ARFÈ B., CASELLI M.C., DEGASPERI L., DEEVY P. & LEONARD L.B., 2006, Clinical markers for language impairment in Italian., International Journal of Language and Communication Disorders, **6**, 695-712.

BOUCHER,J., & LEWIS, V., 1997, *Preschool Language Scale -3 UK Edition* (London: Psychological Corporation)

BRADFORD, A., & DODD, B., 1994, The motor planning abilities of phonologically disordered children. *European Journal of Disorders of Communication,* **29**, 349-369.

BROOMFIELD, J., & DODD, B., 2004, Children with speech and language disability: caseload characteristics. *International Journal of Language and Communication Disorders,* **39**, 1-22.

BROWN,R., & FRASER,C., 1963, The acquisition of syntax. In Cofer, C.N. & Musgrave, B., *Verbal behavior and learning: Problems and processes*. (New York: McGraw-Hill.Brown & Fraser)

Brownell, R., 2000, *Receptive and Expressive One Word Picture Vocabulary Tests* (Illinois: Linguisystems)

CARROW-WOOLFOLK, E, 1999, *Test for Auditory Comprehension of Language (TACL-3)* (3rd edition ed.) (Austin, Texas: Pro.Ed).

CHIAT,S., 2001, Mapping theories of developmental language impairment: premises, predictions and evidence. *Language and Cognitive Processes*, **16**, 113-142.

CHIAT, S., & ROY,P., 2008, Early phonological and socio-cognitive skills as predictors of later language and social communication outcomes. Journal of Child Psychology and Psychiatry, **49**, 635-645.

CONTI-RAMSDEN, G., BOTTING, N., & FARAGHER, B., 2001, Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, **42*****(6)**, 741-748.

CRYSTAL, D., FLETCHER, P., & GARMAN, M., 1976, *The grammatical analysis of language ability* (New York: Elsevier).

DEVESCOVI, A., & CASELLI, M.C., 2007, Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language and Communication Disorders*, **42**(2), 187-208.

DODD, B., HUA, Z., CROSBIE, S., HOLM, A., & OZANNE, A., 2002, *Diagnostic Evaluation of Articulation and Phonology* (London: The Psychological Corporation).

DRUKS, J., and MASTERSON, J., 2000. *An Object and Action Naming Battery* (UK: Psychology Press Ltd).

FENSON, L., 1993. *MacArthur Communicative Development Inventories* (California: Singular Publishing).

HOWELL, 1997, *Statistical Methods for Psychology* (4th edition) (Boston: Duxbury Press).

KAMHI, A. G., CATTS, H. W., & DAVIS, M. K., 1984, Management of sentence production demands. *Journal of Speech and Hearing Research,* **27**, 329-338.

LEONARD, L., 1998, *Children with Specific Language Impairment* (London: MIT Press).

LEVELT, W.J.M., ROELOFS,A., & MEYER,A.S., 1999, A theory of lexical access in speech production. *Behavioural and Brain Sciences*, **22**, 1-75.

KLEE, T., 2008, Considerations for appraising diagnostic studies of communication disorders. *Evidence-based Communication Assessment and Intervention, **28**,* 155-165.

MCGREGOR, K.K., & LEONARD, L.B., 1994, Subject pronoun and article omissions in the speech of children with specific language impairment – A phonological interpretation, *Journal of Speech and Hearing Research*, **37**, 171-181.

NEWCOMER, P. L., & HAMMILL, D. D., 1997, *Test of Language Development. Primary (3rd Ed)(TOLD-P:3)* (Austin, Texas: Pro-Ed).

PANAGOS, J. M., & PRELOCK, P. A., 1984, Comments on the interaction of syntactic and phonological disorders. *Journal of Speech and Hearing Research,* **27**, 318-319.

PAUL, R., & SHRIBERG, L. D., 1982, Associations between phonology and syntax in speech delayed children. *Journal of Speech and Hearing Research, 25* (December), 536-547.

RENFREW,C., & MITCHELL,P., 1997, Renfrew Word Finding Vocabulary Test (UK: Speechmark Publishing Ltd).

SACKETT, D.L., & HAYNES, R.B., 2002, Evidence base of clinical diagnosis: the architecture of diagnostic research. *British Medical Journal, **324**,* 539-541*.*

SEMEL, E., WIIG, E., & SECORD, W., 1994, *Clinical Evaluation of Language Fundamentals-Revised* (San-Antonio, Texas: The Psychological Corporation).

SHRIBERG, L. D., & AUSTIN, D., 1998, Comorbidity of speech-language disorder. In R. Paul (ed.), *Exploring the Speech-Language Connection* (Vol. 8), (Baltimore: Paul H. Brookes Publishing Co), pp. 73-117.

STOEL-GAMMON, C., 1998, Sounds and words in early language acquisition: The relationship between lexical and phonological development. In R. Paul (ed.), *Exploring the Speech-Language Connection* (Baltimore: Paul H. Brookes Publishing Co), vol. 8, pp. 25-52.

STOKES, S.F., WONG, A., FLETCHER, P., & LEONARD, L.B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research*, **49**, 219-236.

STURNER, R. A., KUNZE, L., FUNK, S. G., & GREEN, J. A., 1993, Elicited imitation: its effectiveness for speech and language screening. *Developmental Medicine and Child Neurology,* **35**, 715-726.

WECHSLER, D., 1990, *Wechsler Preschool and Primary Scale of Intelligence - Revised (WPPSI-R$^{uk)}$* (London: The Psychological Corporation).

WIIG, E. H., SECORD, W., & SEMEL, E., 1992, *Clinical Evaluation of Language Fundamentals - Preschool (CELF-Preschool)* (The Psychological Corporation, Harcourt Brace Jovanovich, INC).

Table 1: Profile of participants

	N	Gender		Age (months)		Receptive language (quotient)		Non-verbal screen (scaled score)	
		M	F	Mean	Range	Mean	Range	Mean	Range
SLI	13	10	3	56.5	48-72	100.9	85-124	11.9	8-14
CPD	14	12	2	61.1	48-73	108.4	91-128	12.5	8-17
IPD	14	11	3	61.9	48-71	100.6	85-124	11.3	8-16
TD	33	19	14	58.1	48-75	108.9	91-128	12.3	8-14

Table 2: Performance on SIT[61]: TD and SLI groups

	Content words		Function words		Inflections	
	TD	SLI	TD	SLI	TD	SLI
Mean	99.8%	90.1%	99.6	76.8%	99.3%	80.8%
SD	.3	10.4	.7	17.3	1.3	17
Range	98.7-100%	65-99.4%	96.9-100%	34.6-97.4%	95-100%	33.3-97.7%

Table 3: Performance on SIT[61]: CPD, IPD and SLI groups

	Content words			Function words			Inflections		
	CPD	IPD	SLI	CPD	IPD	SLI	CPD	IPD	SLI
Mean	96.5%	89.4%	90%	90%	71.9%	76.8%	78.8%	68.9%	80.8%
SD	4.4	12.6	10.4	12.8	22.1	17.4	18	20.3	17.0
Range	85.4-100%	58-100%	65-99%	49.7-100%	29.6-98.7%	34.6-97.4%	29.6-98.7%	35.9-100%	23.1-97.4%

Figure 1: Comparison of Content word, Function word and Inflections for the TD and SLI groups

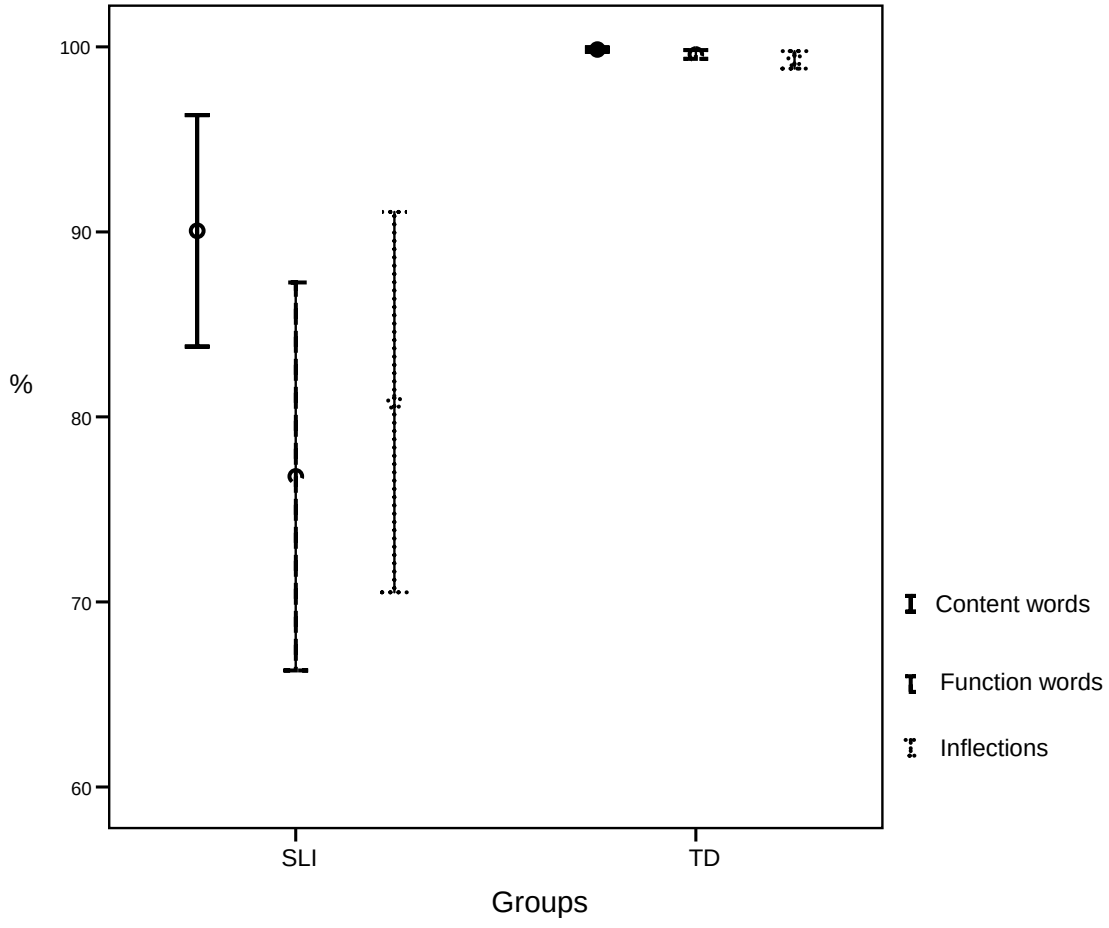

Figure 2: Comparison of Content word, Function word and Inflections for the CPD, IPD and SLI groups

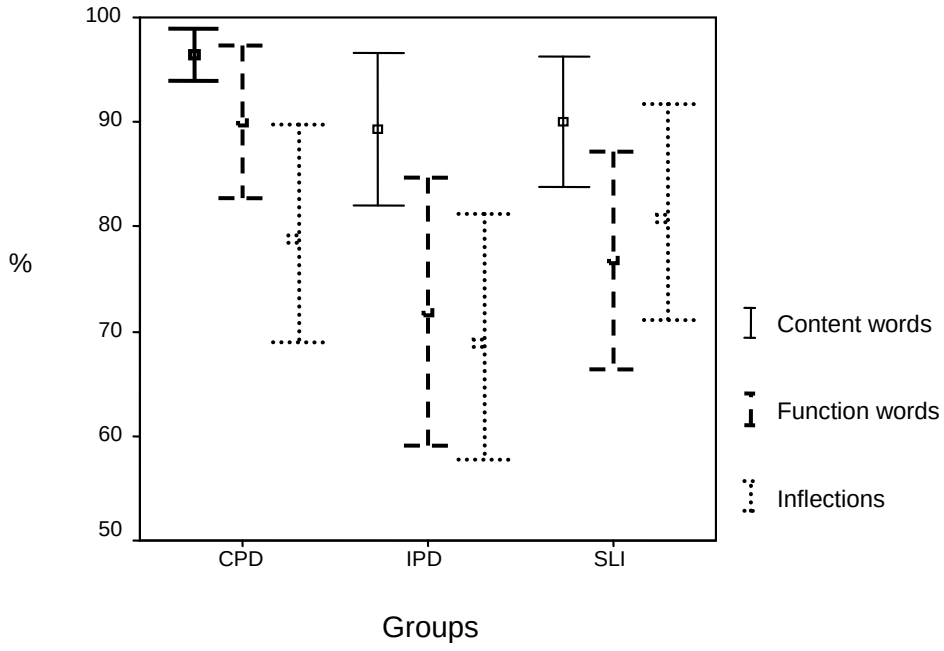

Figure 3: Phonological accuracy scores as measured on the Phonology Subtest of the DEAP: CPD and IPD groups

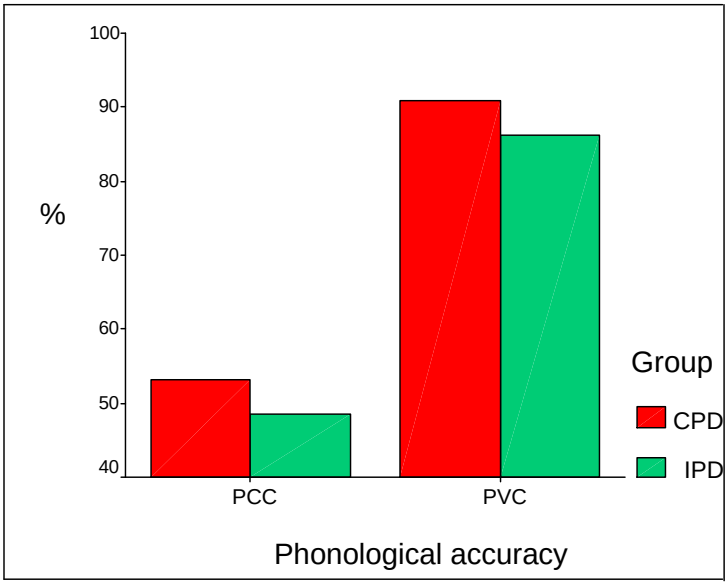

Figure 4: Error types as a percentage of total content- and function- word errors: IPD and SLI groups