THE EFFECTIVENESS OF THE OTAGO SCREENING

PROTOCOL IN IDENTIFYING SCHOOL-AGED STUDENTS

WITH SEVERE SPEECH-LANGUAGE IMPAIRMENTS

_____

A thesis submitted in partial fulfilment of the

requirements for the Degree

of Master of Speech-Language Therapy

in the University of Canterbury

by Jane Musgrave

University of Canterbury

2007

_____

Acknowledgements

Abstract

This study examined the effectiveness of the Otago screening protocol in identifying school-aged children with severe speech and language impairments. In order to do so, the results of the Otago screening protocol were compared with those of comprehensive language assessment as determined by best practice protocol (Gillon & Schwarz, 1998, Kennedy, 2002). Following the completion of the screening and the comprehensive assessments, an evaluation of the true positives and false positives was calculated, and an analysis of the false negative outcomes made.

Findings indicated that fourteen of the twenty participants were true positives, three were true negatives, three were false positives, and none were false negatives. The Positive Predictive Value and Negative Predictive Value of the screening protocol was 100%. Test Sensitivity and Specificity were very high at 82% and 100%. Inter-rater reliability was very high, generally ranging from 92-100%. Adding a standardised measure of phonological awareness would improve efficiency of the screening protocol. Consideration of alternative screening tools, such as the GAPS test (Gardner et al, 2006) and the CELF-4 screening test (Semel, Secord & Wiig, 2004), should be made. Additional factors which could influence a screening protocol are discussed.

The Otago screening protocol is a valid procedure to detect severe speech and language impairments in school-aged students referred to Special Education.

CONTENTS

FIGURES

Figure 1: Referral process for school-aged children at GSE Otago (p 16)

INTRODUCTION

Service delivery is an issue that speech-language therapists must address in their practice. Often there are too few therapists to address the needs of the large number of children presenting with speech and language disorders. In New Zealand, it is estimated that 18% of school-aged children require speech-language therapy services (Gillon & Schwarz, 1998). Often, service providers, such as the Ministry of Education, are constrained by budgetary issues and only a percentage of children requiring treatment can receive it. One way speech-language therapists address this issue is by over-delivering services relative to the contract. For instance, the speech-language therapy service to school-aged students who have severe speech-language impairments in the Otago district of Special Education has consistently over-delivered in relation to the contractual number. That is, too many students have received input from speech-language therapists compared to the contractual number for whom funding is received. In 2005-6, speech-language therapists at Special Education in Otago actually provided a service to 471 school-aged students with severe speech-language impairments, 170% of the contractual figure (Ministry of Education, 2006).

In order to address the level of over-delivery, it is important to examine what factors are contributing to the high numbers of children identified. One possible reason for over-delivery is that the screening protocol utilized may be too sensitive and may be identifying children as language impaired who are not. Other factors may include ineffective models of service delivery, inconsistent closure criteria, or that the contractual number for the target population may be too low. This study will examine one factor, the effectiveness of the Otago screening protocol in identifying school-

aged students with severe speech-language impairments.

Screening

Purpose of screening

Screening tools are those designed to determine whether someone is at risk of a disorder (Salvia & Ysseldyke, 1995), and should be able to face critical appraisal (Bamford, Davis, Boyle, Law, Chapman, Brown & Sheldon, 1998). Although screening tools are an initial part of an assessment, they are not designed for clinicians to make decisions about presence or absence of a language disorder (Kennedy, 2002). Rather they provide cursory information that allows a clinician to determine whether a child should be seen for a more in-depth assessment of speech and language (Gardner, Froud, McClelland, van der Lely, 2006, Kennedy, 2002).

Although a cursory measure, screening protocols are often used by schools or other service providers where large numbers of children are referred but where it is unreasonable to provide a full assessment (Paul, 2001). In those cases, screening procedures may entail administration of tests or tasks that evaluate a wide assortment of behaviours (Paul, 2001). When screening tools are used to inform decisions about service eligibility, factors that influence the outcome of the screening, such as psychometric properties (Paul, 2001), nature of the tasks used (Paul, 1995, Masterton, Bernhardt & Hofheinz, 2005) and timing of the assessment (Summers, Larson, Miguel & Terrell, 1996) should be considered.

Nature of the tasks used in a screening protocol

Kennedy (2002) recommended that not only the examination of all aspects of speech,

language, and communication skills is warranted in a screening protocol, but also the assessment of connected domains, such as fluency, voice, oral-motor and hearing. Additionally, Kennedy (2002) considered that related skill areas such as play and cognition should be examined.

It is recommended that valid screening protocols should use an assessment battery of standardised tools for language skills that examines verbal understanding in structured testing, and in the child's natural environment (Gillon & Schwarz, 1998). Information should also be provided about the range and complexity of grammatical structures used by the child, the vocabulary and concepts expressed, the pragmatic skills of the child and phonological skills, including phonological awareness development (Gillon & Schwarz, 2001). Consideration should also be given to any marked inconsistency between the child's receptive and expressive language abilities, and their symbolic play skills and auditory memory skills.

Recently, Gardner, Froud, McClelland & van der Lely (2006) claimed that language impairment can be detected by screening a child's grammatical and phonological skills only. A number of researchers claim that core syntactical deficits (grammatical, morphological and phonological markers) are often present in children with specific language impairment (Bishop, 1999, Conti-Ramsden, Botting & Faragher, 2001, Conti-Ramsden & Hesketh, 2003, van der Lely, 2005, van der Lely, Rosen & Adlard, 2004). On this basis, Gardner, Froud, McClelland & van der Lely (2006) developed a standardised screening tool to identify possible language impairment and likely reading impairment, known as the Grammar and Phonology Screening (GAPS) test. GAPS was normed on a large cohort (668) of children aged 3;04-6;06 years, across

many districts in the UK. GAPS has two subtests (sentence imitation and non-word repetition), takes only 10 minutes to administer, and can be carried out by professional and non-professional people. Gardner et al (2006) found that GAPS was an effective screening tool for language impairment in children in the early years of education.

A screening protocol should include information gathered from sources knowledgeable about the child, such as parents and teachers. Some research suggests that parental report should be the primary source of information. For example, The General Language Screen, developed by Stott, Merricks, Bolton & Goodyear (2002), is a series of twelve questions, to be completed by the student's parent/carer to detect possible speech and language impairment.

In addition to objective measures, using clinical judgment to identify speech and language difficulties is also emphasised in the literature (Tyler, 2005, Broomfield & Dodd, 2004, Lees & Urwin, 1997). Provocatively, Glascoe (1991) claimed that most developmental speech-language impairments could be reliably identified through professional judgments alone, and standardised assessments used sparingly. This is supported, too, by Rescorla & Alley (2001), who demonstrated the reliability and validity of the Language Development Survey, a teacher checklist, in accurately identifying preschool children with expressive language impairments. The GAPS test (Gardner et al, 2006) too is able to be administered by educators and carers alike, and the Social Communication Questionnaire (SCQ), a parental report, was used by Baird, Simonoff, Pickles, Chandler, Loucas, Meldrum & Charman (2006) to confirm autistic diagnoses in young school-aged children.

The outcome of a screening protocol is simply expressed as a 'pass' or 'fail', based on a predetermined 'cut off' eligibility score (Kennedy, 2002). This score is varied amongst researchers. Law, Lindsay, Peacey, Gascoigne, Soloff, Radford, Band, & Fitzgerald (2000), reported that some studies used a figure of -1.5SD on standardised tests as evidence of the presence of severe speech-language impairment. In contrast, other studies recognised -2.0SD as the level of reliable evidence of severe speech-language impairment (Law et al, 2000, Sanger, Aspedon, Hux & Chapman, 1995).

Factors Influencing Screening Results

Psychometric Properties of Screening Tools

Screening batteries should include reliable and valid assessment tools, appropriately scored and interpreted, in order that accurate prognoses and recommendations for the target group of students can be identified. Norm-referenced measures used in screening should fulfill reliable and valid psychometric criteria (Paul, 1995). Key factors that clinicians must consider when evaluating their screening protocols are the validity and reliability of the tools they are using.

Validity is the extent to which a test measures what it is designed to measure (Kennedy, 2002). There are several validity measures that could be considered. The examiner must be confident that a standardised tool accurately represents the language area being tested (content validity), and how that is achieved (construct validity). For example, if a language assessment purports to assess a child's semantic skills then the subtests of the assessment should provide details of the number of vocabulary items the child knows, the range of vocabulary items the child can use as

well as the child's skills in recalling and classifying vocabulary. The child's performance on a particular test should be comparable to that on a criterion-referenced measure (criterion-related validity), such as language sampling. For example, the examiner could expect to gain similar information on a child's semantic performance whether a standardised test or a language sample was used. Concurrent validity considers how the child's performance on the standardised and criterion-referenced tools is associated, and predictive validity examines the test's predictive value for future test performances. If semantic skills were being assessed, for example, then concurrent validity would describe the nature of the relationship between the child's semantic skills on a standardised measure and in a language sample. Whilst predictive validity ensures that the standardised test predicts the likely future semantic performances of the child and the path of the child's semantic skill development.

Reliability is an important psychological property. Reliability is the likelihood that the test can be relied upon to measure what it is claiming to. There are several types of reliability. These include test-retest reliability, that is, the likelihood that a child's test performance would be replicated if the test were repeated. Another type of reliability, inter-rater reliability, ensures that two raters would score the test performance similarly, and that there is association between the test's subtests (equivalent reliability or internal consistency).

The CELF-4 Screening Test (Semel, Wiig & Secord, 2004) is an example of a well standardised screening tool. The earlier CELF-3 Screening Test (Semel, Wiig & Secord, 1992) had undergone extensive validity and reliability measures, and was

further developed to produce the CELF-4 Screening Test (Semel, Wiig & Secord, 2004). The CELF-4 Screening Test was standardised on a sample of 1,200 American students aged 5;00 to 21;00 years with a wide range of demographic characteristics.

Several validity measures are provided. Convergent validity, comparing scores on the CELF-4 with scores on the CELF-4 core language standard scores, revealed mean scores ranging from 93.2 to 99.9 across age ranges. Test sensitivity, the likelihood that a test result would be positive in the presence of a disorder, and test specificity, the likelihood that a test result would be negative in the absence of a disorder, were both high at 0.88. The CELF-4 Screening Test also demonstrated excellent positive predictive power and negative predictive power.

Reliability measures of test-retest reliability and internal consistency were given for the CELF-4 Screening Test too. Whilst the test-retest reliability scores were high, ranging from 0.82 to 0.90, internal consistency was relatively low at 0.70 and 0.72. This difference was explained by the variation in language skills assessed in a screening tool.

Another concept important to the selection of a screening protocol is that it must identify the individuals who are impaired, but not miss those who are not. In order to judge the effectiveness of a screening assessment instrument, Feeney & Bernthal (1996) recommended that a high accuracy of correct identification must be found. That is, to be a valid screening protocol, the percentage of true positives found must be well above 50%. Additionally, there should be a low number of false-positives and false-negatives. In Feeney & Bernthal's (1996) study, multiple examiners and a

varying test battery were used, thus compromising reliability and validity of the screening protocol.

Timing of Testing

Not only is the choice of assessment tools important in a screening protocol, but the timing of the protocol can be relevant to the outcome. Summers, Larson, Miguel & Terrell (1996) compared the results of the CELF-Revised Screening Test (Semel, Secord & Wiig, 1989) with those of the Bankson Language Test- Second Edition (Bankson, 1990) on 211 students during their first year at school. Both tests yielded similar results, but up to one third of the children were inconsistently identified as having speech and language impairments according to the timing of their assessment. Students scored higher on later administered tests over a seven month period. If children are assessed under a screening protocol early in their first year at school, the outcomes may provide a high number of false positive results.

In summary, screening protocols should include standardized assessment tools with strong psychometric qualities. Screening should be comprehensive, sampling all language domains, and information gained from all sources relevant to the child.

Speech and language screening

It has been suggested that screening protocols should cover a breadth of speech and language domains (Gillon and Schwarz, 1998, Kennedy, 2002). Although it may be sufficient to sample a small range of linguistic behaviours in screening (Gardner et al, 2006), it is generally accepted that some assessment of phonology and language behaviour should be included in a speech and language screening.

Speech and language screening: phonology

Phonological impairment may present in a variety of ways. There may be evidence of motor speech impairment, trouble with storing, planning and retrieving speech sound representations, or problems in processing speech sound knowledge. The screening protocol used to identify phonological problems must take account of the range of possible speech sound difficulties, and be sensitive enough to recognise them (Tyler, 2005, Paul, 1995).

To make a preliminary differential diagnosis and to decide if an in-depth evaluation is called for, Gillon & Schwarz (2001) supported the use of a screening measure as the first stage in the assessment process. A single word elicitation task may be sufficient to provide representative information of a student's phonological system (Masterton, Bernhardt & Hofheinz, 2005). Whilst there are several examples of single word assessment tools, Kennedy (2002) suggested that the Quick Screen of Phonology (Bankson & Bernthal, 1990) was a particularly useful standardised phonological assessment tool for a screening protocol.

Screening for phonological impairment should include not only a description of the student's speech patterns and identification of the underlying deficit, but also an evaluation of phonological awareness skills and literacy achievements. Literacy is widely accepted as an essential component of language by many (Beitchman, Nair, Clegg & Patel, 1986, Catts, Fey, Tomblin & Zhang, 2002, Gillon and Schwarz, 2001, Law, Garrett & Nye, 2003, Nathan, Goulandris, Stackhouse & Wells, 2004), and as such, phonological awareness skills, as prerequisite skills for literacy development, should be ascertained in both comprehensive and screening assessments. In contrast,

some researchers (Tyler, 2005, Paul, 2002) consider phonological awareness as a related language skill, an optional extra piece of data. The Preschool and Primary Inventory of Phonological Awareness (Dodd, Crosbie, McIntosh, Teitzel, & Ozanne, 2000) is an example of a standardised screening tool with sound psychometric properties.

Speech and language screening: language

Screening protocols for language often include samplings of behaviours across all language domains of syntax, semantics and pragmatics in both expressive and receptive modalities. Standardised screening protocols for language include the Fluharty Preschool Speech and Language Screening Test (Fluharty, 1978), and the Clinical Evaluation of Language Fundamentals-3 Screening Test (Semel, Wiig & Secord, 1996). This latter assessment has recently been superseded by the Clinical Evaluation of Language Fundamentals- 4 Screening Test (Semel, Wiig & Secord, 2004) which may offer similarly useful information.

At present, reliable, consistent, diagnostic standards for language impairment are not available. Some aspects, such as pragmatic elements (Tomblin, Records & Zhang, 1996), are often not included in assessment measures whilst other areas are not as thoroughly assessed as they ought to be, such as semantic development (Brackenbury & Pye, 2005). Since very few children have social or semantic language impairments that are not manifest in the syntactic and phonological domains (Gardner et al, 2006), screening tools that examine grammar and phonology may be sufficient to detect all children with language impairment.

Several researchers noted the particular sensitivity of the Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals, 4th edition (Semel, Secord & Wiig, 2004) in identifying specific language impairment (Lloyd & Paintin, 2006, Conti-Ramsden, Botting & Farragher, 2001). The first subtest of the GAPS test, developed by Gardner et al (2006), is a sentence imitation task on the basis that short term memory recall is a significant feature of language impairment. The scores of the Recalling Sentences sub-test by participants in this study may provide useful information in detecting the presence of language impairment.

The New Zealand Context

In New Zealand, the main provider of speech-language therapy services to school-aged children is Special Education, part of the Ministry of Education. Special Education comprises sixteen districts in New Zealand. The Ministry of Education contracts Special Education to provide specialist services to school-aged children who have severe behaviour needs, to children who have severe speech and language impairments, and to children who have complex special educational needs (these students are verified and on the Ongoing and Reviewable Resources Schemes [ORRS]). Special Education also provides specialist services to pre-school children who have moderate or severe behaviour needs, moderate or severe speech-language impairments, or complex special educational needs.

Screening in NZ Special Education

At Special Education in New Zealand, screening protocols are used to determine eligibility for service for school-aged children. Under the current contractual arrangements, children are eligible for service if they present with severe speech-language impairment.

Population-based funding, based on international prevalence rates, determines the contractual numbers for each district of Special Education. This is translated into the number of school-aged children within a district that are funded to receive a comprehensive speech-language therapy service.

The Otago situation

In the Otago district of Special Education, 8.3FTE speech-language therapists are currently employed to provide speech-language therapy services. Approximately 2.56FTE of these speech-language therapists deliver service to those school-aged children who have severe speech-language impairments.

During the financial year 2005-6, speech-language therapists at Special Education in Otago were contracted to provide services to 246 school-aged children with severe speech-language impairments. This contractual figure is made up of current cases and children who receive a screening assessment, and is sourced from population-based funding formula of primary school roll figures of 5-8 year old students, the target population.

In recent years, the speech-language therapy service to school-aged children who have severe speech-language impairments in the Otago district of Special Education has consistently over-delivered in relation to the contractual number (Ministry of Education, 2006). That is, too many children have received service from speech-language therapists compared to the contractual number for whom funding is received. In 2005-6, speech-language therapists at Special Education in Otago actually provided a service to 471 school-aged children with severe speech-language

impairments, 170% of the contractual figure (Ministry of Education output reports, 2006).

Consideration of the current screening protocol used in the Otago district of Special Education may not only highlight aspects of the process that are working well, but expose some weaknesses in the process, in order that the over-delivery issue may be addressed.

The Referral process

Under the Ministry of Education's communication initiative, school-aged children may be referred for an evaluation of their speech-language skills to their local Special Education office. Districts vary in their acceptance of referral agent source. In some districts, referrals are accepted from schools and educational professionals only. In other areas, including Otago, an open referral system operates. Under these conditions, referrals may come from concerned school staff, parents/carers/whanau, health practitioners or other education professionals, such as Resource Teachers of Learning and Behaviour (RTLB), and Special Education colleagues. In Otago, referrals are accepted in writing, which could be an e-mail, or by phone.

The referral form provides demographic identifying information, optional further details regarding the nature of the concern, parental consent and to which service the referral is made. Increasingly, documentation, as evidence of school concern, such as an Individual Education Plan (IEP) and a completed Junior Oral Screening Tool (JOST), is being recommended to accompany the referral. Upon receipt of the referral, it is date stamped and passed to a service manager who decides if the referral

is likely to meet the eligibility criteria for a speech-language therapy service (sometimes in discussion with the key speech-language therapist, lead practitioner, in the district). Then the referral is allocated to the relevant speech-language therapist, depending on locality of the school facility. It is processed by a member of the administration team, who opens a case file, and sends an acknowledgement letter to the referring agent.

In Dunedin, the main Otago office, the referral is presented and discussed at a weekly meeting of speech-language therapists, chaired by the lead practitioner. Outcomes of previous assessments are discussed here too, to maximize inter-rater reliability of the screening protocol. The lead practitioner maintains a database of referral information for the Dunedin area.

If Maori or Pacific Island ethnicity is recorded on the referral form, a copy of the referral is also passed to the district Maori advisor in the Dunedin office. The district Maori advisor passes the referral on to a kaitakawaenga, who then engages in discussion with the relevant speech-language therapist. The kaitakawaenga and speech-language therapist collaboratively decide on the most appropriate cultural measures to establish a positive working relationship with the whanau concerned. The kaitakawaenga or speech-language therapist may make the first contact with the whanau, or a joint visit may be arranged. The kaitakawaenga will maintain involvement with the whanau as long as they wish. This helps ensure that cultural issues are addressed.

Adherence to professional guidelines, provided by the Royal College of Speech and

Language Therapists (2001), ensures that within six to eight weeks of receiving a referral, some contact is made by a speech-language therapist. Usually, this will be an appointment to attend for an initial screening of the child's speech and language skills. The speech-language therapist will collaboratively arrange an appointment with the child's facility and parents/carers. This usually takes place in the child's school facility. Parents/carers are invited to attend, and informed of the purpose and nature of the appointment, and the child's teacher is consulted, if possible. If the screening protocol is likely to be delayed, this will be communicated to the child's parents and school, and the likely timing of the appointment will be advised (See Figure 1).

Figure 1: Referral process for school-aged children at GSE Otago

Referral made

↓

Referral received by GSE Otago office

↓

Administration team send acknowledgement letter

↓

Case file created

Referral added to data base ← Case file created → Maori/ Pasifika ethnicity

File given to service manager

Copy of referral given to district Maori advisor

↓

File allocated to relevant speech-language therapist

Collaboration

File allocated to kaitakawaenga

↓

Referral discussed at weekly SLT referral meeting

↓

SLT contacts parents & school within 6-8 week timeframe

Kaitakawaenga contacts whanau

↓

SLT screening protocol administered

Referred to other agency ← SLT screening protocol administered → File to service manager

↓

Ineligible- file closed

Eligible- await comprehensive service

The priority checklist

Decisions regarding eligibility for speech-language therapy service may not be dictated wholly by severity of need (Whitmire, 2002, ASHA, 2000). Other factors such as the child's age, significant background information, the extent of parental and teacher concern, difficulty in accessing the curriculum, and support available, may be considered.

At Special Education in New Zealand a 'priority checklist' (see Appendix A) is used to determine eligibility for a comprehensive service. At, or immediately subsequent to, the screening appointment, a priority checklist is completed by the speech-language therapist. The checklist document currently used was devised by speech-language therapy managers at Special Education in the late 1990s, and adopted nationally in the SE2000 [Special Education 2000] policy by all speech-language therapists working with school-aged students in Special Education, under the communication initiative. It was intended to guide speech-language therapists to identify those children who have severe speech-language impairments, and are therefore eligible for a comprehensive speech-language therapy service, those children who have moderate speech-language impairments, and are eligible for a home or school programme, and those children who have mild speech-language impairments and who are not eligible for service, but may receive advice. The priority checklist has never been formally validated. Instead the checklist drew on the knowledge of experienced speech-language therapists, and not an evidence base.

The checklist also reflected the priorities of the contract that Special Education held with the Ministry of Education for school-aged children with severe speech-language

impairments. Namely, the targeting of services to the 5-8 year old population, to addressing difficulties in accessing the curriculum of students with speech-language impairments, and to the specific communication needs of language impairment, phonology problems, and voice and fluency difficulties.

Using this form, the speech-language therapist can score the child one point for the presence of each of several factors; age of 5-8 years, the unlikelihood of spontaneous recovery, significant background information, the extent of parental and teacher concern, and the impact of the child's speech-language impairment on accessing the curriculum. The child's specific communication needs can score up to three points each in the areas of speech, language, fluency and voice. These scores are derived from the speech-language therapist's interpretation of the child's performance on the standardised measures used. However, no guideline attainment scores are provided, such as -1.0SD test score translates to a score of 2 for language impairment on the checklist or -1.5SD scores 3. Neither is there a description of how to judge, for example, the extent of parental or teacher concern. Thus, the priority checklist may be significantly flawed by subjective bias.

Recently, there has been national variability amongst Special Education districts in the eligibility points required on the priority checklist. Originally, a score of seven was needed for the child to qualify for a comprehensive therapy service from a speech-language therapist at Special Education, but currently there is evidence of some districts using eight or nine points as the eligibility criteria. This seems to have been a response to increased referral rates, and fixed contractual numbers. In Otago, the benchmark figure of seven is still used. A score of between four and six inclusive,

identifies the child as having moderate speech-language difficulties, and results in a home or school programme being set, and a score of three or less equates to a child with mild speech-language difficulties and the provision of advice only. A recommendation to refer on, such as to private providers or other services, may also be made.

The priority checklist has the advantage of being a process available nationally within Special Education, and of taking account of significant factors other than communication needs that may affect the child's progress in speech-language therapy. Its limitations include inconsistent use across districts, and within districts, and the major drawbacks associated with subjectivity. In addition, 'priority checklist' is a confusing title for an instrument which is in fact a determinant of eligibility, and little to do with prioritising the nature of the comprehensive speech-language therapy service to be offered.

Assessment practices

A screening protocol is used to determine the presence of severe speech-language impairment, and to recommend eligibility for service provision. Both objective measures and clinical judgment are involved. The Otago screening protocol samples both receptive and expressive language skills, including a phonological test. Further investigation, using a comprehensive assessment battery, would be needed to provide detailed diagnostic information of all domains of language development.

Referral information

Information provided on the referral document guides the speech-language therapist to the nature of the screening protocol to be offered. Comments made by the teacher on the referral form expressing concern about a child's language skills, such as 'does not follow class instructions' or 'uses one or two word sentences', would result in an appraisal of the child's receptive and expressive language skills. If the referrer gives information on the referral form suggesting the child has speech sound difficulties, such as 'says t for k, and w for r' or 'unintelligible', then a phonology assessment only may be completed. Referral comments suggesting inappropriate voice or fluency issues, such as 'harsh voice' or 'stumbles at the beginning of words' would result in a screening protocol related to these issues being followed. Thus the information provided by the referrer is important in guiding the assessment carried out by the speech-language therapist.

The accuracy of the referral information is dependent on the skills and knowledge of the referring agent, usually a teacher, parent or, occasionally, GP or health colleague. If the speech-language therapist has an ongoing relationship with the school, a teacher may discuss a child's speech and language skills with the therapist prior to making a referral. Such discussion is more likely to ensure that an eligible referral is made for speech-language therapy to Special Education, that is, the child does have severe speech-language impairment.

Investigating appropriate speech-language therapy referrals by teachers, Sanger, Aspedon, Hux and Chapman (1995) found that such referrals were often appropriate. Teachers' judgments of a child's speech and language skills were measured through

the completion of an observational checklist, and confirmed by the researchers'
completion of a battery of standardized language measures. One year later, the
researchers found that 75% of those at-risk children had been referred to the local
speech-language therapy service. There was a reported high correlation between those
children who were identified by the teachers as having marked speech-language
difficulties, and those who were subsequently receiving speech-language therapy.

Standardised Tools

In Otago, the screening protocol for language impairment examines verbal
understanding using the Bureau Auditory Comprehension Test (Bureau of Maternal
and Child Health, 1999), and expressive language using the Renfrew Action Picture
Test, 4th Edition, (Renfrew, 1997). A phonological assessment is also administered to
screen for possible speech sound difficulties. Speech-language therapists use the
Goldman-Fristoe Test of Articulation-2 (Goldman & Fristoe, 2000), or the Tauranga
Articulation Test (De Candole, 1986) or more recently, the New Zealand Test of
Articulation (Ministry of Education, 2004), which was developed and standardised on
the New Zealand primary-school population in 2004. A full description of the tests
follows:

Bureau Auditory Comprehension Test [Bureau] (Bureau of Maternal and Child
Health, 1999)

This is a test of a child's understanding of verbal language structures, suitable for
children aged 2;00-7;00 years of age. The test comprises four pages of pencil drawn
pictures of familiar objects, and a coin or counter is needed. Scoring consists of one
point for each correct answer, with deductions made for the numbers of repetition of
each instruction (0.2 each for boys and 0.34 for girls). The Bureau was standardised

on 768 children in Australia, with 44 more girls than boys within the cohort. Using the tables provided, the child's score can be compared to the mean score for their chronological age, and using 3.32 as 1 SD (standard deviation), the average range of scores for the child's age can be calculated.

There is a lack of any other statistical information regarding reliability and validity constructs, so the robustness of this tool is called into question. Having considerable gender differences within the standardised population is a concern which may cause bias to the normative data. Since speech-language therapists typically have far more males than females on their caseload, reflecting the gender bias in the international incidence of language impairment, the standardised population may not have been a representative sample.

The Renfrew Action Picture Test, 4th Edition [RAPT] (Renfrew, 1997)

The Renfrew Action Picture Test (Renfrew, 1997) was designed as a screening test for expressive verbal language impairment. It aims to evaluate the information and grammatical structures used by the student, in response to questions asked. Information is examined according to the student's use of nouns, verbs and prepositions. Grammatical aspects assessed include regular present, past and future verb tenses, irregular past verb tense, passive sentence forms, and complex sentence constructions. Regular and irregular plural noun use is also measured.

The Renfrew Action Picture Test was standardised for use with children aged 3;00 to 8;00 years. Ten coloured pencil-drawn pictures are used with stimulus questions or

directions, and the child's verbal responses are recorded. Scoring is made by comparing the child's utterances to the two tables of acceptable responses for Information and Grammar. Statistical data is provided that enables comparison of the child's scores to the mean score, the middle half of the range and the standard deviation for each six month band between 3;06 and 8;05. Recommendations are made according to the child's emerging profile, such as if there is a particular discrepancy between the Information and Grammar scores. The test was standardised on 594 children within the United Kingdom, of class III socio-economic status, with equal numbers of boys and girls.

In terms of reliability evidence, no test-retest reliability was conducted, except for some random testing where the number of participants was not disclosed. Inter-rater reliability was attempted through two processes. Firstly, recording reliability was ascertained by two examiners simultaneously recording responses to 341 pictures. The manual stated that 30 discrepancies were found, 11 of these making a difference in scoring, but no statistical data was applied to these results. Secondly, scoring reliability was examined by four speech-language therapists scoring sample sets of responses from 12 children. Whilst 3% discrepancy was found initially, the author made such corrections to the scoring table, that 100% agreement in scoring Information and 98% in Grammar was subsequently reported.

Validity data is sparse. A small unpublished study is recounted of comparisons between the performances of 5-6 year old students with 'moderate learning difficulties' and age-matched peers on the Renfrew Action Picture Test and the Carrow Elicited Language Inventory (Brown, 1988). Since no specific information

about this study is given, such as the number of participants, demographic data, how the participants were recruited, the method used and specific results found, it must be discounted as validity evidence.

Whilst the Renfrew Action Picture Test may be efficient in identifying syntactic errors, it may not recognise semantic deficits. There is a relatively small vocabulary set required to score at age-appropriate levels. The Renfrew Action Picture Test has undergone minor review since it was developed in the 1970s. The line drawn pictures now appear outdated, and unappealing to current primary aged children.

Further, the UK standardisations may not transfer accurately to the New Zealand context. Whilst there are similarities in the cultures of the UK and New Zealand, for example both are considered to be Western European societies, there are also critical ethnic and cultural differences in the population sets. The Renfrew Action Picture Test has not been standardised on Maori and Pacific Islanders so cannot be used reliably on these populations. With differences in ethnicities, come linguistic and cultural differences. The vocabulary and sentences structures of New Zealand English may not be accurately represented in the Action Picture Test. Finally, the lack of any reliability and validity constructs call into question the robustness of the Action Picture Test.

.

New Zealand Articulation Test [NZAT] (Ministry of Education, 2004)

The NZAT (Ministry of Education, 2004) was developed as a screening tool for articulation in 2004. It was standardised on 1014 New Zealand children aged 5;00 to

8;00 years, attending 53 school facilities, covering all geographical areas. It is a single word elicitation test, using a vocabulary of 100+ words. There are five subtests; single consonant sounds, blended initial consonants, vowels, multi-syllabic words, and a conversational speech sample (obtained using a complex picture). Only the single consonant sounds and blended initial consonants subtests are normed, as these are the subtests that are intended to be used as a screening assessment measure. The remaining subtests may be used to provide qualitative information. Additionally, stimulability testing can be carried out to assist therapy planning.

Rationale is given as to why the vowels, multi-syllabic words and conversational sample subtests were not subject to standardisation. Vowels are rarely in error after the age of 5;00 years (the target group for the test) and are intended for further assessment only. The multi-syllabic and conversational sample subtests are considered important in the clinical decision-making process to guide therapy, but not within a screening process.

Statistical data are strong. Acceptable internal reliability scores are reported, using co-efficient alpha. Median scores were 0.93 for boys and 0.92 for girls. To examine test-retest reliability, 47 students were reassessed on the NZAT within 2-3 weeks of the initial assessment. The median percentage agreement of speech sounds was 100%. Inter-rater reliability was not measured. Validity information is given for each subtest. Construct validity results are reported to be comparable to the results of other articulation tests, with similar ages of speech sound acquisition.

The NZAT is presented as a set of coloured clip-art pictures, with as many as 12 pictures to a page. For some children with attention difficulties or younger children, the number of pictures on the page may prove too distracting. The vocabulary has been carefully chosen to be well known to New Zealand students within the target age range of 5;00 to 8;00 years, and to avoid phonemic contrasts. Occasionally, cloze procedure phrases are recommended for the examiner to stimulate a particular word.

Raw scores can be converted to standard scores and percentiles for each 12 month band in the target age group. P Values are detailed for each consonant sound and blended initial consonant by gender. Further statistical analysis by the author allows comparison of a child's speech sound profile with a typical child's speech sound development (Moyle, 2005). For example, the age at which 95% of children have acquired a particular speech sound.

The NZAT is likely to be a sound phonology assessment tool and provide valid information, as it was standardised recently on the target group in the New Zealand population, provides strong statistical support, and has current familiar vocabulary items.

Parental Report

Parental concern may be ascertained through face-to-face contact at the screening appointment, or by phone. In Otago, up to 50% of screening appointments go ahead without parental attendance. A small number of parents choose not to attend because they feel that the speech-language therapist will obtain a more accurate picture of their child's communication skills in their absence. These parents make their feelings

known to the speech-language therapist beforehand, and often request information and discussion around the speech-language therapist's findings afterwards. Another minority group of parents are not able to attend due to work commitments, and make their anxieties clear.

Of concern are the considerable numbers of parents/carers who do not attend the screening appointment nor make contact, and therefore, who are not engaged in accessing the speech-language therapy service. The reason for this non-contact is unclear. It could be that these parents have poor experiences of the education system themselves, or that cultural expectations are different. Alternatively, it could be that their child's communication skills are not considered a priority issue, or that the parents/carers do not recognise the extent of their child's difficulties. However, parental involvement should be considered an integral component of service provision, encompassing the screening protocol and decision-making process (Lindsay & Dockerell, 2004, Paul, 2002).

Summary

Several factors which may affect the outcome of a screening protocol are described. A broad sampling of language behaviours must be taken (Gillon and Schwarz, 2001), using standardised tools with strong psychometric properties (Kennedy, 2002), from all sources relevant to the child. It may be possible for non-speech-language therapists to identify language impairment using a standardised screening tool of grammar and phonology (Gardner et al, 2006), and timing of the screening may be important too (Summers, Larson, Miguel & Terrell, 1996).

In the current Otago district screening protocol, professional standards may be compromised. There may be shortcomings within the standardised tools and in the referral process. Under circumstances of variability and confusions around the referral process for school-aged children, high numbers of referrals may be made which are likely to be ineligible for comprehensive speech-language therapy service.

For preschool children, Special Education holds an early intervention contract with the Ministry of Education to provide speech-language services to children who have severe or moderate speech-language impairments. This differing level of speech-language needs served by therapists working in early intervention compared to school-focus therapists can be confusing for parents/carers and schools. False expectations of the school-based speech-language therapy service may be experienced by parents/carers and teachers once the child starts school.

Under the current regime of over delivery, an increased waiting time for a comprehensive speech-language therapy service by speech-language therapists at Special Education in Otago results. At present, the waiting time for a comprehensive therapy service is approximately 10-20 weeks. A bottleneck of existing cases awaiting comprehensive service is created, increasing pressure on the speech-language therapist with their current caseload. This, in turn, may influence the types of service delivery models being operated, leading to less efficient services, and impacting on a therapist's time available to screen new referrals.

In Otago, for several years now, speech-language therapy services at Special Education have consistently and significantly over-delivered in relation to the

contractual number, consisting of all new and existing cases, for each financial year. In 2005-6, the contractual figure stood at 246 whereas speech-language therapists provided a service to 170% more school-aged children (Ministry of Education output reports, 2006). There could be several explanations for this. One explanation may be that the screening protocol may not be appropriate. Given some of the limitations of the tests and method used in the screening protocol, it was deemed important to investigate the usefulness of the Otago screening protocol more closely. This research aims to examine the effectiveness of the Otago screening protocol in identifying school-aged children with severe speech-language impairments.

METHOD

This study examined the effectiveness of the screening protocol used by speech-language therapists working with school-aged students at the Otago office of Special Education. The screening protocol was administered to twenty primary-aged students and compared with the results of a battery of comprehensive assessments performed on the same twenty students. Statistical analysis was applied to the results that provided test sensitivity and specificity, and positive and negative predictive values. A qualitative analysis was made of the false negative outcomes. Inter-rater reliability was applied to 20% of the assessments.

Ethics

Approval in principle was sought and obtained from the Regional Manager, District Manager and Service Managers at Special Education in October 2005. Ethical approval was obtained from the Human Ethics Committee at the University of Canterbury in February 2006.

Participants

Twenty participants were selected, by order of referral, from referrals to the school-age speech-language therapy service at Special Education in Otago during January to May 2006. Referrals of students aged 5;00-8;00 years, where concern about the student's language skills was noted in the referral, voluntarily participated in the study. All participants lived in urban areas. Exclusion criteria were applied. The participants had no known hearing or visual difficulties, no physical difficulties, no

difficulties with fluency or voice production, and no diagnosed medical syndromes

(See Table 1).

Table 1

 Demographic details of participants

Participant	Age	Gender	Referrer	Presenting problem
1	5;05	M	GSE-Psychologist	Language concerns
2	5;04	F	Parent/school	Language & speech concerns
3	5;04	M	GSE-EI	Language concerns
4	5;03	M	GSE-EI	Delayed language & speech
5	5;02	F	Health	Language concerns
6	5;03	M	Health	Delayed language
7	5;01	M	GSE-EI	ASD/SLI?
8	7;05	F	GSE-Canterbury	SLI
9	6;04	M	Parent/school	Language & speech concerns
10	5;00	M	GSE-EI	Delayed language & DVD
11	6;00	M	School	Delayed language
12	5;10	F	School	Delayed language
13	5;08	M	School	Delayed language
14	5;05	M	School	Delayed language & speech
15	5;00	M	GSE-EI	Delayed language
16	5;01	M	Parent/school	Delayed language & speech
17	5;03	M	Parent/school	Language concerns
18	6;10	M	RTLB	Language concerns
19	5;02	M	GSE-EI	Delayed language & speech

20	5;06	F	Parent/school	Delayed language & speech

Note: M=male, F=female, GSE=Group Special Education, EI=Early Intervention, ASD=Autistic Spectrum Disorder, SLI=Specific Language Impairment, DVD=Developmental Verbal Dyspraxia, RTLB=Resource Teacher of Learning & Behaviour

Materials

Screening protocol

The assessment tools used in the screening protocol were the Bureau Auditory Comprehension Test (Bureau of Maternal and Child Health, 1999), the Renfrew Action Picture Test, 4th Edition (Renfrew, 1997), and the New Zealand Articulation Test (Ministry of Education, 2004). The priority checklist, an internal document of Special Education, was completed for each participant (see Appendix A).

In using the Bureau Auditory Comprehension Test and the Renfrew Action Picture Test, an age-equivalent of 1;06 years below a student's chronological age level is considered by speech-language therapists at Special Education in Otago to indicate severe language impairment for school-aged students, following the priority checklist criterion. This criterion of 1;06 below chronological age level was used in this study to indicate severe speech-language impairment.

Using the single-word phonological assessment, the NZAT, evidence of three or more phonological error processes, such as stopping, cluster reduction, fronting, is used to indicate severe speech-language impairment (using the priority checklist criterion).

This criterion of evidence of three or more phonological error processes was used in this study to indicate severe speech-language impairment.

Comprehensive assessment

For the battery of comprehensive assessment, the Clinical Evaluation of Language Fundamentals, 4th Edition, Australian adaptation (Semel, Wiig & Secord, 2004) and The Preschool and Primary Inventory of Phonological Awareness (Dodd et al, 2000) were used. In addition a structured language sample of 50+ utterances was taken. This was recorded on a high quality recording device, a Digitor AC/DC desktop recorder, model A-4031. The priority checklist was reapplied for each participant.

Severe language impairment was considered to be indicated by a performance where there was a percentile rank of 5 or less on the Core Language Score of the CELF-4 (Semel, Secord & Wiig, 2004), or a percentile rank of 5 or less on the majority of subtests of the PIPA (Dodd et al, 2000). Within the language sample, severe language impairment was considered to be indicated by an MLU considerably below the level expected for the student's chronological age, restricted vocabulary or possible word finding difficulty, several grammatical errors, a lack of or few complex sentences, a lack of or few conversation initiations, or any unusual language features (Crystal, 1986, Ministry of Education, 1998, Speech Pathology Australia, 2005).

The Clinical Evaluation of Language Fundamentals, 4th edition [CELF-4] (Semel, Secord & Wiig, 2004)

The CELF-4 is designed to provide an accurate tool for the diagnosis of language and communication impairment in 5-21 year old students. It is used and recommended

extensively in research studies as an accurate measure of language impairment (Condouris, Meyer & Tager-Flusberg, 2003, Frazier Norbury & Bishop , 2003, Stott, Merricks, Bolton & Goodyear, 2002). The CELF-4 is a comprehensive assessment tool that examines both comprehension and production of language. The first level of the test, the Core Language Score, is made up of four subtests: Concepts and Following Directions, which assesses comprehension of linguistic concepts and syntactic structures such as conjunctions; Word Structure which evaluates comprehension and production of syntactic structures; Recalling Sentences which assesses imitation of sentences of increasing length and complexity; and Formulated Sentences which examines the creation of sentences around a given word. Subtests are scored individually and then a Core Language Score is derived. Through calculation of the subtest age-equivalent, the subtest scaled score and percentile figure, the Core Language Score is calculated.

The CELF-4 has strong psychometric properties. In 2004, the CELF-4 was standardised on a normative sample of 825 children and adolescents in Australia. There was an equal gender bias, participants came from all Australian states and territories, and from urban and rural backgrounds.

In terms of reliability, the average reliability coefficient of the Core Language Score is very high at 0.96. Within the subtests, the average reliability coefficients range from 0.79 for Word Structures to 0.92 for Recalling Sentences (See Table 2).

Table 2

Average reliability co-efficient of CELF-4 subtests

CELF-4 subtest	Average reliability co-efficient
Concepts & Following Directions	0.86
Word Structure	0.79
Recalling Sentences	0.92
Formulated Sentences	0.85
Core Language Score	0.96

Confidence levels at 68%, 90% and 95% are given for each age group from 5-21 years. For the ages of participants in this study, with respect to the Core Language Score, confidence levels are given for ages 5;00-5;11, and for ages 6;00-7;11 (See Table 3).

Table 3

Confidence levels for age groups in this study

Confidence Interval	68%	90%	95%
Age 5;00-5;11	3	4	5
Age 6;00-7;11	3	6	7

The CELF-4 manual provides three sources of validity evidence supporting the Australian standardisation results. These are based on content validity, response process and internal structure of the test. In terms of content validity, there are several

examples of cited research evidence to support the domains of language sampled in the CELF-4.

The internal structure of the CELF-4 is documented through inter-correlational and factor analysis evidence. The mean correlation scores are given for all ages. The Core Language Score has a high correlation, ranging between 0.65 and 0.78, with other language measures. Evidence of construct validity is taken and extended from research obtained in the CELF-3 development process.

The Preschool and Primary Inventory of Phonological Awareness (Dodd, Crosbie, McIntosh, Teitzel, & Ozanne, 2000)

The identification of children who have poor phonological awareness skills is the stated purpose of the Preschool and Primary Inventory of Phonological Awareness (PIPA). The target population is children aged 3-7 years. It has coloured line drawings, is quick to administer, taking approximately 30 minutes to complete. There are 6 subtests; Syllable Segmentation, Rhyme Awareness, Alliteration Awareness, Phoneme Isolation, Phoneme Segmentation and Letter Knowledge. The first five subtests examine critical phonological awareness skills such as the recognition, detection and manipulation of phonemes, whilst the latter assesses phoneme-grapheme correspondence (naming letters and phonemes). Scoring for the PIPA is done by converting the child's raw score into a normalised standard score and percentile score, using the most appropriate norms (Australian or UK). A standard score of 10 is given for the mean, and standard deviation is 3.

The PIPA was standardised on two populations of Australian and UK children. The Australian sample comprised 583, and the UK population numbered 595 children aged 3;00 to 6;11 years. Similar gender and socio-economic status of the two groups was sought. However, whilst the geographical information for the UK sample is provided, showing widespread distribution across the UK, there is no such information for the Australian norms. This may create some bias to the Australian sample norms. Neither is there evidence of urban versus rural spread in both population samples.

Reliability measures of internal consistency, test-retest reliability and inter-rater reliability are given. Internal consistency data for the UK sample is provided. The reliability coefficients for internal consistency range from 0.70 for phoneme segmentation to 0.98 for letter knowledge (see Table 4). Thus, evidence of acceptable levels of internal consistency for the UK sample was present.

Table 4

Internal consistency reliability coefficients for PIPA subtests

PIPA subtest	Alpha*
Syllable Segmentation	0.84
Rhyme Awareness	0.83
Alliteration Awareness	0.84
Phoneme Isolation	0.91
Phoneme Segmentation	0.70
Letter Knowledge	0.98

*Alpha scores rounded to two decimal places

Test-retest reliability was obtained from 42 tests conducted over an interval of two

weeks by the same tester (see Table 5). Whether this data was collected from the UK

or Australian sample or both samples is unclear. Whilst all these scores are

significant, the score for phoneme segmentation suggests this is a less reliable subtest.

Table 5

Test-retest correlations for PIPA subtests

PIPA subtest	Pearson correlation*
Syllable Segmentation	0.69
Rhyme Awareness	0.87
Alliteration Awareness	0.80
Phoneme Isolation	0.95
Phoneme Segmentation	0.33
Letter Knowledge	0.98

*correlation scores rounded to two decimal places

Inter-rater reliability was measured on a very small sample of six children. The two

examiners' test scores were subjected to independent t-tests (see Table 6). The results

revealed no significant differences between the examiners' scores.

Table 6

Inter-rater reliability for PIPA subtests

PIPA subtest	t value	Significance (2 tailed)
Syllable Segmentation	-1.150	0.277
Rhyme Awareness	0.000	1.000
Alliteration Awareness	-0.143	0.889
Phoneme Isolation	-1.581	0.145
Phoneme Segmentation	-0.632	0.341
Letter Knowledge	0.099	0.923

 Several measures of validity were carried out; content validity, concurrent validity, criterion-related validity and construct validity. Since the PIPA subtests were based on findings of international research, content validity is provided. Concurrent validity was obtained by comparing the scores of the two subtests that the PIPA and the Phonological Abilities Test [PAT] (Muter, Hulme & Snowling, 1997) have in common (see Table 7). These showed significant correlations, and thus concurrent validity, between the two tests

Table 7

Concurrent validity between PIPA and PAT scores

	PAT	
PIPA subtest	Rhyme detection	Letter knowledge
Rhyme detection	0.631*	0.612*
Letter knowledge	0.590*	0.916*

* Correlation is significant at the 0.01 level (2 tailed)

Criterion validity was measured by comparing the performance of speech-disordered children on the PIPA and the Test of Early Reading Ability [TERA] (Reid, Hresko & Hammill, 1989). The performances of thirty Australian children, referred to the University Speech and Language Therapy clinic, were correlated using the TERA and five of the PIPA subtests (see Table 8), showing significant criterion validity.

Table 8

Criterion validity of the TERA and PIPA subtests

PIPA subtest	TERA correlations
Syllable Segmentation	0.440*
Rhyme Awareness	0.473**
Alliteration Awareness	0.499**
Phoneme Isolation	0.389*
Phoneme Segmentation	0.457*

 *Correlation is significant at the 0.01 level
**Correlation is significant at the 0.001 level

An evaluation of the inter-correlations between the PIPA subtests was used to ascertain construct validity (see Table 9). All correlations reached a level of significance. Letter knowledge was highly correlated with phoneme isolation and segmentation, rhyme and alliteration awareness. Syllable segmentation showed a weaker correlation.

Table 9

Inter-correlations of the PIPA subtests

PIPA subtest	SS	RA	AA	PI	PS	LK
Syllable Segmentation (SS)		0.416	0.414	0.462	0.327	0.485
Rhyme Awareness (RA)			0.703	0.604	0.539	0.672
Alliteration Awareness (AA)				0.645	0.588	0.728
Phoneme Isolation (PI)					0.556	0.788
Phoneme Segmentation (PS)						0.628
Letter Knowledge (LK)						

The manual provides examples of children with age-appropriate phonological awareness skills, and those who are at risk of literacy difficulties. All show high confidence values. The standardised scores for Letter Knowledge are not available. Nonetheless, the PIPA is a well researched assessment tool of phonological awareness skills.

The test was scored according to the test instructions. Raw scores were converted into standard scores and percentiles according to the instructions in the manual, using the Australian norms.

Procedure

Screening protocol

The researcher, a trained speech-language therapist of 26 years experience, five of which have been gained at Special Education in Otago, administered the tests during an assessment session of approximately 45 minutes duration in the participant's school facility. Each participant's parents/carers were informed and invited to attend the appointment. Participants completed the Bureau Auditory Comprehension Test (Bureau of Maternal and Child Health, 1999), the Renfrew Action Picture Test (Renfrew, 1997), and the New Zealand Articulation Test (Ministry of Education, 2004). The standardised tests were administered according to the instructions of the test. The researcher recorded the participant's responses on each of the assessments.

With the Bureau Auditory Comprehension Test, as stated in the manual, the participants scored 1 point for each verbal instruction completed successfully. For each instruction needing to be repeated, 0.34 for a female and 0.2 for a male was deducted from the final score, according to the instructions in the manual. The points were totaled and compared to the mean score for the participant's chronological age. The participant's performance in terms of standard deviation from the mean can be calculated using 3.32 as 1 SD.

Each participant's responses on the Renfrew Action Picture Test were recorded and scored for information (content) and grammar, as stated in the manual. The information score was added, and compared to the age equivalent for that score. The grammar score was treated similarly.

The single word responses of each participant on the NZAT were recorded. The single sound and blended initial consonant subtests were administered, according to the instructions reported in the manual. Each participant's raw score can be converted to a standard score and percentile, using the tables provided in the manual. In this study, a qualitative analysis, examining the number of phonological processes operating, was used.

The researcher then completed the priority checklist. The score was recorded. The child scored one point for age of 5;00-8;00 years, and 0.5 or 1 point for factors of; the unlikelihood of spontaneous recovery, difficulty accessing the curriculum, parental/teacher anxiety, significant background information (eg hearing problems, family history, medical difficulties, ASD diagnosis). Communication needs scored 1-3 points each in the domains of speech, language, fluency and voice. The points were totaled. A score of 7 points or more allows the child access to comprehensive speech-language therapy service, a score of 4-6 points gives the child a home/school programme for an educator or parent/carer to carry out, and less than 4 points provides advice to the teacher and parent/carer.

Inter-rater reliability of the assessments and priority checklist results was completed on 20% of the sample, which had been video recorded. Three speech-language therapists, who have 14-30+ years experience of working at Special Education in Otago, and who work with school-aged students in Otago utilising the same assessment tools, provided inter-rater reliability ratings. The inter-rater was blinded to the assessment results of the participants, and to the identity of the participants.

Comprehensive assessment

The comprehensive assessment was based on best practice in the assessment of speech and language for young school-aged children (Gillon & Schwarz, 1998). The following test battery was used: the Core Language Score of The Clinical Evaluation of Language Fundamental, 4th edition [CELF-4] (Semel, Secord & Wiig, 2004), The Preschool and Primary Inventory of Phonological Awareness (Dodd, Crosbie, McIntosh, Teitzel & Ozanne, 2000), and a language sample was taken.

Language sample

A structured language sample of 50+ utterances from each participant was recorded on a tape recorder. A high quality recording device, a Digitor AC/DC desktop recorder, model A-4031 was used. The structured oral narrative language sample was elicited using two picture story books, 'I've lost my yellow zebra' and '"Not me" said the monkey'. The examiner shared each story with the child (the child was allowed to choose the order), asking minimal questions and making a few descriptive comments about the pictures to enable the child to understand the story and to encourage the child to make responses. The data took approximately 90 minutes to transcribe and 60 minutes to analyse. Mean Length of Utterance (MLU), grammatical errors and semantic impairments were recorded. MLU was calculated by totaling the number of words in each utterance and dividing by the number of utterances. A grammatical error was noted when a child produced a morphological or syntactic error. A semantic error was identified when a child used a vocabulary item that did not make sense or was inappropriate for the name of an item. MLU and error scores were compared with those of a typical child using Communicate to Participate checklists (Ministry of

Education, 1998), 'Learning to speak and listen' fact sheets (Speech Pathology Australia, 2005) and information in Crystal (1986) Other significant linguistic features were noted, for example, number of initiations, jargon, and echolalia.

The researcher administered the tests in the participant's school facility within four weeks of the screening protocol. Each participant's parents/carers were informed and invited to attend the appointment. The comprehensive assessment took approximately 90 minutes.

After considering the comprehensive evaluations, the researcher reapplied the priority checklist. Each participant's score was compared with the earlier application of the checklist. Any deviation in eligibility was noted.

Comprehensive assessment reports were written by the researcher, providing information about the participants' performance on the comprehensive assessments. These were distributed to the participants' parents and school facility. A comment regarding eligibility for a speech-language therapy service from Special Education was given in these reports.

Inter-rater reliability of 20% of the comprehensive assessments and priority checklists was carried out, with one of the three speech-language therapy colleagues used earlier, who have 14-30+ years experience of working with school-aged students at Special Education in Otago. The therapists were asked to score the comprehensive assessments and priority checklist, blinded to the name and identity of the

participants. Different participants' scores were scored than those rated in the screening protocol.

Data analysis

The researcher carried out two types of analysis of the results: firstly, comparing the outcome of the screening protocol with that of the comprehensive assessments. The number of True Positives, False Positives, True Negatives and False Negatives was calculated. Test Sensitivity and Specificity, and the Positive and Negative Predictive values of the screening protocol were ascertained. Secondly, a qualitative analysis of the false negative results was conducted.

RESULTS

To investigate the effectiveness of the Otago screening protocol in identifying school-aged students with severe speech and language impairments, a comparison was made between the screening protocol and a battery of comprehensive assessment procedures. Specifically, the number of false positives and negatives was identified. In addition, a qualitative analysis was carried out whereby those children who were identified as either a false positive or a false negative were examined.

Test Results

For each participant, the researcher completed a screening protocol and a comprehensive assessment. The results of the screening tools are shown in Table 10.

Table 10

Results of screening measures for each child

Participant	CA	Bureau*	RAPT *		NZAT**
			Info	Gram	
1.	5;05	4;07	5;00	4;05	1
2.	5;04	3;08	4;00	4;00	2
3.	5;04	5;08	4;06	4;06	0
4.	5;03	3;11	5;00	3;06	2
5.	5;02	2;07	<3;06	<3;06	1
6.	5;03	3;08	4;00	3;06	0
7.	5;01	2;07	<3;06	<3;06	1

8.	7;05	8;00	7;00	4;00	1
9.	6;04	8;00	8;00	8;00	1
10.	5;00	4;11	6;06	4;00	4+
11.	6;00	8;00	8;00	7;06	1
12.	5;10	4;08	4;00	3;06	1
13.	5;08	3;03	<3;06	<3;06	0
14.	5;05	8;00	8;05	5;05	2
15.	5;00	2;05	<3;06	<3;06	2
16.	5;01	4;04	3;06	<3;06	3
17.	5;03	4;06	<3;06	<3;06	2
18.	6;10	6;00	6;06	6;00	1
19.	5;02	2;09	<3;06	<3;06	3
20.	5;06	4;00	5;06	4;00	2

*= age equivalent in years, **= number of atypical processes,

Info= Information, Gram= Grammar

The results of the comprehensive assessments are shown in Table 11.

Table 11

Results of comprehensive assessments

Participant	CA	CELF-4*	PIPA*	Language Sample**
1.	5;05	16	5	4.6
2.	5;04	5	9	2.3
3.	5;04	55	16	6.5
4.	5;03	1	9	2.2
5.	5;02	0.1	5	1.9
6.	5;03	1	9	3.4
7.	5;01	0.1	0.1	1.9
8.	7;05	0.1	9	4.7
9.	6;04	96	63	8.6
10.	5;00	8	9	3.4
11.	6;00	5	25	3.6
12.	5;10	3	25	3.2
13.	5;08	0.1	1	2.7
14.	5;05	42	63	7.6
15.	5;00	0.1	5	2.1
16.	5;01	12	63	3.2
17.	5;03	5	9	3.7
18.	6;10	0.1	5	3.2
19.	5;02	0.1	0.1	1.8
20.	5;06	8	5	4.2

*= percentile ranks, **=Mean Length of Utterance

The outcomes of the priority checklists are shown in Table 12.

Table 12

Outcomes of the priority checklists

Participant	CA	M/F	Screen Result	Comp Result	Referrer	Parent Attendance	Consent
1.	5;05	M	negative	positive	GSE	no	yes
2.	5;04	F	positive	positive	parent/sch	no	yes
3.	5;04	M	negative	negative	GSE	yes	yes
4.	5;03	M	positive	positive	GSE	yes	yes
5.	5;02	F	positive	positive	health	no	yes
6.	5;03	M	positive	positive	health	yes	yes
7.	5;01	M	positive	positive	GSE	yes	yes
8.	7;05	F	positive	positive	GSE	no	yes
9.	6;04	M	negative	negative	parent/sch	no	yes
10.	5;00	M	positive	positive	GSE	no	yes
11.	6;00	M	negative	positive	school	no	yes
12.	5;10	F	positive	positive	school	yes	yes
13.	5;08	M	positive	positive	school	no	yes
14.	5;05	M	negative	negative	school	yes	yes
15.	5;00	M	positive	positive	GSE	yes	yes
16.	5;01	M	positive	positive	parent/sch	yes	yes
17.	5;03	M	positive	positive	parent/sch	yes	yes
18.	6;10	M	negative	positive	RTLB	yes	yes
19.	5;02	M	positive	positive	GSE	yes	yes

| 20. | 5;06 | F | positive | positive | parent/sch | yes | yes |

Note: CA= Chronological Age, M= male, F= female, positive= score of 7+ on the priority checklist, negative= score of less than 7 on the priority checklist, GSE=Group Special Education, RTLB= Resource Teacher of Learning & Behaviour

This information can be presented as the positive predictive power (PV+) and negative predictive power (PV-) of the screening protocol by comparing the results of the outcomes of the screening protocol with the results of comprehensive assessment. Table 13 shows these results.

Table 13

The diagnostic accuracy of the screening protocol

Screening Assessment	Comprehensive assessment	
	Positive	Negative
Positive	14	0
Negative	3	3

n = number of participants

True Positives

Findings indicated that fourteen of the twenty participants were true positives. Fourteen participants were positively identified as having severe speech and language impairments through the screening protocol, and this was confirmed by the comprehensive assessments.

True Negatives

Three of the participants were true negatives. That is, three participants were identified as not having severe speech and language impairments in the screening protocol, and the comprehensive assessments supported these findings.

False Negatives

Three of the participants' performances were false negatives. The screening protocol identified three participants as not having severe language impairment, and thus ineligible for service, and the comprehensive assessments identified them as having severe speech and language impairments and thus eligible for a comprehensive service.

False Positives

No participants' performances were false positives. No participants were identified positively by the screening protocol as indicators of severe speech and language impairment that were later identified as ineligible by the comprehensive assessments.

Positive Predictive Value

The positive predictive value of the screening protocol was 100%. There was maximum probability that a student has severe speech-language impairment when there was a positive result of the screening protocol. This was calculated by dividing the number of participants with true positive outcomes on both assessments, with the total number of positive outcomes on the screening protocol (14/14), and converting

the result into a percentage.

Negative Predictive Value

The negative predictive value was 100%.There was maximum probability that a student does not have severe speech-language impairment when there was a negative result of the screening protocol. This was calculated by dividing the number of participants with true negative outcomes on both assessments, with the total number of negative outcomes on the screening protocol (3/3), and converting the results into a percentage.

Two additional measures of validity were calculated; test sensitivity and test specificity.

Test Performance

Test Sensitivity

The probability of the screening protocol result being positive in the presence of severe language impairment was very high at 82% (Kennedy, 2002). This was calculated by dividing the number of participants with true positive outcomes on both assessments with the total number of positive outcomes (14/17) on the comprehensive assessments, and converting the result into a percentage.

Test Specificity

The probability of the screening protocol result being negative in the absence of

severe language impairment was perfect at 100%. This was calculated by dividing the number of participants with false negative outcomes on the comprehensive assessments with the total number of negative outcomes (3/3) on both assessments, and converting the result into a percentage.

Inter-rater Reliability

Inter-rater reliability was administered on 20% of the sample. Comparison of scores on the screening protocol showed high reliability. For the Bureau Auditory Comprehension Test, inter-rater reliability ranged from 90-100%. For the information subtest of the Renfrew Action Picture Test, inter-rater reliability ranged from 91.67-96.88%, and for the grammar subtest from 75-100%. Inter-rater reliability on the New Zealand Articulation Test ranged from 92-100%.

There was high inter-rater reliability on the comparative sample of comprehensive assessments. For the CELF-4, the inter-rater reliability ranged from 90-100%, and for the PIPA inter-rater reliability ranged from 90-100%. The transcription and interpretation of the language sampling showed high inter-rater reliability, with similar quantitative and qualitative comments made on the data.

The Otago screening protocol was a totally valid procedure to use to detect severe speech and language impairments in school-aged students referred to Special Education.

Qualitative Analysis

Three participants showed false negative outcomes. That is, the priority checklist used
at the screening protocol scored them at less than 7, identified as not having severe
speech-language impairment (negative), and therefore not eligible for speech-
language therapy service. When the priority checklist was re-applied after the
comprehensive assessments, these three participants were identified as having severe
speech-language impairment (positive), and deemed eligible for service. Their test
scores are shown in Table 14.

Table 14

Test scores of false negative results

Part't Language Sample#	CA	Bureau*	RAPT * Info	Gram	NZAT**	CELF-4***	PIPA***	
1.	5;05	4;07	5;00	4;05	1	16	5	4.6
11.	6;00	8;00	8;00	7;06	1	5	25	3.6
18.	6;10	6;00	6;06	6;00	1	0.1	5	3.2

Part't= participant, *= age equivalent in years, **= number of phonological processes,
***= percentile rank, #=MLU.

Using the screening tools, the age equivalent scores of these participants were less
than 12 months below their chronological age level, and all three had only one
phonological error process operating. When the comprehensive assessments were
carried out, two participants scored below the 5th centile on the CELF-4 and had
MLUs below 4, and two scored below the 7th centile on the PIPA. For participant 1,
the PIPA and MLU scores identified him as having severe language impairment. For

participant 11, the CELF-4 and MLU scores identified him as having severe language impairment, and for participant 18, the CELF-4, PIPA and MLU scores all identified him as having severe language impairment.

DISCUSSION

Main Findings

This study examined the effectiveness of the Otago screening protocol in identifying school-aged children with severe speech and language impairments. In order to do so, the results of the Otago screening protocol were compared with those of comprehensive language assessment as determined by best practice protocol (Gillon & Schwarz, 1998, Kennedy, 2002). Following the completion of the screening and the comprehensive assessments, an evaluation of the true positives and false positives was calculated.

Feeney & Bernthal (1996) noted that an effective screening tool should have high numbers of true positive outcomes, and low numbers of false positive and false negative outcomes. The Otago district screening protocol fulfils the measures of effectiveness of a screening assessment protocol. Analysis of the results revealed that positive and negative predictive values are very high. All children who were deemed to be at risk of language impairment on the screening protocol were confirmed to be severely language impaired on the comprehensive assessment. Three of the twenty children assessed, however, were deemed to be eligible for service following the comprehensive assessment but would not have been eligible had only the results of the screening protocol been considered.

In the Otago district of Special Education, the likelihood of the screening protocol correctly identifying school-aged children with speech and language impairment is high, however a reduction in false negative results would be desirable. A discussion of the results relative to specific aspects of the screening protocol follows.

Screening

Nature of the Tasks in a Screening Protocol

Best practice (Gillon & Schwarz, 1998, Kennedy, 2002) recommends that a screening protocol should incorporate assessment in a wide range of domains across receptive and expressive language, and phonology, a single word task being enough to provide a representative sample of a child's phonology (Masterton, Bernhardt, & Hofheinz, 2005).

Gardner et al (2006) argue that screening in the domains of grammar and phonology is sufficient to identify children with possible language impairment. The GAPS test (Gardner et al, 2006) was designed to identify specific language impairment and/or specific literacy impairment in preschool and young school-aged children, through sampling key grammatical and phonological skills. Two subtests of sentence imitation and non-word repetition are used. Gardner et al (2006) cite extensive research claiming that difficulties in syntax and phonological manipulation are key features in children with specific language impairment (Bishop, 1999, Conti-Ramsden, Botting & Faragher, 2001, Conti-Ramsden & Hesketh, 2003, van der Lely, Rosen & Adlard, 2004, van der Lely, 2005).

The current assessment tools used in the screening protocol at the Otago district of Special Education do sample grammar and phonology, as well as a number of other domains. This study finds that, many children are accurately identified as language impaired given the screening test results, but, of the three children with false negative outcomes, none would have been identified as having language impairment through sampling only the syntactic and phonological domains. Whether young school-aged

children with severe language impairment, the target population for Special Education, could be reliably identified using the GAPS test would require further investigation. School children aged above 6;06 years would require an alternative assessment measure such as the CELF-4 screening test (Semel, Wiig, & Secord, 2004).

Many researchers accept that phonological awareness skills are essential prerequisites to a child's literacy development (Beitchman, Nair, Clegg & Patel, 1986, Catts, Fey, Tomblin & Zhang, 2002, Gillon & Schwarz, 2001, Law, Garrett & Nye, 2003, Nathan, Goulandris, Stackhouse & Wells, 2004), and should be incorporated in all screening protocols. In the Otago protocol, an assessment of phonological awareness is not consistently used within the screening protocol, and, in this study, two of the three false negative outcomes were identified through the completion of a phonological awareness assessment tool, the PIPA (Dodd et al, 2000). If this tool had been used in screening, the screening protocol would have been more efficient in identifying children with speech-language impairment.

The Otago screening protocol takes approximately 45-60 minutes to administer. The GAPS test (Gardner et al, 2006) suggests that effective screening could be administered in 10 minutes, and that the screening could be carried out by non-speech-language therapists. Support of this finding requires replication of Gardner et al's study in New Zealand.

Screening Tools

Several shortcomings in the tools used in the screening protocol are highlighted with

reference to best practice (Gillon & Schwarz, 1998). The Bureau Auditory Comprehension Test (Bureau of Maternal and Child Health, 1999) has weak psychometric properties, and is often considered inadequate in its assessment of severe receptive language difficulties. In particular it uses a restricted range of vocabulary and concepts, and has uninteresting material for school-aged children. Nevertheless, in this study, the presence or absence of receptive language difficulties was accurately identified in 85% of children using the Bureau. However, of the three children showing false negative outcomes, none were identified as having language impairment from the results of the Bureau test.

The Renfrew Action Picture Test (Renfrew, 1997) too has little reliability and validity data, and uses a narrow vocabulary range. The RAPT is often criticised for its inability to identify severe semantic deficits in school-aged children, and of using dated visual material. Once again, in this study, the presence or absence of expressive language impairment was accurately identified in 85% of children, using the RAPT. Of the children showing false negative outcomes, none were identified as having language impairment from their performance on the RAPT.

Whilst the New Zealand Articulation Test (Ministry of Education, 2004) is a robust screening tool, having extensive reliability and validity data relevant to the New Zealand context, it fulfils an adequate assessment purpose at a single word level only. The NZAT does have material that could be used to examine phonological representations at a sentence level through a picture description task, but recommends that this should be used in comprehensive assessment. In the Otago screening protocol, there is no mandatory examination of the child's phonological performance

at a functional level, as recommended by best practice (Gillon & Schwarz, 1998). Within this study, three children were positively identified as having phonological impairment by the NZAT in the screening protocol, and were eligible for speech-language therapy service.

Information from all Sources

Research recommends that for a screening protocol to be effective, information must be gathered from all those who are knowledgeable about the child, such as parents and teachers (Gillon & Schwarz, 1998, Kennedy, 2002). Parental report may be considered the primary source of information (Stott, Merricks, Bolton & Goodyear, 2002). In this study, 60% of parents/carers attended the screening appointment, enabling positive relationships to be developed, and information to be conveyed between the parent/carer and speech-language therapist. In the remaining 40% of cases, the parent/carer did not attend the screening appointment. The parents/carers were contacted by the researcher in most cases, but communication could be limited to the parent/carer receiving reports from the researcher. Perhaps significantly, of the three children who showed false negative outcomes, the parent was present in only one case. In the remaining two cases, the researcher attempted but could not contact the parent/carer. Further research is recommended to consider strategies to improve parental contribution to the screening protocol.

Priority checklist

An examination of the priority checklist currently used in the Otago district of Special Education is made in this study. In response to a difficulty in identifying an

appropriate process to access speech-language therapy services at Special Education (Gillon & Schwarz, 1998), a priority checklist was devised by Special Education speech-language therapy managers in the late 1990s, and intended for national use by speech-language therapists working with school-aged children at Special Education.

The results of this study reveal that the priority checklist is suitable for its purpose, despite its apparent subjectivity. The Otago district continues to use an eligibility figure of seven on the priority checklist, as recommended by the SE2000 initiative (Ministry of Education, 1997). In contrast, some districts, as a response to high referral numbers and staffing difficulties, have decided to use an eligibility figure of eight or nine on the priority checklist. Such variations influence the number and type of speech-language difficulties accepted within a district. A closer examination of the consistency of the screening protocol across districts may reveal more widespread variations. National consistency in the application of the priority checklist is called for.

The priority checklist is a misnamed document. It should be re-named as an eligibility checklist, since it determines eligibility for speech-language therapy service and not priority. At present, the needs of a child with a score of ten on the checklist could be construed as a higher priority than a child who scores seven. In fact, both children would be eligible and, to determine priority, further assessment required. The checklist is simply to gain eligibility to a comprehensive speech-language therapy service. In this study, a score of seven on the priority checklist is used, and the results support this as a valid eligibility figure. However, both the title and use of the priority checklist remain confusing.

False negative outcomes

Feeney & Bernthal (1996) recommend that to judge the effectiveness of a screening assessment tool, there should be findings of a high accuracy of correct identification, true positives, and a low number of false positives and false negatives. This study finds the screening protocol has a high accuracy rate of 85%, no false positives, and a low percentage, 15%, of false negatives.

The results of this study show a clear contrast in the performances of children with positive and false negative outcomes on the screening protocol to the performances of children with negative outcomes. Using the Core Language Score of the CELF-4, total percentile ranks of less than 0.1 to 16 on the four subtests could be used as indicators of severe language impairment. The critical subtest contributing to their low score appears to be Concepts and Following Directions. Further research with greater numbers of participants is recommended to establish support of these findings.

The PIPA manual claims that percentile ranks below 16 are significant in identifying severe phonological awareness difficulty. In this study, the performances of those children who have false negative outcomes indicate that rhyme awareness, alliteration awareness and phoneme segmentation may be important subtests of the PIPA to indicate severe speech-language impairment. Within the group of children showing true positive outcomes, the performance of each child on the rhyme awareness task seems to be an important indicator of severe speech-language impairment. Overall, this study supports research that claims rhyme awareness is a critical task of phonological awareness indicating severe language impairment (Gillon & Schwarz, 2001), and the consequential implications of difficulties in this domain on academic

progress (Catts, 1993, Gillon & Schwarz, 2001, Hesketh, 2004, Logemann & Baum, 1998, Shriberg, Tomblin & McSweeney, 1999). The results are inconclusive with respect to the recognition of phoneme identification as a key phonological awareness task (Gillon & Schwarz, 2001).

Using the language samples taken in this study, there are a number of linguistic features that clearly indicate the possibility of severe speech-language impairment. The language samples of children with false negative outcomes show linguistic features of several grammatical errors, and few conversation initiations in their language samples. Children with true positive outcomes show these features in addition to the syntactic difficulty of few complex sentence constructions, and semantic difficulties, of restricted vocabulary or possible word finding difficulty. Key linguistic features to indicate severe language impairment seem to be a 2-4 MLU, syntactic difficulties of several grammatical errors and few complex sentences, semantic difficulties of restricted vocabulary or possible word finding difficulty, and pragmatic difficulties of few conversation initiations.

The limitations of the conclusions that can be drawn from the language sampling procedure could be attributed to the lack of using a standardised tool, such as the New Zealand version of the Systematic Analysis of Language Transcripts (Miller, 1992). The use of a standardised language sampling tool may provide additional useful information. The validity of a child's linguistic performance could be measured, and the child's performance compared to that of their peers.

Participants

Twenty children took part in this study. To provide more reliable evidence an increased cohort of at least 100 children should be used (McCauley & Swisher, 1984). A limitation of this study in terms of sample size is acknowledged. Therefore, a recommendation that future research should consider a greater number of participants is made.

Eighteen of the twenty children in this study are English-speakers of European ethnicity. Whilst the results of this study can claim validity with respect to English speakers no such claim can be made for speakers of other languages and, indeed, other ethnicities. This is an issue, in New Zealand where English and Te Reo are official languages, and for consideration in those Special Education districts of New Zealand where there are a number of speakers of languages other than English. Future research should explore the validity of the screening protocol with children of other ethnicities and languages. In the New Zealand context, the effectiveness of the screening protocol for children of Maori ethnicity, and Te Reo speakers should be examined. Additionally, the screening protocol should be validated for other relevant cultures in New Zealand, such as Asian ethnicities, and where English is not the primary language of the children.

The sample in this study has a gender ratio of 3:1 males to females, reflecting a typical speech-language therapist's caseload. The mean age of participants is 5;07 years with a spread from 5;00 to 7;05. The median age is 5;04 years. The contract held by Special Education with the Ministry of Education is to target school-aged children in the 5;00-8;00year age band. This sample fits the target age criteria of the

contract.

Additional factors impacting on screening

Several additional factors may impact on the speech-language therapist's decision-making, including the screening process. There is no data available to indicate the prevalence and incidence of speech-language impairment in New Zealand. Currently, international data and population-based funding informs staffing numbers. Caseload size and management, the changing scope of practice of speech-language therapy, the efficacy of speech-language therapy, and the appropriateness and increasing numbers of referrals all add professional pressures to the role and workload of the speech-language therapist which, in turn, could exert influence on the screening process.

Prevalence and incidence of speech-language impairments

Population-based funding of the predicted number of school-aged children with severe speech-language impairments determines the contractual numbers for Special Education services. The formula for funding is based, in part, on international prevalence rates of speech-language impairment.

There are wide variations in the estimates of prevalence of speech-language impairment in international research. Approximate international prevalence figures range from 5-16% (Broomfield & Dodd, 2004, Gillon & Schwarz, 1998, Law et al, 2000). Assuming a similar prevalence rate, Gillon & Schwarz (1998) estimated that there were about 70,000 pre-school and school-aged children in New Zealand requiring speech-language therapy services. Gillon & Schwarz (1998) claimed that this figure was far higher than the current speech-language therapy staffing levels

could manage.

The long term impact of speech-language impairment on academic and social outcomes is far reaching and well documented (Catts, Fey, Tomblin & Zhang, 2002, Gillon & Schwarz, 2001, Goulandris & Snowling, 2004, Law, Garrett & Nye, 2003, Nathan, Stackhouse, Beitchman, Nair, Clegg & Patel, 1986). Despite the extensive data on the prevalence for speech and language impairments within school-aged populations internationally, there is a marked lack of such data in New Zealand. Currently, translating overseas data must be relied upon, but should be done with caution. Demographic and political factors, diagnostic criteria, workplace practices and service delivery options available can all be discrepant, so the validity of using international figures is called into question. However, these international prevalence and incidence figures are frequently used to advise service delivery and set target numbers in many countries.

In New Zealand, the contract for speech-language therapy services between the Ministry of Education and Special Education is determined, in part, on international prevalence data and a population-based funding formula. This formula is based on the total population of a district, irrespective of density and age factors of the population. Since these numbers are translated from international data, and the unique population and demographics of New Zealand are not considered, inappropriate contractual numbers for the speech-language therapy service in Special Education may result. Low contractual numbers adds considerable pressure to the screening protocol used to access speech-language therapy services at Special Education.

Speech-language therapy staffing

In 1997, Special Education Services (now Special Education) launched the SE2000 initiative (Ministry of Education, 1997). At that time, Specialist Education Services employed 134 FTE speech-language therapists to work with pre-school and school-aged children (Gillon & Schwarz, 1998). Gillon & Schwarz (1998) calculated that this staffing level was grossly inadequate to meet the needs of the target population, pre-school children with moderate or severe speech-language impairment, and five to eight year olds with severe speech-language impairment, and called for an increase in speech-language therapists.

The SE2000 initiative did result in a change in government policy to double the funding to service children with severe speech-language impairment, with the intention that a more intensive or comprehensive service would be provided. In 1997, 12,000 children were receiving input from speech-language therapy, within Special Education Services, 71% of those aged 5-8years. The primary focus of the SE 2000 initiative was to be on developing communication skills within the early childhood sector. In particular, communication difficulties in the domains of language, phonology, fluency and voice were to be targeted.

Children with moderate special educational needs, including moderate speech-language impairments, were not eligible for Special Education Services and expected to have their needs met through programmes of professional development to teachers, and use of the school's Special Education Grant (SEG). Each school was expected to decide on the most appropriate use of the SEG to support the children in their facility. A review of the Special Education Grant (ERO, 2004) reported that, in general, SEG

was used to fund a narrow range of support programmes, with most schools opting to fund literacy programmes, despite the individual needs of the children on their roll. Under this funding regime, it seems unlikely that the needs of all children with moderate speech-language impairments are being met through the use of SEG, and therefore speech-language therapists at Special Education may feel under pressure to provide services to those children by admitting them in the screening protocol.

If speech-language therapists employed by Special Education were expected to provide comprehensive services to school-aged children with moderate speech-language impairments, substantially greater staffing numbers would be required. The number of school-aged children with moderate speech-language impairments may be two-and-a-half times that of those with severe speech-language impairments, within the American population (ASHA, 2000). Given expected similar incidence and distribution trends in New Zealand, then approximately a further 17,000 school-aged children could have moderate speech-language impairments.

Whilst some studies report international figures regarding staffing ratios (Lindsay, Soloff, Law, Band, Peacey, Gascoigne & Radford, 2003, Van der Gaag, McLoone & Reid, 1999), there are no recommended guidelines of the appropriate, and most effective, staffing levels. If the number of speech-language therapists employed by Special Education continues to remain too low to cope with the demand, a therapist's professional decision-making may be compromised in judging a child's eligibility for service. In some districts of Special Education, tighter criteria or higher eligibility figures are applied in an attempt to cope with high referral rates, and staffing difficulties. Speech-language therapists working in rural areas may feel particularly

overwhelmed with demand, since there may be no alternative service providers.

Caseload size and management

In 1993, ASHA recommended a maximum caseload size of 40, for full-time speech-language therapists working in schools, with reduced numbers, of 25, if pre-school and/or complex cases are involved. Law et al (2000) supported the use of 40 as a benchmark figure for experienced speech-language therapists as having demonstrable validity. The Royal College of Speech Language Therapists (1996) recommended that a prescriptive caseload number is undesirable. Instead, managers should select and apply the most suitable methods of caseload analysis to best meet the needs of their communities (RCSLT, 1996). Despite the expansion of speech-language therapists' responsibilities and workload, there seems to have been no progress made towards these targets a decade later, with most therapists operating much larger caseloads (Ministry of Education, in progress).

The impact of high caseload numbers is critical in reducing the service delivery options available, and the effectiveness of therapy. Under caseload pressure, therapists tend to operate tighter eligibility criteria, engage in more withdrawal group work, with groups of increasing size, despite the individual needs of the children. They engage in less collaboration and professional development. With high caseloads, speech-language therapy service provision is likely to become poorer in quality and ineffective, leading to worse outcomes (Cirrin et al, 2003, Russ, Chiang, Rylance & Bongers, 2001, Schooling, 2003). Education staff, too, feel frustrated not only by the lack of support available for those children with moderate speech-language impairments (Ukrainetz & Fresquez, 2003), but also by the perceived insufficient

support available to them for the children who are (Sadler, 2005).

International literature often stresses the importance of a balanced caseload, in terms of severity of need, for the well-being of the speech-language therapist (ASHA, 1993). At present in New Zealand, speech-language therapists employed by Special Education work exclusively with school-aged children who have severe speech-language impairments. Positive therapy outcomes are much more difficult to achieve and take considerably longer, under these circumstances, resulting in a slower throughput of cases and a bottleneck of children on a speech-language therapist's caseload. This slower throughput adds pressure to the speech-language therapist's eligibility decision-making in the screening protocol.

In addition, a significant shift towards a collaborative service delivery model in inclusive settings is recorded empirically (ASHA, 1991, 1992, Elksnin & Capilouto, 1994, Peters-Johnson, 1996, Throneberg et al, 2000). Collaboration, together with the burden of large caseloads, has led to several alternative management approaches being sought. These are numerously described in the literature. Alyson Portch (2003) outlines a Care Aims model, Blosser & Kratcowski (1997) the PACs [Providers, Activities, Contexts] process, and Chandler & Pickering (2004) the MACS [Maroondah Approach to Clinical Services] system. Some are medically-based, and are problematic to apply to the education sector, whilst others rely on training parents/carers or other professionals to take responsibility for implementing the speech-language therapy programme (Law, 2004, Ruscello, Cartwright, Haines & Shuster, 1993). This high level of collaboration, to educational colleagues and parents/carers, places greater demands on the speech-language therapist working in

the school, especially when the time and systems supports may not be in place. Tighter eligibility criteria are likely to be imposed on the screening protocol if time is not available to provide high levels of collaboration for children on the therapist's caseload.

Changes in scope of practice

The expanding scope of practice in the speech-language therapy profession now includes areas such as swallowing disorders, literacy difficulties, augmentative and alternative communication, and with children showing emotional and behavioural disorders. Furthermore, the scope of practice is continually evolving. This is reflected in Gillon & Schwarz (2001) updated review of best practice which encompasses these emerging therapy domains. It is arguable whether the contract between the Ministry of Education and Special Education for speech-language therapy services has similarly changed to take account of these wider domains. Such a change would require a major review of the screening protocol currently used in Special Education to determine eligibility for speech-language therapy service.

There are greater expectations of the speech-language therapist's role, including evidence-based practice, measuring outcomes, training support personnel, providing in-service training to parents/carers, whanau and teaching staff, and responding to diverse cultural needs. Additionally, they are expected to be service co-coordinators, engage in research and offer supervision to greater numbers of graduate speech-language therapy students, assistants and teaching staff. The speech-language therapist is encouraged to offer a full range of service delivery options in a variety of environments through integrated service provision, and collaboration (Bankson,

Diefendorf, Elman, Forsythe, et al, 2002, Wright, 1992). These increased demands on time may influence the therapist's judgment in determining those eligible for a comprehensive speech-language therapy service.

Efficacy of Speech-Language Therapy

Given funding constraints, speech-language therapy services should be targeted to where they can be most effective, and thus inform eligibility criteria. Finding evidence of efficacy in speech-language therapy is problematical. Some sort of evaluation of effect is required, and the use of outcome measurements to judge the value of speech-language therapy is problematical. The variability and influence of diverse referral policies, base populations, and the skill and work practice of differing speech-language therapists makes comparison across impairment groups very difficult (Enderby and John, 1999). Significant issues affecting efficacy are highlighted including intensity and rate of treatment regimes (Gillon and Schwartz, 2001, Schooling, 2003), programme content (McCartney et al, 2004) and support of parents/carers and teachers, the age of children at referral, diagnosis (Goorhuis-Brouwer & Knijff, 2002, Law, 2004) and severity of need (Lindsay, Soloff, Law, Band, Peacey, Gascoigne & Radford, 2003).

Robey (2004) challenges the definitions of efficacy and effectiveness given in many studies, and recommends a five stage paradigm to provide more accurate outcomes in the future. This system may help in the search for empirical data to guide service provision, identify the target populations with which speech-language therapists can be most effective, and inform eligibility criteria.

Law, Garrett & Nye (2003) provide some limited evidence of effectiveness, through meta-analysis, across a range of disorder groups. They find that there is more support for speech-language therapy to address difficulties with expressive phonology (Gierut, 1998) and vocabulary, than syntactic deficits. There is disputed evidence of the efficacy of speech-language therapy in treating receptive language difficulties (Law, 2004, Whitehurst & Fischel, 1994). For other diagnostic groups, such as stuttering and voice disorder, there is little evidence to support speech-language therapy intervention (Law, 2004).

Whitmire (2002) calls for more comparative research into differing service delivery models, and with different diagnostic groups. Robust efficacy data would enable more accurate speech-language therapy service delivery decisions to be made, influencing the types of referrals that would be eligible for comprehensive service from Special Education.

Increasing number of referrals

The number of referrals of pre-school and school-aged children to the speech-language therapy service has increased significantly in recent years. Enderby & Petheram (1998) describe an almost three fold increase across a ten year period from mid- 80s to mid-90s in health care districts in the UK. Law et al (2000) report similar findings.

Certain diagnostic groups have increased more than most. For example, the apparent increase in diagnosis of Autistic Spectrum Disorder may more accurately reflect changes to the diagnostic criteria used (Doury and Nash (2003). Consequently,

'milder' forms of autism such as Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) are more often identified, and constitute the highest diagnostic sub-group (Doury & Nash, 2003). Whilst the screening protocol currently used in Special Education may detect severe speech-language impairment, the results are inconclusive with regard to the specific needs of the autistic population.

Appropriateness of referral

Since there is sound evidence of increasing numbers of referrals, Special Education faces considerable pressure to ensure that appropriate referrals are made. Some researchers have considered the appropriateness of referral, with false positive figures ranging from 10-33% (ASHA, 2000, Broomfield and Dodd, 2004, Glascoe, 2001).

In 2004-5, of the referrals of school-aged children made to the Special Education speech-language therapy service in Otago, up to one third were ineligible to receive a comprehensive service. Many of these ineligible children were given a home or school-based programme for parents/carers or teaching staff to support. In the current study, 15% of referrals were ineligible for comprehensive speech-language therapy service. However, a small sample size was used.

The cost of ineligible referrals could be questioned. Washington and Craig (2004) reported that, whilst recognising the expense of screening protocols to service providers, children who pass screening procedures differ from those who do not. The use of a screening protocol to identify speech-language impairment may be beneficial not only to focus intervention, but to provide advice and facilitate referral to other services. Washington and Craig (2004) concluded that screening protocols should be

considered more valuable as a 'screening out' tool than in accurately identifying speech-language impairments.

An article by Sunderland (2004) passed the responsibility for accurate referral in the direction of school staff, by providing information for teachers to recognise the characteristics of speech-language impairment. If information about speech-language impairment, and the eligibility criteria for speech-language therapy service from Special Education, were available and consulted in all educational facilities, the percentage of eligible referrals made may increase.

CONCLUSIONS

The screening protocol used for referrals of school-aged children in the Otago district of Special Education is valid for the purposes of identifying speech-language impairment, and determining eligibility for speech-language therapy service. Implications of the findings of this study are limited by the small sample size, and demographic factors of the participants. The study validates the Bureau Auditory Comprehension Test, the Renfrew Action Picture Test, and the New Zealand Articulation Test as valid tools to use in the screening protocol. The priority checklist fulfills the purpose of a screening protocol.

Despite being valid, the Otago screening protocol failed to identify language impairment in 15% of cases and took 45 minutes to administer so was not a particularly efficient tool given the limited resources available. To increase the effectiveness of the screening protocol, a measure of phonological awareness skills

should be added, and alternative standardised screening assessment tools considered, such as the GAPS test (Gardner et al, 2006) for younger school-aged children and the CELF-4 screening test (Semel, Secord & Wiig, 2004) for older school-aged children.

The number of school-aged children who receive a speech-language therapy service is based on translating international incidence figures to the New Zealand context. This could be problematical. The needs of school-aged children with moderate speech-language impairments may not be being met under the current system. In addition, several factors may influence the speech-language therapist's judgment of a child's eligibility for a comprehensive speech-language therapy service. These include staffing levels, caseload size, workload, changes to the scope of practice of the speech-language therapy profession, increasing numbers of referrals and efficacy of speech-language therapy. The results of this study suggest that the source of over-delivery of comprehensive speech-language therapy services to school-aged students in Otago is not the screening protocol.

Implications for Future Research

Future research is recommended to replicate the findings of this study with a greater population sample, and with diverse population groups, such as children of Maori ethnicity and Te Reo speakers. There is an urgent need to find the incidence of speech-language impairment within the pre-school and school-aged populations in New Zealand, and to examine the services available for school-aged students with moderate speech-language impairments. To increase the effectiveness of the Otago screening protocol, a standardised test of phonological awareness skills should be added to the screening measures, and consideration given to the use of the GAPS test

(Gardner et al, 2006) and the CELF-4 screening test (Semel, Secord & Wiig, 2004) as alternative screening tools. In addition, the priority checklist requires re-naming as an eligibility checklist, and consistent eligibility criteria applied in all Special Education districts. Parents/carers and schools need clear written information about the speech-language therapy service at Special Education for school-aged students, and parental/carer and teacher consultation should be evident in all screening protocols. The development of a pre-referral checklist would be useful not only to inform referring agents about the nature of speech-language impairment, but also to advise the speech-language therapist of the nature of concern regarding a student's speech-language difficulties. Finally, future research is recommended to explore alternative sources of the over-delivery issue facing the speech-language therapy service for school-aged students in Otago.

APPENDIX A   Priority Checklist

Name _____ Date of Birth _____

No.	Question	No		Yes	
1.	Is the communication difficulty unlikely to resolve itself within six months?	0		1	
2.	Is the difficulty causing an inability to communicate effectively?	0		.5	1
3.	Are there significant circumstances that influence priority ratings? (Comments, if any) • Medical • ESOL • Hearing Impaired • Asperger's Syndrome / ASD • Parental Support • Significant Family History	0		.5	1
4.	Does the communication difficulty affect the child's ability to access the curriculum? .5 = Emerging Skills 1.0 = Adaptation to curriculum necessary	0		.5	1
5.	Is the client's difficulty causing concern for the caregiver of teacher?	0		.5	1
6.	Is the difficulty causing a reluctance to speak?	0		.5	1
7.	Is the child aged between 5 - 8 years?	0		1	
8.	Degree of severity (choose 0,2,2) for each of the following				
	Phonology / Articulation:	0	1	2	3
	Language Impairment, incl. Phonological Awareness	0	1	2	3
	Fluency: do not pick up if no family support	0	1	2	3
	Voice:	0	1	2	3

Phonology Articulation:

1. Clusters / k;g / mild disorder	2. Developmental speech sound problems (hard to understand	3. Significant disorder / Unintelligible

Language Impairment:

1. Mild Delay Phonological awareness alone (mild) High Level language Difficulties	2. Moderate delay Language disorder affecting work	3. Significant audit processing Extremely limited functional Expressive language

Fluency:

1. Intermittent and specific to certain circumstances Advice to parents	2. Generalised: pick up	3. Reluctance, anger, frustration

Voice:

1..Husky – mild / behaviour No medical	2. N/A	3. Medical

Total Score: _____   Date: _____

APPENDIX B

Inter-rater reliability

Screening protocol assessments; raw scores

Participant	Bureau Auditory Comprehension Test		RAPT				NZAT *	
	rater 1	rater 2	rater 1 info	rater 2 info	rater1 gram	rater 2 gram	rater 1	rater 2
1.	27	26	31	32	19	18	6	6
2.	9	10	14	15	4	3	23	25
3.	12	12	17	16	9	10	6	6
4.	11	10	11	12	3	3	29	30

Note: *= number of errors

Comprehensive assessments

CELF-4; raw scores

Participant	C&FD		WS		RS		FS	
	rater 1	rater 2	rater 1	rater 2	rater 1	rater 2	rater 1	rater 2
1.	1	0	8	10	11	13	1	0
2.	2	2	1	1	3	3	0	0
3.	1	1	12	11	9	10	1	1
4.	47	47	28	26	50	48	41	39

Note: C&FD= Concepts and Following Directions, WS= Word Structure, RS=

Recalling Sentences, FS= Formulated Sentences

CELF-4; scaled scores

Participant	C&FD		WS		RS		FS	
	rater 1	rater 2	rater 1	rater 2	rater 1	rater 2	rater 1	rater 2
1.	1	1	4	5	3	4	3	2
2.	4	4	3	3	1	1	4	4
3.	3	3	8	7	4	4	6	6
4.	16	16	15	13	12	11	17	17

Note: C&FD= Concepts and Following Directions, WS= Word Structure, RS=

Recalling Sentences, FS= Formulated Sentences

CELF-4; percentile rank

Participant	C&FD		WS		RS		FS	
	rater 1	rater 2	rater 1	rater 2	rater 1	rater 2	rater 1	rater 2
1.	0.1	0.1	2	5	1	2	1	0.4
2.	2	2	1	1	0.1	0.1	2	2
3.	1	1	25	16	2	2	9	9
4.	98	98	95	84	75	63	99	99

Note: C&FD= Concepts and Following Directions, WS= Word Structure, RS=

Recalling Sentences, FS= Formulated Sentences

CELF-4, comparison of total scaled scores

Participant	rater 1	rater 2
1.	11	12
2.	12	12
3.	18	20
4.	59	57

PIPA; raw scores

Participant	SS		RA		AA		PI		PS		LK	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
1.	3	3	3	3	0	0	0	0	0	0	0	0
2.	0	0	0	0	0	0	0	0	0	0	0	0
3.	9	10	1	1	1	1	0	0	1	1	1	1
4.	11	11	10	10	10	10	12	11	6	6	27	27

Note: SS= Syllable Segmentation, RA= Rhyme Awareness, AA= Alliteration

Awareness, PI= Phoneme Isolation, PS=Phoneme Segmentation, LK= Letter

Knowledge, R1=Rater 1, R2=Rater 2.

PIPA; standard scores

Participant	SS		RA		AA		PI		PS		LK	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
1.	3	3	3	3	3	3	3	3	7	7	-	-
2.	3	3	3	3	6	6	6	6	8	8	-	-
3.	12	12	3	3	7	7	6	6	11	11	-	-
4.	12	12	10	10	9	9	12	12	11	11	-	-

Note: SS= Syllable Segmentation, RA= Rhyme Awareness, AA= Alliteration

Awareness, PI= Phoneme Isolation, PS=Phoneme Segmentation, LK= Letter

Knowledge, R1=Rater 1, R2=Rater 2.

PIPA; percentile ranks

Participant	SS		RA		AA		PI		PS		LK	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
1.	1	1	1	1	1	1	1	1	16	16	-	-
2.	1	1	1	1	9	9	9	9	25	25	-	-
3.	75	75	1	1	16	16	9	9	63	63	-	-
4.	75	75	50	50	37	37	75	75	63	63	-	-

Note: SS= Syllable Segmentation, RA= Rhyme Awareness, AA= Alliteration

Awareness, PI= Phoneme Isolation, PS=Phoneme Segmentation, LK= Letter

Knowledge, R1=Rater 1, R2=Rater 2.

References

Adams, C., & Lloyd, J. (2005). Elicited and spontaneous communicative functions and stability of conversational measures with children who have pragmatic language impairments. *International Journal of Language and Communication Disorders, 40, 3, 333-347.*

Adams, C., Lloyd J., Aldred C., & Baxendale, J. (2006). Exploring the effects of communication intervention for developmental pragmatic language impairments: a signal generation study. *International Journal of Language and Communication Disorders, 41, 1, 41-65.*

American Speech-Language Hearing Association (1991). A model for collaborative service delivery for students with language-learning disorders in the public schools. *ASHA, 33., supplement 5, 44-50.*

American Speech-Language Hearing Association (1992). Why now? Changing school speech-language service delivery. *ASHA, 34, 11, 40-41.*

American Speech-Language Hearing Association (1993). Guidelines for caseload size and speech-language pathology service delivery in the school. *ASHA, 35, supplement 10, 33-39.*

American Speech-Language Hearing Association (2000). *2000 Schools Survey.* Rockville, MD: ASHA.

Bamford J., Davis, A., Boyle, J., Law, J., Chapman, S., Stewart Brown, S., & Sheldon, S. (1998). Preschool hearing, speech, language, and vision screening. *Quality in Health Care, 7, 240-247.*

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). *Prevalence of disorders of the autism spectrum in a population cohort of children in south Thames: the special needs and autism project (SNAP).* The Lancet, 368, 9531, 210-215.

Bankson N., Diefendorf, A., Elman, R., Forsythe S., et al. (2002). Scope of practice in speech-language pathology. *Communication Disorders Quarterly, 23, 2, 77-84.*

Beitchman, J. H., Nair, R., Clegg, M., & Patel, P. G. (1986). Prevalence of speech and language disorders in 5-year old kindergarten children in the Ottawa-Carleton region. *Journal of Speech and Hearing Disorders, 51, 95-110.*

Bishop, D. V. M. (1997). *Uncommon Understanding: Comprehension in Specific Language Impairment.* Hove: Psychology Press.

Bishop, D. V. M. (2003). *The Children's Communication Checklist-2.* London, UK: Psychological Corporation.

Bishop, D. V. M., & Leonard, L. B. (Eds.) (2000). *Speech and language impairments in children: causes, characteristics, intervention and outcome*. Hove, UK: Psychology Press.

Blood, G. W., Ridemour, V. J., Qualls, C. D., & Hammer, C. S. (2003). Co-occurring disorders in children who stutter. *Journal of Communication Disorders, 36, 427-448.*

Bloom, L., & Lahey, M. (1978). *Language Development and Language Disorders.* New York, John Wiley & Sons.

Blosser, J. L., & Kratcoski, A. (1997). PACs: A framework for determining appropriate service delivery options. *Language, Speech and Hearing Services in Schools, 28, 2, 99-108.*

Bothe, A. K., Davidow, J. H., Bramlett, R. E., Frank, D. M., & Ingham, R. J. (2006). Stuttering treatment research 1970-2005: II systematic review incorporating trial quality assessment of pharmacological approaches, *American Journal of Speech Language Pathology, November.*

Bothe, A. K., Davidow, J. H., Bramlett, R. E., Frank, D. M., & Ingham, R. J. (2006). Stuttering treatment research 1970-2005: I systematic review incorporating trial quality assessment of behavior, cognitive and related approaches, *American Journal of Speech Language Pathology, November.*

Brackenbury, T., & Pye, C. (2005). Semantic deficits in children with language impairments: issues for clinical assessment. *Language, Speech and Hearing Services in Schools, 36, 1, 5-16.*

Broomfield, J., & Dodd, B. (2004). Children with speech and language disability: caseload characteristics. *International Journal of Language and Communication Disorders, 39 (3), 303-324.*

Broomfield, J., & Dodd, B. (2004). The nature of referred subtypes of primary speech disability. *Child Language Teaching and Therapy, 20 (2), 135-151.*

Brown, C. (1988). An investigation to compare responses of 'normal' children and those with learning difficulties, to a revised version of the Renfrew Action Picture Test. *Unpublished report of a project for BSc (Hons).*

Bureau of Maternal and Child Health (1999). *Bureau Auditory Comprehension Test.* NSW, Australia: Bureau of Maternal and Child Health.

Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *Journal of Speech and Hearing Research, 36, 948-958.*

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language and Hearing Research, 45, 6, 1142-1157.*

Chandler, R., & Pickering, C. (2004). Making the most of what you have. *Royal College Of Speech and Language Therapy, Bulletin, February, 10-11.*

Cirrin, F., Bird, A., Biehl, L., Disney, S., Estomin, E., Rudebusch, J., Schraeder, T., & Whitmire, K. (2003). Speech-language caseloads in the schools: a workload analysis approach to setting caseload standards. *Seminars in Speech and Language, 24, 3, 155-180.*

Condouris, K., Meyer, E., & Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measures of spontaneous speech in children with autism. *American Journal of Speech-Language Pathology, 12, 3, 349-358.*

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: longitudinal considerations. *Journal of Speech, Language and Hearing Research, 42, 1195-1204.*

Conti-Ramsden, G., Botting, N., & Farragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Speech, Language and Hearing Research, 40, 765-777.*

Conti-Ramsden, G., & Hesketh, A. (2003). Risk markers for SLI: a study of young language-learning children. *International Journal of Language and Communication Disorders, 38, 251-263.*

Crystal, D. (1986). *Listen to your child.* Penguin Books, London, England.

De Candole, R. (1998). *Tauranga Test of Articulation.* GSE Tauranga office, Ministry of Education, New Zealand.

Dodd, B., Crosbie, S., McIntosh, B., Teitzel, T., & Ozanne, A. (2000) *Preschool and Primary Inventory of Phonological Awareness.* The Psychological Corporation, London, UK.

Dodge, E. P. (2004). Communication skills: the foundation for meaningful group intervention in school-based programs. *Topics in Language Disorders, 24, 2, 141-150*

Doury, D. L., & Nash, P. L. (2003). Epidemiology and aetiology of autistic spectrum disorders difficult to determine. *Thorofare, 32, 10, 696-705.*

Education Review Office (2004). Review of the Special Education Grant. *Ministry of Education, Wellington, New Zealand.*

Enderby, P., & John, A. (1999). Therapy outcome measures in speech and language therapy: comparing performance between different providers. *International Journal of Language and Communication Disorders, 34, 4, 417-429.*

Enderby, P., & Petheram, B. (1998). Changes in referral to speech and language therapy. *International Journal of Language and Communication Disorders, 33 supplement, 16-20.*

Enderby, P., & Phillipp, R. (1986). Speech and language handicap: towards knowing the size of the problem. *British Journal of Disorders of Communication, 21, 151-165.*

Feeney, J., & Bernthal, B. (1996). The efficiency of the revised Denver developmental screening test as a language screening tool. *Language, Speech and Hearing Services in Schools, 27, 330-332.*

Fluharty, J. (1978) *Fluharty Preschool Speech and Language Screening Test* in Paul, R. (Ed) (2002) *Introduction to clinical methods in communication disorders.* Baltimore, USA: Paul H. Brookes Publishing Company Incorporated.

.

Forness, S., Ramey, S. L., Ramey, C. T., Heu, C., Brezsausek, C. M., MacMillan, D. L., Kavale, K. A., & Zima, B. T. (1998). Head start children finishing first grade: preliminary data on school identification of children at risk for special education. *Behavioural Disorders, 23, 2, 111-124.*

Frazier Norbury, C., & Bishop, D. V. M. (2003). Narrative skills of children with communication impairments. *International Journal of Language and Communication Disorders, 38, 3, 287-313.*

Frazier Norbury, C., Nash, M., Baird, G., & Bishop, D. V. M. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the Children's Communication Checklist-2. *International Journal of Language and Communication Disorders, 39, 3, 345-364.*

Gardner, H., Froud, K., McClelland, A., & van der Lely, H. K. J. (2006). *Development of the grammatical and phonology screening (GAPS) test to assess key*

markers of specific language and literacy difficulties in young children. International Journal of Language & Communication Disorders, 41, 5, 513-540.

Gierut, J. A. (1998). Treatment efficacy: phonological disorders in children. *Journal of Speech, Language and Hearing Research, 41, S85-S100.*

Gillon, G. (2004). *Phonological awareness: from research to practice.* New York: The Guildford Press.

Gillon, G. T. (2005). Prologue: phonological awareness: evidence to influence assessment and intervention practices. *Language, Speech and Hearing Services in Schools, 36, 4, 281-284, 350.*

Gillon, G., Moriarty, B., & Schwarz, I. (2005). *An international literature review of best practices in speech and language therapy (update).* University of Canterbury, Christchurch, New Zealand.

Gillon, G., & Schwarz, I. (1998). *An international literature review of best practices in speech and language therapy.* Wellington: New Zealand Ministry of Education. ISBN:0-477-05092-1 (90pp).

Gillon, G., & Schwarz, I. (2001). *An international literature review of best practices in speech and language therapy (update).* Wellington: Special Education Services, New Zealand Ministry of Education.

Gillon, G., Schwarz, I., Miller, J., & Westerveld, M. (2002). The spoken language skills of New Zealand children. *Paper presented at NZSTA biennial conference, Wellington.*

Gillon, G., & Westerveld, M. (1999/2000). Narrative language sampling in young school-age children. *New Zealand Journal of Speech and Language Therapy, 53/54, 34-41.*

Glascoe, F. P. (2001). Can clinical judgement detect children with speech-language problems? *Pediatrics, 87, 3, 317-322.*

Glasoce, F. P. (2001). Are over referrals on developmental screening tests really a problem? *Archives in Pediatric and Adolescent Medicine, 155, 1, 54-9.*

Goldman, R., & Fristoe, M. (2000). *Goldman Fristoe Test of Articulation-2*. Circle Pines, MN, USA: American Guidance Service Incorporated.

Goorhuis-Brouwer, S. M., & Kniff, W. A. (2002). Efficacy of speech therapy in children with language disorders: specific language impairment compared with language impairment in co-morbidity with cognitive delay. *International Journal of Pediatric Otorhinology, 63, 2, 129-136.*

Goorhuis-Brouwer, S., & Kniff, W. (2003). Language disorders in young children: when is speech therapy recommended? *International Journal of Pediatric Otorhinolaryngology, 67 (5), 525-529.*

Guendouzi, J. (2003). 'SLI', a generic category of language impairment that emerges from specific differences: a case study of two individual linguistic profiles. *Clinical Linguistics and Phonetics, 17, 2, 135-152.*

Hall, N. E. (2004). Lexical development and retrieval in treating children who stutter. *Language, Speech and Hearing Services in Schools, 35, 1, 57-69.*

Harris, F. (2003). Continuing to climb. *Speech and Language Therapy in Practice, Winter, 20-21.*

Heins, K. (2003). Collaborating for communication. *Speech and Language Therapy in Practice, Winter, 11-13.*

Hesketh, A. (2004). Early literacy achievement of children with a history of speech problems. *International Journal of Language and Communication Disorders, 39, 4, 453-468.*

Hirst, E., & Britton, L. (1998). Specialised service to children with specific language impairment in mainstream schools. *International Journal of Language and Communication Disorders, 33, supplement, 593-598.*

Hogan, T. P., Catts, H., & Little, T. D. (2005). The relationship between phonological awareness and reading: implications for the assessment of phonological awareness. *Language, Speech and Hearing Services in Schools, 36, 285-293.*

Kennedy, M. (2002) in *Introduction to Clinical Methods in Communication Disorders, Edited by Paul, R.* Baltimore, USA: Paul H. Brookes Publishing Company Incorporated.

Kilmon, C. A., Barber, N., & Chapman, K. (1991). Instruments for the screening of speech/language development in children. *Journal of Pediatric Healthcare, 5, 2, 61-70.*

Law, J. (2004). The implications of different approaches to evaluating intervention: evidence from the study of language delay/disorder. *Folia Phoniatrica Logopaedia, 56(4), 199-219.*

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35(2), 165-188.*

Law, J., Garrett, Z., & Nye, C. (2003). *Speech and language interventions for children with primary speech and language delay or disorder.* London, ŪK: The Nuffield Foundation.

Law, J., Lindsay, G., Peacey, N., Gascoigne, M., Soloff, N., Radford, J., & Band, S. (2002). Consultation as a model for providing speech and language therapy in schools: a panacea or one step too far? *Child Language Teaching and Therapy, 146-163.*

Law, J., Lindsay, G., Peacey, N., Gascoigne, M., Soloff, N., Radford, J., Band S., & Fitzgerald, L. (2000). Provision for children with speech and language needs in England and Wales: facilitating communication between education and health services. *Nottingham: DfEE Publications. ISBN: 1-84185-398-4, Research Brief 239.*

Lees, J. & Urwin, M. (1997). *Children with Language Disorders, 2nd Edition,* London, UK: Whurr.

Lindsay, G., & Dockerell, J. E. (2004). Whose job is it? Parents' concerns about the needs of their children with language problems. *The Journal of Special Education, 37, 4, 225-235.*

Lindsay, G., Soloff N., Law J., Band S., Peacey N., Gascoigne, M., & Radford, J. (2002). Speech and language therapy services to education in England and Wales. *International Journal of Language and Communication Disorders, 37, 3, 273-288*

Lloyd, H., & Paintin, K. (2006). Performance of children with different types of communication impairment on the Clinical Evaluation of Language Fundamentals (CELF). *Child Language Teaching and Therapy, 22, 1, 47-67.*

Lonigan, C. L., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental psychology, 36, 5, 596-613.*

Lubker, B. B., & Tomblin, J. B. (1998). Epidemiology: informing clinical practice and research on language disorders in children. *Topics in Language Disorders, 19, 1, 1-26.*

McCartney, E., Boyle, J., Bannatyne, S., Jessiman, E., Campbell, C., Kelsy, C., Smith, J., & O'Hare, A. (2004). Becoming a manual occupation? The construction of a therapy manual for use with language impaired children in mainstream primary schools. *International Journal of Language and Communication Disorders, 39, 1, 135-148.*

McCauley, R. J. and Swisher, L. (1984). Psychometric Review of Language and Articulation Tests for Preschool Children. *Journal of Speech and Hearing Disorders, 49, 34-42.*

Marshall, J., Stojanovik, V., & Ralph, S. (2002). "I never even gave it a second thought": PGCE students' attitude towards the inclusion of children with speech and language impairments. *International Journal of Language and Communication Disorders, 37 (4), 475-489.*

Masterton, J. J., Bernhardt, B. H., & Hofheinz, M. K. (2005). A comparison of single words and conversational speech in phonological evaluation. *American Journal of Speech-Language Pathology, 14, 229-241.*

Meline, T., & Paradiso, T. (2003). Evidence-based practice in schools: evaluating research and reducing barriers. *Language, Speech and Hearing Services in Schools, 34, 273-283.*

Miller, J. F. (1992). Systematic Analysis of Language Transcripts [Computer software]. Retrieved from http://www.LanguageAnalysisLab.com

Miller, J. F. (1992). *Salt Reference Database Project.* Madison, WI: Language Analysis Laboratory, University of Wisconsin.

Ministry of Education (1997). Special Education 2000. *Wellington, New Zealand: Ministry of Education.*

Ministry of Education (1998). Communicate to Participate. *Wellington, New Zealand: Ministry of Education.*

Ministry of Education. (2002). Special Education 2000: Monitoring and evaluation of the policy. Final Report. *Wellington, New Zealand: Ministry of Education.*

Ministry of Education. (2004). *New Zealand Articulation Test (NZAT).* Wellington, New Zealand: Ministry of Education.

Ministry of Education. (2006). Special Education output reports. *Wellington, New Zealand: Ministry of Education.*

Ministry of Education (in progress). National Review of Speech-Language Therapy Services to Children and Young People. *Wellington, New Zealand: Ministry of Education.*

Muter, V., Hulme, C., & Snowling, M. (1997). *Phonological Abilities Test.* London, UK: The Psychological Corporation.

Nathan, L., Stackhouse, J., Goulandris, N., & Snowling, M. (2004). Educational consequences of developmental speech disorder: Stage 1 National Curriculum assessment results in English and Mathematics. *British Journal of Educational Psychology, 74, 2, 173-186.*

New, E. (1998). An effective model for a speech and language therapy service in mainstream schools. *International Journal of Language and Communication Disorders, 33, supplement, 602-607.*

Paul, R. (1995) *Language disorders from infancy through adolescence: assessment and intervention, 2nd edition.* St Louis MO: Mosby-Year Book Incorporated.

Paul, R. (2002). *Introduction to clinical methods in communication disorders.* Baltimore, USA: Paul H. Brookes Publishing Company Incorporated.

Pena, E. D., & Quinn, R. (2003). Developing effective collaboration teams in speech-language pathology: A case study. *Communication Disorders Quarterly Forum, 24, 2, 53.*

Peters-Johnson, C. (1996). Action: school services. *Language, Speech and Hearing Services in Schools, 27, 185-186.*

Portch, A. (2003). Out of the frying pan into the fire? *Speech and Language Therapy in Practice, Winter, 16-19.*

Reid, D. K., Hresko, W., & Hammill, D. (1989). *Test of Early Reading Ability.* Austin, Texas: Pro-ed..

Renfrew, C. E. (1997). *Renfrew Action Picture Test, 4th Edition.* Nottingham, England: Renfrew C. E.

Rescorla, L., & Alley, A. (2001). Validation of the language development survey (LDS): a parent tool for identifying language delay in toddlers. *Journal of Speech, Language and Hearing Research, 44, 2, 434-445.*

Rice, M. L, Sell, M. A., & Hadley, P. A. (1991). Social interactions of speech and language impaired children. *Journal of Speech and Hearing Research, 34, 1299-1307.*

Robey, R. R. (2004). A five-phase model for clinical-outcome research. *Journal of Communication Disorders, 37, 5, 401-411.*

Roulstone, S. (2001). Consensus and variation between speech and language therapists in the assessment and selection of preschool children for intervention: a

body of knowledge or idiosyncratic decisions? *International Journal of Language and Communication Disorders, 36, 3, 329-348.*

Rousseau, I., Packman, A., Onslow, M., Harrison, E., & Jones, M. (2006). An investigation of language and phonological development and the responsiveness of preschool age children to the Lidcombe program. *Journal of Communication Disorders, November 20.*

Royal College of Speech and Language Therapists (1996). *Communicating Quality, 2ⁿᵈ Edition.* London, UK: Royal College of Speech and Language Therapists.

Ruscello, D. M., Cartwright, L. R., Haines, K. B., & Shuster, L. I. (1993). The use of different service delivery models for children with phonological disorders. *Journal of Communication Disorders, 26, 3, 193-203.*

Russ, S., Chiang, B., Rylance, B. J., & Bongers, J. (2001). Caseload in special education: an integration of research findings. *Exceptional Children, 67, 2. 161-172.*

Russell, S. C., & Kaderavek, J. N. (1993). Alternative models for collaboration. *Language, Speech and Hearing Services in Schools, 24, 76-78.*

Rutter, M., Mahmood, L., & Howlin, P. (1992). *Language delay and social development.* In P. Fletcher & D. Hall (Eds.), Specific Speech and Language Disorders in Children. London: Whurr.

Sadler, J. (2005). Knowledge, attitudes and beliefs of the mainstream teachers of children with a preschool diagnosis of speech/language impairment. *Child Language, Teaching and Therapy, 21, 2, 147-163.*

Salvia, J. & Ysseldyke, J. E. (2004). *Assessment: in special and inclusive education.* Boston: Houghton Mifflin.

Schooling, T. L. (2003). Lessons from the national outcomes measurement system (NOMS). *Seminars in Speech and Language , 24, 3, 245-257.*

Semel, E. M., Wiig, E. H., & Secord, W. A. (1996). *Clinical Evaluation of Language Fundamentals-3rd Edition, Screening Test*. San Antonio, Texas, USA: Harcourt Assessment Incorporated.

Semel, E. M., Wiig, E. H., & Secord, W. A. (1996). *Clinical Evaluation of Language Fundamentals-Third Edition*. San Antonio, Texas, USA: Harcourt Assessment Incorporated.

Semel, E. M., Wiig, E. H., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals, 4th Edition, Screening Test*. San Antonio, Texas, USA: Harcourt Assessment Incorporated.

Semel, E. M., Wiig, E. H., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals, 4th Edition, Australian adaptation*. San Antonio, Texas, USA: Harcourt Assessment Incorporated.

Semel, E. M., Wiig, E. H., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals, Preschool, 2th Edition.* San Antonio, Texas, USA: Harcourt Assessment Incorporated.

Shriberg, L. D., Tomblin, J. B., & McSweeney, J. L (1999). Prevalence of speech delay in 6-year old children and co-morbidity with language impairment. *Journal of Speech, Language and Hearing Research, 42, 1461-1481.*

Speech Pathology Australia (2005). Learning to speak and listen: fact sheets. *Speech Pathology Australia, Melbourne.*

Stanton-Chapman, T. L., Chapman, D. A., Bainbridge, N. L., & Scott, K. G. (2002). Identification of early risk factors for language impairment. *Research in Developmental Disabilities, 23, 6, 390-405.*

Stott, C. M., Merricks, M. J., Bolton, P. F., & Goodyer, I. M. (2002). Screening for speech and language disorders: the reliability, validity and accuracy of the general language screen. *International Journal of language and Communication Disorders, 37, 2, 133-151.*

Summers, P. A., Larson, G. W., Miguel, S. A., & Terrell, S. L. (1996). Test-retest comparisons using the CELF-RST and BLT-2S with kindergartners. *Language, Speech and Hearing Services in Schools, 27, 324-329*

Sunderland, L. C. (2004). Speech, language and audiology services in public schools. *Intervention in School and Clinic, 39, 4, 209-217.*

Swenson, N. C. (2000). Comparing traditional and collaborative settings for language intervention. *Communication Disorders Quarterly, 22, 1, 12-18.*

Throneburg, R. N., Calvert, L. K., Sturm, J. J., Paramboukas, A. A., & Paul, P. J. (2000). A comparison of service delivery models: effects on curricular vocabulary skills in the school setting. *American Journal of Speech-Language Pathology, 9, 1, 10-21.*

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairments in kindergarten children. *Journal of Speech, Language and Hearing Research, 40, 6, 1245-1260.*

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39, 1284-1294.*

Tyler, A. (2005). *Assessment for determining a communication profile.* In Phonological Disorders in Children, Kamhi, A. G. (Ed) & Pollock, K. E. (Ed), Baltimore, MD: Paul H Brookes Publishing.

Ukrainetz, T. A., & Fresquez, E. F. (2003). "What isn't language?" A qualitative study of the role of the school speech-language pathologist. *Language, Speech and Hearing Services in Schools, 34, 284-298.*

Van der Gaag, A., McLoone, P., & Reid, D. (1999). Speech and language therapy caseloads in seven districts in the UK. *Journal of Managing Medicine, 13, 1, 23-32.*

Van der Lely, H. K. J. (2005). Domain-specific cognitive systems: insight from grammatical-specific language impairment. *Trends in Cognitive Sciences 9, 53-59.*

Van der Lely, H. K. J., Rosen, S., & Adlard, A. (2004). Grammatical language impairment and the specificity of cognitive domains: relations between auditory and language abilities. *Cognition, 94, 167-183.*

Washington, J. A., & Craig, H. K. (2004). A language screening protocol for use with young African American children in urban settings. *American Journal of Speech-Language Pathology, 13, 329-340.*

Westerveld, M., & Gillon, G. (1999/2000). Narrative language sampling in young school-age children. *New Zealand Journal of Speech-Language Therapy, 53-54, 34-41.*

Whitehurst, G. J., & Fischel, J. E. (1994). Practitioner review: early developmental language delay: what, if anything, should the clinician do about it? *Journal of Child Psychology and Psychiatry, 35, 4, 613-648.*

Whitmire, K., & Dublinske, S. (2000). Provision of speech-language services in the schools: working with the law. *Seminars in Speech and Language, 14, 2, 147-152.*

Whitmire, K. (2002). The evolution of school-based therapy services: a half century of change and a new century of practice. *Communications Disorders Quarterly, 23, 2, 68-77.*

Wright, J. A. (1992). Collaboration between teachers and speech therapists with language impaired children. In Fletcher, P. & Hall, D. (Editors), *Specific speech and language disorders in children,* London, UK: Whurr.