UCL

Assessing Speech Fluency Problems in Typically Developing Children Aged 4 to 5 Years.

Avin Mirawdeli

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

University College London

ACKNOWLEDGEMENTS

There are numerous people who have contributed extensively to this work.

Professor Peter Howell, Thank you! Not only have you provided help, direction and guidance whenever needed, but you have also enabled me to be free and independent. By offering me this PhD post, I was able, for the first time, to live outside of the four walls that my culture had imposed on me. For this, no amount of thank you and not a single word exists that would sufficiently express and show my gratitude. Your dedication and the hours you devote to making sure the work is absolutely perfect is incredible and truly inspiring.

Mr and Mrs Barker and the Dominic Barker Trust, I am forever indebted to you, not only for funding this work, but also for welcoming me into your home and providing much support and assistance. Again, without your funding, I would not have been able to do this work. I am truly honoured to have met you all. Thank you for everything.

UCL EP and Impact Fund, Thank you. You have enabled me to meet the most interesting and most successful people around in this field. The knowledge I have gained here is invaluable.

Dr Alastair McClelland and Professor Charles Hulme, thank you for all your help and support.

The schools, their participation and warm welcome are greatly appreciated. I would like to thank all the staff, teachers, pupils and their parents for taking part in this project. With their help I have managed to screen 1300 children in 11 schools, so a very big thanks to:

Beecholme Primary School, London
Bond Primary School, London
Hatfeild Primary School, London
Hollymount Primary School, London
Lonesome Primary School, London
Poplar Primary School, London
Priory Primary School, London
Stanford Primary School, London
St Thomas of Canterbury Catholic Primary School, London

And

St. Matthew's CoE Primary School, Ipswich
St. Helen's Primary School, Ipswich.

Student assistants, thank you so much for accompanying me to the schools and helping with bringing the children and keeping them calm and entertained and for all your help in the office too.

The PhD students who helped by provided initial guidance when I started: Dr Lucy Riglin, Dr Terry Ng-Knight and Dr Chris Street.

A big thank you to all my friends and colleagues for their invaluable support and encouragement throughout: Terza and her parents, Nina, Celine, Matthew Jones, Eryk Walczak, Naheem Bashir, Dr Laura Speed and Dr Anthony Buhr.

Finally, I would like to thank my uncle, Dr Kamal Mirawdeli and my sister, Sophia-Dila for always being there.

Declaration

I, Avin Mirawdeli, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Avin Mirawdeli.

Abstract

This thesis addresses the identification of children with expressive speech difficulties with a focus on stuttering. It is based on theoretical work that investigated the symptoms associated with stuttering (Howell, 2013). It also has a practical goal: The procedures that have been developed should help determine the risk of a child acquiring some form of speech difficulty. The children examined had just entered school (4-year-olds). To ensure reliable results were obtained, large, representative samples of children were required. Most of the children do not have speech difficulty. A sample of speech was obtained and analysed. The approach taken in analysis was to use an instrument that has been standardised and is currently used in research (Riley, 1994) and to apply it to the assessment of speech difficulty. Howell (2013) showed that this instrument is effective in screening for stuttering.

The background to the screening work with stuttering is given in the literature review in Chapter 1. The challenges that arise when screening (a form of risk factor modelling) a real-world sample are discussed. Definitions and general features of stuttering are presented and various theories concerning how stuttering symptoms arise are reviewed. Chapters 2 to 6 report background studies, fieldwork and analyses that were conducted. Chapter 2 reports the results of a survey that was conducted to determine whether there was a need for a screening instrument and, if so, what form it should take. Chapters 3 and 4 report studies that were conducted to balance the need to keep assessments in schools short with ensuring the procedures are reliable and valid when used to identify children with speech difficulty. The assessments were based on Riley's (1994) Stuttering Severity Instrument. Chapter 3 determined the minimum length a sample of speech needed to be and whether a spontaneous speech sample was sufficient when using Riley

(1994) for assessing children for speech difficulty. The Stuttering Severity Instrument has three components (percentage of syllables that are not fluent, duration of selected long stutters and a measure of physical concomitants to stuttering). Chapter 4 addressed whether all three components are required to identify children with speech difficulty, since assessing fewer components would keep the procedure simple for use in schools. Chapter 5 reports an extensive field study that used Riley (1994) for identifying children with speech difficulty. Chapter 6 examined whether adding additional symptoms to those available in the Stuttering Severity Instrument that are appropriate for other common paediatric forms of speech difficulty would enhance accuracy of screening performance. Chapter 7 summarizes the work, draws conclusions and identifies future directions this research should take.

CONTENTS

Figures

Tables

Table of abbreviations:

A list of abbreviations used throughout this report

Abbreviation	Meaning
AWS	Adults who stutter
CRH	Covert repair hypothesis
CWS	Children Who Stutter
ECAT	Every Child a Talker
EAL	English as an Additional Language
HL	Hearing Loss
PC	Physical Concomitants
PDel	Phonological Delay
PDis	Phonological Disorder
SEN	Special Education Needs
SLCN	Speech, Language & Communication Needs
SLT(s)	Speech and Language Therapist(s)
SFS	Speech Filing System
SSI	Stuttering Severity Instrument

Chapter 1

Literature Review and Introduction to the Topic

The Bercow Report (2008) steered a change in the way specialist speech and language services for children are delivered in the U.K. In addition, the financial climate has led to funding cuts that have resulted in resources being reduced and distributed unevenly. The inequities in access to these resources have been partly disguised because of limited communication between speech and language therapy services, schools and parents of children with communication difficulties. This has resulted in schools playing a bigger role in identifying children who are considered to be at-risk of expressive speech difficulties (speech difficulty for short) and in deciding whether to refer them to already over-stretched SLT services (Mirawdeli, Dockrell & Howell, in prep). The teachers' roles in this respect are put under additional pressure because of time constraints that some local government authorities have imposed on referral and treatment. For example, in the London Borough of Merton and Sutton, children cannot be referred for SLT after they have completed the reception class (i.e. the first school year when children are aged between 4 and 5 years of age). Consequently, children who develop speech difficulties at later ages or who are not identified in reception classes would not receive help, unless the parents were prepared to secure intervention privately or could successfully argue for access to funded services.

Research has shown that children referred for intervention at later ages are less likely to report a successful outcome than those referred at an early age (Howell, 2010). Absence of intervention increases the risk of adverse consequences on a child's later quality of life and educational attainment (Department for Education, 2013; Marmot Review, 2010). For these reasons, identifying children with speech difficulty at the ages of four to five years is recommended so that effective early intervention can take place (Mirawdeli, 2015).

One solution to inequity of access would be to implement a national screening program. This would meet the teachers' requirement of early detection of problems. Missing children with speech difficulty is a risk when the current subjective approach is applied whereby children are selected by the teachers without any guidance. A screening program would reduce the chances of missing children. The benefits of screening overall would be that the children could be targeted early before the chance of recovery declines and that children's later educational attainment and quality of life would improve. This is of obvious importance since the Marmot Review (2010) and the report from the Department for Education in the United Kingdom (https://www.gov.uk/government/publications/early-years-foundation-stage-profile-results-2012-to-2013) reported that only 52% of children had reached a "good level of development" in their readiness for entry into year one of school in the UK. Many of the recommendations for improving this statistic have focused on improving the speech and language skills of the children. In order for this approach to be inclusive, provision of support services needs to be improved. Then children in need of extra help would not pass through the system with undetected problems.

To approach the question of screening reception class children for speech difficulty, this thesis used Riley's (1994) Stuttering Severity Instrument version 3 (SSI-3), is a standardized instrument for assessing fluency that is conducted on a short, spontaneous speech sample. Riley (1994) indicated that it could be used for identifying children with speech (specifically fluency) difficulty, and Howell (2013) developed procedures to use it for this end. Its brevity and the fact that it can be based on spontaneous speech alone, make it suitable as a basis for screening young children (Mirawdeli et al., in prep, Chapter 2). Sample length and validation that spontaneous speech alone suffices for assessment are addressed in Chapter 3 of this thesis (Todd, Mirawdeli, Costelloe, Cavenagh, Davis & Howell, 2014). Other reasons for considering that SSI-3 can be used effectively as a screening tool are discussed in the present chapter and Chapter 4 (and the latter addresses whether SSI-3 can be simplified for screening purposes by focusing on certain components of the measure). Results on the use of SSI-3 as a screening tool in the field (i.e. in schools) are reported in Chapter 5. The SSI-3, as mentioned above, is a tool primarily for assessing disfluency. However, a screening tool would need to take into account more than just disfluency symptoms. Chapter 6 considers the type of symptoms that should be used when screening children for speech difficulty including stuttering and other problems that affect speech.

Analyses in all chapters (except for Chapter 2) were conducted on speech samples taken from over 900 children in 11 schools (9 in London and 2 in Ipswich). All children were audio-recorded and their speech samples were evaluated. The samples were recorded in English for all the children including those children who used English as an additional language (EAL). Spontaneous monologues were elicited from each child using the picture stimuli that are provided in the SSI-3 manual. Sensitivity and specificity in identifying children with speech difficulty were assessed relative to teachers' judgments. The teachers indicated which children: 1) they believed to be at risk based on regular contact with each child or where the child or parents had expressed concern; 2) were already in an intervention program or 3) were in the process of being referred for an intervention. It is important that a screening tool that is used by teachers is practical, not time consuming but takes into account a comprehensive symptom set appropriate to common types of speech difficulty that children at this age can exhibit. With this in mind, the work in Chapter 4 looked at ways of simplifying SSI-3. This work supports focusing on the frequency of symptom measure in SSI-3. A systematic review intended to identify an extended set of symptoms is reported in Chapter 6. The review showed that an alternative method of identifying additional symptoms was required. Additional symptoms were identified from expert sources and these were used for classifying selected cases in Chapter 6. These results provide recommendations about how to improve symptom assessment when screening for speech difficulty. In the concluding chapter, the effectiveness of this screening approach for all children, including children with English as an additional language (EAL) is discussed. Some drawbacks and limitations of the screening program and ways of addressing these are considered.

1.1 Use of SSI-3 as the Instrument Chosen to Screen Children in this Thesis

In this thesis the SSI-3 was used to identify children who may be at risk of speech difficulty, in a typically developing population; this is termed screening. Typically developing children aged 4 to 5 years were examined. The intention was not to categorize the children into severity classes (mild, severe etc.) but to use raw SSI-3 scores (standard ones and in the future in modified forms based on the findings of Chapters 4 and 6) to establish whether this approach to screening for speech difficulties works. The use of SSI-3 in screening children was first raised in a study by Howell and Davis (2011). They assessed a group of 300 children who stuttered (CWS), and followed them up until teenage, at which point the children were divided into those who persisted and those who recovered from stuttering. Models that predicted prognosis were developed. Howell and Davis (2011) looked at a range of measures taken at age 8 (before it was known whether the CWS would recover or persist). The measures included gender, age of onset, family history of language problems, handedness, whether the child spoke more than one language, previous head injury and SSI-3 score. All of the measures have been shown to lead to significant differences when groups of fluent speakers and speakers who stutter are compared (Ajdacic-Gross, Vetter & Muller, 2010; Brosch, Haege & Kalehne, 1999; Dworzynski, Remington & Rijksdijk, 2007; Howell, Davis & Williams, 2008; Mannson, 2000; Seider, Gladstone & Kidd, 1983; Segalowitz and Brown, 1991; Yairi & Ambrose, 2004;). The analysis that was conducted used a backward stepwise procedure. This method started with all the variables included in the first step, then each predictor was removed (one after the other) starting with the least useful predictor to see if its removal significantly affected the model fit. The results showed that SSI-3 was the only factor that predicted which children would persist and which would recover (Howell & Davis, 2011). The model with SSI-3 score at age eight alone as predictor performed well, achieving around 80% specificity and sensitivity.

Howell (2013) validated this model and extended it to show that it could be applied successfully in distinguishing CWS from fluent children. In this validation archived data were used. SSI-3 was calculated in various ways including variants where one symptom sometimes considered being stutters (whole-word repetitions) were either included or excluded in the SSI-3 severity score calculations. The reason for assessing with and without whole-word repetitions was that there is disagreement about whether whole-word repetition (WWR) should be included or excluded. Riley (1994) argued that whole-word repetitions should be excluded when calculating the severity ratings; this view is also shared by Howell (2010). Since the standards were developed with WWR omitted, WWR should always be excluded when an SSI-3 score is calculated. However, other researchers suggest that WWR should be included as a symptom of dysfluent speech (Yairi & Ambrose, 1992). The analyses of the archived data with WWR included and excluded showed that models which excluded WWR distinguished between CWS and fluent children and also between CWS who recovered and those who persisted. The results for the models with WWR excluded were better than those for the models where WWR were included. It has also been shown that children who have many WWR in their speech are more likely to recover than those who display other symptoms of stuttering such as prolongations, part-word repetitions and blocks (Howell, Bailey & Kothari, 2010). This suggests that WWR have a role in recovery of stuttering as opposed to being a symptom of that speech difficulty.

The role of WWR in early speech development of fluent children and children who stutter is also of particular importance when considering children whose first language is not English; the presence of WWR in their speech probably reflects word-finding difficulties rather than stuttering. Thus a child who says “my, my, my, apple” may be experiencing word-finding difficulty on ‘apple’ and be repeating ‘my’ to retain the conversation turn. The screening work done by Howell (2013) to identify the best method for use by different groups of professionals (be it screening by teachers or speech language therapists, SLTs, in schools) validated the screening for fluency with WWR excluded and it classified 84.4% of fluent children, and 88.0% of CWS, correctly. A strong feature of the validation was that some of these children differed in age from those used to develop the model, and analysis of their data showed that the SSI-3 was successful at screening for fluency for them. Therefore children outside the range used in modelling (younger children in particular) can be screened using SSI-3. This (along with other reasons discussed later) is why this thesis assessed reception class children (i.e. 4-5 year olds).

In the case of SLTs, it is useful to know which CWS will persist or recover so that this can be taken into account when deciding about service provision. They may wish to concentrate on those children who are likely to persist through to adulthood.). Howell (2010; 2013), argued that factors that precipitate the onset of stuttering or those that predict prognosis would not necessarily work when examining screening. Moreover, the factors that increase the risk of persistence are not necessarily the same as those that increase the risk of starting to stutter. The same argument was made by Reed and Wu (2013), who also pointed out that when risk factor analysis is conducted some studies reverse the relationship between dependent and independent variables: The outcome becomes the independent variable, while the predictor becomes the dependent variable. They consider that risk factor modelling using logistic regression should benefit research into speech difficulties because it could lead to improvements in diagnosis, prognosis and assessment of treatment outcome (Howell, 2010; Reed & Wu, 2013). This approach has particular relevance to the present thesis as the aim was to single out the factors that successfully distinguish children with speech difficulty from typical children.

The preceding arguments indicate that SSI-3 should be explored as a screening instrument. However, there is some ambiguity about what severity score should be used as a cut-off value to use for fluency (see Chapter 4 for further discussion). No designated cut-off score was given by Riley (1994; 2009) even though he stated that the instrument can be used to aid diagnosis, which implies fluent children, can be separated from CWS. Howell and Davis (2011) estimated that an approximate threshold SSI-3 score of eight would separate fluent children from CWS. Howell (2013) subsequently revised the value to 13 based on empirical work (see section 4.3.2 for further discussion concerning what threshold is appropriate). Howell (2013) also noted that the threshold needed to be flexible depending on the type of screen undertaken and that the exact threshold may depend on the SSI-3 assessment procedure that was used because Riley described several procedures that can be employed when obtaining an SSI-3 score that lead to different SSI-3 scores

(Howell, Soukup-Ascencao, Davis & Rusbridge, 2011; Jani, Huckvale & Howell, 2013).

A further fact to note is that Howell (2013) used a time-intensive computer-based annotation method to obtain frequency and duration measures. This raises an issue since when large numbers of school-aged children are screened, time-consuming procedures would not be possible (or necessary when children are fluent). Instead, children's SSI-3 scores could be obtained with a quicker clinical assessment procedure (allowed by Riley, 1994; 2009). Another procedural variant that Riley (1994) allowed was to use a spontaneous sample alone versus use of a spontaneous and read sample. Some four- and five-year old children can only provide a spontaneous sample. Although this should give reliable results (a table for non-readers is given by Riley, 1994, 2009) it was decided to check this in Chapter 4.

Another procedural issue concerning use of SSI-3 is whether 200 syllables give a reliable measure of symptom frequency or whether a longer, or conceivably a shorter, sample could be used. The shorter the sample length the quicker and more efficient the procedure would be. Sample length is a further issue that was addressed (Todd et al., 2014 and Chapter 4). Research has shown that around 6-10% of young children may suffer from a paediatric speech problem (Broomfield & Dodd, 2004b; Howell, 2010). Even though this rate is high, these studies could be underestimating incidence rates as Broomfield and Dodd (2004b) excluded stuttering and Howell (2010) only refers to studies that have estimated stuttering incidence. If the incidence of speech difficulty is around 6-10%, screening of children requires large samples so that statistical analyses are valid (Howell, 2010; Reed & Wu, 2013).

1.2 Demand on Services vs. Demand for Services

One factor stressed earlier is that this PhD was conducted with samples of children where small numbers have speech difficulty. The reasons why this is necessary are given in the Bercow (2008) report. The report investigated the services that were available for children with communication disorders in general (including speech difficulty) and the impact these services have on the children and families. It also delved into the co-ordination of services between schools, local councils and referral services. It found that services were limited and not easily accessible; parents were in a state of confusion about services available to them and sometimes did not understand the information that they were given. Overall, there was a lack of co-ordination between the schools and the services to which children were referred. Finally, the report suggested that early intervention is vital for successful recovery from communication disorders. This provides further justification as to why screening should take place at an early age and perhaps at a national level. Concern has been expressed because children are usually referred to clinics when they are older than the age at which they start school (Landa, Holman & Garrett-Mayer, 2007). Early detection of communication disorders has also been shown to reduce complications in a child’s education (McLean & Cripe, 1997).

However, early selection and national screening of children would require an expansion of currently available resources (Bercow, 2008); Services are already stretched and over-subscribed with waiting lists of six months and higher in South-West London (http://www.smcs.nhs.uk/files/smcs-service-directory.pdf). However, whilst demand for services may increase in the short term if screening was conducted, subsequently there would be less initial demand and less need for protracted call on services after effective intervention (assuming early intervention to be more effective than later intervention).

At a general level, a screen for children with speech difficulties might seek to distinguish the children who are fluent from those with a problem based on the dysfluent characteristics in their speech. Disruption to speech could be indicative of speech or hearing loss and, possibly, also language disorders. Here the focus is on speech difficulties, including those that originate from hearing loss (the term *speech difficulty* is used for all of these for brevity). In more refined analyses of speech the different speech symptoms may permit identification of particular types of speech difficulty. Chapter 6 describes work on extending the SSI-3 symptom set in the analyses to incorporate symptoms from other types of speech difficulties. This could potentially allow different types of speech difficulty to be identified.

Part of the reason for Bercow's recommendations was that communication disorders are a barrier that prevents children achieving their full potential in school. Specialist remedial support can be obtained if children have a Statement of Special Educational Need (see http://www.bdadyslexia.org.uk/about-dyslexia/parents/staments.html for advice about how to obtain one). In the situation under consideration (that screening be provided), a statement could lead to support from SLT services. Provision for SLT (and for healthcare in general) in the UK's National Health Service is free at the point of delivery. To receive support, in principle, a child can be referred for SLT by the parents, by self-referral in older individuals, by the school if a statement is available or by a general medical practitioner. However, in practice, access to SLT is rationed because of heavy caseloads, reduced central government support and the shift toward commissioning services. This need for rationing has led some boroughs within London to restrict the ways in which children are referred for therapy; earlier the examples of the London Borough of Merton and Sutton was given, where restrictions based on age and the school year that the child is from are in place when making a referral. In detail, in the London Borough of Merton and Sutton, a child may only be referred up until the end of the reception year (4 to 5 years old) http://www.smcs.nhs.uk/files/smcs-service-directory.pdf (page 42, point number two under the sub-heading “inclusion Criteria”). The net effect is that although clinics, central and local government and others have good intentions, currently no procedure is available to ensure the equity of access to services that Bercow recommended. A screening instrument for speech difficulties would allow schools, as well as parents, to be informed about each child’s communication problems and to seek further advice.

In sum, time constraints are imposed on how quickly local government services need to be alerted if they are required to provide support to a child with speech and ;language communication needs (SLCN). A child starting school enters a reception class between the ages of four and five. The child would have to be screened, the data analyzed and presented in the form of a preliminary report for each child within 11 weeks of the child starting school (i.e. the first term they attend school). There are reasons why screening for speech difficulty at age four is advisable (see Chapter 2 for further discussion). School entry is a convenient age at which to check for speech difficulties since the children are in class groups for the first time. DSM-IV and DSM-5 note that speech difficulties are usually diagnosed during childhood or adolescence. For instance, in the case of stuttering, the disorder is not congenital, but starts at different points in early childhood (Andrews & Harris, 1964). Yairi and Ambrose (2005) reported 85% of stuttering started before age of 3.5 years and 65% started before the age of 3 years. They also mentioned that around 5% of pre-school children have stuttering-like symptoms, however, most children recover from their stuttering, as prevalence (the total number of cases of a specific disorder or disease found in a population during a specific time) in the adult population is around 1%.

1.3 Taking Screening Beyond Stuttering and Demand for Services

A screening program may need to identify speech difficulties other than stuttering. The usefulness of a screening instrument that selects only those children with a particular type of disfluency (e.g. stuttering) is valuable but limited. A comprehensive instrument needs to be able to efficiently select those children with speech difficulties of any type. Thus, it would need to include speech difficulties that are related to hearing, phonological delay, and disfluency, all of which affect speech. To this end the most apparent feature for assessing speech difficulties (including hearing) is disruption to fluency. As mentioned, a short screen with known accuracy would meet the timing constraints required by schools. It could help provide indications of the efficacy of early intervention that the Bercow report and SLTs emphasize as important, and SLTs want to contact young children with speech difficulties as early as possible (Mirawdeli et al., in prep; Chapter 2). A screen would provide some documentation on the incidence of speech problems in general in the school population at large and, if a screen is made in more depth, it could identify children with specific speech difficulties; the relative incidence of different types of speech difficulties could be documented. As mentioned, the incidence study of Broomfield and Dodd (2004b) on a UK sample that excluded stuttering found that 6.4% of typically developing children suffered from speech difficulties. The types of difficulties they found were classified as phonological delay (57.5%), consistent non-developmental errors (20.6%), inconsistent errors on the same lexical item (9.4%) and articulation disorders (12.5%). Having such information available on more speech difficulties and more sub-types would provide essential data for health policy providers.

Clearly the timing considerations imposed by schools and the high number of children that need to be assessed place considerable constraints on what can be performed when conducting a screen (Chapter 2). For instance, full clinical assessments, that are appropriate when speech or hearing problems are suspected, are not feasible. On the other hand, any procedure that is used needs to be backed up by scientific research and results that indicate that it meets required levels of performance. As discussed in section 1.1, SSI-3 looks like a reasonable starting point as it is supported by scientific research that has shown that it can screen children with one form of speech difficulty (stuttering) off from fluent children (Howell, 2013). It would take a short time to administer in school. Looked at from another perspective, conducting and analyzing SSI-3 bears some resemblance to what teachers do informally when monitoring children for communication disorder. They have ancillary information from parents in some cases which they use in making decisions about children (the teachers' decisions are used in the present work for case validation of children with speech difficulty). These factors all support formulating the hypotheses that SSI-3 may help screen children for all speech difficulties or to separate types of speech difficulties. Although the eventual idea is to identify children with any type of speech difficulty without necessarily identifying the specific type of speech difficulty a child has, here stuttering symptoms were used as the starting point. This is justifiable as stuttering is one of the commonest types of childhood speech difficulty, its symptoms are comorbid with other speech difficulties (as the results in Chapter 5 show) and the starting point of the screening procedure was the established Stuttering Severity Instrument (Riley, 1994). In the next section, definitions of stuttering and selected theories of stuttering that account for how the symptoms arise are reviewed. In a later chapter, a systematic review attempted to identify what symptoms should be added to the SSI-3 list to make it applicable to all types of pediatric speech difficulties, and explored their use briefly (Chapter 6).

1.4 Definitions and Theories of Stuttering

Stuttering has been investigated for many years and has led to views on how it is caused. For instance, Bloodstein (1995) indicated that Aristotle had a theory about stuttering in which an abnormality of the tongue prevented fluent speech. Bloodstein (1995) categorized the theories which attempt to explain stuttering into three types; first, there are theories which explain the etiology of stuttering, second, there are those that concentrate on the instant of stuttering and finally, there are attempts that reformulate existing theories to address both previous topics.

Theories of stuttering are important as they stimulate interest into the phenomenon, have potential benefit for its treatment and suggest how likely the chance of recovery from stuttering is. However, despite many years of research, there is still no consensus with respect to stuttering definitions and symptoms. In turn, the lack of consensus leads to different opinions when attempts are made to pinpoint the etiology of stuttering especially when the person who stutters may also present with social and emotional disorders as well as speech problems. For instance, Howell (2007) gave an example of how some children who are referred to clinics occasionally meet all but one of the DSM-IV criteria for the diagnosis of selective mutism. The similarity between stuttering and selective mutism arises because children who stutter may refrain from speaking because of the inconvenience of their stutter. To be diagnosed with selective mutism, a person must not have an additional diagnosis as stuttering (APA, 2000). Howell (2007a) argued that the latter criterion is included as a means of dissociating between the two separate disorders and it is not an inherent characteristic of mutism nor stuttering. Most definitions of stuttering emphasize the importance of language factors in stuttering whilst at the same time allowing emotional or psychological factors to play a role. Wingate (1964) presented a "standard" definition of stuttering which took into account the speech factor in stuttering and allowed other related factors to have an influence. His definition has seven points which are arranged under three category headings:

> "I. (a) disruption in the fluency of verbal expression, which is (b) characterized by involuntary, audible or silent, repetitions or prolongations in the utterance of short speech elements, namely sounds, syllables and words of one syllable. These disruptions (c)

> usually occur frequently or are marked in character and (d) are not readily controllable. II. Sometimes the disruptions are (e) accompanied by accessory activities involving the speech apparatus related or unrelated body structures or stereotyped speech utterances. These activities give the appearance of being speech related struggle. III. Also, there are not infrequently (f) indications or report of the presence of an emotional state, ranging from a general condition of "excitement" or "tension" to more specific emotions of a negative nature such as embarrassment, fear, irritation or the like. (g) “The immediate source of stuttering is some incoordination expressed in the peripheral speech mechanism; the ultimate cause is presently unknown and may be complex or compound.” (Wingate, 1964, p. 488).

This definition from Wingate (1964) clearly states that factors other than speech may have an influence on stuttering. For example, Wingate mentions accessory and socio-emotional states. By accessory Wingate means factors that are not necessarily speech-related and are idiosyncratic in nature. Wingate also referred to motor factors in stuttering that can have an adverse effect on fluency and these are given under (g) in his definition.

Subsequently, Howell (2010a) categorized the different models of stuttering into three separate categories by focusing on language and speech-motor factors that he considered paramount: (a) those that attribute stuttering to language factors, (b) those that attribute stuttering to motor factors, and (c) those that propose an interaction between motor and language factors. These three categories are discussed briefly before moving on to discussing the symptoms of stuttering.

1.4.1. Models that Attribute Stuttering to Language Factors.

Here three main theories are described, all of which attribute stuttering to language factors alone. The three theories are Levelt's blueprint theory (1989), the Covert Repair Hypothesis (CRH) by Kolk and Postma (1997) and the Vicious Cycle Account by Bernstein Ratner and Wijnen (2007).

1. Levelt's blueprint.

Levelt argued that perception and production of sounds are both involved in speech output control. In Levelt’s Blueprint, production starts at conceptualization (see Figure 1.1) and then the information is sent to the formulation stage where the speech ‘concept’ goes through a hierarchy of linguistic processing steps. The final formulation stage is to produce a phoneme string. Speech is then output by sending instructions to the speech-motor system. Levelt’s blueprint considers that errors can occur at any stage and that performance is then affected.

Any errors that ensue are detected by the perceptual system. The perceptual system can receive error information as input via two separate routes: the internal and external loops. The external loop operates post speech production and the perception system in the model can recover the error after the speaker has heard it. The internal loop operates prior to speech output by sending linguistic information to the perception system immediately it is available (i.e. before the sound is output). After the perception system has decoded the information received about what language message has been formulated it is then sent back to a monitor in the conceptualizer (a feedback process). If during feedback the information concerning intended and actual speech achieved correspond, then language has been correctly formulated and the next extract of speech is processed. If there is a discrepancy, however, then an error has occurred and speech may be interrupted so that the problem can be fixed.

If a speech error is evident in the external loop then an overt error has occurred, which the monitor detects then corrects speech output. If, however, the speech is detected through the internal loop then the error can be covert (no sign of the actual error in the output). There will be signs of the covert error present in the speech, which indicate an error had occurred internally but was corrected prior to production. For example, there may be hesitancy in the speaker's utterances as in 'I went to the, the garden'. In this case, the speaker has interrupted the speech, as indicated by the comma, and then repeated the word 'the', which can be seen as a sign of hesitancy or delay. These disfluencies indicate that the perception system noticed the error that was about to happen over the internal loop and corrected it prior to production. Whole-word repetition, that occurs in this example, is a symptom that some authors consider as stuttering (Howell, 2010a), and Levelt (1989) offers one account concerning why this symptom arises. However, it should be noted that Levelt (1989) did not use his blueprint to illustrate stuttering as such. Therefore, it is not an account of how all speech disfluencies occur in stuttering.

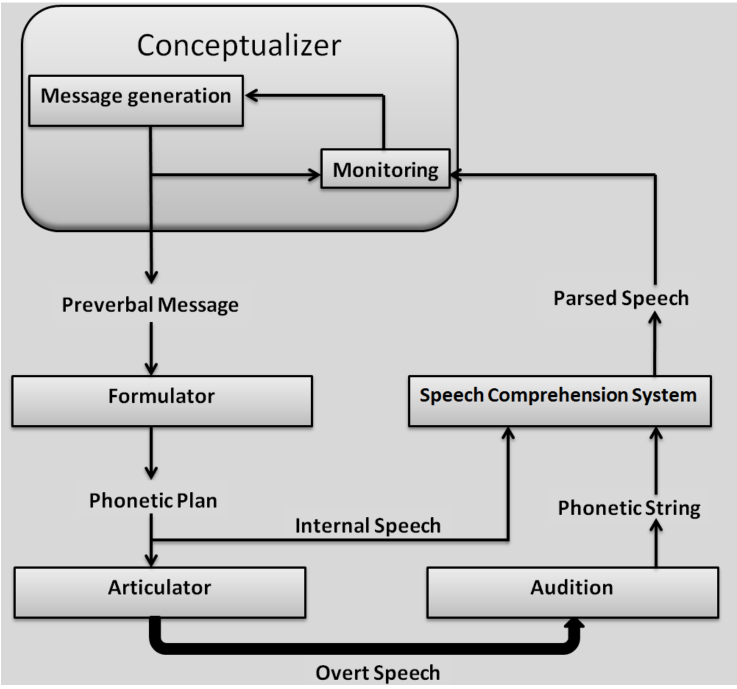

***Figure 1*. Levelt's Blueprint for control of speech production**

2. Covert repair hypothesis (CRH).

Kolk and Postma (1997) used Levelt's model and developed it to account for stuttering. They argued that when an error occurs in the process of transforming thought into speech, this causes disfluency in speech, and stuttering is an indication of breakdown in this process. Their theory argues that the disfluencies are instances of the speaker's problems in phonological encoding and their failed attempts at adaptation and self-repair (Postma & Kolk, 1993). Furthermore, they suggested that people who stutter make many repairs because they have a slow rate of word activation compared to fluent speakers because of the problems in their phonological encoding system. Because the words are not fully activated, the chance of an error in lexical selection increases. The speaker who stutters, once aware of this error over the internal loop, (i.e. before the speech is output, termed covert repairing) may try and repair the error, leading to a high incidence of hesitancy in their speech (Postma & Kolk, 1993). A feature of the Levelt and Kolk and Postma models is that they both consider whole-word repetitions to be indications of errors having been made. As discussed, this contrasts with the position taken by Riley (1994).

3. The vicious cycle account.

Bernstein Ratner and Wijnen (2007) argued that CWS have impaired language ability. They also have a strict or overactive monitoring system which invests too much time in monitoring for disfluencies in their speech. As Bernstein Ratner and Wijnen (2007, p.88) put it, they have an intolerance for ‘less-than-perfect’ speech output. In earlier work, Bernstein Ratner (1997) argued that the weakness in language was a syntactic issue. In the vicious cycle account it is also argued that a weak phonological system accounts for the development of stuttering at later stages (Bernstein Ratner & Wijnen, 2007).

1.4.2. Models that attribute stuttering to motor factors.

Several studies support a motor component as underpinning stuttering (Smith & Kleinow, 2000; Van Lieshout, Hulstijn, & Peters, 2004). Some employ models such as that of Kelso (1995), which adopted a nonlinear theory for motor control namely the dynamic systems concept. At the other end of the spectrum there are studies that argue for an interaction between motor and language components (Peters & Starkweather, 1990). These argue that linguistic variables have an influence on stuttering but the people who stutter have incompetency in motor control systems that are taxed when linguistic demand is high (Peters, Hulstijn & van Lieshout, 2000).

In a more detailed model, Howell (2004a) suggested that if linguistic plans, for subsequent words, arrive late and they are not completely ready for output, then the motor system may delay execution of the word itself by repeating whole words or phrases that precede the word that is not ready. There are many studies which support this hypothesis including those which have argued for categorization of disfluencies into either motor or linguistic types (Manning & Shirkey, 1981; Wexler & Mysak, 1982). Areas where there is evidence that motor factors are involved in stuttering are reviewed briefly next.

1) Lee's delayed auditory feedback (a motor influence)

Lee (1950) argued that delayed auditory feedback (DAF) in fluent speakers slows speech by elongating vowels, the speech is characterised by high amplitude and a monotone pitch. These features are evident on the vowel sounds in DAF speech. Lee also argued that the speakers listen to their speech and conduct a complete linguistic analysis of it and use the results to establish if there are discrepancies between the output speech and the original intended speech, in a similar way to the feedback process discussed in connection with Levelt's theory. Here, if there is a discrepancy, then the speaker tries to correct the output by altering the motor timing of the speech. In the case of DAF, timing is altered although no correction was needed on the output speech (Fairbanks, 1955; Lee, 1950; Levelt, 1989) – the experimenter 'tricked' the speaker into thinking an error had been made. Furthermore, in Lee's model, DAF can disrupt all levels of language monitoring that are carried out by speakers (Lee, 1950). It was initially proposed that DAF in adults who stutter caused the speech to become fluent while in fluent adults the speech became dysfluent (Lee, 1950). However, as mentioned above, delayed auditory feedback disruptions appear in the form of prolongation on medial vowels (Howell, Wingfield & Johnson 1988) but it has been shown that prolongations in stuttered speech occur on initial consonants (Brown, 1945).

2) Local and global rate change (another motor influence).

Speakers can alter their mean speech rate, allowing it to either go up or down. When all speech is adjusted, this is referred to as global rate change. When selected parts within an utterance are altered, this is called local rate change (Howell, 2010a). Much research has shown that disfluency rates reduce dramatically when speech rate of speakers who stutter is decreased (Johnson & Rosen, 1937; Perkins, Kent & Curlee, 1991; Starkweather, 1985; Wingate, 1976); moreover, increasing the global speech rate seems to have the opposite effects on the frequency of disfluencies (Bloodstein, 1987;

Johnson & Rosen, 1937). It has been argued that speech rate could affect stuttering in one of two ways, i.e. severe cases of stuttering could: 1) exhibit slower rates of speech or (2) produce speech at too fast rates, thus causing stuttering to occur. There is support from empirical data to support the latter hypothesis, for example it has been found that people who have a severe stutter do in fact, exhibit fast speech rates (Andrade, Cervone & Sasssi, 2003; Arcuri, Osborn, Schiefer & Chiair, 2009). Local rate change, where the rate of the speech is changed at the single word or phrase level, has also been shown to reduce stuttering. For example if a person who stutters can slow down the rate of speech in the local region of a single uttered disfluency, this could be sufficient to prevent plans not being ready. Empirical support for this comes from Howell, Au-Yeung and Pilgrim (1999). They reported that tone units (metrical phrases) spoken rapidly, involving more than five syllables per second, were more likely to be stuttered than those that contained four and five syllables per second, and there were even fewer stutters for those that contained fewer than four syllables per second.

3) Articulatory variability and coordination (two more motor influences).

In fluent speech, speakers can accomplish accurate control of the articulators rapidly, thus minimising disfluencies. This ability is affected in people with speech-motor difficulties such as stuttering (Howell, 2010a). There has been much interest in developing indices that could quantify the amount of variation that is considered natural by testing fluent adults.

Moreover, these indices have been used to assess sensitivity in demanding tasks and across speaker groups with different levels of fluency. Variability, however, should not be confused with coordination, where gestures that are related to speech production have to be precisely timed with respect to each other, such as the velar opening and tongue contact for the phoneme /n/). Here, disfluency occurs when timing of the movements are not synchronized in production.

1.4.3. Models that propose an interaction between motor and language factors.

EXPLAN theory

EXPLAN theory (Howell, 2010a) argues that fluent speech control involves appropriate handling of two separate processes, language and motor. EXPLAN is the symbolic name that incorporates each process: - PLAN represents language control that takes different amounts of time to generate a word depending on complexity; EX represents the motor function that takes output from the language planning system and converts the representation to vocalized output. EXPLAN divides the symptoms of stuttering into two separate classes: Stallings and Advancings. Stallings include phrase repetitions, whole-word repetitions and pauses and are not specific symptoms of stuttered speech as they seem to occur in fluent speech too (Howell, 2007a; 2010a; 2010b). Advancings include part-word repetitions, broken words and prolongations and these are characteristics of stuttered speech as they occur more frequently in stuttered than in fluent speech (Howell, 2007a;

2010a). The attribution of whole-word repetitions to the stalling class separates them from true stutters and is consistent with Riley's (1994) procedure that excludes whole-word repetitions as stutters.

EXPLAN's connection to other theories

In the previous sections, separate Language and Motor theories were discussed in relation to stuttering. EXPLAN argues that language and motor factors connect together and that both factors are important in explaining stuttering symptoms. EXPLAN does not assume that language and/or motor factors are error prone in people who stutter, but that the two have to be timed together precisely at the interface. Levelt's theory (1989) used a perceptual process that consisted of a feedback loop linking motor output to linguistic conceptualization, that is, feedback information about the accuracy of production was sent via perception to the language mechanism. Since errors are not responsible for how speech is controlled according to EXPLAN, such a feedback process is not required here.

EXPLAN argues for the interdependency of Language and Motor mechanisms

As explained above, PLAN in EXPLAN stands for the symbolic generation of the language, which can take different amounts of time to generate depending on how complex the representation is. EX is the motor mechanism that derives vocal output. Similar to PLAN, EX requires time and this would depend on the complexity of sections of speech required to output. As with Levelt's theory, EXPLAN maintains that language and motor factors are independent, but EXPLAN argues that fluency depends on the way these two independent mechanisms interface. The literature has reported that people who stutter have various language deficits, however, people who have other language deficits do not stutter. Likewise, some people who stutter have various motor deficits, but other people with the same motor deficits do not stutter. These observations suggest that a language or motor deficit per se does not lead to stuttering and could mean that the interdependency of the language and motor deficits is implicated in people who stutter. It is unlikely that either mechanism is solely responsible for stuttering. EXPLAN embodies a multifactor account of stuttering, and although Max, Guenther, Gracco, Ghosh and Wallace (2004) argue that such models only exist because evidence attributing stuttering to either the language or motor factors is lacking, and not because of supporting evidence for a multifactor model. This point does not apply to EXPLAN where there is evidence for language and motor factors jointly determining stuttering.

EXPLAN is distinct from many other theories. For example, EXPLAN assumes that stuttering is a result of a timing issue that occurs with the need to handle accurate language plans that are predominantly error-free, but that are not yet complete; this stuttering is not a result of linguistic errors during language processing. If errors were frequent and covert repairs were an implication of that, then some sort of monitoring would be required to account for speech control. However, the suggestion that language could indeed be monitored for errors in real time on a continuous basis, as suggested by Levelt (1989), is problematic. Within EXPLAN, errors do not govern speech control and therefore information such as feedback and perceptual monitoring, which are sent to the linguistic system are considered unnecessary (Howell, 2010). Nonetheless, a simpler feedforward system is included within EXPLAN and provides answers to the issues raised about auditory feedback by Borden (1979). Borden argued auditory feedback monitoring would not allow speech to proceed rapidly, because of the amount of time required to process the feedback to obtain a full language representation. EXPLAN proposes that an efferent copy of the speech is made at the time it is executed and the actual execution form is obtained from the auditory signal. The two are subtracted to determine if they are identical (in which case they cancel each other. Subtraction takes minimal time and would not slow speech (Howell, 2010). Furthermore, it was suggested that the speech control mechanism must be able to operate without receiving veridical information about speech output because bone-conducted sound masks the sound of the voice. Hence, the auditory input channel does not provide information about how speech is articulated but rather operates as a serial input to the timekeeper. Here emphasis is on the speech activity and making sure there is no disruption to the timing pattern. To produce the sounds required, the speaker does not retrieve any information about the location of articulators, thus, whether the speaker’s speech is veridical or not is not relevant.

When compared to CRH, EXPLAN does not differ very much on the linguistic components, for example they both agree on the existence of a linguistic processing hierarchy and that syntactic constituents can affect fluency. CRH has retained Levelt’s theory’s argument that disfluencies can occur at any level of the linguistic hierarchy, however, EXPLAN argues that the prosodic form in particular (EXPLAN’s phonological words) is linked to disfluency. It is argued that any factors that affect the timing of language planning (usage factors or any level of the linguistic hierarchy) can have an impact on stuttering. That said, research supports the importance of phonological words whereas the evidence that other levels of the linguistic hierarchy affect stuttering is not strong (Howell, 2010). Phonological words can be described at two levels: 1) Usage terms, the usage properties of the components in phonological words result in the levels of them varying in complexity and 2) linguistic terms, where content word and a function word are the salient properties of phonological words. Howell (2010) argues that the evidence showing that syntax is important is not compelling. Thus at present, EXPLAN favors linguistic factors other than syntax. Although, Bernstein Ratner (1997) has argued that children who stuttered displayed syntactic deficits, evidence from Reilly et al (2009 & 2010) suggests that children who stutter have precocious language abilities by the age of 2 years.

In CRH, motor processes receive language plans and proceed to implement them as motor programs for output (Howell, 2010). However, it is argued that motor processes do not contribute to whether a word is output as fluent or not. In EXPLAN, motor processes are involved in fluency failures at the point where motor and language factors combine temporally. In EXPLAN, If rate is too high prior to a word (motor factor) that is difficult to plan (linguistic factor), then there will be a failure in fluency as the plan is not ready in time for execution and output.

1.4.4. Neuro-computational model of speech production: The Directions into Velocities of Articulators (DIVA) model

The DIVA model (see figure 2) proposes that the neural networks within cortical structures and their interactions are responsible for the execution and the control of articulation (Guenther, 2008; Guenther & Vladusich, 2012). The model is made up of specific components that are associated with specific areas of the brain. This has supported analyses and associations during simulations of brain activity of humans during speech in neuroimaging studies; for example, some studies have found that DIVA accounts for the activation found during the production of simple speech utterances (Guenther et al., 2006; Tourville et al., 2008). The DIVA model comprises of a control circuit for somatosensory targets, it also argues that an auditory target reference frame guides the neural network articulation (Guenther et al., 1999).

In this model, it is proposed that an interaction at various levels occur between the feedforward and feedback subsystems in order to control articulation. The feedback control circuit results in the production of motor commands. It is argued that during normal rate of speech, humans can produce up 5 syllables in one second (Tsao & Weismer, 1997) the rate at which normal speech is produced is arguably outpacing the sensory feedback (Guenther, 2006). For this reason, the feedforward circuit is also incorporated within the model together with feedback control, they account for all motor output. Guenther et al. (2009) argue that the process, from the generation of the command to the execution of the motor output, takes on average 40 milliseconds. The feedforward circuit can offset the delay to speech processing this would cause by computing targets before the onset of the movement.

The DIVA model argues that the left ventral premotor cortex contains the speech sound map, which is suggested to be responsible for all articulatory executions (Guenther, 2006). This speech sound map is responsible for the construction of speech as it contains the representational units of the target reference frame, which are proposed to be either a word, a syllable or a phoneme (Guenther, 2006, Guenther & Vladusich, 2012). This is similar to Levelt et al’s (1999) mental syllabary.

The feedforward system contains two routes for commands, first the direct route, which allows commands from the speech sound map to go through to the articulatory velocity and represents these maps in the primary motor cortex. This is then translated into motor movements (Beal, Quraan, Cheyne, Taylor, Gracco & Luc, 2011). The second route transmits information to the articulatory velocity from the speech sound map. This is then represented in the subcortical or the cerebellar level (Beal et al., 2011). However, the DIVA model does not account for the interaction between subcortical motor pathways and supplementary motor pathways (Beal et al, 2011).

Information from the auditory and somatosensory systems are both incorporated within the feedback control system. The perception of the speaker's own voice within the auditory cortex, affects the auditory feedback loop and the signals transmitted from subcortical auditory pathways. Expected sound signals are represented in the auditory state map and the information is shared with the auditory error map. Both of these maps are represented in the superior temporal cortex within the DIVA model (Guenther, 2006). Somatosensory information is shared in the same manner within the DIVA model, for example, somatosensory information is sent to the somatosensory state map and the information is shared with somatosensory error map. Both of these maps are represented in the inferior parietal cortex

The somatosensory error is the difference between the status of the vocal tract and the target oral sensation. This difference is then mapped onto the motor cortex. Articulatory velocity and position maps are adjusted according to the information retrieved from a comparison between the feedback signal and the 'efference' copy. This process happens in the auditory and somatosensory feedback systems (Beal et al., 2011, Tourville et al., 2008). Corrections to the signal are then made by measuring the difference against the internal feedforward model (Tourville et al., 2008). Some researchers have argued that the efference copy is inhibitory and this is what causes speech induced auditory suppression (Houde et al., 2002).

In terms of stuttering, many researchers have argued that stuttering may be due to an abnormal sensory feedback system where people who stutter differ from people who do not stutter by relying too heavily on sensory feedback (Tourville et al., 2008; van Lieshout, Peters, Starkweather, & Hulstijn, 1993). On the other hand, some researchers have argued that people who stutter may in fact benefit from the reliance on sensory feedback (Max, 2004; Max et al., 2004; van Lieshout, Hulstijn, & Peters, 1996a, 1996b).

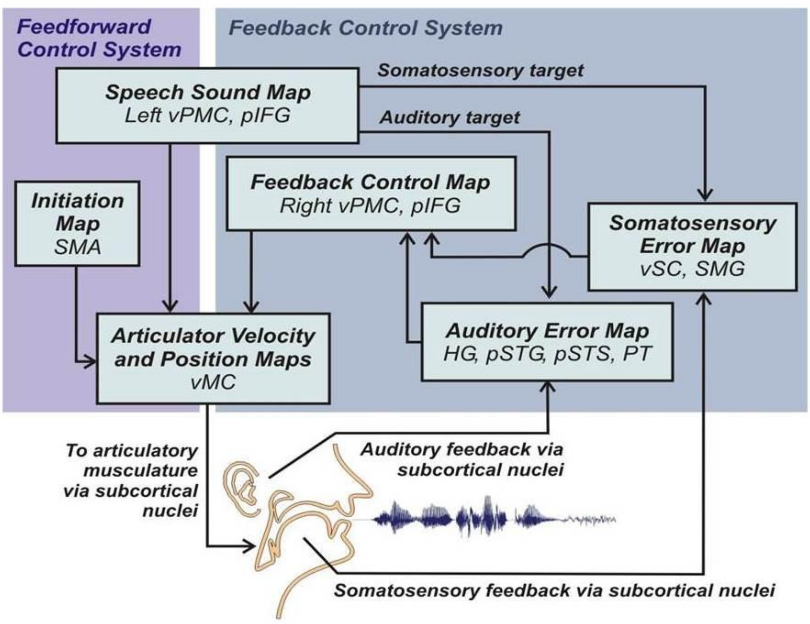

Figure 2. **The DIVA model: The cortical structures and their associated connections**[1]

[1] The figure is taken from Guenther, Ghosh and Tourville, (2006)

1.5. Assessment of Stuttering

This section focusses on research that provides evidence for the assessment of children who stutter (CWS) and those with EAL. Whilst its primary focus is on stuttering, some reasons for assessing other speech difficulties (including hearing) are considered. Since the symptom set from SSI-3 is used during screening, and given the interest in children with EAL, the next section considers assessment of stuttering in languages other than English.

1.5.1. Assessment of Children Who Stutter (CWS).

In order to understand the work that follows, further details about Riley's Stuttering Severity Instrument are needed (procedural features in particular). It is the only statistically reliable estimator for predicting which children will be fluent and which stutter (the essence of screening), which it does with high sensitivity and specificity. Its use in a school population is explored in this thesis. Most published peer reviewed research on stuttering in general characterizes participants by reporting scores on the SSI-3 instrument of Riley (1994). The assessment was updated and renamed as the SSI-4 in 2009. Riley (2009) has ensured that the SSI-4 is downward compatible with SSI-3, so that the derived scores are not altered from one version to the next. The most significant change in SSI-4 was inclusion of computer count procedures for syllables, stutters and their duration. These were previously obtained manually in SSI-3. Since the computer technique was not used when the SSI-3 was standardized and as it produces different results from manual counts (Jani et al., 2013), the computer technique cannot be used until re-standardization has been performed. Therefore, in this research project, the manual counting techniques employed in SSI-3 were used. The following sections describe how SSI-3 is administered manually.

1.5.2. Administration of SSI-3.

The SSI-3 can be used with adults as well as children. The assessment procedure with adults is that speech is elicited and a reading sample is taken using age-appropriate materials provided with the instrument. When children cannot read, age appropriate pictures are supplied that are used to prompt the child to discuss some of the themes that are covered. Three assessments are made for each child/adult (one at the time of recording and the others offline). The following procedure for assessment was followed throughout this thesis and follows Riley (1994):

1) A score of the frequency of stuttering is obtained in the first analysis: The frequency score is obtained by first counting the total number of syllables in the speech sample then marking each stuttered event. Precise instructions are given on what is counted as stutters. For example, “you are no-no-no-not listening”, has 6 syllables (the beginning of each syllable is marked with an underline) and there is one stuttered event, which includes all the part-word repetitions on the word “not”. Such counts are done on the whole speech sample and according to Riley (1994) the total number of syllables should not be less than 200 (Todd et al., 2014; Chapter 3). After syllables and stuttered events have been counted, the percentage of stuttered

syllables is calculated by dividing the number of stuttered events by the total number of syllables and multiplying by 100. In the example above it would be (1/6) x100= 16.7%. The task scores associated with this percentage are obtained from Table 1. The task score is combined with two other scores (described later) to give the overall SSI-3 score.

Table 1. Conversion of obtained % SS (stuttered syllables) to task scores for non-readers (left) and readers (right).

Nonreaders		**Readers**			
Speaking		Speaking		Reading	
% SS	Task Score	% SS	Task Score	% SS	Task Score
1	4	1	2	1	2
2	6	2	3		
3	8	3	4	2	4
4-5	10	4-5	5	3-4	5
6-7	12	6-7	6	5-7	6
8-11	14	8-11	7	8-12	7
12-21	16	12-21	8	13-20	8
22 & up	18	22 & up	9	21 & up	9

Note. Table adapted from Riley (1994)

2) A duration score is obtained in the second analysis: the duration score is the average of the three longest stutters. For example, if a child's longest three stutters were 2.1, 2.3, and 3.1 s, this would give an average of 2.5 s. The duration score is converted to a task score using Table 2.When there are two samples (i.e. when a child can read and provides a read sample as well as a spontaneous sample) the three longest stutters irrespective of sample are used.

Table 2. Conversion of duration score associated to a task score.

Category	Time in seconds	Associated SSI-3
Fleeting	0.1-0.4	2
One-half second	0.5-0.9	4
One full second	1.0-1.9	6
2 seconds	2.0-2.9	8
3-5 seconds	3.0-4.9	10
5-9 seconds	5.0-9.9	12
10-29 seconds	10.0-29.9	14
30-59 seconds	30.0-59.9	16
60 seconds +	60.0 and up	18

3) A score for physical concomitants is obtained at the time of recording. This can be obtained either from watching a videotape of the interview or during the interview itself (the latter was the procedure adopted in the present study). Four categories of physical concomitants are assessed: 1) Distracting sounds: constant clearing of throat, whistling noises, frequent noisy swallowing, blowing and clicking sounds. 2) Facial grimaces: any abnormal movement or tension of the face muscles. Examples include pressing of the lips, pursing of the lips, protruding the tongue, uncontrolled eye, or jaw movements. 3) Head movements: turning the head to avoid eye contact, scanning the room, looking down at the floor or up at the ceiling. 4) Movements of the extremities: movements of other parts of the body. For example, kicking the feet, swinging of hands or arms, fidgeting and shifting around in the chair. Each of these is assessed on a five point scale (Table 3) and the scores for each of these categories are summed to give the overall score.

Table 3. Five scale point to assess physical concomitants for the four aspects described above.

Score	Description for this score point
0	None
1	Not noticeable unless looking
2	Barely noticeable to casual
3	Distracting
4	Very distracting
5	Severe and painful looking

4) The task scores for %SS and duration and the raw score for physical concomitants are added to obtain the total overall raw SSI-3 score. This raw score value can be associated with a severity descriptor (given in Table 4). Only the raw scores are used in this thesis.

Table 4. The total overall score and its associated severity level.

Total overall SSI-	Percentile	Severity
0-8	1-4	Very mild
9-10	5-11	
11-12	12-23	Mild
13-16	24-40	
17-23	41-60	Moderate
24-26	61-77	
27-28	78-88	Severe
29-30	89-95	
32 and up	96-99	Very severe

The example employed earlier can be used to illustrate how a raw SSI-3 score and severity descriptor are obtained. There was 16.7% SS, which gives a frequency score of 16 from Table 1. The duration of 2.5 seconds was the average of the three longest stuttered events, which gives a task score of 8 using Table 2.). Finally, if this child scored 0 for physical concomitants, the three quantities are added to give the overall raw SSI-3 score: 14+8+0= 22. A raw score of 22 has the descriptor “moderate” in Table 4.

1.5.3. Stuttering and additional languages.

In the EU it is estimated that around 10% of all school-aged children speak a first language that is not the language that the majority of citizens in that country speak (Siegel, Iozzi, & Surian, 2009). Stuttering can occur in any language and culture and, due to increased globalization, the number of children speaking more than one language is increasing. Early studies showed that bilingual speakers had higher rates of stuttering than monolingual speakers (Stern, 1948; Travis, Johnson & Shover, 1937). However, later studies do not show the same trends and tend to agree that the later the second language was acquired the less likely a child is to stutter (Au-Yeung, Howell, Davis, Charles & Sackin, 2000; Howell, Davis & Williams, 2009). Furthermore, it has been recommended that a second language is introduced after the first language was acquired completely so as not to aggravate problems in children who may be vulnerable to stuttering (Howell et al., 2009).

As mentioned above, multilingualism (which includes bilingualism) is widely encountered; some children may be exposed to a second language at the same time as the first (termed early bilingualism) while others may be exposed to the second language once in school (termed second language or later bilingualism). Bilingualism is best described as a continuum where efficacy in each language can vary across expression and auditory versus written comprehension (Roberts & Shenker, 2007). Working with bilingual children poses many questions to educators and clinicians. For example, can the examiner identify stuttering in an unfamiliar language; does the multilingual child's linguistic proficiency have an effect on stuttering; does treatment require different approaches for bilingual children; and finally, can the stuttering in a bilingual child be treated effectively? A lack of research in this area means that these questions are far from being answered (Bernstein Ratner, 2004; Van Borsel, Maes & Foulon, 2001).

The challenges in investigating these issues include aspects such as stuttering severity not being constant across all the languages spoken by the individual, which is a particular problem in children (Roberts, 2002; Van Borsel & Britto Pereiria, 2005). Uncertainty remains with respect to the quality of an assessor's judgments when assessing a child who speaks an unfamiliar language. A possible consequence of this is that most research in this area has focussed on Indo-European languages (Bernstein Ratner & Benitez, 1985; Jankelowitz & Bortz, 1996; Nwokah, 1988). Some studies have shown that a determining factor in assessing stuttering in an unfamiliar language and the accuracy in doing so depends on how close the language is to the assessor's native language (Van Borsel, Leahy & Britto Pereiria, 2008).

Experienced clinicians who are not familiar with a target language have been shown to be able to report reliably stuttering events in 3-5 year old Icelandic children (Einarsdottir & Ingham, 2009). In this study, the speech samples were divided into 5-second intervals for the clinicians, making it easier but perhaps more time-consuming to assess them (Einarsdottir & Ingham, 2009). Nonetheless, these findings can be useful for clinicians who experience such situations. Another way to assess judgement reliability is to form agreement about ambiguous and unambiguous stuttering events with a person who is familiar with the language but who is not necessarily a clinician. Furthermore, seeking validation from parents or others who are familiar with the language of the child can help to determine whether these disfluencies are due to stuttering or a lack of proficiency in the child's newly adopted language (Shenker, 2011).

Due to the complexities in learning a language to a proficient level (i.e. proficiency in vocabulary, pronunciations and understanding of grammar), children who experience two or more languages may develop uneven proficiency (Shenker, 2011). This is why it is important to consider that proficiency may offer an explanation for the disfluency in the child's speech. It has also been argued that children who were exposed to two or more languages from birth can still show delay in linguistic properties of both languages and this can lead to misdiagnosis of language delay or even over-identification of stuttering (Bedore & Pena, 2008; Shenker, 2011).

It must also be remembered that languages differ and linguistic complexity varies between languages. Also there are different milestones in acquisition in different languages (e.g. age on tone acquisition represent milestones for Mandarin, but not English). These make it more difficult to identify with certainty the dysfluent events in the unfamiliar language

(Thordardottir, 2006). Also, some symptoms may be signs of word-finding difficulty (e.g. the whole-word repetitions discussed above) that are considered by some to be symptoms of stuttering as well. This has been reported to lead to different forms of stuttering being reported in the dominant and non-dominant languages (Howell, Ruffle, Fernández-Zúñiga, Gutiérrez, Fernández, O'Brien, Tarasco, Vallejo-Gomez, & Au-Yeung, 2004). It has also been shown that children can compromise the structural integrity of one language so as to keep the more dominant language. For example, applying the grammatical rules from a dominant language to the second language (e.g. a French child may say 'the dress long' in English rather than 'the long dress'). Shenker (2011) argued that this should not be seen as a grammatical error.

Another example, not specifically about stuttering, is code switching or code mixing, where the vocabulary of two or more languages are used interchangeably (Genesee, 2001). Again, it is argued that this is not a dysfluent event (avoidance of stuttering in a language) or a word retrieval problem. In fact, many studies suggest that this is a natural part of the learning process for a bilingual child (Pert & Letts, 2006; Shenker, 2011). Shenker (2011) provided two examples of how to improve proficiency in identifying problems in a bilingual child's speech: 1) Using interpreters who are familiar with the cultural aspects of the language as well as the linguistic properties so that they can confirm language proficiency of the child and validate any dysfluent events. However, this is costly and there is no guarantee the interpreter will be able to help with any dysfluent utterances if they are not clinically trained. 2) A detailed language history report which includes i) when each language was acquired, ii) first exposure to each language, iii) all other concerns regarding speech and language, and iv) where and how is each language used. Furthermore, Roberts and Shenker (2007) argued that this should be used alongside parent reports of daily experiences about the child, and that the SSI-3 by Riley (1994) is not a valid measure with bilingual speakers due to lack of evidence in cross-cultural differences in linguistic properties. They argued that a more valid measure would be to obtain speech samples and conduct assessments in language proficiency and stuttering events in both languages.

Further difficulties are faced when dealing with children migrating with parents. In this context the child may be introduced to a second language only after starting school. At this point it becomes more difficult to differentiate between early onset and late onset of stuttering. These children, in addition to learning a second language, are also faced with dealing with psycho-social factors and cultural adaptation, which may add further complexities to clinical judgment. It has been suggested that children who are exposed to a second language at school level, and who are exposed to mixed linguistic inputs (i.e. different language in school and home) could be a factor leading to persistence of stuttering (Van Borsel et al., 2001).

When treating stuttering, it is important to take into account cultural differences in treatment. As suggested by Westby (1990), ethnographic interviews may be used to identify a family’s perception of stuttering and the treatment plan and delivery needs to take this into account. For example, Wong-Fillmore (1991) suggested that some parents believe that treating stuttering in the acquired language may result in loss of, or undermine ability in, the first language. This is deemed threatening as it is considered that the child will lose their identity or will not be able to communicate with family and/or extended family. As part of the present project, some discussion between the researcher and members of the Tamil community was undertaken. Different perceptions were discussed and taken on board during feedback. For example, a Tamil teaching assistant raised issues within the Tamil population in the school about socio-economic class and regional differences in interpretation of instructions as well as about proficiency in the Tamil language for the children aged 4-5 years old. This may explain differences within this population (i.e. why some Tamil children experience more difficulty in learning a new language, or why some exhibit more disfluencies than others).

The concerns the issues that have been discussed raise for the present study are about assessment in children with EAL. As mentioned above, how can a conclusive decision be made about which form of assessment is most effective and reliable in screening for disfluency in children with EAL? Another issue was whether other symptoms, such as those related to phonological delay, add to the complexity of finding an appropriate method to screen children with EAL. Furthermore, Van Borsel et al. (2001) investigated the efficacy of assessing a speech sample which is not in the assessor’s mother tongue, however, in our samples the children with EAL speak English. How effective are assessors in judging if a child is dysfluent when the child is attempting to speak a language they have not mastered yet? It is important to investigate this and find solutions to this problem, as it is vital in multicultural societies to be able to find an effective way to assess children who do not speak the national language as their first language. What also needs to be assessed, are symptoms related to hearing problems. Van Borsel details processes which involve disfluencies in children with EAL but does not discuss speech issues related to children with EAL or, indeed, speak English only who also have hearing problems.

Therefore, a screen test needs to be versatile enough to use with children who speak English as a first language but also those who have English as an additional language. This can become an issue when children with EAL are also known to display speech symptoms that may be interpreted as disfluencies, when they actually are not.

1.6. Summary

In this chapter literature was reviewed that provided ideas on how a screening assessment should be made. Some general issues that arose were also discussed, e.g. children with EAL and their assessment. In the next chapter the results of a survey are reported. This was conducted with teachers and SLTs to assess their needs and opinions concerning a comprehensive way of assessing children in schools and the discrepancies between what each party perceives to be most applicable and practical. The discrepancies between these services were raised, but not evaluated, in the Bercow report (2008). In this chapter, models and theories of stuttering were discussed, this was important to address as the tool, which is used as the screening tool in later chapters of this thesis, is the Stuttering Severity Instrument (SSI-3) (Riley, 1994). However, one of the aims of this research is to see if the SSI-3 can also pick up children with speech problems other than those who stutter. This is important if the tool is to act as a general speech difficulty screen for children. These children may display various speech symptoms and not necessarily those associated with stuttering. For this reason, this thesis also explores whether extending the speech symptom set enhances the performance of the SSI-3 as a screening tool. In chapter six, a systematic review is reported that explores the symptoms which are generally associated with speech difficulty in children. This was the basis of the extended symptom set in the analyses that were then conducted to see whether extending the symptom set improves predictions when distinguishing fluent children from those with a range of speech problems.

Chapter 2

A survey of UK Teachers' and Speech and Language Therapists' Joint Roles with Respect to Reception Class Children's Speech Language Communication Needs[2]

2.1. Introduction

It is generally agreed that many aspects of educational and behavioural experience (Conti-Ramsden & Simkin, 2012; Johnson, Beitchman & Brownlie, 2010) and subsequent employment opportunities (Clegg, Hollis, Mawhood, & Rutter, 2005) are affected by a child's speech and language communication needs (SLCN). Consequently, schools need to be vigilant about SLCN (Bishop & McDonald, 2009), especially at early ages, in case speech and language intervention is required (Skeat, Wake, Ukoumunne, Eadie, Bretherton & Reilly, 2014).

Teachers are an important source of referrals to SLTs (Dockrell & Lindsay, 2001). Teachers consider whether factors such as anxiety and bullying have any effect on children's educational and behavioural outcomes as well as their language and communication skills when considering children's SLCN (Antoniazzi, Snow & Dickson-Swift, 2010). The SLCN of the growing number of children who have EAL is an increasingly important topic

[2] A version of this chapter is to appear as Mirawdeli, A, Dockrell, J. & Howell, P. (revision submitted). A survey of UK Teachers' and Speech and Language Therapists' joint roles with respect to reception class children's Expressive Speech Difficulties. *International Journal of Language & Communication Disorders.*

for both teachers (Murphy, 2015, Strand, 2015) and SLTs (McLeod, Verdon & Bowen, 2013). It is likely that referrals of all children giving cause for concern would be improved if SLTs were informed of the process teachers use to assess a child. SLTs may also provide advice to schools on how children with SLCN could be identified (Marshall, Ralph & Palmer, 2010; Wium & Louw, 2015). However, referral by schools is one-directional at present and SLTs have little input to this process (Wium & Louw, 2015). Feedback post-intervention is not necessarily two-way but could be improved (Bercow, 2008). At present, a general framework for such discussions is lacking partly because there is no standard procedure for schools to follow (Broomfield & Dodd, 2004a) and teachers select children subjectively (Mirawdeli, 2015).

The present chapter reports the results of a survey concerning procedures for identifying children with SLCN in general and speech difficulty in particular. The primary goal was to establish whether schools and SLTs see a need for a scheme to assess speech difficulty in the early school years. In the next section, reasons are given as to why speech difficulties at school entry were singled out for attention. Then information about SLCN and speech difficulty, available from a systematic review of the literature on screening in pre-school populations, is given (Nelson, Nygren, Walker & Panoscha, 2006). The findings of the review were updated and the results used to design the current survey. In addition, findings from structured interviews that were conducted with teachers and SLTs were used in the design of the survey. The survey addressed what form procedures for identifying speech difficulty for all children should take and how feedback received from, and provided to, schools should proceed. This shows that there is a need for the screening programme in schools such as that developed in this chapter (described in later chapters of his thesis) and the form the procedure needs to take to meet the requirements of schools and SLTs.

2.1.1. SLCN and Speech Difficulty When Children Start School

Whilst the importance of considering all SLCN in the early years in school is generally recognized, there has been little consideration about whether or not types of SLCN have different requirements. Less attention has been paid to speech difficulty compared to other SLCN (McLeod & Goldstein, 2012) although some initiatives are currently attempting to redress this imbalance (e.g. McLeod et al., 2013). Two arguments are given that suggest speech difficulties should be assessed as early as possible (at, or before, school entry).

First, the onset of speech difficulties (Dodd et al., 2002) is earlier than that for language difficulties (Conti-Ramsden, Botting & Faragher, 2001). Speech difficulties often start before the age of school entry although cases arise throughout childhood (Dodd et al., 2002). Thus, surveillance for speech difficulties needs to start early and continue throughout childhood. Health visitors pick up speech difficulties, *inter alia*, in the pre-school period (Nelson et al., 2006), whilst teachers monitor for speech difficulties when children start school. Teachers need guidance on how to identify speech difficulties that is appropriate for children entering reception classes and at subsequent intervals. In contrast, literacy difficulties are evident later in primary school years, and progress and improvement in identifying this has been more rapid than is the case with speech difficulty (Kohnert, Yim, Nett, Kan & Duran, 2005). Second, early intervention is essential (Bishop, Reilly & Tomblin, 2014) as there may be a critical period when speech difficulties can be treated most effectively (Hakuta, Bialystok & Wiley, 2003; Lenneberg, 1967; Penfield. & Roberts, 1959).

Taken in conjunction with differences in age of onset, the need for early intervention suggests that speech difficulties may need to be identified at younger ages than is the case with other SLCN (Reilly et al., 2010).

2.1.2. Review of work on screening pre-schoolers for speech difficulty.

Nelson et al. (2006) conducted a systematic review to establish what work had been conducted on screening children for SLCN covering the age range up to five years. This provided a useful starting point for information about how speech difficulty might be identified when children enter schools. They reviewed research on how primary health care workers, rather than teachers, screen children for SLCN. They did not single out speech difficulty nor consider how children with EAL should be assessed. Furthermore parts of the review that addressed work on intervention are not relevant. Consequently, only the first four issues that Nelson et al. (2006) considered are reported here. These were: 1) The effectiveness and feasibility of screening children aged five years and younger for speech and language delay in primary care settings; 2) Whether screening evaluations work; 3) Adverse effects of screening; and 4) The role of enhanced surveillance in primary care. The review revealed that there was little literature on screening. For instance, Nelson et al.'s (2006) review did not locate any literature at all on issues 1, 3 and 4.

To check for subsequent developments, a systematic review of the literature was conducted on the PSYCINFO database. Searches were conducted for publications that appeared from the date of Nelson et al.'s systematic review to the present and were not restricted to work on pre-schoolers. Comprehensive sets of search terms were used for: types of speech difficulty; general features associated with non-fluent speech; and behaviours in schools that could be affected when a child has SLCN. The search terms were applied to any field in articles (not just keywords) and the search was limited to peer-reviewed journal articles in English. Little pertinent literature was returned, as Nelson et al. (2006) found. Inspection of the articles returned by the search showed that they were mainly experimental studies intended to inform readers about aetiology of a disorder. There were two exceptions - studies by Archibald, Joanisse and Shepherd (2008) and Barbarin (2007).

A second approach was taken – identifying any peer-reviewed articles on UK PubMed that cited Nelson et al. (2006). Any returns were potentially relevant themselves and they may have reviewed other work which had been missed in the search. This search returned 16 articles, nine of which were relevant to current concerns (articles by Ek et al., 2012; Northam et al., 2012; Peyre et al., 2014; Ribeiro et al., 2011; Ruandaragon et al., 2009; Samelli et al., 2014; Sim et al., 2013; Skarżyński et al., 2012; and articles by Wake et al., 2011 and 2012). The two literature searches indicated that the situation concerning paucity of evidence on screening young children has improved little since 2006 and underlines the need for the current survey. The opinions and needs of the teachers are considered important as they are the first point of contact a child with speech difficulties has. Therefore, it was considered important to know what teachers' needs are. Since the teachers make the most referrals, it was also considered important to check for any disparity between the two ends of the service; the schools (as represented by the teachers) and the SLTs. None of the 16 papers looked at these issues, even though the importance of this issue was highlighted in the Bercow report (2008).

2.1.3. Design of survey for teachers and SLTs.

The review work confirmed that information needs to be solicited from teachers and SLTs on how SLCN and speech difficulties are identified and acted upon in schools. The next step toward setting up the survey was to conduct structured interviews with three reception class teachers and five SLTs (Lee, 2014). Information based on Nelson et al.'s (2006) four issues mentioned above was elicited. The interviews extended beyond Nelson et al.

(2006) as they addressed how speech difficulty is assessed by teachers. Care was taken to ensure attention was paid to how children with EAL are assessed for SLCN, given the need for equitable assessment (Bercow, 2008). It emerged that teachers often employed practices for identifying speech difficulty that differed from those which SLTs would recommend, due to practical considerations. For example, procedures need to be short for use in schools so that they could be repeated periodically. Also, it was clear that some speech symptoms that may indicate speech difficulty were discounted by teachers (e.g. whole-word repetitions, WWR, and pauses) because children with EAL would be expected to have word-finding difficulties and exhibit these symptoms (Abdullah, 2012; Bada, 2010). These features (particularly WWR) are considered to be signs of speech difficulty by some SLTs. Consequently, it is important to be aware what the position in schools is on this matter since SLTs would not necessarily ignore WWR and pauses when assessing children's speech. The interviews also showed that teachers would prefer children with EAL to be assessed in English rather than their alternative language. Teachers wanted to identify whether or not a child was fluent, not whether a child has one or another type of speech difficulty. This is because they wish to refer any children who give cause for concern, and rely on SLTs to perform more detailed assessment for the type of SLCN. The teachers indicated that any assessment needed to be short and based on an audio (not video) sample. As many of these decisions are contentious for

SLTs, a separate section was included in the current survey to focus on, and verify, these impressions with a large sample of respondents.

2.1.4. The present study.

Teachers and SLTs have a common concern with children's welfare and there are some areas for which both sides need to be apprised of current practices. Information is lacking about issues that Nelson et al. (2006) raised concerning identification of children with SLCN, whether speech difficulties should be distinguished from language difficulties and, assuming a speech assessment is desirable, how the needs of children with EAL can be incorporated. The survey was conducted to seek information from teachers and SLTs regarding how speech difficulties are identified and acted on as a basis for recommendations for improvement that are agreeable to all parties.

The survey was designed to elicit information about the four topics raised by Nelson et al. (2006) and two additional areas of concern raised in the structured interviews. First, information was sought from teachers about identification and referral of children, and SLTs' views on how such decisions are made (Nelson et al., 2006). Better understanding of how teachers make, and how SLTs view, these decisions are timely as policy about SLCN is a current focus of debate (Bishop, 2014; Reilly, Tomblin, Law, McKean, Mensah, Morgan, Goldfield, Nicholson & Wake, 2014).

Second, what resources are available for identifying SLCN in general and speech difficulty in particular? As argued earlier, a one-size-fits-all solution to identifying children with all types of SLCN may not be appropriate. Therefore, information was sought about the assessment of speech difficulties, in particular, and whether they should be considered separately from other SLCN. The third topic was whether teachers and SLTs were aware of potential risks involved when identifying SLCN and whether EAL status contributed to such risks.

The fourth topic derived, in part, from the focus on speech difficulty. The goal was to obtain information on a procedure for identifying speech difficulty in children entering reception classes. This was extended to include provision for children with EAL.

The two additional topics were based on the results of the structured interview. Firstly, specific information was solicited about a speech assessment procedure for use in schools. In addition to procedural features, some aspects of the sound patterns considered to be indications of speech difficulty were examined (i.e. the symptoms teachers listen for when attempting to detect speech difficulty). A list of the symptoms used throughout this thesis is given in appendix I. The final section sought information on provision for SLCN in general, and speech in particular, for children with EAL in comparison with provision for children whose first language is English

2.2. Method

2.2.1. Procedure.

An online survey was set up that applied to reception class teachers and SLTs (http://www.qualtrics.com/). Qualtrics software has display logic so that presentation of material can be based on answers given to previous questions. This was used in the present survey when questions were only appropriate for one profession, and where follow-up questions were contingent on a given answer. As an example of the latter, there was a question whether the schools had resources for assessing speech difficulty and only respondents who answered 'yes' were asked to provide details (text entry). Qualtrics was used to obtain both information about participants who filled in the questionnaire and the substantive information reported in the results section. Preliminary analysis used the internal report formatter in Qualtrics (http://www.qualtrics.com/). The data were downloaded into SPSS for further statistical analysis. Sign test analyses were performed to analyse the data. The total number of answers given in one category was compared to the total number of answers given in the same category by the others (i.e. the teachers or the SLTs).

2.2.2. Participants.

All participants were from the UK and were not paid. E-mail lists for both groups of professionals were used to identify and contact potential participants, and the people contacted were also asked to pass the link on to other appropriate parties. Thirty-five teachers and 35 SLTs completed the survey. As participation was voluntary, there was no control over who completed the survey. However, all participants provided background information about their professional status. This included details about gender, years in post, contact time with children with EAL (only for SLTs) and school type. This information is summarized separately for teachers and SLTs in Table 5. Details about the classes that the teachers and SLTs reported on are summarized in Table 6, again separately for teachers and SLTs.

Table 5. Details of the teachers (column two) and SLTs (column three) who participated.

	Teachers	SLTs	Response options allowed (where appropriate)
Number.	35	35	-
% female.	83%	80%	-
% indicating modal option for time in post in year.	10-20 years, 36%.	Less than 5 years, 56%.	0-5, 5+-10, 10+-20, or more than 20 years).
% of SLT's indicating modal option for time in contact with children with EAL.	-	0-20% of time, 49%.	0-20%, 20+-40%, 40+-60%, 60+-80%, 80+-100%.
% who worked in the modal type of school for participants who worked in schools.	Mainstream, 97%.	Mainstream, 91%.	Mainstream, special, private.

Note: Information sought is listed in the first column. For forced choice responses, the options available are given in column four.

Table 6. Details of the reception classes of schools in which respondents worked.

	Teachers	SLTs	Response options allowed
% respondents giving the modal option	20+-30, 54%	20+-30, 75%.	10-20, 20+-30 or more than 30.
Main social class of children	Working class, 76%	Working class, 52%.	Lower, middle, upper class.
% respondents and modal answer for referrals to SLTs	Up to 5%, 73%.	5+-10%, 50%.	0-5%, 5+-10%, 10+-15%, 15+-20%, more than 20%.
% children who have EAL with modal answer (only for SLTs)	-	0-20%, 53%.	0-20%, 20+-40%, 40+-60%, 60+-80% and 80+ to 100%.
% children with EAL in reception classes with modal answer (only for teachers)	60+-80%, 50%.	Not asked.	0-20%, 20+-40%, 40+-60%, 60+-80% and 80+ to 100%.

Note: Information reported on is given in column one. Responses are given separately for teachers (column two) and SLTs (column three). Column four gives the available response options for each row.in the second column, the percentage value indicates the percentage of respondents who gave the designated answer. For example, in the bottom row, 50% of teachers gave the answer 60+-80% and this was the modal answer.

2.3. Results

2.3.1. Communications between teachers and SLTs about children with speech difficulties subsequent to intervention.

The results on exchange of information (feedback for short) between teachers and SLTs concerning children with speech difficulties are summarized in Table 7. The numbers of teachers who received feedback from SLTs and the SLTs who gave feedback to teachers post intervention are given in row one. Post-intervention, significantly more teachers received feedback from SLTs ($p < .001$) than those who did not receive feedback, and a significant number of SLTs gave feedback to teachers ($p < .001$) compared to those SLTs who did not give feedback. Significance was assessed by two-tail Sign tests (a test which is used throughout this chapter). Here the sign test was used to determine whether the number of participants within one group who gave one answer (e.g. received feedback), differed from the number of respondents who gave the alternative answer (i.e. did not receive feedback). Additionally, as necessary, the Sign test was used to determine whether the respondents in one group (e.g. SLTs) differed from those who gave the same answer in the other group (e.g. teachers).

The numbers of teachers and SLTs who provided or received feedback, respectively, post-intervention are given in row two. The number of teachers who provided feedback to SLTs after intervention was not significantly higher than the number who indicated that they did not provide feedback ($p > .05$). Consistent with this, there was no significant difference concerning whether or not SLTs received feedback from schools after intervention ($p > .05$). Overall, both groups responded that SLTs provided feedback, but the number of teachers who subsequently provided feedback to SLTs after intervention was complete, was not significantly higher than those who did not provide feedback.

The responses in rows one and two of Table 7 were followed up only for those respondents who indicated that they gave or provided feedback to the other party. The type of feedback teachers sent was examined first (row three of Table 7). Teachers (column two) and SLTs (column three) chose one of five types of feedback (indicated in column four, row three of Table 7). The majority of teachers (22 of the 27) indicated that their feedback was about 'improvements in speech'. The number of teachers who gave this response was significantly higher than the number who gave one of the remaining responses (*p* <.001). Most of the SLTs who indicated that they received feedback from teachers post-intervention also indicated 'improvements in speech' (row three, column three of Table 7). However this response was not given by significantly more respondents than the number of respondents who gave one of the remaining five responses ($p > .05$).

Next, the type of feedback SLTs sent was examined (row four of Table 7). Respondents chose one of four types of feedback (indicated in column four, row three of Table 7). The majority of teachers who received feedback from SLTs indicated that this was 'improvements in speech' (row four, column two of Table 7). However, the number of teachers who gave this response was not significantly higher than for those who gave one of the remaining responses ($p > .05$). For the SLTs who provided feedback to schools post-intervention (row four) the majority chose 'all of these' (26 of the 33). This response was given by significantly more respondents than the number of respondents who gave one of the remaining four responses ($p < .001$).

To summarize, both groups of professionals indicated that they provided and received feedback about *improvements in children's speech*. The exception was that SLTs indicated that they provided a *full range of feedback*.

Details were next sought about feedback specifically concerning children with EAL post-intervention. These addressed whether feedback was received or provided (last two rows of Table 7). Row five shows that significantly more teachers received no feedback about children with EAL from SLTs compared to those who did receive feedback from SLTs ($p < .001$). However, the number of SLTs who received such feedback from teachers was not significantly higher than those who did not receive feedback ($p > .05$). Row six shows that the number of teachers who indicated that they gave feedback to SLTs about children with EAL was not significantly higher than the number who did not give feedback ($p > .05$) whereas the number of SLTs who gave such feedback was significantly higher than those who did not ($p < .001$). The pattern of results was similar to that reported for all children and was again consistent across teachers and SLTs.

Table 7. Responses to questions concerning feedback between teachers and SLTs.

	Teachers	SLTs	Response options allowed (where appropriate)
Is feedback provided by SLTs post intervention? % yes.	77%	94%	-
Is feedback provided by teachers post-intervention? % yes.	37%	41%	-
Type of feedback sent (teachers) or received (SLTs) from teachers.	Improvements in speech, 80%.	Improvements in speech, 57%.	1) Anti-social behaviour; 2) Improvements in speech; 3) Changes in affective reactions; 4) Educational attainment; 5) All of these.
Type of feedback received (teachers) or sent (SLTs) from SLTs.	Improvements in speech, 63%.	All of these, 79%.	1) Anti-social behaviour; 2) Improvements in speech; 3) Changes in affective reactions; 4) All of these.
Does any feedback you receive post-intervention address issues about children with EAL? % yes.	14%	37%	-
Does any feedback you give post-intervention address issues about children with EAL? % yes.	40%	85%	-

Note: the questions asked are given in column one, teachers and SLTs responses are given in columns two and three and response options for the question in row four, where appropriate.

2.3.2. Resources available for identifying SLCN in general and speech difficulty in particular.

Five questions were included to obtain general information about SLCN identification, two of which were fact-finding for use in the design of future surveys (text input), and as such the results are not reported here. The results for the remaining three questions are summarized in Table 8. The first row of Table 8 shows that both groups of participants indicated that there were resources for SLCN not specific to children with EAL; significantly more teachers and SLTs indicated that they had such resources than those who did not ($p < .001$ and $p < .01$ respectively). Row two of Table 8 shows that there was general agreement that the resources they had for assessing language and literacy were unsuitable for use with children with EAL (63% of each group). Although the number of ‘not suitable’ responses was not significantly higher than ‘suitable’ responses for the individual groups, it was significant when the groups were combined ($p = .041$). Row three provides data on responses concerning whether there were any resources for language and literacy used specifically with children with EAL. Significantly more teachers, but not SLTs, indicated that they had specific resources for using with children with EAL than those who indicated that they did not have such resources ($p < .001$).

Table 8. Questions concerning general language and literacy resources and whether these are suitable for children with EAL.

	Teachers	SLTs
Resources for language and literacy? % yes.	91%	77%
If so, are they suitable for children with EAL? % yes.	37%	37%
Any specific resources for children with EAL? % yes.	83%	43%

Nine questions sought information about identifying speech difficulty in particular. One of these was fact-finding so is not reported here. The remaining eight questions addressed available resources for speech assessment and whether there was a need for more resources in this area. The results, summarized in Table 9, show that generally speaking there was good agreement between both professions. The first three rows in Table 9 showed that there were significantly more respondents who indicated that speech difficulty should be assessed separately from language for both groups of participants ($p < .001$ in both cases), that there was nothing for assessing speech difficulty in particular ($p < .005$ and $p < .01$ for teachers and SLTs respectively) and that more was needed for assessing speech difficulty ($p < .001$ for both groups).

The next five questions (summarized in rows four to eight of Table 9) sought information on the informal assessments teachers make when deciding whether a child needs to be referred to an SLT. Row four gives data concerning how children get to be seen by SLTs. The most frequent answer for both groups of participants was that teachers used their personal experience when referring children to SLTs. This response was given significantly more often than any other answer by teachers ($p < .001$), but the difference was not significant for SLTs. Row five indicates that SLTs estimated that their assessments took longer than did the teachers (more than 15 minutes and 10-15 minutes respectively), but the option chosen was not given significantly more often than other responses for either group. Row six shows that significantly more respondents indicated that assessment time varied than those who did not ($p < .001$ for both groups of participants). The reasons for any variation was raised with those respondents who indicated assessment time varied. Row seven shows that the most frequent response given by teachers' and SLTs' concerned amount of time children with EAL needed, and the complexity of difficulties children presented with, respectively. The number of teachers who gave the most frequent response versus those who gave any other response was not significant ($p > .05$). However, for the SLTs (all of whom indicated that assessment time varied) the number of respondents who indicated that the complex forms of difficulties was the reason for this variation was significantly higher than those giving any other response ($p < .001$). Row eight shows that, for both groups, significantly more participants indicated that a short systematic procedure for assessing speech difficulty would be useful than the participants who did not consider such a procedure would be useful ($p < .001$ and $p < .025$ for teachers and SLTs respectively).

Table 9. Responses to questions concerning resources for identifying SLCN in general and speech difficulty in particular.

	Teachers	SLTs	Response options allowed (where appropriate)
Assess speech separate from language? % "yes".	91%	80%	-
Anything for assessing speech? % "yes".	23%	26%	-
More needed for assessing speech? % "yes".	100%	100%	-
How do children get seen by SLTs? The predominant response is indicated with its %age.	Teachers identify, 83%.	Teachers identify, 56%.	1) SLTs assess all; 2) Teachers identify children, based on personal experience; 3) SLTs advise teachers; 4) Other.
How much time do assessments take? The most frequent choice with its %age is given.	10-15 minutes, 40%.	More than 15 minutes, 64%.	1) 1-2 minutes; 2) 3-5 minutes; 3) 5-10 minutes; 4) 10-15 minutes; 5) More than 15 minutes.
Does assessment time vary? % "yes".	89%	100%	-
Why does assessment time vary (if indicated it does)? The most frequent choice with its %age is given.	children with EAL take more time, 38%.	The range of difficulties a child has, 83%.	1) Children with EAL take more time; 2) Depends on severity; 3) Academic ability of the child; 4) Whether a child has speech difficulty or not; 5) The range of difficulties a child has.
Would a short systematic procedure for assessing speech be useful? % "yes".	100%	71%	-

Note: The question and valence of the response are given in row one and options in row four, where appropriate.

2.3.3. Awareness of potential risks involved when identifying speech difficulties.

A question was included to establish whether one or both parties were aware of potential risks when an error arises in identifying whether a child has speech difficulties. Respondents chose one of four responses to indicate which had highest risk: 1) Missing a child with speech difficulty; 2) Misidentifying a child as having speech difficulty; 3) Making errors about children with EAL (mistaking poor English for speech difficulty); and 4) No risk.

The most frequent response was ‘missing a child’ for both teachers (71%) and SLTs (88%) and this response was given significantly more often than the remaining responses in both cases (p = .0167 for teachers and p < .0001 for SLTs).

2.3.4. Role of an identification procedure for speech difficulty.

An indication was obtained from both groups of respondents of the percentage of all children (irrespective of language spoken) who were referred to SLTs by *schools* as opposed to other sources of referrals (GPs, self-referral etc.). The five response options available were 0-20%, 20+-40%, 40+-60%, 60+-80% and 80+ to 100%. The most frequent response was the same for both teachers and SLTs, at 60+-80% (indicated by 48% of teachers and 43% of SLTs). Responses outside this class were mainly skewed towards higher values. The 60+-80% response was not given significantly more often than the remaining responses by either group separately nor when the groups were combined (p > .05). However, it was given significantly more often when this option was combined with the higher estimate of 80+ to 100% (p < .01 for both groups of respondents).

Both groups of participants then ranked six factors for their order of importance when making a referral. Table 10 gives the order in which teachers and SLTs placed these factors based on the mean ranking over participants. This shows that, generally speaking, the rank order of factors when making a referral was similar across teacher and SLT groups.

Table 10. Rank order (left-hand column) of six factors for their importance when making a referral by teachers (column 2) and SLTs (column 3).

	Teachers	Speech and Language Therapists
1	Speech is affected (1.70).	Educational attainment (2.71).
2	Educational attainment (2.70).	Speech is affected (3.14).
3	Child anxious (3.10).	Child bullied (3.29).
4	Antisocial behaviour (3.70).	Antisocial behaviour (3.43).
5	Child bullied (4.00).	Child anxious (3.3).
6	Child has EAL (5.80).	Child has EAL (5).

Note: The numbers in parentheses are the mean ranks for the response option indicated.

2.3.5. Specific information solicited about a speech assessment procedure for use in schools.

The need for an assessment specifically for speech difficulty was a topic raised in the structured interview, and its importance was confirmed with the data shown in the last row of Table 9. Here information was sought about what form a speech assessment should take, bearing in mind that the system has to be workable in schools. This section sought further information following up comments that were also made in the structured interviews. The questions chosen, concerned desirable features to include in a procedure for identifying children with speech difficulty and whether there were any particular considerations about children with EAL when these assessments are made. These two issues revealed some discrepancies between teachers' and SLTs' views in the structured interviews. The relative importance of these factors concerning the information a speech test could provide, based on average rankings across participants, and are given in Table 11.

Table 11. Rank order (left-hand column) of six factors concerning the importance of a short speech test.

	Teachers	SLTs
1	External validity (2.22).	Short test can be repeated (3.21).
2	All children can do (2.39).	All children can do (3.43).
3	Short test can be repeated (2.67).	Index of intervention (3.45).
4	Standardized tests are available (4.11).	Provides an objective measure (3.56).
5	Provides an objective measure (4.83).	Standardized tests are available (3.57).
6	Index of intervention (4.89).	External validity (4.01).

Note: The rankings for teachers (column two) and SLTs (column three) are given separately. Average ranks are given in parentheses.

Table 11 shows that repeat testing was ranked third by teachers and first by SLTs. A further question addressed the optimum frequency of repeat testing in primary school. Six response options were allowed (row one of Table 12, column 4). The response given by teachers indicated a test should not be conducted any more frequently than once a year whereas both of the responses given by SLTs involved repeat testing (either at yearly or more often than six-monthly intervals). Teachers' second most popular response also indicated that testing should not take place more often than once a year (option 1). The number of teachers and SLTs who gave their respective top two options was significant ($p < .01$ and $p < .025$ respectively). However, although there were 5 options to choose from, three out of the 5 options asked about frequency of once per year, but at different stages of the academic year. For example, option 1 was "Once, when a child starts school", option 2 was "Once, towards the end of the first year in school" and option 4 was "once a year". This was included to investigate time preferences and whether they differed between SLTs and teachers. The reasons governing how often a test should be repeated were examined next. Respondents had to choose one option from three (row two of Table 12). Teachers' selections were equally split between options. However, SLTs were more focused in their choice with 74% selecting ‘identify when children develop a speech problem’ (which was significant $p < .01$).

Other questions also revealed that teachers’ views about identifying children with speech difficulty were partly influenced by considerations about potential classroom disruption, although teachers were not specifically asked about this in the questionnaire, it was discussed while testing in the schools. The teachers did not want the child to be away from the classroom for more than 10-15 minutes as they felt that they would miss some important tuition in the classroom. This aspect of concern for teachers’ should be addressed formally in the future. Respondents were asked which specialist should do the assessment for speech difficulty (row three) and which should deal with phonological delay in the first instance (row four). Most respondents to each of these questions indicated their own profession and this response was significant throughout ($p < .001$ and $p < .0001$ for teachers’ responses to speech difficulty and phonological delay and $p < .0001$ to both questions for SLTs).

Teachers who thought that teachers should do the assessment for speech difficulty and SLTs who thought that SLTs should do the assessment for speech difficulty were next asked to choose one reason from four alternatives as to why this was the case (column four of row five in Table 12). The most frequent responses for teachers and SLTs differed: Knows all the children was given by 23/28 teachers ($p < .001$) and the professional training option was given by 30/30 SLTs ($p < .0001$).

As mentioned, the last row of Table 9 indicated that any test for assessing speech difficulty should be *short*. Test length was followed up by the question in row one of Table 13: Teachers wanted the sample to be as short as possible, and this response was given significantly more often than the remaining choices ($p < .005$); SLTs indicated that a long test was needed so that symptoms were not missed. This response was given significantly more often than the remaining ones ($p < .01$).

Table 12. Responses to questions concerning administration of a test for identifying speech difficulty.

	Teachers	SLTs	Response options allowed (where appropriate)
Optimum frequency of repeat testing.	Once a year, 40%.	Once a year; more often than once in six months tied at 26%.	1) Once, when a child starts school; 2) Once, towards the end of the first year in school; 3) Once every six months; 4) Once a year; 5) More often than once in six months; 6) Less often than once a year.
Reasons given about how often a test should be repeated.	No clear preference.	Option 3, repeat periodically to identify when problems start, 74%.	1) Not too often to avoid disruption to school activities; 2) Just once, when children enter school, to ensure that children who start school with a speech problem are identified early; 3) Repeat testing periodically to identify when children develop a speech problem.
Specialist who should assess speech difficulty.	Teachers, 80%.	SLTs, 86%.	-
Specialist who should deal with phonological delay.	Teachers, 91%.	SLTs, 86%.	-
Reason for indicating which specialist should do the assessment for speech difficulty.	Knows all the children, 82%.	Has the professional training for identifying speech difficulties, 100%	Teachers: 1) Knows all the children; 2) Has the necessary skills to communicate with children; 3) Appreciates the difficulties associated with children with EAL; 4) Has the professional training for identifying speech difficulties. For SLTs, option 1 replaced by: Knows in general about children's speech difficulties. (i.e. has been educated about a range of speech difficulties)/.

Note: The question and valence of the response are given in row one, teachers and SLTs responses in rows two and three and response options in row four.

Results on type of speech sample that should be collected are summarized in row two of Table 13. Once again teachers and SLTs gave different responses; 77% of teachers chose ‘spontaneous sample alone’ (p < 01) whereas 100% of SLTs chose ‘representative samples’ (p < 001). The structured interviews also supported this difference which reflected either the need for a short test (teachers) or a thorough more time-consuming assessment of all children (SLTs).

The types of speech recording that should be taken are given in row three of Table 13. Sixty-seven percent of teachers chose audio, which was not quite significant ($p = .0895$). Eighty percent of SLTs chose audio-visual (significant $p < .001$).

The language that should be used when assessing children with EAL are given in row four of Table 13. Teachers responded English alone ($p < .001$). SLTs rarely chose English alone as an option. Alternative language alone (46%) or alternative language plus English (40%) were given about equally often as responses by SLTs. The two choices that included alternative language (alone or in combination with English) were given significantly more often than English alone ($p < .001$).

Teachers and SLTs were asked whether they would count WWR (row five of Table 13) and pauses (row six of Table 13) as signs of speech difficulty; The majority of teachers indicated that they *would not* consider WWR as an indication of speech difficulty whereas the majority of SLTs indicated that they *would* include them ($p < .001$ in both cases). Similarly, most teachers indicated that they *would not* consider pauses as an indication of speech difficulty whereas a small majority of SLTs indicated they *would* include them. Exclusion of pauses was only significant for teachers ($p < .005$).

The reason why participants who indicated they would exclude WWR) or pauses was examined next (rows seven and eight of Table 13). Participants were required to choose one of three alternatives: 23/28 teachers and 5/7 SLTs indicated 'WWR are shown by all children' (option 3) and 22/27 teachers and 14/17 SLTs indicated the same option for pauses ($p < .01$) for SLTs

Table 13. Questions addressing practical details of an identification procedure (first four rows), symptoms to include (rows five-six) and reasons for including these symptoms (rows seven and eight).

	Teachers	SLTs	Response options allowed (where appropriate)
Length of speech sample.	A short sample, 77%.	A long sample, 74%.	1) As short as possible to avoid disruption in class; 2) A sample sufficiently long to provide a stable measure.
Type of speech sample.	Spontaneous sample alone, 77%.	Representative range of speaking situations, 100%.	1) Spontaneous sample alone; 2) Representative range of speaking situations.
Recording mode	Audio alone, 67%.	Audio-visual, 80%.	1) Audio; 2) Audio-visual.
Type of language sample from children with EAL.	English, 80%.	Alternative language, 46%; English and alternative language, 40%.	1) English; 2) Alternative language; 3) English and alternative language.
Include WWR as indications of speech difficulty?	Would not include WWR, 80%.	Would include WWR, 80%.	-
Include pauses as indications of speech difficulty?	Would not include pauses, 77%.	Would include pauses, 51%..	-
Reason why WWR should be excluded.	Option 3, All children show them, 80%.	Option 3, All children show them, 66%.	1) WWR are indications of word-finding difficulty which teachers should deal with; 2) children with EAL show a lot of WWR as they do not speak English; and 3) all children show WWR, so they are not unusual.
Reason why pauses should be excluded.	Option 3, all children show them, 77%.	Option 3, all children show them, 85%.	As above with 'pauses' substituted for 'WWR').

Note: Choices are indicated for teachers (column two) and SLTs (column three) and percentage giving these responses is indicated in parentheses where appropriate. Options available for multiple choice questions are indicated in column four (as appropriate).

Overall, the findings highlight differences of opinion between teachers and SLTs concerning: How long a sample should be, whether or not the sample needs to be widely representative, the type of recording that should be taken, the language used to obtain a speech sample with children with EAL and whether or not WWR and pauses should be considered as indications of speech difficulty.

2.3.6. Provision for language in general and speech difficulty in particular for children with EAL in comparison with provision for children whose first language is English.

Four questions concerned the adequacy of current SLT provision and, in particular, whether such provision is satisfactory for children with EAL. Rows one and two of Table 14show that significantly higher numbers of respondents from both groups indicated that there was insufficient SLT provision ($p < .001$) and a lack of resources for children with EAL in particular (($p < .01$ for teachers and $p < .025$ for SLTs).

The issues on provision were followed up with two further questions that addressed how provision could be improved for children with EAL. These concerned what provision should be provided when *identifying speech difficulties* (one from the three options shown in row three of Table 14) and for *speech intervention* (one from the five options shown in row four of Table 14) for children with EAL. For teachers, ‘one-to-one guidance’ was the most frequent answer for both identification and intervention with children with EAL, and was given significantly more often than the remaining responses for identifying ($p < .001$) and intervention ($p < .01$). For SLTs, 'interpreter' was the most popular answer for both identification and intervention. However, this response was not given significantly more often than the remaining options, either for identification or for intervention.

Table 14. Sufficiency of current SLT provision and future needs.

	Teachers	SLTs	Response options allowed (where appropriate)
Is SLT provision for language adequate for all children? % yes.	6%.	3%.	-
Is there provision for SLT intervention for children with EAL? % yes.	26%.	29%.	-
What could be provided for children with EAL who have speech difficulty?	One-to-one guidance, 100%.	Interpreter; 66%.	1) Interpreter; 2) One-to-one guidance; 3) Culturally-appropriate material.
What provision should be available for children with EAL for speech intervention	One-to-one guidance, 74%.	Interpreter; 63%.	1) Interpreter; 2) One-to-one guidance; 3) Culturally-appropriate material; 4) Linguistically-appropriate material; 5) An SLT who speaks the alternative language

Note: The question and valence of the response are given in row one and responses are indicated for teachers (column 2) and SLTs (column 3). Response options (where appropriate) are given in column 4.

Finally, ten factors often associated with risk for speech difficulty were ranked in order of importance to obtain and feasibility of collecting, by teachers and by SLTs. Table 15 gives the order in which teachers and SLTs put these factors based on the mean ranking over participants. There were discrepancies between teachers and SLTs when factors were ranked either in order of importance, or how feasible they were to collect. Teachers ranked intellectual ability, handedness and use of other languages higher than SLTs whereas the opposite was true of family history and gender. Teachers ranked comorbid difficulties higher than SLTs whereas family history was ranked higher by SLTs than teachers for feasibility. Here the lower ranking given to family history by teachers than SLTs probably reflects the difficulty teachers experience with communications to parents (which also applies to other factors given a low rank by teachers). The principal factor of note from Table 15 is that there are large discrepancies between importance and feasibility. Thus, whereas both groups of participants ranked speech first in importance, they both ranked it fourth in terms of feasibility to collect.

Table 15. Rank order of ten factors for importance (top section) and feasibility (bottom section) separately for teachers (column two) and SLTs (column three).

	Teachers	SLTs
Importance		
1	Speech (2.79)	Speech (3.57)
2	Intellectual ability (3.61)	Gender (3.93)
3	Other languages (3.94)	Family history (4.14)
4	Gender (4.39)	Comorbid difficulties (5.29)
5	Family history (4.63)	Other languages (5.71)
6	Handedness (5.44)	Intellectual ability (5.86)
7	Closed head injury (7.28)	Closed head injury (6.29)
8	Birth difficulty (7.32)	Extraneous movements (6.36)
9	Comorbid difficulties (7.39)	Handedness (6.64)
10	Extraneous movements (7.72)	Birth difficulty (7.64)
Feasibility		
1	Gender (2.33)	Intellectual ability (2.00)
2	Handedness (3.56)	Gender (4.00)
3	Intellectual ability (3.78)	Family history (4.33)
4	Speech (4.11)	Speech (5.17)
5	Other languages (4.50)	Other languages (5.14)
6	Comorbid difficulties (6.25)	Handedness (5.50)
7	Extraneous movements (6.25)	Extraneous movements (6.17)
8	Closed head injury (7.50)	Comorbid difficulties (6.33)
9	Family history (7.89)	Closed head injury (8.10)
10	Birth difficulty (8.00)	Birth difficulty (8.20)

Note: The numbers in parentheses are the average ranks across participants.

To summarize, teachers and SLTs agreed that provision for language intervention (identification of problems and subsequent intervention) were not adequate for children with EAL. However, the participant groups indicated different preferences as to how provision could be improved. Most teachers wanted some form of one-to-one guidance whereas most SLTs wanted an interpreter who spoke the child's alternative language. Discrepancies were also noted when factors were ranked in order of importance and for feasibility of collecting both across these types of judgments and between the participant groups. Many of the discrepancies were attributed to practical considerations concerning what activities take place in schools (time for assessments is limited as expressed by the teachers during the screening) and what the different groups of professionals focus on (education or a wide range of communication issues).

2.4. Discussion

The current survey showed that teachers play an important role in deciding whether a child is referred for SLCN intervention. It would be impractical for SLTs to assess every child in a school and doing so would risk the possibility of over-referral (Skeat et al., 2014). Such considerations make it imperative that information is exchanged between teachers and SLTs concerning how teachers attend to children with speech difficulties (Bercow, 2008). Here, information was sought concerning the way that teachers and SLTs currently identify children with SLCN and speech difficulty (including children with EAL), what communication occurs between the professional groups and what could be done to improve matters. The picture that emerged from the survey was that there was agreement between the parties about feedback that was supplied post-intervention, the availability of resources for dealing with speech difficulty, awareness of risks inherent in making an incorrect identification that a child has SLCN, the purpose of an identification procedure and referral practices for all children and, more particularly, for children with EAL. Minor differences between teachers and SLTs concerning these issues are noted below as well as some more extensive disagreements about the form a surveillance procedure for identifying children with speech difficulties should take. The latter discrepancies were mainly due to the different constraints under which the two groups of professionals work and arise because preliminary identification is carried out by teachers with little information provided by SLTs.

2.4.1. Communications between teachers and SLTs about children with speech difficulties subsequent to intervention.

The present survey documented what feedback was being supplied between teachers and SLTs: Teachers and SLTs agreed that SLTs provided feedback, but that teachers did not provide feedback post-intervention., the teachers who reported that when they received feedback from SLTs, this concerned speech improvement. SLTs indicated that they provided a full range of feedback about all children, and those who received feedback from schools indicated that the main form of feedback concerned improvement in speech specifically. Lee’s (2014) structured interviews indicated that SLTs would have found receipt of general feedback about the children useful: In particular, they would have liked to know about any impact that their intervention had on educational attainment and behavioural engagement (Durkin et al., 2012; Johnson et al., 2010) and because this could influence SLT's treatments (Broomfield & Dodd, 2004a). The latter is particularly pertinent because an issue currently being discussed by SLTs concerns the appropriate basis for referral so that the case-load can be serviced (Skeat et al., 2014). Case-load levels are regulated locally if teachers use variable referral criteria, and SLTs need to be informed of any such variation. Both groups of participants agreed that something specific was needed for identifying speech difficulty in children for whom English is their first language (Dodd et al., 2002) and for children with EAL (McLeod et al., 2013). They also agreed that feedback from SLTs to schools was delivered but post-intervention feedback from schools to SLTs was lacking.

2.4.2. Resources available for identifying SLCN in general and speech difficulty in particular.

The section of the survey on resources was intended to establish whether schools and SLTs have what they need to identify SLCN and whether a procedure is needed to identify speech difficulty in particular. Significant numbers from each group indicated that resources for SLCN were available and both groups indicated that these resources were not suitable for children with EAL. Significant numbers of teachers, but not SLTs, indicated that there were resources for SLCN that were used specifically with children with EAL. In Lee's (2014) structured interviews the main resource mentioned by SLTs was Dodd et al.'s (2002) Diagnostic Evaluation of Articulation and Phonology (DEAP) for identifying speech difficulty, which is suitable for monolingual children. McLeod et al. (2013) are finessing requirements concerning what is needed when dealing with children with EAL. Non-word repetition is one test that could be used to identify children with SLCN (Conti-Ramsden et al., 2001) and speech difficulty that arises because of impaired phonological processing (Gathercole, Willis, Baddeley & Emslie, 1994). However, when the instruments are designed for a target language, they are not suitable for use with other languages (Masoura & Gathercole, 1999; Windsor, Kohnert, Lobitz & Pham, 2010). Other types of instrument are being developed that could be used to identify speech difficulty in children who speak English in UK schools (Holm, Dodd, Stow & Pert, 1999) and for children with EAL (McLeod, Harrison & McCormack, 2012). Some attention has been given to interventions SLTs could use with children who speak English (Crosbie, Holm & Dodd, 2005) and for use with children with EAL (Holm & Dodd, 1999; Holm, Ozanne & Dodd, 1997). Attention is also being given in educational research to interventions for children with EAL (Murphy, 2015), but at present, these are mainly addressed at children's attainment and adjustment.

The part of the survey that concerned the availability of any formal procedures and whether such procedures were needed for identifying speech difficulties, revealed the following important findings: Language assessment should be separated from speech assessment, there is no routine procedure available for speech, and that both groups considered such a procedure desirable. Recent reviews (Bishop, 2014; Reilly et al., 2014) and empirical work (Conti-Ramsden et al., 2001; Fricke, Bowyer-Crane, Haley, Hulme & Snowling, 2013) focus on tests with English-speaking children. Many available *language* tests (e.g. ones for reading) are also English-based. This is acceptable for use in schools in the UK, the US and elsewhere if children are taught in English starting from when they enter reception classes, since even children with EAL would have some proficiency in that language by the time reading tests are administered. However, this situation is not acceptable when instruments are required for assessing children with EAL in reception classes. This is a further argument showing assessments for language and literacy difficulties have different requirements to those for speech difficulty.

Information was sought on how children are referred to SLTs by schools. Significantly more teachers and SLTs indicated that teachers identified children with potential speech difficulties than any other option. Surprisingly, this appears to be the first source of information about rate of referrals by schools to SLTs, this is unexpected as researchers have assumed that speech difficulties are easy to identify and perhaps only indirectly related to language. This issue was followed up in the section where estimates of relative rates of referrals from different sources were obtained. School referrals are obviously dominant at school-entry age although many pre-school referrals are made by health visitors (Nelson et al., 2008). An implication from the survey is that the way identification takes place in schools needs to be thoroughly understood; aspects of this were examined in the next two sections of the survey.

The time taken for teachers' and SLTs' assessments varied, and SLTs indicated that their assessments took longer than did the teachers'. The difference in assessment times reported is not surprising, as SLTs would be expected to assess the comparatively small number of children they see from schools in more depth than the teachers (e.g. using Dodd et al.'s 2002, DEAP). The reason for any variation given by the majority of teachers was 'needs of children with EAL' and for the SLTs 'the range of difficulties a child has'. The children sent to SLTs from UK schools that are identified as having speech difficulties by teachers have been implicitly 'pre-screened' during the identification process (identification is one-way). During this process, teachers might ensure that children with EAL do not have a higher chance than children who speak English alone of being referred to SLTs (Strand, 2015). This would leave SLTs to deal with more complex and time-consuming cases, both children who speak English as their alternative language and children with EAL. There were significantly more respondents in both participant groups who indicated that a short systematic procedure for assessing speech would be useful than those who considered one was not necessary. Discussion about issues associated with using a short procedure for identifying one form of speech difficulty, stuttering, has appeared in Ward (2013). The issue of test length is considered in more detail in the section, where the issue of what format a speech procedure for use in schools should take was examined.

2.4.3. Awareness of potential risks involved in identifying speech difficulty.

Responses from SLTs and teachers indicated that they were aware of potential risks when identifying children. The responses of both groups of participants emphasized the importance of not missing cases. Although this supports referral to SLTs whenever a child is suspected of having speech difficulty, there is the potential for over-referral, as discussed in Skeat et al. (2014). Issues about missing cases should be followed up in focus groups of people who were either missed or conversely falsely identified as having speech difficulty to see whether they agree or whether the latter group, in particular, has a different opinion from the teachers and SLTs. Although the survey documented that both parties were aware of difficulties, at present there is no information about relative rates of missed and falsely-identified cases.

2.4.4. Referral for speech difficulty.

Referral by schools is an important route for a reception class child with speech difficulty to be seen by an SLT relative to other sources of referral (parents, GPs etc.). Most teachers and SLTs indicated that referral by schools was the most common source of referral when children enter schools.

Teachers and SLTs agreed about the importance of the factors they take into account when making a referral. In particular, teachers and SLTs ranked six factors for importance when making a referral, in a similar order. Both groups placed ‘child has EAL’ in the last position. This probably reflects the fact that they were balanced in their judgments (no positive or negative attitude about speech difficulties for children with EAL) as discussed earlier. There is a need for continuing scrutiny as there is recent evidence to suggest that children who speak certain EAL languages in UK schools (e.g. Somali, Lithuanian and Lingala) are at some disadvantage in terms of educational attainment (Strand, 2015).

Other possible aims regarding a speech identification procedure are: to ensure that not too many children are sent to SLTs (Skeat et al., 2014); how to ensure universal provision (Murphy, 2015; Strand, 2015), whilst at the same time keeping costs down; the impact of any new procedure on work in schools and how to ensure that SLTs’ advice can be included so that children with speech difficulty are not missed by streamlined procedures (Bercow et al., 2008).

2.4.5. Form of speech assessment in schools.

Questions concerning the form a short speech assessment should take revealed some disagreement between the two groups of respondents, for instance concerning who should make the assessment. The difference was mainly due to practical consideration (e.g. Minimizing disruption in the classroom, maximizing learning time) that teachers took into account regarding how speech assessment would work (Howell, 2013) versus the need for comprehensiveness by SLTs (Bishop, 2014). A major task is how to meet these antithetical demands. The following areas of disagreement can be interpreted as the teachers' concerns about disruption to core activities planned for the class that may affect the child's learning outcomes. The language in which children with EAL should be recorded was considered to be English for teachers versus children's alternative language for SLTs. The teachers' choice was expressed forcefully in the structured interviews (Lee, 2014). The reason was that, unless there was an alternative speaker of the child's language on the staff, bringing in outside people to assess children was prohibitive in terms of disruption to class activity and expense. Equally strong views were expressed by the SLTs for their choice in the structured interviews where considerations about fair and thorough assessments were mentioned for English-speaking (Dodd et al., 2002) and children with EAL (Crosbie et al., 2005; Holm & Dodd, 1999; Holm et al., 1999; Holm et al., 1997; McLeod et al., 2012; Stow & Dodd, 2003).

The practicality issue also relates to the answers elicited about a speech assessment procedure for use in schools, where six factors were rank ordered. The structured interviews highlighted that teachers wanted something which complied with their notion of whether a child has speech difficulty. This discrepancy was reflected in the survey in terms of the ranking teachers gave for external validity (ranked first by teachers but last by SLTs).

Areas of agreement in the rankings were the fact that a speech test can be done by all children (2nd for teachers and SLTs) and the importance of repeat testing so that late onset of speech difficulty can be identified (3rd for teachers and 1st for SLTs). Regarding frequency of repeat testing, teachers would prefer tests to be repeated less frequently than would SLTs, because assessing speech is not a core activity in schools.

Two other areas of disagreement were based on considerations other than practicality. First, teachers considered that they should intervene when there is phonological delay because they see this as part of their educational responsibility whereas SLTs considered that *they* should address phonological delay because they have had the necessary professional training. Once again, this is an issue that needs to be resolved. Thus, it is possible that the way the term ‘phonological delay’ is interpreted differs across the participant groups (Dodd et al., 2002). The second area concerned the report that teachers preferred audio recordings whereas SLTs considered that videos were preferable. The structured interviews indicated that the audio mode was preferred by teachers so that child confidentiality could be ensured (Lee, 2014). The reason SLTs wanted a video was so that body movements or tics associated with speech difficulty could be detected (Leckman, Walker & Cohen, 1993).

To summarize, both groups of respondents see the usefulness of a short test for teachers, and the apparent divergence of views about testing is because they work in different environments and under different constraints.

Whilst the detailed assessments that SLTs can offer are essential for the roles they play, something short is needed for teachers. A short test needs to minimize the chance of missing children with speech difficulty (Skeat et al., 2014) and should ensure that children with EAL are not disadvantaged (Strand, 2015), particularly when they are assessed in English. Most importantly, SLTs need to know how assessment of children's speech difficulty is made in schools. The situation whereby informal, *ad hoc* and variable assessments are made is rather unacceptable given the teachers' expression for the need for something to assess speech difficulty and other issues that teachers are currently finding difficult to deal with. Moreover, the requirement for teachers to assess children was a recommendation of the Bercow report, which suggested that a national screening program is needed. This would reduce the stress on teachers to adopt different methods and ad hoc assessments in order to check if they have identified children with speech difficulties. The next section also supports a speech-based assessment because it is an important risk factor in its own right and collection is feasible whereas other risk factors are either less important or are considered by teachers to be less feasible to collect.

Some symptoms that could be used to identify speech difficulty tended to be included by SLTs, but excluded by teachers. For instance, 80% of teachers indicated that they would not include WWR whereas 80% of SLTs indicated that they *would* include them. Similarly regarding pauses, 77% of teachers indicated that they would not include them and 51% of SLTs indicated they *would* include them. The teachers' reason for excluding WWR and pauses was that they are not unusual speech features. The exclusion of these two types of symptoms is related to the disagreement about who should assess children with EAL and children with phonological delay discussed earlier. Both these groups of children have high rates of word repetition and pausing due to word-finding difficulty (Bada, 2010; Rad & Abdullah, 2012). Therefore, the exclusion of these symptoms could be a result of teachers' views on how children with EAL or phonological delay should be assessed, and is a factor that ensures not too many children with EAL are referred to SLTs (Skeat et al., 2014).

2.4.6. Provision of resources for language in general and speech in particular that are appropriate for children with English as their first language and children with EAL.

Research studies assessing SLCN have excluded children who do not speak English as their first language (e.g. Conti-Ramsden et al., 2001; Fricke et al., 2013). The exclusion of children with EAL from consideration cannot be entertained in the UK as they are present in large numbers (Strand, 2015). Teachers and SLTs agree about EAL provision, except (again) where practical considerations governed answers. Provision for SLT services in general, and for children with EAL, were considered inadequate. When asked what provision was needed for identifying speech difficulty and subsequent intervention in children with EAL, the answers given by teachers and SLTs differed. The most frequent answer given by teachers was one-to-one guidance for both identification and intervention for children with EAL. The answer given most often by SLTs was access to an interpreter for both identification and intervention for children with EAL.

The rank order of ten factors assessed by teachers and SLTs for both importance and feasibility was given in Table 2.11. This also revealed differences between the groups of respondents arising from practical considerations and differences in professional roles. For importance, intellectual ability (2nd for teachers and 6th for SLTs) and comorbid difficulties (9th for teachers and 4th for SLTs) showed some discrepancy. The emphasis on intellectual ability by teachers and on comorbid difficulties by SLTs is not surprising, as they reflect aspects of their respective professional roles such as educating pupils or dealing with problems like stuttering. Family history was the main factor showing discrepancy when assessed for feasibility (9th for teachers and 3rd for SLTs). In this case, the difference between the rankings suggests that SLTs underestimated the problem of obtaining accurate information from children's families including family history of speech difficulties. The teachers face difficulties obtaining this information perhaps because the parents may be reluctant to give this information to someone other than professionals who deal with speech problems. Agreement about what is important and what is feasible was generally good, other than for the exceptions noted. Looking across importance and feasibility rankings, speech was ranked high on both of these for teachers and for SLTs (first for both groups for importance and 4th for both groups for feasibility). Birth difficulty was ranked low for importance (8th for teachers and 10th for SLTs) and for feasibility (10th for both groups). Difficulty around the time of birth is one of the main factors noted by Nelson et al. (2008), as frequently reported in their systematic review of research to be a risk factor for language communication difficulties. However, teachers consider this information to be difficult to collect.

2.5. Summary, Limitations and Future Work

This survey provided the first impression of UK practice for identifying children with speech difficulty in schools. The major areas where information was lacking prior to this survey concerned whether teachers and SLTs thought that speech should be assessed separately from language, whether school referral is an important route for referring children to SLTs, how children are identified by teachers, what additional factors teachers take into account when deciding whether to refer a child to SLTs, what factors each group feeds back to the other party post-intervention and whether the feedback differs between teachers and SLTs. One issue that has not been discussed in sufficient depth in prior literature is children with EAL's speech difficulties (McLeod, et al., 2013).

Some of the questions may have benefitted if other response options were offered. For example, in terms of contact with children with EAL, 0-20% was the bottom option. 'Not experienced' may have been more informative when addressed to SLTs.

The main focus of the survey was on *identifying* speech difficulty, rather than intervention. The current survey was partly based on Nelson et al. (2006) but went beyond that in examining issues associated with speech difficulty separate from language difficulty, and in its attention to children with EAL. The survey was completed by reception class teachers. The role of other educational workers (e.g. special needs coordinators, teaching assistants) may also have an impact and should be examined in the future.

Chapter 3

Length and Type of Recording[3]

As was mentioned in the previous chapter, a procedure for identifying children with speech difficulties that could be used by teachers or by trained staff in schools, needs to be short and easy to use. The screening work developed in this thesis is based on the SSI-3 procedure. SSI-3 is fully documented (a review of some features that shows this is given in the next section). However, there is a later version (SSI-4, Riley 2009) where certain changes that would affect its usefulness if employed for a screen (change in sample length requirement and the suggestion to use video samples). A general problem concerning the introduction of SSI-4 is that these and other suggested changes are so extensive that they would call for a re-standardization. Since this has not been done, the SSI-3 procedure should be employed allowing the problematic features in SSI-4 to be avoided.

Riley (1994) specified that a 200-syllable long sample was sufficient for obtaining an SSI-3 score, making this version of the instrument appropriate for use in schools. However, although Riley's length specification was explicit, no details concerning how he arrived at this value have been given. He also

[3] A version of this chapter has been published as: Todd, H., Mirawdeli, A., Costelloe, C., Cavenagh, P., Davis, S.., & Howell, P. (2014). Scores on Riley's Stuttering Severity Instrument versions three and four for samples of different length and for different types of speech material. *Clinical Linguistics and Phonetics, 28,* 912-926.

stated that the instrument can be used to identify children with disfluency. Therefore, the question arises as to whether 200–syllable long samples are sufficient when SSI-3 is used as the basis of a screening tool. Checks also need to be made concerning whether using a spontaneous sample alone in assessments (all that can usually be obtained when assessing reception class children) is satisfactory. The empirical work in the current chapter assessed whether the length specification is satisfactory to give stable SSI-3 estimates and how reader and non-reader SSI-3 procedures compare. If these procedural details are confirmed as appropriate, the SSI-3 non-reader procedure would provide a suitable basis for use in screening in schools.

3.1 Standardization of the SSI-3

From a general perspective, the design features of SSI-3 (Riley, 1994) and SSI-4 (Riley, 2009), generically SSI, have led to them being the most frequently used instruments in fluency assessment; SSI has been employed in over 350 peer reviewed publications and has been translated into, but not standardized for, other languages (Bakhtiar, Seifpanahi, Ansari, Ghanadzade, & Packman, 2010). The design features of SSI that have led to its widespread use for reporting details about people who stutter include: clear description of events that are counted as stutters; flexibility of the procedures with formats for use in clinics versus research laboratories as well as non-reader and reader versions of the assessment package; and specification of the minimum sample length that is required. Providing a clear-cut and brief, but comprehensive, list of symptoms makes it easy for an assessor to conduct the test allowing it to be used in clinical, laboratory or home/school settings. The length of sample required to make an SSI-3 score (200 syllables) is the same across the reader and the non-reader forms. All of these features of SSI were identified as desirable features to include in an in-school assessment for speech difficulty in Chapter 3.

The symptoms sets designated as non-fluent events are identical across all versions of SSI. That said, there has always been some disagreement about the specification concerning what events are and are not stutters (i.e. not everyone agrees with Riley) and this has led some researchers to use Riley's procedure with their own symptom set. To illustrate the discrepant views about what events are non-fluent, Table 16 summarizes the differences in stuttering symptom sets between Johnson and Associates (1959), Howell (2010) and Riley (1994, 2009). In this thesis, following Riley (1994; 2009) single syllable WWRs, were not counted as stutters in analyses.

Table 16. Stuttering symptoms (column one) with illustrative examples (column two) from Johnson and Associates (1959).

Johnson (1959)	Illustrated examples	Counted as stuttering by Riley (1994, 2009)?
Incomplete phrases (or abandonments	That co... that animal	No
Revisions	The sis... the nurse.	Not explicitly stated
Interjections	Um, erm, ah,	No
Whole-Word Repetition (WWR)	He, he, he went	Not counted unless it includes other difficulties, such as prolongations, it's shortened, staccato or tensed.
Phrase repetition	To the, to the house	No
Part-word repetition	Lo-lo-london	Yes
Prolongations	Sssssleeping	Yes
Broken words	Vol-cano	Not explicitly stated

Note: The right-hand column indicates whether the symptom specified in column one is counted as stuttered by Riley (1994; 2009).

Statistical assessments of SSI are available. The reliability and validity of the SSI-3 were reported in Riley (1994). The reliability assessments were conducted on the converted raw scores that were obtained following the procedure detailed in the SSI-3 manual (details of how this is done were given in section 2.2). The statistical evaluations that were conducted were as follows:

Intra-judge reliability: This concerns the level of agreement between repeated administrations of the SSI-3 by the same assessor. Two experienced examiners and a research assistant each scored 17 samples twice. For intra-judge reliability, the agreements across the occasions that the judges repeated the assessments were compared. Agreement for each of the samples for all judges for both duration and frequency components were good (the minimum agreement was 71.4%). No intra-judge results were given for physical concomitants other than mention that a video tape sample could be used for obtaining this.

Inter-judge reliability: this concerns the level of agreement across different judges for the same samples. The two experienced examiners mentioned above served as judges as well as 15 trained graduate students. Table 17 gives the percentage agreement for each SSI component as well as for the overall SSI-3 score for the experienced examiners and the team of 15 trained graduate students.

Table 17. Data from Riley (1994), showing the percentage (%) inter-judge agreement (left) between two experienced assessors (middle column) and between 15 trained graduate students (right column).

Component of SSI-3	Experienced assessors (N=2) %	Trained students (N=15) %
Frequency	91.0	91.4
Duration	84.4	87.8
Physical concomitants	82.9	Not available
Total overall SSI-3 score	93.4	Not available

Criterion validity: This concerns how SSI-3 relates to other independent measures of severity. Riley (1994) argued that his Stuttering Prediction Instrument, SPI (Riley, 1981) provided an appropriate measure for comparison with the SSI-3. Riley (1994) reported that the correlations between SSI and SPI ranged from .72 to .83 and that this was adequate for criterion validity.

Construct validity: This concerns whether the internal components of the SSI-3 (frequency, duration and PC measures) measure what they claim to be measuring. Riley (1994) reported two tests. The first was based on the argument that the longer a person has stuttered, the more severe the problem is. Therefore, if SSI-3 has construct validity, scores for a group of individuals should correlate with the length of time that they have stuttered. This was evaluated and support for all three components for three groups of participants (pre-school children, school age children and adults). The other test for construct validity (part-whole comparisons for SSI-3 components) is reported in Chapter 4 (sections 4.2.2 and 4.6.3).

3.2 Procedural Changes in the SSI-4

For the SSI-4 (Riley, 2009), all the components were retained but, as mentioned, some procedural changes were made. Despite the changes, no new statistical analyses were conducted over and above those just discussed. This is problematic as it is apparent below that the changes that were made were substantial and would require re-standardization if adopted. The SSI-4 requires the examiner to count the syllables using a computer mouse, using one side to count the syllables (each click counts as one syllable) and the other side to indicate that the syllable is disfluent. When signalling disfluenct syllables, the examiner also holds down the right key for the entire length of the disfluency to measure the duration. This marked change in procedure may affect the final counts when compared to the 1994 manual standards. No comparisons were made between the procedures, at least using statistical tests. And this raises questions about the validity of the new method.

The minimum number of syllables required was changed from 200 syllables for the SSI-3 (Riley, 1994) to 150 minimum syllables for the SSI-4 (Riley, 2009). Riley (2009) also suggested that for improved reliability, samples of between 300 and 500 syllables should be used. That said, no evidence was provided that long samples have better reliability nor was any other support given for this suggestion. It is reasonable to assume that different length speech sample may be needed when SSI-3 and SSI-4 are used for different purposes. For example, they may need to be long for clinical work (Meltzer, & Wilding, 2009; Yairi & Seery, 2011). Conversely, when there is time pressure, as is the case when the instrument is used for screening large cohorts of children in schools most of whom are fluent, the shorter the sample needed the better. The proposed change in number of syllables in SSI-4 is one factor that would call for re-standardization.

Riley (2009) suggested that audio recordings of samples should be replaced with audio-visual samples to make the scoring of physical concomitants more objective. However, teachers consider videoing children to be inadvisable (see Table 13) for two reasons; 1) it is difficult to obtain timely consent from the parents (parents may seek further advice or delay reply on the basis of uncertainty and 2) the teachers, as well as the parents, would worry about the potential misuse and the ensuing harm if the videos were not appropriately dealt with (i.e. who exactly would have access to them and for how long/ how can they be sure that they will be deleted, etc.). This would also be a major change in SSI-4 over SSI-3 and would specifically affect the PC score. Therefore, this would also call for re-standardization before SSI-4 procedures are adopted.

Another concern is about the use of the reader and non-reader forms of assessment allowed in the SSI-3 and SSI-4. For non-readers only a spontaneous sample can be obtained but for someone of reading age, two separate samples are required (reading and spontaneous). It is necessary to investigate the relationship between the scores obtained with the two forms. For example, if the children screened in this thesis were followed up at later stages when they can provide samples of their read speech, it is necessary to know whether two scores obtained according to the different forms would give stable values. Only if this is the case can changes in scores over occasions be taken as an indication of recovery (scores drop) or that speech difficulty has arisen (scores increase). The alternative is that changes in scores over the two forms of SSI result in differences for spurious reasons, precluding comparison across the period of time over which the child has learnt to read. Analyses to rule out the latter possibility are reported in this chapter.

The main procedural change in SSI-4 (Riley, 2009) concerned the way stutters and syllable counts were made. This changed from a method using paper and a pencil in SSI-3 to a computer program in SSI-4. In the new counting procedure, the assessor uses a computer mouse and makes a right click to increase the fluent syllable count or a left click to increase the stutter count. The software automatically logs the syllables and stutters. To measure stutter duration, the user holds the stutter (left) key down for however long the stutter lasts. There are no reports of the reliability or validity of this new method. As this is a drastic change to procedure, the reliability and validity statistics reported for SSI-3 and reviewed earlier would not apply.

3.3 Research Questions

A change in length of sample needed for obtaining a score was made between SSI-3 and SSI-4. The reason for this change was not commented on in Riley (2009). In particular, no justification was given nor were any new analyses offered to support the change in sample length required (Riley, 2009). In addition, for both SSI-3 and SSI-4, no assessments concerning sample length specifications have been reported. Therefore, to address these issues, the impact on SSI-3 scores of using samples of different length needs to be investigated.

The consistency of the measures obtained with the reader form (a spontaneous and a read sample are used) and the non-reader form (only a spontaneous sample is used) was also evaluated. One reason for doing this is that the relationship between these two forms of assessment has not been reported previously. More importantly for current concerns, if SSI is to be used in longitudinal work in schools, children will start on the non-reader form and progress to the reader form. Therefore, it is essential to check that the two forms provide equivalent results when applied to a given child's speech. Otherwise, when scores from a child above and below reading age differ, it would not be known whether this resulted from a change in the child's performance or whether the two forms of the test are not comparable.

3.4. Methods

3.4.1. Participants.

There were two separate age groups in this study; a younger age group (age range from 2 years and 8 months to 6 years and 3 months; mean= 4 years 7 months, SD= 1 year 0 months; N=23; male=18) and an older age group (age ranges from 10 years and 0 months to 14 years and 7 months; mean= 13 years and 0 months, SD= 1 year and 1 month; N=31; male=22).

3.4.2. Recordings.

An SLT based in Ipswich, Suffolk, obtained speech samples from the younger children. None of these children could read and only provided a spontaneous sample. All these children were categorised as children who stutter by the SLT. The spontaneous samples were elicited using topics of interest, such as school, sporting activities, TV and hobbies. These same topics were used to elicit spontaneous speech for the older age group. For both groups, speech was recorded on a Sony DAT audio-recorder using a Sennheiser K6 microphone. Physical concomitants were scored at the time that the recordings were made by the SLT following the instructions given in Riley (1994).

The older age group provided spontaneous and read samples of speech. For the read samples, age-appropriate material provided by Riley (1994) was used. The assessor took PC at the time of the recording; in this case, the assessor was a trained academic who was not an SLT.

3.4.3. Annotations of the speech samples.

All the audio samples were uploaded into Speech Filing System (SFS) (Huckvale, 2013). Free software, available for download from University College London (http://www.phon.ucl.ac.uk/resource/sfs/). Trained members of staff at UCL transcribed and coded the SFS files in a format that allowed syllable counts to be made; all stuttering events were also marked on the software for subsequent analysis.

Eight speech sample files were selected at random to estimate inter-judge reliability. They were transcribed and coded by two trained members of staff at UCL. Agreement was well above chance with Kappa coefficients of .92 (96%) for fluency counts and .89 (93%) for syllables (Fleiss, 1971). The three stuttering events with the longest durations were identical within 5 ms for both judges. As PCs were obtained at the time of testing, and no audio-visual recordings were available, inter-judge reliability was not assessed for this component.

3.4.4. Administration of SSI-3.

This followed the procedure detailed in section 2.2.

3.4.5. Procedure.

For each age group, 250 syllables in each sample were counted. Each sample was divided into extracts that increased by 50 syllables at each step (first 50 syllables; the first 100 syllables including the first 50 syllables; the first 150 syllables, which also included the first 100 syllables; the first 200 syllables, which included the first 150 syllables; and all 250 syllables which included the first 200 syllables) SSI-3 scores were calculated independently for all extract lengths.

3.4.6. Analyses.

As the procedures differed across the two age groups, the data from each age group were analysed separately. IBM SPSS Statistic 21 (SPSS Inc., Chicago, IL) was used to perform all statistical analyses.

3.5. Results

3.5.1. Sample length for the younger age group.

A one-way ANOVA was used to analyse the data from the younger group of participants using length of the extract in syllables as the factor (i.e. 50-, 100-, 150-, 200- and 250-syllable long extracts). Mauchly's test of sphericity was violated so the degrees of freedom were adjusted using the Greenhouse-Geisser correction. Results revealed a main effect of sample length $F(2.354, 51.784) = 8.515$, $p<.01$. The means and 95% confidence intervals are shown for each sample length in Figure 3.

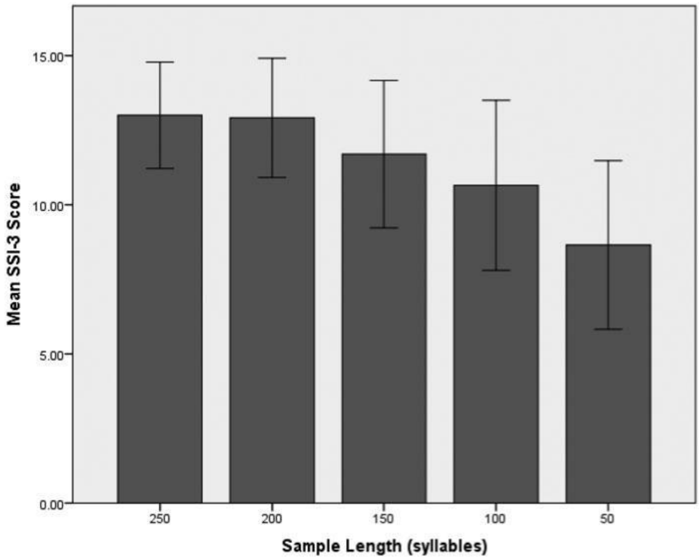

***Figure 3*. Mean SSI-3 scores and ±1 standard error for different sample lengths for the younger group of children**

Table 18. Post-hoc t-tests comparing the 200-syllable-long sample with other sample lengths for the younger group of children.

Comparison	T	Significance
250/200	0.20	0.847
200/150	1.95	0.064
200/100	2.80	0.010*
200/50	3.65	0.001*

*Significant at p<0.01.

Figure 3 shows that as the sample size decreased so did the SSI-3 scores. However, the 200 syllable long sample had a similar overall SSI-3 score to that of the 250-syllable long sample. The post hoc related *t*-tests in Table 18 give support for these impressions. Bonferroni corrections were There were significant differences between samples that were 100-syllables or less and the 200-syllable sample but there was no difference between the 150-syllable, 200-syllable and the 250-syllable long samples.

3.5.2. Sample length and test form (reader/non-reader) for the older age group.

An ANOVA was conducted with two factors (length and form of procedure reader/non-reader). Mauchly's test showed that sphericity was violated again, so degrees of freedom were adjusted with the Greenhouse–Geisser estimates of sphericity. The main effect of sample length was significant ($F_{2.15, 64.38} = 10.08$, $p < 0.001$). However neither procedure ($F_{1, 30} = 0.026$, $p = 0.873$) nor the interaction between sample length and procedure ($F_{3.138, 94.137}=0.558$, $p = 0.652$) were significant. Therefore, no differences were detectable between the reader and non-reader forms.

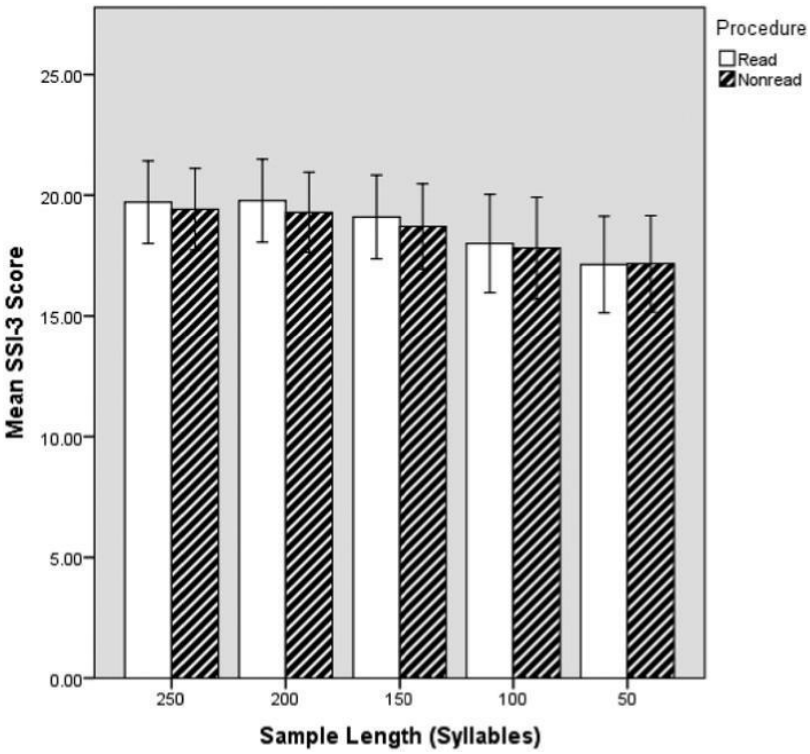

***Figure 4*. The mean SSI-3 scores with ±1 standard errors for the different sample lengths for the older group of children.**

Note: The SSI-3 scores from the reading procedure are indicated by the white bars on the left while the black diagonal stripes on the right indicate the SSI-3 score using the non-reader procedure.

The sample length factor was followed up with post hoc related *t* tests as with the younger age group and Bonferroni correction was made. Table 19 shows that there was a significant difference between the SSI-3 scores of 200-syllable long samples with samples of 150-syllables and shorter for the reader procedure. However, this was not the case for the non-reader procedure. In the non-reader procedure, the only significant difference was between the 50-syllable and 200-syllable long samples. This suggests that samples of any length above 50 syllables provide stable estimates of SSI-3.

Table 19. t-tests for the older children. Tests were conducted for both the reader and non-reader procedure.

Procedure	Sample length	*t*-test statistic	*P*
Non-reader procedure	250/200	0.57	.72
	150/200	1.43	.163
	100/200	2.50	.018
	50/200	3.04	.005*
Reader Procedure	250/200	-0.34	.738
	150/200	3.09	.004*
	100/200	3.79	.001*
	50/200	4.88	.001*

Note: The SSI-3 scores for a 200-syllable-long sample against other sample lengths. The results for the reader and non-reader procedures are given at the top and bottom respectively.
*****Significant at $p<0.01$.

3.6. Discussion

For the younger group of participants, 200 syllable long samples gave the same SSI-3 scores as 150- and 250-syllable long samples whereas shorter samples differed significantly. This suggests that the recommended minimum 200 syllable long sample is appropriate for these participants. SSI-3 overall scores reduced significantly when syllable lengths were shortened below 150 syllables suggesting that these sample lengths are too short. For the older participants, SSI-3 scores reduced when sample length was reduced below 200 syllables for both the reader and reader procedures. This suggests that samples should be 200 syllables long at minimum for stable SSI-3 score to be obtained for the older participants. Together, these results confirmed that the 200 syllables suggested by Riley (1994) as a minimum is an appropriate sample length for obtaining stable SSI-3 scores for both groups of participants.

The reader and non-reader samples at all sample lengths were compared for the older group of participants. In the two factor ANOVA, the main effect of sample length (discussed above) was significant. However, there was no effect of test format either as main effect or in interaction with sample length procedure. This is consistent with the view that reader and spontaneous samples give similar SSI-3 scores (Todd et al., 2014). This is the first report that shows that the reader and non-reader forms provide equivalent scores. These results support comparisons being made across ages where children tested at one age use one form (non-reader at an early age) and at another age using the alternative form (reader at older ages).

3.6.1. The age effect and the length of the speech samples.

The suggestion that 200-syllable long speech samples are adequate to provide a reliable SSI-3 score does not mean that samples of this length need to be adopted for all applications. For example, different sample lengths have been deemed necessary in other research: Thus, Sawyer and Yairi (2006) argued that research has been varied when it comes to sample length needed for clinical assessments of stuttering. They pointed that speech sample size has varied greatly from study to study, as well as among subjects in the same study. They gave the examples of Johnson et al. (1959), who included samples that ranged in length from 31 to 2044 words, whereas Schwartz and Conture (1988) used 85–650 words. They remark further that many studies in the past two decades were based on samples of 300–350 words (e.g. Conture & Kelly, 1991). Some samples have been even smaller, with Yaruss

(1997) employing 200-syllable samples, and Onslow, Costa, and Rue (1990) using samples as short as 1 minute. Here, the authors tacitly imply that longer sample sizes are necessary for appropriate assessment. However, here it is argued that different sample lengths are required when there are different applications. In particular, samples as short as 200-syllables would appear appropriate for screening for fluency in schools.

The oldest speaker in this study was just under 15 years of age. This highlights the need for length of sample to be investigated with adults who stutter. Some previous research on adults who stutter that used longer samples has been reported. For example, Logan and Haj-Tas (2007) showed that the overall %SS does not change in sample lengths of 300 to 1800 syllables even though a more extensive symptom set than that employed in SSI-3 was used (Roberts et al., 2009). This would suggest that even for evaluation of stuttering, samples do not need to be longer than 300-syllables. Indeed shorter samples may be adequate too as was found here with the younger children.

3.6.2. Reader, non-reader procedures.

The SSI scores of the reader procedure did not differ significantly from the SSI scores from the non-reader procedure. This indicates that a spontaneous speech sample alone can give a comparable estimate to the reader calculation that includes the extra speech sample. This may suggest that the read sample could be discarded. However, before this suggestion is adopted further investigation is necessary. In cases where time is short, or a read sample cannot be made, then scores from the spontaneous speech could be used, albeit with caution. It is possible that a read sample could provide information about avoidance, as stutters can be concealed using this gambit in a spontaneous sample, but not in a read one. Since young children can only provide non-reader samples, avoidance strategies are less likely to be picked up (Ward, 2013). However, the current findings did not reveal any significant difference between the reader and non-reader forms. Hence, support is lacking for the idea that the read sample could provide information about avoidance.

3.6.3. Limitations.

Only the minimum duration of speech samples was investigated in this study. Riley (1994) maintained that longer samples should be used to provide more detailed information where possible. The use of long samples merits investigation across ages since there appears to be some discrepancy between research into adults (Logan & Haj-Tas, 2007; Roberts et al., 2009) and children (Sawyer & Yairi, 2006). Another issue with the SSI-3 and SSI-4 is the difficulty in making objective assessments about physical concomitants and the duration of the three longest stutters. These are topics considered further in Chapter 4.

3.6.4. Conclusion.

Riley (1994) suggested that the minimum sample length should be 200 syllables. The present research confirmed that this sample length is satisfactory (i.e. gives a stable SSI-3 score compared to longer samples) for children. This research also showed that the reader version and the non-reader versions of the SSI-3 give similar scores when applied to samples from the same participants. In the previous chapter, teachers indicated that a short sample was desirable for an instrument used in schools and SSI-3 appears to fulfil this requirement. The teachers also need something that can be administered to children who may not be able to read if screening is undertaken in reception classes. For SSI-3 to be suitable and to make it appropriate for testing older children (who can read), the reader and non-reader forms need to give corresponding results. This was found to apply in the analyses that compared reader and non-reader forms.

Chapter 4

Is It Necessary to Assess Fluent Symptoms, Duration of Disfluent Events and Physical Concomitants When Identifying Children Who Are At Risk of Speech Difficulty?[4]

4.1. Introduction

A short procedure could be used to assess all children again at later stages in school in order to identify fluency problems that arise post-entry. In some respects, SSI-3 has the required efficiencies built in, which commends its use as a basis for identifying children with speech difficulty. For example, the minimum length of speech sample that is required is 200 syllables (Riley, 1994) and this was validated in Chapter 3 (Todd et al., 2014). SSI-3 allows users to base the scores on analysis of a spontaneous speech sample alone. The use of a spontaneous sample alone was also validated in Chapter 3. These features make SSI-3 appropriate for children starting school who may not be able to read (Howell, 2013).

[4] A version of this chapter to appear as: Mirawdeli, A. & Howell, P. (in prep). Is it necessary to assess fluent symptoms, duration of dysfluent events and physical concomitants when identifying children who are at risk of speech difficulty? *Clinical Linguistics and Phonetics*.

However, to date, there has been limited investigation concerning whether each structural component (frequency, duration and physical concomitant measures) of the SSI is necessary and none concerning this question as it applies to screening. A speech screening tool needs to be short but effective, easy to train teachers to use, yet reliable in the results it provides. Currently the three structural components of SSI are combined to establish an overall score. The overall scores are successful in screening (Howell, 2013; Mirawdeli, 2015; Chapter 5). However, is the overall SSI score the ideal measure to use in a screening tool? This question can only be answered affirmatively if each component is measured in a robust way, the procedure for combining them is principled and all of them have validity when assessed against independent assessments of the fluency of children. These questions are investigated in the current chapter.

Just to recap, the first of the three components required to obtain an SSI-3 score is the frequency of non-fluent events. Counts of non-fluencies are expressed relative to the number of syllables in a passage and converted to a percentage (percentage of stuttered syllables, %SS, also called the frequency measure). Duration of the three longest non-fluent events in the sample is the second component. The final component measures physical movements of the face and body that occur at the time the speech is collected (termed physical concomitants, PC). When identifying at risk children in schools, time would be reduced and the assessment made simpler if one or more of these components of SSI-3 could be discarded. PC is most likely to be redundant as concern has been expressed about how objectively it can be scored (Bakhtiar et al., 2010; Todd et al., 2014). Also, the procedures that Riley (1994) outlined for assessing PC are ambiguous (Todd. Mirawdeli et al., 2014) and PC has lowest reliability of the three components of SSI-3 (Bakhtiar et al., 2010; Lewis, 1995). In addition, it may be possible, in principle, in the near future to save assessment time by automating %SS and possibly the duration component as well, but this looks to be unlikely in the case of PC. Evaluation concerning whether all components are necessary are given next and then the research questions investigated are presented.

4.1.1. Evaluation of Riley's Stuttering Severity Instrument

1. Assessment of components.

All the components for obtaining an SSI-3 score (%SS, duration and PC) are obtained at the time that either one or two samples of speech are collected. The raw PC score is added to the transformed %SS and duration scores to obtain the total overall SSI-3 score. The SSI-3 scores can be reported as: the raw numerical values (used throughout this chapter); a percentile; or a severity descriptor (very mild to very severe).

In clinical samples, although assessment of fluency on multiple dimensions of behavior is desirable because it provides a picture over and above that obtained from speech *per se*, a single composite score that summarizes severity prevents users from identifying which specific components lead to a speaker's problem (Lewis, 1995). Consequently, speakers with the same SSI-3 score can weight differently on the three components. The ambiguity about what a particular SSI-3 score means is a further reason why the contribution that individual SSI-3 components make to overall scores needs examining. Information about separate components is also useful for deciding whether one or more of them can be dropped, thus reducing ambiguity in the interpretation of SSI-3 scores.

2. Frequency measure (percentage of syllables stuttered, %SS).

A measure of %SS is an essential component when stuttering severity or fluency in general is assessed. The procedure for obtaining %SS in SSI-3 is precise. Procedures to count syllables, and stutters are described and clear and unambiguous definitions of what events are considered as stutters are given in the SSI-3 manual (Riley, 1994). The ratio of stutters out of all syllables is then converted to a percentage to give %SS. Transformation of raw %SS is simple (a conversion table is provided in the manual, see section 1.5.2 for a full description). The way this is subsequently used in computation of the total score is straightforward.

3. Duration.

Bloodstein and Ratner (2008) questioned whether duration was useful when fluency is assessed. Moreover, the procedure for estimating raw duration scores in SSI-3 is dated. A manual method involving a stopwatch was used. Making the measurement this way requires pre-emptive judgements about what words will be stuttered. Anticipating when a stutter is coming up would impact on the reliability of the assessor's duration measurement because they have to: 1) concentrate on establishing whether the next syllable is stuttered; and 2) perform various multitasking activities such as using the stopwatch (it takes time to start and stop), monitoring the child's speech and movements (for PC), and preparing prompt questions when spontaneous speech samples are collected. All of these factors risk duration measurements being inaccurate. These problems would be particularly acute when the instrument is used in schools (teachers are unlikely to work from recordings).

Duration can be measured objectively from digitally-captured oscillograms of speech (Jani et al., 2013). However, as noted, computers were not used to measure duration when the standards were developed. Therefore, the more accurate computer-based methods should not be used if users want to reference results to the published standards. If duration measurements cannot be made accurately even though procedures to achieve this are widely available, it should be considered whether duration should be omitted when SSI-3 scores are obtained.

4. PC (physical concomitants).

The reasons why procedures for making a PC score lacked objectivity were touched upon by Todd et al. (2014). Breathing noises, jaw movements, eye movements and other bodily movements are each assessed by the user. A short description is given indicating which of the five scale points is appropriate (the scales and further details how to use these to calculate SSI3 scores are given in Chapter 1 section 1.5.2 p39 onwards. Here essential details and some critical points are considered.). The separate PC estimates for each body area are summed to give an overall PC score. No transformation is applied to the PC score. Todd et al. (2014) noted that the guidance for points one and two are difficult to distinguish ("not noticeable unless looking for it" versus "barely noticeable to casual observer"). Similarly, points three, four and five also seem to overlap ("distracting", "very distracting" and "severe and painful looking"). These difficulties probably explain why Bakhtiar et al. (2010) reported that PC measurement was not reliable. Also, PC cannot be reassessed when audio recordings are used nor when a clinician makes the measurements whilst a client is observed live, which are the formats that were used when the standards were derived. In both cases, no check on PCs can be made unless there is a supplementary video recording, which is not always possible. The inclusion of PC in severity scores because of their reported poor reliability and validity and because of the ambiguous descriptors is a major weakness in the SSI-3 assessment procedure.

4.1.2. Combining the three components: Checks of component distributions and transformations of scores.

Usually, when tests are developed, checks are made as to whether the data are normally distributed. If not, transformations can be applied to make them normal. When more than one component is incorporated into an overall score, each component may need to be scaled so that they all have the same means and variances. The scores can then be combined and a normal distribution should arise. Riley may have needed to transform %SS and duration but not PC to achieve this. This information is lacking, but the conversion tables provide some information on what these transformations achieve.

4.2. Transformations

The raw (x-axis) versus transformed (Y-axis) %SS table (Riley, 1994, p.10) is plotted in Figure 4. Riley's (1994) tables give a single or a range of %SS that converts into one task score value. For example, the non-reader table indicates 6-7 %SS has a task score of 12. The average raw %SS is plotted on the X axis against the corresponding task score on the Y-axis in Figure 4. The task scores are the tabulated values from the non-reader conversion table (x is 6.5% and y is 12 in the preceding example). This shows that up to 7 %SS, an increase of 1-2 %SS leads to a two point increase in the task score. At values above 8 %SS, an increase of 4 % or more is needed to increase the task score by two points. Figure 4 also shows that the functions are monotonic, but not linear.

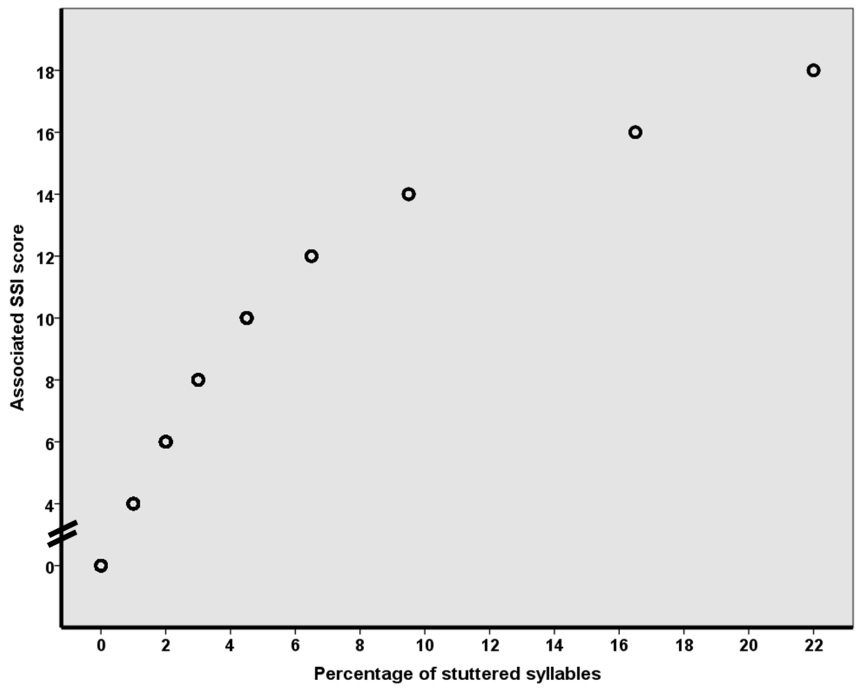

Figure 5. **X-Y Plot for the raw % of stuttered syllables and its associated task score for the frequency component of the SSI-3.**

The non-linearity changes the sensitivity of the severity measure over different ranges of %SS. When fitting a linear and a quadratic curve to the data in figure 4, the R-squared values for both equations are very high (linear= .71, quadratic= .95). The superior fit of the quadratic is because the rate of increase of the task scores levels off as %SS increases. Also, the task scores reach asymptote (do not increase further) above 18 %SS. Both these features give low scores more weight relative to high scores. Sensitivity is lost for high task scores (high %SS) at the expense of giving more sensitivity to low task scores. The transformation may have been deliberately designed to achieve the different sensitivities in different ranges or to give the data a normal distribution.

Similar observations apply to the transformation appropriate for duration. The duration measures are also transformed non-linearly, as shown in Figure 5. Linear and quadratic curves were fitted to the data in figure 5 and once again the R-squared values were both very high (linear= .85, quadratic= .90). The superior fit of the quadratic occurred because a change in a unit of time at short durations increases the overall task score more than the same change at longer durations. An additional concern when duration is measured is that the sample length should have been fixed in terms of number of syllables, otherwise the duration scores (specified in absolute time units) would be problematic. This is because the duration score could vary with the length of the sample. For instance, if SSI-3 scores are obtained on 200 and 400 syllable-long samples, the chance of including longer disfluent syllables is greater in the 400 syllable-long samples.

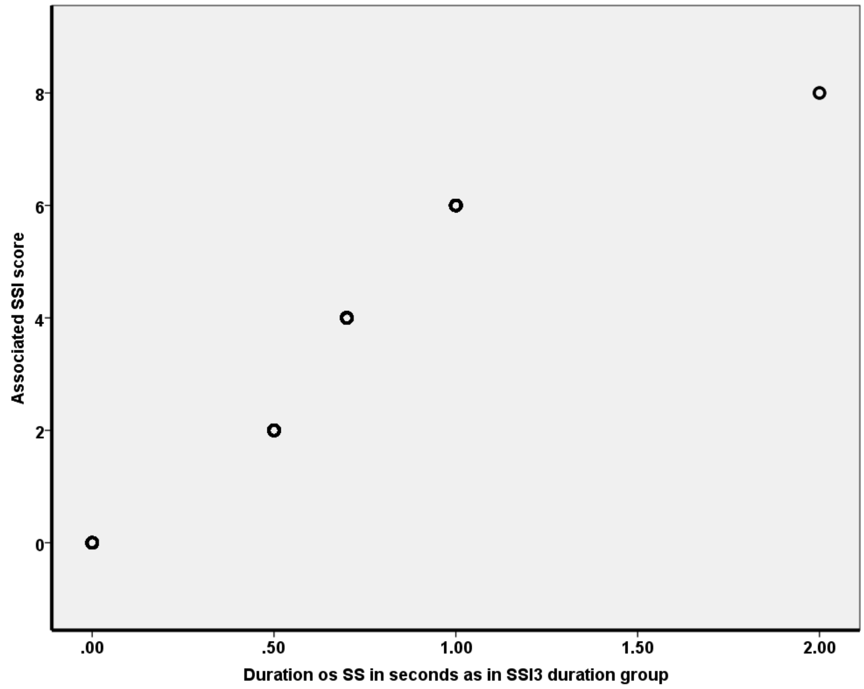

***Figure 6*. X-Y plot for the raw stuttered syllable length in seconds.**

Note: The average is for the three longest stuttered events and its associated task score for the duration component of the ssi-3.

4.2.1 Use of raw physical concomitant scores.

Raw PC scores are added to transformed %SS and duration task scores. Using raw PC scores leads to them having an additive effect across the range of values this measure can take (linear weighting). Taken in conjunction with the levelling off of transformed %SS and duration scores at higher values, PC scores effectively make a bigger contribution to high total overall SSI scores than do %SS and duration scores.

4.2.2. Part-whole correlation analysis as a justification for including all components

Riley (1994) justified the inclusion of all components based on correlation analyses. This was his second test for construct validity (the first was described in section 3.1). Riley (1994) correlated each component score with each of the remaining ones and also with the total overall score. The correlation matrix showed that coefficients were lower when individual component scores were compared (e.g. %SS and duration had a correlation coefficient of 0.37) than when any component score was compared with the total score (e.g. %SS and total overall score had a correlation coefficient of 0.83). The set of correlation coefficients for pairwise comparison of individual components is given in Table 20 and the set of correlation coefficients for individual components with total scores is given in Table 21. Based on these analyses, Riley (1994) argued that all components should be included because “None of the parameters used alone will produce the same severity indications as the combination of all three” (Riley, 1994, p.18). The argument justifying inclusion of all components that Riley offered is fallacious since correlating a measure with another quantity that contains that measure inflates the correlation coefficient (Barry, 1983; Snedecor, 1956). For example, when the frequency measure is correlated with frequency and duration combined the correlation coefficient is .96, when duration is correlated with frequency and duration combined the coefficient is .78 and when frequency and duration are correlated with the overall SSI score, the coefficient is .97 (data from table 27). However, when duration and %SS were correlated, the correlation coefficient was .58, when %SS and physical concomitants (PC) were correlated the correlation coefficient was .40 and when duration and PC were correlated correlation coefficient was .29 (data from table 26). The latter coefficients were even lower when correlations were made using the raw data (not scaled as per SSI-3, table 26).

Table 20. Correlation matrices for %SS, duration and physical concomitants.

Child, non-reader

	%SS	Duration
Duration	.37	
Physical concomitants	.28	.46

Child, reader

	%SS	Duration
Duration	.40	
Physical concomitants	.37	.46

Adult reader

	%SS	Duration
Duration	.41	
Physical concomitants	.29	.45

Note: The separate sections from top to bottom are for children who cannot read, children who can read and adults who can read. The data are re-tabulated from Riley (1994).

Table 21. Correlation matrices for total overall SSI-3 scores and individual component scores (%SS, duration and physical concomitants columns two, three and four respectively).

	%SS	Duration	Physical concomitants
Child, non-reader	.83	.74	.68
Child, reader	.79	.77	.76
Adult reader	.74	.78	.77

Note: The first row is for non-reader children, and the second and third are for readers, children and adults respectively (labelled in column one). The data are re-tabulated from Riley (1994).

An additional feature to note is that Riley used converted scores, rather than the raw score in his correlation analyses. To obtain an SSI-3, raw scores from the frequency and duration components are converted to a standard SSI score. These converted scores were used in the correlations. Previous research has shown that when the raw scores were used for the correlations the agreement results were lower (Lewis, 1995). This could be because a single converted score value encapsulates a range of raw score findings. For example, in the duration component for non-reader samples, the average of the three longest stutters for a score of 8 can range between 2.0 to 2.9 seconds. For this reason Lewis (1995) concluded that a severity indicator on the SSI-3 (e.g. moderate, severe, etc.) could, in fact, reflect a wide range of stuttering behaviours and, therefore, was not a valid or reliable measure of stuttering severity.

4.3. Research Questions

A streamlined procedure is required when large numbers of participants are assessed for fluency (Howell, 2013). Riley's (1994) SSI-3 instrument has several procedural features that commend its use in such applications. This includes its brevity (Todd et al., 2014, Chapter 3) and allowance for assessments to be made with minimal equipment. As seen in the above analysis of SSI-3, it is not clear why some components were included, whether their inclusion serves a purpose in obtaining an SSI-3 score and why the transformations were made. The following research questions address whether all components are needed for applications where children at risk of speech difficulty are identified in schools.

4.3.1. Research question 1: Distribution and transformation of individual components.

The component scores (raw PC, transformed %SS and duration) should be normally distributed with similar means and variances to allow them to be combined to give normally distributed total overall SSI-3 scores. Whether the distribution of components meet these requirements is examined on a sample of UK children entering education. To assess the effects of the transformations on %SS and duration, the distributions of unselected school children's scores were obtained after Riley's (1994) transformations were made. For PC, the distribution of raw scores was obtained (Riley does not given a transformation for PC). The distributions of all components and overall SSI-3 scores were assessed for normality and to obtain their mean and higher moment statistics. Then the distributions of the individual components were compared with the distribution of the overall SSI-3 scores. Together, these results provide an indication of whether the data were processed appropriately and indicate any problems introduced when the components were combined.

4.3.2. Research question 2: Scatter plots to ascertain how cutoffs based on individual components compared to cutoffs based on overall SSI-3 scores.

A fixed threshold of 16 was applied to overall SSI-3 scores to separate the children who were fluent from those who were deemed to be at risk of fluency problems (Mirawdeli, 2015, Chapter 5). This assumes that SSI-3 gives unambiguous case classification. To see how individual components compared, the selected component was plotted on the X axis, and overall SSI-3 scores on the Y axis. These plots show how well that component classifies children with respect to fluency relative to the overall SSI-3 score standard (i.e. how well the participant groups were separated by the individual component compared to the overall SSI-3 score). Contingency tables (classification based on SSI-3 score versus classification based on the individual threshold) were obtained for all children, including statistics to establish classification performance and to compare classifications when different components were used.

Riley (1994) indicated that SSI-3 can be used in the differential diagnosis of fluent children and those at risk of speech difficulty. This suggests that there is a threshold SSI-3 score that can separate fluent children from the rest. However, Riley (1994; 2009) did not specify what this threshold should be, as mentioned in section 1.1. Table 3 in Riley (1994) shows that 6 is the lowest SSI-3 score that has an associated severity estimate (a label of ‘very mild’). This may indicate that a child who scores below 6 should be considered fluent. Subsequent estimates were made by Howell and Davis (2011), Howell (2013) and Mirawdeli (2015), who proposed scores of 8, 13 and 16 respectively as the threshold SSI-3 value for the fluent/at risk cutoff. These thresholds are higher than that deduced from Riley (1994), probably because sensitive computer-based procedures were used to obtain SSI-3 in these studies. More non-fluent events are detected when computer-based methods are used compared to other methods (Jani et al., 2013). Therefore, the threshold for classifying a child as fluent needs to be higher when computer-based methods are used.

4.3.3. Research question 3: Part-whole correlation, and related analyses.

Riley’s (1994) part-whole correlation analysis was conducted on the present data. Although these analyses also showed that correlations between individual component scores were lower than those between each individual component and the overall score, as pointed out above, correlating a score with something that includes that score is fallacious. To assess this problem, correlations of individual components with the remaining components were computed.

4.3.4. Research question 4: Evaluation of performance using different numbers of scale values for the PC component.

The previous literature indicates that PC is probably not a useful component of SSI-3 severity estimates. To check this, analyses were made to ascertain whether the scale descriptors Riley gives can be used consistently. The PC measured using Riley's five-point scale was compared with PC obtained when the scale was collapsed in different ways. The PC scale was collapsed into two or three categories that were suggested by which scale points seemed to overlap, as reported by Todd et al. (2014). The questions are whether or not the two- and three-point scales give the same results as Riley's (1994) five-point scale. If the five point scale provides a better estimate than two or three point scales, better correlation coefficients should occur between PC measured this way and the other components of SSI-3 than those found when the two or three point scales are used.

The research questions provide indications as to whether components of SSI-3 can be dropped. As indicated above, PC and duration could potentially be dropped. The research questions were examined in an unselected sample of reception class children.

4.4. Method

4.4.1. Participants.

Children from the reception classes of 11 London and Ipswich schools were assessed on SSI-3 in the academic years 2012-2013 and 2013-2014. See appendix A, B and C for the materials sent to schools, before and after the screening. There were 879 children in total (448 males and 431 females); 315 were aged four years, 562 were aged five years and two were aged 6 years. Approximately 57% of the children spoke English as their first language (native language), 9.2% did not disclose language information and 34.2% had English as an Additional Language (EAL). These data are presented in the pie chart in Figure 6.More details about the children are given in the method of Chapter 5.

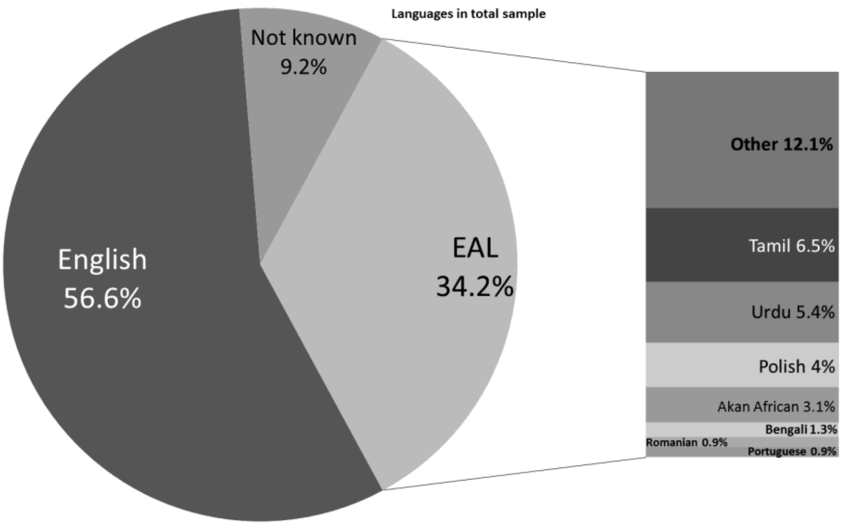

***Figure 7*. The distribution of English, undisclosed and children with EAL in the sample (pie chart at left).**

***Note*: The frequency of different first languages of the children with EAL is given on the right in the bar chart.**

Table 22. Age distribution (top section) and gender (bottom section) of the 879 participants.

AGE		
in years	**Number**	**%**
4	315	35.8
5	562	63.9
6	2	0.2
Gender		
Male	448	51
Female	431	49

Note: Statistics are given as numbers (second column) and percentages (third column).

4.4.2. Recordings.

Parents' consent was obtained before recordings of the children took place, see appendix D, E, F and G for the material that was sent to the parents and guardians. Picture stimuli from the SSI-3 manual were used to elicit speech. Topics that children could readily talk about were also employed. Examples of the latter included favorite TV programs, hobbies, days out and school activities. All speech samples were recorded using a Zoom H4N recorder with an internal microphone. The samples were 10 to 15 minutes long and were in English irrespective of the child's first language. PCs were scored by the first author at the time the recording was made according to the procedure given in Riley (1994).

4.4.3. Scoring %SS and duration.

All audio recordings were uploaded and annotated using Speech Filing System (SFS) software (Huckvale, 2013). SFS allows extracts to be selected and played whilst judgments were made. The numbers of syllables spoken in an extract were counted both for those that were fluent and for those that were not according to the guidelines in Riley (1994). Duration of the three longest stutters was measured from the oscillographic display of the speech sample using SFS.

4.4.4. Scoring PC using different numbers of scale points.

The PC data were pre-processed so that they were scored on three different scales. 1) five point scale (Riley, 1994); 2) two point scale where points one, two and three on Riley's (1994) scale were assigned to new scale point one, and points four and five on Riley's (1994) scale were assigned to new scale point two; 3) three point scale where points one and two on Riley's (1994) scale were assigned to new scale point one, points three and four on Riley's (1994) scale were assigned to new scale point two and point five on Riley's (1994) scale was assigned to new scale point three. The collapsing for the new scales (2 and 3) was based on the ambiguous descriptions concerning scale points discussed above (Todd et al., 2014).

The %SS, duration and PC components were used to calculate SSI-3 scores according to the non-reader procedure given in Riley (1994).

4.4.5. Reliability.

Two trained transcribers independently assessed the speech materials from eight speakers chosen at random. Inter-judge reliability for non-fluencies was 96% and for syllables was 93%. The associated kappa coefficients were 0.92 and 0.89, which represents agreement well above chance (Fleiss, 1971). The durations of the three longest non-fluencies were identical to within 5 ms for the two judges. This was because the SFS has a calibrated time display on the x-axis of the oscillographic display. The average durations of the three longest stutters for the eight speech samples did not differ significantly by *t* test across the judges. Inter-judge reliability estimates were not performed as they were not available for one of the components, PC, because only the first author was present when the recordings were made.

4.5. Results

4.5.1. Research question 1: Distribution characteristics of each component (%SS, duration and PC) and their appearance in the distribution of overall SSI-3 scores.

The frequency scores after conversion using Riley's (1994) tables and conversion to *z* scores, are plotted in Figure 7 (standardized scores on the X axis and frequency counts on the Y axis). Figure 7 shows that the %SS scores are positively skewed and that the transformed scores peak approximately at the mean. The duration scores after conversion are plotted in a similar way in Figure 8. This shows that the scores have a positive skew and the transformed scores have a peak below the mean. The raw PC scores were plotted in a similar way to frequency and duration scores in Figure 9. This shows the scores are not normally distributed. Low scores dominated and there is a long tail through to high values (positively skewed). The distribution of the overall scores is given in Figure 10. The overall scores are not normally distributed. The distribution has the separate features noted for each of the components. Thus, there is a peak approximately at the mean that reflects that seen in the distribution of transformed symptom %SS scores (Figure 7); There is a second peak below the mean that occurred in the transformed duration scores (Figure 8); There is a long positive tail as observed for the raw physical concomitant scores (Figure 9). Combination of the components according to Riley's (1994) procedure does not yield a distribution that could be used to provide standard severity estimates.

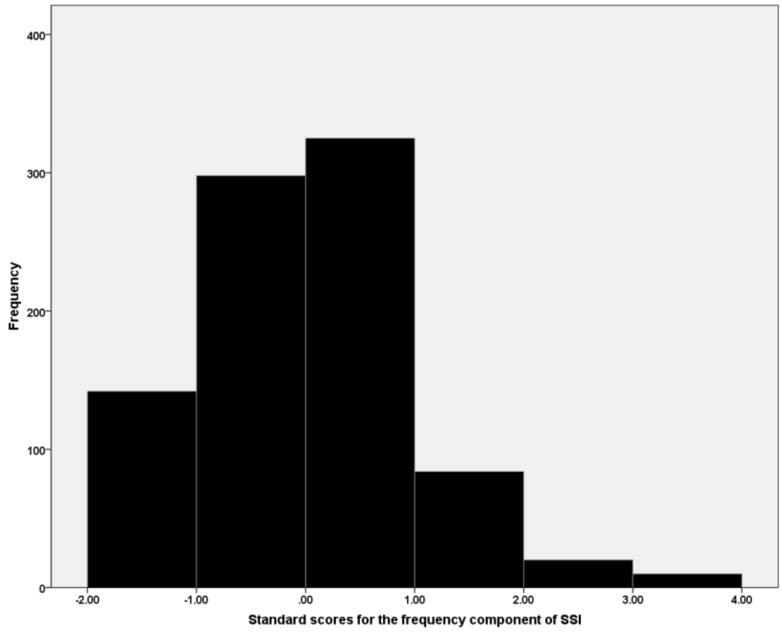

***Figure 8*. Distribution of frequency scores after conversion using Riley's (1994) table.**

***Note*: Standard scores are plotted on the x axis and counts are given on the y axis.**

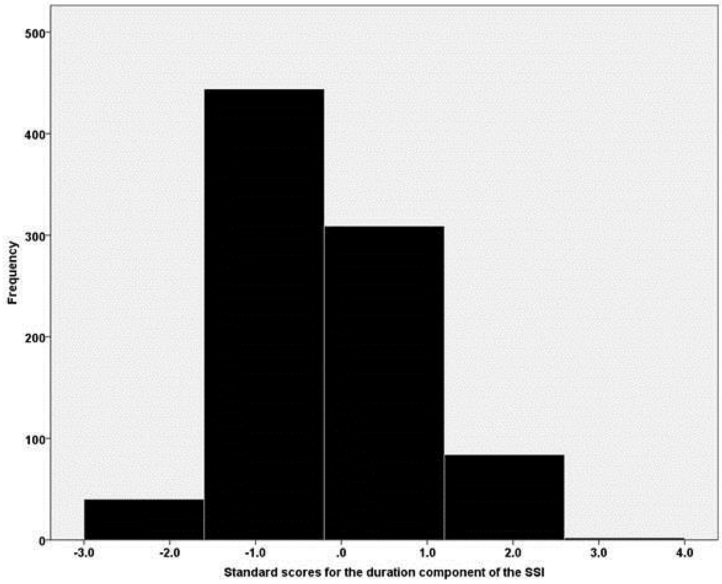

Figure 9. **Distribution of duration scores after conversion using Riley's (1994) duration table.**

Note: Standard scores are plotted on the x axis and counts are given on the y axis.

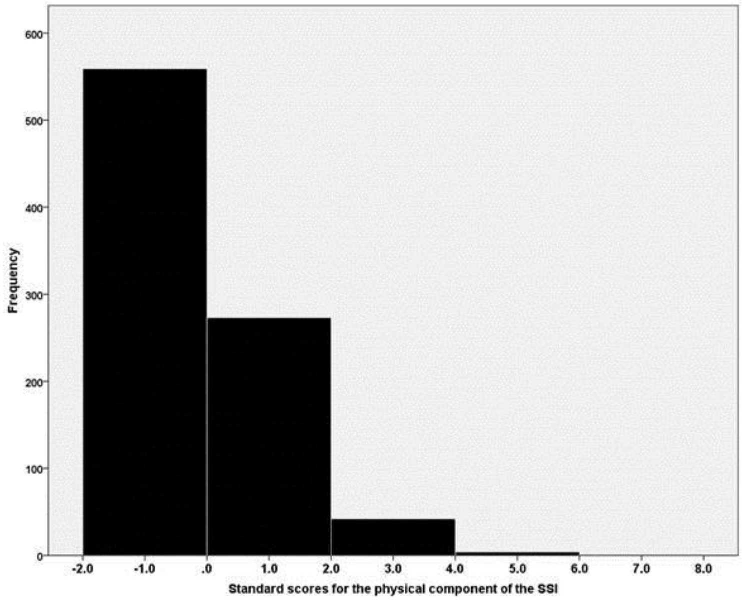

***Figure 10*. Distribution of raw PC scores according to Riley (1994).**

Note: Standard scores are plotted on the x axis and counts are given on the y axis.

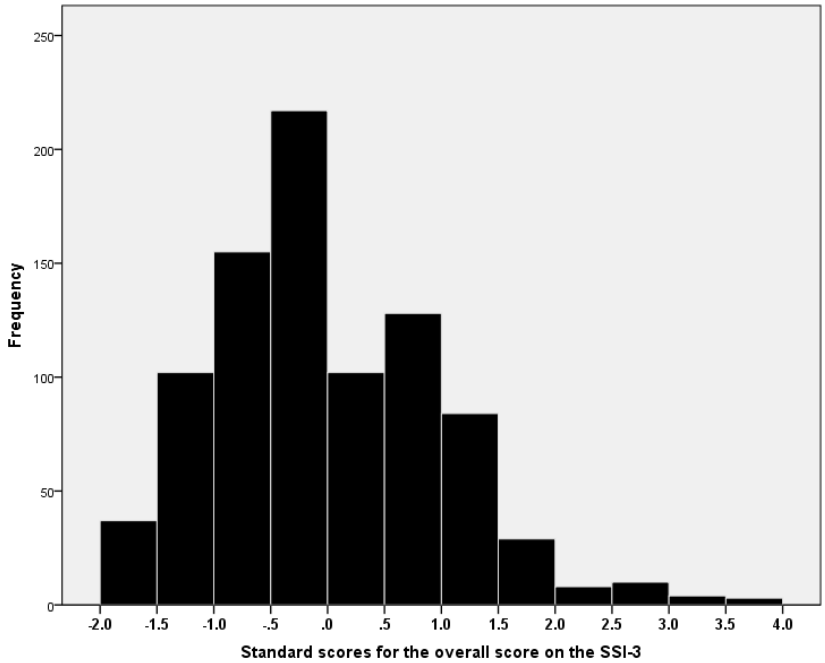

***Figure 11*. Distribution of overall SSI-3 scores using Riley's (1994) procedure.**

Note: Standard scores are plotted on the x axis and counts are given on the y axis.

Skewness and kurtosis statistics were obtained to quantify the deviation of the distributions from normality. Skewness reflects asymmetry about the mean, kurtosis the peakiness of the distributions. After *z* score conversion, a set of scores that are exactly normally distributed would have skewness and kurtosis values of zero. The statistics given in table 23, show that none of the distributions of either the individual components or the overall scores were an acceptable fit to a normal distribution. Table 23 also shows that the distributions have different means, which precludes them from having a normal distribution when they are combined.

Table 23. Descriptive statistics for SSI-3 components with and without tabular conversion (where appropriate) and total overall SSI-3 scores.

	Frequency	Duration	Phys. Con.	Overall
Mean	5.27	0.71	0.78	9.11
Skewness	0.44	28.58	2.28	0.72
Kurtosis	0.815	836.56	6.61	1.70

4.5.2. Research question 2: Scatter plots to ascertain how cutoffs based on individual components compare to cutoffs based on overall SSI-3 scores.

It is known that an SSI-3 threshold provides reasonable results for identifying children at risk of speech difficulty (Howell, 2013; Mirawdeli, 2015). The following analyses evaluated how well each raw individual SSI-3 component performed relative to overall SSI-3 score. Mirawdeli's (2015) cutoff value of 16, which was used to separate children at risk of speech difficulty from those who were fluent, was applied here to classify children as fluent or not. There were occasional exceptions where children were deemed at risk even though their SSI-3 scores were below 16 (see 5.2.3 for a full description of how schools classified children as at risk of speech difficulty). These cases were classified by the schools as at risk (these were usually cases where children spoke very little). Scatter plots are given in Figures 11-13 each with one raw SSI-3 component (%SS, duration and PC respectively) given on the X axis and overall SSI-3 score on the Y axis. Individual children's scores are shown as circles (at risk) or squares (fluent). The built-in cutoff criterion ensured overall SSI-3 scores (Y axis) separated the two groups almost perfectly (except for the cases mentioned). Examination of the distribution along the X axis indicates how well groups were separated when individual raw component scores were used.

To quantify how well each of the components classified children in comparison with a threshold value of 16 applied to overall scores, classifications were compared with the fixed SSI-3 score setting and each component for a range of values on the individual components. The results are summarized in Table 24. Number of cases for overall SSI-3 classification and component by threshold classification are given in columns three to six and sensitivity and specificity in columns seven and eight. The cutoff frequency of 3% is often used on percentage speech-like disfluencies (Yairi & Ambrose, 2015). However, the threshold value of 5%SS was superior (100% sensitivity and 89% specificity). Sensitivity never reached 80% for the duration threshold (80% sensitivity and specificity are conventional values for satisfactory classification of medical cases).

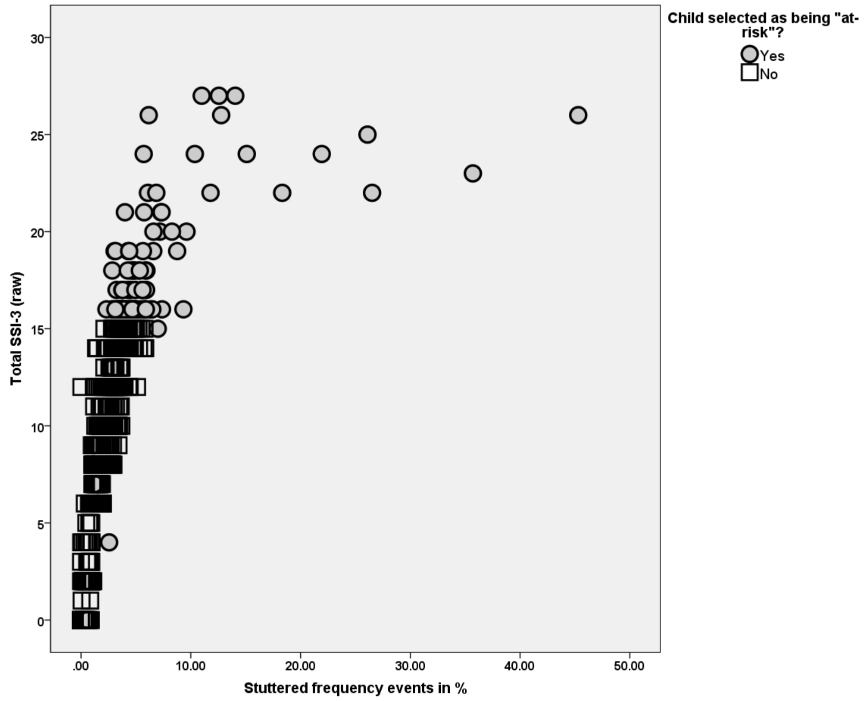

***Figure 12*. Scatter plots of raw %SS (frequency) scores (x axis) against overall SSI-3 scores (y axis).**

Note: Circle symbols were used for plotting cases where children were at risk (ssi-3 score of > 16) and squares for children deemed fluent.

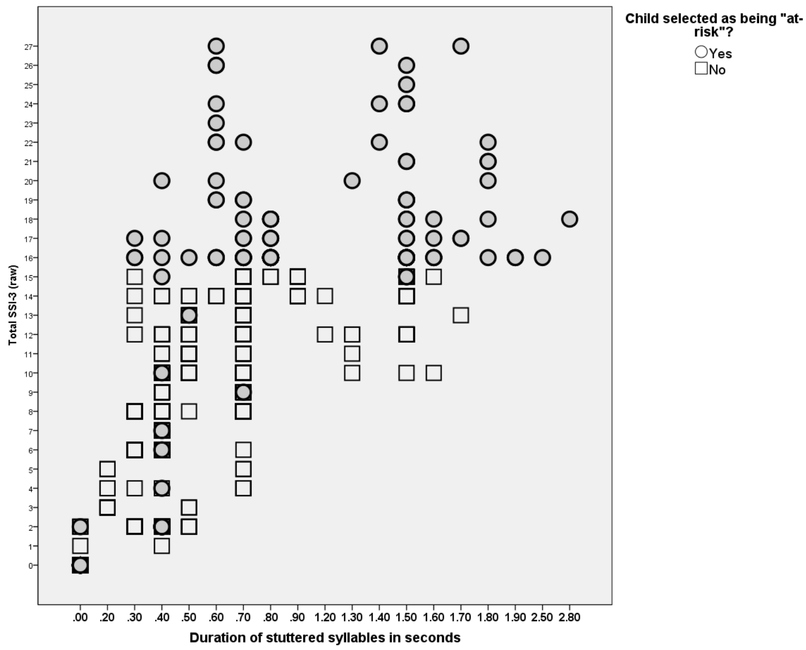

***Figure 13.* Scatter plots of raw duration scores (x axis) against overall SSI-3 scores (y axis).**

Note: Circle symbols were used for plotting cases where children were at risk (SSI-3 score of > 16) and squares for children deemed fluent.

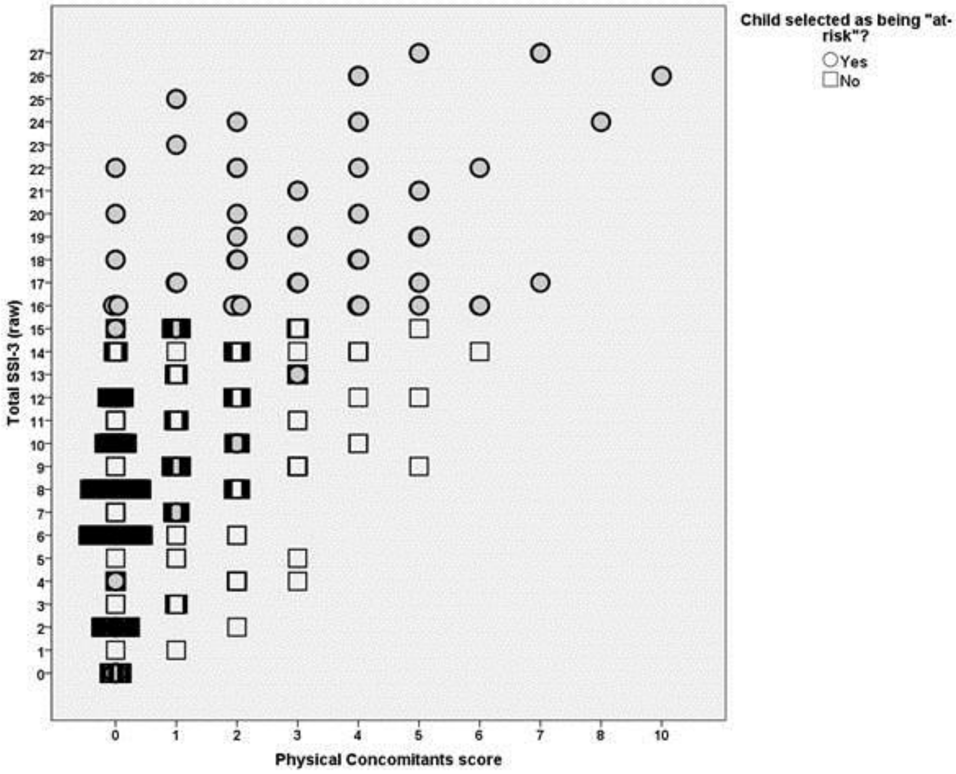

***Figure 14*. Scatter plots of raw physical concomitant scores (x axis) against overall SSI-3 scores (y axis).**

Note: Circle symbols were used for plotting cases where children were at risk (SSI-3 score of > 16) and squares for children deemed fluent.

Table 24. Classification results comparing those obtained when a threshold was applied to individual SSI-3 components against a threshold that used a value of 16 on overall SSI-3 score.

Component	Threshold	+/+	+/-	-/+	-/-	Sensitivity	Specificity
Frequency	2%	98	25	129	627	43.1%	96.2%
"	3%	75	48	38	718	66.3%	93.7%
"	5%	30	93	0	720	100%	89%
Duration	0.6s	82	41	265	491	23.6%	92.5%
"	0.9s	50	73	36	720	58.1%	90.8%
"	1.2s	95	73	36	720	72.5%	90.8%
PC	1	78	45	119	637	39.6%	93.4%
"	2	47	76	32	724	59.5%	90.5%
"	3	37	86	10	746	78.7%	89.7%
"	4	19	104	4	752	82.6%	87.9%

Note: The first symbol in the labels +/+, +/-, -/+ and -/- refers to classification according to the overall SSI-3 score of 16 (+ = has speech difficulty, - = does not have speech difficulty according to this criterion). The second symbol in the labels +/+, +/-, -/+ and -/- refers to classification according to the threshold applied to a component of SSI-3. The component can be identified by the label in column one and the threshold in column two. For example, the 2% frequency threshold means that any child scoring up to 2%SS is designated as not having speech difficulty, whereas any child scoring above 2%SS is designated as having speech difficulty. +/+ and -/- are correct identification of children with speech difficulty and fluent children respectively. Columns 7 and 8 give sensitivity and specificity statistics respectively.

Using duration rather than frequency would lead to more children being referred (as pointed out by Dr Frisson: the frequency component is about 30% more stringent than the duration component). This is true when absolute counts are examined. However, when statistical analysis was performed, the sensitivity did not reach the statistically agreed upon threshold of 80%. The highest achieved was 72.5% and this was only when considering the longest possible disfluency of 1.2 seconds. Furthermore, the total 168 children combines those children that were selected as being disfluent by AM and agreed upon by the schools as well as those selected by AM but not agreed upon by the schools but ignoring the children that the schools picked up but were not corroborated by AM (36 in total). Therefore, duration is deemed to be less sensitive than frequency for these data.

4.5.3. Research question 3: Part-whole and correlations and related analyses.

Riley (1994) justified inclusion of all components based on the part-whole correlation analyses reviewed in the introduction to this chapter. It emerged, when the data distributions were examined, that the overall score deviated from normality. Strictly speaking, Pearson product moment correlations require normally-distributed data as it is a parametric test. However, Riley’s results are replicated for completeness, followed by ancillary analyses.

4.5.4. Replication of Riley's (1994) part-whole correlation analysis.

Table 25 gives the correlations between total overall scores and %SS, duration and physical concomitant scores. The values ranged from .57 to .93 (all were significant by parametric and non-parametric tests and for raw and transformed data, N= 879, *df* = 877 $p < .001$ in all cases. Table 25 gives the correlation matrix for the three components of SSI-3. These analyses replicate the results in Riley's (1994) report (i.e. the coefficients in Table 25 are higher than those in Table 26).

The correlations between each individual component with each of the two remaining components (non-overlapping components) are given in Table 27 for the two age groups (i.e. children of reading age and non-reading age children). Table 27 shows that the correlations between pairs of components and overall SSI-3 scores, where overlap occurs, were high (e.g. .97, .97 & .85; column five). Correlation coefficients along the ascending diagonal in columns two to four were lower than those for the other cells in these columns, for example duration and physical concomitants combined correlated with frequency .61, frequency and physical concomitants combined correlated with duration .57 and finally, frequency and duration combined correlated with PC .41. These low values represent cases where components were not shared between the components that were correlated. The correlation coefficients in the cells that are off the ascending diagonal all involve correlations where a component was shared. This should increase the coefficients. This was confirmed as the correlation coefficients in the ascending diagonal are always lower than in the off-diagonal cells. The lower correlations, when individual components are compared with the remaining components, is not subject to this problem, implying that not all components may need to be included. It is also noteworthy that PC showed the lowest correlations (the next section investigated whether the procedure for measuring this has any objectivity).

Table 25. Correlations between total overall scores and the component scores (%SS, duration and PC columns two, three and four respectively) for non-reader children.

	%SS	Duration	Physical concomitants	*P*	*N*
Scaled parametric	.93	.75	.62	p<.01	879
Scaled non-parametric	.92	.77	.57	p<.01	879
Raw parametric	.68	.69	.61	p<.01	879
Raw non-parametric	.90	.74	.57	p<.01	879

Note: The data were scored as transformed and raw scores using parametric and non-parametric tests (the combination of type of score and type of test used are given in column one).

Table 26. Correlations matrix between individual SSI-3 components (%SS, duration and PC).

	%SS		Duration		N
Duration	.58	p<.01			879
Physical concomitants	.40	p<.01	.29	p<.01	879

Scaled non-parametric

	%SS		Duration		N
Duration	.58	p<.01			879
Physical concomitants	.39	p<.01	.31	p<.01	879

Raw parametric

	%SS		Duration		N
Duration	.34	p<.01			879
Physical concomitants	.33	p<.01	.25	p<.01	879

Raw non-parametric

	%SS		Duration		N
Duration	.57	p<.01			879
Physical concomitants	.40	p<.01	.30	p<.01	879

Note: The separate sections from top to bottom are scaled parametric, scaled non-parametric, raw parametric and raw non-parametric.

Table 27. Correlations of pairs of components labelled in the left column with the remaining individual component (%SS, duration and PC)

Parametric test				
Parameter	**Frequency**	**Duration**	**P.C.**	**Over-all SSI score**
Frequency + duration	.96 p<.01	.78 p<.01	.41 p<.01	.97 p<.01
Frequency + P.C.	.95 p<.01	.57 p<.01	.66 p<.01	.97 p<.01
Duration + P.C.	.61 p<.01	.83 p<.01	.78 p<.01	.85 p<.01
Nonparametric test				
Parameter	**Frequency**	**Duration**	**P.C.**	**Over-all SSI score**
Frequency + duration	.92 p<.01	.80 p<.01	.40 p<.01	.97 p<.01
Frequency + P.C.	.92 p<.01	.58 p<.01	.62 p<.01	.95 p<.01
Duration + P.C.	.63 p<.01	.87 p<.01	.72 p<.01	.85 p<.01

Note: where the individual component can be identified by the labels given in columns two, three and four, respectively. Column five gives the correlations between the pairs of scores and overall SSI-3 scores (*N*=879 in all cases).The top and bottom sections give parametric and non-parametric statistics respectively.

4.5.5. Research questions.4: Evaluation of performance using different scale values for the PC component.

Parametric and non-parametric correlations were performed between PC scores obtained according to the three scales (five, two and three point) and the remaining two SSI-3 components (%SS and duration). The results are given in Table 28. Since the distribution of all components deviated significantly from normality, the parametric coefficients (in brackets) are not appropriate but are given for completeness. All correlation coefficients were significant as N was large (879). The principal issue was whether the five point scale provides useful information, bearing in mind the confusion in the scale-point descriptions. Hence, the pattern of correlations across the three forms of scale and between PC versus %SS or duration were important. Correlations between PC and %SS were better than those between PC and duration whatever scale was used. This underlines, once more, the status of %SS in characterizing severity. The main feature to note in the comparisons is that there was little difference across the different PC scale formats. The non-parametric coefficients for the correlations between PC and %SS ranged from .386 to .391 across scale formats. The non-parametric coefficients for the correlations between PC and duration ranged from .315 to .316 across scale formats. Thus it appears that the correlation patterns were not sensitive to the changes in the imposed scale values. Therefore, in terms of the prediction for this research question, the five scale-points did not provide a better estimate than two or three point scales.

Table 28. Correlation coefficients for comparison between PC (column one) and: 1) %SS (column 2); and 2) duration (column 3).

	%SS	Duration
PC as Riley (five point)	.391** (.401**)	.315** (.122**)
PC.a. (1,2&3 = 1/ 4&5= 2) (two point)	.386** (.401**)	.316** (.320**)
PC.b. (1&2=1; 3&4= 2; 5=3) (three point)	.386** (.399**)	.316** (.302**)

Note: The three rows are for PC measured according to the different scale points indicated in column one. The first coefficient is non-parametric and the coefficient in parentheses is Pearson's *r* (parametric value). All correlations were significant $p < .01$ (indicated by **).

4.6. Discussion

The main goal of this chapter was to evaluate whether all SSI-3 components are necessary when Riley's instrument is used to identify children starting schools who have speech difficulties. The implications that the findings have for this process are considered. The process of identifying children with speech difficulty and the role a speech assessment has in it are discussed. Finally, implications that the current assessment of SSI-3 has for other situations where the instrument is employed are considered.

4.6.1. Should all SSI-3 components be included when Riley's procedure is used to identify children starting in schools who have speech difficulty?

Frequency, duration and PC components of SSI-3 were not normally distributed. %SS peaked around zero after *z* transformation and had a positive skew. The transformed duration scores after *z* transformation peaked below zero and again had a positive skew. The raw PC scores were not normally distributed. Most PC scores were zero and extended to positive values and there was a long positive tail on the distribution. The deviations from normality and positioning of some of the modal values away from zero after *z* transformation implies that combining the components would not lead to a composite that was normally distributed. This was confirmed as the distribution of overall scores after *z* transformation had the features noted in the distribution of the separate components.

These observations on the distributions question the appropriateness of combining the components of SSI-3 to yield unambiguous interpretations of an individual's fluency level (Lewis, 1995). One solution would be to drop any component(s) that showed marked departures from normality. The component with the closest match to a normal distribution that also had a peak near zero after *z* transformation was %SS. However, these observations alone are not sufficient grounds for only using that component in identifying children at risk of speech difficulty as %SS may not achieve satisfactory classification of children. To address whether any of the components on their own provided a basis for classifying children as compared to overall SSI-3 scores, scatter plots were made between scores on each component and the overall SSI-3 scores. The overall SSI-3 score has been successfully applied to identifying children with speech difficulty elsewhere (Howell, 2013; Mirawdeli, 2015). Classifications based on overall SSI-3 scores were evaluated by establishing the degree of correspondence with children about whom schools had expressed concern. Although SSI-3 classifications correspond well with schools' judgments about children, they do not provide a 'gold standard'. Thus, it is possible that any of the components alone could provide better case-classification in some instances than overall scores (although this is unlikely given the close correspondence with schools' judgments). Case classifications were evaluated by applying a threshold SSI-3 score of 16 (Mirawdeli, 2015) to overall scores, and a range of thresholds was applied for each component. A 3% threshold (Yairi & Ambrose, 2005) applied to %SS produced classification with good sensitivity and specificity. However, a 5% SS threshold was even better (100% sensitivity and 89% specificity). When similar analyses were applied to duration and PC, they did not achieve good classifications relative to the score of 16 applied to overall SSI-3 (sensitivity and specificity were lower than for %SS). Thus overall, %SS not only had the most appropriate distribution of scores, but a threshold value applied to this component alone reached satisfactory levels in classifying children as fluent or not (relative to that achieved by an SSI-3 score of 16).

The classification analysis suggests that %SS alone contributed to the success of overall scores for identifying the fluent children. The implication is that duration and PC could be dropped. Research questions three and four also addressed whether each of the components could be dropped, but in different ways. Research question three replicated and extended Riley's (1994) part-whole correlation analyses. The implication that the part-whole procedure led to spuriously high correlations was confirmed (Barry, 1983; Snedecor, 1956). Therefore, Riley's justification for including all three components was not appropriate. PCs are the most problematic of the components (Bakhtiar et al., 2010; Lewis, 1995; Todd et al., 2014) and were singled out for evaluation in research question four. PC scores did not change their pattern of correlations with either of the remaining components when the scales were collapsed to two or three points. This lack of sensitivity confirmed the concerns expressed about how PC are measured in SSI-3. On this basis and in the light of other reports (Bakhtiar et al., 2010; Lewis, 1995; Todd et al., 2014), PC should be omitted or their measurement improved in future revisions of SSI-3.

4.6.2. Comments about identifying children with speech difficulty in general.

Teachers are in contact with children from the time they enter school and are sensitive to their needs and problems including any speech difficulties. They want a systematic method that can be used to assess children's speech (Dockrell & Howell, 2015, Chapter 2). Schools also report that any proposed tool has to be applicable for use with children with EAL (Mirawdeli, Dockrell & Howell, in prep, chapter 2).

SSI-3 was chosen as the preferred contender as a starting point for such a tool as it has a short, efficient and clear procedure. In addition basing assessments on audio samples maintains children's confidentiality and it does not require children being able to read (Todd et al., 2014). Any further simplifications of the procedure would be welcomed by schools. In the light of this, the present work confirmed that PC and duration can be dropped without affecting performance of a revamped procedure.

Riley's (1994) symptoms focus on breaks, repetition of parts of words or prolongation of the initial part of words, but excluded repetitions of whole words (Howell, 2013). Two features to highlight are: 1) that the fragmentary disfluencies Riley selected are appropriate for identifying stuttering and other pediatric speech difficulties (Howell, 2013; Mirawdeli, 2015); 2) that the exclusion of whole-word repetitions allows them to be used to identify other forms of difficulty, namely word-finding difficulty. When children cannot find a word, they use hesitation phenomena such as whole-word repetition (Fehringer & Fry, 2007). Word-finding difficulty is a particular problem in the speech of children with EAL (Bada, 2010). Riley’s procedure effectively does not confuse children with EAL who have word-finding difficulty but are otherwise fluent with those who have speech difficulty - a desirable feature for schools.

Schools would use a procedure that is based on SSI-3 to identify a range of pediatric speech difficulties, not stuttering alone as they are not primarily interested in what form of speech difficulty a child has (which is a matter for SLTs). That said, a suitable instrument should include a set of symptoms that is appropriate for identifying all forms of speech difficulty. This may call for an extension of the symptom set, which would require redesign of SSI-3 so that the symptoms are comprehensive. When this is done, it is important that whole-word repetitions are left out from the speech difficulty set, allowing them to be used to identify word-finding difficulty separately. As mentioned above, duration and PC could be left out at this stage too in order to make the teacher’s task simpler (at present these measures do not appear to fulfil a role in assessing children’s fluency).

Identification using an objective speech-based procedure is desirable to schools, but they take other factors into account when deciding whether to refer a child for intervention (the main factor being educational attainment, Chapter 2). It is often not recognized that these other factors schools take into consideration bias the samples of children that present to SLTs. Work is taking place on improving co-ordination which will make both sides aware of features like this so that they can be incorporated into feedback between schools and SLTs (Bercow, 2008, Mirawdeli et al.,, in prep, Chapter 2).

4.6.3. Relevance of results for wider applications of SSI-3.

A problem with the above proposal on the next stage in development of a suitable instrument to identify children with speech difficulty in schools is that if the suggestions were adopted, they would distance the procedure from ones which SLTs would want to implement in any revised version of SSI-3. For instance SLTs would want to upgrade the recording format from audio to audio-video and take longer and more representative samples (Mirawdeli et al., in prep, Chapter 2).

There are some common considerations that apply to teachers and SLTs. PC, and possibly duration, should be dropped in new forms of the instrument. This would call for re-standardization. Whilst re-standardization is taking place, some attention should be given to examining distributions, tabular conversions and what they achieve, and combination of components as was done here. Also, consideration of ways in which teachers' and SLT's assessments could maintain some compatibility is essential. The syllable and disfluency count procedures used in SSI-3 are appropriate and should be retained. The exclusion of whole-word repetitions also seems desirable, although these should be counted separately for the information they provide about word-finding difficulties. If a measure of PC is considered essential, a better option might be to use one based on the premonitory urge for tics scale (Woods, Piacentini, Himle & Chang, 2005).

If it is desirable for teachers and SLTs to coordinate their activities (Bercow, 2008; Mirawdeli et al., in prep) it would help if they used scores that were translatable. One possibility would be to embed %SS counted as indicated above into all forms. SLTs could transform the school results to interpret them in their own terms and vice versa. For instance, transforming of %SS and duration in the current form of SSI-3 squashes the scale at the end where teachers score most children, the majority of whom are fluent (least sensitivity where it is most needed). A different transform might be used in schools from the one used in clinics. Software could easily be written that converts scores from one form to another if the scores only differ in respect of the transform that was used. Similar adaptations could be made for versions with and without additional components. If a child is referred to clinicians, the appropriate %SS or SSI-3 score could then be recalculated.

4.6.4. Limitations.

The application to identification of children at risk of speech difficulty is an important, but not a major use of SSI-3. SSI-3 needs an equivalent set of performance evaluations in clinical settings to those conducted here. This should examine distributions of scores, the roles of all components etc. Decisions as to whether any of the components can be discarded for clinical application could then be made.

Cases of speech difficulty were identified by teachers. At minimum, this is not as thorough as SLT's methods of assessing a child. Presently, SLTs do not take educational attainment into account in deciding on interventions, but teachers do so when identifying children for intervention. The impact of educational attainment on assessments needs further examination.

The children in schools had a range of speech problems, although stuttering was most prevalent. Other speech features may need to be taken into account. If the issue of revising symptom is addressed, new goals of identifying children with speech difficulty should be set and appropriate symptom sets defined depending on these goals. For example, is the procedure intended to identify stuttering alone or a range of pediatric speech difficulties? The issue of using whole-word repetitions to identify word-finding difficulty in children with EAL needs further attention. Some of these issues are addressed in Chapter 6.

4.6.5. Conclusions.

SSI-3 has many features that make it suitable for use in schools (see also Chapter 3), and a procedure for assessing speech difficulty is desirable to schools (Chapter 2). It can be simplified by using %SS alone. These alterations would mean that the instrument has to be re-standardized. Schools would want many features present in SSI-3 retained. Ironically, SLTs would want to change some of these. It is possible that any SSI-3-based procedure for schools would diverge from clinical forms of SSI-3. This would undermine coordination between school and Speech and Language Therapy services. Alternatively, the importance of schools as a source of referral (Mirawdeli et al., 2015) may require retention by SLTS of at least some of those features that schools would like to keep. If this is the case, some way of converting scores between schools' and SLT's results would be desirable.

Chapter 5

Identifying Children Who Stutter or Have Other Difficulties in Speech Production in School Reception Classes[5]

5.1. Introduction

In this chapter, the effectiveness of the SSI-3 in selecting children, who are “at-risk” of speech difficulties, is examined. SSI-3 is used with all its components (frequency, duration and PC). The results of chapter four suggest that an evaluation with %SS alone would be successful. However, before this is done, the issue whether the symptoms used in the frequency component need extending so that they are applicable to a wider range of pediatric speech difficulties needs examining (work has commenced on this and is reported in Chapter 6).

This chapter used Riley’s (1994) Stuttering Severity Instrument, third edition (SSI-3) to identify children who may be at risk of speech difficulties in a cohort of typically-developing children. However, the intention was not to categorize the children into severity classes (mild, moderate, severe) which is

[5] A version of this chapter is published as: Mirawdeli, A. (2015). Identifying children who stutter or have other difficulties in speech production in school reception classes. *Procedia - Social and Behavioral Sciences*, *193*, 192-201.

the main application of the SSI-3, but to screen children for fluency as proposed in Howell and Davis (2011).

Clearly the timing constraints imposed by the schools and the high number of children that need to be assessed place considerable constraints on what can be performed when conducting a screen. For instance, full clinical assessments, that are appropriate when speech or hearing problems are suspected, are not feasible to conduct if all children in schools are screened. On the other hand, any procedure that is used needs to be backed up by scientific research and results that indicate that it meets required levels of performance. As discussed previously, SSI-3 seems a reasonable starting point as it is based on research studies that show it can distinguish between children who are fluent and those with one form of speech difficulty (stuttering). Given the high rates of comorbidity (Yairi, 2007) between stuttering and other speech difficulties, it seem likely that SSI-3 symptoms would identify other types of speech difficulties too. Looked at from another perspective, conducting and analyzing SSI-3 bears some resemblance to what teachers do informally when monitoring children for speech difficulty. They have ancillary information in some cases (e.g. information from parents and information about educational attainment) that they use in making decisions. The teachers' judgments about children in their care are used in the screening work reported below for validating cases, where teachers judge children as either fluent or dysfluent. Together, these observations lead to the hypothesis that SSI-3 may be a starting point for screening children for all types of speech difficulty.

In this chapter, analyses were performed to investigate the effectiveness of the SSI-3 as a tool for screening speech difficulties. SSI-3 scores were obtained from recordings as allowed by Riley (1994; 2009) when the instrument is used in clinics.

5.2. Method

5.2.1. Participants.

Children from the reception classes of 11 schools were assessed on SSI-3 in the academic years 2012-2013 and 2013-2014. There were 730 children in total (369 males and 361 females); 246 were 4 years old, 482 were 5 years old and 2 were 6 years old. In Chapter 4, there were 879 participants. The smaller number here arose because, when this analysis was conducted, not all the schools had replied with their feedback on agreement on the children selected as being at risk.

5.2.2. Speech Samples and SSI-3 scoring.

All speech samples were recorded using a Zoom H4N recorder with an internal microphone. The speech was elicited by using picture stimuli included in the SSI-3 manual. A sample of 10 to 15 minutes of spontaneous monologue speech was taken. All samples were recorded in English.

5.2.3. Administration of SSI-3

As most of the children involved in this work cannot read, age-appropriate pictures supplied with the SSI-3 manual were used to prompt the child to discuss the themes that were illustrated. As indicated in other chapters, the three features that have to be obtained to calculate an SSI-3 score for a child are physical concomitants obtained at the time of recording, frequency of dysfluent symptoms and duration of stutter-measures that were obtained on subsequent analysis of the data.

Reports about the children selected as being at risk of speech difficulties that were categorised by the screen were then given to teachers in the form of reports. The teachers were then asked to report back with their feedback on whether they agree with the selection. This was done by asking them to provide information on three sets of different factors: 1) believed to be at-risk by the teacher or the Special Education Needs Coordinators (SENCOs); 2) already on SLT intervention; 3) in the process of referral by either parents, schools or specialists/professionals services. These criteria identified any children the teachers indicated had speech difficulties that were missed by AM (i.e. not selected by the researcher as being as-risk). These were used to validate the method.

5.3. Results

5.3.1. Incidence of native English speakers versus those with EAL in the cohort.

The sample included a large number of children with EAL. The diversity of first languages used and their percentages in the overall samples are shown for all the school children in the pie chart in Figure 3 in Chapter 4. A χ^2 test was performed to see if there was any association between the children with EAL and native-English speaking children and whether or not they were identified as at risk of speech difficulty. This was done separately for the author's and school's designations. The results shown in Table 29, indicate that there was no association between language and fluency designation (child was or was not fluent) for either AM or the schools. Thus, there did not appear to be any greater tendency for children with EAL to be identified as having speech difficulty compared to English children for either group of judges.

Table 29. χ^2 results for the contingency table with child's language (EAL vs native-English) and designation type (fluent/speech difficulty) for AM and for the schools.

	χ^2
AM selection	χ^2 (1 df, N= 711)= .30, p= .6
School's selection	χ^2 (1 df, N= 711)= .07, p= .8

5.3.2. Sensitivity and Specificity of selection.

Next, the data were cast into a 2 x 2 contingency table (AM vs school judgments against child being fluent vs manifesting speech difficulty) to examine sensitivity and specificity of children's identification of speech difficulty. The judgments by AM were designated 'correct' as they were based on a formal procedure. AM/school judgments are indicated by the + and – operators in the quadrants of Table 30 (e.g. '+-' in the top right quadrant indicates AM considered the child fluent, but the school did not). Table 30 gives the number of children selected in each group by AM and by the schools.

Table 30. The percentages and raw counts (in brackets) of children in the sample with respect to AM's and teachers' designations of potential risk for speech difficulties.

		School selected the children as being "at risk"?		
		Yes	**No**	
AM selected the children as being "at risk"?	**Yes**	9.2 % (67) ++	4% (29) +-	Total selected by AM= 96
	No	2.9% (21) -+	84% (613) --	
		Total selected by school = 88		

Note: Each child falls into one of four categories: 1) selected by the researcher and by the school as being at risk (yes, yes/ ++); 2) selected by researcher as being at risk, but not by the school (yes, no/ +-); 3) selected by the school, but not by the researcher (no, yes/-+); and 4) selected by neither researcher nor the school (no, no/, --).

The specificity and sensitivity were calculated to measure the rate at which each child was identified as being at risk of speech difficulties (i.e. sensitivity, True Positive, TP, ++ in Table 30 above) and the rate of those correctly identified as not being at risk (i.e. specificity, True Negative, TN, -- in Table 30 above). These statistics show how well each child was categorized into each group. The positive predictive value (PPV) and negative predictive value (NPV) were also calculated for this sample. Table 31 shows that 96.7% of the fluent children were correctly identified as being fluent, while 69.8% of the children at risk of speech difficulties were correctly identified as not fluent. The test produces a high PPV, indicating that the test does select “at-risk” children with high confidence when the children scored highly on SSI-3. The test was most effective for discriminating fluent children, identifying 95.5% of them. Thus if a child scored low on SSI-3, the likelihood that they were fluent was high. The overall accuracy of identification was 93.2% (calculated as $\frac{\Sigma \text{TP}+\Sigma \text{TN}}{\Sigma \text{total population}} = \frac{67+613}{730}$ = .932) and the prevalence in the research sample of n=730 was 13.2% (calculated as $\frac{\Sigma \text{ condition positive}}{\Sigma \text{ total population}} = \frac{TP+FN}{730} = \frac{67+29}{730}$ = .132).

Table 31. Results for sensitivity and specificity as well as the PPV and NPV.

		Did the school select the children as being “at risk”?		
		Yes	**NO**	
Researcher selected children as being “at risk”?	**Yes**	TP = 67	FN = 29	**Sensitivity** =TP/(TP+FN) = 67/(67+29)x100 =69.8%
	No	FP = 21	TN = 613	**Specificity** =TN/(FP+TN) =613/(21+613)x100 96.7%
		PPV =TP/(TP+FP) =67/(67+21)x100 =76.1%	**NPV** =TN/(FN+TN) =613/(29+613x100) =95.5%	

5.4. Discussion

The findings suggest that the SSI-3 is a reliable starting point for screening children, as it successfully distinguishes fluent from dysfluent children. The results showed a high specificity for fluent children, where the agreement between AM's judgment and the school's judgment as to which children are fluent was almost 97%. The sensitivity between the agreements of the two type of judgment was moderately high; where the agreement on which children were at risk of speech difficulties was almost 70%.

The difference between sensitivity and specificity could be due to the teachers' lack of training concerning what symptoms to listen for when assessing speech difficulties. For example, the total number of children selected by the screen as being "at-risk" of SLCNs was 96, but the teachers did not corroborate 29 of these cases. On the other hand, the schools selected 88 total children that they thought were "at-risk", but the screen did not corroborate 21 cases. The reason why the teachers may have selected more cases than were identified by the screen may be due to teachers having more information about the children, such as the history of the child, parents' fears, longer and regular access to the child's speech etc. Judging a child's ability based on a 15-minute monologue is different to what teachers do when making their decisions about whether a child has speech difficulty, but the current assessments were more objective, as a validated test was used (Howell & Davis, 2011). However, this may raise the question of who is to be considered more accurate, the teachers who have access to more information about the child or the screen test with its objective manner of testing. Nonetheless, the fact that neither the teachers nor the screen test agreed on 100% of the cases raises the dire need for a screening tool that is validated and objective.

Despite the limited nature of this assessment, the results indicated that the test selected the cases of children in the “at risk” group with good confidence, as indicated by the PPV (76.1%), and the NPV (95.5%). The teachers’ selection is arbitrary and is not based on explicit criteria (i.e. no particular test, assessment inventory or screen is available). Absence of selection criteria leads to variation from school to school and teacher to teacher. This causes uncertainty when interpreting the results from the PPV and NPV calculations. This would have led to the high number of children out of the total sample that was missed by the school compared to the researcher. Helping the teachers to be more aware of the symptoms associated with speech difficulties would reduce this problem.

Currently, many teachers use their own initiative regarding what symptoms they should look for in a child with speech difficulty, although some teachers and Special Education Needs Coordinators (SENCOs) receive training. Some teachers indicated that they do not accept WWRs as a valid symptom of speech difficulties (Mirawdeli et al., in prep, Chapter 2), which is consistent with Riley’s (1994) instructions for the use of SSI-3. If the number of teachers following this procedure increased, the specificity and sensitivity of the results may improve. On the other hand, the practice of omitting WWR could explain why there did not appear to lead to children with EAL to be identified as having speech difficulty by the teachers. Children with EAL display word-finding difficulties which can be mistaken for disfluency-like symptoms (WWR) and if the teachers had interpreted these as a sign of speech difficulty, more children from the EAL group would have been picked up as being “at risk”. Children with EAL are also expected to show phonological delay when learning an additional language. But this may be a normal transitional phase for children with EAL as they learn English or should these children be considered as having a phonological delay.

Allocating more children to the “at risk” group, to be referred for further clinical assessment, is better than leaving these children to be identified at later stages when success rates in intervention would decline. This will inevitably put pressure on the currently-available resources, and expansion of services would seem to be ruled out as it would require increased funding.

The data discussed in this report were based on a clinical form of assessment which the SLTs perform on children referred to them for stuttering. The SSI-3 is not a screening tool, rather a severity instrument, which informs the SLT of the stuttering severity of the child. In order to make this instrument more practical as a screening tool, it will need to be less time consuming to administer and thus easier for use by teachers. Chapter 4 suggests %SS alone may be sufficient for correct identification. Further analyses will also extend to other speech symptoms than those of stuttering alone. The latter is investigated in Chapter 6 by carrying out analyses on how well children with speech difficulties other than stuttering have been identified by the SSI-3 during the screening. A screening tool for speech difficulty that takes into account a wider set of symptoms than those used for assessing stuttering would be preferable to having several screening tools for different types of speech difficulty. A tool that selects all possible cases of speech difficulties on one test occasion will hopefully be an essential instrument for teachers to use.

Chapter 6

Investigation into Whether Extending the Symptom Set May Improve Detection of Speech Difficulties

6.1. Introduction

As discussed in section 5.4, a screening tool for identifying children with speech difficulties, may need to take more symptoms into account than those that are included in SSI-3-based analyses. The symptoms associated with speech difficulties in general are not restricted to those associated with stuttering. The absence of these symptoms could lead to children with other forms of speech difficulty being missed. Also, as has been seen when considering whole-word repetitions, there are disagreements about what symptoms are associated with stuttering.

As preparation for the empirical work in this chapter, systematic reviews of the literature were conducted to identify what symptoms have been associated with stuttering and other speech difficulties. Any additional symptoms that emerge could potentially be used to extend the set employed in future assessment procedures, if they prove to be significant predictors. Speech samples collected and analyzed based on the extended set of symptoms might then improve identification of children with speech difficulties other than stuttering. Unfortunately, the systematic review had limited usefulness in generating possible symptoms, so another approach was taken.

The symptoms that were identified according to the adopted approach, along with the SSI-3 set of stuttering symptoms, were used in the second part of this chapter: Children were selected so that the sample included cases with various types of speech difficulty. Their speech samples were analyzed using the extended symptom set. The analyses were exploratory and the results were examined to see whether any symptoms looked promising for identifying speech difficulties. If so they may be used to improve symptom-based diagnoses of speech difficulty in future versions of the screening instrument for use in schools.

6.1.1. Systematic review to identify what symptoms are used to classify speech difficulties in addition to stuttering in the beginning-school population.

It was argued above that the symptoms used in SSI-3 may need to be extended if they are to be used to screen for the complete range of speech difficulties that affect otherwise typically-developing young children. Here, the way that systematic reviews were conducted are described. As a caution, it needs mentioning that the outcomes of the reviews for the purpose at hand were disappointing.

Three categories of speech difficulty in addition to stuttering were the subject of separate reviews; phonological disorder and phonological delay Dodd (1995), and auditory disorders (hearing problems affect speech production). Dodd, Zhu, Crosbie, Holm & Ozanne (2002) sub-divided the first two types into four sub-categories: 1) Phonological delay. This applies when children's speech systems develop normally, but at a slower rate than typically developing children. As a consequence, typical early speech errors are produced at older ages; 2) Articulation disorders. Here, a child has difficulty producing certain speech sounds in all contexts; 3) Consistent phonological disorder. A child's speech contains atypical patterns not usually found in typically-developing children where these abnormal patterns do not vary between productions; 4) Inconsistent phonological disorder. A child's speech shows a high degree of variability and he or she produces the same word in a number of different ways.

The review was initially intended to identify symptoms for the three main types of speech difficulty indicated above, as well as stuttering. The subsequent plan was to go on to examine the additional three types in Dodd et al. (2002) – i.e. sub-categories 2 – 4 above. Separate systematic reviews were conducted that addressed step one in the initial plan. However these revealed several shortcomings. Consequently for the quantitative analyses, expert sources rather than results of the systematic reviews were used to identify the symptoms appropriate for the four main categories alone (phonological delay, phonological disorder, hearing loss and stuttering). The procedure followed, and the issues raised are detailed below for completeness.

1. General search strategy for the systematic reviews.

Several search criteria were considered. The final search criteria involved limits on the database used, year of publication of results, language that the article was published in, other speech difficulty types subsumed under our superordinate category (e.g. phonological disorder), permitted methodologies used in any studies that were selected, age group of participants, and factors related to quality of the evidence. The specifics of the search criteria were as follows:

1) Database used to conduct the search: PsychInfo

2) Years allowed - past ten years

3) Had to be published in English

4) Target terms -Type of Search: Multi-field search. To illustrate for phonological disorder, the following were selected: Bar 1 = “communication disorder”, Bar 2 = “articulation disorder”, Bar 3 = “phonological disorder”, Bar 4 = “phonetic disorder”, Bar 5 = “phonemic disorder”, Bar 6 = dyspraxia, Bar 7 = dysarthria, Bar 8 = apraxia. Term separation: Bars 1-8 were separated by “OR” in the left hand tab and searched for in “Key Concepts”.

5) One of the following descriptors of the methodology was required in any paper that was selected: 0400 Empirical Study, 0430 Follow up Study, 0450 Longitudinal Study, 0800 Literature Review, 0830 Systematic Review, 1200 Meta-Analysis.

6) Age Groups: 160 Preschool Age <age 2 to 5 yrs>, 180 School Age <age 6 to 12 yrs>

7) Quality - Articles had to be peer reviewed.

2. Phonological Disorder.

The disorders that fall in this category do not have agreed definitions (Forrest, 2003). However, there are some general characteristics about the symptoms. For example, apraxic speech has segmental errors whereas dysarthrias lead to overall, suprasegmental distortion of the signal that cannot easily be described on a segmental basis.

The titles (and abstracts) of the papers that were generated by the search were examined next in order to get a general idea of the content of each paper. For phonological disorder, ten of the 79 articles were excluded as they were not related to this form of difficulty. The next two phases looked at quality criteria and reviewed those papers that passed these criteria for descriptions of symptomology. During these phases a further 11 of the 68 articles were excluded based on the content not being relevant or symptom criteria being inadequately reported. Finally, it was found that a number of the articles arising from the search were not published in full text and were therefore not used in search; this is because the criteria set initially was that all articles had to be published in full text in peer reviewed journals. After this process, there were 50 papers (of the original 79) available from the PsycInfo database that met all criteria.

A table was made of authors of the paper, year published, title, number of participants, ages of participants, and type of study, although these categories were not complete for all papers. For example, out of the 50 papers only 16 had all this information available, three had none of the information available, 10 did not indicate what type of study it was, 13 did not indicate the age groups, seven had age and type of study missing and one had age and number of participants missing. Each paper was scrutinized for symptoms associated with specific speech difficulties. After reviewing each paper individually, the data were tabulated according to type of difficulty, total number of studies in which the difficulty was mentioned, and the symptoms associated with the disorder. Within the "Symptoms" column, symptoms were listed by the number of papers that mentioned them as a characteristic of the given disorder, from most to least frequent. Symptoms that were identical but used different names were collapsed at this stage.

Twenty disorders were mentioned in the papers: dysarthria, spastic dysarthria, dyskinetic dysarthria, ataxic dysarthria, phonological dyslexia, surface dyslexia, developmental dyslexia, dyspraxia, developmental dyspraxia, dysphagia, apraxia, childhood apraxia of speech, anarthria, aphasia, pseudobulbar palsy, congenital syndrome, motor speech disorder, phonological disorder, inconsistent phonological disorder, and phonological delay (the latter was the subject of a separate search).

A number of symptoms applied to more than one of the disorders discussed. For example, disordered articulation was associated with dysarthria, ataxic dysarthria, apraxia, childhood apraxia of speech, congenital syndrome, and phonological disorder. Other overlaps were seen for reduced intelligibility, disordered or atypical prosody, inconsistent vowel and/or consonant errors, reduced literacy abilities, impaired speech sequencing or planning, and groping, among others. These overlaps occurred for a variety of symptoms and for a number of the disorders. It was not a problem that individual symptoms were associated with several different speech difficulties in the phonological disorder category, as this may show that the superordinate groupings are appropriate. However, given that only 16 of the papers included all the required information the systematic review approach may not give valid results and expert sources were used instead.

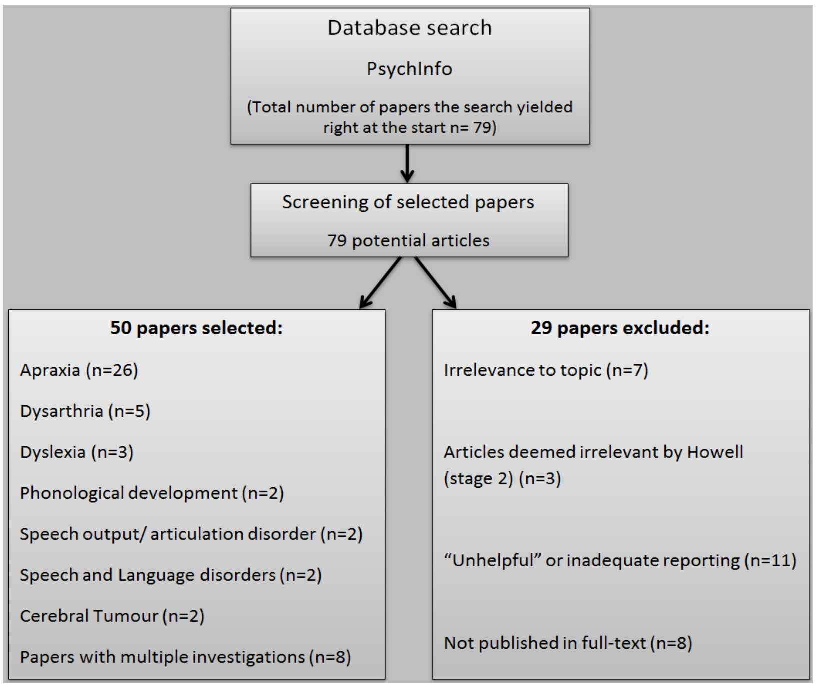

***Figure 15*. The total number of papers generated and those included and excluded.**

3. Limited research concerning symptoms of phonological delay.

For phonological delay, no symptoms were returned. This is curious since Broomfield and Dodd (2005) were able to categorize successfully 320 children with speech difficulties into articulation disorder, phonological disorder, and phonological delay. The majority (57.5%) fitted their criteria for phonological delay. As phonological delay is common, there should be some symptoms returned for this form of difficulty from the literature. The absence suggests that there were complications with use of the terms, or authors exercised caution because of disagreement about what constitutes delayed phonology. Search terms could be adjusted to account for this, and again, further searches should be conducted to examine phonological delay further.

Absence of identifiable symptoms associated with phonological delay, or the limited amount of research on this particular aspect, raises several questions about the relevance of this category with respect to the present research: 1) should phonological delay be considered as a speech difficulty? In Broomfield and Dodd (2005), 57% of 320 children were categorized as suffering from phonological delay. Furthermore, the proportion of the sample that were children with EAL is not known. This relates to the following point. 2) Children with EAL are expected to show phonological delay as they are learning the additional language later than native speakers of that language. It is debatable whether the problems that children with EAL experience are the same as the problems that arise when there is phonological delay in a child who only speaks the native language. Furthermore would the problem clear up quickly for children with EAL and be more protracted for children who speak only English and who have phonological delay? This brings us to the final point: 3) What distinguishes phonological delay as a speech difficulty from delay that resolves after a short transitional learning period? One possible way of distinguishing these forms may be to look at different patterns of morpheme and phoneme delays and see whether there are differences between children with EAL versus those with English only.

4. Symptoms related to hearing loss.

The systematic reviews concerning symptoms to identify children whose speech was affected when there is hearing loss returned a high number of papers. However, there was little agreement about defining symptoms, as was found with phonological disorder. A narrative review of some sources and papers of interest yielded some symptoms that recur in papers. A systematic review would give a more objective list and would be treated as a summary of peer reviewed research thus far into hearing-related symptoms. However, the research in this area is vast and terms are not standardized. Hence, a systematic review would not be conclusive. Assessing hearing-related problems in children with EAL is another obstacle, which needs to be addressed.

6.1.2. The present study.

The results of the systematic reviews were disappointing and did not lead to a set of symptoms that seem appropriate for identifying the main types of speech difficulty targeted. A major reason for this appeared to be that symptoms were described that precisely identified participant characteristics sometimes at the expense of general symptom characteristics appropriate to the types of speech difficulty as a whole (i.e. an emphasis on individual differences). Instead of abandoning this quest for characteristic symptoms of different types of speech difficulty, authoritative sources were identified that sought to specify such general symptom characteristics. The extended symptom types were used to examine whether they improved categorization in the study that follows study. The results of this study should inform future developments of the screening instrument.

The purpose of the following analyses was to see if individual or grouped symptoms accurately predicted: 1) the classification of children as fluent or dysfluent, termed *fluency*; and 2) the classification of children into the SLCN categories of fluent, stuttering, phonological delay, phonological disorder and hearing loss, termed *fluency type*.

6.2. Method

6.2.1. Participants.

Speech samples from 28 children were employed. There were two main groups. Twenty were children who were judged by schools as “at-risk” of speech difficulty, where the type of speech difficulty was established by two independent judges (see below). The children were selected at random from the sample of such children identified in Chapter 5. Eight fluent children were included as controls. Again these were selected at random from the children available from Chapter 5.

6.2.2. Procedure for identifying symptoms appropriate to a more comprehensive class of speech difficulties.

The symptom groupings for different type of speech difficulty were based on past research (Broomfield & Dodd, 2004b; Howell, 2010; Merritt and Rowland, 2000). The symptom groups were: phonological disorder, phonological delay, Riley's stuttering symptoms, and hearing loss symptoms. The symptom groupings used are listed in Tables 32 to 35 along with the source from which were obtained.

Table 32. Phonological Disorder Symptoms.

Phonological Disorder Symptoms	Source	Defined as
Backing (B)	Dodd, Hua, Crosbie, Holm & Ozanne (2002)	Producing a sound that is made further back in the mouth in lieu of the target eg. /t/ - /k/.
Affrication (A)		Replacing a non-affricate with an affricate eg. /t/-/tʃ/.
Initial consonant deletion (ICD)		Deletion of the initial consonant in a target word eg. 'toy' – 'oy'.
Medial consonant deletion (MCD)		Deletion of the middle consonant in a target word eg. 'faster' – 'faser'.
Intrusive consonant (IC)		The use of an extra consonant in a target word eg. 'animal' – 'aninmal'.
Denasalisation (DN)		The lack of nasal airflow during a nasal sound eg. /m/ and /n/.
Sound preference substitution (SPS)	Small (2005)	The consistent use of a particular sound in lieu of a target eg. 'yellow tractor' – 'dedo dacdor'
Devoicing (DV)		To use the voiceless consonant when the place and manner are kept constant with the target eg. /v/ - /f/.
Stops replacing glides (SRG)		Replacing a stop consonant with a glide consonant eg. /t/-/w/.
Fricatives replacing stops (FRS)		Replacing a fricative consonant with a stop consonant eg. /t/-/s/.
Palatal fronting (PF)		To produce a palatal target further forward in the mouth eg. /y/ - /t/.
Liquid simplification (LS)		To replace a liquid consonant with a glide consonant eg. /r/-/y/.

Table 33. Phonological Delay Symptoms.

Phonological Delay Symptoms	Age at which the symptom is 'delayed'	Source	Defined as
Deaffrication (DA)6	5	Dodd, Hua, Crosbie, Holm & Ozanne (2002)	Replacing an affricate with a non-affricate eg. /tʃ/ - /t/.
Cluster reduction (CR)	4		The deletion of a consonant in a cluster eg. /kr/ - /k/ such as in 'cream' – 'ceam'.
Fronting of velars (FV)	4		Fronting sounds made at the velum eg. /k/ - /t/.
Weak syllable deletion (WSD)	4		Deletion of the non-stressed syllable in a target word eg. 'telephone' – 'telphone'.
Stopping of fricatives (SF)	3,6		Producing a stop consonant in lieu of a fricative eg. /s/ - /t/.
Voicing (V)	3		Using the voiced version of a target consonant when the place and manner are consistent eg. /k/ - /g/.
Final consonant deletion (FCD)	2,1		Deletion of the final consonant in a target word eg. 'pretty' – 'prett'.

Table 34. Stuttering Symptoms.

Stuttering Symptoms	Source	Defined as
Prolongation (P)	Riley (1994)	The prolongation of a sound eg. 'aaand'.
Part Word Repetition (PWR)		Repetition of part of a word eg. 'du-du-duck'
Broken Word (BW)		Pausing in between a target word eg. 'du-ck'
Phoneme Repetition (PhoR)		Repetition of a phoneme eg. 'd-d-d-duck'
Blocking (Bl)		No airflow or sound produced when attempting a target eg. 'and the k (pause) p (pause) duck'

[6] Deaffrication was only applicable to some of the children as the age of delay is five. Therefore, if a child of four produced this particular phonological process it was not recorded as a symptom of phonological delay.

Table 35. Hearing loss Symptoms.

Symptom	Reference
Smaller vocabularies	Stiles, D. J., Bentler, R. A. and McGregor, K. K., (2012), 'The speech intelligibility index and the pure-tone average as predictors of lexical ability in children fit with hearing aids.' Journal of Speech, Language, and Hearing Research, 55, 764–778.
Deficits in phonological discrimination/non-word repetition	
Hyper/hyponasality (esp. hypernasaility on all vowels)	Merrit, R. Deafness.
Manner contrasts reduced (fricative/plosive)	
Voicing Errors	
Reduced/addition of syllables	
A-typical intonation(including lack of nuclear pitch changes/no nuclear shifts(at phonetic level))	
Airstream mechanism difficulties (e.g. Use of implosives)	
Raised pitch on closed vowels	
Silent articulation	
High frequency sounds omitted	
Omission of copular/auxiliary verbs	
Omission/confusion of determiners	
Omission of tense endings/ 's' endings (genitive/plurals)	
Overall loudness/voice quality/pitch	

6.2.3. Annotating Samples.

The audio speech samples from the children were transcribed as follows. Individual dysfluent speech symptoms (individual symptoms for brevity) for the 28 children were coded in the transcriptions. The individual symptoms were then assigned to groups based on types of speech difficulty (grouped symptoms). The symptom names in the Tables given in the previous section were added, as appropriate, to annotations on digital audio versions of the speech. To do this, the audio files were converted into Speech Filing System format. Speech Filing System was employed to display an oscillographic version of speech (Figure 15) that can be zoomed in on and an extract of speech selected and played (this can be repeated as and when necessary when the type of symptom is judged). An ‘orthography’ row was added to Speech Filing System so that annotations were available in alignment with the oscillogram. All individual words were marked in this row using a vertical line to mark the beginning and an ‘x’ to mark the end with the word itself in between. For example, ‘|and |x’ would appear under the oscillogram at the point where the word ‘and’ appeared (Figure 15). A ‘disfluencies’ row was added below the orthography row specifically to locate dysfluent events and their type. The procedure was similar to that used for word annotation. Disfluencies were coded using vertical lines to delimit the word and the type of disfluency as entered between them, for example, |PDel|. Table 36 gives the codes and their abbreviations.

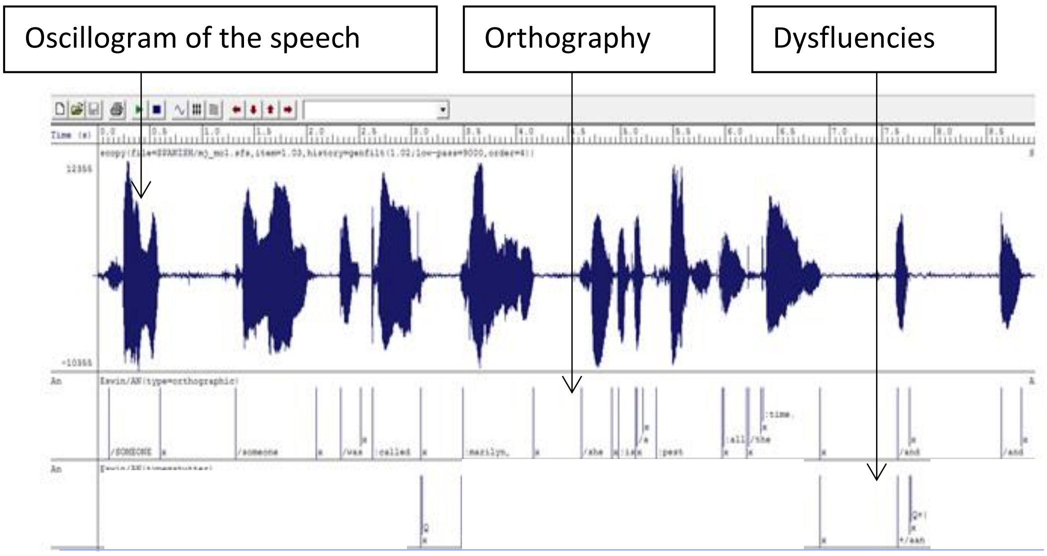

***Figure 16*. SFS display showing oscillogram, orthographic transcriptions and disfluency annotations.**

Table 36. Speech difficulty groups (SLCN) and the individual symptoms associated with them.

SLCN Group	Individual symptoms within the group	Transcription code used
Stuttering	Prolongations	Pro
	Blocks	Blo
	Hesitation	Hes
	Revisions	Rev
	Broken words	BW
	Part-word repetitions (PWR)	PWR
	Whole-word repetitions (WWR)	WWR
	Phrase repetitions	PhR
	Interjections	( )
Phonological Disorder	Backing	B
	Initial consonant deletion	ICD
	Middle consonant deletion	MCD
	Sound preference substitution	SPS
	Intrusive consonants	IC
Phonological Delay	Gliding (r-w)	G
	Deaffrication	D
	Poor phonological awareness	PHA
	Final consonant deletion	FCD
	Weak syllable deletion	WSD
Hearing Loss (HL)	Voicing errors	VE
	Reduction/addition of syllables	RS/AS
	Raised pitch on closed vowels	RPCV
	Omission of copular auxiliary verbs	OCA
	Omission or confusion on determiners	OCD
	Omission of tense endings/ ‘s’ endings	OTE
	Pitching issues	PI
	Phonological discrimination failings	PDF
	Poor digit recall	PDR
	Hypo/hyper-nasality	HyperN/ HypoN

Note: The final column gives the codes used in the transcriptions (as they appear in SFS displays like that shown in Figure 15).

The individual and grouped symptoms were expressed as the percentage of syllables affected by individual symptoms or groups of symptoms (depending on the analysis) out of all syllables as is done in SSI-3 with symptoms of stuttering. The percentage scores for each set of grouped symptoms for each child are given in Table 37.

Table 37. Percentage (%) symptom values for each participant for sets of symptoms appropriate for four types of speech difficulty.

Student No.	Stuttering (%)	PDis (%)	PDel (%)	HL (%)
Student 1	4.34	.11	2.56	1.11
Student 2	5.55	.32	1.49	3.42
Student 3	5.96	.24	2.92	2.19
Student 4	7.2	0	3.90	5.0
Student 5	5.58	0	4.0	5.06
Student 6	6.6	0	1.05	2.76
Student 7	22.80	0	4.50	10.27
Student 8	18.36	0	3.94	5.07
Student 9	41.95	0	7.47	27.59
Student 10	2.09	0	5.64	1.88
Student 11	1.89	3.16	5.47	1.89
Student 12	.37	4.60	6.25	.55
Student 13	.95	2.70	4.45	2.70
Student 14	4.34	0	4.58	3.76
Student 15	.54	.81	5.15	15.45
Student 16	3.86	.20	2.44	8.94
Student 17	5.24	2.62	4.19	8.90
Student 18	5.0	2.5	5.0	11.50
Student 19	5.0	.32	3.19	5.86
Student 20	7.50	1.40	5.96	13.10
Student 21	1.55	.69	2.41	.34
Student 22	1.75	.52	1.57	.17
Student 23	.99	0	0	0
Student 24	.32	0	.32	0
Student 25	1.03	0	0	.90
Student 26	2.45	0	1.43	.20
Student 27	.88	0	.88	.66
Student 28	1.80	.22	2.02	2.02

Key: PDis = Phonological disorder**, PDel** = Phonological Delay**, HL**= Hearing Loss

Judgements were then made by two separate judges on whether each child was at risk of each of the speech difficulties, using the data in Table 37. Each judge's decisions are given in Table 38. Inter-judge reliability between the two judges was 89.3%. The associated kappa coefficient was 0.85 which represents agreement well above chance (Fleiss, 1971).

Table 38. Diagnoses as determined by two independent judges.

Student No.	Suspected diagnosis Examiner 1	Suspected diagnosis Examiner 2	Teacher/parent feedback
Student 1	Stuttering	Stuttering	Parents and Teachers concerned/ referred for SLT
Student 2	Stuttering	Stuttering	working with ECAT* group- Teachers concerned
Student 3	Stuttering	Stuttering	No follow up data provided
Student 4	Stuttering	Stuttering	Teachers NOT presently concerned- no action taken
Student 5	Stuttering	Stuttering	Child had been referred already as Teachers and Parents concerned
Student 6	Stuttering	Stuttering	Teachers NOT presently concerned- no action taken
Student 7	Stuttering	Stuttering & HL	Child had been referred already as Teachers and Parents concerned
Student 8	Stuttering	Stuttering	Teachers NOT presently concerned- no action taken
Student 9	Stuttering	Stuttering & HL	Several referrals made previously to various services- under monitor
Student 10	PDel	Stuttering	Child under monitor- no significant concerns for referral
Student 11	PDel	Stuttering	Child under monitor- no significant concerns for referral
Student 12	PDel	PDel	Working with SLT- NHS
Student 13	PDel	PDel	Child under monitor- no significant concerns for referral
Student 14	PDel	PDel & HL	No follow up data provided
Student 15	HL	HL	Initial concerns- Teachers feel child improved now
Student 16	HL	HL	Parents and Teachers concerned/ Seen GP for Stutter-referred for SLT
Student 17	HL	HL	Working with SLT- NHS
Student 18	HL	HL	Diagnosed with autism, having speech therapy- working with ECAT* group
Student 19	HL	HL	Parents and Teachers concerned/ referred for SLT
Student 20	HL	HL	Worked with SLT- discharged
Student 21	Fluent	Stutter	Fluent
Student 22	Fluent	Fluent	Fluent
Student 23	Fluent	Fluent	Fluent
Student 24	Fluent	Fluent	Fluent
Student 25	Fluent	Fluent	Fluent
Student 26	Fluent	Fluent	Fluent
Student 27	Fluent	Fluent	Fluent

***ECAT- (Every Child a Talker)** an organisation working with children to improve speech.

6.2.4. Overview of analyses.

In the analyses, the independent variables (IVs) are either the % of the individual symptoms or the % of the grouped symptoms, and the dependent variable (DV) is either fluent/not fluent or individual fluency type (depending on the analysis conducted). Discriminant function analysis (DA) was used in the first analyses of the data. DA uses the independent variables (IV) as predictors of the DV. The statistical question is how well the DV is predicted by the IVs. The IVs are first tested in the DA to see whether they differed between groups. When the means between groups were significantly different, that variable was included in the classification of cases (children) into groups. The selection of which IVs are used as predictors takes place automatically (analyses performed using SPSS).

Hierarchal cluster analyses were then conducted. Hierarchal cluster analysis was used to determine how the cases grouped, or clustered, together. This is usually done by assigning each case to be clustered together with its sequentially most similar case. However, in the current analyses, Ward's (1963) method of clustering was used. In this method, cases are clustered together based on how much the clustering decreases the overall group variance, Comparisons of cases to their group centroid means can also be made using Ward's method (Appendix H). This shows how distant each case is from the centroid of its cluster. Hierarchal cluster analysis allows users to impose a specific number of clusters based on theory or on previous research. Alternatively, hierarchal cluster analysis allows number of clusters to be determined automatically. For the present analyses, four clusters were found first and then analysis was made where no imposition about number of clusters was given. The output of both of these analyses produced exactly the same Dendograms; the same four clusters were found when there was no imposition compared to when there was an imposition of four clusters.

6.3. Results

6.3.1. DA for fluent vs. dysfluent.

DA was used to test the hypothesis that children can be classified as either fluent or dysfluent (fluency group) based on scores on the four IVs % symptoms associated with: stuttering, phonological disorder; phonological delay; and hearing loss. Note that, by chance, no child had a diagnosis of phonological disorder), but scores were available for all children for the IV %PDis symptoms. A stepwise method was used in the logistic regression where any group of symptoms could be selected. The phonological delay group of symptoms was the only IV that was extracted by the logistic regression. A contingency table of the classification of phonological delay based on agreed versus that based on delay symptoms was significant by χ^2 (Wilks λ = .498, χ^2= 17.8, df = 1, Canonical correlation = .709, p <. 001). The significant χ^2 shows that fluency group could be classified better than chance using % phonological delay scores alone. The single function using this IV accounted for 50.2% of the variance in the assignment of the children into fluency groups. Phonological delay group symptoms correlated with the function (Table 39), that is the linear combinations of the predictor variables (the discriminant functions). Table 40 shows the group centroids for function 1 (the only function in this case), which was successful in classifying the cases into the originally-determined groupings (Table 38), and this achieved 85.7% correct classification. The way individual cases are classified involves evaluating them with respect to any functions available and assigning them to the group that has the closest centroid to that of the case.

Table 39. Standardized Canonical Discriminant Function Coefficients.

	Function 1
Phon.Delay	1.0

Table 40. This table shows the centroids or the predicted group means

Fluent / Dysfluent	Function 1
Fluent	-1.531
Disfleunt	.612

Figure 16 shows the effectiveness of the function in making the classifications. Children classified as fluent (top left) all have scores that separate them from the children deemed to be dysfluent. Thus, it is apparent that there is no overlap between the classifications of these children into the two groups.

It is possible that all children who are dysfluent would display some symptoms of phonological delay, even when they are diagnosed as having another speech difficulty. Broomfield and Dodd (2004b) found that this speech difficulty was the most prevalent in their incidence study of clinical cases, with a rate of 57.5%. If a child is suffering from hearing loss (HL), they are likely to present with phonological delay symptoms, as their ability to discriminate sounds is affected by the loss. The overlap with phonological delay could also apply to children who stutter or who have phonological disorder. In addition, it was shown in the systematic literature review, summarized earlier in this chapter, that there was considerable symptom overlap between different types of speech difficulty. Moreover, children who are fluent may also display some disfluency, either because of their young age or because they are children with EAL. Thus, some overlap is expected between the children who are fluent and those who stutter. The Pearson's correlation statistic was used to test for any relationship between the symptoms in the phonological delay group and the symptoms in the hearing loss group. A significant positive correlation was found, *r*= 0.516, *p*≤0.01.

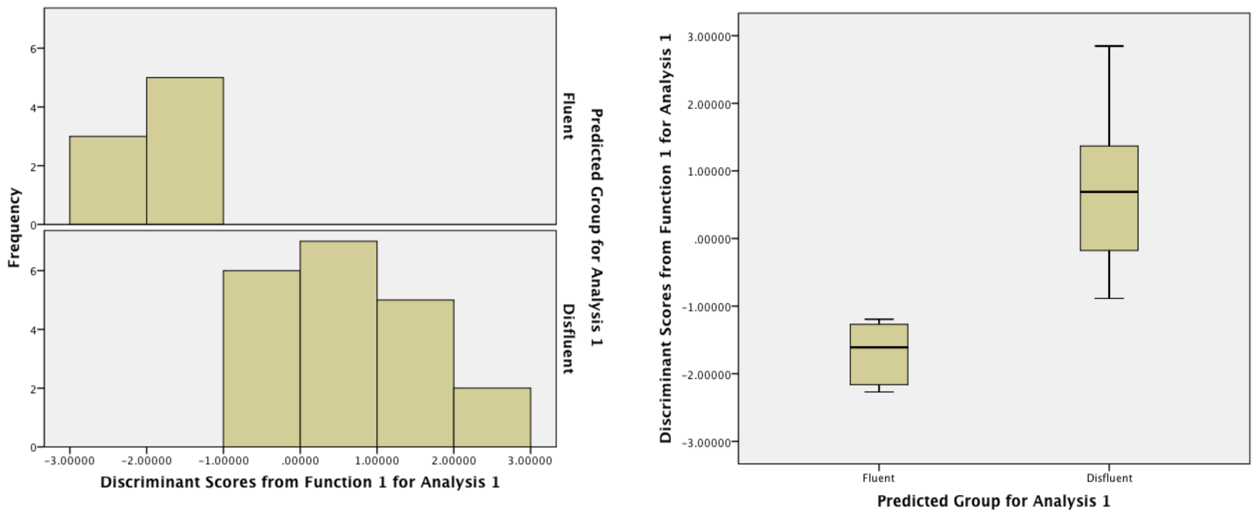

***Figure 17*. Histogram (left) and Boxplots (right) showing that the discriminant function using % phonological delay symptoms classified children as either fluent or dysfluent successfully.**

Note: The histogram is plotted in successive rows where the top one has the cases classified as fluent and the bottom one has the cases classified as dysfluent. The function scores on the X-axis in the histogram and on the Y axis in the boxplots represent the discriminant score calculated by DA. The boxplots at the bottom were plotted using the same data but show the mean of each group (the box of the band in the middle is the median) and variability in the groups (the whiskers), where the whiskers are the extent of the scores in that group.

6.3.2. DA analysis using pre-determined symptom types for specific SLCN.

The same procedure (DA) was used next to see whether children could be correctly classified into fluency type (stuttering, phonological delay, phonological disorder or hearing loss) using group symptom scores associated with these speech difficulties as IVs. The DV was the child's classification category as either: stuttering, phonological delay, phonological disorder, hearing related or fluent (no speech difficulty). Each child's score on the discriminant functions that used the IV group symptom scores for stuttering, phonological disorder, phonological delay, and hearing loss were employed to classify the children. The DA classifications were compared with the independent classifications given in Table 38 to determine the overall accuracy in classifying fluency types.

Here, the DA extracted three functions: Function one used the group of phonological delay symptoms alone; Function two used the phonological delay and stuttering symptoms; Function three used the Hearing Loss group of symptoms alone. The phonological disorder group of symptoms was not selected for use in any of the functions, probably because none of the children selected presented with this type of speech difficulty. The overall χ^2 for speech difficulty classification was significant (Wilks λ = .06, χ^2= 66.19, df = 9, Canonical correlation = .857, p <. 001). This showed that stuttering scores, phonological delay scores and hearing loss scores were strong predictors of the type of speech difficulty a child had with 94% of the variance was accounted for. Table 41 presents the standardized discriminant function coefficients and Table 42 shows the group centroids for functions 1 to 3. The new canonical variables correctly classified 89.3% of the children into the groups given in Table 38.

Table 41. Standardized Canonical Discriminant Function Coefficients.

	Function 1	Function 2	Function 3
Stuttering	-1.434	1.303	.688
Phon.Delay	1.138	.998	.052
HL	.599	-2.160	.301

Note: Column two shows that the most important predictor in function 1 is stuttering score with a coefficient of -1.434, followed by phonological delay score then hearing score.

Table 42. Functions at group centroids.

	Function 1	Function 2	Function 3
Stuttering	-.991	.618	.795
Phon.Delay	2.052	1.966	-.472
HL	1.772	-1.922	.242
Fluent	-1.496	-.482	-.781

This table shows the centroids or the predicted group means.

The predicted grouping variables were used to plot the histograms and boxplots as previously (Figure 17). Here 10.7% of the classifications did not correspond and these are the cases where scores overlapped in the histograms at the top of Figure 17. The misclassification arose because some cases from one fluency type were closer to the centroid of another fluency group than they were to their own centroid. Table 41 shows that function 1 is best at predicting group membership where stuttering symptoms group scores have the highest coefficients (-1.434), followed by phonological delay (1.138) and hearing loss (0.599). The coefficients for each symptom group are lower in function 2 and 3.

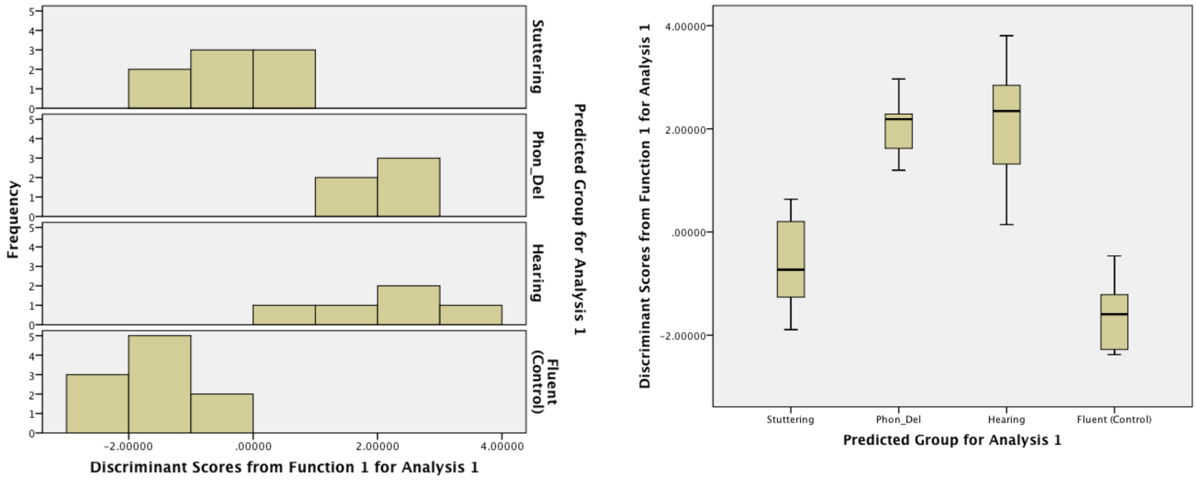

Figure 18. **Histogram (left) and Boxplots (right) showing the discriminant function for stuttering, phonological delay and hearing loss.**

Note: These were successful at classifying children as either stuttering, phonologically delayed, experiencing hearing loss or fluent. The function score on the x axis at left represents the distance from the centroid of each group (see Appendix H for an illustrated example of centroids). The boxplot at the bottom shows the means, medians and variability in the data of each group.

As can be seen from Figure 17, there is some overlap between the symptom group classifications, namely, between the stuttering and the fluent groups, and between the hearing loss and the phonological delay groups. This was expected as previous research (Broomfield & Dodd, 2004b; Howell, 2010) indicates such overlap. Thus, many children who show one type of speech difficulty also have symptoms associated with other speech difficulties. Therefore, symptoms for different types of speech difficulty may not be exclusive and this may lead to misclassification. For example, a child may have high and similar scores for all categories of speech difficulty, making it difficult to classify the child as having one specific speech difficulty. This problem increases when there is more than one child who exhibits deviant patterns, because the discriminant functions adjust to these cases.

6.3.3. Cluster analysis for fluent vs. dysfluent.

Cluster analyses using Dendograms as a display format were performed next to establish children's classification into speech difficulty classes. The group symptoms were used for this analysis (stuttering, phonological delay, phonological disorder and hearing loss). As mentioned, cluster analyses determine how each case is grouped to the next sequentially similar case in that group. Here the input variables were the grouped symptom scores for each case, and the analysis specified that cases should be classified into the supplied speech difficulty types using Ward's method.

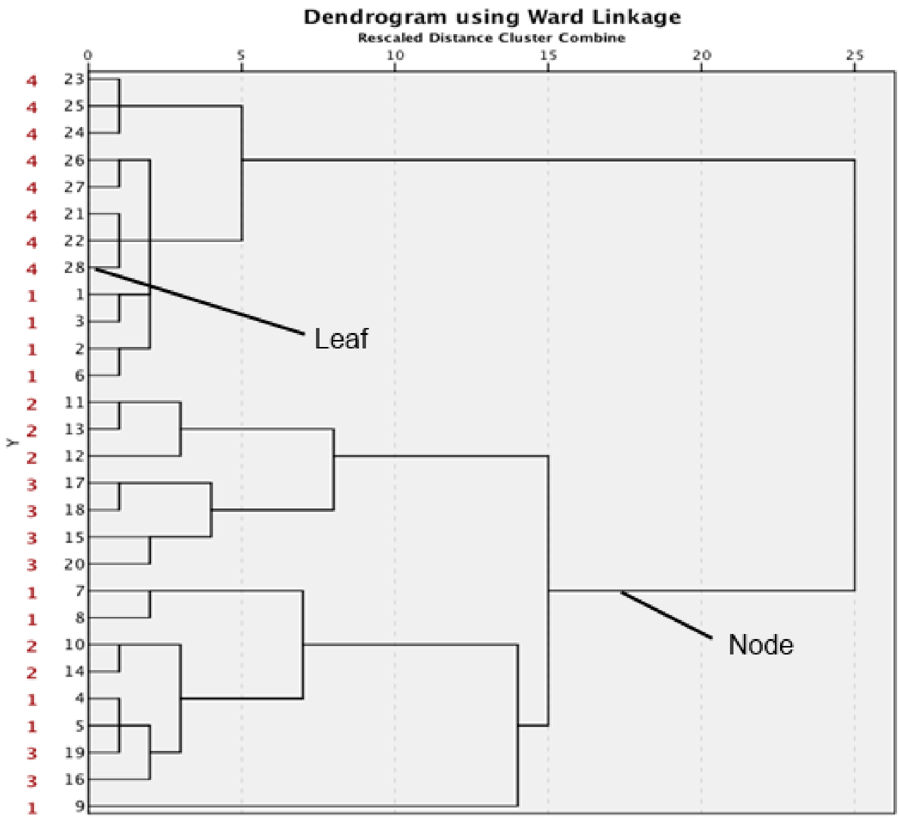

Figure 19. **Dendogram Showing How the Individual Cases are Grouped.**

Note: These plots are also termed tree diagrams, the longest lines are termed nodes and the end of the line at the y axis is termed a leaf. A leaf represents a case (in this analysis, a child) and a node represents the difference between individual cases or clusters of cases, where the highest node indicates the biggest difference. The Y axis has two columns of numbers, the numbers in black are the child identification codes and the numbers in red are codes for type of speech difficulty.

Key: In first column on far left of the Y axis are the groups which the children/cases belong to where 4= fluent; 3= hearing loss, 2= phonological delay; 1= stuttering.

The results are shown in Figure 18. This indicates that some of the children classified as stuttering cluster close to the fluent children (numbers 23 down to 6 in black on the Y axis). This indicates that the difference between each case and the centroid for ‘stuttering’ and the centroid for ‘fluent’ were similar (example illustration in Appendix H). Cases 11 to 20 clustered together and belonged to the hearing loss and phonological disorder fluency groups. Cases 7 to 16 included a mixture of stuttering, phonological delay and hearing loss and case 9 was a child who stutters.

Figure 19 shows the same Dendogram as that in Figure 18 except that this time ellipses are included to highlight two important points. First, the bigger ellipse in red shows how different the two ends of the Dendogram are. This is indicated by the length of the node, which indicates the level of difference. The smaller red ellipse includes all the fluent children as well as a few stuttering children (presumably, those with lower scores on the individual symptoms associated with stuttering). The larger red ellipse does not include any of the fluent children but encompasses all the dysfluent children. This is further support that the children are classified correctly as dysfluent. The grouping of children with different forms of speech difficulty may arise because, as discussed in earlier chapters, some children have co-morbid speech difficulties. The second point is that there are four brown ellipses at the next node down. The top one encompasses only the fluent children. The second one encompasses children with high scores on phonological delay and hearing loss symptoms. This demonstrates again that hearing loss and phonological delay are closely related but differ significantly from the form of speech difficulty encompassed by the third brown circle which is associated with the second longest node. This ellipse encompasses children who are mainly those who stutter, but include a few cases where there is phonological delay or hearing loss.

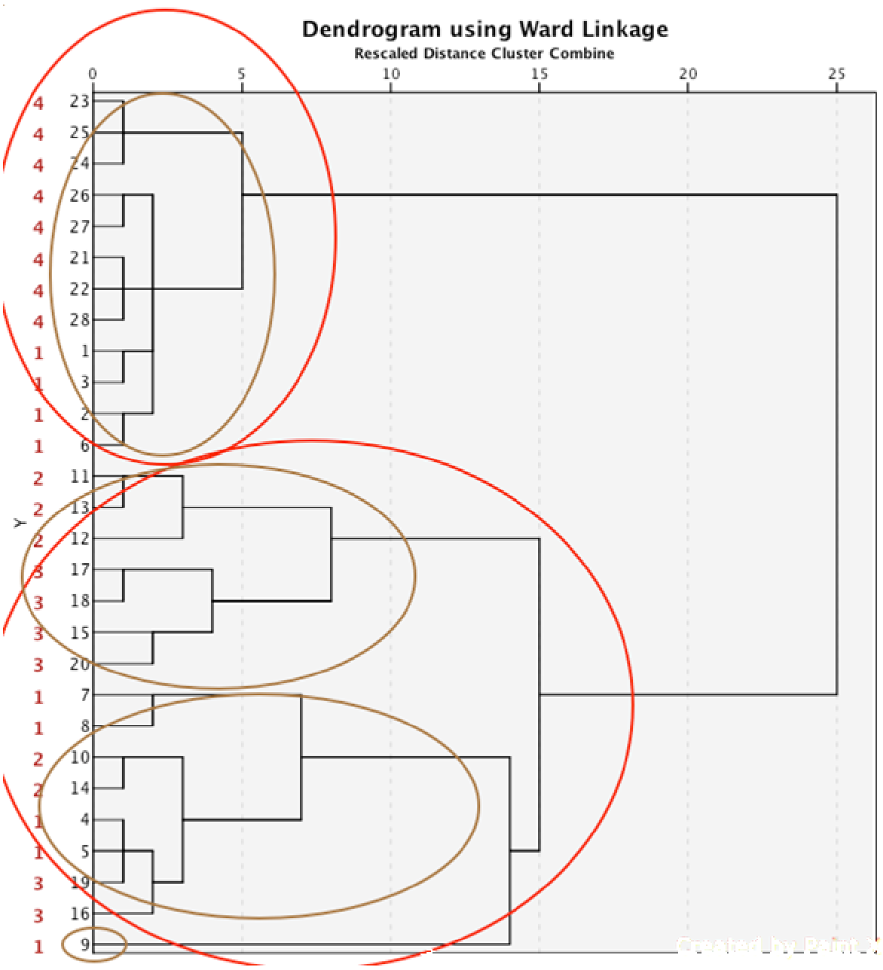

***Figure 20*. Dendogram showing the clustering of the different speech difficulties.**

Note: This is the same as the Dendogram in Figure 18, but has ellipses imposed (the reasons for this are given in the text).

Key: In far left column, on the Y axis, are the groups which the children/cases belong to where 4= fluent; 3= hearing loss, 2= phonological delay; 1= stuttering.

6.3.4. Cluster analysis using individual symptoms.

Again, using DA, individual symptoms were used to see how well they classify children for fluency, and subsequently for fluency type. To be able to determine which symptoms are best for classifying children as having or not having speech difficulty, the symptoms that are redundant are removed. This would minimize the list of symptoms that need to be included in a screening procedure, which would make implementation of the procedure and training of it easier. Twenty-nine separate symptoms were available to categorize each child into one of the four speech difficulty types in Table 36 using DA. Note that there were no cases of children with phonological disorder in this sample. The DV was the child's classification of speech difficulty type (stuttering, phonological delay, hearing loss, or fluent). The IVs were individual scores on the symptoms listed in Table 38. In other respects, the analysis was the same as that in the previous section.

The analysis revealed that six of the 29 symptoms (seen in table 36) had significant predictive power. These were: hesitations from the stuttering group of symptoms; phonological discrimination failings, pitching issues and hypo/hyper-nasality from the speech difficulties due to hearing loss group; backing from the phonological disorder group; and gliding from the phonological delay group. The overall χ^2 test was significant (Wilks λ = .026, χ^2= 80.04, df = 18, Canonical correlation = .908, p <. 001), showing that scores on these six symptoms accurately classify children into their type of speech difficulty. Three functions were generated in this analysis and they accounted for 97.4% of the variance in the classification of the children. Table 43 presents the standardized discriminant function coefficients and Table 44 shows the group centroids for functions 1 to 3 for this analysis. The new canonical variables correctly classified 89.3% of the original groupings. Figure 20 summarizes the classifications that were obtained.

Table 43. Standardized Canonical Discriminant Function Coefficients.

	Function 1	Function 2	Function 3
Hesitations	.342	.822	-.519
Backing	.481	1.261	1.264
Gliding	.001	1.016	.032
Pitch issues	.738	-.415	.424
Phon. Discrimination failings	1.057	-.302	.167
Hypo/Hyper-nasality	.593	-1.102	-.675

Table 44. Functions at group centroids.

	Function 1	Function 2	Function 3
Stuttering	.830	1.193	-1.055
Phon.Delay	-1.289	1.664	1.598
HL	2.997	-1.348	.678
Fluent	-2.361	-1.371	-.321

The first function has most discriminant power. The coefficients for function one show that symptoms for phonological discrimination failings have the highest predictive power, followed by pitch issues, hypo/hyper-nasality, backing, hesitations and gliding.

The boxplot at the right of Figure 20 right shows the groups and their means, medians and variability using box and whiskers plots. It also shows that there are two outliers, one in the stuttering group and the other in the fluent group. Apart from the outlier in the fluent group there is no overlap between the fluent group and the speech difficulty groups, which means that the function using the six symptoms mentioned above discriminates well between children who are fluent and those who have different forms of speech difficulty. However, the discrimination of the speech difficulty groups is not as clear. For instance, the discriminant function scores overlap for children with phonological delay and those who stutter. This could arise because the D scores for some of the children were close to the centroids of the two types of speech difficulty.

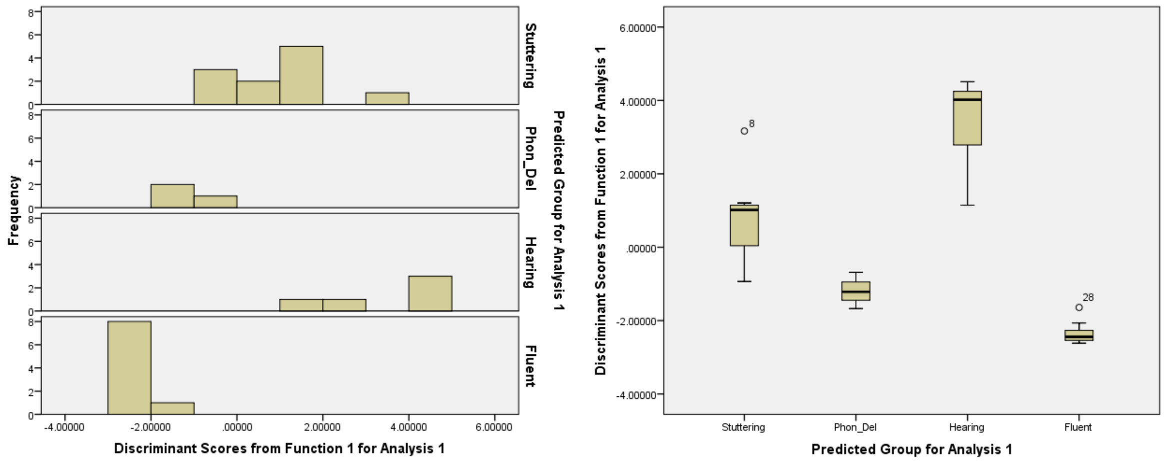

Figure 21. **Histogram (left) and boxplots (right) showing the discriminant function using individual symptoms.**

Note: Hesitations, backing, gliding, pitch-issues, phonological discrimination failings, and hypo/hyper-nasality, which were able to successfully classify children as either stuttering, phonologically delayed, hearing loss or fluent.

The function score on the x axis at left is a D score calculated by DA and represents the distance from the centroid of each group (See Appendix H for an illustrated example of centroids).

6.4. Discussion

The discriminant function analyses showed that the symptom sets were useful at classifying children as fluent or having speech difficulty and, in further analyses were performing well at classifying types of speech difficulty. The results suggested that the symptoms for phonological delay were particularly useful in selecting children with speech difficulty in this respect. The finding that phonological delay was the best predictor for being considered as being “at-risk” of speech difficulties is in line with previous research by Broomfield and Dodd (2005). They found that most children fitted into the phonological delay category (57.5%), out of the three main categories they found. Although there were suggestions that symptoms for phonological delay should be useful for classifying children into types of speech difficulty due to comorbidity, there may be a tendency to classify ambiguous symptoms as belonging to the phonological delay class. Furthermore, this chapter used a sub-sample of the original larger sample, and although they were selected at random, the outcome may change nonetheless, once the whole of the cohort is used.

Whilst the results show some promise, it should be remembered that case selection was not made to be representative of the population at large and that this can affect classification results (Howell, 2010; Reed & Wu, 2013). Thus, whilst the present results suggest that it is worth pursuing development of symptom sets to improve classifications relative to the SSI-3 set used in previous chapters, much work remains to be done.

The Dendogram analyses that were used identified six symptoms that were successful at classifying children with particular forms of speech difficulty. There was some overlap between stuttering and phonological delay and between hearing loss and stuttering. Again, whilst this is worth pursuing, future analyses need to be conducted where samples with accurate representation of the incidence of the various types of speech difficulty in the population are taken into account.

It is important in the future to analyze all the data collected in the way described in this chapter to evaluate properly specific symptoms in order to improve reliability and validity as well as to have an analysis on a sample representative at the population level. These symptom sets could then be employed as part of a training program to enable the teachers to do the screen.

Extending the number of symptoms that need to be used may be taxing when teachers perform these analyses (e.g. issues may arise when considering children with EAL, as discussed) and controversial, as this may require a change in policy that would require some redistribution of available resources.

Chapter 7

Summary and Next Steps

It is generally agreed that a child's language skills at school entry are an indicator of the child's later academic performance (Bercow Report, 2008; Department for Education: UK, 2013; Marmot Review, 2010). Furthermore, if a child has SLCN, this has an impact on behavioural and educational experiences (Conti-Ramsden & Simkin, 2012; Johnson, Beitchman & Brownlie, 2010) and their future employability (Clegg, Hollis, Mawhood, & Rutter, 2005). Given this general consensus in the literature and the recommendations and conclusions of the Government commissioned Bercow Report (2008) and subsequent work for the Department for Education in the UK (2013), it is surprising that no headway has been made towards ensuring that the issues associated with addressing SLCN in schools have been fully addressed. The seriousness of this matter is underlined by the Marmot Review (2010) report that around 53% of children are not school ready at the point at which they enter education, especially with respect to their communication and language skills.

The current thesis began by trying to reconcile alternative viewpoints in the literature concerning how the issue concerning children's SLCN could be resolved. The Bercow Report is the sole report commissioned by the UK Government that has specifically investigated the SLCN of children and the responsibilities of service providers in this field. The literature review made a clear recommendation for a national screening program for pre-school children, similar to what is available for eyesight and hearing. However, implementation of a national SLCN screening program would face many complications For example, the test used must be standardized and there should be general agreement concerning use of the test. It is no use implementing a test that the teachers find helpful but that SLTs do not trust or vice versa. It is also important to decide who will conduct the test, as this can have considerable impact on what funding for the service is needed. For example, if teachers are to be trained to conduct the test, then the costs will be lower over the long term than if an SLT conducts the test. The original research questions addressed by this thesis raised further questions and identified complications that need to be addressed. For example, cost to the government was a significant concern raised in the Bercow Report (2008). Consequently, it was suggested in this thesis that teachers should be responsible for testing. The questions this suggestion raises are considerations about how many teachers in each school need to be trained, e.g. whether all the teachers should be trained or a selection of them be given specific responsibility? Further questions then arise about teacher turnover rates and what the impact of this would be if not all teachers within each school were trained?

The issue concerning how long the test should be was raised in the survey reported in Chapter 2. Teachers preferred a short assessment whereas SLTs considered longer assessments were needed. These different views about length of a test resulted from the different goals of the test as seen by the two groups of professionals. The length of a sample in order to determine whether a child has a SLCN has been debated for some time

(Howell, 2010; Howell, 2013). It is important for schools to know: 1) how long a child will be absent from the class, if the screen is conducted in schools; and 2) that the recording obtained is long enough to provide a reliable estimate of speech difficulty. Empirical work that assessed the stability of estimates of speech analyses (SSI-3 scores) for samples of different length to inform used of such assessments in schools, was examined in the next chapter of this thesis (Chapter 3).

The SSI-3 is an instrument that fulfils many of the requirements school would impose, it also has a specification of minimum sample length needed although no check of this appears to have been made. Consequently, in Chapter 3, SSI-3 scores for portions of the same speech samples of speech which differed in length were obtained. Riley (1994) suggested that samples need to be 200-syllable long at minimum. The results in Chapter 3 showed that 200 to 250 syllables led to stable SSI-3 scores. It was also found that when sample length was decreased to 150 syllables the overall SSI-3 scores changed significantly. Together these findings showed that samples between 200 and 250 syllables provide stable SSI-3 scores. This indicated that 200 syllables can be used as the length of a speech sample for identifying speech difficulty test, especially in schools where timing constraints apply. It is important to not keep children away from the classroom for long periods of time, a point that was raised by teachers during the Survey (Chapter 2). The SSI procedure has forms that are to be used by people who can and those who cannot read. It was also shown in Chapter 3 hat the SSI scores obtained when the reader procedure was applied to two samples (spontaneous and read speech) did not differ significantly from the SSI scores obtained with the the non-reader procedure (spontaneous speech sample only was used). This indicated that a spontaneous speech sample can give an equivalent estimate to that obtained with the non-reader form. As well as suggesting s that the non-reader form that uses a spontaneous speech sample alone (all that can be obtained before children can read) is valid, it also suggests that an SSI-based procedure can be used at later ages and the results compared across the two ages and that only a spontaneous speech sample is needed at all ages.

Riley (1994) suggested that the SSI-3 can be used for distinguishing fluent children from those who stutter, which is why it has been used as the basis of a screening tool (Howell, 2013). However, neither Riley nor anyone else outside Howell's group have conducted any tests to verify that an SSI-3-based procedure can be employed in screening. The next major issues addressed was whether all components of SSI-3 were necessary when it is used as the basis of a screening tool . It is known, for example, that the ways of measuring physical concomitants component is ambiguous. It was also argued that PC may be more useful for assessing children after they have been categorized as being at-risk of speech difficulty, but not necessarily during screening. The lack of reliability of the physical concomitants measure poses serious problems if teachers were required to make them as part of screening. Thus, even when trained SLTs make physical concomitant judgments, they are unreliable and subjective (Bakhtiar et al., 2010; Lewis, 1995; Todd et al., 2014). In Chapter 4, all three components were assessed to determine whether all of them are necessary in a screening test based on SSI-3.

The assessments in Chapter 4 showed that %SS alone could screen as well as measures incorporating duration and physical concomitants since this component alone successfully discriminated fluent children from those with speech difficulty. Alteration of SSI procedures by dropping duration and physical concomitants would require that the instrument be re-standardized before it can be used when screening. It was noted that the schools would want many features present in SSI-3 retained whilst SLTs would want to change some of these. If such changes are implemented, some way of converting in-school scores to clinical SSI-3 scores would be desirable so that coordination between school and Speech and Language Therapy services can be maintained (e.g. ways of comparing scores).

It is important to bear in mind that this thesis was not concerned with standardizing the new reduced component SSI-3. Instead, the goal was simply to establish whether SSI-3 could be used to distinguish fluent children from children with speech difficulty reliably. Standardizing procedures and implementation all need to be addressed in the future once an instrument has been agreed. To this point, the thesis tackled questions about length of the speech samples to be taken from the children, whether the some components of SSI-3 can be dropped in order to make the instrument easier to implement and to train teachers on. These efficiency considerations were based on issues raised in the survey conducted with SLTs and teachers as well as a review of the research literature.

Next the thesis examined the effectiveness of the SSI-3 in distinguishing or separating children who were at-risk of speech difficulties from those who were fluent. The instrument selected children with difficulties successfully and results agreed with those provided by schools. Hence this showed that SSI-3 was a good starting point for developing a screening tool. The instrument showed a high level of specificity for fluent children where the agreement rate was almost 97% between the school and AM.

As mentioned previously, teachers require a screening instrument that is not time consuming, as their priority is to keep the children in class and not to interrupt the school curriculum. To these ends, Chapter 4 suggested that %SS alone may be sufficient for making an identification about whether a child is fluent or not.

So far, it has been shown that SSI is an effective tool for distinguishing fluent children from those who may be at risk. However, this result may not apply if an untrained person was asked to perform this test, without prior knowledge of what symptoms to look out for. Therefore, it is extremely important to establish a comprehensive but easy to train, set of symptoms that can be easily identified with some basic training. The SSI-3 includes a specified symptom set associated with stuttering. However, a screening tool for speech difficulties must also take into account a wider set of symptoms that are appropriate for other pediatric difficulties where speech is affected as well as those used for assessing stuttering. Once available, this would provide a single tool that after standardization would be used by teachers as opposed to having separate screening tools for different types of speech difficulty. A tool that selects all possible cases of speech difficulties on one test occasion will also help coordinate across school and SLT services.

The illustrative analyses in Chapter 6 examined ways in which the symptom set could be extended along these precepts. One additional question that was raised was whether identification of individual types of speech difficulties was possible if a refined symptom set was used. It is important to note that this is not something that schools would want to do (their role is education, not to identify types of SLCN). Its relevance to teachers would be if this method improved classification of children as fluent versus having speech difficulty. The application whereby types of speech difficulty are identified is potentially more relevant to SLTs (for example Dodd et al.'s, 2002, DEAP does something similar to this for clinical samples). This approach needs extensive attention in the future where samples with representative proportions of all types of speech difficulties are included. In this sense, the present work serves as a pilot for more extensive analyses.

Another criticism found in the Bercow Report (2008) was that much of the funding into the service provision for SLCN with children was not optimally used. In order to optimize the effectiveness of the screening tests, future research will need to concentrate on distinguishing children who persist with their speech difficulties from and those who will recover from them by seeing whether screening procedures can separate children into these two sets. Howell (2013) showed that for CWS severity, as measured by the SSI-3, was the most reliable factor in predicting persistence. Knowing if a child will persist or not in speech difficulties in general is of great importance here too, as at some stage SLTs may decide that it is necessary to concentrate on children who are most likely to persist with their speech difficulty. Furthermore, restrictions on funding for service provision by the NHS may make it necessary to concentrate on only those children who are likely to persist, assuming that this is a cost-effective way for the NHS. Therefore, developing a reliable test that SLTs can use to estimate likelihood of persistence, would be valuable in the future.

Another factor that should be examined further is bilingualism and multilingualism in children in the UK. Working with bilingual children poses many questions to educators and clinicians. For example, can the examiner identify symptoms unambiguously in an unfamiliar language; does the multilingual child’s linguistic proficiency have an effect on speech symptoms; are there suitable interventions for bilingual children; and finally, should intervention be conducted on one or more language sin bilingual children? The lack of research in this area leads to no answers to these questions (Bernstein Ratner, 2004; Van Borsel, Maes & Foulon, 2001). Furthermore it has been argued that children who were exposed to two or more languages from birth can still have delayed linguistic processing in all languages and this can lead to misdiagnosis of language delay or even over-identification of stuttering (Bedore & Pena, 2008; Shenker, 2011).

Conclusion

This thesis first identified the need (expressed by schools and SLTs) for a systematic procedure that could be used to identify children who have speech difficulty when they start school. Efficiency considerations were raised by schools and, following Howell (2013) the SSI-3 was identified as an appropriate starting point. Checks were made that the minimum length Riley specified was adequate to provide a stable measure. Also, for future developments, the critical component that worked when identifying children with speech difficulty was %SS. A field test was conducted using SSI-3 and this showed reliable and acceptable levels of classification of children as having speech difficulty compared to case identification by schools. The final chapter explored whether refined symptom sets may improve classification and allow some enhancement in identification of children with different types of speech difficulty.

Much remains to be done before a screening tool for use by teachers becomes available. The directions this thesis suggests are that this should be based on symptom measures. Further questions that need attention are to ensure that what is done in schools when identifying children with speech difficulty is fully understood and closely scrutinized by SLTs. Similarly, if in-school interventions are conducted, care should be taken to ensure they are compatible with SLT interventions.

References

Abdullah, A. N. (2012). Effect of context on types of hesitation strategies used by Iranian EFL learners in L2 Oral Language Tests. *English Language Teaching*, *5,* 102-109.

Ajdacic-Gross, V., Vetter, S., Müller, M., Kawohl, W., Frey, F., Lupi, G. & Rössler, W. (2010). Risk factors for stuttering: a secondary analysis of a large data base. *European archives of psychiatry and clinical neuroscience*, *260*(4), 279-286.

American Psychiatric Association. (2000). Stuttering. In M. B. First (Ed.), *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (4th ed.) (pp. 67-69). Arlington, VA: American Psychiatric Association.

Andrews, G., & Harris, M. (1964). *The syndrome of stuttering, Clinics in developmental medicine, No. 17*. London: William Heineman Medical Books Ltd.

Antoniazzi, D., Snow, P., & Dickson-Swift, V. (2010). Teacher identification of children at risk for language impairment in the first year of school. *International Journal of Speech-Language Pathology, 12*(3), 244-252.

Archibald, L. M., Joanisse, M. F., & Shepherd, M. (2008). Associations Between Key Language-Related Measures in Typically Developing School-Age Children. *Journal of Psychology, 216*(3), 161-171.

Arcuri, C. F., Osborn, E., Schiefer, A. M., & Chiari, B. M. (2009). Speech rate according to stuttering severity. *Pro-Fono*, *21*, 46-50.

Au-Yeung, J., Howell, P., Davis, S., Charles, N., & Sackin, S. (2000). UCL survey on bilingualism and stuttering. In H.-G. Bosshardt, J. S. Yaruss, & H. F. M. Peters (Eds.), *Fluency disorders: Theory research treatment self-help* (pp. 129–132). Nijmegen, the Netherlands: Nijmegen University Press.

Bada, E. (2010). Repetitions as vocalized fillers and self-repairs in English and French interlanguages. *Journal of Pragmatics, 42*(6), 1680-1688.

Bakhtiar, M., Seifpanahi, S., Ansari, H., Ghanadzade, M., & Packman, A. (2010). Investigation of the reliability of the SSI-3 for preschool Persian-speaking children who stutter. *Journal of Fluency Disorders*, *35*, 87–91.

Barbarin, O. A. (2007). Mental health screening of preschool children: Validity and reliability of ABLE. *American Journal of Orthopsychiatry, 77*(3), 402-418.

Barry, W. J. (1983). Some problems of interarticulator phasing as an index of temporal regularity in speech. *Journal of Experimental Psychology*, *9*, 826-828.

Beal, D. S., Quraan, M. A., Cheyne, D. O., Tayloe, M. J., Gracco, V. L., & Luc, F. D. N. (2011). Speech-induced suppression of evoked auditory fields in children who stutter. *NeuroImage, 54,* 2994-3003

Bedore, L. M.,& Pena, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11, 1–29.

Bercow Review Advisory Group. (2008). *The Bercow Report. A review of services for children and young people (0–19) with speech, language and communication needs.* Nottingham: DCSF Publications.

Bernstein Ratner, N. (1997). Stuttering: A psycholinguistic Perspective. In R.F. Curlee & G. Siegel (Eds.), *Nature and treatment of stuttering: New directions* (2nd edition) (pp.99-127). Needham, MA: Allyn & Bacon.

Bernstein Ratner, N. (2004). Fluency. In B. Goldstein (Ed.), *Language development: A focus on the Spanish–English speaker* (pp. 287–310). Baltimore, MD: Brookes.

Bernstein Ratner, N., & Benitez, M. (1985). Linguistic analysis of a bilingual stutterer*. Journal of Fluency Disorders*, 10, 211–219.

Bernstein Ratner, N., & Wijnen, F. (2007). The Vicious Cycle: Linguistic encoding, self-monitoring and stuttering. In J. Au- Yeung (Ed.), *Proceedings of the 5th World Congress of fluency disorders* (pp. 84-90). Dublin, Ireland: International Fluency Association.

Bishop, D. V. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders, 49*(4), 381-415.

Bishop, D. V., & Mcdonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Language and Communication Disorders, 44*(5), 600-615.

Bloodstein, O. (1995). *A handbook on stuttering*. San Diego, CA: Singular.

Bloodstein, O., Ratner, N.B., (2008). A Handbook on Stuttering, 6th ed. Thomson Delmar Learning, Clifton Park, N.Y.

Borden, Gloria J. (1979). "An Interpretation of Research on Feedback Interruption in Speech.” *Brain & Language,* 7, 307-319.

Braun, A.R., Varga, M., Stager, S.,Brocklehurst, P.H., (2013). Stuttering prevalence, incidence and recovery rates. *Journal of Fluency Disorders, 38*, 290-293.

Brown, S. F. (1945). *The loci* of stutterings in the *speech sequence*. *Journal of Speech* Disorders, *10*, 181-192.

Brocklehurst, P. H., (2013). Stuttering prevalence, incidence and recovery rates. *Journal of Fluency Disorders, 38*, 290–293.

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

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

Brosch, S., Haege, A., Kalehne, P., & Johannsen, H. S. (1999). Stuttering children and the probability of remission-the role of cerebral dominance and speech production. *International Journal of Pediatric Otorhinolaryngology*, *47*(1), 71-76.

Campbell, T. F., Dallaghan, C. A., Rockette, H. E., Paradise, J. L., Feldman, H. M., Shriberg, L. D., Sabo, D. L. & Kurs-Lasky, (2003). Risk factors for speech delay of unknown origin in 3-year-old children. *Child Development*, *74*, 346-357.

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders - a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry, 46*(2), 128-149.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for Specific Language Impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*(6), 741-748.

Conture, E. G., & Kelly, E. M. (1991). Young stutterers' nonspeech behaviors during stuttering. *Journal of Speech and Hearing Research, 34,* 1041–1056.

Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders, 34*(2), 61-71.

De Andrade, C.R., Cervone, L. M. & Sassi, F. C. (2003). Relationship between the stuttering severity index and speech rate. *Sao Paulo Medical Journal*, *121*, 81-84.

Department for Education. Early years foundation stage profile results: 2012 to 2013 2013 [07/04/2014]. Available from: https://www.gov.uk/government/publications/early-years-foundation-stage-profile-results-2012-to-2013.

Dockrell, J. & Howell, P. (2015). Meeting the needs of children with Speech Language and Communication Difficulties: Challenges and opportunities. *British Journal of Special Education*, *42*, 411-428.

Dodd, B. (1995). Children with speech disorder: Defining the problem. In B.Dodd (Ed.), *Differential Diagnosis and Treatment of Children with Speech Disorder.* 2nd ed. (pp. 1-19). London: Whurr.

Dodd, B. (2007) Evidence-based practice and speech-language pathology. *Folia Phoniatrica Logopedia*, *59*,118-129.

Dodd, B., Zhu, H., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic evaluation of articulation and phonology DEAP*. London, England: Psychological Corporation.

Durkin, K., Conti-Ramsden, G., & Simkin, Z. (2012). Functional Outcomes of Adolescents with a History of Specific Language Impairment (SLI) with and without Autistic Symptomatology. *Journal of Autism and Developmental Disorders, 42*(1), 123-138.

Dworzynski, K., Remington, A., Rijksdijk, F., Howell, P., & Plomin, R. (2007). Genetic etiology in cases of recovered and persistent stuttering in an unselected, longitudinal sample of young twins. *American Journal of Speech-Language Pathology.16,* 169–178.

Einarsdottir, J., & Ingham, R. J. (2009). Does language influence the accuracy of judgments of stuttering in children? *Journal of Speech, Language and Hearing Research, 52*, 766–779.

Ek, U., Norrelgen, F., Westerlund, J., Dahlman, A., Hultby, E., & Fernell. E. (2012). Teenage outcomes after speech and language impairment at preschool age. *Neuropsychiatric Disorders Treatment*, *8*, 221-7.

Fairbanks, G. (1955). Selective vocal effects of delayed auditory feedback . *Journal of Speech and Hearing Disorders*, *20*, 333-346.

Fehringer, C. & Fry, C. (2007). Hesitations phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica*, *41*, 37-72.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*, 378–382.

Forrest, K. (2003). Diagnostic criteria of developmental apraxia of speech used by clinical speech-language pathologists. *American Journal of Speech-Language Pathology*, 12(3), 376.

Fricke, S., Bowyer-Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry, 54*(3), 280-290.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*(2), 103-127.

Genesee, F. (2001). Bilingual first language acquisition: Exploring the limits of the language faculty. In M. McGroarty (Ed.), *21st annual review of applied linguistics* (pp. 153–168). Cambridge, UK: Cambridge University Press.

Guenther, F. (2008). Neuroimaging of normal speech production. In R. J. Ingham (Ed.), *Neuroimaging in communication sciences and disorders* (pp. 1-52). San Diego, CA: Plural Pub.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*(5), 350-365.

Guenther, F. H., Brumberg, J. S., Wright, E. J., Nieto-Castanon, A., Tourville, J. A., Panko, M., et al. (2009). A wireless brain-machine interface for real-time speech synthesis. *PLoS ONE, 4*(12), e8218.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modelling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*(3), 280-301.

Guenther, F. H., Husain, F. T., Cohen, M. A., & Shinn-Cunningham, B. G. (1999). Effects of categorization and discrimination training on auditory perceptual space. *Journal of the Acoustical Society of America, 106*(5), 2900-2912.

Guenther, F.H. and Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*. *25*, 408-422.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*, 31–38.

Holm, A., & Dodd, B. (1999). An intervention case study of a bilingual child with phonological disorder. *Child Language Teaching and Therapy, 15*(2), 139-158.

Holm, A., Dodd, B., Stow, C., & Pert, S. (1999). Identification and differential diagnosis of phonological disorder in bilingual children. *Language Testing, 16*(3), 271-292.

Holm, A., Ozanne, A., & Dodd, B. (1997). Efficacy of intervention for a bilingual child making articulation and phonological errors. *International Journal of Bilingualism, 1*(1), 55-69.

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience, 14*(8), 1125-1138.

Howell, P. (2004). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communication Science and Disorders*, *31*, 122-139.

Howell, p. (2007a). Signs of developmental stuttering up to age 8 and at 12 plus. *Clinical Psychology Review*, *27*, 287-306.

Howell, P. (2007b). A model of serial order problems in fluent, stuttered and agrammatic speech. *Human Movement Science*, *26*, 728-741.

Howell, P. (2010a). Language processing in fluency disorders. In J. Guendouzi, F. Loncke, & M. Williams (Eds.). *The handbook of psycholinguistics and cognitive processes: Perspectives on communication disorders*. London, U.K.: Taylor and Francis.

Howell, P. (2010b). *Recovery from Stuttering*. New York: Psychology Press.

Howell, P. (2013). Screening school-aged children for risk of stuttering. *Journal of Fluency Disorders, 38*(2), 102-123.

Howell, P., Au-Yeung, J., & Pilgrim, L. (1999). Utterance rate and linguistic properties as determinants of speech disfluency in children who stutter. *Journal of the Acoustical Society of America, 105*, 481-490.

Howell, P., Bailey, E., & Kothari, N. (2010). Changes in the pattern of stuttering over development for children who recover or persist. *Clinical linguistics and phonetics*, *24*(7), 556-575.

Howell, P., & Davis, S. (2011). Predicting persistence of and recovery from stuttering at teenage based on information gathered at age eight. *Journal of Developmental and Behavioral Pediatrics, 32*, 196–205.

Howell, P., Davis, S., & Williams, R. (2008). Late childhood stuttering. *Journal of Speech, Language and Hearing Research, 51*, 669–687.

Howell, P., Davis, S., & Williams, R. (2008). The effects of bilingualism on speakers who stutter during late childhood. *Archives of Disease in Childhood, 94*(1), 42–46.

Howell, P., Ruffle, L., Fernández-Zúñiga, A., Gutiérrez,R., Fernández, A. H., O'Brien, M. L., Tarasco, M., Vallejo-Gomez, I., & Au-Yeung, J. (2004). Comparison of exchange patterns of stuttering in Spanish and English monolingual speakers and a bilingual Spanish-English speaker. *In Theory, Research and Therapy in Fluency Disorders*. Pp. 415-422. A. Packman, A. Meltzer & H. F. M. Peters (Eds.) Nijmegen: Nijmegen University Press.

Howell, P., Soukup-Ascencao, T., Davis, S., & Rusbridge, S. (2011). Comparison of alternative methods for obtaining severity scores of the speech of people who stutter*. Clinical Linguistics and Phonetics*, *25,* 368–378.

Howell, P., Wingfield, T. &Johnson, M. (1988). Characteristics of the speech of stutterers during normal and altered auditory feedback. *In Proceedings Speech 88*, vol. 3. W.A. Ainsworth& J.N. Holmes (Eds.). pp. 1069-1076. Edinburgh: Institute of Acoustics.

Huckvale, M. (2013). *Speech filing system Windows version 1.75 [computer software].* London: University College London.

Jani, L., Huckvale, M., & Howell, P. (2013). Procedures used for assessment of stuttering frequency and stuttering duration. *Clinical Linguistics and Phonetics*, *27*, 853–861.

Jankelowitz, D. L., & Bortz, M. A. (1996). The interaction of bilingualism and stuttering in an adult. *Journal of Communication Disorders*, 29, 223–234.

Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-Year Follow-Up of Children With and Without Speech-Language Impairments: Family, Educational, Occupational, and Quality of Life Outcomes. *American Journal of Speech-Language Pathology, 19*(1), 51.

Johnson, W., Boehmler, R., Dahlstrom, W., Darley, F., Goodstein, L., & Kools, J. (1959). *The onset of stuttering: Research findings and implications*. Minneapolis: University of Minnesota Press.

Johnson, W., & Rosen, P. (1937). Studies in the psychology of stuttering: VII. Effects of certain changes in speech pattern upon stuttering frequency. *Journal of Speech and Hearing Disorders*, *2,* 105-109.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT.

Kohnert, K., Yim, D., Nett, K., Kan, P. F., & Duran, L. (2005). Intervention With Linguistically Diverse Preschool Children: A focus on developing home language(s). *Language Speech and Hearing Services in Schools, 36*(3), 251-263.

Kolk, H. H. J., & Postma, A. (1997). Stuttering as a covert-repair phenomenon. In R. F. Curlee & G. Siegel (Eds.), *Nature and Treatment of Stuttering: New directions* (pp. 182-203). Boston, MA: Allyn & Bacon.

Landa, R., Holman, K. C., & Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and alter diagnosis of autism spectrum disorders. *Archives of General Psychiatry*, 64, 853–864.

Leckman, J. F., Walker, D. E., & Cohen, D. J. (1993). Premonitory urges in Tourette's syndrome. *American Journal of Psychiatry, 150*(1), 98-102.

Lee, B. (1950). Effects of delayed speech feedback. *Journal of the Acoustic Society of America*, *22*, 824-826.

Lee, X. (2014). *A structured interview with teachers and speech language therapists to establish how children with speech difficulties are identified in schools and referred for intervention* (Unpublished project). Division of Psychology and Language Sciences, University College London (UCL).

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Logan, K., & Haj-Tas, M. (2007). Effect of sample size on the measurement of stutter-like disfluencies. *Perspectives on Fluency and Fluency Disorders, 17*, 1–6.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-75.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lewis, K. E. (1995). Do SSI-3 scores adequately reflect observations of stuttering behaviors? *American Journal of Speech-Language Pathology, 4*, 46–59.

Manning, W. H., & Shirkey, E. A. (1981). Fluency and the aging process. In D. S. Beasley & G. A. Davis (Eds.), *Aging: Communication processes and disorders* (pp. 175–189). New York, NY: Grune & Stratton.

Månsson, H. (2000). Childhood stuttering: Incidence and development. *Journal of Fluency Disorders*, *25*(1), 47-57.

The Marmot Review Team. (2010). *Fair Society, Healthy Lives: Strategic review of health inequalities in England post-2010*. London: Marmot Review Team.

Marshall, J., Ralph, S., & Palmer, S. (2002). 'I wasn't trained to work with them': Mainstream teachers' attitudes to children with speech and language difficulties. *International Journal of Inclusive Education, 6*(3), 199-215.

Masoura, E. V., & Gathercole, S. E. (1999). Phonological Short-term Memory and Foreign Language Learning. *International Journal of Psychology, 34*(5-6), 383-388.

Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H. F. M. Peters, P. H. H M. Van Lieshout & W. Hulstijn (Eds.). *Speech motor control in normal and disordered speech* (pp. 357–387). Oxford, UK: Oxford University Press.

Max, L., Guenther, F. H., Gracco, V.L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Sciences and Disorders, 31*, 105-122.

McLean, L. K., & Cripe, J. W. (1997). The effectiveness of early intervention for children with communication disorders. *The effectiveness of early intervention*, 349-428.

McLeod, S., & Goldstein, B. (2012). *Multilingual aspects of speech sound disorders in children*. Bristol: Multilingual Matters.

McLeod, S., Harrison, L. J., & McCormack, J. (2012). The Intelligibility in Context Scale: Validity and Reliability of a Subjective Rating Measure. *. Journal of Speech, Language, and Hearing Research, 55*(2), 648-656.

McLeod, S., Verdon, S., Bowen, C., & The International Expert Panel on Multilingual Children's Speech. (2013). International aspirations for speech-language pathologists' practice with multilingual children with speech sound disorders: Development of a position paper. *Journal of Communication Disorders, 46*(4), 375-387.

Merritt HH, Rowland LP, ed. (2000). *Merritt's neurology*. 10th ed. Philadelphia: Lippincott Williams & Wilkins.

Mirawdeli, A. (2015). Identifying children who stutter or have other difficulties in speech production in school reception classes. *Procedia - Social and Behavioral Sciences*, *193*, 192-201.

Mirawdeli, A, Dockrell, J. & Howell, P. (in Prep). A survey of UK Teachers' and Speech and Language Therapists' joint roles with respect to reception class children's Expressive Speech Difficulties.

Mirawdeli, A. & Howell, P. (in prep). Is it necessary to assess fluent symptoms, duration of dysfluent events and physical concomitants when identifying children who are at risk of speech difficulty? *Clinical Linguistics and Phonetics.*

Murphy, V.A. (2015). *A Systematic Review of Intervention Research Examining English Language and Literacy Development in Children with English as an Additional Language (EAL).* Reading: NALDIC [online]. Available: http://www.naldic.org.uk/Resources/NALDIC/Research%20and%20Information/Documents/eal-systematic-review-prof-v-murphy.pdf [February, 2015]. Website: National Association for Language Development in the Curriculum (NALDIC).

Nelson, H. D., Nygren, P., Walker, M., & Panoscha, R. (2006). Screening for speech and language delay in preschool children: Systematic evidence review for the US preventitive services task force. *Pediatrics, 117,* e298-e319.

Northam, G. B., Liégeois, F., Chong, W. K., Baker, K., Tournier, J., Wyatt, J. S., Baldeweg, T., Morgan, A. (2012). Speech and oromotor outcome in adolescents born preterm: Relationship to motor tract integrity. *The Journal of Pediatrics, 160*(3), 402-408.

Nwokah, E. E. (1988). The imbalance of stuttering behaviour in bilingual speakers*. Journal of Fluency Disorders*, 13, 357–373

Onslow, M., Costa, L., & Rue, S. (1990). Direct early intervention with stuttering – Some preliminary data. *Journal of Speech and Hearing Disorders, 55*, 405–416.

Pampel, F. C. 2000. *Logistic Regression: A primer*. Thousand Oaks, CA: Sage.

Penfield, W., & Roberts, L. (1959). *Speech and brain-mechanisms*. Princeton, NJ: Princeton Univ. Press.

Pert, S., & Letts, C. (2006). Codeswitching in Mirpuri speaking Pakistani heritage preschool children: Bilingual language acquisition. *International Journal of Bilingualism*, 10, 349–374.

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research, 36*, 472–487.

Peters, H. F. M., & Starkweather, C. W. (1990). The interaction between speech motor coordination and language processes in the development of stuttering: Hypotheses and suggestions for research. *Journal of Fluency Disorders*, *15*, 115–125.

Peters, H. F. M., Hulstijn, W., & van Lieshout, P. H. H. M. (2000). Recent developments in speech motor research into stuttering. *Folia Phoniatrica et Logopaedica*, *52*, 103–119.

Peyre, H., Bernard, J.Y., Forhan, A., Charles, M. A.., De Agostini, M., Heude, B., & Ramus, F. (2014). Predicting changes in language skills between 2 and 3 years in the EDEN mother-child cohort*. PeerJ*, 1;2:e335.

Rad, S. K., & Abdullah, A. N. (2012). Effect of context on types of hesitation strategies used by Iranian EFL learners in L2 oral language tests. *English Language Teaching, 5*(7), 102-109.

Reilly, S., Onslaw. M., Packman, A., Wake, M., Bavin, E. L., Prior, M., Eadie, P. ... Ukoumunne, O.C. (2009). Predicting stuttering onset by the age of three years: A prospective, community cohort study. *Pediatrics, 123*, 270-277.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., . . . Bretherton, L. (2010). Predicting Language Outcomes at 4 Years of Age: Findings from Early Language in Victoria Study. *Pediatrics, 126*(6), e1530-e1537.

Reilly, S., Tomblin, B., Law, J., Mckean, C., Mensah, F. K., Morgan, A., . . . Wake, M. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders, 49*(4), 416-451.

Ribeiro, L. A., Zachrisson, H. D., Schjoberg, S., Aase, H., Rohrer-Baumgartner, N., & Magnus, P. (2011). Attention problems and language development in preterm low birth weight children: Cross-lagged relations from 18-36 months. *BioMed Central, Pediatrics*, *11*, 59.

Riley, G. (1981). *Stuttering prediction instrument*. Austin, TX: PRO-ED.

Riley, G. (1994). *The Stuttering Severity Instrument for Adults and Children (SSI-3)* (3rd ed.). Austin, TX: PRO-ED.

Riley, G. (2009). *The Stuttering Severity Instrument for Adults and Children (SSI-4)* (4th ed.). Austin, TX: PRO-ED.

Roberts, P. M. (2002). Disfluency patterns in four bilingual adults who stutter. *Journal of Speech-Language Pathology and Audiology*, *26*, 5–19.

Roberts, P., Meltzer, A., & Wilding, J. (2009). Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of Communication Disorders, 42,* 414–427.

Roberts, P. M., & Shenker, R. C. (2007). Assessment and treatment of stuttering in bilingual speakers. In E. G. Conture, & R. F. Curlee (Eds.), *Stuttering and related disorders of fluency*. New York: ThiemeMedical Publishers.

Rousseau, I., Onlsow, M., & Packman, A. (2008). Comparison of audio and audio visual measures of stuttering frequency and severity in preschool children. *American Journal of Speech Language Pathology, 17*, 173–178.

Ruangdaraganon, N., Chutapisith, J., Mosuwan L, Kriweradechachai, S., Udomsubpayakul, U. & Choprapawon, C. (2009). Television viewing in Thai infants and toddlers: impacts to language development and parental perceptions. *BioMed Central, Pediatrics*, *9*(1), 34.

Samelli, A., Rondon, S., Oliver, F., Junqueira, S., & Molini-Avejonas, D. (2014). Referred speech-language and hearing complaints in the western region of São Paulo, Brazil. *Clinics, 69*(6), 413-419.

Sawyer, J., & Yairi, E. (2006). The effect of sample size on the assessment of stuttering severity. *American Journal of Speech-Language Pathology, 15*, 36–44.

Schwartz, H. D., & Conture, E. G. (1988). Subgrouping young stutterers: Preliminary behavioral observations. *Journal of Speech and Hearing Research, 31*, 62–71.

Segalowitz, S.J., & Brown, D. (1991). Mild head injury as a source of developmental disabilities. *Journal of Learning Disabilities. 24,* 551–559.

Seider, R., Gladstein, K., & Kidd, K. (1983). Recovery and persistence of stuttering among relatives of stutterers. *The Journal of speech and hearing disorders. 48*, 394–402.

Shenker, R. C. (2011). Multilingual children who stutter: Clinical issues*. Journal of Fluency Disorders*, *36*, 186–193.

Siegel, M., Iozzi, T., & Surian, L. (2009). Bilingualism and conversational understanding in young children. *Cognition*, *110*, 115–122.

Sim, F., O'Dowd, J., Thompson, L., Law, J., Macmillan, S., Affleck, M., . . . Wilson, P. (2013). Language and social/emotional problems identified at a universal developmental assessment at 30 months. *BioMed Central Pediatrics, 13*(1), 206.

Skarżyński, H., & Piotrowska, A. (2012). Prevention of communication disorders – screening pre-school and school-age children for problems with hearing, vision and speech: European Consensus Statement. *Medical Science Monitor, 18*(4), SR17-SR21.

Skeat, J., Wake, M., Ukoumunne, O. C., Eadie, P., Bretherton, L., & Reilly, S. (2013). Who gets help for pre-school communication problems? Data from a prospective community study. *Child: Care, Health and Development, 40*(2), 215-222.

Smith, A., & Kleinow, J. (2000). Kinematic correlates of speaking rate changes in stuttering and normally fluent adults. *Journal of Speech, Language, and Hearing Research*, *43,* 521–536.

Snedecor, G. W. (1956). Statistical methods, fifth edition. Iowa State University Press, Ames, Iowa.

Stern, E. (1948). A preliminary study of bilingualism and stuttering in four Johannesburg schools. *Journal of Logopaedics*, 1, 15–25

Stevens, J. P. (2002). *Applied Multivariate Statistics for the Social Sciences*, Fourth Edition. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Stiles, D. J., Bentler, R. A. and McGregor, K. K., (2012), 'The speech intelligibility index and the pure-tone average as predictors of lexical ability in children fit with hearing aids.' *Journal of Speech, Language, and Hearing Research*, 55, 764–778.

Stow, C., & Dodd, B. (2003). Providing an equitable service to bilingual children in the UK: A review. *International Journal of Language and Communication Disorders, 38*(4), 351-377.

Strand, S., Malmberg, L. and Hall, J. (2015). *English as an Additional Language (EAL) and Educational Achievement in England: an Analysis of the National Pupil Database*. London: EEF [online]. Available: http://educationendowmentfoundation.org.uk/uploads/pdf/EAL_and_educational_achievement2.pdf [February, 2015]. Website: Education Endowment Foundation (EEF).

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Thordardottir, E. (2006). Language intervention from a bilingual mindset. *The ASHA Leader*, *11*, 20–21.

Todd, H., Mirawdeli, A., Costelloe, C., Cavenagh, P., Davis, S.., & Howell, P. (2014). Scores on Riley's Stuttering Severity Instrument versions three and four for samples of different length and for different types of speech material. *Clinical Linguistics and Phonetics, 28,* 912-926.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage, 39*(3), 1429-1443.

Travis, L. E., Johnson, W., & Shover, J. (1937). The relationship of bilingualism to stuttering. *Journal of Speech Disorders*, *2*, 185–189.

Tsao, Y., & Weismer, G. (1997). Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech, Language, and Hearing Research, 40*(4), 858-866.

Van Borsel, J., & Britto Pereiria, M. (2005). Assessment of stuttering in a familiar versus an unfamiliar language. *Journal of Fluency Disorders,* 30, 109–124.

Van Borsel, J., Maes, E., & Foulon, S. (2001). Stuttering and bilingualism: A review*. Journal of Fluency Disorders*, 26, 179–205.

Van Borsel, J., Leahy, M. M., & Britto Pereiria, M. (2008). Judging stuttering in an unfamiliar language: The importance of closeness to the native language. *Clinical Linguistics and Phonetics*, 22, 59–67.

van Lieshout, P. H. H. M., Hulstijn, W., & Peters, H. F. M. (2004). Searching for the weak link in the speech production chain of people who stutter: A motor skill approach. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 313–355). New York, NY: Oxford University Press.

Van Lieshout, P. H. H. M., Hulstijn, W., & Peters, H. F. M. (1996a). From planning to articulation in speech production: What differentiates a person who stutters from a person who does not stutter? *Journal of Speech and Hearing Research, 39*(3), 546-564.

Van Lieshout, P. H. H. M., Hulstijn, W., & Peters, H. F. M. (1996b).Speech production in people who stutter: Testing the motor plan assembly hypothesis. *Journal of Speech and Hearing Research, 39*(1), 76-92.

Van Lieshout, P. H. H. M., Peters, H. F. M., Starkweather, C. W., & Hulstijn, W. (1993). Physiological differences between stutterers and nonstutterers in perceptually fluent speech: EMG amplitude and duration. *Journal of Speech and Hearing Research, 36*, 55-63.

Wake, M., Levickis, P., Tobin, S., Zens, N., Law, J., Gold, L., . . . Reilly, S. (2012). Improving outcomes of preschool language delay in the community: Protocol for the Language for Learning randomised controlled trial. *BioMed Central Pediatrics, 12*(1), 96.

Wake, M., Tobin, S., Girolametto, L., Ukoumunne, O. C., Gold, L., Levickis, P., . . . Reilly, S. (2011). Outcomes of population based language promotion for slow to talk toddlers at ages 2 and 3 years: Let's Learn Language cluster randomised controlled trial. *British Medical Journal, 343*, D4741-D4741.

Ward, D. (2013). Risk factors and stuttering: Evaluating the evidence for clinicians. *Journal of Fluency Disorders, 38*(2), 134-140.

Ward, J. H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, *58*, 236–244.

Westby, C. (1990). Ethnographic interviewing, asking the right questions to the right people in the right way. *Journal of Childhood Communication Disorders, 13*, 101–111.

Wexler, K. B., & Mysak, E. D. (1982). Disfluency characteristics of 2-, 4-, and 6-year-old males. *Journal of Fluency Disorders*, *7*, 37–46.

Windsor, J., Kohnert, K., Lobitz, K. F., & Pham, G. T. (2010). Cross-Language Nonword Repetition by Bilingual and Monolingual Children. *American Journal of Speech-Language Pathology, 19*(4), 298-310.

Wingate, M. (1964). A standard definition of stuttering. *Journal of Speech Hearing Disorders, 29*, 484-489.

Wium, A., & Louw, B. (2015). The South African national school curriculum: Implications for collaboration between teachers and speech-language therapists working in schools. *South African Journal of Childhood Education, 5*(1), 19-41.

Wong-Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly*, 6, 323–346.

Woods, D. W., Piacentini, J., Himle, M.B. & Chang S. (2005). Premonitory urge for tics scale (PUTS): Initial psychometric results and examination of the premonitory urge phenomenon in youths with tic disorders. *Journal of Developmental and Behavioral Pediatrics*, *26*, 397–403.

Wright, R.E. (1995). Logistic Regression. In L. G. Grimm & P. R. Yarnold, (Eds.), *Reading and Understanding Multivariate Statistics* (Pp. 217-244). Washington, DC: American Psychological Association.

Yairi, E. (2007). Subtyping stuttering I: A review. *Journal of Fluency Disorders, 32,* 165-196.

Yairi, E., & Ambrose, N. (1992). A longitudinal study of stuttering in young children: A preliminary report. *Journal of Speech and Hearing Research*, 35, 755-760.

Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering. I: Persistency and recovery rates. *Journal of Speech and Hearing Research*, 42, 1097–1112.

Yairi, E., & Ambrose , N. (2004). *Early Childhood Stuttering for Clinicians by Clinicians*. Austin, TX: Pro-Ed.

Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering for clinicians by clinicians*. Austin, TX: PRO-ED.

Yairi, E., Ambrose, N. G., Paden, E. P., & Throneburg, R. N. (1996). Predictive factors of persistence and recovery: Pathways of childhood stuttering. *Journal of Communication Disorders*, 29, 51–77.

Yairi, E., Ambrose, N., Watkins, R., & Paden, E. (2001). What is stuttering? *Journal of Speech, Language, and Hearing Research, 44*, 585–592.

Yairi, E., & Seery, C. H. (2011). *Stuttering: Foundations and clinical applications*. Upper Saddle River, NJ: Pearson.

Yaruss, J. S. (1997). Clinical implications of situational variability in preschool children who stutter. *Journal of Fluency Disorders, 22*, 187–203.

APPENDICES

Appendix A

The school recruitment letter

«School_Name»
«Address_Line_1»
«Address_Line_2»
«City»
«Post_Code»

06 October 2016

Ref: **Assessing speech fluency problems in children aged 4 to 5 years.**

«GreetingLine»

The Speech team, headed by Professor Peter Howell, in the Department of Psychology and Language Sciences at the University College London (UCL), has secured funding for the next three years to investigate the assessment of speech fluency in typically developing children aged 4 to 5.

We are hoping to assess children during their reception or first year of school for any communication problems and to investigate the effectiveness of instruments currently used for this purpose. Data will be collected from 1000 pupils, their parents and/or guardians and the teachers. Data from the children will be collected in the form of voice and/or video recordings. «School_Name» has been identified as a school that fits our inclusion criteria and we would like to invite you to work with us on this project.

This study has been approved by UCL's ethical committee and has a data protection registration number. Data will be handled by the UCL Speech team who have CRB clearing and will be dealt with confidentially. Throughout the duration of the study, personal information will not be revealed, and teachers', parents', or pupils' names will not be disclosed in any reports. In the unlikely event that this may be altered then the originators' permission and consent will be requested.

For specific information about this study and the benefits it may have for your school, we have enclosed a leaflet. If you would like further details on what participation in the project would entail, please return the attached reply slip or contact Ms Avin Mirawdeli (details below) by phone or email. Please also note that this letter may be followed up by a phone call from Avin Mirawdeli at a later date.

Yours sincerely,

Avin Mirawdeli

Ref: **Assessing speech fluency problems in typically developing children aged 4 to 5 years old.**

Name: ………………………………………………………………………………………………….

My school might be interested in participating in this study of assessing children for fluency problems.

School name AND address:
……
………………………………………………………………………………………………….

School head teacher:
………………………………………………………………………………………..

The best way of contacting me is:

☐ Phone me on …………………………………………………….

☐ Phone the school office to make an appointment to phone me on ………………………………………………………………………….

☐ E-mail me at: …………………………………………………….

My full contact details are:

(Please include names with titles, position, telephone number &/or email address)

……
……
……
……
…………………………………………

Appendix B

Information leaflet sent to schools with recruitment letter.

Final Feedback

At the end of the study, the following will be provided to all participating schools:

- A summary document: detailing the procedure carried out, how the data were used and analysed and the implications of the results so that you are aware of the impact of the research.
- A booklet for the schools: explaining how to identify children who may be at risk of communication disorders, how to speak to the parents and advice on referrals.
- A leaflet for parents: this can be given out by the school to parents. It will detail communication disorders, symptoms, signs to watch out for and advice on what to do next.

Further information

This research has ethical approval from UCL No.: 4374/001

Data protection act registration number from UCL: Z6364106/2012/10/31

All researchers involved in this project have recent enhanced CRB checks.

This research is funded by the UCL IMPACT award and the Dominic Barker Trust.

For more information about the funding body please see the website bellow:

http://www.dominicbarkertrust.org.uk

The Dominic Barker Trust for research into stammering

Dom's Fund

The Dominic Barker Trust Hon DCL (East Anglia)

Department of Psychology
University College London
Gower Street
London WC1E 6BT

Phone: +44 (0) 20 7679 7566 Prof. Howell
Phone: +44 (0) 20 7679 5401 Ms. Mirawdeli

http://www.ucl.ac.uk/psychlangsci/research/CPB/people/cpb-staff/p_howell

UCL

Assessing speech fluency problems in children aged 4 to 5 years.

Aim:

We aim to screen children between the ages of 4 to 5 for speech disfluency including things like stuttering, cluttering or any other speech or communication problems. Through analyses we hope to predict which factors are likely to lead to long-term difficulties.

Why is this research important?

It is hoped that, as a result of this research, it will be easier to identify those factors that are important in determining those children who will, and will not, recover from early speech dysfunctions.

Fluency problems can have a detrimental impact on the individual's life if not resolved early. Ineffectiveness of communication can impact on personal relationships and social interactions as well as employment opportunities and quality of life.

How will participation in this research help the children?

Pupils we identify as "at risk" will have the benefit of early treatment, as we would recommend to the parents to have their children seen by a GP or a Speech and Language Therapist. Why are the points highlighted above important?

Why are the children being screened so young?

Much recent research has focussed on the importance of early identification of, and intervention on, young children with communication disorders. Intervention is most effective when provided at early childhood, rather than during adolescence. (Bercow Report)

The Bercow report identified shortcomings in the ability of child services and specialist services working together with schools and parents to identify and help children with speech and communication needs.

How will participation in this research help the school?

By allowing this research the school is ensuring the best possible outcome for the children, who are identified with speech disorders. This will be considered 'best practice' under the 'Every Child Matters' agenda.

UCL will produce a small "guide" with tips and signs on how to pick out speech and communication disruptions in children, enabling your school in the identification of children who are at risk of language and communication disorders. This "guide" can also be used as a leaflet to give to parents as a means of outreach from your school.

At the end of the study another leaflet will be given to the school highlighting ways in which the school can work closely with other services so that children identified as "at risk" of communication problems can be referred.

What does participation in this study entail for your school, the parents and the children?

- A meeting will be set up with the head teachers and other nominated staff of the school by Prof. Peter Howell (Primary researcher) prior to the commencement of the research, to outline step-by-step procedures and answer schools' questions.
- All Year One pupils will be interviewed individually in an age-relevant manner by trained and CRB vetted personnel provided by UCL. Sound recordings of their speech (plus a video recording) will be obtained for later analysis.
- During the data collection, each child will be away from the classroom for 10-15 minutes. The school is asked to provide up to two days for the interviews to take place.
- The recordings will be analysed in acoustic labs at UCL by trained UCL researchers. All materials will be anonymised and handled according to both the Data Protection Act and UCL's ethical standards.
- Parents' consent will be needed. The school *may* be asked to send out consent forms and information sheets on behalf of UCL, as the schools would have address databases for the children. Costs will be covered by the research team.
- If "at risk" children are identified then the school will be informed and it will be suggested that the child sees a General Practitioner or a Speech and Language Therapist.
- These children and their parents will be asked if they are willing to participate in a follow-up study, which will take place post-therapy so that we are able to confirm our initial findings.

Timeline

We hope to start the tests as soon as possible. The first phase will be over the period October to December 2013, so there is flexibility in the dates.

Appendix C

Feedback provided to schools after the screening.

«School_Name»
«Address_Line_1»
«Address_Line_2»
«City»
«Post_Code»

14th July 2013

«Title» «First_Name» «Last_Name»

The UCL Speech Research team were at your school to screen all the reception class children for any speech fluency problems. As well as the children who we considered were "at risk", we also found several who we considered borderline at risk. A separate letter describes those children who we consider are at risk. We can explain the details of the scores if required, but at this busy time for you with the end of term looming, we just list the names of the children we want to alert you to.

Please remember that this is based on research findings, but further validation is needed. Also, these are based on a 10-minute speech recording taken from the child during a single visit and these children may have been more fluent on other days. Consequently, your feedback about these is requested here and will be very much appreciated as well as necessary for this project. We also need indications of children who you think may have communication difficulties who are not mentioned in case we have missed any children.

The decision as to whether to refer these children for speech therapy or keep them monitored is at your discretion, but we are prepared to help in any way you require

Details of the children:

Name(s)	Gender	Reception Class ID
«Child_Name»	«gender»	«class»

Finally, please can you fill in the word document attached and email it back to me. We want this to be emailed, so it does not get lost in the post and cause more work. If you prefer to post then please use the following address:

Avin Mirawdeli (PhD Student)

UCL
FREEPOST University College London
Department of Psychology and Language Sciences
26 Bedford Way
WC1H 0AP

Please do not hesitate to contact me if you have any questions. Thank you very much for taking part in this project and we hope to continue working with you next September at the start of the school year.

Yours sincerely,
Avin Mirawdeli

CHILD NAME	**Were you ever concerned about this child? YES/NO**	**Details if you wish:**	**Your action plan?**
«child_name_1b»			
«child_name_2b»			

Appendix D

Leaflet sent to parents alongside the consent form.

What happens if I change my mind after I agreed to take part?

You can withdraw from this study at any point as your participation is voluntary and you are not required to give a reason. If you wish for your data to be destroyed after withdrawal then please let us know.

What will happen at the end of the research project?

Everything collected and the data generated will be anonymised and used as part of a PhD report. The report, parts of the report, or publications which may come out of the research may be used in public presentations, conferences, seminars, lectures, public engagement activities and with policy makers and professionals. No information about the participants will be given out at any point. You will not be identified at any point without your prior consent.

What else should I know?

- More info: You can contact the researcher at any point for more information and clarification. Please email Ms Avin Mirawdeli a.mirawdeli@ucl.ac.uk or call on 02076795401
- Concerns/Complaints: You can contact the Head of the research team Professor Peter Howell by emailing p.howell@ucl.ac.uk or call 02076797566

This research is funded by the UCL IMPACT award and the Dominic Barker Trust.

For more information about the funding body please see the website bellow:

http://www.dominicbarkertrust.org.uk

Further information

This research has ethical approval from UCL No.: 2798/001

Data protection act registration number from UCL: Z6364106/2012/10/31

All researchers involved in this project have recent enhanced CRB checks.

Department of Psychology
University College London
Gower Street
London WC1E 6BT

Phone: +44(0)20 76797566 Prof Howell
http://www.ucl.ac.uk/psychlangsci/research/CPB/people/cpb-staff/p_howell

Phone: +44(0)20 7679 5401 Ms. Mirawdeli
http://www.ucl.ac.uk/speech-research-group/people/phd-students/avin-mirawdeli

Assessing Speech Fluency Problems in Children Aged 4 to 5 Years Old

Parent/Guardian Information Sheet

Researchers at the University College London's Speech Research Team are requesting your assistance on a project that may have various benefits, so please take a moment to read the information in this leaflet carefully and if you have any further queries please do not hesitate to contact the researcher (details given in the "More info" section).

What is the aim of the study?

The aim of the study is to screen children between the ages of 4 to 5 (reception year) for speech fluency problems including things like stuttering, cluttering, dyspraxia or any other speech or communication problem. We aim to screen all children in the reception classes in as many schools as possible.

Why is this research taking place?

We believe that the findings of this research will have an impact on the development of policy and practice within specialist services for children in this age group.

Why have we been chosen?

You have been chosen because your child is 4-5 years old and in the reception year of one of the schools which fitted in our inclusion criteria.

Do I have to take part in this research?

You do NOT have to take part in this research as taking part is voluntary. This research will be conducted in many schools and all children in the reception years will be screened, however, if you do not wish to have your child screened please complete the "opt-out" slip given to you by the school alongside this leaflet.

How will this research be carried out?

Once you agree for your child to take part (you do this by not sending back the "opt-out" letter), we will go into schools and screen every child in the reception year of that school. Your child will be screened by speaking to one of the research team members for an approximate of 5 minutes so that a 2-3 minute speech is recorded.

What you and your child are asked to do?

You will be informed of the day the researchers will come to visit your child's school. You will be asked to bring your child to school as normal and your child will be participating in her/his classes and activities as normal. Each child in the reception year will be taken away from the class for a few minutes to have a chat with one of our researchers. This conversation will be recorded on an audio recorder and the speech sample will be taken back to the acoustic labs within UCL to be analysed. In the lab we will be looking for speech disfluencies. All children show some level of disfluency in speech but we will be looking out for children who display more frequent disfluencies. The parents of those children will be asked at a later stage, to provide us with some more information about the child including things like family history of communication disorders, other medical history and if they had noticed anything about the child's speech. Please note that you do NOT have to give this information.

Are there any risks or disadvantages?

The research is not invasive or intrusive and will hold no such disadvantages. During the interviews the children will be asked questions related to class, learning, pets, toys, daily activities and any other topics you think your child is interested in and would instigate speech in the child. We do look for facial expressions and grimaces, distracting sounds, head movements and movements of the extremities (general fidgeting and body movements). During the interviews, if the child becomes upset, emotional, distressed or wishes to not take part then the interview will be terminated and the data destroyed.

What are the benefits to me and my child?

We hope to identify children who could be "at-risk" of developing speech disfluency. Once these children are identified, the school will be informed so that the child is referred to the appropriate specialist services. We hope that this research contributes to future policy making and practices within relevant fields and you will be contributing to this process of knowledge sharing by taking part.

What will happen to the data collected?

The data will be taken back to the Acoustic labs within UCL where it will be analysed in depth. The speech samples will be listened to by trained, CRB cleared members of staff within the Speech Team. The data will be dealt with strict confidentiality. Data collected will be anonymised so that only members of the research team will know which speech sample belongs to which child, otherwise, and on official records or tables, each child will be given an ID number. Names and other information will not be disclosed. All data are stored in accordance to the Data Protection Act 1998 and we have a registered Data Protection number within UCL. At the end of the study, you will be given the opportunity to consent to given UCL the permission to archive and use the data in future research and if you wish to have this data available online for the use of other educational institutions.

Appendix E

Consent form in the form of an opt-out option for parents.

UCL

INFORMED CONSENT FORM

Please complete this form after you have read the Information Sheet. Please hand back to the school by ……………………………. indicating your consent.

The study has been approved by the UCL Research Ethics Committee (Project ID Number): **2798/001**

If you have any questions arising from the information sheet with the title: ***Assessing speech fluency problems in children aged 4 to 5 years*** please email the researcher on email a.mirawdeli@ucl.ac.uk

By NOT returning the "opt-out" slip below, you are agreeing to take part in this research project and agree for your child to be screened and interviewed by the researchers from the Speech Research Group from UCL's Department of Psychology and Language Sciences.

By NOT returning the "opt-out" reply slip you will also be agreeing with the statements below:

- I have read and understood the information sheet/ leaflet provided which explians in detail what will happen during the process of this research and the process after the completion of the research and I am happy for my child to take part in the study.
- I understand that my child will be audio recorded, and I am happy for you to use the recordings during and after the end of the project.
- I understand that if at any time or point in the future, I no longer wish for my child to take part in this project, I can let the researchers know and they will stop taking data from my child and if I wish will no longer use data already collected.
- I am happy for you to know and detail my child's age, gender and languages spoken for this research.
- I understand that my child's personal information (those mentioned above and data collected: audio recordings) will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- I am happy that this research has been properly explained and I agree to take part in the study.

Return Slip

I (insert your full name in capitals please)……………………………………………………………do NOT wish to take part in the research project titled: ***Assessing speech fluency problems in children aged 4 to 5 years.*** I have read and understood the leaflet provided and the information and statements above and wish to opt-out from taking part in this project. Please do NOT screen my child (insert full name of child in capitals please)…………………………………………………………………………………………………

Signature ________________________ Date ____________________________

Name_________________________

Appendix F

Child's history information sheet.

Child information sheet

ID Number: ....

School: .... Class:..................................

Child Name:.. Age:............................

Gender: (please circle) M F Date of Birth: ....

Place of Birth: ....

Languages spoken:

L1 :.. Approximate % time spoken: ....
L2 :.. Approximate % time spoken: ....
L3 :.. Approximate % time spoken: ....

Family Structure: (please circle) 1) Single parent home 2) 2-parent home

Number of siblings: .... Order of child to siblings:..............................

Has the child ever had or does the child suffer from hearing problems? (please circle)

Yes No Wish not to say Don't know

if **YES**, when was this? ....

Has the child ever had or does the child currently receive Speech Therapy? (please circle)

Yes No Wish not to say Don't know

if **YES**, when was this? ....

Has the child ever suffered from trauma? (i.e. Psychological trauma, physical wound/injury to the head)

Yes No Wish not to say Don't know

if **YES**, when was this? ....
What was the trauma? ....

As far as you know, are there any members of the child's family that have had in the past, or currently have a communication or speech problem? (please circle)

Yes No Wish not to say Don't know

If **YES**, please fill in below:

1) Relation to child: .... Name of disorder if known:..
2) Relation to child: .... Name of disorder if known:..

Folder stored on recorder: .... File Number: ....

Avin Mirawdeli (PhD student), UCL, FREEPOST University College London (PaLS, 26 Bedford Way), Gower Street, WC1E 6BT
a.mirawdeli@ucl.ac.uk

Appendix G

Child's history information sheet cover letter to parents.

UCL

Ref: ***Assessing speech fluency problems in children aged 4 to 5 years***

Dear Parent/Guardian,

As you know, The UCL research team has now started to screen the children in this Primary school. The team needs some more information about each child and I've enclosed an A4 sheet with some questions, which I hope you can answer with as much detail as possible.

May I also remind you that you are under no obligation to answer these questions, and if you wish not to, then please feel free to do so. If there are questions which you are willing to answer, and others which you prefer not to, that is also possible.

Once you have filled in the sheet, please return it to the class teacher as soon as possible. The preferred return date is ………………………………………….. May 2013.

Thank you for your participation and your time.

Yours truly,

Ms Avin Mirawdeli
PhD student/ researcher UCL
Please see below for contact details

Appendix H

DA function centroids

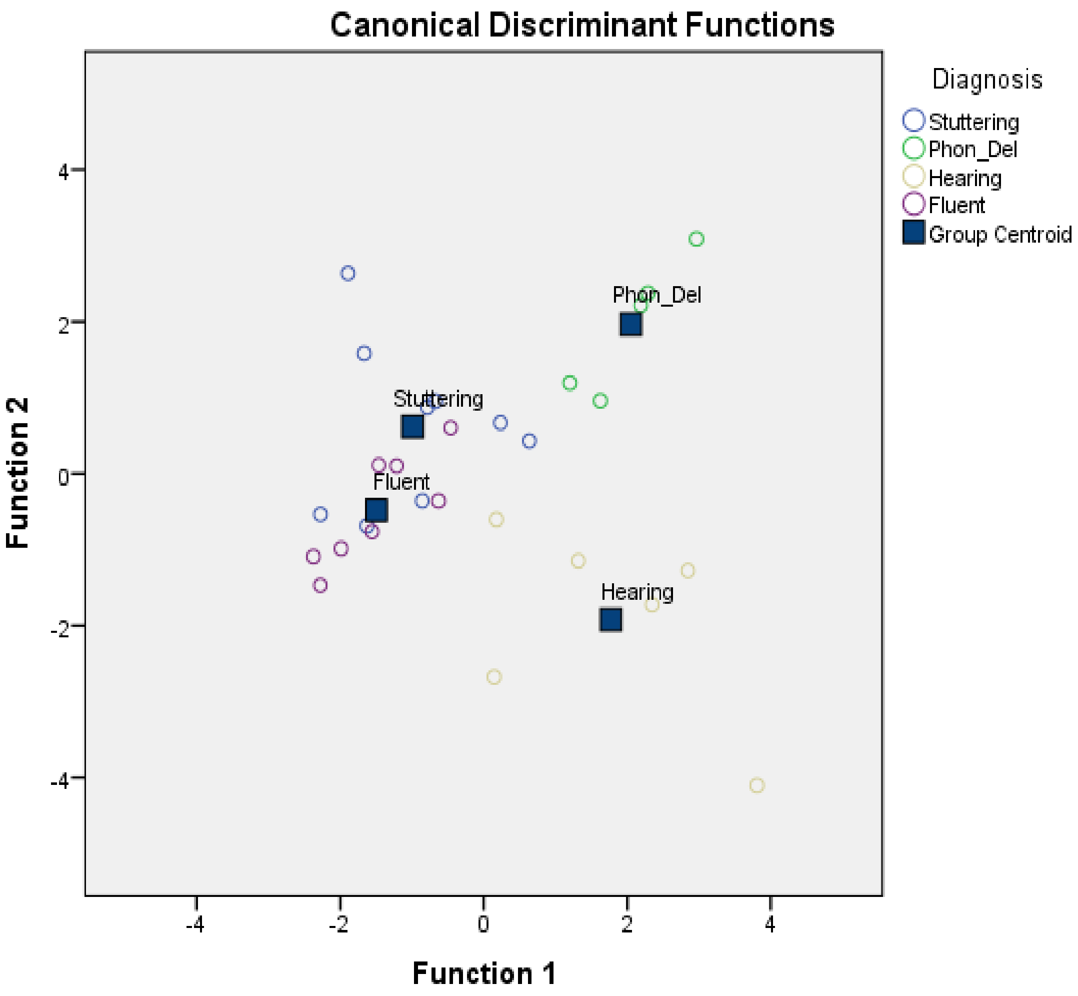

Figure.1. (ABOVE) the figure shows an example of data clustering as a means of demonstrating centroids of each cluster group. This is for illustrative purposes only. There are two ellipses:

1) Blue: this encapsulates cases, which are deemed stuttering and the blue square within the blue ellipse is the centroid of the stuttering cluster.
2) Orange: encapsulates the hearing loss cases with the blue square representing the centroid of that group.

In Ward's Hierarchal Cluster Analysis, cases are clustered together based on how close they are to the centroid of that cluster.

Figure.2. (BELOW) Shows an example of a datasheet from SPSS, where the D scores have been saved under Dis1_1 and their associated reclassification in Dis_1, where 1 is fluent and 2 is dysfluent. Column “Fluent_VS_Dysfluent” is the original observed classification. As can be seen from this datasheet, some of the cases originally deemed dysfluent have been reclassified as being fluent (i.e. some of the stuttering children).

Diagnosis	Fleunt_VS_Disfluent	Dis_1	Dis1_1	child_ID
Stuttering	2	2	-.51628	1
Stuttering	2	1	-1.24934	2
Stuttering	2	2	-.26964	3
Stuttering	2	2	.40177	4
Stuttering	2	2	.47028	5
Stuttering	2	1	-1.55079	6
Stuttering	2	2	.81283	7
Stuttering	2	2	.42917	8
Stuttering	2	2	2.84760	9
Phon_Del	2	2	1.59385	10
Phon_Del	2	2	1.47738	11
Phon_Del	2	2	2.01177	12
Phon_Del	2	2	.77857	13
Fluent	2	2	.86764	14
Hearing	2	2	1.25815	15
Hearing	2	2	-.59849	16
Hearing	2	2	.60045	17
Hearing	2	2	1.15538	18
Hearing	2	2	-.08466	19
Hearing	2	2	1.81309	20
Fluent	1	2	-.61904	21
Fluent	1	1	-1.19453	22
Fluent	1	1	-2.27015	23
Fluent	1	1	-2.05092	24
Fluent	1	1	-2.27015	25
Fluent	1	1	-1.29045	26
Fluent	1	1	-1.66726	27
Fluent	1	2	-.88623	28

Discriminant Function Analysis (DA) procedure:
There are two steps in the DA:

1) Significance of the set discriminant functions: Here F tests are used to determine if there are significant differences in observed group means. This is done by comparing the matrices of variance-covariance with the pooled within-groups variance-covariance matrices.

2) Once significance is observed, DA will automatically begin the classification of variables, this is done by finding the optimal combination of variables. DA will return with the Functions, in which the first function is the optimal discriminant, followed by the second, then third and so on. Canonical correlations (correlation which tests the cross-covariance matrices cov(X, Y)) between the functions is performed, which allows for classifications to be determined, the larger the standardized coefficients the greater the contribution is from that variable in its discriminatory power.

Example.1.

The table below shows the standardized canonical Discriminant function coefficients for three speech difficulty groups and their power in discriminating between the DV.

	Function 1	Function 2	Function 3
Stuttering	-1.434	1.303	.688
Phon.Delay	1.138	.998	.052
HL	.599	-2.160	.301

Function 1 has the highest values for the three IV groups. Stuttering is -1.434 in functions 1 compared to 1.303 in function 2 and .688 in function 3. This shows that stuttering had the highest discrimination power in the DV followed by Phonological delay and then HL.

How is the discriminant function coefficients reached?

Scores from each case in the IV are multiplied by the coefficient of the independent variable. A composite score for that case is then created by adding the sum of these multiplications and adding it to the constant. In DA this is called the discriminant score. These scores are used for predicted group membership.

Visually, in Discriminant Function Analysis (DA), the D Scores are represented by the distance from the centroid. The function created maximises the difference between groups and returns a score for each case depending on its distance from the mean (centroid) of that group.

Appendix I

This Table shows the SLCN groups and the individual symptoms associated with it. In the final column the codes used in the transcriptions are shown (as seen in figure 1).

SLCN Group	Individual symptoms within the group	Transcription code used
Stuttering	Prolongations	Pro
	Blocks	Blo
	Hesitation	Hes
	Revisions	Rev
	Broken words	BW
	Part-word repetitions (PWR)	PWR
	Whole word repetitions (WWR)	WWR
	Phrase repetitions	PhR
	Interjections	( )
Phonological Disorder	Backing	B
	Initial consonant deletion	ICD
	Middle consonant deletion	MCD
	Sound preference substitution	SPS
	Intrusive consonants	IC
Phonological Delay	Gliding (r-w)	G
	Deaffrication	D
	Poor phonological awareness	PHA
	Final consonant deletion	FCD
	Weak syllable deletion	WSD
Hearing Loss (HL)	Voicing errors	VE
	Reduction/addition of syllables	RS/AS
	Raised pitch on closed vowels	RPCV
	Omission of copular auxiliary verbs	OCA
	Omission or confusion on determiners	OCD
	Omission of tense endings/ ‘s’ endings	OTE
	Pitching issues	PI
	Phonological discrimination failings	PDF
	Poor digit recall	PDR
	Hypo/hyper-nasality	HyperN/ HypoN