Profiling developmental speech sound disorders: differences, deficits and outcomes

Profiling developmental speech sound disorders: differences, deficits and outcomes.

Aileen A. Wright

Thesis submitted for the degree of Doctor of Philosophy in Speech and Language Therapy

Supervisor: Professor Sue Franklin

Submitted to the University of Limerick, August 2017

Declaration

Title: **Profiling developmental speech sound disorders: differences, deficits and outcomes.**

I hearby declare that this is entirely my own work and has not been submitted to any other University or Higher Education Institution, or for any other academic award in this University.

Where use has been made of the work of other people, it has been fully acknowledged and referenced.

Signed: _____

Aileen Wright

Abstract

Profiling developmental speech sound disorders: differences, deficits and outcomes.

Aileen Wright

There is evidence that many pre-schoolers with speech sound disorders (SSD) will resolve their SSD without therapy. The identification of children who are likely to resolve without help would avoid resources being expended on children who do not need therapy. Conversely, identification of sub-groups with specific underlying deficits that lead to persisting SSD would allow intervention to target the source of the difficulty.

Research questions:

Are profiles of SSD associated with other cognitive difficulties?

Do sub-types of children classified by symptomatology remain stable?

Which children's SSD do not resolve?

Eighty-five children referred to local SLT services with suspected SSD were assessed on a battery of tests for speech, language and other cognitive abilities. Seventy-five children were reassessed two years later to establish changes in diagnostic profile and outcomes.

At time 1 exploratory factor analysis identified clearly defined factors (of which one was speech accuracy), suggesting speech accuracy in pre-schoolers is largely independent of language, motor or cognitive-linguistic abilities. Sub-groups of children with SSD defined by symptomatology or proposed aetiology were not differentiated by profiles of underlying deficits. Symptomatology changed as children aged, suggesting a continuum of delay mediated by age and severity rather than qualitatively different disorders.

Longitudinal follow up showed that 79% of children normalised their speech after 2 years, with or without therapy. Logistic regression indicated that the most clinically useful predictor of recovery was the child's score on a test of Non-Word Discrimination. Lateralised fricatives did not resolve.

Clinical implications are discussed.

Table of Contents

Table of tables

Table of Figures

Abbreviations

General Abbreviations		Abbreviations of test names	
AD	Articulation Disorder	BEERY VMI	Beery-Buktenica Developmental Test of Visual-Motor Integration, Sixth Edition, (Beery and Beery 2010)
CAS	Childhood Apraxia of Speech		
CLS	Core Language Score		
CPD	Consistent Phonological Disorder		
DLD	Developmental Language Disorder		
DDK	Diadochokinetic Rate	BPVS	British Picture Vocabulary Scale: Third Edition. (Dunn *et al*. 2009)
EHC	Education, Health and Care Plans		
EPG	Electropalatography		
F	Female	CELF-4 (UK)	Clinical Evaluation of Language Fundamentals - Fourth Edition UK (Semel *et al*. 2004
HSE	Health Service Executive, Ireland		
IAx	Initial Assessment		
IPD	Inconsistent Phonological Disorder		
IQ	Intelligence Quotient	CELF-P	Clinical Evaluation of Language Fundamentals-Preschool-2 (Wiig *et al*. 1992)
LI	Language Impairment		
M	Male		
MSD-NOS	Motor Speech Disorder-Not Otherwise Specified	CELF-P2 (UK)	Clinical Evaluation of Language Fundamentals-Preschool 2 (UK) (Semel *et al*. 2006)
NWD	Non-Word Discrimination		
OME	Otitis Media with Effusion		
OR	Odds Ratio		
PA	Phonological Awareness	DEAP	Diagnostic Evaluation of Articulation and Phonology, (Dodd *et al*, 2002)
PCC	Percent Consonants Correct		
PD	Phonological Delay		
PSD	Persistent Speech Disorder	FIST	Flexible Item Selection Test (Jacques and Zelazo 2001)
PSSD	Persisting Speech Sound Disorder		
RCSLT	Royal College of Speech and Language Therapists, UK	PAT	The Phonological Awareness Test (Robertson & Salter, 1997).
RD	Reading Difficulty		
S&L	Speech and Language	PIPA	The Preschool and Primary Inventory of Phonological Awareness (Dodd *et al.* 2000)
SD-DPI	Speech Delay-Developmental Psychosocial Involvement		
SD-GEN	Speech Delay-Genetic		
SD-OME	Speech-Delay-Otitis Media with Effusion	TERA	Test of Early Reading Ability-2 (Reid *et al.* 1989)
SE	Speech Errors		
SI	Speech Impairment	TEWL	Test of Early Written Language-2 (Hresko *et al.* 1996)
SLCN	Speech, Language and Communication Needs		
SLT	Speech and Language Therapist or Speech and Language Therapy	TOLD-P2	Test of Language Development: Primary, 2nd Edition (Newcomer & Hammill 1988)
SSD	Speech Sound Disorder		
TD	Typically Developing		
UK	United Kingdom		
WNL	Within Normal Limits		

Acknowledgements

My sincere thanks go to Professor Sue Franklin, who has supervised this work with wonderful clarity, generosity, responsiveness and patience.

Thanks also go to all my colleagues in the Department of Clinical Therapies for their marvellous support, encouragement, advice and constant, unflagging optimism. Thanks also to Dr. David Howard for emergency statistical help and to Michelle O'Donoghue for help with transcription reliability, and to Professor Barbara Dodd, whose work first inspired me to work with speech sound disorders, for her help with the study design and comments on early drafts.

Many thanks to my friends Steve Wilson and Karen Gomoluch, who looked after all my corporeal needs for months while I explored factor analysis on their sofa, and to my Newport friends, for their encouragement and meals.

Thanks also to all the Speech and Language Therapists in Clare, Limerick and Tipperary who supported this work by referring children to the study, and providing clinic space where needed.

Sincere thanks are offered to the children and their families who participated in the study.

My parents, who were determined that I would go to university, and who gave me the best start in life to ensure that I could, also deserve many thanks, as does my husband, Joe, who took my lovely daughters out every Sunday so I could get my Master's. Also my love and gratitude to my daughters, Ciara and Rebecca, who kept me going with laughs.

Lastly, an acknowledgement to Joe Wallis, who told me many years ago, before I even became an SLT, that one day I would have a PhD.

In memory of my mother,

June Wright (née Wragg),

1928-1995

who would have loved to have seen this.

1 Chapter 1 Introduction

1.1 Background

In the area of speech and language therapy (SLT), the largest group of clients presenting to Speech and Language Therapy services is that of children with functional speech sound disorders (SSD) (Broomfield and Dodd 2004b, McLeod *et al.* 2004), also called SSD of unknown origin: that is, children whose speech impairment is not due to a structural or organic cause, nor secondary to some other condition. A systematic review of studies by Law *et al.* (2000) found the estimated prevalence of functional speech impairment (termed primary speech delay in that study), ranged from 2.3% to 24.6% in children under 16 years of age and additional studies since that time support prevalence rates like these for speech impairment in pre-schoolers (Eadie *et al.* 2015) and school-aged children (McLeod and McKinnon 2007, Jessup *et al.* 2008, Wren *et al.* 2016), with the wide range due to such factors as criteria for diagnosis, age at identification, sample types and sizes. Children with functional SSD are typically diagnosed between 2 and 4 years of age (Broomfield and Dodd 2005b, Claessen *et al.* 2016), and present with speech that is unusual or difficult to understand, without apparent sensory, structural, neurological or psychological impairment.

Most children follow a very similar course of speech development, using similar patterns of error that gradually resolve, until by the age of between 5 and 6 they are intelligible to most people and produce over 90% of speech sounds accurately (Dodd *et al.* 2003, Wren 2015). Speech sound disorders (SSD) are diagnosed when children's speech accuracy develops more slowly than their peers, so that they present with speech similar to that of a much younger child, or in some cases with patterns of errors or consistency in production that is quite different from that seen in the normal course of development. Speech sound disorders range from the very mild, which rarely affect intelligibility, e.g. minor distortions of single sounds, usually /s/ or /ɹ/ in English, to very severe, where children may use only a very limited range of sounds and/or structures in their speech, or where they are very inconsistent so that a single word may be pronounced in very many different ways. In these latter conditions, intelligibility may be severely reduced.

Research shows that those whose speech deficits persist are at risk for a range of negative consequences: they are rated more negatively by peers on likeability, intelligence and personality (Freeby and Madison 1989, Crowe Hall 1991, Hitchcock *et*

al. 2015), are at risk of failing to reach their academic and social potential (Overby *et al.* 2007), and suffer negative consequences for work and social life in adulthood (Silverman and Paulus 1989, Felsenfeld *et al.* 1992, Veríssimo *et al.* 2011). Rooke *et al.* (2013) found that 2.2% of children in Ireland aged 9 were reported by teachers as having speech impairment severe enough to limit their activity, according to a survey carried out as part of the Growing Up in Ireland longitudinal cohort study of 8,500 children (Murray *et al.* 2011).

Children with a history of phonologically based SSD have an increased risk of problems with phonological awareness and acquisition of literacy, even if the SSD resolves by school age (Preston *et al.* 2013), and this risk is even greater where the SSD persists into school-age (Bird *et al.* 1995, Nathan *et al.* 2004b, Raitano *et al.* 2004). The 'modified critical age hypothesis' (Nathan et al. 2004a) suggests that the risk for reading difficulty is much reduced if the child has achieved age appropriate speech, and phoneme awareness skills are emerging, by the time children begin formal reading instruction. In their longitudinal study of 3 groups of children (n=57) assessed at 4 and 7 years, the group with persistent SSD at 7 performed more poorly on reading measures than those whose SSD had resolved, or who had no history of SSD. What is not clear from this study is whether children whose SSD resolved due to substantial therapeutic input had equally good outcomes as children whose SSD resolved naturally, i.e. without therapy. It may be that the latter group are "slow starters" with only transient SSD, showing no particular underlying phonological deficit, so that we would not expect literacy to be affected. In contrast, where children need substantial therapy to resolve their speech problems it may indicate an underlying deficit in phonological processing which may still go on to effect literacy development. Nevertheless, ensuring that children begin school with age appropriate speech and phonological processing skills is important to support both social and academic success later in life.

Meanwhile, in Britain and Ireland at least, speech therapy is difficult to access. SLT services are under-resourced and waiting lists are long: according to the Irish Health Service Performance Report in December 2016 (HSE 2016a), there were 13, 801 people awaiting assessment by public, i.e. Health Service Executive (HSE), SLT services for up to one year, with 456 waiting more than one year, plus 7,411 on the waiting list for therapy up to one year, and 621 waiting more than a year. In addition, the SLT service remains understaffed, with only 993 Whole Time Equivalent Speech

And Language Therapists (SLTs) working in the public health system in Ireland in June 2016 (HSE 2016b), which is 292 fewer than the number recommended in the Bacon report published 16 years ago (Bacon and Associates 2001). In the UK, the Royal College of Speech and Language Therapists (RCSLT) recently revealed that Speech and Language Therapy services in England face severe capacity challenges, with only 40% of SLTs reporting that they had capacity to deliver services to children without EHC (education, health and care) plans, despite the fact that only a minority of children with speech, language and communication needs in England have such a plan (RCSLT 2017). Given this context of limited resources, long waiting times, but a limited timeframe in which to resolve deficits (especially given the very young school starting age in Britain and Ireland), it is vital that services are prioritised for those who are likely to have persisting deficits and thus suffer long term consequences.

Not all children require intervention. There is evidence that a number of children diagnosed with SSD in preschool resolve their difficulties in a reasonable timeframe without treatment, e.g. Roulstone *et al.* (2003) and see Section 1.5.3 later in this chapter. Shriberg *et al.* (2005) assert that 75% of children normalize their SSD by 6 years of age, with most of the remaining children treated for SSD normalizing by 9 years of age (Shriberg *et al.* 1994a, Shriberg *et al.* 1994b). Children whose SSD persists past the age of 9 typically show limited spontaneous improvement and poor response to intervention. However, evidence that younger children can "grow out of" their difficulties highlights the fluid profile of early speech and language development. Given the evidence regarding the negative outcomes for children with unresolved SSD, therapists are under pressure to provide therapy as early as possible, so that there has been a shift to providing intervention for very young children, often at the expense of older children (Dockrell *et al.* 2006), although this has not been based on research evidence (Glogowska *et al.* 2000). However, treatment given to children who are likely to resolve without it represents an opportunity cost, i.e. for the service, a lost opportunity to give therapy to someone else who needs it, for clients and carers, a lost opportunity to spend their time in leisure, school or economic activity, and for those who need services but do not fit criteria for prioritisation, a lost opportunity for normal speech. Knowing which children have only transient difficulties would allow therapists to target their resources most effectively.

For those children whose SSD persist, Shriberg (2003) argues that "accurate differential diagnosis of a patient's disorder, including information on both original and maintaining causes, is necessary to determine the optimum form and content of treatment" (p.502). In 2013, in an Update to the 2004 WHO report on Priority Medicines for Europe and the World, the World Health Organisation called for more research into stratified medicine (Kaplan *et al.* 2013). Historically, disease has been treated on a 'one-size-fits-all' basis. Guided by evidence-based information and professional guidelines, a treatment is described for a condition, and expected to suit all clients with the said condition. However, heterogeneity of aetiology, course and responsiveness to treatment of disease is a well-recognised phenomenon. Certain clients fail to respond to the best-recognised treatment for their condition and clinicians then frequently have to resort to a 'trial-and-error' approach, adapting the dosage or moving down a list of possible alternatives. The aim of stratified medicine, in contrast, is to ensure that each client gets the right treatment right from the onset. To do this, it aims to identify sub-groups of clients within a certain presentation of disease or diagnosis, who show distinct mechanisms of disease, or particular responses to treatments. By identifying and gaining a deeper understanding of such subgroups, specific treatments can be better targeted, and novel targets and therapeutic strategies identified, in order to improve outcomes for the client and cost-effectiveness for the service. In speech therapy, the practice of stratified medicine requires a reliable method to separate out homogeneous sub-groups of children with SSD. Identification of distinct sub-groups would have clinical benefits if it allows researchers to tailor specific intervention approaches to their deficits (Rvachew 2007).

The following review will briefly outline a theoretical approach to speech development, and examine the literature on potential original and maintaining causes for functional SSD. It will then consider current theories of sub-typing children with functional SSD, and finally investigate factors shown to be predictive of outcome in longitudinal studies, which might provide validation for any putative sub-types and their usefulness in the stratified management of children with SSD.

1.2 Acquisition of speech

1.2.1 Factors influencing speech development

In their first five or six years, children learn to pronounce most of the sounds of their native language in a way close enough to adult versions to be perceived as accurate and

intelligible. While early word forms in very young children can be very variable, as children get to age 3, the process of phonological development in most children is so similar that we can recognise “normal” patterns of acquisition (Williams and Elbert 2003, Claessen *et al.* 2016). Children begin to acquire speech and language while the speech organs and necessary ability to control them for speech production are still developing. The shape and size of the vocal tract and articulators changes substantially during the first two years of life, and continue to develop up until adolescence. Speech motor control also develops rapidly throughout infancy, and continues to develop through childhood (Walsh and Smith 2002, Nip *et al.* 2009). Thus, differences between adult and child speech have been attributed to limitations arising from such motor and anatomical differences, and development in speech attributed to development in these areas.

Across many languages it has been observed that most children produce most of the vowels of the ambient language by age 3 (Donegan 2013, Selby et al. 2000, Pollock and Berni 2003) with stops and glides being introduced before fricatives, affricates and liquids (Li et al. 2009, Bernhardt and Stemberger 2007). Jakobson (1968) suggested that these similarities in the course of acquisition were due to universal principles that structure the phoneme inventories of all languages, and also determine the order of acquisition of sounds, because they reflect universal physical and cognitive constraints on speech production. For example, stops are considered easier to produce as they require only a more-or-less accurate ballistic gesture, while fricatives require protagonist and antagonist muscles to work together to produce a precisely located narrow constriction (Kent 1992, Altvater-Mackensen and Fikkert 2010). However, when viewed cross-linguistically, differences between order of acquisition of phonemes and syllable structures in different languages become evident. For example, /ʧ/ is a late acquired consonant in English, considered difficult to produce due to the need to combine two gestures, i.e. a stop followed by a fricative release. In Spanish, however, /ʧ/ is acquired earlier than in English. In contrast, codas, (less frequent in Spanish than English), are acquired by English speaking children much earlier than by Spanish speaking ones (Macken 1995). Zamuner (2004) shows evidence that, at the phoneme level, children who are learning different languages acquire phonemes at different rates depending on the frequency of the sounds in the ambient languages rather than as an effect of universal sequence or articulatory ease. This influence is evident even before the child acquires meaningful speech. Although early babbling is similar across

languages, probably due to the infants' similar anatomy and physiology, a longitudinal study of 20 children acquiring English, French, Japanese, or Swedish followed between the ages of 9 and 17 months, showed that, with age, the sounds in babbling are increasingly influenced by the ambient language (de Boysson-Bardies and Vihman 1991) . Similarly, a study of 23 infants exposed to Canadian French and 20 infants exposed to Canadian English, aged between 10 and 18 months, showed that the older children, but not the younger ones, had differences in the vowel space which were associated with the infant's ambient language environment (Rvachew *et al.* 2006), again attesting to the influence of the input even before words develop. Such findings may thus point more logically to an input-led constructivist-emergentist approach to phonological acquisition, where frequency and functional load in the input influence phonological acquisition, rather than Jacobson's determinist, nativist approach. However, other factors must also be in play. For instance, in English, /s/ is both a frequent sound and carries a high functional load, marking plurals, verbs and possessives. However, one study showed that /s/ is not accurately acquired by 90% of children until after age 9 (Smit *et al.* 1990). This suggests that neither ease of articulation nor frequency of input alone direct the course of phonological development.

If input from the ambient language directs development, it is logical that perception of the input must also influence development. It is widely accepted that perception precedes production in phonological acquisition (Altvater-Mackensen and Fikkert 2010), so that any limitations in perception would influence the course of acquisition and production. However, speech development is not limited by perception alone, as a body of research in infant speech perception, and the fact that children often reject their own inaccurate productions of words as incorrect (Smith 1973), has shown that children may have accurate perception of sounds even when they cannot produce them accurately. A large number of experiments testing infants' speech perception abilities have shown that, in the first six months of life, they are able to discriminate a wide range of phonetic contrasts in both the ambient language and in languages they are not exposed to (Eimas *et al.* 1971), see also Rvachew and Brosseau-Lapré (2012), Chapter 2, for a review. During the second six months of life, evidence demonstrates that infants gradually lose the ability to distinguish, or at least stop attending to, those contrasts not used in the ambient language (Werker and Tees 1999). This occurs at the period where they first start attending to words.

Such evidence regarding their perceptual abilities led to the view that infants are able to perceive and store all necessary phonetic detail, so difficulties in production could be not the result of deficits in perception. However, subsequent research showed that not all contrasts are easily perceived. For example, separate studies of French and Spanish infants showed that at 4 months both groups discriminated only the long-lag voicing contrast as used in English, discriminating the short lag contrast used in Spanish and French only when they reached 6 to 8 months of age (Lasky *et al.* 1975, Eilers *et al.* 1979, Hoonhorst *et al.* 2009). Nittrouer and colleagues have demonstrated that, at least for some phonetic distinctions, children weight the various acoustic properties upon which phonetic decisions are made differently than adults do, resulting in potential mismatches in phonetic categorisation. Specifically for fricative vowel combinations, they have shown that when young children make judgements about whether an initial fricative is /s/ or /ʃ/, they weight formant transitions more, and fricative-noise spectra less than adults do (Nittrouer 1992, Nittrouer and Miller 1997). It is only as they get older that this developmental strategy shifts to fit the adult pattern, so that noise is weighted more than formant transitions. Thus, they would argue that children are not born with in-built phonetic categories, but must organise the information in the speech input to derive the phonetic structure, thereby implicating the importance of a statistical learning mechanism. They assert that phonetic segments do not exist as neatly packaged and sequenced units, with distinctive boundaries in the speech signal indicating where one segment ends and the next begins, but must derive categories from the input via statistical learning (Nittrouer 2002a). If, as shown, they initially use cues different to those used by adults, it is likely that their representations of phonetic categories develop gradually, and are not wholly identical to adult representations in the early stages, meaning that perceptual as well as motor limitations may play a part in speech development.

However, the work described above has been concerned with the acquisition and perception of individual segments rather than the perception and production of words. From the naturalistic, deterministic perspective, the process of phonological acquisition is described as adding sounds in a regular and systematic way to fill out the structure of a formal phonological system (Werker and Tees 1999). In contrast to this, usage-based, constructivist-emergentist approaches see phonological acquisition set in a social context involving the need to communicate. In contrast to determinist approaches which specify innate categories, constructivist-emergentist approaches propose that children

do not have any innate knowledge of linguistic grammar, but must abstract knowledge of the ambient phonological system from the input language (Ambridge and Lieven 2011). The grammar then emerges from the generalisations that children form from their experience of the input language. In this approach, children are viewed as active learners who must form their own hypotheses about the language system from the input, rather than be provided with initial hypotheses from an innate grammar that they just need to set up to fit the input. Phonetic categories are not in-built but derived from statistical probabilities in the speech signal. Thus they argue that phonological acquisition begins with words, rather than minimal units of phonetics or phonology such as features and segments (Velleman and Vihman 2002). In order for children to begin to recognise words at about 6 months old, they must have laid down some lexical representation of the word in long-term memory, which must include a phonological representation that is adequate to distinguish it for recognition purposes (Stackhouse and Wells 1997). As mentioned above, evidence about infants' perceptual abilities led to the assumption that phonological representations at the word level were fully specified and adult-like. However, research has shown that infants, as well as not using the same cues as adults for phonetic categorisation, may also omit to include all the phonetic detail available to them when forming phonological representations at the word level. For example, Stager and Werker (1997) contrasted the ability to discriminate phonetically similar "bih" and "dih" syllables in sixteen 8-month-old and sixteen 14-month-old children in two different tasks: a sound discrimination task and a word discrimination task. They found that, when they taught the 14-month-old children to associate the syllable "bih" with a novel object, i.e. to lay down a lexical and phonological representation for it as a word, they did not appear to notice a switch when it was labelled "dih" rather than "bih", while the 8-month-olds did notice such a switch. However, when represented with phonetically dissimilar words ("lif" and "neem"), the 14-month-olds identified the switch, suggesting failure in the previous task was due to the fine phonetic difference. Nevertheless, when changed to a simple sound discrimination task, the 14-month-olds identified the switch from "bih" to "dih", showing they were still able to perceive the fine phonetic detail, but were not using it at the word discrimination level, suggesting it was not included in the phonological representation of the word. Stager and Werker proposed that this might be due to a functional reorganisation taking place when children start the demanding task of learning words, whereby inattention to phonetic detail may be beneficial, so that they need to lay down only that information necessary to discriminate new words from the

small number already in the lexicon. As word learning becomes more practiced, they suggest, and presumably, as more detail is needed to differentiate words as the lexicon expands, they will begin to access more of the perceptual information available to them when forming the phonological representation. Thus, it would be argued that phonological representations are not adult-like at early stages, but become more detailed with development of cognitive/linguistic skills and pressure from the input. (Altvater-Mackensen and Fikkert 2010), showed a similar finding in a study of Dutch 14-month-old infants, but, coming from a more nativist perspective, explained it in terms of markedness. Although the children in their study were able to recognize a change from stop to fricative in a pure discrimination task, when tested on newly learned words, they were sensitive to the stop-fricative contrast in word-initial, but not in word-final position. Even in word-initial position, while they detected changes from fricative to stop, they did not notice changes in the other direction, from stop to fricative. They also concluded that children's early phonological representations are not adult-like, but proposed that particular features were left unspecified, namely that marked but not unmarked features are specified, and specification proceeds gradually, not to all positions at once. This explanation would account for the fact that fricatives (marked as continuants) are often substituted with stops (unmarked defaults), but not vice versa, and also explains how contrasts may be achieved earlier in some word positions than others. Again, with maturation, representations gradually become more detailed until they match the adult representation.

If we then consider that the phonological representation of a word stores the information that is used to construct a motor programme for output of the word (Stackhouse and Wells 1997), inadequately specified phonological representations will necessarily lead to inaccurate output forms. However, adult-like phonological representations will not guarantee accurate output forms as limitations in phonological planning, motor planning, programming and execution abilities may still constrain output forms.

Supporting the view that phonological acquisition begins with words rather than sounds, is the evidence from acquisition studies that first word forms produced by children show little evidence of the systematic substitutions or omissions suggested by generative and natural phonology (Waterson 1976). The following section outlines the hypotheses regarding the possible course of acquisition in the constructivist-emergentist view.

1.2.2 The whole word phase

Studdert-Kennedy (1987) gives three pieces of evidence from early word production to support the view that the earliest phonological structure is the whole word: firstly that phonetic forms may appear in one word but not another, e.g. *no* produced with initial [n], but *night* produced with [m] substituting for /n/, whereas /m/ is substituted with [b] in *moo*; secondly the inconsistency in word production on different occasions seen in very young children (e.g. Ferguson and Farwell (1975), which he suggests happens because the child analyses the adult model into a collection of gestures rather than a sequence of integral segments, but has not mastered the ability to apply correct timing/sequencing of the gestures; and lastly the difficulty in switching place or manner in articulation within one syllable which is seen in the first stages of word production, and which leads to consonant deletion or consonant harmony, where features apply to the word as whole, rather than to individual segments.

Velleman and Vihman (2002) suggest that words develop from babble. It has been shown that the sounds the child produces in late babbling are the sounds included in first words, suggesting that babbling and speech are a continuous and overlapping process, where babble first serves to discover and practice auditory and articulatory mappings for speech. Velleman and Vihman suggest that, during babble, the child develops vocal motor schemes (VMS) i.e. phonetic patterns that she/he can produce at will. Such frequently occurring babble patterns then become a type of "articulatory filter" so that when similar patterns occur in the input they are particularly salient to the child, so that he/she attends to words containing such patterns earlier. The child can then combine frequently used VMS that are salient in the ambient language to form templates for words that he/she can easily produce, then preferring for her/his first words those that fit the template relatively well. Thus in early development, both lexical selection, where children attempt to produce only those words which fit into their current production inventory, and lexical avoidance, where words that contain sounds or structures not in the current inventory are not attempted, are seen (Ferguson and Farwell 1975, Velleman and Vihman 2002, Stoel-Gammon 2011).

As the child needs to acquire more words in order to increase communicative ability, he/she must adapt some words to fit into the template, leading to the types of production that appear to be far from the target, which are seen at the later single word level (Waterson 1976), and which cannot be accounted for by segmental or structural

correspondence rules. The view that the child acquires words in a holistic manner rather than as a sequence of segments thus accounts for the lack of systematicity in the first stages of word production. Stackhouse and Wells (1997) describe phonological representations of words at this stage as consisting of the most acoustically salient features of the adult form without a segmented internal structure, with the motor programme for production also consisting of a single undifferentiated gestalt of gestures.

1.2.3 Transfer from holistic to segmental representations

It is generally accepted that it is increase in the size of the lexicon that leads to further development of the phonological representation (Rose 2014). Walley *et al.'s* (2003) Lexical Restructuring Model suggests that larger vocabularies force changes to take place in the storage of phonological information in long-term memory, whereby phonological representations become more refined and are gradually segmented into smaller sub-lexical units such as syllables, onset/rime components and eventually phonemes.

How segmentation occurs is viewed in different ways. If phonological representations of words contain only enough specification to allow for the discrimination of one word from another, when a child has few words, little specification is needed to discriminate them. Werker and Tees (1999) suggest that an increase in vocabulary could result in the need to add more phonetic detail to the phonological representations of similar-sounding words so that they can be better differentiated, i.e. a refining of perceptual characteristics. In contrast, Nittrouer *et al.* (1989) see the emergence of phonetic segments as the result of gradual reorganization of the gestures forming early words or syllables, i.e. a motoric process. While children only have a few words, they need only a few articulatory routines to produce them. As vocabulary increases, new routines must emerge to enable production of adult models not previously attempted, so that development occurs as a process of elaborating the articulatory routines, but also gradually narrowing the domain within the word to which the articulatory routine applies, until there is a single articulatory routine for each phonetic segment. They supported this hypothesis by showing that as children increased in age from three upwards, they made greater differentiation between /ʃ/ and /s/ in production, while the extent of co-articulation with the following vowel decreased, i.e. the domain of the gesture became smaller.

Usage-based phonology offers a third, cognitive perspective, where segments and features emerge from generalisations over exemplars stored in memory at a more holistic level; i.e. as vocabulary increases, segments can be abstracted from stored representations of words when children notice areas of acoustic and articulatory overlap amongst those representations (Byun and Tessier 2016). Byun and Tessier support this view using evidence that children produce words from dense neighbourhoods more accurately and consistently than words from sparse neighbourhoods, suggesting they have more detailed segmental representations of words from dense neighbourhoods because areas of overlap are more evident where exemplars are more densely stored. Further support is given by Munson *et al.* (2005a), who report evidence that young children repeat non-words containing frequent, high probability sequences more accurately than those with low probability sequences. However, as children's vocabularies increase, this effect becomes smaller, which, they suggest, shows that children with larger vocabularies have more developed, more detailed, segmental representations, which would arise as areas of overlap become more frequent as more words are stored. Thus repetition becomes more accurate as individual segments within sequences become more salient.

In fact, Byun and Tessier (2016) manage to incorporate both the motoric and usage based perspectives in development of segmental representations by positing that representations of words that have been both heard and produced are more robustly encoded than those that have only been heard, so that both production ability and the number of stored exemplars influence the development of the phonological grammar.

The ability to represent phonological representations at the segmental level is closely linked to the acquisition of literacy. Segmental phonological representations are needed to allow the development of phoneme awareness, which underpins the acquisition of decoding skills (Carroll and Snowling 2004). Difficulty in developing such segmental phonological representations may thus be common to some children with speech difficulties and those with literacy difficulties, explaining the increased risk for literacy difficulties in children with SSD (Anthony *et al.* 2011).

Once phonological representations have become segmented, the word templates of the whole word phase can become expanded and loosened, and the relationship between child and adult forms becomes more systematic, with more regular patterns of

correspondence, and less variability in production. Stackhouse and Wells (1997) term this phase the systematic simplification phase.

1.2.4 Systematic Simplification Stage

During this phase, particular patterns of segment deletion and substitution are observed in children's speech frequently enough to be identified as normal patterns of development at specific ages. However, the transition between the phases emerges gradually, with patterns affecting word structure, such as consonant harmony and final consonant deletion gradually giving way to patterns affecting individual segments, e.g. stopping of fricatives, velar and post-alveolar fronting, deaffrication and gliding (Dodd *et al.* 2003), although cluster reduction continues to affect word shape for longer.

Such error patterns are commonly termed "phonological processes", following the theory of Natural Phonology described by Stampe (1979). Such terminology reflects the assumption that such error patterns are phonological, i.e. grammatical, in nature, indicating that children's knowledge of the phonological structure of the input language is still developing. Evidence supporting this is seen in natural class effects, where children spontaneously eliminate a speech pattern that affects segments at different places of articulation. For example, performance accounts given for the process of stopping of fricatives frequently put forward the difficulty of maintaining the narrow constriction required for fricative airflow (Altvater-Mackensen and Fikkert 2010, Byun 2011). However, as Byun and Tessier (2016) point out, in many children, stopping occurs across several places of articulation, which require differing structures and muscles, but nevertheless there are many accounts where children eliminate stopping across all places at once, e.g. (Rose 2014), which would not be expected in a purely performance based account. This must therefore indicate that the change is due to an addition of the continuant feature to the grammar, which is then instantiated across the board in production. Further evidence that error patterns are a grammatical rather than articulatory phenomenon, is the frequent occurrence of U-shaped learning, where children first produce a target with a high level of accuracy, then shift for a time to inaccurate production, before resuming accurate production, meaning the error pattern cannot be due to physical constraints alone (Byun and Tessier 2016).

However, other research shows that different error patterns can originate from different sources, and the same surface pattern can have a different source in different children.

Children's planning and control of motor gestures is still developing during this phase, and instrumental analyses have shown evidence that children's productions may not change categorically from incorrect to completely accurate. Instead they make a covert contrast that is sub-phonemic, but which can be measured instrumentally, although it may not be evident perceptually, or at least not evident as different enough to produce an "out of category" judgement in a listener. Munson *et al.* (2010) report a number of studies where covert contrast for a range of different contrasts has been measured acoustically, and EPG and ultrasound studies have also provided evidence both of covert contrast, and neutralisation of contrasts due to motoric rather than grammatical limitations, i.e. undifferentiated lingual gestures (Cleland *et al.* 2016, Gibbon and Lee 2016). Using evidence from perceptual rating tasks that require a gradient rather than categorical response, Munson *et al.* go as far as to suggest that covert contrast is the rule rather than the exception in speech acquisition, and Richtsmeier (2010) even claims that there is no reliable evidence for substitution in child speech errors, suggesting that all substitutions are really covert contrasts.

In speech errors that show covert contrast then, it would seem that the child must have phonological knowledge of the contrast, which he is not yet able to produce in performance. Thus for one child, for example, final consonant deletion may indicate failure to attend to the final consonant in the input, or failure to store it in the phonological representation, while in another, covert traces that can be identified by acoustic analysis indicate that the perceptual lack of a final consonant is due to an imperfect attempt to produce it, rather than lack of phonological knowledge (Kent 1992, Stackhouse and Wells 1997).

Thus from the account of children's speech acquisition given here, we can see that children have good perceptual abilities at birth, but must learn to map perceptual representations with articulatory routines for production. According to usage-based theory, they use implicit learning to organise the data from the input language to abstract phonetic categories and recognise word boundaries. When they begin to form lexical representations that match meaning to word forms, phonological representations are built first at the holistic, whole word level, and subsequently progressively segmented, allowing the child to construct the phonological structure of the input language. However, the child may not attend to the same perceptual data as the adult in building perceptual categories, and organisational, categorical and motor planning and

execution abilities continue to develop throughout early childhood and beyond. Speech development follows a similar pattern for most children, so that by the age of 6, most English-speaking children have developed a full inventory of word structures, and have acquired most segments, although alveolar fricatives and /ɹ/ may not be fully acquired (Shriberg *et al.* 1994b).

However, in some children speech is slower to develop, or develops atypically, resulting in a SSD. The following section looks at possible causes for SSD.

1.3 Possible causes for SSD

The description of speech acquisition above shows that several different types of knowledge are needed for successful acquisition. Children need to perceive, store and replicate sounds and sound sequences in words, which entails segmenting the speech stream into syllables and words, recognising and categorising sounds that they hear as exemplars of sounds that convey meaning in the ambient language, encoding and storing representations of sounds and words, and being able to map such acoustic representations to accurate articulatory plans for execution. Munson *et al.* (2005a) suggest 4 types of knowledge are necessary: perceptual, i.e. knowledge of the acoustic and perceptual characteristics of speech sounds; articulatory, i.e. knowledge of the articulatory gestures needed to produce speech sounds; higher-level phonological knowledge of the ways that words can be divided into sounds, and sounds can be combined into sequences in words, and finally social-indexical knowledge of the ways that variation in pronunciation can be used to signal social identity. This section is concerned with the first three of these as being the most important in the investigation of developmental SSD, together with consideration of sufficiency of input to the process.

1.3.1 Lack of Input

If word forms emerge from the input via statistical learning, reduced or deficient input should affect learning. The link between social economic status (SES) and language learning is well attested (Rowe 2008), and thought to be mediated by the tendency for lower SES mothers to talk less overall and less responsively to their children (Hoff *et al.* 2002, Hoff 2003). However, in speech acquisition, the number of words that are needed to start the transition from the whole word to segmental state of phonological representations is reported as being quite small: between 50 words (Ingram 1976) and

200 (Sosa and Stoel-Gammon 2006). In addition, many children with SSD have normal vocabulary (Carroll and Snowling 2004, Sutherland and Gillon 2007), suggesting lack of input is not a causal factor. Indeed, Waring *et al.* (2016) in a study of fourteen 3 to 5 year-olds with phonological delay from high SES backgrounds, found that, although the SSD group had somewhat smaller receptive vocabulary compared to matched typically developing (TD) children, both groups scored above the average on a standardised test, thus implicating a learning mechanism, rather than an absolute number of words, as an association with SSD. McNally and Quigley (2014), in the large Irish cohort study, Growing Up in Ireland, found that higher levels of maternal education predicted better language development in children aged 3, but at 9 months of age, the children of mothers with the lowest level of education were the most likely to pass the communication subscale of the Ages and Stages Questionnaire. It may be that input needs to be sufficient at a particular stage of learning (Kuhl 2004, Kuhl *et al.* 2005), so that the input needed for speech development, being a more constrained system, is different in nature to that needed for language development.

It is well evidenced that mothers use a specific register when talking to pre-lexical children ("motherese") which is more clearly articulated, and so thought to aid in learning to distinguish the phonetic units of speech. Goldstein and colleagues have shown in studies of sixty 8 and 9 month-old children, that contingent, but not non-contingent, maternal responses to their children's vocalisations increased the frequency of the infants' vocalisations and facilitated development of more advanced sounds. They concluded that infants learned new vocal forms from phonological patterns heard in their mothers' contingent speech and were then able to generalise from these patterns (Goldstein *et al.* 2003, Goldstein and Schwade 2008). Thus the form and timing, as well as amount, of input may be important. Research considering SES as a risk factor for SSD has conflicting findings. As Rowe found that use of less facilitatory communication styles with toddlers was mediated by parental knowledge of child development (Rowe 2008), it might be expected that speech development is slower in children whose mothers are less well educated. This is in fact what Campbell *et al.* (2003) found in their study of 539 3 year-olds with SSD. However Roulstone *et al.* (2002) did not find such an association in their study of 1,087 25 month-olds, although they did report that the sample in the study had a somewhat higher educational level than the whole population sample. Eadie *et al.* (2015) found an association between SSD and SEIFA disadvantage score, but not for level of maternal education, in a

community sample of 1,494 four-year-olds, and Wren *et al.* (2016) found an association between rental tenancy and presence of persistent SSD in a community sample of 14,676 eight- year-olds, though again not for maternal education level. It may be that children manage to catch up from the effects of low level of maternal education once they start receiving more input from peers, school and pre-school etc. Associations between children with SSD and SES measured by factors other than maternal education might reflect other influences such as access to remedial services, or physiological or neurological deficits linked to conditions such as inadequate nutrition or exposure to environmental hazards. Generational disadvantage, as discussed by Broomfield and Dodd (2005b), whereby families with a history of communication difficulties gravitate to lower economic strata due to the effect of speech, language and communication needs (SLCN) on social mobility and academic achievement, may also play a part if they are more likely to have children with SLCN, so that increasing disadvantage associates with increasing SLCN needs and vice versa. Dollaghan *et al.* (1999) found that language measures, but not speech measures (Percent Consonants Correct) were affected by level of maternal education, in a sample of 240 normally developing three-year-olds. Another population study, the Iowa and Nebraska Articulation Norms study (Smit *et al.* 1990), of almost 1000 children (including a representative proportion with SSD) similarly found that maternal education had no significant effect on age of acquisition of phonemes, and this has been largely supported by a number of other studies: (Harrison and McLeod 2010, To *et al.* 2013, Keating *et al.* 2001).

This evidence of little association between level of maternal education and SSD might just mean that it is a poor proxy for facilitatory input received. However, the fact that many children with SSD have normal vocabulary suggests that at least some children with SSD receive adequate input. Alternatively, it could be that the right input at the right time influences the development of a learning mechanism that is involved both in SSD and in vocabulary acquisition; for example, the ability to build well-specified phonological representations has been implicated in word-learning, SSD and reading difficulty. Studies showing that, as a group, children with SSD have poorer vocabulary than TD children have, but which also show a large overlap between the SSD and TD groups, may reflect that an interaction between input and learning is a factor in SSD for at least some children.

Another factor that reduces language input to children with developing sound systems is otitis media. However, most children suffer at least one episode of otitis media, with over 10% having recurrent episodes (Bradford and Dodd 1996), but not all children with a history of otitis media have SSD. Nor do all children with SSD have a history of otitis media, meaning it cannot be a sole cause. A study of seventy-eight 3 to 5 year-old children with SSD and 87 controls (Dodd and McIntosh 2008), found only approximately 6% to 7% of children had a history of ear infections, with no difference in proportion between groups. The influence of fluctuating hearing loss caused by otitis media on SSD is discussed further in sections 1.5.2 and 3.2.4, but there is little evidence that it has a large effect on SSD.

1.3.2 Deficits in perception

A diagnosis of functional SSD rules out children with hearing that does not meet normal thresholds, but the idea that more subtle perceptual deficits may be a causal factor in phonological disorder has been investigated since Locke proposed it in 1980 (Locke 1980). Research previously cited in section 1.2.1 highlights infants' ability to perceive many non-native contrasts in speech when they are under 6 months of age; however, Nittrouer (2001) points out that in such experiments, there are frequently attrition rates of over 40% of infants who are unable to complete the experiment, giving rise to the possibility that at least some of them do not complete the experiment precisely because they are unable to perceive the distinctions. In fact Kuhl *et al.* (2005) showed that children who could still perceive non-native language contrasts at 7 months had worse language outcomes than those who did not. A possible mechanism for this, suggested by Velleman and Vihman (2002), is that a child with impaired implicit learning ability could get overwhelmed by "innumerable phonetic cues", affecting their ability to focus on the parts of the speech stream that allow them to associate particular word forms with meanings. Thus efficient language learning involves "learning what to ignore as well as what to attend to" (Bishop 2000, p.136).

Thus perception of the speech signal as such may not be as important as what is done with the perception, i.e. what happens where auditory perception interfaces with higher linguistic or cognitive processes (Ziegler 2008). While children need to encode enough phonetic detail for accurate production, they also need to make generalisations over the phonetic parameters to discover the categorical structure of the phonological system, as an implicit learning task. If we impute the learning of phonetic/phonemic categories to a

statistical learning mechanism, then a reduced ability to learn probabilistically from the input might cause slower progress in speech acquisition or mislearning, such as idiosyncratic setting of category boundaries, which might be interpreted as deficits in perception. A failure to attend to necessary parts of the signal, or to extract the correct cues would do the same. Similarly, restricted phonological memory might result in difficulties in encoding perceived speech so that deficient phonological representations result. This could then affect performance in word discrimination tasks, as well as mispronunciation detection tasks. Therefore it is difficult to differentiate whether poor performance in tasks used to assess perception results from poor speech perception or from further processing of the percept at higher cognitive/linguistic levels. For example, Nittrouer *et al.* (2011) in a study of 14 eight-year-olds with language impairment, and 14 age-matched peers, found that the children with language impairment that included phonological processing deficits, showed equivalent perception but poorer category formation for non-speech sounds. They suggested that children with language and phonological processing deficits may have an underlying deficit in organizing sensory information to form coherent categories rather than poorer attention to the relevant acoustic properties. This would seem to implicate cognitive/linguistic processing rather than perception per se. However, speech perception is used in the literature to express abilities relating to word recognition, sound categorisation and sound and word discrimination, so I will follow that practice in this section.

It is clear that categorical discrimination of speech sounds is still developing during the preschool years and early school age (Nittrouer 2002b, Munson *et al.* 2005a). Nittrouer (2001) tested 41 children aged 2;6 to 3;4, with typically developing speech and selected to have no risk factors for speech or language impairment, on discrimination of /s/ versus /ʃ/ contrasts. Only 59% were successful, despite all completing the same task with vowel contrasts successfully. Thus this could be one way (but not the only way) to account for the error pattern of fronting /ʃ/ to /s/ which is commonly seen in normally developing children at this age, where the phoneme boundary has perhaps been set using the wrong parameters, or the child has not yet completed the segmentation process for more similar sounds.

There is evidence that some children with SSD have difficulty discriminating sounds they do not produce (Rvachew and Jamieson 1989, Shuster 1998). In a study of 26 children who had difficulty with the production of /ɹ/, Shuster found that they had most

difficulty in judging whether their own recorded productions were correct. Jamieson and Rvachew (1992) demonstrated that training children who had both poor production and perception of /ʃ/ to discriminate it from other sounds, improved production of the sound as well as discrimination, although it had no effect on the production of a child who could already discriminate it well, suggesting that deficient perception is a factor in SSD for some children. However, if perception difficulties are limited to those contrasts the child does not produce, the direction of influence may also be reversed: the child may not perceive a contrast well that he or she has never produced, according to the reciprocal development of motor and perception abilities suggested by Velleman and Vihman (2002) and Byun (2012).

There is also evidence of more generalised deficits in speech perception for some children with SSD. Edwards *et al.* (1999) in a study of 6 children with mild to severe phonological SSD, matched to 6 TD children and 6 adult controls, found that the group of children with SSD were significantly less accurate than the group of TD children in identifying backward-gated CVC words (where portions of the final consonant had been deleted), and in identifying CVC words with degraded vowels, even though they performed as well as the TD group in live voice presentation. Only one of the children with SSD omitted final consonants in production, showing that the majority had adequate phonological knowledge and motor ability for CVC structures. Edwards et al. concluded that the children with SSD had more fragile phonological representations of final consonants, and they needed more acoustic information to recognize segments than TD children did. However, at least one child with SSD performed as well as the TD children, indicating that not all have difficulties with speech perception. For example, Nathan and Wells (2001), in a study of five-year-olds, found that some children with SSD performed just as well as TD children in tests of non-word discrimination. In contrast, results in studies of older children and adults do not show such an overlap in discrimination abilities between impaired and non-impaired groups. Gósy and Horváth, (2014) tested non-word discrimination (NWD) abilities in sixty 7 and 8 year-old children with SSD, and found they had significantly poorer scores compared to age matched controls, with no children with SSD performing as well as a TD child. In a study of 9 adults with persisting SSD, Kenney et al. (2006) found all had deficits in discrimination of fine acoustic cues for word identification when compared to 20 controls. The fact that there is no overlap at older ages might suggest that development of perception abilities could differentiate children with transient as opposed to persistent

SSD. Those children with good perception abilities whose deficits lie at the output side of speech production might catch up with their peers in terms of speech accuracy, whereas those whose perception abilities are deficient might not, as poor perception leads to phonological representations inadequately well specified to form a basis for accurate motor programmes for production. In support of this hypothesis, Stackhouse et al. (2007a) found that children whose SSD persisted performed less well on measures of auditory discrimination than those who resolved at all assessment times. A longitudinal study by Rvachew (2006), which followed the progress of 47 prekindergarten children with SSD at ages of 4 to 5, found that speech perception skills in prekindergarten did indeed predict some of the variance in both articulation and phonological awareness skills at the end of kindergarten. They found that about half the children with SSD had speech perception deficits, and that these children also had phonological awareness deficits.

In a review of 14 papers on speech perception in phonological disorder, Shiller *et al.* (2010) found that 13 of them reported evidence of significantly reduced perception skills in children with SSD. Thus it may be that some children with SSD lack the perceptual accuracy and thus the interface with cognitive/linguistic functions for encoding that is needed to allow the construction of accurate phonological representations for sounds that is the necessary basis for accurate motor output. Children who are able to construct adequate phonological representations, however, may still have deficits further along the line to output.

1.3.3 Cognitive/linguistic deficits

As discussed in the preceding section, it is surmised that the building of accurate phonological representations requires an interaction between perceptual abilities and implicit learning mechanisms that are cognitive in nature. As pressure mounts for organisation of the growing lexicon, children must abstract higher-level phoneme categories from phonetic categories, along with the rules that govern how sounds can be sequenced, where they can occur in a word and alternations due to phonetic context. As this phonological grammar is specific to each language, it must emerge, or be constructed, from the input by each child (Ramus *et al.* 2010), using cognitive abilities such as statistical and probabilistic learning of categories and prototype abstraction, to implicitly extract regularities and derive the rules of the ambient phonological system

(Dodd and McIntosh 2010). Thus any impairment to this learning mechanism could result in difficulties with acquisition of the phonological system.

Evidence that some children with SSD show an ability to produce speech sounds that they do not use in appropriate contexts, e.g. "mashing" pronounced /mæsɪŋ/ but "passing" as /pæhɪŋ/ (Fey 1992), is taken to show that it is the failure to acquire the correct rules governing use of speech sounds in specific contexts or word positions which underlies the SSD in those children, rather than any perceptual or motoric constraints. Such children are considered to have a phonologically based disorder, that is, a SSD which reflects a cognitive or neurological deficit in how phonemes are represented and organised within the language specific phonological system, or grammar (Fey 1992). Some accounts, then, see phonological SSD arising, in at least some cases, from an impaired understanding of the structure of the phonological system, due to delays or deficits in abstracting the rules that govern said system (Dodd 2011).

Dodd and colleagues have carried out a number of investigations to establish whether children with SSD do show the difficulties in rule abstraction and knowledge of phonological grammar thus hypothesised. Dodd *et al.* (1989) tested 41 children aged between 3;5 and 5;2 on a phonological legality task, where they had to choose which of 2 animals had the best name, where one had a phonologically legal non-word name, and the other an illegal non-word name that differed by one phoneme e.g. /ʃroupi/ and /ʒroupi/. One group of children with SSD, termed the consistent phonologically disordered (CPD) group, defined by use of atypical phonological errors, showed no preference for legal over illegal names, whereas the children with delayed and inconsistent phonology performed similarly to TD peers, showing a preference for legal non-words. Dodd *et al*. thus concluded that the CPD group showed a deficit in the cognitive processes involved in deducing the system of contrasts specific to the phonological system of their ambient language. However, although significant, the difference in mean scores was not great, with the CPD group choosing six legal non-words out of 12 (i.e. performance at chance), and the TD group only 7.5. In addition, the sample size was small, with the CPD group containing only nine children.

Torrington Eaton and Ratner (2016) concur with Dodd and colleagues in suggesting that a deficit in domain-general cognitive processes could be involved in difficulties in

constructing an accurate phonological system. For example, they point out that working memory is used to hold and manipulate the input while constructing and updating phonological rules, inhibitory control might be involved in suppressing outdated speech patterns, and cognitive flexibility is needed to switch from an old to a new, more accurate production pattern as maturation occurs. This was investigated by Dodd and colleagues, who examined the ability of children with SSD on tasks of cognitive flexibility and rule abstraction in three studies. Dodd and McIntosh (2008) investigated these abilities in 78 children with SSD and 87 matched controls, aged 3;0 to 5;7. The rule abstraction task was similar to the Wisconsin Card Sorting Test (Heaton 1993), but modified to be appropriate for the age group and presented on a computer touch screen. The cognitive flexibility task was the Flexible Item Selection Test (Jacques and Zelazo 2001). In this, the children were shown a set of three cards containing line drawings of coloured objects, and asked to select two that matched, on any of the three dimensions of colour, size or object shape. After making the first selection, they were asked to find two from the same set that match in "a different way". The first selection tested the ability to abstract a categorisation rule, and the second showed the ability to demonstrate cognitive flexibility in changing the selection rule to a different parameter. Children with SSD performed significantly worse than the controls as a group on the first rule abstraction task, deriving a mean of 1.1 rules out of four, as opposed to 1.6 in the control group. On the FIST, both groups performed better than chance and equally well on the first selection, but on the second, the children with SSD performed significantly worse, with 17% of children with SSD performing at chance level, in contrast to only 3% of control participants. Dodd & McIntosh suggested that the difficulty in deriving rules, shown by the children with SSD in the results from the first task, might also reflect difficulty in deriving the rules of the ambient phonological system, such that the wrong rules may be acquired, resulting in atypical phonological errors. The lack of cognitive flexibility shown in the results of the second task by the children with SSD, they suggest then, may reflect a persistence in the use of inappropriate rules that also affects persistent use of outdated phonological rules. However, there was considerable overlap in results between the SSD group and controls on all tasks, indicating that not all children with SSD have rule abstraction or cognitive flexibility deficits.

To investigate whether such deficits affected an identifiable sub-set of children with SSD, Crosbie *et al.* (2009a) compared performance on the FIST as described above, in

three groups of 5 year old children: fifteen were typically developing (TD) controls, 15 had SSD characterised by consistent atypical errors, termed the consistent phonological disorder group (CPD) , and 15 made inconsistent errors, that is, they produced whole words in more than one error form, showing inconsistent patterns of substitution and omissions for the same word. This group were termed the inconsistent phonological disorder group (IPD). Results from the FIST showed that the children with CPD performed worse on both selections than the other two groups, who did not differ from each other, with large effect sizes. Crosbie et al. concluded that children with CPD showed a generalised difficulty with rule abstraction and cognitive flexibility in categorization tasks, which underlay their speech difficulties, whereas children with IPD did not. Such an explanation was in accordance with the fact that the speech of the CPD group was characterised by the use of atypical error patterns, i.e. abstraction of the "wrong" rules, as well as with research that showed that children with CPD have poorer phonological awareness than children with SSD characterised by other types of error, and are at most risk for literacy difficulties (Leitão *et al.* 1997, Leitão and Fletcher 2004, Preston and Edwards 2010), thus suggesting that some children with SSD may have wider difficulties with rule based systems. However, a further study of younger children did not show an association between performance on a rule derivation task and speech accuracy: Dodd and McIntosh (2010) tested rule abstraction skills and speech accuracy in 62 TD 2 year olds. A task where the child had to work out which farm animal was required by the researcher pretending to be a farmer, from feedback on their selections, showed no correlation between ability to derive non-verbal rules and speech accuracy. However the mean number of rules derived was low, 1.4 out of 4, suggesting the task may have been too difficult for many of them, although some achieved the maximum score.

Further evidence that deficits in rule derivation and cognitive flexibility are not evident in all children with SSD was shown in Torrington Eaton and Ratner (2016)'s study of a group of 4 to 6 year-olds, where 42 were in the high average range on the Goldman Fristoe test of articulation, 11 were in the low average range, and 9 had SSD. The researchers found that the children with SSD performed best on the FIST, and were statistically indistinguishable from TD children who scored in the high average range. In this study, the small number of children with SSD may have meant that a sample biased to a particular sub-type of SSD without such deficits was recruited. If such deficits are limited to an identifiable sub-group, then investigating whether different

sub-groups have different outcomes, would show potential for a test with good predictive power.

The findings of the investigations described above thus provided preliminary evidence to support the proposition that some children with SSD, specifically those who use atypical error patterns, may have an underlying deficit in understanding or constructing the rules of the phonological system.

However, rules that result in atypical productions could also be viewed from a different perspective, not as a deficit, as much as a novel and creative attempt to maintain contrasts that the child perceives in adult speech, but is not yet able to produce. For example, Fey (1992) presents the example of a child who produced word-final voiceless stops correctly, but word-final voiced stops were produced as oral stops released with an accompanying nasal, e.g. “back"→ /bæk/ but "big" → /bɪgŋ/. He explains this unusual pattern as the child’s attempt to preserve the phonological contrast between voiced and voiceless stops in word final position, whilst not yet being able to articulate such a distinction in the adult manner. In the example above, ("mashing" pronounced /mæsɪŋ/ but “passing” as /pæhɪŋ/), if the child is not able to produce /ʃ/, the strategy of using /s/ for /ʃ/, and /h/ for /s/ in a context where /h/ cannot normally occur, similarly maintains the contrast between words (although doing it the other way round would have resulted in a greater number of accurate targets). These explanations suggest atypical productions might show a knowledge of the phonological system and attempts to signal it under constraints arising from other factors, rather than showing a deficit in rule abstraction ability, i.e. children abstracting the “wrong” rules (Dodd and McIntosh 2010).

In order to investigate whether the production problems in children with phonological SSD did, in fact, result from difficulties in making abstractions over lexical items, Munson *et al.* (2011) compared phonotactic knowledge in forty 3 to 6 year-old children with SSD, to that of 40 same age TD peers, using a non-word repetition (NWR) task, where performance on non-words containing low frequency phoneme sequences was compared with performance on non-words containing high frequency sequences. Children in the SSD group performed worse overall, as would be expected given their speech difficulties, but both groups showed a similar effect of frequency, with low frequency sequences being repeated less accurately than high frequency ones.

Frequency did not affect children with SSD proportionally more than TD children. Instead the size of the frequency effect was correlated with vocabulary size, but not with speech perception, (as measured by a real word minimal pair discrimination task), nor articulatory ability, leading to the conclusion that the production difficulty with low frequency sequences was associated with size of vocabulary, but not with the development of articulatory or perceptual ability, i.e. not with SSD status. They concluded therefore that children with SSD did not have more difficulty with abstracting phonological knowledge from lexical items than TD children, because they did not show relatively more difficulty in manipulating segments from low frequency sequences than TD children did. These results did not altogether concur with those of McKean *et al.* (2013), who found in a longitudinal study, that, although there was no difference in overall NWR performance or frequency effect once the level of vocabulary size of the two groups was accounted for, the gap between accuracy of NWR between high and low frequency words did not narrow with increasing vocabulary as fast for children with language impairment as it did for TD children. However, McKean et al's participants all had primary language impairment, whereas Munson et al's all had vocabulary scores within the normal range, suggesting normal language abilities. As there is some evidence that children with language impairment have more difficulty in learning words with low phonotactic probability (Alt and Plante 2006), their lexicons might be qualitatively different despite being of the same size, containing fewer words with low frequency sequences and thus accounting for the difference in NWR of words with low phonotactic probability, rather than showing a difficulty with phonological abstraction.

So might other domain-general cognitive mechanisms contribute to speech accuracy? Some studies have found reduced short-term verbal memory in children with SSD, compared to TD children. Short-term memory would logically be involved in forming stable phonological representations: memory traces of speech input need to be sustained long enough for acoustic features for phonological encoding to be abstracted, so if perceptual representations decay too quickly, the phonological representations that form templates for motor planning and production may be inaccurate. In addition, reduced storage would impair the child's ability to compare her/his own articulations with adult targets, in order to refine and revise both phonological representations and motor programmes, as perceptual and motor abilities mature, and input/vocabulary increases. The latter task involves simultaneously storing and processing verbal information, so

that working memory is also implicated. Non-word repetition (NWR) tasks are frequently used to investigate memory and executive function, but results are difficult to interpret in children with SSD, partly because speech output errors are likely to confound results, but also because NWR tasks involve speech perception, encoding and motor planning abilities, as well as memory, making it difficult to pinpoint the locus of deficit. To obviate this confound, Waring *et al.* (2016) used a different task to investigate short term and working memory in 14 children with the phonological delay (PD) type of SSD, where children make errors that are typical, but only of much younger children, and 14 matched TD controls, aged 3 to 5 years. Short-term memory was tested using a pointing task, where the child had to remember a list of items spoken by the researcher, and point to them in the sequence spoken. The number of items increased from two to eight. Working memory was tested by two tasks: a similar pointing task, where the child had to point to the items in **reverse** order to presentation, and the standardized Numbers Reversed subtest from Woodcock–Johnson III Tests of Cognitive Abilities (WJ-III) (Vought and Dean 2011). Their findings showed that the children with PD had poorer performance on all tasks, indicating poorer phonological short-term memory (STM) and working memory (WM) abilities, compared to their peers without SSD. The difference between the STM and WM scores was the same for both groups, indicating that the poorer WM scores were due to holding, rather than manipulating, items in memory. However, there was almost complete overlap in scores between groups, where some children with SSD scored as well as those without, and some TD children scored as poorly as those with SSD, suggesting that poor STM and WM is neither necessary nor sufficient as a causal factor in SSD. Waring's results concurred with those of Torrington Eaton and Ratner's (2016) study, referred to above, who also found that the SSD group performed worse on a task of forward digit span, with a significant positive correlation between number of digits remembered and speech accuracy. A larger study by Lewis *et al.* (2011) which included 69 TD children, 73 with SSD alone and 53 with SSD and co-morbid language impairment, all aged 4-7 years, found that children with SSD alone did not differ from TD children in measures of phonological memory, but those with the co-morbid condition had significantly poorer scores than the other groups. However, phonological short-term memory was measured using a battery of tests including digit span, sentence repetition and the repetition of multisyllabic real and nonsense words, meaning that other constraints on speech perception, encoding, output and/or syntactic ability could have influenced results. In a study of older children with SSD, aged 7 to 12, Murphy *et al.* (2014) found that 18

children with SSD performed significantly worse than 37 TD children on a forward digit span test, but the range of scores was not given, so it is not clear whether all had STM deficits. The children with SSD also had significantly worse language skills than the TD children, which may be a confounding factor in evidence for whether STM has a causal or maintaining effect on SSD. Kenney *et al.* (2006) studied nine adults with residual speech sound disorders and found significant differences in forward and backward digit span between the adults with SSD and the 20 controls. There was a only small overlap between scores of those with SSD and controls, but scores of adults who made reduction errors were at the bottom of the range compared to adults with distortion errors (more likely to be motor rather than phonologically based), indicating that STM deficits could be a factor in persistence of phonologically based SSD. Thus, the differences found in STM between people with and without SSD at older ages might show that STM has a predictive value in identifying transient versus persistent SSD.

An alternative theory is put forward by Keren-Portnoy *et al.* (2010). They argue that, rather than STM underlying speech development, speech production practice drives the development of the phonological loop, so that it grows as children learn to categorise sounds, segment speech and plan and execute speech. In a longitudinal study of fifteen 26 month-olds, they found that ability in NWR tasks was better for consonants that had been practised longer, i.e. acquired earlier, and that children's production of nonwords constructed of consonants they used frequently was more accurate than productions of comparable nonwords made up of consonants they used infrequently, but which were "developmentally equally accessible". They also showed that children made holistic rather than segmental errors in repeating nonwords containing infrequently used consonants, suggesting that the infrequent consonants were not yet available to be used as segments. However, this effect would seem to be better explained by a delay in the segmentation of phonological representations than increase in phonological memory, and repetition effects could equally as well be due to better availability of gestural plans for syllables containing consonants produced frequently. However bi-directional effects cannot be ruled out.

Torrington Eaton and Ratner's (2016) study was the only one that investigated inhibitory control. Evidence from intervention studies, e.g. Gierut *et al.* (2010), has shown that using non-words as intervention targets rather than familiar words can facilitate success in treatment. Torrington Eaton and Ratner therefore hypothesised that

difficulties with inhibition of early production patterns for familiar words may interfere with speech development in children with SSD. However, results showed no differences in performance on tasks meant to capture this between TD children and children with SSD. Limitations included, as mentioned above, the small size of the SSD group, and the fact that the tasks also relied on visual-spatial processing, suggesting more research is needed in this area before conclusions can be drawn.

The findings outlined above suggest that some children with SSD show deficits in domain general cognitive functions that could feasibly affect their acquisition of the structure of their phonological system. However, as they are domain general, it would appear unlikely that they affect the acquisition of phonology alone. Given the fact that not all children with SSD demonstrate such deficits, three explanations are possible:

1. SSD arises from a multiplicity of sources, perceptual, cognitive, or motor, with different children having different deficits. This explanation accords with the heterogeneity of presentation and course of development of children with SSD.
2. Domain general cognitive deficits originally contributed to difficulties with the acquisition of phonology, but have since been outgrown, leaving the phonological structure, however, fixed at an immature stage. McKean *et al.* (2013) give an explanation, based on neuroconstructivist theory (Karmiloff-Smith 1998), for how this might happen, explicated in section 1.3.4 below.
3. For some children, cognitive deficits accompany, rather than underlie, SSD. However, such accompanying deficits might cause SSD to be more resistant to change/treatment. This explanation accords with the findings that some children perform as badly on tests of STM, cognitive flexibility or rule abstraction as children with SSD, but have normally developing speech. It also helps to explain why some, but not all, children with SSD also have difficulties with phonological awareness (Rvachew *et al.* 2003), reading (Lewis *et al.* 2015) or maths (Nathan *et al.* 2004b).

Whichever theory holds, the presence of cognitive deficits might provide a way of sub-typing children with SSD that indicates whether the SSD is persistent or transient, and whether the child is likely to go on to have difficulties in other areas.

1.3.4 Deficits in phonological representations

As discussed in section 1.2, when children begin to speak, their phonological representations of words are holistic, i.e. based on large parts of the speech signal (Vihman and Croft 2007, Anthony *et al.* 2010). In normal development, as children acquire more vocabulary, their representations become increasingly more refined and segmented into smaller and smaller units until they reflect the precise phonemic structure of the word. One explanation given for SSD is that children may become "stuck" in an early-developed template at the whole word level so that all words have to be adapted to fit a specific template or set of templates. Stackhouse and Wells (1997) suggest that children whose speech output is constrained by a small set of possible word structures have not progressed from the Whole Word phase to the Systematic Simplification phase, which would mean that they have not progressed to segmenting phonological representations, but are still storing them as gestalt, whole word shapes. This has similarities with Bishop's suggestion that children with developmental language disorder (DLD) become "stuck in immature modes of processing language, which hamper new learning" (Bishop 2000, p.139). She suggests that some children may stick at the level of syllabic analysis, without progressing to sub-syllabic units, so that analysis of the input is inefficient, as children cannot make use of phoneme level knowledge from words already learned to facilitate the acquisition of new words. For word learning this makes storage and retrieval more difficult, whereas at the phonological level it would obscure the phonological structure at the segment level.

McKean *et al.* (2013) describe one mechanism for how children might become stuck at an immature level of processing. Following developmental neuroconstructivist theory as developed by Karmiloff-Smith and Thomas (Karmiloff-Smith 1998, Thomas and Karmiloff-Smith 2005), they describe the process of learning and development as an "iterative cycle" of uptake, abstraction and reorganisation, where the child learns new information using whatever processing mechanisms are available at that stage of her/his development, and finds regularities in the new data from which rules or representations can be abstracted. The addition of new rules or abstracted patterns then instigates reorganisation and change in the processing system, such that, with every iteration, it becomes more sophisticated, moving from domain-general to domain specific, so that what is learned in each cycle of the uptake phase is altered as a result of these iterative changes. However, in the child who has deficits in such genetically determined domain-general abilities, such as implicit learning, or memory, as discussed in the section

above, these interact with his/her experience to produce a perturbation in the process of development.

Viewed using a connectionist model of language acquisition, if the model settles on a particular pattern of processing before it has received sufficient input, and continues to process subsequent input the same way, it cannot learn further from additional input, and development stops. McKean *et al.* (2013) suggest then, that if children's word learning is slow, the speech processing mechanism fixes on a pattern of processing derived from too few lexical representations, where the process for abstraction of sublexical representations is then undeveloped. The acquisition of more lexical items does not then effect the further sophistication of phonological processing that occurs in TD children, because a critical mass of vocabulary was not reached within the developmental window for developing an efficient processing system, resulting in phonological representations that are insufficiently defined, incomplete or inaccurate.

One factor that needs explanation is that most children with SSD gradually improve, if not overcome, their SSD over a number of years (Shriberg *et al.* 1994a, Preston *et al.* 2013), which would not seem possible if learning and phonological representations become fixed at an early developmental stage. One explanation might be redundancy in the input, where children hear words so many times that they can continue to collect new information about them (Snowling and Hulme 1994). Alternatively, change in the speech processing mechanism might be reactivated at a later stage, by another factor, such as feedback in communication, or reaching a critical level of literacy attainment, which stimulates phoneme knowledge (McKean *et al.* 2013) and increases phonological awareness.

Incomplete or inaccurate phonological representations are suggested to be the common factor that underlies the higher risk of difficulties with the acquisition of literacy reported for children with SSD (Bird *et al.* 1995, Carroll and Snowling 2004). It is posited that phonological representations that are not defined at the segmental level result in less efficient word learning (Bishop 2000), leading to impaired language ability, and, if they remain undefined by the time a child learns to read, also result in phonological awareness difficulties and reading difficulty (Boada and Pennington 2006). Regarding speech, unsegmented phonological representations would lead to difficulties with abstracting the structure of the phonological system, as well as in

producing accurate motor programmes for speech output. However, if immature phonological representations underlie SSD, as well as LI and RD, it is difficult to see why so many children with SSD do not also show language and reading difficulties. In fact, a number of studies have shown that children with isolated SSD have almost no risk for reading difficulty (Lewis *et al.* 2000, Sices *et al.* 2007, Carroll 2008, Pennington and Bishop 2009), although others have contradicted this (Carroll and Snowling 2004, Nathan *et al.* 2004a). However, there is evidence that children with SSD are at risk for spelling difficulties, particularly those whose SSD persists after age 6 (Lewis *et al.* 2000), where spelling is another ability that would logically depend on well-defined phonological representations. However, it is children with SSD+LI who have a large risk of reading difficulties, as do children with LI alone, suggesting that language skills may be more strongly related to reading abilities than speech skills (Sices *et al.* 2007, Carroll 2008, Skebo *et al.* 2013).

However, not all children with SSD show signs of having been arrested at the phase of unsegmented phonological representations. Where children make systematic substitution errors of individual segments, it implies that they have reached the stage of segmented phonological representations. The fact that not all children with SSD have difficulties with phonological awareness (Rvachew 2007), which requires knowledge of the sublexical structure of phonological representations, also provides evidence that not all children with SSD have phonological representations that are insufficiently segmented. Difficulties with speech perception, encoding and processing the acoustic-phonetic features of speech input, as discussed in section 1.3.2, could result in phonological representations that are segmented, but still inadequately defined at the segment level. If, as Stackhouse and Wells (1997) suggest, the motor programme for production depends on the phonological representation for its own specification, inadequately defined segments in phonological representations would lead to production targets that were poorly defined and poorly differentiated from each other. The relationship between perception and production deficits may not necessarily be a direct one: where perceptual targets or phonological representations for segments are ill defined, children may default to patterns in production that are more frequent, well learned, or more easily articulated (Munson et al. 2010a).

Macrae *et al.* (2014) proposed that unstable representations at the whole word level are an underlying deficit in children with SSD characterised by high word variability, and

are linked to lower levels of vocabulary, while children who produce particular speech sounds, rather than whole words, in a variety of ways, have indistinct phonological representations for errored sounds at the categorical phonemic level. They used a NWR task, but one designed specifically for children with SSD, the Syllable Recognition Task (SRT) (Shriberg *et al.* 2009), and a receptive vocabulary test to investigate this. However the participants in the study comprised only 18 children with SSD over a wide age range (3 to 5), and although correlations in the results showed trends supporting the proposition, they were non-significant.

One problem with testing hypotheses about the quality of phonological representations is that it is difficult to design tasks specific enough. Non-word repetition tasks are argued to reflect the quality and persistence of phonological representations and test children's ability to construct immediate phonological representations (Gathercole 2006) while eliminating word familiarity effects which could act as a confound. Children with SSD have been shown in several studies to perform more poorly than TD children on these tests (Munson *et al.* 2005b). However as (Smith 2006) points out, it is not clear how accurately input, memory and output factors can be distinguished in NWR performance as it involves auditory, cognitive, phonological and output processing, so a failure in any one of these, as well as a combination of several factors, could all lead to errors in the task. This problem is compounded in testing children with SSD, as problems with representation are difficult to differentiate from errored productions due to output constraints, and it is especially prominent when testing pre-schoolers, who have difficulties with producing multi-syllabic words regardless of speech status.

To avoid the influence of output difficulties, some studies have used receptive-based judgement tasks where children have to judge the correctness of spoken real words. Rvachew *et al.* (2003) compared the performance of thirteen 4-year-old children with SSD but normal receptive vocabulary, with that of 13 same-aged TD children, on a mispronunciation detection task using the SAILS test of phonemic perception, where four single syllable words are produced in both correct versions, or versions with a changed first phoneme. The children with SSD performed significantly worse than the controls on both that test and a test of phonological awareness, allowing Rvachew et al. to conclude that they had phonological representations that were imprecisely specified and inadequate at the level of segmental organisation. Sutherland and Gillon (2007)

also used receptive tasks to judge the quality of phonological representations. They followed the progress of nine children aged 3 to 5, with severe SSD at initial assessment, and 17 same age TD children for 12 months. They were assessed at three intervals using mispronunciation detection tasks with real and newly learnt non-words, as well as on speech accuracy and phonological awareness and reading measures. Children with SSD were significantly worse at mispronunciation detection and non-word and real word repetition at all times than TD children, but their ability on the mispronunciation detection task, as a group, improved as their PPC improved. PPC at time 2 and 3 was significantly and highly correlated with performance on mispronunciation tasks at times 1 and 2, but not at time 3, possibly because PPC growth slowed as some children approached ceiling. The study identified stronger, though non-significant, correlations between performance on PA tasks and mispronunciation detection and non-word learning tasks, than between PA tasks and speech production ability, thus suggesting that both PA and speech production ability are associated with the ability to form precise and detailed phonological representations of words. Vocabulary scores did not differ between groups, suggesting that the children with SSD were not able to gain as much detail about the specification of phonological representations from learning words as the TD children, indicating the possible contribution of perception/cognition in forming phonological representations.

Phonological awareness (PA) ability also demonstrates the quality of phonological representations, as children require the representation to be increasingly well defined in order to isolate number of syllables, onset and rime, and individual segments. As the term suggests however, it requires explicit, metalinguistic awareness which even TD children under the age of 4 frequently do not have access to (Carroll *et al.* 2003, Hesketh *et al.* 2007), and is amenable to explicit instruction, as well as having a bi-directional relationship with reading ability, both of which can act as confounds. For this reason, findings must be interpreted with caution: it would be wrong to impute a deficit in a skill that it would not yet be reasonable to consider acquired. Nevertheless, a number of studies have shown that children with SSD have poor phonological awareness relative to TD children (Rvachew *et al.* 2003, Leitão and Fletcher 2004, Raitano *et al.* 2004), particularly at levels of segmentation below the syllable, e.g. onset and rime, phoneme identification (Carroll and Snowling 2004, Harris *et al.* 2011), providing support for the proposition that children with SSD have inadequately segmented representations. Poor PA has been found even in children with only mild

SSD (Gernand and Moran 2007, Rvachew *et al.* 2007) and children without concomitant language impairment (Rvachew *et al.* 2003, Gernand and Moran 2007). However, again, not all children with SSD show PA difficulties. Rvachew *et al.* (2007) found that 39/58 (67%) of four-and-a-half year-olds with SSD scored below 1 SD from the mean on the Phonological Awareness Test (PAT; Bird *et al.* 1995), falling to 27/58 (47%) at age five-and-a-half.

To account for children with SSD who do not show deficits in tasks designed to reveal the quality of phonological representations, a further possibility is that while phonological representations might be intact, errors might occur when they are accessed from memory. Munson and Krause (2016) considered whether SSD is associated with a deficit in phonologically encoding words for output that have been accessed from memory. They tested this in a cross-modal priming task with thirty-six children aged 3 to 7, 18 with TD speech and 18 with SSD, who named pictures while listening to interfering words which varied in similarity to the picture's name. They found that the facilitatory or inhibitory effect of the interfering words was the same for both groups, leading them to conclude that there was no difference in encoding ability between children with SSD and TD children, so speech errors in SSD do not appear to be due to poorer abilities for phonologically encoding words for output, after they have been accessed from memory.

The quality of phonological representations has been researched more often in children with reading rather than speech sound disorders, but the evidence for poor phonological representations in some children with SSD outlined here suggests that poorly specified representations may be the factor that leads to increased risk for phonological awareness difficulties and reading disorders. However, what is not explained is why children with language and reading difficulties do not also have SSD, as numerous studies have shown that many of them also perform poorly on tasks claimed to test the quality of phonological representations (Ramus *et al.* 2013, Loucas *et al.* 2016). Carroll and Snowling (2004)'s study of 17 children, aged between 3;11 and 6;06, with SSD only, each matched to a child at family risk of dyslexia and a control child, did, in fact, show that speech production difficulties were more common in the group of children at risk of dyslexia than in the control group, with 6 of the 17 obtaining PCC scores, in a test of expressive phonology, lower than the lowest scoring control child. However, a number of the children in the SSD group had higher PCC scores than some control or family

risk children. In their study, both the children with SSD and those with a family history of dyslexia performed very similarly on a mispronunciation detection task, and the difference between those two groups and the control group was small (d = .13), but they performed much more poorly than controls on a new word learning task, with those with SSD performing worst of all. They concluded, then, that the risk factor shared by children with SSD and children with a family history of dyslexia was poorly specified phonological representations.

Phonological representations gradually mature throughout the pre-school years and longer, so underspecified phonological representations could thus be a cause or a symptom of SSD. If a child's representations fail to develop past the holistic stage, constraints on form or template shape might lead to structural errors in output, or underspecified motor programmes resulting in inaccurate articulatory parameters for words. The segmentation process is gradual (Fikkert 2010), and a delay in this process might lead to substitution errors, as underspecified representations at the segment level, possibly due to inefficient perceptual/cognitive processes, result in poorly differentiated motor outputs, or reliance on default values. Such a gradual shift from holistic to segmented representations would predict a preponderance of structural errors at early stages, giving way to more substitution errors as the overall shape became segmented and faithful to the number of segments, while still lacking the detail to differentiate segments adequately. Although a very small sample of only 9 children with SSD, Sutherland and Gillon (2007)'s study showed just such a trend over the three time points of the study in 8 of the 9 of the children. Shriberg *et al.* (1994b) also showed that children whose SSD persisted after 1 year had more omission errors at initial assessment than children whose SSD normalised, suggesting their phonological representations might have been at an earlier stage of development when first tested.

Thus the argument that children with SSD have holistic or underspecified phonological representations provides a logical account of their presentation, and the link to risk for PA and reading difficulties. However, measuring the quality of phonological representations is difficult as the tasks devised also rest on other cognitive, perceptual and output abilities. From the evidence outlined here, it seems that, if we are measuring it accurately, some children with SSD do have deficits in phonological representations, but again, only a subset. An alternative explanation for others may be that they have difficulties with motor programmes, planning or control.

1.3.5 Motor deficits

Traditionally, SSD was considered to be due to a deficit in articulatory ability (Davis 2005). However, a diagnosis of functional speech disorder rules out frank motor deficits such as those found in dysarthria, and deficits in motor planning and/or programming which are generally accepted to be at least one component of deficits in childhood apraxia of speech (ASHA 2007). Nevertheless, there is evidence that children with purely functional SSD may present with sub-clinical signs of difficulties with motor control: for example, Edwards (1992) found that children with SSD were less able to compensate for articulatory perturbation, e.g. in producing vowels when their jaws were fixed in place, than TD children.

A number of studies have demonstrated an association between developmental language disorders (DLD) and poor motor skills in general (see Hill, 2001, for a review), and some studies have looked at generalised motor difficulties in children with SSD compared to those with DLD. Bishop (2002) investigated motor ability in a sample of mono and dizygotic twin pairs aged 7-14 where at least one of the twins had DLD, and 171 TD controls. Seventeen of the twins had SSD alone, 70 had DLD alone, and 29 had SSD+DLD, 18 had resolved language difficulties and 13 were unaffected. Motor skill was assessed using a tapping task where the child was instructed to tap on a tally counter as fast as possible for 30 seconds. Results showed that the SSD, SSD+DLD and DLD alone groups had mean scores significantly lower than the control group, with a trend for SSD and SSD+DLD groups to do more poorly than the DLD alone group, but this was not significant. Unaffected twins and children whose SSD had resolved were not distinguishable from controls, suggesting the possibility of motor deficits as a factor in persistence of SSD.

Visscher *et al.* (2007) and Visscher *et al* (2010), in two large studies of 125 and 105 children aged 6-9 who attended special schools for speech and language disorders in the Netherlands, found that gross motor skills measured on the manual dexterity, ball skills, and balance subtests of the Movement Assessment Battery for Children (Movement-ABC), and locomotor and object control tests of the Test of Gross Motor Development 2, had poorer scores relative to age-matched, typically developing children, and that children with SSD alone or with SSD+DLD performed significantly worse than those with DLD alone. However, these children had quite severe impairments, with at least some of those in the SSD group appearing to have CAS,

which is known to be associated with motor difficulties. Nevertheless, only 51% of the speech and/or language impaired children in the first study had borderline or definite motor problems according to norms on the test, indicating that only a subset of children have gross motor difficulties. The age of the children is also a critical factor. Both studies comprised only children over the age of 6, by which time many children have normalised their SSD (Shriberg *et al.* 1994b). This may then indicate that motor difficulties are a marker of persistent rather than transient SSD. Wren *et al.* (2012) found that the group of children who were identified as having persisting SSD in the ALSPAC study of a population sample of 7390 children aged 8, performed more poorly on tests of diadochokinetic rate than those identified as having age appropriate or near age-appropriate speech, also suggesting there may be a motoric element to SSD that persist. However their sample did not exclude children with SSD that was not functional in nature, i.e. secondary to some other condition, or due to dysarthria or apraxia.

Instrumental studies have allowed further investigation of oral motor control for speech. EPG studies of older children with residual SSD frequently report motor difficulties, such as undifferentiated lingual gestures, as discussed in section 1.2.4 above, or otherwise aberrant lingual kinematics (Gibbon 1999, Goozée *et al.* 2007), but samples are small and there are none with pre-schoolers with functional SSD. However, Vick *et al.* (2014) tested 97 preschool children with SSD aged 3;0 to 4;11 on a battery of tasks including measurements of jaw and lip control in speech and non-speech chewing tasks, and measures of accuracy in non-word repetition and lexical stress tasks. They found 2 distinct sub-groups: 76% of the children had typical motor control, according to norms collected from a large sample of TD children in Vick *et al.* (2012), but 10% showed significantly higher variability in measures of articulatory movements, and poorer ability to imitate words with iambic stress, suggesting deficits in speech motor control; (the remaining children were not clear members of any group). Both groups had similar mean length of utterance (MLU). The kinematic variables were the feature that most distinguished the two groups, who were not distinguished by factors of gender or family history of communication disorder, suggesting an idiopathic rather than genetic basis. However, Vick et al. highlighted that these differences were subclinical and could not be identified using measures typically used in a clinic situation. One measure that they highlighted as being worth further development was the ability to imitate lexical stress, as the atypical motor control group differed from the larger group in this ability.

In contrast to Vick *et al.*'s (2014) finding that oro-motor deficits in children with SSD are subclinical, Broomfield and Dodd's (2004a) study of children aged 0-16 referred to SLT services in one area of the UK, found that 23.9% of those with SSD had oro-motor deficits, compared to around 40% of those with language deficits, although it is not clear how this was assessed, nor how many children had both speech and language deficits.

Namasivayam *et al.* (2013) studied 12 children aged three to 6 with moderate to profound SSD not due to CAS or dysarthria, and normal language. Nine of the 12 had oro-motor difficulties on at least one subtest of the Verbal Motor Production Assessment for Children (VMPAC) (Hayden 1999). They found that speech intelligibility at both the word and sentence-level was significantly correlated with speech motor control as assessed by the VMPAC, but not articulatory proficiency as assessed by the GFTA-2 (Goldman and Fristoe 2000). They also found that children with more severe speech motor control difficulties made less progress in intervention. This sample was small, and had been referred to the study because of suspicions of oro-motor difficulties. However it shows that motor control appears to be a factor in SSD of at least some children, and may affect their ability to benefit from intervention, suggesting that it may be a factor in persistence of SSD.

Findings where differences are significantly different from the norm, but small, might reflect the influence of only a small sub-set of children: Bradford and Dodd (1996), in a study of 51 children with SSD aged between 3 and 7, in tasks assessing volitional and non-volitional oral movements, fine motor skills and novel word learning, found that only the 9 children who made inconsistent speech errors, and the five with CAS, performed differently from the 51 matched controls. Children with CAS performed as well as controls on making single oral movements in context, suggesting that the ability of the articulators to move appropriately was not impaired, but their ability to imitate a sequence of two oral movements was significantly worse than the other children's. They also had difficulty on tasks assessing fine motor skills, which, Bradford and Dodd suggested, reflect "deficits at the levels of integrating sensory information into a plan of action, and at the level of co-ordinating speed and dexterity of intricate movements" (p.95). Children who made inconsistent speech errors performed as well as controls on tasks assessing volitional movement of the lips and tongue, and the ability to imitate a sequence of oral movements, but were significantly poorer than controls on fine motor

tasks which required speed and dexterity, leading Bradford and Dodd to conclude that they had more difficulty in organising complex sequences of movement under time constraints.

In a sample that included no children with inconsistent disorder, CAS, or neurological or cognitive deficits, very few children were found to have oro-motor difficulties. Dodd and McIntosh (2008) investigated oro-motor skills in 78 children aged 3 to five-and-a-half who made consistent speech errors, and 87 matched controls, using the standardized oro-motor screening task from the DEAP (Dodd *et al.* 2002). All children had normal language development. The children with SSD performed as well as controls on the isolated and sequenced oro-motor movement tasks, but did worse on the diadochokinetic rate (DDK) task. However, this was due to substitution of /k/ with [t] in the "patacake" task used to assess DDK, rather than any difficulty with fluency. No child in the control group performed outside the normal range in any task, but two children with SSD performed below normal limits on the sequenced movements task.

However, very different findings are reported in a more recent study. McLeod *et al.* (2017) reported that only 16.4% of children in their sample of 263 children aged 4 to five-and-half, identified by parents or teachers because of concerns about talking and making speech sounds, had typical oral functioning, with scores outside of normal limits on tasks such as lip retraction or rounding, DDK tasks and producing polysyllabic words. The assessment, the Oral and Speech Motor Control Protocol is more extensive than the DEAP oro-motor screener, testing movements of lips, tongue, diadochokinesis and prosody, and includes a multi-syllabic DDK task, which is vulnerable to the types of speech error made by children with SSD, which might account for some of the findings. In the Bradford and Dodd study, the same assessment was used to decide on group assignment, and 72% of children who made speech errors typical of younger children made two or more errors on the DDK task, although very few made errors on the other tasks. Children who made atypical errors, or whose word production was inconsistent, however, showed a higher rate of errors in all tasks than controls, again indicating that motor deficits may play a part in SSD for a sub-group of children. There is no breakdown of task results in McLeod et al. 2017, so it is not known how many children had deficits in tasks other than the DDK task.

Thus findings of the studies cited above show mixed outcomes regarding the number of children with SSD who show motor deficits, but it appears that at least some proportion presents with motor deficits that might contribute to SSD. The variability in findings may be due to sampling issues where motor speech disorders are more common in children with neurological, cognitive or language difficulties. Where these were excluded in Dodd and McIntosh (2008), very few children with motor deficits were found. Vick (2014) similarly excluded children with language difficulties or impairments to oral structure and found few children with identifiable difficulties, whereas the large ALSPAC study had no exclusions and found motor deficits to be a significant predictor. Age may also be a factor if motor deficits are only characteristic of SSD that persists: the studies of Visscher, Broomfield and Dodd, Bishop and Wren all included children over the age of six, whereas Vick's study had the youngest children, all under 5. However Namasivayam did find distinct motor deficits in a small group of young children without language impairment or CAS or dysarthria, indicating that some children even without additional problems do have motor impairments. The type of assessment may make a difference: some assessments may not be as sensitive to motor difficulties as others. McLeod et al.'s findings are difficult to explain, but may be due to most children failing only the multisyllabic syllable repetition sub-test, due to speech errors.

As motor deficits are frequently found in older children with persistent SSD, e.g. (Gibbon 1999, Goozée *et al.* 2007, Preston *et al.* 2016), presence of a motor deficit could predict persistence of the condition.

1.3.6 Interaction of deficits

Difficulties experienced by children with SSD may be more to do with the interaction of deficits at different levels rather than gross deficits in separate domains. As perceptual, cognitive and motor systems continue to develop throughout the period of speech acquisition (Nittrouer 2002b, Smith 2006), a delay or deficit in any one of these may affect development of the others. Velleman and Vihman (2002) suggest that lack of prelinguistic motoric practice may result in failure to make production/perception linkages early on. Evidence for covert contrasts discussed in section 1.2.4 above may indicate that, for some children, it is the gradual mastery of the acoustic-phonetic and articulatory cues that signal the phonological contrasts in the target structure, and mapping between them, which poses a central difficulty (Gibbon and Lee 2016). Byun

and Tessier (2016) suggest that speech motor development and phonological grammar are "intertwined", so that motor production experience plays a key role in deriving segment-level representations in the phonological structure. Finally domain general cognitive processes may operate at the interface of perceptual and cognitive processes used to derive the phonological system by statistical learning from the input.

It is a problem for clinical decision-making that, in much of the research reported above, results are often reported with means and standard deviations, but not the range of scores which would allow us to see how far the results apply to individual children, i.e. to see the overlap in abilities between TD children and children with SSD. However, where they are reported it shows that there is frequently an overlap, i.e. in individual areas, some children with SSD perform as well as TD peers, and some TD children perform as badly as children with SSD. This suggests that all functional SSD does not arise from a single cause, and reflects the heterogeneity of the population. It also underlines the importance of establishing those deficits that are pertinent to each child, so that distinct sub-groups can be identified in order to allow researchers to tailor specific intervention approaches to their deficits. The following section describes current models of sub-typing of children with functional SSD.

1.4 Classification of children with functional SSD

The studies presented in Section 1.3 above have failed to identify any factor, or combination of factors, which characterizes the group of children with functional SSD. Different children appear to have different deficits, or no identifiable deficit. Such findings have led to a growing consensus that children with functional SSD do not form a single, homogeneous group (Baker 2006, Bowen 2009, Dodd 2005, Grunwell 1981, Shriberg 1997, Stackhouse and Wells 1997). They differ in severity, type of speech errors, degree of involvement of other aspects of the linguistic system, risk factors, and course and responsiveness to treatment (Dodd 2005). This has motivated several groups of investigators to look for sub-groups within the group of children with functional SSD, but at present, no system of classification is universally agreed on. Waring and Knight (2013) have reviewed the most commonly used approaches to classification, and identified three systems as potentially useful for classifying children with SSD of unknown origin into homogeneous subgroups: Shriberg's Speech Disorders Classification System (Shriberg et al. 2010), based on aetiology of genetic and/or environmental factors, Dodd's Differential Diagnosis system (Dodd 2005), a descriptive

linguistic approach based on characteristics of surface speech patterns and Stackhouse and Wells (1997)'s Psycholinguistic Framework, based on specific points of breakdown in the speech processing system. The latter system does not claim to be a classification system, but could be used as such to group children by common points of breakdown.

All three systems suggest the same three broad groups among children with SSD of unknown origin: firstly, a group with peripheral speech errors that are articulatory in nature, i.e. mostly distortions of sibilants and rhotics; secondly a group with segmental and supra-segmental features of speech consistent with dysarthria and/or dyspraxia, and thirdly a group identified as having phonological delay/disorder, whose errors are primarily substitutions and omissions. This last group are by far the largest, representing over 80% of children referred to SLT services for SSD (Broomfield and Dodd 2004b), and showing a wide range of severity and outcome (Tyler 2011). The quest is now to break down this group into more homogenous sub-types in order to discover more accurate predictors of outcome, to better prioritise children for treatment and to provide intervention more specifically targeted at underlying deficits. It is in the discovery and description of sub-types in this group that classification systems differ. There is general agreement that children in the second group, with dysarthria and dyspraxia, have underlying deficits in motor abilities (ASHA 2007, Bernthal *et al.* 2009), so that only the first and third group, with articulation or phonological disorders, can be considered truly functional. It is with these types that this thesis is concerned. The next section looks at how these two groups are classified in current systems.

1.4.1 Phonetic versus phonological disorder

The first and generally accepted (Bowen 2014) sub-typing of children with functional SSD is into those with purely phonetic (articulatory) disorders, as opposed to those with phonological disorders. An articulation disorder has been defined as an impaired ability to produce specific speech sounds, with the child producing the same substitution or distortion irrespective of context, e.g. in words or isolation, spontaneously or in imitation (Dodd 2005b). Articulation disorders are considered to be peripheral disorders due to deficits in motor planning or execution (Holm *et al.* 1997, Gierut 1998, Dodd and Crosbie 2005). They generally affect a few specific phonemes, most usually /s/ and /ɹ/ and may be distortions, e.g. a lateral lisp where air is allowed to escape across the sides of the tongue rather than down the central groove, or they may involve the substitution

by another phone that also functions as a separate phoneme, e.g. substitution of /ɹ/ with [w], or /s/ with [θ].

In contrast, a phonological disorder is considered to be a language disorder that represents a difficulty with the organisation and representation of the phonological system at the cognitive-linguistic level (Grunwell 1981, Gierut 1998). The difficulty is not with the ability to produce speech sounds, but with the ability to use speech sounds appropriately to convey meaning, (e.g. where /s/ is substituted with [tʰ], both *sea* and *tea* are pronounced as [tʰi], thus neutralising the contrast between words). Phonologically based SSD is characterised by the occurrence of substitutions and omissions rather than distortions. When such difficulties are associated with limitations to phonological processing abilities, thus affecting children's ability to construct accurate phonological representations for words in the lexicon (Carroll and Snowling 2004), it puts children at greater risk of difficulties with phonological awareness and the acquisition of literacy than children with articulation disorder (Rvachew *et al.* 2007, Preston *et al.* 2013), whose difficulties lie at a more peripheral level.

However, children frequently present with distortion errors, suggesting articulation disorder, as well as patterns of phonemic substitutions and omissions that suggest a phonologically based SSD. In cases of substitution errors especially, it may be difficult to determine if an error has an articulatory or phonological basis. For example, in cases where /k/ is substituted with [t], it is normally considered a phonologically based error, as it changes the meaning of words. However substitutions of /ɹ/ with [w], or /s/ with [θ], are frequently considered to be articulation errors, although they also change a word's meaning, thus the classification can appear to be rather arbitrary. It also depends on the availability of the substitution as a phoneme in the child's language: in English substitution of /s/ by a voiceless lateral fricative could not be classified as a phonological error as the voiceless lateral fricative does not occur in English. However in a Welsh-speaking child, it presumably could be classified as a phonological error, as that language contains both phones functioning as phonemes (Ball *et al.* 2001).

The evidence discussed in section 1.2.4 also shows that, when measured instrumentally, phones used as substitutes may have different acoustic or articulatory characteristics when they are used correctly as phonemes, compared to when they are used as substitutions for a different phoneme (Gibbon and Lee 2016), so in these cases, it is, in

fact, a difference in articulation rather than a phoneme substitution that accounts for the error (Dodd and Crosbie 2005). Other studies have shown that perceptual deficits may also contribute to errors, in that some children with SSD have difficulty discriminating correct from incorrect versions of the sounds they cannot make correctly (Locke 1980, Rvachew and Jamieson 1989), even in cases which present like classic articulation errors, i.e. where errors on /ɹ/ sounds are the only error the child makes (Shuster 1998). It has been suggested that production errors may in fact influence the child's phonological system, such that incorrect productions of the target sound become "phonolologized" so that they are included inside the phoneme boundary and become acceptable target productions for the child (Ball *et al.* 2004, Byun 2012). Therefore, both deficits at the perceptual/cognitive level, and peripheral motor execution errors, may account for errors that are commonly classified as articulation errors.

Ball and Müller (2002) have pointed out that the distinction made between substitution ("phonological") and distortion ("phonetic") errors is problematic. As noted above, if an English-speaking child realizes /s/ or /z/ as a sound which operates as a different phoneme within the phonological system of English, e.g. [θ] or [ð] respectively, it is traditionally considered to be a substitution error, which is a phonological error because a contrast has been lost. However if the same sounds are realised as a phone outside the English phonological system, e.g. [ç] or [ʝ], then it is considered a distortion error because it does not cause a neutralisation of contrast. This is despite the fact that the differences in the sound produced result from somewhat similar gestures, i.e. loss of grooving of the tongue in both cases, and movement of the place of constriction forward in the first case, and backward in the second. In another case, if the realisation remains within the phonological system, but an allophonic error results, e.g. realisation of /s/ as dental [s̪] in all positions, this is also classified as a distortion, because contrastivity is maintained. As Ball and Müller point out, this division is made despite the fact that the articulatory difficulty acknowledged in the production of sibilant sounds is likely to be the common cause of all the above-described errors. When regarding the classification of substitution errors as phonological errors, evidence regarding covert contrasts shows that, in some cases, what is perceived by the listener as a sound substitution resulting in a loss of contrast e.g. between velar and alveolar plosives, may in fact be two separate gestures in production, resulting from quite different phonetic sources. Furthermore, where two or more phonemes are realised as the same sound from outside the target phonological system, e.g. both /s/ and /ʃ/ realised as [ç], we have what would be

classified as distortion errors actually resulting in a loss of contrastivity: a phonological problem. To deal with the problems resulting from this simple binary classification, Ball and Müller (2002) suggest a much more comprehensive classification system that incorporates firstly the source of the error, i.e. whether phonetic (articulatory) or phonemic (organisational), secondly, the effect of the error, again whether the effect on the listener is of a phonetic or phonemic nature, i.e. whether it affects contrastivity, and thirdly, whether the realisation of the sound is within or without the target phonological system. In the resulting classification, articulation errors are subsumed under the phonetic source, (Ball 2016), and may result in either allophonic or phonemic effects for the listener, further sub-classified as to whether the realisations are internal or external to the target phonological system. Thus this proposed typology allows a much richer description and much finer classification of speech errors, with resultant implications for treatment. The caveat is that it is not always easy to assign some speech errors to a source category: although wider availability of instrumental measures may help with discrimination of phonetic from phonological source errors, Ball and Müller (2002) point out that it may still be difficult to distinguish immediate from original source in certain cases, where, for example, speech errors originally resulting from motor or structural constraints (e.g. in cleft palate, subsequently repaired), become incorporated into the individual's phonological system: see Wright *et al.'s* (2017) case study of an adolescent with a severe persistent SSD for a case in point.

Thus, although the division between articulation and phonological SSD is widely accepted in the speech therapy literature, it may not be entirely clear cut, and difficult to establish in individual presentations. In fact, researchers such as Gibbon (2007) and Richtsmeier (2010) have argued that phonetic difficulties might underlie many of the surface patterns that are currently classified as resulting from phonological deficits.

Reflecting the problematic terminology of distortion versus substitution errors, the term distortion is frequently used in the literature regarding SSD without a clear definition. Ruscello (2003) describes distortions as "sound variations not within the phonetic boundaries of the intended target sound" p. 157. Shriberg (1993) gives perhaps the most comprehensive description, listing common distortion errors as dentalisation of voiced or voiceless sibilant fricatives or affricates, derhotacized /ɹ/, /ɚ/ or /ɝ/, lateralization of voiced or voiceless sibilant fricatives or affricates, and velarized or labialized /l/ or /ɹ/. Wren (2016) also refers to Shriberg's (1993) paper, but Wren (2012) defines children

classified as having common clinical distortions as those whose errors were "exclusively misarticulations of /ɹ / and/or /s/, consistent with Shriberg's [1993] description of common clinical distortions" p.3. Dodd does not define distortion, but uses it in her description of an articulation disorder, as:

> the inability to produce a perceptually acceptable version of particular phones, either in isolation or in any phonetic context. In most cases of 'functional' articulation disorder, the most commonly affected phones are /s/, /r/, /θ/ and /ð . Children may consistently:
>
> - produce a specific distortion (e.g. a lateral lisp where the aistream in the production of /s/ escapes bilaterally rather than centrally);
> - substitute another phone (e.g. [f,v] for /θ,ð/, [w] for /r/). Note however, that the first substitution pattern may be dialectal and that there is some evidence that, when measured instrumentally, the [w] for /r/ has different acoustic characteristics from [w] for /w/, emphasizing the phonetic nature of the disorder. (Dodd 2005c, p.150).

In this text, I will attempt to provide the definition of distortion used by the original author. Where common clinical distortions are referred to without a definition, I intend Shriberg's 1993 definition. Where articulation disorder is used, if not otherwise defined, Dodd's description above is intended.

1.4.2 Shriberg's Speech Disorders Classification System

Shriberg and colleagues have sought to classify children with SSD by aetiology, which they argue offers the possibility of a better understanding of the nature of the problem as well as the ability to target treatment approaches directly at the cause of the problem. They have put forward five potential sub-groups based on aetiological factors, with 2 of these having further sub-groupings (Shriberg *et al.* 2005).

The first sub-group, termed Speech Delay-Genetic (SD-GEN) comprises children who have a family history of one or more relatives who also had speech and language difficulties. Shriberg et al. propose a distal genetic cause in these children, which presents as a proximal cause in a cognitive/linguistic deficit, resulting in a large number of omission errors as they have difficulty identifying the phonological structure of the

language. Lewis *et al.* (2007) report findings supporting the genetic cognitive/linguistic deficit theory: in a study of parents of children with SSD, they found that the parents who themselves had a history of therapy for SSD performed significantly more poorly on standardised measures of speech and language than parents with no history did. Felsenfeld *et al.* (1995) investigated this in the other direction: they found in a sample of 24 adults with a history of moderate phonological and language disorder, and 28 TD controls, that the children of parents with a history of SSD had significantly worse scores on tests of articulation and expressive language functioning, and were significantly more likely to have received articulation treatment, than children of adults with no history. Such evidence of familial aggregation of speech and language disorders lends support to a genetic foundation for at least one sub-group of functional disorders, but may not be predictive of disorder in an individual child, as at least as many or more children with functional disorders have no reported family history, while other children with a family history do not show speech or language deficits (Campbell *et al.* 2003, Wren *et al.* 2016). In addition, it might be expected that SSD with a genetic foundation should be a qualitatively different disorder to SDD without such foundation, but there appears to be no evidence at present to support this (Lewis and Freebairn 1997).

Shriberg's second sub-group, termed Speech-Delay-Otitis Media with Effusion (SD-OME) comprises children with histories of frequent episodes of OME who, he proposes, have difficulty extracting relevant phonemic details from the ambient language due to auditory/perceptual difficulties, in this case due to intermittent hearing loss (Shriberg *et al.* 2000). Shriberg and colleagues found that children in this sub-group presented with poorer intelligibility in connected speech than would be predicted by the accuracy of consonant production in words (Shriberg *et al.* 2003a), and a high number of backing errors (Shriberg *et al.* 2003b), but this finding has not been further investigated by other researchers, and there is conflicting evidence regarding the overall effect of OME in the aetiology of SSD, as discussed in sections 1.5.6.4 and 3.2.4 .

The third group, Speech Delay-Developmental Psychosocial Involvement (SD-DPI) is reported to comprise children who have temperaments that impede their ability to obtain the feedback necessary to develop and normalise their speech skills. They are described as aggressive, angry or control seeking, or socially withdrawn, anxious or taciturn (Hauner *et al.* 2005), as well as easily frustrated by difficulties in communication and less likely to persist in communicative attempts. They have also been described as

having more severe involvement than the other groups, which suggests the possibility that the causal link could equally well be in the other direction, i.e. they are angry, frustrated or withdrawn because they are unable to communicate. A longitudinal study that documented measures of temperament before and during speech development would be necessary to test this hypothesis. Hauner *et al.* suggested that these behaviours, especially when combined with lack of sustained attention or task persistence might increase severity or persistence of an SSD, due to difficulties in feedback and correction, highlighting the importance of the skill set of SLTs in dealing with children who are reluctant to engage.

The fourth group, Motor Speech Disorders, have deficits in speech motor control. This group comprises sub-groups of apraxia and dysarthria, as well as a category added later (Shriberg *et al.* 2010), Motor Speech Disorder-Not Otherwise Specified (MSD-NOS), to cover children with symptoms consistent with motor speech impairment (e.g., slow rate, imprecise consonants), but which do not fit criteria for apraxia or dysarthria.

The last group, Speech Errors (SE) includes children whose speech errors consist of age-inappropriate distortions limited to specific sounds or feature classes, usually distortions of sibilants and derhotacizing of target rhotics. It has two sub-groups, termed Speech Errors-Sibilants and Speech Errors-Rhotics, and it is suggested that they adopt a production pattern for the sounds when very young which reflects an immature motor ability, and retain this immature pattern despite subsequent maturation of production abilities.

These groups are not mutually exclusive, as children could conceivably present with factors that include them in more than one group. Conversely, the classification is not exhaustive, as children with SSD frequently present without signs of any of these aetiologies (Fox *et al.* 2002), and so cannot be classified in this system. This reduces its utility in identifying sub-groups, and it is not clear how genetic sub-typing would contribute to the development of specific targeted interventions (Dodd 2005, Tyler 2011). A difficulty in looking for causes of SSD is that the original, distal cause of the problem may have led to other problems, which are now maintaining the disorder, while the distal cause may have since disappeared. For example, perceptual limitations may have led to the establishment of faulty phonological representations for early-acquired words. The inaccurate representation results then in inaccurate output for those words,

despite the maturation of perceptual abilities, which now appear to be normal (Bernthal *et al.* 2009). In addition, differential long-term outcomes have not been reported for the sub-types described, which also limits their utility (Lewis *et al.* 2006a).

1.4.3 The Psycholinguistic Framework

Another approach that could be used to differentiate children with functional SSD is to look for current underlying deficits that could be affecting how they process speech. Stackhouse and Wells (1997) argue that sub-groups based on aetiology do not necessarily suggest specific treatment approaches, whereas focusing on current deficits does.

The Psycholinguistic Framework (Stackhouse and Wells 1997*)* proposes a speech processing chain that contains a lexical representation with phonological, semantic, and motor representations at the centre, with input arriving through stages of peripheral auditory processing, discrimination of speech versus non-speech sounds, recognition of phonological forms specific to the language in question, and the phonetic discrimination of speech sounds. The output processes then proceed through motor programming, motor planning, and motor execution. Stackhouse and Wells propose that children with SSD present with breakdowns at one or more of these points in the speech processing chain, and that targeting the level(s) of breakdown will improve speech output. This allows a profile of strengths and weaknesses to be developed for each child, which suggests a specific programme of remediation.

Waring and Knight (2013) suggest that further research using the Psycholinguistic Framework may help identify patterns of strengths and weakness in children that would allow classification into more homogenous sub-groups for which one could design optimal interventions, and predict the course of development and response to therapy. A possible candidate for this could be deficits in speech perception, as defined in section 1.3.2. It is known that some, but not all, children with SSD have deficits in speech perception (Rvachew and Jamieson 1989, Edwards *et al.* 2002). Rvachew (2006) found that speech perception skills in 47 pre-schoolers at age 4 predicted unique variance in scores of speech accuracy one year later, suggesting that good speech perception skills predict better speech outcomes. Rvachew (2007), in a study of 33 pre-schoolers followed up 2 years later, found that preschool measures of phonological processing, as measured by a composite of scores on tests of speech perception, rime awareness, and

onset awareness, differentiated two distinct subgroups of children with SSD, one whose scores were indistinguishable from the control group, and one whose scores were significantly worse, despite both groups having equally severe SSD. These sub-groups were validated by the finding that scores on this composite predicted significant variance in non-word decoding 2 years later, suggesting that speech perception abilities also have a role in the acquisition of literacy. More such evidence would establish the predictive power of the Framework.

Currently, however, the individual profiling limits the Psycholinguistic Framework's utility in delineating homogenous sub-groups. Furthermore, it has been criticised for only considering deficits at the peripheral input and output levels, whereas deficits at a more central level, e.g. in executive function, could affect the operation of the speech processing chain (Dodd *et al.* 2005). As well as this, a deficit at one level of the speech processing chain could affect operation of processes further down the speech processing chain, so that a deficit in output occurs as a consequence of a deficit in input processing, rather than in output processing itself. This framework also requires extensive assessment, (Stackhouse *et al.* 2007), which may limit its feasibility for use in practice.

1.4.4 Dodd's Differential Diagnosis system

Dodd's Differential Diagnosis system is also based in psycholinguistic theory, but assigns children to specific sub groups based on surface symptoms, which she suggests reflect underlying cognitive linguistic deficits at a more central level (Dodd *et al.* 2005). Children with functional SSD are divided into 4 sub-groups. The first sub-group comprises those with articulation disorder, as defined above, and the second those with phonological delay (PD). Children with PD are those who use only error patterns that occur during typical speech development at a younger chronological age. This is the largest sub-group, and in experiments reported in Dodd (2005), they appear to have no specific cognitive deficit, performing similarly to, although slightly behind, typically developing (TD) controls on tests of vocabulary knowledge, knowledge of phonological legality, phonological awareness, motor accuracy in tracing, nonsense word learning, and phonetic variability in phoneme production. The third sub-group are children who use one or more error patterns that they claim are atypical of normal development in English (e.g. initial consonant deletion, backing, favoured sound). They are classified as having consistent atypical phonological disorder (CPD). As well as making atypical errors, they frequently also make more errors than children classified as delayed. These

children also performed similarly to controls on all tests listed above, except those assessing knowledge of phonological legality (Dodd *et al.* 1989) and phonological awareness (Harris *et al.* 2011). In experiments reported in Dodd (2011), children with CPD also performed less well than a group with PD on measures of cognitive flexibility, and rule abstraction, leading Dodd and colleagues to conclude that children in this sub-group have an impaired ability to derive rules, to exhibit cognitive flexibility in rule abstraction and implementation, and thus to organise knowledge about the nature of the phonological system (Crosbie *et al.* 2009). The fourth small sub-group are children classified as having Inconsistent Phonological Disorder (IPD). These are children who produce multiple error types (e.g. zebra as [dzeuwa], [jeiua] [jedwa]) for the same lexical item (Holm *et al.* 2007), with unpredictable variation between a relatively large number of phones, such that no consistent pattern of substitutions or omissions can be seen in any context. Despite producing multiple and varying error forms, they do not fit criteria for a diagnosis of childhood apraxia of speech (Dodd *et al.* 2006), having no deficits in prosody or oro-motor abilities, no observations of groping, and greater accuracy on imitation than in spontaneous speech. It is suggested that these children either have a deficit in phonological assembly, i.e. storing and/or retrieving a phonological output plan (Holm *et al.* 2008), or under-specified plans for word production that lead to phonetic programmes with articulatory parameters that are too broad (Dodd & McCormack 1995). Bradford and Dodd (1996) found that these children have a more general motor planning difficulty, identifying difficulties in planning sequences of motor output for speech, as well as for sequences of fine motor behaviour not involved in speech. Dodd *et al.* (2005) reported that children with inconsistent disorder performed more poorly on tests of vocabulary than other groups, but they were found to perform similarly to controls on recognition of phonological legality and phonological awareness. Macrae *et al.* (2014) also suggested an alternative explanation, namely that high word variability reflects unstable lexical representations, which may be due to small vocabularies and which may be the cause of problems with phonological assembly. They predict that the representations will become more stable as children learn more words.

The evidence base for Dodd's Differential classification approach is extensive, showing that the classification can assign all children with SSD to discrete sub-groups, and can be applied cross-linguistically, (Hua and Dodd 2000, Fox and Dodd 2001). Assessments have been developed for this system, making it feasible for clinical and research use,

which has led to wide scale adoption in English speaking countries (Joffe and Pring 2008, Dodd 2014).

Additional support for these sub-groups has been found in the fact that they show differential outcomes regarding phonological awareness and acquisition of literacy. Children with a history of atypical errors, like those in the CPD group, have been found, in a number of studies, to have poorer phonological awareness and be at increased risk of literacy difficulties, compared to those with typical errors or distortions (Leitão and Fletcher 2004, Preston *et al.* 2013). However those who make inconsistent errors appear to be more likely to have difficulty with spelling than reading (Holm *et al.* 2008).

As proposed by stratified medicine, the identification of these different sub-groups have allowed Dodd and colleagues to propose different types of intervention to target the supposed underlying deficits of each group. Core vocabulary therapy has been shown to be more effective than phonological contrast therapy for children with IPD, (Crosbie *et al.* 2005), but both the PD and CPD group respond to the latter (Dodd *et al.* 2005). More recently, (Holm and Dodd 2010) introduced a novel intervention for children with CPD, targeting underlying cognitive difficulties rather than surface speech symptoms. However, this intervention was not shown to be more effective than phonological contrast therapy alone, and phonological contrast therapy was administered alongside it, so that there is no evidence of its efficacy when used alone.

As cited above, children classified into sub-groups using this system have been shown to have differing profiles of performance on tasks set to examine underlying deficits. However, it relies on the tests used actually being valid measures of the underlying processes assumed to be the source of the sub-group's difficulties (Tyler 2010). As Waring and Knight (2013) point out, none of the supporting studies investigated all the purported deficits in the same group of children, so no children were assessed on all the profiling tasks. Neither have there been any studies following a cohort classified this way in the long term, to see whether the sub-groups remain stable and are predictive of outcome in terms of SSD.

Children with SSD also show different profiles of strengths and weaknesses in other abilities not fully considered by the classification systems described above, suggesting other ways be which they may be sub-typed.

1.4.5 Classification by language ability

Another way that children with SSD can be differentiated is by presence of a co-morbid language impairment (LI). Prevalence of comorbid LI in children with SSD has been reported to be from 40–60% (Shriberg and Kwiatkowski 1994, Broomfield and Dodd 2004b, Eadie *et al.* 2015). However, there are issues in the measurement of language in children with SSD that can affect language measures, especially where the SSD is severe. For example, English morphology is largely represented by fricatives and clusters at the ends of words, both of which are frequently substituted, reduced or omitted by children with SSD, meaning that they may do poorly on tests of morphology, or mean length of utterance in morphemes, due to the speech difficulty alone (Pharr *et al.* 2000). For example, the widely used CELF series of language assessments (Wiig *et al.* 2000, Semel *et al.* 2004) include two sub-tests (word structure and sentence recall) which require accurate production of morphology to achieve maximum scores (Leitão *et al.* 1997). Furthermore, children with much reduced intelligibility often use strategies to enhance communicative function, which affect measures of expressive language, i.e. decreasing length and complexity of utterances, or using generic rather than specific vocabulary (Dodd 1995a). Some studies have shown that children with SSD+LI have more severe SSD than children with SSD alone (Sices *et al.* 2007), but the fact that children with severe SSD score badly on some tests of language impairment for reasons due to speech rather than language impairment might be at the root of this finding, in some cases at least. When receptive language impairment, which is less vulnerable to speech deficits in assessment, is considered, the rate of comorbidity drops: in a review of the literature, Shriberg and Austin (1998) found that up to 60% of children with SSD have a language disorder overall, but only 20% had a receptive difficulty. Similarly, Broomfield and Dodd (2004b) found that between 34% to 67% of children with SSD had expressive deficits, depending on sub-group, but only 23 to 40% had receptive deficits. This may, however, merely reflect the lower incidence of receptive deficits in the population as a whole.

Nevertheless, many children with SSD are not classified as language impaired, and several studies demonstrate better outcomes in both speech and literacy acquisition for children with isolated SSD compared to those with combined SSD+LI (Bishop and Adams 1990, Lewis *et al.* 2000, Nathan *et al.* 2004a, Sices *et al.* 2007). Children who have language as well as speech difficulties have been shown to go on to experience greater academic difficulties compared to those who have SSD alone (Lewis *et al.*

2000a, Nathan *et al.* 2004b). However, presence of a concomitant language disorder does not appear always to predict worse outcomes for speech normalisation per se. (Shriberg *et al.* 1994b) in a study of 54 4 year-olds with SSD followed up after one year, found that language status was not a predictor of normalisation of SSD. Lewis *et al.* (2000b), in a 2 year follow up of children presenting with SSD or SSD+LI at ages 4-6, found no difference in the persistence of SSD between the two groups. However, there was a difference for literacy. The SSD alone group scored within the normal range on assessments of reading and spelling, but the SSD+LI group scores were more than 1 SD below the norm for age. As previously cited in section 1.3.4, several studies have shown that children with isolated SSD go on to have normal literacy achievement. In a longitudinal study by Bishop and Adams (1990) of 83 children with SSD and/or LI at age 4, the 12 with isolated SSD had normal reading when followed up at age 8;6, whereas those with isolated or co-morbid language impairments demonstrated much worse performance in reading than controls. Hesketh (2004), in a study of 35 children diagnosed with moderate to severe SSD but normal language between the ages of 3;6 and 5;0, found that almost all the children had gone on to show both PA and literacy development within the normal range by age 6;6 to 7;6. Only four children had scores more than 1SD below the mean, and only one score on one sub-test fell more than 2SD below the mean, which she suggests reflects the wide range of performance among even TD children at early stages of literacy achievement. In Sices *et al.'s* (2007) study of 125 children aged 3 to 6 years with moderate to severe SSD, 53% had comorbid language impairment (defined as a score >1.5 SD below the mean on the Test of Language Development: Primary, 2nd Edition (TOLD-P2; Newcomer and Hammill 1988) or Clinical Evaluation of Language Fundamentals-Preschool-2 (CELF-P; Wiig *et al.* 1992). They found that language skills, but not articulation skills, were associated with development of early literacy skills. However, literacy was measured by the Test of Early Reading Ability-2 (TERA; Reid *et al.* 1989) and the Test of Early Written Language-2 (TEWL; Hresko *et al.* 1996), which do not assess decoding and spelling, which are proposed to be affected by a core phonological deficit shared by children with phonological and literacy disorders, but rather knowledge of print concepts and reading and writing readiness. Peterson *et al.* (2009) investigated literacy outcomes at age 7 to 9 in 86 children with a history of SSD at preschool age, and 37 TD controls. The children with a history of SSD were classified as having either persistent or resolved SSD at age 5 to 6. They found that persistent SSD was associated with PA impairment, but literacy outcome was predicted by language skill rather than SSD or PA. As a whole group,

22% of children with SSD were classified as having reading disability (RD), compared to 5.4% of controls, but when split into sub-groups defined by presence of LI, only 10.3% of those with a history of SSD alone were in the RD group, as opposed to 66.7% of children with SSD+LI. Peterson et al. concluded that a large part of the increased odds for RD seen in children with SSD was due to the additional factor of LI. This finding is better explained by a multiple deficit account than the core phonological deficit account, which holds that children with a history of SSD are at greater risk of RD because both conditions involve a deficit in phonological processing. Instead, the multiple deficit account holds that pre-schoolers with SSD have a phonological deficit putting them at risk for RD, but that outcomes for literacy are determined by how this deficit interacts with other risk and protective factors, with poor language skills increasing the risk of RD further.

Differences in results might occur due to the way speech and language impairment is measured. Many studies use only single word samples, whereas others use connected speech or extensive testing of multisyllabic words. It would also be affected by the size and make-up of the sample: children whose speech deficits are articulatory rather than phonological in nature are at less risk of literacy difficulties, as their deficits are not supposed to be phonological in nature. The studies reviewed here comprised children aged from 3 to 6, which is the age of interest in determining whether SSD may be persistent or transient, as after this age it is slower to resolve spontaneously (Shriberg *et al.* 1999). The large effect of co-morbid LI on literacy outcomes shown by the studies cited above, suggests it may be clinically useful to sub-type children with SSD according to presence or absence of a co-morbid language impairment (Lewis et al. 2006a), to aid in prioritisation and focus of therapy. It may be that children with SSD+LI need especial monitoring for PA and literacy difficulties.

Thus the different classification systems have proposed sub-types of children based on aetiological factors, deficits at different stages of the speech processing chain, type of errors and co-morbidity with language difficulties. No studies, however, have investigated all of these factors in the same group of children. In addition, they have all left a large group who are difficult to distinguish from TD children on any obvious aetiological, motoric, linguistic or organic basis: for Dodd the PD group, for Shriberg the group called SD-delay who do not have a family history, history of OTE or any psychosocial involvement, for Stackhouse and Wells those children who are not

particularly weak at any task. Despite this, the children in these groups are differentiated by course and outcome. The following section investigates the literature to show whether we can predict which children are likely to resolve their SSD.

1.5 Predictors of persisting difficulties in pre-schoolers with SSD

Once subgroups with distinct profiles have been identified, they can be validated by longitudinal research to see whether the different profiles make a difference to the course of development and outcomes. The reported extent of negative academic, social and psychosocial consequences of SSD outlined in Section 1.1 highlights the importance of remediating these conditions before such consequences can take effect. However, a number of studies have shown that speech and language difficulties in pre-schoolers can be transitory. Shriberg *et al.* (1994b) suggest that 75% of pre-schoolers with SSD will have attained age appropriate speech by age 6.

It is widely assumed that the earlier intervention is provided, the more effective it is (Paul 2000, Harrison and McLeod 2010); however, there are few studies to provide evidence about this in the SSD population. It seems that early intervention is often confounded with intervention at a young age, rather than soon after the identification of a problem. Intervention may not always be necessary for young children. Roulstone *et al.* (2008) showed that almost 40% of children aged less than 3;6, and referred to therapy at initial assessment, had resolved their difficulties by follow up one year later, despite receiving no therapy.

Broomfield & Dodd (2005a) showed that intervention for different conditions was actually most effective at different ages, e.g. articulation disorder is best treated after the age of seven, and therapy for children with phonological delay was most effective at the age of five. As children with SSD are usually referred at around the age of two or three, we may be giving intervention too soon. Given the pressure on SLT services, where many older children are not able to access intervention for proven needs, it is inefficient to give intervention to children too early, or to those whose problems are likely to resolve spontaneously, as well as unethical to subject children to treatment they may not need. Thus clinicians need information that will help them decide, for the individual child, whether they are likely to resolve spontaneously, or whether they are a priority for intervention.

The beginning of formal literacy education seems to be a critical age for resolution: Bishop and Adams (1990) found that literacy development proceeded normally if speech and language problems had resolved by 5;6, whereas those with speech difficulties after that age had difficulties in learning to read. This effect of persisting SSD was supported by a later study of 39 children with SSD at age 4 who were followed up at age 6;8 (Nathan *et al.* 2004b). They found that children with persisting speech problems scored less well than TD controls in tests of reading, reading comprehension, spelling and maths, whilst those whose SSD had resolved were only distinguished from controls in the case of spelling. Thus, finding predictors which identify children who go on to present with persistent rather than transient SSD would inform SLTs on how to prioritise their services to target those most in need. In addition, it is important to know at what age those features actually become predictive.

A number of researchers have looked for possible predictors of normalisation/ persistence in children with SSD. Evidence for possible predictors of outcome comes both from longitudinal studies which track children with SSD over time, to establish specific profiles which differentiate children whose SSD persist from those who resolve, as well as from cross-sectional studies which identify specific deficits in older children and adults whose SSD persist into and past school age. The following section looks at features that have been suggested as possible risk factors and predictors, and synthesises findings from longitudinal and cross-sectional studies.

1.5.1 Prevalence and Rates of resolution

Prevalence studies at different ages could provide evidence about the proportion of affected children whose SSD persists into school age. For example, Campbell *et al*. (2003) found the prevalence of SSD to be 15.6% in a sample of 3 year olds, whereas (Shriberg *et al.* 1999) found this had dropped to only 3.8% in 6 year olds. McKinnon *et al.* (2007) in a study of 10,425 Australian children in primary school, found that prevalence of SSD reduced with grade level from 3.04% in Kindergarten (around 5 years old) to 0.37% at age 12. Keating *et al.* (2001), using Australian Health Survey data for 12,174 children without developmental delay, found a gradually declining prevalence from 1.8% at age 3 to 0.4% in 10 to 14 year olds for girls, but in boys the prevalence rate increased from 4.2% in 3 year olds to 6.6% at age 5, possibly reflecting a slower course of development in boys, followed by a steep drop at age 6 to 2.6%, going down to 1.0% in 10-14 year olds. Recent figures from Eadie *et al.* (2015), also in

Australia, were in line with this, finding a prevalence rate of 3.4% in a cohort of 1494 4-year-old children. However, in results from a large population study, the ALSPAC study of 7390 children in the UK, Wren *et al.* (2013) found a much higher prevalence rate of 13.4% in 8 year olds. These large differences in prevalence may be due to methodological differences. For instance, most samples excluded children who had a known aetiology for the SSD (e.g. cleft palate, cerebral palsy, Down Syndrome), whereas Wren's sample included all children who presented. The way SSD was defined would also affect findings: Wren *et al*.'s sample used a relatively broad criteria of -1.2 SD from norms collected from 47 TD children in the sample and they stated that it was possible that not all those they defined as having speech difficulties would reach criteria for a clinical diagnosis of speech sound disorder. In fact, when Wren *et al.* (2012) further analysed the sample, they found that 141 children did not reach criteria either for Persistent Speech Disorder (PSD), or the other category of SSD defined, namely common clinical distortions (CCD). Children with CCD were defined in this paper as "Those children whose errors were exclusively misarticulations of /ɹ / and/or /s/, consistent with Shriberg's [1993] description of common clinical distortions" (Wren 2012 p. 3). Once the 141 children were excepted, the prevalence figure for children with PSD and/or CCD fell to 11.5%, although this is still much higher than other prevalence studies. Age at assessment might also play a part, in that minor distortions such as dentalised fricatives or gliding of /ɹ/ may be accepted as appropriate productions for target phonemes at preschool and early school age, but at older ages fully adult like speech is expected, so the same production persisting at age 8 or 9 becomes a speech error. This could result in a seeming increase in the prevalence of SSD, without any increase in the number of people with speech errors. Other differences would stem from the way data was collected, i.e. by parent or teacher survey or by direct assessment, whether a continuous speech sample was used or a single word citation sample, and in how SSD was defined. Apart from cut-off levels, variations in the type of error classified as a characteristic of SSD might also vary: some samples might not classify small distortions as errors. Nevertheless, the majority of large studies do show prevalence declining as age increases, supporting the contention that a large number of SSD resolve over time. However, only longitudinal studies can provide accurate information about resolution rates, and persistence of SSD for individual children.

1.5.2 Risk factors for SSD

Factors that put children at risk of SSD could also prove to be maintaining factors, leading to persistent SSD. Identification of risk factors for SSD in older children and comparison with risk factors in younger children could suggest which factors are linked to persistent as opposed to transient SSD. Risk factors for speech and language delay have been identified from correlational or group mean comparison studies, but most studies have not differentiated risks for children with SSD as opposed to language delay. Those studies that have focused on risk factors for children with SSD, have most frequently mentioned male gender, socio-economic status, family history of speech and language disorders and history of persistent otitis media.

Campbell *et al.s* (2003) population study of 639 apparently healthy 3 year olds, excluded children who had experienced adverse perinatal circumstances, had serious illness, lived in a home where a language other than English was spoken, or if they had a mother who was under 18, seriously ill, a substance abuser or dead. Children were part of a sample recruited for a larger trial (Paradise *et al.* 2000), from eight primary care practices in a socio-demographically diverse area of Pennsylvania. One hundred children were classified as having an SSD, using the criteria for the Speech Disorders Classification System (Shriberg and Austin 1997), and data from transcriptions of speech samples. Logistic regression was used to calculate odds ratios (OR) for 7 risk variables: gender, maternal education, history of communication disorder in a first degree relative, health insurance status, race, cumulative duration of otitis media and hearing status. They found that low maternal education, male gender, a family history of communication disorder and Medicaid health insurance all increased the risk for having SSD, but only low maternal education (OR 2.58) and being male (OR2.19) increased the odds ratio above two, which has been suggested as a threshold representing a clinically significant level of risk (Sackett et al. 1991). However, when children had the three most significant factors: male, positive family history and low maternal education, they were almost 8 times more likely to have SSD, showing the importance of considering the aggregation of factors over one single factor. For example, considering family history alone, only 36% of children with SSD in their sample had a positive family history of speech and language difficulties, compared to 25% of children with typically developing speech, which is not a large enough effect to aid in decision making for prioritisation for the individual child.

Campbell et al. did not consider co-morbid language impairment and excluded children with pre- and perinatal complications from their analysis. Fox *et al.* (2002), in a smaller study of all children, regardless of birth history, referred for speech and language assessment to one therapist in an 8 month period, compared 65 children, aged 2 to 7, with SSD and normal language development, to 48 TD controls. Logistic regression was also used to calculate OR for the risk factors of pre- and peri-natal problems, sucking habits, otitis media and other hearing problems, and family history. A higher OR of having SSD was found for positive family history (OR 4.38) than in Campbell's sample, which may be explained by the fact that this was a clinical rather than population sample: parents with a history of difficulties themselves may be more likely to refer their children sooner. Perinatal and birth risk factors (OR 8.86), and bottle or thumb usage (OR 2.42), were significant predictors, but not hearing variables, age or gender. However, no measure of SES were included, demonstrating that consideration of different factors can lead to different conclusions.

Delgado et al. (2005) included both birth history and SES in their study of 959,148 pre-schoolers born in Florida between January 1, 1994 and December 1998. Their average age was 3;10. Information was taken from entries in the Florida Department of Health and linked to preschool exceptionality records from the Children's Registry and Information System (CHRIS), which tracks all children in Florida who receive services through the federal Individuals with Disabilities Education Act (IDEA). This Act provides for special education and related services for children with disabilities. The CHRIS database included 6,835 children with primary speech impairment, (SI), 2,357 with primary language impairment (SLI), and 3,607 with comorbid SI+SLI (CSLI). The 946,177 children without speech or language impairment acted as the comparison group. The prevalence rate of 13% for speech and/or language impairment, is therefore substantially lower than the 15.6% reported in Campbell's study for SSD alone, which is possibly due to the fact that not all children who present with speech and language difficulties fulfil selection criteria for IDEA Section B. Delgado also included in the SI group children with fluency or voice difficulties, so it is not restricted to SSD alone, although the former disorders are substantially less prevalent than SSD.

In contrast to Campbell *et al.,* Delgado *et al.* found the likelihood of a child's being classified with SI increased with higher rather than lower maternal educational attainment. In Delgado's sample of children with CSLI, results for maternal education

less than 12 years (RR = 0.84), maternal age younger than 18 years (RR = 0.74), and unwed maternal marital status (RR = 0.81) were in the opposite direction to that expected, indicating a decrease in risk for CSLI associated with these factors, indicating again that low SES did not increase risk, unlike Campbell *et al.'s* sample. They posited that the finding may indicate a bias toward the identification of SI in more educated families with higher socioeconomic status, as risk was also increased for children of mothers who were married and/or older than 18. It is possible that, as the children in this sample had to qualify for state services, higher SES parents may be better able to advocate for their children, and so secure this qualification. In contrast to Fox *et al*, they found that although adverse birth history was a significant risk factor for LI, it was not for SI alone. For SI, they found that presence of a congenital abnormality and multiple birth were the factors associated with the largest increased risk, but no factor increased odds ratios by more than 1.69, where 2.0 has been suggested as a threshold representing a clinically significant level of risk (Sackett *et al.* 1991). For CSLI, congenital abnormality was also the largest risk factor and the only one which was greater than 2.0. In the co-morbid condition, aspects of adverse birth history also increased the odds somewhat. As Fox *et al*. excluded children with "a more general language problem than a speech disorder", the heightened risk for SSD for children with birth difficulties in their cohort cannot be put down to children with a co-morbid condition. An alternative explanation may be that SES is not a risk factor, but a protective factor. In large-scale population studies such as Smit et al. (1990), SES is not reported as having an effect on phonological acquisition, so it appears that, in contrast to language development, low SES is not a handicap to speech developing normally. However, where speech development falters due to some other reason, children with high SES might receive more stimulation or other assistance than those with low SES, such that more high SES children overcome their difficulties at an earlier stage and are thus not recorded as having speech difficulty.

The proportion of males in the groups of those with SSD (68%) or SSD+LI (70%) in Delagado *et al.*'s study was, as in Cambell *et al.*'s study, greater than that in the group of TD children (51%). This was also reflected in the epidemiological study by Broomfield and Dodd (2005). In their sample of 320 children referred to SLT services at all ages for difficulties with speech and/or language, 62.8% were male. Being male is frequently also cited as a risk for language impairment (Choudhury and Benasich 2003, Ellis and Thal 2008, Harrison and McLeod 2010), following the trend for all

developmental disorders to occur more frequently in boys than girls (Kraemer 2000), which may point to a genetic component that could be a characteristic of a persistent SSD.

A positive family history of speech and language difficulties might also reflect a genetic basis to SSD that affects likelihood to resolve. Studies of family aggregation have shown greater prevalence of family members with a history of speech and language impairments in children with SSD than in controls (Felsenfeld *et al.* 1995, Lewis and Freebairn 1997, Eadie *et al.* 2015). However the rate of history of speech and language disorders among first-degree relatives in studies of children with SSD averages only about 30% (Campbell et al. 2003, Felsenfeld and Plomin 1997, Fox et al. 2002), and the difference in presence of family history between children with and without SSD is small: in Campbell *et al.* (2003) it was 36% as opposed to 25%, indicating it would not be a good predictor of persistence for the individual child. In fact, in Broomfield and Dodd's sample, the proportion of children with SSD who had a positive family history of speech and language disorder was smaller than the proportion in Campbell et al.'s unaffected group, at 22%, suggesting it is not useful for identifying children whose SSD might persist. Lewis and Freebairn (1997) hypothesized that children with a positive family history of SSD might constitute a sub-group within children with SSD, but their findings did not support this as they found no differences in measures of speech accuracy, language, oral-motor skills or literacy between those children with a family history of SSD and those without.

A history of poor motor skills, whether in general or in specific oro-motor skills, as discussed in section 1.3.5, might also indicate a more intransigent SSD. Eadie *et al.* (2014) have reported a history of concerns about motor co-ordination at age 2 as a strong predictor of SSD at age 4 (OR 4.8), (where 41% of the children with SSD had co-morbid language disorder), suggesting that slow development of motor skills may affect speech development. Wren *et al.* (2012), in the large scale ALSPAC study, found that children with common clinical distortions (see definition on page 72), and persistent speech errors at the age of 8, performed more poorly on tasks of DDK than the children who made few or no speech errors, suggesting that weakness in oro-motor skills may indeed be a factor in persistence of SSD. However, their study had no exclusion criteria, meaning that results may have been influenced by a sub-set of

children with CAS, or SSD secondary to some other condition, and may not affect children with functional SSD.

In these studies, then, there is no clear indication of what factors put children at risk of developing a SSD. As the factors considered vary in each study, each may be missing the effect of non-included variables that might affect the results. Added to variations in methodology and sampling procedures, this means it is difficult to draw conclusions about which factors might predict and/or maintain SSD.

Neither is it known whether factors that put a child at increased risk of developing a SSD in the first place, also play a part in the persistence of the SSD past school-age. Once a child is diagnosed with an SSD, do risk factors enable us to differentiate transitory from persistent impairments? Studies of older children with persisting SSD might give some indications of factors to consider. Wren et *al.'s* (2012) investigation of the ALSPAC sample at age 8 was useful in that it distinguished children with common clinical distortions only, i.e. articulation disorders, from those with more typically phonological difficulties. It showed that the prevalence rate for the former (7.9%) was much higher than that for the latter (3.6%). As we have a UK incidence study that shows that the rate of articulation disorders, which include CCD, is much lower than phonologically based SSD at earlier ages, we can deduce that articulation disorders, especially distortions of /s/ and /ɹ/, are less likely to resolve than phonologically based SSD. Broomfield and Dodd (2005) showed that articulation disorders made up only 12.5% of the SLT caseload, with phonological disorders making up the rest, and cross linguistic studies have supported this lower prevalence figure for articulation disorders at younger ages: Fox and Dodd (2001): 5% in German 2 to 7 year olds; Hua and Dodd (2000): 3% in Putonghua-speaking 2 to 7 year olds; 12% in Cantonese-speaking 3 to 6;6 year olds. If this proportion changes in favour of articulation disorders at older ages, as shown by Wren et al (2012), it would suggest that phonologically based SSD are more likely to resolve than articulation disorders.

However, unless we know the distribution of factors of interest in the groups of children who resolve their SSD as well as in those with PSSD, they may not distinguish which pre-schoolers in a clinical sample will go on to have a persistent SSD. For example, in Wren et *al.'s* (2016) sample of children with persisting SSD at age 8 in the ALSPAC study, predictors of SSD (not including common clinical distortions, defined in this

paper as per Shriberg 1993) at age 8 were found to be male gender, early language delay as a pre-schooler, weak sucking as a baby, unintelligibility to strangers at 3, difficulty with speech sounds at school age, history of co-ordination problems or hearing difficulties, and low SES. As all of these have been noted to be risk factors for having SSD to start with, unless children with these characteristics resolved their SSD faster than those without them did, we might expect the group with persistent SSD to show these characteristics, compared to the non-SSD population. For example, the proportion of boys with PSD in Wren et al (2016) is 64.3%, very similar to the 62.8% found by Broomfield and Dodd (2005) to be the proportion referred to SLT services, and lower than the proportions found by other studies to have SSD at younger ages (Delgado et al. 2005, Campbell et al. 2003), suggesting that, although male gender is a risk factor for an SSD to start with, boys are not less likely to resolve their SSD than girls. Similarly, in large population studies, Wren et al. (2016) found that suspected motor coordination difficulties increased the odds of having SSD at age 8 by 2.05. However, Eadie et al. (2016), in a similarly large population study, found that the odds ratio for SSD at age 4 was 4.8 if children had a history of concerns about using arms and legs, allowing the possibility that motor coordination difficulties decreased as a risk factor for SSD as children got older.

Thus, to find predictors useful in clinical samples, it may be more useful to look at longitudinal studies of the outcomes of particular children with SSD. The tables in Appendix 1-A and 1-B summarize 30 studies that have followed cohorts of children with SSD over time. The studies had a variety of aims, but all report outcomes at a later time for children with SSD (with or without LI), which are related back to earlier characteristics or measures. Eight studies had a primary aim of investigating the natural history or outcomes of children with speech and language impairments, and factors that were associated with outcome, while a further seven focused on academic achievement, and three aimed to examine both speech and language outcomes and literacy/academic achievement. Two were studies of clinical efficacy, and four investigated types of speech error and their relationship to outcome. One examined the relationship between speech perception and PA abilities, and literacy outcomes, while another looked at predictors and outcomes of decisions to offer intervention to children with speech and language impairment. Two population studies which examined the same cohorts at different intervals investigated prevalence and outcomes for children with SSD. The number of variables taken into account in these studies varied from one or two, to 167

variables covering demographic, treatment, anatomical, hearing, cognitive linguistic and psychosocial factors. The following section evaluates their findings concerning the number of children who resolve their SSD, and factors predictive of transient or persistent SSD.

1.5.3 Age of participants and time course of study

The studies report a wide range of outcomes. In the 21 studies which report rates of resolution of SSD, the percentage of children who attain age appropriate speech range from 12% to 94%. The study with the lowest resolution rate, Broomfield and Dodd (2011), followed the cohort over the shortest interval, 6 months, and included children from one to 16 years of age. Table 1-1 below shows that, generally, the longer the interval, the more children are likely to resolve. However, there is an effect of age, with far fewer children resolving within 1 year in studies recruiting children older than 3;6. In fact, studies of children recruited between 3 and 6 years of age, take 2 to 3 years to show similar resolution rates to those that children recruited under 3;6 can show in one year. Those recruited after the age of 4;6 take over 3 years to achieve comparable outcomes. The effect of age, whereby those who are older at diagnosis have poorer outcomes, has also been reported by Paul (1993) and Rescorla and Schwartz (1990) for toddlers with late language emergence, and potentially reflects the wide range of normal development at younger ages, although a possible contributing factor was that the children in Paul's study were categorised as delayed by parent report rather than by direct assessment, allowing the possibility of over- identification.

1.5.4 Effect of co-morbid language impairment on rate of resolution

Two studies reported resolution figures separately for children with SSD and SSD+LI: Lewis *et al.* (2000b), Lewis *et al.* (2015). In both studies, a smaller proportion of children with co-morbid LI resolved, i.e. children with SSD+LI were more likely to have persisting SSD. Bishop and Edmundson (1987) also found that children with more pervasive language difficulties were less likely to resolve their difficulties, observing that a 4-year-old who was impaired only in phonology was likely to be problem-free by 5;6 years.

1.5.5 Effect of therapy

In most studies, receipt of therapy was incidental to the study, and therefore not reported. Two studies, Roulstone *et al.* (2003), and Williams and Elbert (2003), reported

on children who were in a "watch and wait" condition, where no child received therapy, but outcomes were not substantially different to other studies. Two studies were efficacy studies, and both showed that children with SSD who received therapy had better outcomes in terms of normalisation: in a randomised trial over one year, Glogowska *et al.* (2000) compared a therapy as per normal service delivery group with a "watch and wait" group.

Table 1-1: Proportion of children who achieved age appropriate speech by the end of the study, by age at recruitment and interval

Age at start		**≤3;6**				**3 to 6 years**				**4;6+**			
Interval		**≤ 1yr**	**≤ 2yrs**	**≤ 3yrs**	**≤ 3 to 5**	**≤ 1yr**	**≤ 2yrs**	**≤ 3yrs**	**4yrs**	**≤ 1yr**	**≤ 2yrs**	**≤ 3 to 4**	**7yrs+**
Study	**N**												
Roulstone et al. 2003	69	58%											
Williams & Elbert 2003	5	60%											
Paul 1996	31				74%	(normal speech and language)							
Roulstone et al. 2008	191				89%								
McIntosh & Dodd 2008	10	60%											
Bishop & Edmondson 1987	9						77%						
Shriberg et al. 1994a	64					19%							
Shriberg et al. 1994b	29							66%					
Lewis et al. 2000b	28								94%				
Lewis et al. 2000a	24								75%				
Hesketh, 2004	35							80%					
Bernhardt & Major 2005	12							58%					
Lewis et al. 2006	28											54%	
Nathan et al. 2004a	47											51%	
Nathan et al. 2004b	39										29%		
Rvachew 2006	37									22%			
Rvachew et al. 2007	58									26%			
Rvachew 2007	33										52%		
Broomfield & Dodd 2011	710	12%	SSD and/or LI, ages from 0 to 16 years										
Lewis et al. 2015	77	(SSD only)											91%
Lewis et al. 2015	93	(SSD+LI)											70%
Morgan et al. 2017	93								~60%				

A significantly larger proportion of those with SSD who received therapy resolved than those who had no therapy, but actual figures were not reported. Broomfield and Dodd (2011)'s randomised study also reported better outcomes for children with a primary diagnosis of SSD who received therapy: after six months, over 20% of those who had received therapy had age appropriate speech, compared to fewer than 3% who had not. In contrast, two studies where receipt of therapy was tracked, despite being incidental to the study, reported poorer outcomes for those who received therapy: in Bishop and Edmundson (1987), 87% of those who had a persisting SSD had received therapy, compared to only 69% of those who resolved their SSD. In Paul (1996) 39% of children with poor outcome and 29% with good outcome had had some preschool intervention. However, Paul's subjects were children with slow expressive language delay overall, not children with a primary diagnosis of SSD. Shriberg *et al.* (1994b) similarly found that receipt of intervention was not a predictor of short-term normalisation of SSD. As the children in these studies were not randomised to treatment condition, it is likely that worse outcomes were seen with therapy, because therapy was prioritised for those with the most severe and intractable deficits.

1.5.6 Demographic variables

1.5.6.1 Gender

Five studies did not report numbers of participants by gender. Of the rest, all but one, (Eadie *et al.* 2015), (26 girls to 25 boys) showed a larger number of males than females in the group with SSD. Despite the reported gender divide in the sample of children with SSD, Eadie *et al*. reported a bigger OR for SSD in males. Six longitudinal studies considered gender as a predictor of outcome, and five found no effect, suggesting that gender is not a predictor of persistence of SSD. However, Broomfield & Dodd (2005) found mixed outcomes depending on sub-group of SSD, with girls having better outcomes than boys if they had PD, but worse outcomes if they had CPD.

1.5.6.2 Family history

The findings from many studies that children with SSD and/or LI are more likely to have relatives also affected with speech and language deficits than the general population suggests that, for some children at least, SSD is genetically transmitted. As discussed in section 1.4.2, Shriberg *et al.* (2005) posits that these children form a sub-type, with genetically transmitted cognitive-linguistics deficits as the aetiology of the SSD. It might be that SSD that is the result of such deficits is more pervasive than SSD of other origins. Four longitudinal studies considered the presence of a positive family history as a predictor of persistence of SSD. Three found no effect, but the study with the shortest interval, Broomfield & Dodd (2005), found it was associated with poorer outcome. When considering outcomes in terms of language status and academic achievement, Lewis *et al.* (2000) found that family history for speech and language disorders did not predict significant variance in measures of language, reading and spelling at school age, but family history for reading disorders predicted spelling impairment at school age.

1.5.6.3 Socio-economic Status (SES)

Lower SES was associated with persistent SSD at older ages in population prevalence studies by Wren et al (2012) and by Wren et al (2016). In the first case, this was measured by level of mother's education, and in the second by type of residence, i.e. homeowner versus tenant. Two longitudinal studies using maternal education as the measure, however, Shriberg et al., (1994a) and Roulstone et al. (2003), found no effect of SES on persistence of SSD, as did another using the Hollingshead Four Factor Index of Social Class (occupation & education), Lewis et al. (2015). One further longitudinal study, Morgan *et al.* (2017), which did not report the type of measure used, also found no association with persistence of SSD. It may be that results are confounded by the number of children with SSD+LI rather than SSD alone. Beitchman *et al.* (1996) found that children with SSD alone mainly came from higher SES backgrounds (measured by parent's occupation) compared to those with LI, as did Young *et al.* (2002).

1.5.6.4 Developmental factors

A variety of developmental factors was considered in the studies, including motor skills, ear infections and late onset of speech. Only one longitudinal study (Shriberg

1994a) investigated association between motor skills and SSD, and found no influence of general motor difficulties on persistence of SSD. Wren *et al.* (2016) reported that children with SSD at age 8 were more likely to have had weak suck as an infant than controls, but this may have been due to the characteristics of the sample, which included children with SSD secondary to some other condition that could affect sucking, e.g. children with cerebral palsy, cleft palate, or learning difficulties. Perinatal difficulties were investigated in two studies. Broomfield & Dodd (2005) found they were associated with better outcomes for speech and/or language, whereas Shriberg (2004b) found no association with likelihood to resolve the SSD. History of ear infections with associated transient hearing loss is frequently considered a risk factor for SSD. However, Paradise et al. (2005, 2007) carried out a large-scale longitudinal study that showed no association between SSD and otitis media with effusion in otherwise healthy children. This was investigated in two longitudinal studies of children with SSD, where Shriberg (1994b) found no association with persistent SSD, and Broomfield and Dodd (2005b) found a greater likelihood of spontaneous resolution for children with consistent phonological disorder who had a history of intermittent hearing loss. Wren *et al.* (2016) found that tympanostomy tube insertion was a predictor of SSD in 8 year olds. However, Niclasen *et al.* (2016) in a longitudinal population cohort study of 7578 children aged 10 to 12 in Denmark, found that although there was a positive association between 4+ episodes of early otitis media with tympanostomy tubes, and later behavioural and learning difficulties, the association of tympanostomy tube insertion with demographic factors which are also associated with behaviour and learning difficulties, such as social deprivation, time in day care, parental divorce, low maternal education, lower duration of breast feeding and having fathers who smoked, made interpretation of these findings uncertain.

Shriberg et al. (1994b) and Broomfield & Dodd (2005b) also had conflicting findings regarding late talking: Shriberg's study found late talkers were more likely to resolve, whereas Broomfield and Dodd found the converse. However the shorter interval and younger age of the children in Broomfield and Dodd's study might explain the difference.

1.5.7 Perceptual, oral motor and cognitive factors

Various tasks involving perception have been investigated as a predictor of persistent SSD in five longitudinal studies. Lewis *et al.* (2015) found that children with resolved SSD had better non-word repetition (NWR) than those whose SSD persisted. As NWR tasks also involve auditory perception and encoding, it demonstrates the possibility that children who resolve their SSD have better abilities in these areas. Rvachew (2007) did not show an association between good speech perception and likelihood to resolve the SSD, but the measure used was a composite of both speech perception and phonological awareness skills, and the group of children who had had poor speech perception skills at Time 1 had improved their speech perception skills as well as their speech production in relation to the group with good speech perception by Time 2. In Rvachew (2006), where speech perception was measured separately, using a word mispronunciation detection task (SAILS: Speech Assessment and Interactive Learning System; Rvachew and Brosseau-Lapré, 2012), speech perception skills pre-kindergarten predicted about 8% of variance in PCC at the end of kindergarten. Bernhardt and Major (2005) tested the 12 children in their study on the Word Discrimination sub-test of the Test of Language Development 2 (TOLD-2; Newcomer and Hammill 1988). This is a same-different task that requires discrimination of orally presented words differing by one phoneme (e.g. *weak-weep*). Of the seven children who had resolved their SSD, only two had poor scores on this task, compared to 4 out of 5 of those who did not resolve, but it was presented at follow up rather than at initial assessment, so the children's earlier word discrimination abilities are not known. Bernhardt and Major (2005) concluded that the children who scored poorly might have had a more pervasive phonological impairment that affected both production and perception, although the task did not allow determination of whether this was due to phonological processing deficits or inaccurate underlying phonological representations. In the longitudinal study reported by Stackhouse et al. (2007), performance on tasks testing auditory discrimination of words and non-words that differed by one feature or a transposition of phonemes, and mispronunciation detection tasks, administered at 4;6 and 5;8, differentiated children who had resolved their SSD by 6;8 from those with persisting SSD. At each assessment phase, around 20% of children had poor scores on these tasks. By 5;8, all children who resolved their SSD in the course of the study were performing like controls on these input tasks. These

findings therefore suggest that children with input as well as output difficulties might be more likely to have persistent SSD.

Two studies considering oral-motor abilities as a factor in PSSD showed conflicting results: Shriberg *et al*. (1994b), in a sample of 4 year olds, found no association between abilities in isolated and sequenced oral movements or diadochokinetic (DDK) rate and persistence of SSD at age 6, but Lewis *et al.* (2015) found that adolescents with persistent SSD performed worse on the Fletcher Time-by-Count DDK rate task than those with resolved SSD, suggesting a motor-based articulatory deficit might be a maintaining factor in PSSD, supporting the findings of Wren et al. (2012). As previously surmised, it may be that differences in oral motor ability might not become evident until later ages, as it is still developing in childhood.

Low IQ has been implicated in poorer outcomes for children with speech and language disorders. Bishop & Edmondson (1987) found that 44% of speech and/or language impaired 4 year olds with IQ in the normal range had good outcomes at 5;6, compared to 11% of children with below-average IQ. Wren et al. (2012) found that higher IQ decreased the risk of being in the group of children with phonologically based PSSD at age 8, while IQ did not differentiate those who showed only common clinical distortions (i.e. misarticulations of /s/ and /ɹ/) from controls, supporting the hypothesis of a cognitive-linguistic deficit underlying phonologically based disorders. However, children with SSD often have co-occurring language disorder which can reduce scores on the verbal portion of IQ tests, thus presenting a confound. Lewis *et al.'s* (2015) study of adolescents circumvented the confound somewhat by considering only performance IQ scores (PIQ) from the WISC-III or WAIS-IV, which do not include the verbal component. Their findings supported low IQ as a maintaining factor in SSD, in that the group with persistent SSD at adolescence had lower performance IQ than both the resolved SSD group and controls. Conversely, Morgan *et al.* (2017) found that non-verbal IQ at age 4 was not a statistically significant predictor of outcome at age 7. However, age of testing could make a difference. Nathan *et al.* (2004a) found that children whose SSD persisted past the age of 6 performed worse than matched controls on a concurrent task testing non-verbal IQ, while children who had resolved their SSD performed similarly to controls, despite the fact that the children with persistent SSD had performed as well as the other 2 groups in non-verbal IQ at 4;6. This suggests that IQ may not be a factor in SSD at early ages, but a decline

in the rate of development of IQ might be accompanied by or associated with persistence of SSD.

1.5.8 Number and type of speech errors

Severity of impairment is frequently used by SLTs as a criterion for prioritisation for services (Roulstone 2001). Five longitudinal studies investigated whether number of errors was predictive of outcome at a later point. All subjects were under five years of age at initial assessment. They found that number of errors did not predict ongoing speech errors (Bishop and Edmundson 1987, Shriberg *et al.* 1994b, Bernhardt and Major 2005, Rvachew 2006, Roulstone *et al.* 2009). As children get older, however, this may change. Roulstone et al. (2009) found that PCC at age 2 did not predict PCC at age 8, but PCC at 5 years was predictive of ongoing speech errors at the age of 8 years, suggesting that transient SSD can be severe at young ages, but severity declines rapidly, whereas SSD which shows a slower rate of improvement is more likely to persist.

Several studies investigated whether the type, rather than proportion of errors, was predictive of outcome. Shriberg *et al.* (1994b) concluded that a high percentage of omission, as opposed to substitution/distortion errors, at age 4 appeared to provide the strongest predictive power for persistence of SSD in their sample. In contrast, Preston *et al.'s* (2013) study of 25 children with SSD at age 4 who were followed up at age 8, found that a high percentage of distortions at age 4, rather than substitutions and deletions, was the best predictor of poor speech outcomes at age 8. .

Evidence for worse outcomes for children who make atypical as opposed to developmental speech errors comes from several studies. Morgan *et al.* (2017), in the ELVS population study, identified 163 children with SSD at 4 years of age. At age 7, 93 were available for reassessment, and categorized as having resolved (60%) or persistent errors (40%). When family history, gender, socio-economic status, non-verbal intelligence quotient and speech error type (delay vs disorder) at 4 years were examined using logistic regression, only type of speech error was a statistically significant predictor of outcome. Those with atypical error types at 4 years were twice as likely to have poor speech outcomes at seven, compared to those with errors typical of younger children. Broomfield & Dodd (2005b) also found that children with atypical rather than delayed error types made little progress within 6 months, or even

fell behind in speech development without intervention. They classified 320 children referred to SLT services in one local area in the UK because of speech deficits, using Dodd's Differential Diagnosis system (no children were found presenting with CAS). When followed up after 6 months, their findings supported the studies by Preston and Morgan and Williams and Elbert cited above: both the children with articulation disorders (i.e. distortions), and those with consistent or inconsistent disorder (i.e. atypical errors), made less progress than those in the phonological delay group, who used error patterns typical of normally developing younger children. It was concluded that children with phonological disorder did not reorganize their phonological system without intervention when left untreated for six months, whereas those with phonological delay were more likely to resolve (Broomfield and Dodd 2011, Dodd 2011). In contrast, in Preston *et al.'s* (2013) smaller study, a history of a large number of atypical errors was not associated with persisting speech difficulty at age 8. However, it was associated with poor phonological awareness outcomes, and lower scores in reading and spelling than TD controls, although these were still within the normal range.

Claessen *et al.* (2016), in a study of 24 children aged 25-35 months, recruited through local pre-schools and nurseries without inclusion or exclusion criteria (although the pre-schools and nurseries did not take children with pervasive developmental delay), found that the younger group made more atypical errors at initial assessment (mean of 7) than the older group (mean of 3). When followed up 4 to 9 months later, both groups had decreased the number of atypical errors by a similar amount. Of the four children who made the most atypical errors (more than 10) on the Toddler Phonology Test (TPT; McIntosh & Dodd, 2008), at initial assessment, one had age appropriate speech by follow up, whereas three still showed atypical phonological development. Williams & Elbert (2003), in their study of five children under 3, found that the 2 who did not resolve within a year were differentiated from the resolvers by using atypical speech errors and simple structures, and McIntosh & Dodd (2008) found that atypical errors in five 2-year-olds predicted atypical errors at age 3. However, all these children were still under four by the second assessment, so it may be that atypical errors are a sign of a less developed phonological system but one that is still developing, rather than a specific difficulty with acquisition of phonology. Pharr *et al.* (2000), in a study of 20 late talkers (normal receptive language and non-verbal IQ) compared to 15 matched TD children at 24 months, did not track atypical errors as defined by Dodd

(2005), but found that the late talkers used a restricted variety of simpler, less advanced syllable shapes, specifically fewer final consonants and consonant clusters. However, by the second assessment at age 3, the late talkers demonstrated a significant increase in syllabic complexity and had narrowed the gap between themselves and their TD peers. While significant differences continued to be present at age 3, it is possible that the gap would continue to narrow by age 5 and disappear within the period of normalisation. No studies have looked at whether atypical errors at preschool are predictive of atypical errors at school age. Several studies have reported that children who present with atypical errors at school age have worse concurrent outcomes in terms of phonological awareness or literacy (Leitão and Fletcher 2004, Rvachew et al. 2007), so it would be useful to know whether earlier use of atypical errors at young ages is a good predictor of use of atypical errors at later ages. There are conflicting results regarding whether atypical errors at younger ages are predictive of phonological awareness or literacy difficulties: Rvachew et al. (2007) found atypical errors in pre-school did not predict phonological awareness difficulties at the end of kindergarten, whereas (Preston et al. 2013) found that children who made atypical errors at 4;6 had worse phonological awareness at age 8 than those who made delayed errors only. This is discussed further in Section 1.5.9 below.

In fact, it may be that the presence of atypical errors indicates a quantitative rather than qualitative difference, i.e. they are indicators of more severe delay rather than disordered development. Roberts *et al.* (1990), in a study of 147 TD children between the ages of 24 months and 8 years, found that the TD children sometimes used processes commonly considered as disordered, and that in fact, some processes defined as uncommon in normally developing speech (e.g. deletion of medial consonants) occurred with a higher frequency than other processes classified as common. Preston and Edwards (2010) also found that, in a sample of 43 children diagnosed with SSD of unknown origin, aged 4 to 5, all made some atypical speech errors.

Hayiou-Thomas *et al.* (2016) followed 59 children with moderate to severe SSD at 3;6, to investigate PA and literacy achievement at age 5;6 and again at age 8. Of the children with SSD, 73% had delayed errors and 27% atypical errors. They found that atypical speech errors at 3;6 were associated with poorer word reading at age 5;6, but the effect was short-lived and no difference was found at age 8. This might again indicate a more severe delay rather than disordered development. However, the

outcome for children with co-morbid language impairment was poorer than for SSD alone, so that the literacy skills of children with SSD+ LI were significantly impaired relative to controls, and language status continued to predict later development, especially of reading comprehension

From this evidence, it appears that, with regard to error types, persisting SSD may break down into 2 separate types, as reported by Wren *et al.* (2012). Firstly, common clinical distortions (CCD), which rarely impair intelligibility and have little consequence for phonological awareness and literacy, appear to persist, possibly as residual errors, into and past school age, as reported by Shriberg *et al.* (1994a), and Johnson *et al.* (2010). Lewis et al. (2015) also reported that that the persistent SSD group had primarily distortion errors. The second type are phonologically based errors, which become much less frequent by the age of eight, but which may have enduring consequences for phonological awareness, processing and literacy, especially if they are atypical error types, and even when such errors are no longer evident in conversational speech. The two types might have different predictors. For instance, distortion errors may be predicted by poor oral motor skills, or they may reflect the fact that faulty overlearned tongue postures for later developing sounds may be resistant to change in some children (Karlsson *et al.* 2002). Conversely, phonologically based errors may be better predicted by phonological processing deficits in the perceptual, encoding or cognitive-linguistic domain, or may reflect weaknesses in how phonological information is organized or represented, resulting in weak phonological representations that are thought to underlie some difficulties with literacy (Preston and Edwards 2010).

Finally, some children may present with both types of error. This could make identification difficult. For instance, at younger ages when most SSD are phonologically based, motor measures may not predict persistence of SSD because of the preponderance of phonological deficits. For instance, oral motor abilities did not predict persistence of SSD at six (Shriberg et *al.*, 1994a), but did in adolescents (Lewis *et al*., 2015). Similarly, phonologically based deficits may not predict speech errors that persist into older ages, partly because some underlying deficits may resolve in the short term, as exemplified in Rvachew (2007), where children with poor perception initially, began to catch up with their unaffected peers by Time 2. Other factors may be predictive only at a later age, for example in Roulstone *et al*. (2009), number of speech

errors at age 5 predicted speech errors at 8, but speech errors at 2 did not predict speech errors at five. The younger children are, the greater is the variability in the range of normal development. This is seen in the acquisition of skills other than speech and language. For instance, a study of 222 healthy children found that the milestone of sitting without support was reached at ages that ranged between 3 and 13 months, i.e. 4 times as long for the slowest as the quickest, but an analysis comparing the earliest 20% and latest 20% did not show any difference in developmental outcomes measured regularly to adulthood (Jenni *et al.* 2013). This variability in normal development makes it difficult to distinguish persistent from transient impairments. For instance, Roulstone *et al.* (2003) in a sample of children aged under 3;6, found that variables that predicted persistence of SSD when measured in older children, such as maternal education and language status, were not predictive of SSD at age 4. The only factor with a significant association with outcome in their sample was the SLT's judgement of the child's communicative functioning as measured by the Therapy Outcome Measures (TOM) disability scale, (Enderby and John 1997).

It appears that predictors of persistence at different ages are not generalizable. In order to aid in clinical decision-making regarding need for intervention in young children, it is therefore important to look for predictors of outcome in children when they are most useful, i.e. at the age when the majority are referred for services.

1.5.9 Effects of SSD on literacy acquisition

Several studies have shown that children with SSD are at significantly higher risk of difficulties with phonological awareness and problems with acquisition of reading, than children with typically developing speech. This risk has been shown to be higher if they also have language difficulties, but lower if their speech difficulties resolve before school age (Raitano et al. 2004, Lewis et al. 2000). For example, Nathan et al. (2004b), in a study of 39 children who had been diagnosed with SSD at age 4 to 5, reported that, at age 7, 47% of those with SSD alone and 63% of those with concomitant speech and language problems scored more than 1 SD below the mean on measures of reading and/or spelling. They also scored worse at maths. However, the 11 whose speech difficulties had previously resolved scored as well as the control group.

As mentioned above, several studies have found that some error types may be predictive of later difficulties with acquisition of literacy. Harris et al. (2011),

compared the performance of 13 children aged 5;2–7;9, with SSD, and 6 typically developing controls, on tasks measuring phonological awareness, letter knowledge, and real and non-word reading. They found that children with delayed speech, i.e. who used only error patterns typical of normally developing younger children, performed like typically developing controls on all phonological awareness and reading measures, but children who consistently made atypical errors had difficulties on most phonological awareness tasks and showed no measurable emergent reading ability. In line with these findings, Leitão and Fletcher (2004) found that children with atypical errors at age 5-6 had poorer phonological awareness and greater difficulties with literacy at age 7 and later at 12-13, than children who used only delayed error patterns.

Findings that atypical speech errors appear to be linked with phonological awareness have also been seen in preschool children (Mann & Foy, 2007; Preston & Edwards, 2010), but are less consistent. Preston *et al.* (2013), in a study of 25 children, found that a high percentage of atypical errors at age 4 was associated with lower phonological awareness and reading scores at age 8 (although reading level was still age appropriate). In contrast, Rvachew et al. (2007) followed up 58 four year-olds with SSD one year later, and found that, although there was an association between frequent typical syllable structure errors at age 4 and poor phonological awareness at age 4, and between use of atypical segment substitutions at age 5 and poor phonological awareness at age 5, the child's pattern of error types at age 4 was not a reliable predictive indicator of which children would pass or fail the PAT test of phonological awareness (Robertson and Salter 1997) at 5 years. Neither was severity of SSD a predictive indicator. While children who achieved age appropriate speech before the end of the kindergarten year were very likely to have achieved age-appropriate phonological awareness skills, severity of the speech-sound disorder at either age 4 or 5 was not a reliable indicator of phonological awareness skills for the individual child, with 41% of children with only a mild SSD failing the PAT at age 4, and 36% at age 5.

Thus further exploration of factors that distinguish pre-schoolers with SSD who go on to have literacy difficulties from those that do not would indicate a useful sub-type.

1.6 Conclusion of Introduction

A usage-based theory of acquisition indicates the complexity of the task facing infants in acquiring speech. It highlights the role of social context in motivating the

acquisition of gestalt word forms rather than discrete speech sounds, with the former gradually giving way to segmental representations as growth in the lexicon creates a need for more finely differentiated representations. At the same time the ability to make accurate auditory to articulatory mappings is still developing, alongside development of motor control and cognitive abilities. Theory thus suggests various ways that the process could go awry, with resultant SSD: deficient input to the system for appropriate statistical learning; slow development or deficits in perceptual and cognitive-linguistic abilities needed to carry out that statistical learning and to construct and develop phonological representations for words and phonetic representations for segments, with resultant failure to pass from whole-word to accurate segmental representations; failure to set up accurate auditory to articulatory mappings, or deficits in motor control to plan and execute those mappings. It is also possible to conceptualise spontaneous recovery from SSD, as the development of these various abilities continues throughout early childhood, rectifying mismatches in ability that might lead to SSD.

Findings from the literature presented in section 1.3 regarding deficits shown by children with SSD are mixed. There is little support for insufficient input being the sole cause of SSD. Difficulties in assessment arise in tasks that aim to assess the input side of speech processing, where it is difficult to tease out whether deficits lie at the perceptual level, or at the cognitive-linguistic level of what is done with those percepts. However, a number of studies show that at least a sub-set of children with SSD have deficits in the input side of speech processing. Evidence for deficits in more domain general tasks that might influence speech processing and acquisition, such as rule abstraction and cognitive flexibility, provides a basis to explain the higher risk for difficulties with phonological awareness and acquisition of literacy experienced by some children with SSD, but there is not yet a large amount of evidence for such deficits, with findings regarding deficits in higher level phonological knowledge from Munson *et al.* (2011) and Torrington Eaton and Ratner (2016) conflicting with those from Dodd and McIntosh (2008) and Crosbie *et al.* (2009). However, both Waring *et al.* (2016) and Torrington Eaton and Ratner (2016) concurred in findings regarding reduced short term memory in children with SSD, although again this was not supported by the results of a larger study by Lewis *et al.* (2011). Conflicting findings regarding the role of motor deficits in SSD highlight the difficulties of comparing results from different types of assessment, and the need for clarity about what tasks are

assessing. Studies which used tasks containing multisyllabic words to assess oral-motor abilities, e.g. McLeod *et al.* (2017), tended to show high rates of deficits, but these tasks are also vulnerable to phonemic substitutions or structural omissions commonly used by children with SSD. However, tasks using instrumental measures do provide evidence that motor deficits may be a factor in SSD for a sub-set of children, and the fact that this is more evident in the studies of older children cited suggests it may be a factor in persistence of SSD.

Overall, the evidence suggests that SSD is not due to a deficit in one particular area, with different sub-sets of children showing deficits in different areas, confirming the view that children with functional SSD do not form a single, homogeneous group. Identification of pre-schoolers whose SSD will persist is further complicated by the fact that studies looking at younger children find an overlap of abilities, whereby some children with SSD perform as well as TD children in all tasks. This overlap is less evident in studies of older children and adults, suggesting that variability in the course of development may explain transient SSD in pre-schoolers, with children "catching up" in speech development as supporting systems continue to mature.

The undersupply of resources for SLT evidenced in section 1.1, underlines the need to distinguish those children whose SSD will persist without help from those with transient SSD, and to identify how to treat them most efficiently. Thus researchers have sought to classify children on the basis of indicators that suggest underlying deficits or causal and maintaining factors, and predict persistence over transience of SSD. The literature has shown that children identified in the preschool years as presenting with SSD have very variable outcomes. Different studies have shown that factors such as presence of concomitant language disorder, type of speech errors, or deficits in speech perception or motor skills may be factors in those whose SSD persists, or who have later difficulties with phonological awareness and literacy. Longitudinal data can be used to investigate which factors are actually predictive of long-term outcomes for which children, or whether predictive relationships change over time.

When considering treatment decisions for children with SSD, to use clinical resources most effectively, SLTs need evidence as to when and where to focus their efforts in order to achieve the greatest gains. Key to these decisions is the understanding of

individual differences in children's trajectories, and the predictors of those differences. However, as yet, no one predictive instrument or procedure has been found that has adequate sensitivity and specificity to predict which children will resolve their SSD without intervention. One system of classification into sub-groups, Dodd's Differential Diagnosis System, has divided the group of children with phonologically based SSD into those who are merely delayed, and those who have an underlying disorder, and predicted differential rates of resolution and different long-term consequences for each; however, no study has followed a group of children categorised according to this system longitudinally to establish whether the categories are stable over time, or whether they predict outcome.

The study described in this dissertation followed 85 pre-school children, referred to SLT services due to concerns regarding speech development, over 2 years. The first analysis presents the results of an initial cross-sectional study, including factor analysis using a range of different assessments, attempting to profile sub-groups on a variety of potential causal or maintaining factors as discussed in this introduction. The second analysis presents the outcomes of the same cohort after 2 years, to establish whether sub-groups categorised according to Dodd's classification system remain stable, and to look for any other factor that might distinguish transient from persistent SSD.

The next chapter describes the aims and methodology.

2 Chapter 2: Aims and Methodology

2.1 Aims

The research studies in this thesis include a cross-sectional analysis and a longitudinal study. The aim of the cross-sectional study was to investigate the performance of children with SSD of unknown origin on tests of language, phonological processing ability, motor planning and oral motor skills, speech accuracy, cognitive flexibility and rule abstraction ability, to discover whether:

- there are associations among such abilities, and the nature of these;
- children can be classified into sub-types that show distinct profiles of deficits across these factors

The longitudinal study investigated the outcomes of the same children 2 years later, to discover:

- the proportion of children whose SSD resolved within 2 years
- whether children's error profiles as categorised by Dodd's Differential Diagnosis Classification System remained stable, reflecting different underlying psycholinguistic processing deficits
- to investigate whether presence of risk factors, concomitant language disorder, type or number of speech errors, or deficits in speech perception or motor skills, were predictors of outcome in terms of speech accuracy and phonological awareness abilities.

Because the research was intended to provide data that would help SLTs to predict outcome for individual children referred to them, the study was designed to be clinically replicable, i.e. to look for factors that might be predictive of outcome using the type of measures and analysis which are already commonly used or which would be feasible to use in assessment of children referred to SLT services, given the constraints on resources and time available in practice in the publicly funded health service. Similarly, the cohort recruited was not selected using rigid criteria based on test scores, but was simply a group of children who had been referred to SLT services because of concerns that their speech was impaired for no obvious reason. Many research studies only include children in the moderate to severe range of speech impairment. However, children with mild SSD are frequently encountered in the clinic,

and perhaps pose the most difficult question regarding prognosis: do they require therapy at all? Thus this study set out to investigate empirically the range and severity of children presenting to SLT services and to relate these to outcome.

2.2 Methodology

2.2.1 Ethical approval

Ethical approval for the cross-sectional and longitudinal study was granted by the Scientific Research Ethics Committee at the Mid-Western Regional Hospital in Limerick, Republic of Ireland. All parents gave informed consent for their children to be included. Given the young age of the children in the longitudinal study (3;0 to 5;11), only parental consent was gained for this study.

2.2.2 Participants and recruitment

SLT managers from three Health Service Executive (publicly funded body responsible for the provision of healthcare and personal social services in Ireland) areas in the Midwest of Ireland acted as gatekeepers for recruitment. These three areas cover a mix of city, town and rural environments and a range of socio-economic statuses. Parents of children aged between 3;0 and 5;11 awaiting initial assessment following referral owing to speech sound difficulties were invited to take part in the study and sent information sheets about it. At the time of this study, SLT services in Ireland were severely under-resourced and waiting times were as long as 2 years (Hough 2010). Therefore the offer of an immediate assessment was an inducement to parents to take part in the study. Parents of 100 children returned consent forms to take part in the study to their local HSE speech and language therapy office, who forwarded their contact details on to the investigator.

The inclusion criteria specified children who were native English speakers aged between 3:0 and 5:11, referred to SLT paediatric community services owing to concerns about speech. Exclusion criteria excluded children who had SSD secondary to some other disorder such as cleft palate, cerebral palsy, or congenital syndrome, who were growing up as bi- or multi-linguals, who were currently in receipt of speech and language therapy, who had additional diagnoses of intellectual or sensory impairment, or who had significant behavioural issues such as would preclude them from taking part in the assessment tasks. One hundred children were recruited but 15 did not fit the inclusion/exclusion criteria, leaving 85 participants. Reasons for

exclusion are shown in Table 2-1. The excluded children were assessed and then returned to their local service with a report of the assessment findings.

Table 2-1: Reasons for exclusion from research sample

Reason for exclusion from sample	**No. of children**
Growing up in bi or multi-lingual households	4
Outside the age range of the study	4
Referred to Early Intervention Service: Additional diagnosis of intellectual impairment	1
Did not attend for full assessment	2
Unable to take part in formal assessment due to non-compliance (all under 4 years)	3
Currently attending private speech and language therapy	1
Total:	15

2.2.3 Study Design Overview

Participants were assessed at two time points, approximately 2 years apart. At initial assessment a case history was taken and performance was measured on the range of assessment tasks (specified below). Following the assessment, reports of the assessment results, and copies of the standardised assessment forms, were sent to the child's local SLT office where the child was moved to the waiting list for review/intervention. The child then followed the normal pathway of care offered by that service. Participants' parents were contacted again approximately 2 years later and offered an appointment for reassessment by the researcher. Ten children had withdrawn or could not be contacted, leaving 75 children to be assessed at Time 2. At the follow up assessment, performance was measured on a smaller sub-set of tasks, detailed below.

2.2.4 Assessment Measures Used

This section describes the relevant assessment measures that were administered at initial and follow up assessment.

2.2.4.1 Speech measures

Each standardised assessment was administered and scored according to the instructions in the manual for that assessment.

Responses to tests of speech accuracy were transcribed by the researcher (an SLT with a degree in linguistics and extensive experience in teaching phonetic transcription) using narrow phonetic IPA transcription on-line, while a simultaneous recording was

made on an Olympus VN5500PC voice recorder for later review. Where the researcher noted any uncertainty regarding a production, this was reviewed on the recording to decide a final transcription. Five (6%) of the recordings of children's responses to the DEAP phonology sub-test were independently transcribed, also using narrow phonetic transcription, by a second SLT and lecturer, experienced with paediatric clients, from the recordings made. In the case of non-agreement, the recording was reviewed. In all cases, the original transcription by the researcher was retained. The phone-by-phone percentage of agreement between the original and second transcription was 88%, which is considered an acceptable level of agreement (Shriberg and Lof 1991, McSweeny and Shriberg 1995).

2.2.4.1.1 Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd *et al.* 2002)

The DEAP is a comprehensive assessment of articulation and phonology in children and was chosen because it has been normed on a sample of children in Ireland. All children first completed the Diagnostic Screener, where children are asked to name 10 pictures twice to check for consistency of productions. All children then completed the Phonology subtest, where they were asked to name 50 colour pictures which sampled all consonants in syllable-initial and -final positions, as well as all vowels and diphthongs except for /ʊɚ/ as in *cure*. Children who had more than 5 inconsistent productions on the Diagnostic Screener were also administered the Inconsistency sub-test, where they are asked to name 25 pictures on 3 different occasions.

An error was defined as a difference between the child's realization of the speech sounds in each word and the expected realisation of an adult in the same speech community. Hiberno-English is a rhotic dialect and the local dialect also features stopping of dental fricatives. Therefore rhotic vowels and realisations of /θ/ and /ð/ as /t/ and /d/ respectively, were scored correct. Percent consonants correct (PCC) was calculated for each child on the DEAP Phonology sub-test to allow for comparison with norms, and a Scaled Score was assigned following the norms in the Irish manual. Any error patterns used were noted and counted. For children who completed the Inconsistency sub-test, the number of words produced in more than one version across the three trials was counted and converted to a percentage. This resulted in 3 measures for each child: PCC in the Phonology subtest, Scaled Score for PCC in terms of severity, and number of processes used. For those children identified as inconsistent, a fourth measure, percentage of inconsistency was included.

2.2.4.1.2 Test of multisyllabic words

A novel test of multisyllabic words was devised to augment the 5 words of over 2 syllables in the Phonology test. This was included to aid in differential diagnosis of childhood apraxia of speech (CAS) from inconsistent phonological disorder (IPD). Children with CAS have been shown to demonstrate worse scores on imitation of spoken words than in spontaneous production, whereas children with IPD show the opposite presentation (Dodd *et al.* 2005). Pictures of 7 words (dinosaur, motorbike, caterpillar, kangaroo, butterfly, computer, wheelbarrow) were chosen on the grounds of being easily imageable, and likely to be acquired by children of the age of those in the study (Morrison *et al.* 1997). Coloured illustrations for each word were retrieved from Google images, and were presented in a PowerPoint presentation for children to name spontaneously. Following presentation of all the words, the child was then asked to repeat the words again, in the format:

Researcher: "Can you say this one again? Say it the way I say it: (word)". Percent Consonants Correct was calculated for each set of spontaneous and imitative productions in the same procedure as described for the DEAP Phonology sub-test, resulting in 2 measures: PCC spontaneous and PCC imitative.

2.2.4.2 Speech Perception/Discrimination

2.2.4.2.1 The test of non-word discrimination.

This was an input processing task that used a similar procedure to tasks previously described in Reuterskiöld-Wagner *et al.* (2005), Stackhouse *et al.* (2007), and Byun (2012). It was intended to provide insight into the child's phonological encoding ability and/or auditory discrimination as shown in, for example, the representation of the location of boundaries between phonemes. Reuterskiöld-Wagner *et al.* (2005) suggest that non-word discrimination tasks are an effective alternative to non-word repetition in gaining a measure of phonological representation for children with expressive phonology deficits, while controlling for the confounding variables of articulation planning and execution.

Children were asked to make a judgement about utterances produced by two speakers of Hiberno- English. The stimuli were presented in a video recording on a 15.6" computer screen. The video showed two young women side by side. Shoulders, neck and full head were included in the image. The woman on the left presented the first item of the trial and the woman on the right the second. The child was told that the first

girl was going to say a “funny word” and her sister had to copy her and say it “just like her”. After the presentation of each trial the child was asked “Did she say just what her sister said?” by the investigator. Before each trial the investigator said “Ready?” to focus the children’s attention.

The stimuli comprised 16 non-words, 7 of one syllable and 9 of 3 syllables, presented in random order. Eight words (4 of 1 syllable and 4 of three syllables) were repeated identically. The eight items which were not repeated identically were repeated in versions in which one phoneme in word medial or word final position was substituted with another that differed by one feature of place, voice or manner. Children had to answer ‘‘yes’’ or ‘‘no’’ and their response was marked on a score sheet. The score awarded was the number of items that the child correctly discriminated as same or different. The stimuli for the task are shown in Appendix 2-A.

2.2.4.3 Motor measures

2.2.4.3.1 Oro-Motor Test

The Oro-Motor Test used was that included in the DEAP (Dodd *et al.* 2002). It includes a diadochokinetic test where children are asked to repeat the word ‘‘pat-a-cake’’ five times if they are aged under 4;11 and 10 times if they are aged between 5;0 – 6;11, a test of isolated movements where children are required to imitate four demonstrated oro-motor gestures, e.g. ‘‘Can you move your tongue from side to side like this?’’ and a test of sequenced movements, where children are required to imitate three demonstrated sequences of two oro-motor gestures, e.g. ‘‘Can you kiss and cough, like this?’’ Scaled scores were assigned according to norms given in the Irish manual, resulting in 3 measures: DDK score, Isolated Movement Score and Sequenced Movement Score.

2.2.4.3.2 Beery-Buktenica Developmental Test of Visual-Motor Integration, Sixth Edition, (BEERY VMI; Beery and Beery 2010)

The BEERY VMI assesses the individual’s ability to integrate their visual and motor abilities. The visual motor integration test presents drawings of geometric forms in order of increasing difficulty that the child is asked to copy. It provides norms established from testing 1737 children who selected to be a representative sample of the population in USA, and was re-standardised in 2010. It has been found to have good reliability and adequate validity with an American population (Poteat 2014), and the manual provides evidence of association of Beery VMI scores with intelligence

scores and academic achievement. It can be administered by speech and language therapists (SLTs).

2.2.4.4 Language measures

2.2.4.4.1 British Picture Vocabulary Scale: Third Edition. (BPVS; Dunn *et al.* 2009)

The British Picture Vocabulary Scale 3rd Edition (BPVS III) is a widely used assessment of receptive vocabulary. This is the most commonly used vocabulary measure in the UK (Dockrell and Marshall 2015). It provides norms for individuals aged 3-16 years with reliability reported as 0.91. The child is shown an array of 4 colour drawings and asked to point at the drawing corresponding to the word said by the researcher. As it requires no verbal response, it is easy to use with children in early stages of assessment while they are still gaining confidence to speak to the researcher. This was administered and scored according to instructions in the manual and a standard score was calculated for each child according to the norms given.

2.2.4.4.2 The Clinical Evaluation of Language Fundamentals-Preschool 2 (UK) (CELF-P2 (UK); Semel *et al.* 2006)

The CELF-P2 (UK) measures a range of language skills in children aged 3 to 6 and is the most widely used assessment of language skills for this age range in Ireland. Although Irish norms are not available, it was normed on 588 children in the UK. It provides a short form, Core Language Score (CLS), which, according to the technical report, quantifies a child's overall language performance in sufficient depth to identify the presence or absence of a language disorder, which was the purpose of the assessment in this case. This CLS is derived by summing the scaled scores from the subtests that, the authors indicate, best discriminate typical from disordered language performance, namely the Word Structure, Expressive Vocabulary and Sentence Structure sub-tests. This short form allowed assessment to be completed in a reasonable time. Another consideration in its selection was that it was the most widely used formal measure of language in the SLT services that the children were referred from, so was an acceptable measure to report back to them. The assessment was administered according to instructions in the manual and scores were calculated to provide a Core Language Standard Score for each child.

2.2.4.5 Phonological Awareness measures

2.2.4.5.1 The Preschool and Primary Inventory of Phonological Awareness (PIPA; Dodd *et al.* 2000)

The PIPA assesses the nature and extent of a child's phonological awareness development. It was normed on a UK sample of children aged 3 to 6:11, and was considered the most suitable test of phonological awareness available, as UK children start reading instruction at a similar age to Irish children, and no test of phonological awareness with Irish norms exists at present. At the time of testing, it was the only PA test available that had real normative data for 3 year olds as opposed to statistically inferred data. At initial assessment, two subtests: syllable segmentation and alliteration awareness were administered. The scores were summed to make a composite score of phonological awareness at initial assessment. At reassessment, four sub-tests: rhyme awareness, alliteration awareness, phoneme isolation, and phoneme segmentation, were administered and scored according to the instructions in the manual. Letter knowledge was not included as this is dependent on instruction, and not all children had started school or preschool.

2.2.4.5.2 Phonological Legality Task

Another aspect of phonological awareness is knowledge of the constraints of the phonological system being acquired (Harris *et al.* 2011), that is, knowing which phonemes occur in the ambient language and the rules governing how they can be combined to make words. This task was designed to assess whether children could differentiate consonant combinations that were legal from those illegal for English, in order to tap implicit uderstanding of the phonological system and was based on a similar task described in Dodd *et al.* (1989).

Materials constituted 12 pairs of non-words, each containing a sequence of two consonants. The first of each pair had a consonant combination that was legal for English; in the second, this was replaced with an illegal consonant, e.g. /slaʃi/ and /svaʃi/. The assessor engaged the child in a conversation about spacemen, specifically about how they talk with funny words different to “our words”. The child was told that the assessor was going to say some new words that he/she had never heard before, and some of them were like normal words and some would sound like funny spaceman language words. The child was then presented with 2 boxes, one labelled with the picture of a man, and the other with the picture of a spaceman. The non-word pairs were presented in a predefined order. The investigator held one card in each hand and

read the non-word from each card. The child was asked to decide which was the "spaceman word" and to take it and place it in the "spaceman box". The investigator then placed the other word in the box labelled with a man, and the next pair was presented. The words were repeated until the child made a decision. The cards were concealed so that the child would not be able to read them. The child's score was calculated as the number of non-words identified as "spaceman words" which were actually phonologically illegal in English. The stimuli for the task are shown in Appendix 2-A.

2.2.4.6 Cognitive tasks

2.2.4.6.1 Flexible Item Selection Test (FIST; Jacques and Zelazo 2001)

This is a test of rule abstraction and cognitive flexibility, aspects of executive function. The computerised version of the same task used in Crosbie *et al.* (2009) was used, where stimuli are presented in a PowerPoint presentation (Jacques and Zelazo 2005) rather than on paper cards. Other than that, task administration was done as carried out by Crosbie et al. The child is shown pictures of three items that vary by dimensions of size, item pictured, and colour, where two items share one dimension and a different two items share another e.g. a big blue ship, a small red fish and a small blue teacup: two items share 'blue' and two share 'small'. Following graded practice trials that ensure the child can identify all of the dimensions and instructions, 15 test trials are presented, in which the children are instructed to point to two objects that go together in one way: *Show me 2 pictures that go together.* They are then instructed to point to two objects that go together in a different way: *Show me 2 other pictures that go together*. The task requires the child to identify two of the three items that are similar along one dimension (e.g. colour) but then to inhibit that response and identify another two items that are similar along a second dimension (e.g. size). The number of correct responses to the first dimension was recorded as the score FIST1, and the number of correct responses to the second dimension was recorded as score FIST2. The first response is taken as a measure of rule abstraction ability, and the second as a measure of cognitive flexibility (ability to inhibit first response and switch focus to a different dimension).

2.2.4.6.2 Pattern recognition task

Pattern recognition is the ability to identify regularities in the input. Patterning has been identified to be an essential skill in early learning, as it encompasses skills of spatial awareness, sequencing and ordering, comparison and classification (Papic and

Mulligan 2005). The ability to recognise and use patterns has been described as a general cognitive characteristic, and a study of first grader's use of pattern and structure showed that it is involved in a wide range of mathematical tasks (Mulligan *et al.* 2004). Knowledge of patterning has been found to be associated with the development of analogical reasoning in children, and the ability to identify, extend, and generalise patterns is important to inductive reasoning (Papic and Mulligan 2007).

In accordance with Dodd's argument that children who make atypical errors have failed to identify the salient features of the phonological system in the ambient language, Mulligan and Mitchelmore (2009) argue that children who show lack of knowledge of the structure of patterning tend not to detect the significant features of the structure and are more likely to focus on idiosyncratic features.

This assessment was based on the repeating patterns assessments outlined in Papic and Mulligan (2005) and Papic and Mulligan (2007) and used with preschool children of a similar age to this cohort. Children were presented with a pattern of six elements on a computer screen, and an array of 4-7 possible elements that could be dragged from the bottom of the screen to extend the pattern. Following a practice section, the children were presented with 12 trials of increasing difficulty. The first four required identification of the dimension of colour, the second 4 trials required identification of colour and shape, and the last four included three dimensions of colour, shape and size. The assessor focussed the children's attention on the pattern by running her finger along it, and the children were asked to find "*which one comes next?*. When the child indicated their choice of pattern element, the research dragged it into the empty box at the end of the line, and asked "*Is that right*?". If they indicated no, they were allowed one more choice. The score awarded was the number of elements correctly selected to extend the pattern. Test materials are shown in Appendix 2-A.

2.2.4.6.3 Test of Numerical concepts

Like pattern recognition, development of early numerical abilities also requires rule abstraction abilities. Research suggests that some children with speech and/or language impairment have worse performance in mathematical tasks at school than TD peers (Harrison *et al.* 2009). However, the evidence supporting a relationship between phonological abilities and mathematical abilities is not consistent. Passolunghi *et al.* (2007), in a longitudinal study of 170 TD six year olds, found that phonological ability was not associated with mathematical learning ability. In contrast, Nys and Leybaert

(2013) found that, in children with SLI, poor exact arithmetic was specifically related to word repetition accuracy. The evidence for an association between phonological awareness and mathematical abilities seems to be stronger than for speech accuracy and the latter: Leather and Henry (1994) found that a composite phonological awareness measure predicted arithmetic ability better than counting scores. Passolunghi *et al.* (2015), in a further study, found a strong correlation between metaphonological tasks and early numerical abilities, but a weaker one between word repetition accuracy and the same. How phonological abilities would affect numerical abilities is not clear: it has been suggested that phonological abilities might be involved when children use phonological name codes for numbers in counting, or in fact retrieval (Fuchs *et al.* 2006). Leather and Henry (1994) argue that some phonological manipulation tasks require arithmetic processes; for example, addition processes may be involved in phonemic blending tasks, and initial and final sound identification tasks may involve subtraction processes. The possibility of a common underlying deficit in rule abstraction abilities appears not to have been considered, although the roots of mathematical thinking are described to lie in the same areas as the construction of the phonological system, i.e. detecting sameness and difference, making distinctions, classifying and having the ability to ignore irrelevant features (Mulligan and Mitchelmore 2009).

This task was based on checklists and tasks described in Bhargava and Kirova (2006), as appropriate for assessing the development of mathematical concepts in pre-school children, specifically those concepts considered central to the early development of mathematical thinking: matching and one-to-one correspondence; sets and classification; order and seriation. The tasks were designed to be engaging to the children, involving the manipulation of toy animals and cutlery. A description of tasks and scoring is available in Appendix 2-A.

2.2.4.6.4 Auditory Memory Word Span Test

Research has indicated that both children and adults with SSD often show deficits in phonological memory (Kenney *et al.* 2006). A word span test rather than digit span test was devised, as it was thought that for some children digits would have semantic content while others would not be familiar with numbers. There might also be interference from young children's habit of repeating digits in random order when they think they are "counting". A list of 13 monosyllabic animal names was prepared, with pictures for each animal. Children were first asked to point to the picture named by the

researcher, to ensure they were familiar with all the words (i.e. none were non-words to them). All children identified all pictures correctly.

The test was administered following the procedure described in Gathercole and Adams (1994). Children were given a trial, with the instruction "*Now I am going to say 2 animals. I want you to say them back to me, in the same order I said them. Let's try. Ready?".* If the child failed the trial, feedback was given, followed by another trial. If the child was not able to follow the instruction after three trials, the test was discontinued. Following successful trials, the test items were presented by the researcher, who said a sequence of words that the child was required to repeat immediately in the same order. The word lists were constructed by sampling randomly and without duplication from the list of animal names. At least two lists of words were given at each list length.

After two successful repetitions at one list length, the length of the next list was increased by one. If the child correctly recalled only one of the first two lists, a third list that length was given. If the child was successful, lists at the next length were given, but if not, testing was discontinued. Memory span was recorded as the maximum list length that the child correctly repeated twice.

2.2.5 Procedure

The same researcher carried out all assessments at initial and reassessment. All initial assessments took place in clinic rooms either at the University's SLT Clinic or at the child's local SLT office. Reassessment took place at the University SLT clinic or the local office where the child had initially been assessed, except in 5 cases, where no clinic room was available. These children were assessed in their own homes.

2.2.5.1 Initial assessment

Initial assessment took place in 2 or 3 sessions of approximately 45 minutes duration each. In all cases, the child was accompanied by a parent or carer, and occasionally by both parents or parent and another relative. A case history was taken and included details of age, family history of speech, language or literacy difficulties, parent estimation of child's speech and language difficulties, or other concerns, parents' highest level of education, child's birth history, history of speech and language development, medical difficulties, hearing status and number of ear infections. Most children completed all assessments within 2 sessions, but where children tired easily, were not compliant or lost concentration quickly, assessments were divided among 3 shorter sessions.

Table 2-2 shows the schedule of assessments administered at Initial Assessment.

Table.2-2 Schedule of assessments at Initial Assessment

Schedule of Assessments administered at Initial Assessment	
Session 1	Diagnostic Evaluation of Articulation and Phonology (DEAP) Screener DEAP Phonology Sub-test
	DEAP Oro-motor assessment: a) DDK b) Isolated movements c) Sequenced movements
	British Picture Vocabulary Scale III
	Beery-Buktenica Developmental Test of Visual-Motor Integration, 6th Edition
	Test of non-word discrimination
	Test of number concepts
Session 2	CELF-P2 (UK) Core sub-tests
	DEAP Inconsistency Sub-test where indicated by Screener results
	Test of multi-syllabic words
	PIPA syllable segmentation sub-test
	PIPA alliteration awareness sub-test
	Test of phonological legality
	Test of pattern recognition
	Test of auditory memory
	Flexible Item Selection Test

2.2.5.2 Classification into diagnostic sub-groups

Following scoring of the assessments, children were assigned to a diagnostic sub-group, using the criteria outlined in Table 2-3. These criteria were based on those specified in Dodd (1995b), and Dodd and Crosbie (2005). The categorisation process revealed two interesting patterns of performance not previously discussed.

1. Four children (mean age 4;02) presented with only age appropriate error patterns, but they demonstrated so many errors that they fell below the 10th percentile according to the normative sample for their age group. This pattern demonstrates the need to distinguish between a quantitative measure of severity according to PCC, and the qualitative measure of types of errors made, according to error patterns or number of atypical errors. Quantitative measures include omissions, substitutions and distortions, irrespective of age appropriateness and whether the types of errors are typical or atypical of phonological development. Qualitative measures provide data about whether

errors are age appropriate and whether the course of development is following a typical developmental path. There is no reason why these two measures must agree. These children were assigned to the Phonological Delay group because all errors were typical of normal development.

2. Ten children consistently produced more than five words in a way that is atypical of normal phonological development, but this production did not accord with a particular error pattern. The analyses of normative data for the DEAP used the criterion of five occurrences of an error pattern in different lexical items for a child to be credited with having that error pattern. If the number of children within a six-month age band using each error pattern was more than 10% of the sample, the error pattern was considered as 'age appropriate. Determination of disorder cannot be determined in the same way, as there is no 'normative data' on disorder. In this study, the presence of five or more errors atypical of normal development at any age led to a child being assigned to the Consistent Phonological Disorder group.

The profile of performance of all fourteen children will be considered to evaluate whether they constitute special cases and whether they are more likely to resolve than other children in the same sub-type of SSD are. Each child's assignment to a particular sub-group was reviewed and agreed by Professor Barbara Dodd, (Honorary Emeritus Professor), The University of Melbourne.

2.2.5.3 Classification of severity

Severity was measured by two methods; firstly using PCC raw score from the DEAP phonology sub-test, and secondly using the DEAP scaled scores based on norms for the DEAP sample. As the DEAP classification is based on norms for age groups, and since PCC improves with age in both TD and impaired children, this may be considered a more accurate representation of severity of impairment.

Table 2-3 Criteria used to classify children into diagnostic sub-groups

Age appropriate speech	the child uses no error patterns, or only error patterns seen in at least 10% of the normative sample of children of the same age group according to the DEAP manual, and has a PCC score at or above the 10th percentile in the normative sample of children of that chronological age.
Articulation disorder	the child is unable to produce specific phonemes in any context, always producing the same substitution or distortion.
Phonological delay (PD)	the child uses one or more speech error patterns, each observed in at least five different lexical items in the DEAP phonology sub-test, that occur in at least 10 percent of children, but at least one of which is only typical of children of a younger chronological age group (by at least 6 months)
	the child uses only speech error patterns that occur in at least ten percent of children in the normative data for children of that age but has a score of PCC in the in the 13th percentile or below (scaled score of 6 or below) according to the norms for the appropriate age group on the DEAP phonology sub-test
Consistent phonological disorder (CPD)	the child uses at least one speech error pattern, identified as occurring at least five times in different lexical items in the DEAP phonology sub-test, that is atypical of any age group in the normative sample for the assessment, or
	the child consistently produces at least 5 different lexical items in the DEAP phonology sub test in a form which is consistent but atypical of any age group in the normative sample
Inconsistent phonological disorder (IPD)	the child produces multiple forms for at least 10 items when asked to name the same 25 pictures on 3 separate occasions in one session.

2.2.5.4 Criteria for co-morbid language impairment

A child was considered to have a co-morbid language impairment if she/he achieved a standard score of less than 85 (i.e. at least 1SD below the mean) on either the Core Language Score of the CELF Preschool 2 (UK) or on the BPVS III. The criteria of 1 SD below the mean is that used in Lewis *et al.* (2000), thus allowing for comparisons, and is the criteria used to define language impairment in local SLT services.

2.2.5.5 Reassessment at Time 2

The researcher contacted parents approximately 24 months after the initial assessment to arrange for reassessment. Ten children had either withdrawn or could not be contacted, leaving 75 children to be assessed at Time 2. The average time between

initial and reassessment was 24.68 months, (*SD*=3.80, Minimum =17, Maximum =37). The average age at reassessment was 6:2 (*SD*=9.96, Minimum = 4:8, Maximum =8:1).

Of those that did not re-attend, 5 were boys and 5 girls. A Chi-square test for independence revealed no significant difference in diagnostic sub-group of those who did and did not attend: $\chi^2(8, N = 85) = 13.58$, $p = .094$. Mann-Whitney U Tests showed no significant difference in age, (*U*=297, *z*=-1.06, *p*=.29) or PCC between groups, (*U*=339, *z*=-.49, *p*=.62).

At Time 2, a reduced schedule of assessments was carried out, which allowed reassessment to be carried out in one session of approximately 45 minutes, thus requiring only one clinic visit. At this point, many children had been discharged from services with age appropriate speech, therefore the reassessment did not provide any benefit to the child, so for ethical reasons, requirements for re-attendance were kept to the minimum. Table 2-4 shows the assessments administered at reassessment.

Table 2-4: Schedule of assessments administered at reassessment

Assessments administered at Reassessment
DEAP Screener
DEAP Phonology Sub-test
British Picture Vocabulary Scale III
PIPA rhyme awareness sub-test
PIPA alliteration awareness sub-test
PIPA phoneme isolation sub-test
PIPA phoneme segmentation sub-test
DEAP Oro-Motor DDK sub-test
Auditory Memory Wordspan test

All assessments were administered at the beginning of the session. As before, children were accompanied by a parent or carer, and occasionally by both parents, or parent and another relative. Following administration of assessments, a further case history was taken, including details of the child's age, parent report of therapy received in the interim, parent's estimation of child's current speech and language abilities, and parent report of how children were progressing at school or pre-school. This sequencing of events meant that, at the point when the reassessment was carried out, the researcher was blind to whether the child had received therapy or not in all but one case (where the child had attended for therapy with the researcher). The researcher deliberately did

not review any data in the file other than the child's name and address, before the re-assessment, and had had no contact with the children since the initial assessment, excepting the one mentioned above who had attended for therapy. Given the two-year gap between initial and reassessment, the researcher effectively had no knowledge of the children's initial speech status. All children completed the DEAP Phonology sub-test, the DEAP Diadochokinetic rate measure (Dodd *et al.* 2002), the British Picture Vocabulary Scale 3 and the Auditory Memory Wordspan test. All also completed the Rhyme Awareness and Alliteration Awareness subtests of the PIPA, with 74 of 75 completing the Phoneme Identification sub-test and 73 of 75 the Phoneme Segmentation sub-test.

Following reassessment, the children were reassigned to the appropriate diagnostic sub-group according to the same criteria as used at initial assessment (see Table 2-3). However, it is important to note one departure from criteria used in the DEAP for assignment to sub-group.

The DEAP does not classify the dentalisation of alveolar fricatives as a developmental process, unlike gliding of /ɹ/ to [w]. However this presentation did not fit the criteria for an articulation disorder, according to Dodd's classification, as the children were able to produce the sound in isolation, and frequently did use it in some contexts: for example some children used the dentalised production less frequently in word final than in word initial contexts. Smit (1993), in the Iowa-Nebraska norms study of over 900 children aged between 3;0 and 9;0, found that 15 to 30% of under sixes produced /s/ as a dentalised or interdental fricative, and this continued at a rate of about 10% even in 9 year olds. She suggested that as this variant is only rarely seen in the adult population, the rate would reduce further even after the age of 9. The ALSPAC study found that children whose SSD is limited to isolated lisps or difficulties with /ɹ/ made up 7.9% of the sample of 7390 children at age 8, but Wren (2015) reports these children are not considered a clinical group in the UK. Fox and Dodd (1999) also found that 35% of typically developing German 6 year olds in their study used dental or interdental productions for alveolar fricatives, and suggested that, as [θ] and [ð] are not contrastive phonemes in German, this may be considered a normal variation. In Ireland the phones [θ] and [ð] are also not contrastive phonemes in many accents. In some accents they are replaced by dental stops [t̪] and [d̪] respectively, to distinguish them from /t/ and /d/, while in others they are not distinguished from these phonemes,

being produced as alveolar stops (Hickey 2004). This suggests that a dentalised production of target alveolar fricatives may also be considered as age appropriate variation in this sample, rather than as phonological errors. Therefore, as all the children in this study who presented with their only difficulty being a dentalised production of alveolar fricatives, were under the age of 9 (mean age 5;11, oldest 7;2), they were classified for the purposes of further analysis as having age appropriate speech, and as having resolved their phonological difficulties.

The next chapter outlines the characteristics of the sample.

3 Chapter 3: Results 1: Characteristics of the sample

Investigation of the caseload characteristics of clinical populations can provide information that is useful for the planning and delivery of services. While most research into SSD takes place with groups of children who are specially selected for level of severity or absence of co-morbidities that might confound results, to provide insights useful for clinicians operating in the real world, research that looks at populations of clients who actually present for services is needed (Broomfield and Dodd 2005b).

This chapter is a cross-sectional study describing the characteristics of a sample of children who were referred to SLT services in three local areas, due to concerns about their speech. All children who presented with such a referral, and who fitted the inclusion criteria described in section 2.2.2, were invited to attend. Local SLTs acted as gatekeepers and obtained consent for referral to the study, so the characteristics of those who refused the invitation are not known.

The sample are described in terms of demographic variables, co-morbidity, severity of SSD and distribution across Dodd's Diagnostic sub-groups. The primary purpose is to compare the sample with similar samples in the literature in order to establish whether it is representative, such that further results might generalise to the larger population. The second purpose is to consider any factors that might suggest causal or maintaining factors in SSD of unknown cause, and the third is to provide information useful for planning and delivery of SLT services.

For ease of reading, discussion of each section will be included within the section, with concluding remarks to finish the chapter.

3.1 Number of children with and without SSD

Eleven children (13%) were shown on assessment to have age appropriate speech. This is similar to the 10% of inappropriate referrals found by Broomfield and Dodd (2004b) and may be higher because the children had been on the waiting list for assessment longer, allowing for some spontaneous improvement (all of Broomfield and Dodd's referrals were seen within 8 weeks, whereas the children in this sample may have been on the waiting list for up to 4 months). In a study of 354 children under 3;6 referred to

SLT services in the west of England reported by Roulstone *et al.* (2008), 26 (7%) were discharged immediately as having no speech or language problems, and a further 55 (16%) were discharged immediately with advice only, suggesting only sub-clinical difficulties. Seventy-four children (87%) in the current study were diagnosed with SSD of unknown cause, i.e. not related to any other condition, e.g. cerebral palsy, deafness, diagnosed intellectual impairment, anatomical or neurological abnormality, or being bilingual. As in Broomfield and Dodd's (2004b) larger study of a UK caseload of 1100 children, no child was found to have childhood apraxia of speech (CAS).

3.2 Demographic Characteristics

3.2.1 Age and gender

Of the 85 participants in the whole sample, 52 (61%) were male and 33 (39%) female. Of those with SSD, 44 were male (59.5%) and 30 female (40.5%). The higher proportion of males to females is commonly reported in studies of children with SSD, but this is not as high as in other samples, e.g. Fox (2001) found a ratio of 2 males to 1 female, Campbell (2003) found a ratio of 2.3 to 1. However it is comparable to other UK studies: Broomfield and Dodd's (2004a) sample comprised 62.8% male and 37.2% female, and Roulstone *et al.*'s. (2009) sample was almost identically distributed: 59.9% male to 40.1% female. Male gender as a risk factor for SSD has been reported in other large sample studies (Shriberg *et al.* 1999, Campbell *et al.* 2003, Delgado *et al.* 2005). Eleven of the 14 studies of risk factors for speech and/or language impairment reviewed by Harrison and McLeod (2010) found that male gender was a risk factor in preschool and early school age samples, with odds ratios between 1.29 and 1.97, and Shriberg *et al.* (1999) cites population-based studies that have recorded male to female ratios of between 2.1 and 3.1 to one, at 3 years of age. In clinical rather than population samples, at least, the preponderance of males may be due to more frequent referral due to social expectations: immature speech may be considered acceptable or "cute" in girls, but less likely so in boys; however some of the lower ratios are seen in clinical samples. It is also documented in several studies that speech development in girls is somewhat faster than in boys (see Dodd *et al.* 2003 for a review), so that a delay in girls might not be as evident, especially where non-gender specific norms are used to decide on status. Alternatively, male preponderance may be due to the general trend for increased vulnerability to developmental disorders in males (Kraemer 2000), as noted in the introduction, possibly because of increased genetic

risk, or gender related differences in rates of development/maturation. Severity in terms either of PCC raw score, or the DEAP Scaled Score (PCC adjusted for norms for age) did not differ significantly between males (mean PCC 65.4, mean DEAP Scaled Score 4.95) and females (mean PCC 69.6, mean DEAP Scaled Score 5.40), as shown by Mann-Whitney U tests: PCC, $U = 595$, $Z = -.716$, $p = .47$, DEAP Scaled Score, $U = 574$, $Z = -.978$, $p = .33$.

For both the whole sample and those with SSD, mean age was 48 months (minimum 36, maximum 69). Figure 3-1 shows that the majority of children were aged 3 years. As the inclusion criteria allowed only children aged 3;0 to 5;11 to be recruited (so that they were old enough to take part in testing, and to exclude the wide variability in onset of speech below this age), this may not represent the typical age at referral, which could be younger. However, in Broomfield and Dodd's (2004b) study of 320 children referred to SLT services for concerns regarding speech, less than 9% of children were referred under the age of 3 years.

Figure 3-1: Distribution of ages of participants

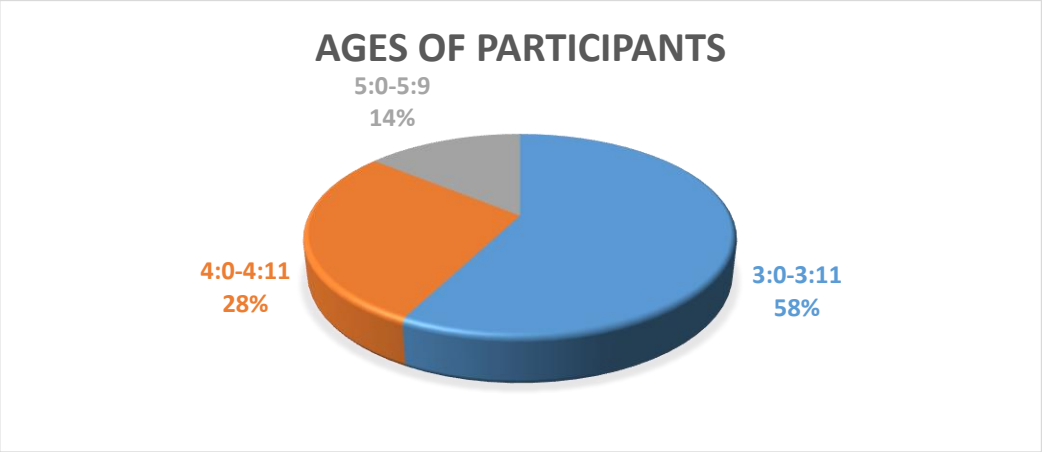


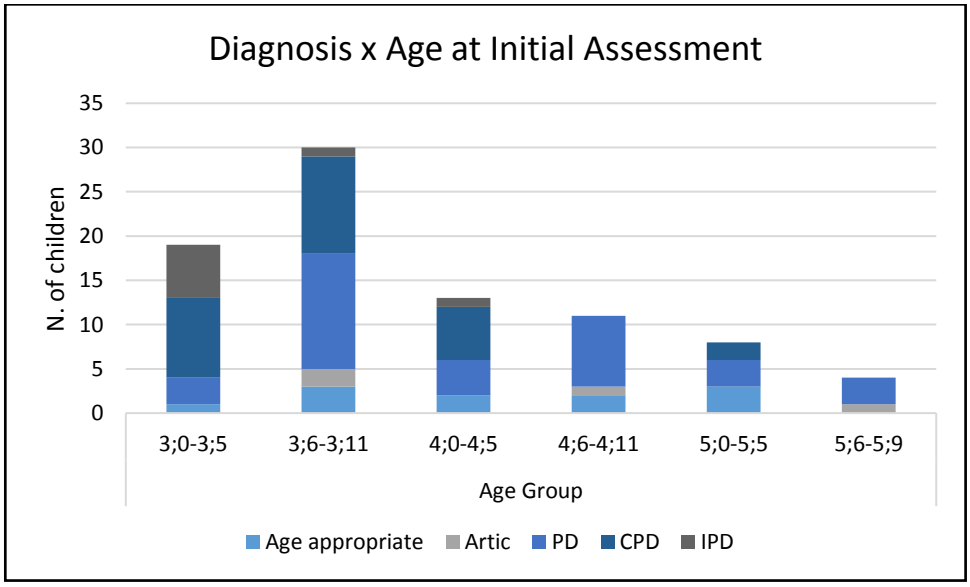
Figure 3-2 shows the breakdown by age group and diagnosis. Again, similar to Broomfield and Dodd (2004b), diagnoses of CPD and IPD were most prevalent between the ages of 3 and 4, with only 2 children over the age of 4;5 being diagnosed with a disorder as opposed to a delay. Broomfield and Dodd account for this with the explanation that children with disorder have a more severe involvement, and are generally more difficult to understand, so that they are referred earlier. An alternative explanation is that disordered error types become less prevalent as children get older.

As in both Broomfield and Dodd's study, and in Fox and Dodd (2001), the age distribution of the phonological subgroups is PD>CPD>IPD (see Table 3-1).

Table 3-1: Age at initial assessment by diagnostic sub-group

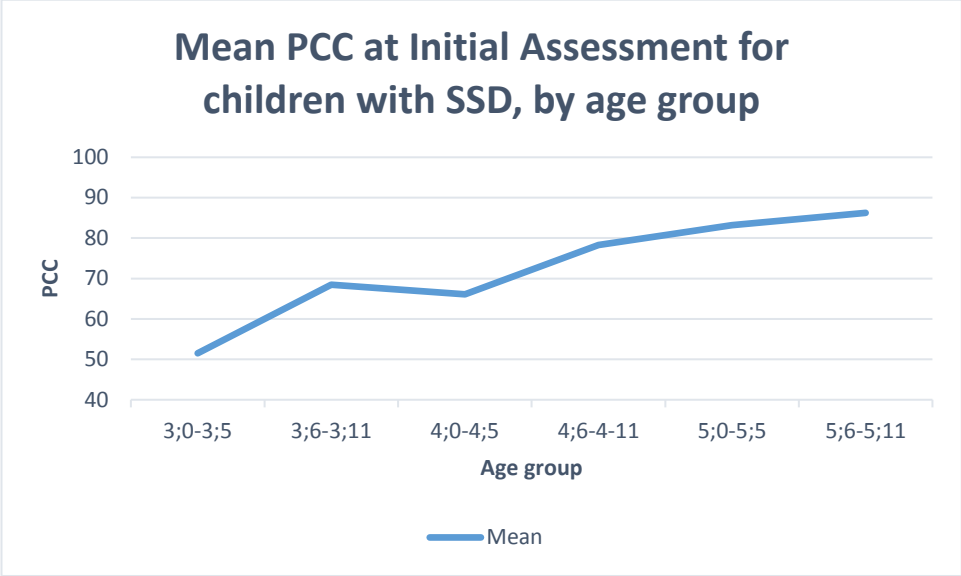
Age at Initial Assessment by Diagnostic sub-group					
Dodd classification	**N**	**Mean Age (months)**	**SD**	**Minimum**	**Maximum**
Age appropriate	11	53.00	8.246	41	65
Artic	4	54.50	10.970	46	69
PD	34	51.35	9.267	36	69
CPD	28	44.89	6.652	36	62
IPD	8	41.00	4.440	36	51
Total	85	48.61	8.907	36	69

Figure 3-2: Age Group and diagnosis at initial assessment



Increase in age was generally associated with increase in mean PCC for age group, with the steepest increase occurring in the youngest age group, a levelling out between the ages of 3;6 and 4;5, followed by another steep rise that gradually evens off, as shown in Figure 3-3.

Figure 3-3: Mean PCC at Initial Assessment, by age group



3.2.2 Family history of speech, language or literacy difficulties.

Forty two children (49%) across the whole sample had a family history of speech, language or literacy difficulties. Of those who had age appropriate speech at initial assessment, 63% had a positive family history. Of those with SSD, 47% had a positive family history. None of those with an articulation disorder only had a positive family history. Forty-four percent of children with SSD alone had a family history, compared to 68% of those with SSD+LI. The incidence of presence of family history of similar conditions is similar to the 56% reported by Shriberg and Kwiatkowski (1994) in their study of 84 children with SSD, the 58.5% reported by Brosseau-Lapré and Rvachew (2016) in their study of 72 children, and the 58% reported by Lewis and Freebairn (1997) in their study of 59 children, but higher than that reported for other studies, (Dodd & McIntosh 2008, 38%; Broomfield & Dodd 2004, 22.3%; Campbell *et al.* 2003, 28.1%, Fox *et al.* 2002, 28%). This may be due to the fact that, as in Brosseau-Lapré and Rvachew's (2016) study, the criteria included second degree relatives, rather than only first degree, and difficulties with literacy were included, rather than only speech and language difficulties. However, as such data is dependent on family member report, and requires recollection and knowledge of past events and conditions that may have long since normalised, the data may not be reliable, and reported data have been shown to both over and underestimate family involvement (Shriberg *et al.* 2005). It may also be that age plays a part by hiding the effect of genetic disposition in the wide variation of development at the youngest ages: of the studies cited here, that

with the youngest children (age 3 years, Campbell et al., 2003) reported the lowest rate of positive family history, whereas the study with the oldest children (age 5;2, Lewis and Freebairn, 1997) reported almost the highest rate. However the trend was not completely linear. Severity of impairment, measured by raw PCC or DEAP Scaled Score, did not differ between those with a positive family history (mean PCC 66.6, mean DEAP Scaled Score 5.18) and those without (mean PCC 67.6, mean DEAP Scaled Score 5.09).

3.2.3 Delayed onset of speech

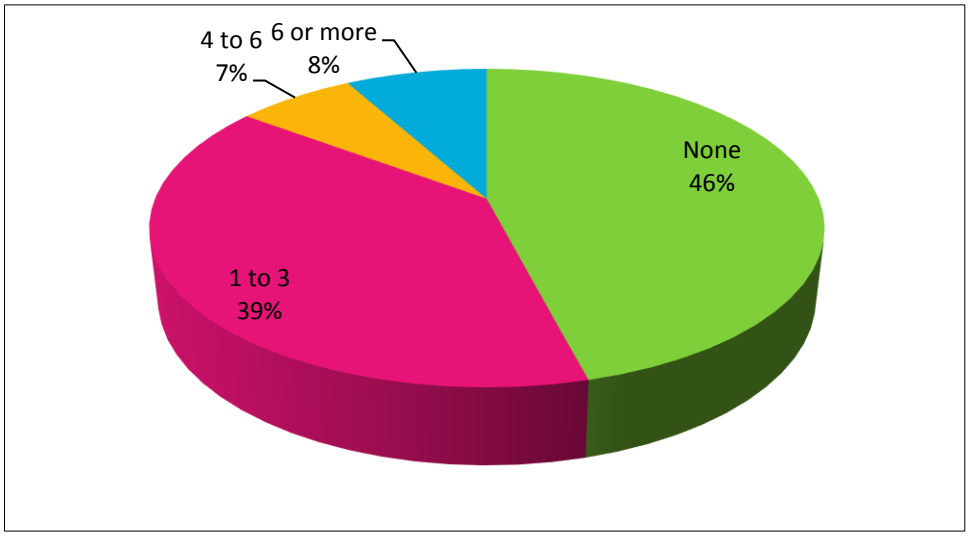
Twenty-two children (26%) were reported to have had delayed onset of speech. Two of these had age appropriate speech and language at initial assessment. Almost twice as many children with SSD+LI (42%) had a history of late-talking, as opposed to 22% of those with SSD alone. This is comparable with the (21.6%) found by Broomfield and Dodd (2004b) for children with SSD alone, and 30.5% for children with speech and/or language difficulties. The prevalence of late talking, as defined by, for example, failure to use at least 50 words, or to use 2 word combinations, at 24 months, has been reported to be around 10% to 15% of toddlers (Rescorla 1989, Hawa and Spanoudis 2014), so the higher figure in this sample supports evidence that a substantial number of them will continue to be at risk of poor outcomes for speech and language at 3 years of age (Paul and Jennings 1992, Dale *et al.* 2003), although they continue to catch up (Whitehouse *et al.* 2011). One late-talker had an articulation disorder and the remaining 19 had phonological impairments, again supporting evidence of slow phonological development in late talkers (Paul and Jennings 1992). Amongst those with SSD, the group of late talkers had significantly more severe impairment in terms of PCC (mean PCC 57.8) than who had not been late talkers (mean PCC 72, $U = 290$, $Z = -2.97$, $p = .003$), but as they were also significantly younger (mean age 44 months compared to 50 months for those with normal age of onset of speech, $U = 308$, $Z = -2.75$, $p = .006$), this was not significantly different for DEAP Scaled Score, i.e. adjusted for age (late talkers, DEAP Scaled Score 4.35, normal onset, DEAP Scaled Score 5.47, $U = 381$, $Z = -1.91$, $p = .056$).

3.2.4 Number of ear infections and hearing status

Three (27%) of the children with age appropriate speech and language and 11 (15%) of those with SSD had experienced recurrent ear infections (defined as more than three,

NHANES III 1988-94). Figure 3-4 below shows the number of ear infections reported for all children with SSD.

Figure 3-4: Number of ear infections reported, children with SSD



This figure is much smaller than the 43% to 48% reported for children aged 3 to 5 by Auinger *et al.* (2003) in the Third National Health and Nutrition Examination Survey in the United States, a survey of approximately 40, 000 children and adults representative of the population. This was also half the number (21.6%) reported by Broomfield and Dodd (2004). This may be affected by the fact that Broomfield and Dodd's study was carried out in a relatively deprived area of the UK, whereas, if levels of parent education can be taken as a proxy for SES in this study, the sample is representative of the country as a whole. It may also represent a better standard of health or access to health care than is true for the sample in the USA, where inequality is greater (OECD 2008).

The low incidence of ear infections in this sample of children with SSD accords with several other studies showing that otitis media with effusion (OME) does not increase the risk of SSD: see Fox et al (2002), Paradise et al.'s (2000) findings of no significant correlation between time with otitis media and scores on measures of speech sound production, Campbell et al.'s (2003) finding of no difference in number of days with OM between the speech delayed and normal group, and Shriberg, Friel-Patti *et al.'s* (2000) findings of no significant association between days with OME and increased risk for speech disorder, in a sample of 35 children followed from birth in a university paediatrics clinic. Shriberg, Flipsen *et al.* (2000) did find an association in one group

of subjects who were all children of Native American background enrolled in a Head Start program on a reservation, but not in another cohort of children followed since birth by a university paediatrics clinic. In a meta-analysis of 32 studies of the effect of otitis media on language development in young children, Casby (2001) found that the effects might be significant but had "markedly low" effect sizes. However, this meta-analysis did not include studies of speech development. Nevertheless, since not all children who have a history of otitis media go on to have SSD, and not all children with SSD have a history of otitis media, it is unlikely to have a causal effect (Bradford and Dodd 1996). In a study of 69 four-year-old children with developmental language disorders, Bishop and Edmundson (1986) compared those with and without a history of recurrent otitis media. They found no differences in severity or presentation of language disorder, but some evidence that otitis media might interact with perinatal risk factors, as both conditions tended to co-occur in the DLD group, but not in the control group. However, they found this only occurred in a minority of cases. As reported for Niclasen *et al.*'s. (2016) study described in section 1.5.6, it may be the otitis media only has an effect on speech development where it interacts with other social or health risk or protective factors, which might explain Shriberg, Flipsen *et al.'s* finding of an association in children enrolled in a Head Start Program, noted above.

In contrast to findings regarding otitis media, Shriberg, Friel-Patti *et al.* found that a diagnosis of hearing loss at 12-18 months was associated with approximately 10-21 times increased risk for later sub-clinical or clinical speech disorder. Only two children in this sample had had earlier diagnoses of hearing loss resolved by subsequent myringotomy: one had an original diagnosis of inconsistent speech disorder at age 3:4, which progressed to phonological delay at 5:5. The other was diagnosed as having a moderate phonological delay at 4:7 and had normal speech at 6:11.

3.2.5 Co-morbid language difficulties

None of the children with age appropriate speech had a language difficulty (as defined by a standard score of below 85, i.e. one SD below the mean, on the CELF-P2 (UK)) or poor receptive vocabulary (as defined by a standard score of below 85 on the BPVS III). Fifteen of the children with SSD (18%) had scores in the impaired range on the

CELF P2 (UK) and eight (9%) had poor receptive vocabulary on the BPVS, with four of these having below normal scores on both instruments. These 19 were identified as children with language difficulties, meaning 26% of children with SSD had co-morbid language difficulties. Thus, even with a relatively tolerant criterion for language difficulties, the sample of those defined as having co-morbid language impairment (LI) is small, limiting the generalisability of any results pertaining to this group. Table 3-2 shows the characteristics of those with LI, and scores on relevant assessments.

Table 3-2: Children classified as having co-morbid language impairment at initial assessment

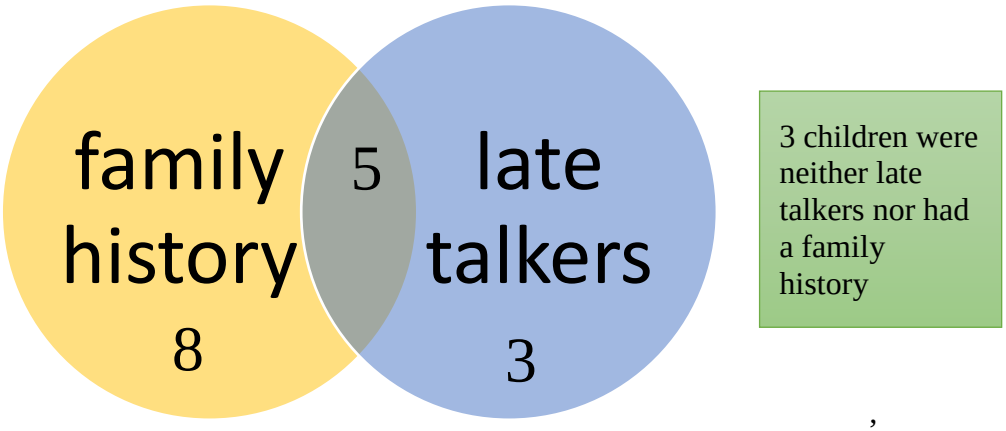
ID number	Gender	Age at Initial Ax (mths)	SSD Diagnosis at Initial Ax	Standard Scores	
				BPVS III	CELF SS
12317	F	36	IPD	86	73
11327	M	36	CPD	70	
11305	M	36	CPD	84	81
12312	F	36	PD	83	75
11312	M	40	IPD	87	55
11322	M	40	IPD	70	94
12318	F	43	PD	87	75
12306	F	44	CPD	99	67
2305	F	44	PD	81	100
11332	M	44	CPD	73	112
11323	M	46	CPD	87	71
11324	M	46	CPD	80	71
1403	M	48	CPD	101	67
2403	F	48	PD	92	81
12401	F	52	PD	90	83
11419	M	54	PD	76	67
11414	M	54	PD	87	75
11420	M	59	PD	95	79
12501	F	62	CPD	91	75

Shriberg *et al*. (1999), in a review of literature regarding the prevalence of co-morbid speech and language difficulties, found results ranging from 11% to 44% for studies of children with primary SSD in a similar age range. The results of the current study are somewhat lower than Broomfield and Dodd's findings that a third of children with SSD had co-morbid expressive language difficulties and half had semantic language difficulties. The difference may again be explained by the fact that their sample came from a disadvantaged area. However, it is also possible that there was a selection bias

in the sample reported here, with SLTs referring only clients who had a primary SSD, as this was the focus of this study. Three year olds with more severe language impairment may not have been referred as SLTs regarded them as not having enough speech for an assessment of SSD: indeed some may not have been talking at all. Thus it is probable that SLTs extended invitations to the study only to parents who expressed concerns about speech sounds, rather than to those who reported children having few words. In fact, of the 19 children classified as having a comorbid language disorder, only 3 fell into the severe range (i.e. a standard score of 70 or below) on the CELF PS2 (UK) core subtests, supporting the conjecture that SLTs were not referring children with more severe language impairment.

Figure 3-5 shows the number of children with co-morbid language difficulties who were late talkers or who had a family history of speech, language or literacy language difficulties.

Figure 3-5: Late talkers and family history in children with language i



,

Similar proportions of boys (25%) and girls (27%) had comorbid speech and language impairment, in contrast to the findings of Eadie *et al.* (2015), who found comorbid speech and language disorder was twice as common in boys as in girls. Their sample was unusual for a sample of children with SSD in that it contained fewer boys (25/51) than girls overall. The speech accuracy of children with and without co-morbid LI did not differ in this study, as shown in Table 3-3. This does not accord with other studies, which have found that children with SSD+LI have more severe SSD than those with SSD alone, e.g. (Nathan *et al.* 2004a, Sices *et al.* 2007). However, as discussed above, the sample in this study may not be representative of either a clinical or a population sample in regard to co-morbidity, due to referral bias.

Table 3-3: Speech accuracy of children with and without co-morbid LI

Comorbid Status	N	Measure	Mean	SD	Min	Max
SSD only	55	DEAP PCC	67.3	18.57	16	95
		DEAP Standard Score	5.02	2.22	3	10
SSD+LI	19	DEAP PCC	66.8	18.92	26	88
		DEAP Standard Score	5.47	2.25	3	9

3.2.6 Other factors

One child with age appropriate speech and language, one with SSD+LI and five with SSD alone were reported to have been delayed in reaching motor milestones. One of these had age appropriate speech and language and the remainder had phonological impairments. This figure of 8% of the sample with SSD having delayed motor milestones is much lower than the 51% found in 6 year olds with severe speech and language difficulties by Visscher *et al.* (2007). Wren *et al.* (2016) found that the performance of children with persistent SSD at age 8 on developmental scales for gross and fine motor skills at ages 18, 30 and 42 months was lower than in children without SSD, but odds ratios predicting SSD were all below 1.3. In contrast, Eadie *et al.* (2015) found that children whose parents expressed concern about limb movement at 24 months were more than 4 times as likely to have a SSD at age 4. In contrast, the figure of 8% of children whose parents reported concern about motor milestones being met is very similar to the 8.2% prevalence of parent concern about gross motor skills in the USA norming sample of the Parents' Evaluation of Developmental Status screener (Glascoe 1997), for TD children. However, it is higher than the 4.6% of parents who reported such concerns for a sample of 262 pre-schoolers aged between 18 months and 5 years in Australia, measured on the same instrument. More evidence is needed to confirm Eadie *et al.'s* findings. However, given the small proportion of the sample made up by children with motor difficulties, it may not help to any large degree with identification of children with SSD.

Eleven children (13%) were born pre-term (before 37 weeks) and/or needed special perinatal care, compared to the national preterm birth rate of 6% in Ireland in 2013 (Healthcare Pricing Office 2016). This accords with Fox et al. (2002) who found that

15% of children with isolated SSD had birth difficulties, their higher figure being accounted for by the fact that their criteria were somewhat wider, including induced birth. It is known that very immature preterm birth is a risk factor for language impairment, (Sansavini *et al.* 2010, Barre *et al.* 2011, Vohr 2014), but less is known about speech, or about less immature preterm births. An older study of 33 pre-term children whose corrected ages ranged from 9 to 22 months found delayed acquisition of speech sounds for the whole group, but this was more severe for the 21 who had suffered intraventricular haemorrhage (Byers Brown *et al.* 1986). Improvements in management of preterm children over the last 30 years may have lessened this risk. A recent review of studies of speech and language outcomes of very preterm infants (Vohr 2014) included only one that assessed speech: Wolke *et al.* (2008), in a study of 198 6 year olds who had been born before 26 weeks in UK and Ireland, found that the preterm children were 2.6 times more likely to have scores 2SD or more below the mean on an articulation screener, even when they controlled for general cognitive ability. However, in a study of 55 five year olds who were less immature, that is, born at or before 32 weeks of gestation, Luoma *et al.* (1998) found that when children with frank neurological difficulties were excluded from the sample, leaving 46 children, there were no significant differences from age-matched controls on measures of speech. Thus it may be that risk of speech impairment is only significant for very immature births, and lessens as gestational age increases, as was found by Kerstjens *et al.* (2011) for communication and other developmental abilities in general.

3.2.7 Parental education

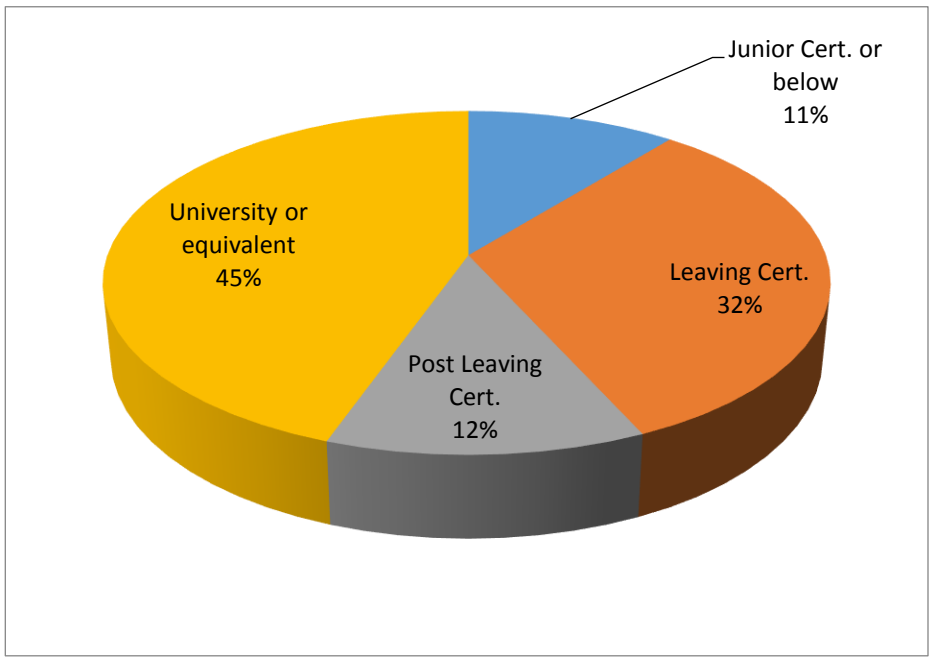
Socioeconomic status (SES) is a multidimensional and complex concept encompassing measures of economic position and social status that are used to place individuals in a social hierarchy. An individual's position in this hierarchy is an indication of their access to financial and other material resources, human capital, for example, educational opportunities, and social capital which may arise from social relationships and be reflected by such measures as number of parents in the home (Bradley and Corwyn 2002). It has been shown in many studies that SES is associated with child development in terms of health, cognitive, personal and social outcomes (Braveman *et al.* 2005), due to its association with access to resources, housing and environmental conditions, and stress, behaviour and lifestyle (Adler and Newman 2002), so it must be taken into account in studies of outcomes for children. However, because of the variety

of factors that make up SES, it is difficult to measure. Measures used include income, occupation, address, housing conditions and education, but all of these are problematic to a degree (Braveman *et al.* 2005, Shavers 2007).

In this study parental education, especially highest level of maternal education, was the preferred measure used. Local SLT services had ceased asking questions related to occupation on case history forms as it was felt to be intrusive, and measures of income were considered too sensitive. Address status is also an unreliable measure of SES outside of urban areas in Ireland, and the sample in this study included many from rural areas. Maternal education also has the advantage of being accurately reported (Bornstein *et al.* 2003), tends to be stable, unlike occupation and/or income for mothers with young children, and was an available measure for all households of children in the sample. It is also associated with environmental variables such as nutrition and healthcare, and availability of books and toys in the home to provide a stimulating environment (Entwisle and Astone 1994), as well as knowledge of child development, and parenting and communication styles (Benasich and BrooksGunn 1996).

Eleven per cent of mothers in the whole sample left school at Junior Certificate or below, at age 15-16 years, and 33% after Leaving Certificate at age 17-18 years. The highest level of education for a further 13% was completing a Post Leaving Certificate course such as a vocational or technical qualification at Level 4 or 5 of the European Qualifications Framework (European Commission 2014), and 43% completed University or equivalent. Figure 3-6 shows the breakdown of highest level of education for mothers of children with SSD

Figure 3-6: Highest level of education in mothers of children with SSD



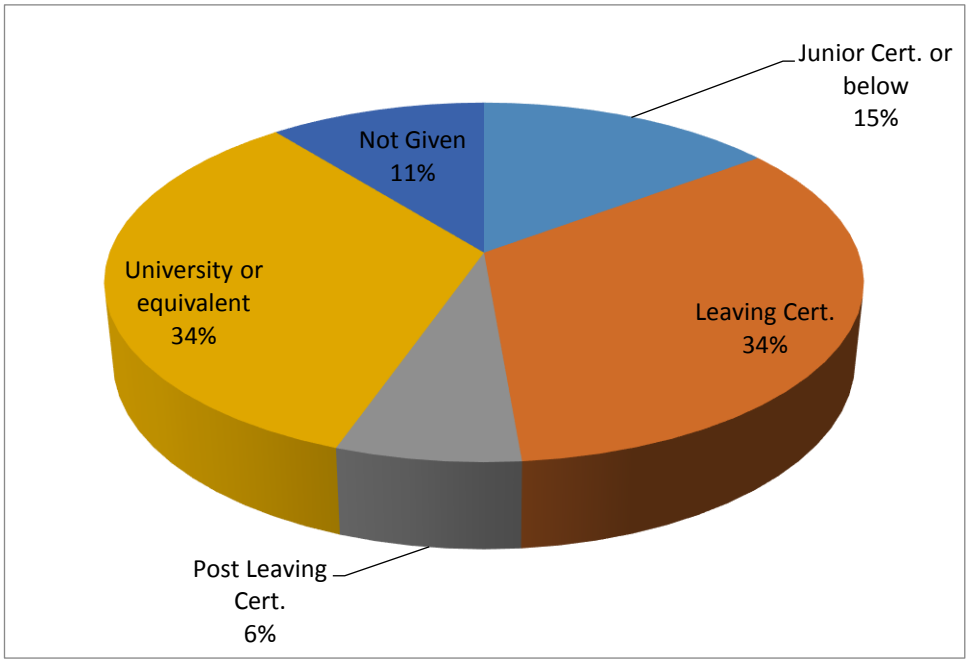
**Junior Certificate =European Qualifications Framework (EQF) level 2, Leaving Certificate = EQF levels 3&4, Post Leaving Certificate = EQF levels 4&5, University = EQF level 6. (National Qualifications Authority of Ireland, 2009)*.

The proportion of post second level educated mothers is much larger than that reported in other studies. In Campbell *et al.'s* (2003) sample, only 7.5% of mothers had completed college, 79.6% had completed high school and 12.8% had not completed high school. Similarly Dollohan *et al*. (1999) reported that in their sample of mothers, 20% had completed college, 70% had completed high school and 10% had not completed high school. However, the distribution in the sample reported here reflects that of the general population in the Republic of Ireland, where just over 45% of women in the age range 25-34 were reported to have a University or equivalent qualification in 2005 (Department of Education and Science, 2007). At 11%, the proportion of mothers of children with SSD who left school before completing the Leaving Certificate is lower than in the general population: in 1994 15% of females in Ireland left school (aged ~16) before the Leaving Certificate, in 2002 it was 13%. (Department of Education and Science, 2007).

Sixteen per cent of fathers in the whole sample left school at Junior Certificate or below, 34% after Leaving Certificate, 7% left after completing a Post Leaving Certificate course and 32% completed University or equivalent. Information was not

given for 11%. The distribution of highest level of paternal education of children with SSD is shown in Figure 3-7 below.

Figure 3-7: Highest level of education in fathers of children with SSD



**Junior Certificate =European Qualifications Framework (EQF) level 2, Leaving Certificate = EQF levels 3&4, Post Leaving Certificate = EQF levels 4&5, University = EQF level 6. (National Qualifications Authority of Ireland, 2009)*

This is similar to the distribution of highest level of education of the population of Ireland as a whole, where 22% of males left at Junior Certificate or below (aged ~16) in 2002, and just under 35% graduated with University or equivalent level qualifications in 2005 (Department of Education and Science, 2007).

The fact that the distribution of highest levels of parental education in a sample of children with SSD reflects the distribution in the general population accords with other findings regarding levels of parental education: Dollaghan *et al*. (1990) found no relationship between level of maternal education and PCC in 240 normally developing three year olds. Harrison and McLeod (2010) in a study of 4,983 children participating in the Longitudinal Study of Australian Children found that low level of parent education was not a risk factor for parent reported concern regarding expressive speech and language, but mothers and fathers having completed more years of education was a protective factor, indicating perhaps that parents with a high level of education were able to access resources to mitigate any developmental problems which occurred early

on. To *et al*. (2013) in a study of 937 Cantonese-speaking children aged 2;4 to 6;7 in Hong Kong also found that maternal education level made only a very small contribution to children's speech acquisition, and paternal education level made no significant contribution. Smit *et al.* (1990), in the Iowa and Nebraska Articulation Norms study of almost 1000 children (including a representative proportion with SSD) similarly found that SES had no significant effect on age of acquisition of phonemes. They also used parental education as the measure of SES, and children were drawn from both rural and urban environments. A study that took place in an area of disadvantage in Dublin, recruited 95 pre-school children from 3 pre-schools, 2 of which catered specifically for children with social emotional and behavioural difficulties and/or children from the Travelling Community (a disadvantaged community similar to, though from different roots than, communities known as "gypsies" in the UK and USA). They found that speech scores followed the normal distribution of the population, showing no effect of disadvantage, although language scores were lower than norms for age, as expected in disadvantaged communities (Hayes *et al.* 2012).

Associations between SES and speech development are also absent when different measures of SES are used: Broomfield and Dodd (2004b) measured SES according to the deprivation index of the district of residence, and found that the distribution of children with SSD across socio-economic status was similar to that of the local population, with a slightly higher number of children from affluent backgrounds. Keating *et al.* (2001), using data from the Australian 1995 National Health Census, found no association between SSD and SES whether this was defined by income, employment status, occupation or highest level of education.

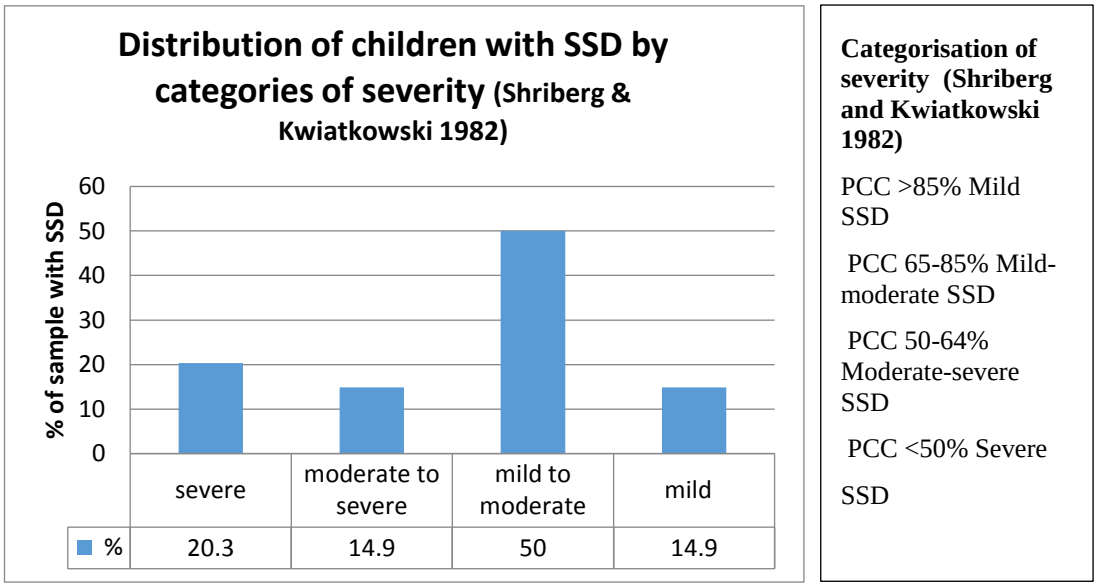
Nevertheless, as mentioned in sections 1.5.2, and 1.5.6.3, some studies have found an association between SES and SSD in children. Campbell et al., (2003), found that low maternal education increased the odds ratio of speech delay to 2.58, but this may reflect the difference in educational experiences between the Republic of Ireland and the USA, as Irish children are more likely to complete second level school education than children from the USA (90.6% of Irish children completed the Leaving Certificate at age 17/18 in 2013/14 (DES 2015), compared to 82% of children in the USA who graduated from public high school in 2014 (Kena *et al.* 2016)). Lower SES was also associated with persistent SSD at older ages in population prevalence studies by Wren

et al. (2012) and Wren et al. (2016). This might reflect poorer access to remedial services for children with low SES, or they may be an effect of intergenerational familial disadvantage, where parents of children with SSD are themselves at higher risk for language and reading difficulties, and consequent lower levels of occupation, income and education (Broomfield and Dodd 2005b, Peterson *et al.* 2009). Results regarding SES that come from different countries may also differ due to different rates of inequality. For example, in the first decade of the 21st century, inequality of income in the UK and Ireland was less than 9% above the average rate for OECD countries, whereas the rate of inequality in the USA was much higher at almost 24% above the average (OECD 2008).

3.3 Speech accuracy/severity of impairment

Figure 3-8 shows the sample of children with SSD broken down by Shriberg's classification of severity. In Shriberg and Kwiatkowski's (1994) sample of 145 children with SSD aged 3-7;11, combined from four studies, 4% were classified as Mild, 54% as Mild to Moderate, 32% as Moderate to Severe and 10% as Severe. In comparison, children in this study comprise more at the Severe and Mild ends, and fewer in the middle. However, there was also considerable variability among Shriberg and Kwiatkowski's 4 studies, with the lowest and highest percentage in each category as follows:

Figure 3-8: Distribution of severity in SSD by Shriberg's system



Mild, 0-10%, Mild to Moderate 33-71%, Moderate to Severe, 16-53% and Severe, 3-15%. The children were recruited in a similar way to this study, i.e. by asking local clinicians to refer them, but the inclusion criteria were children with SSD of unknown origin severe enough to interfere with intelligibility, which accounts for there being fewer Mild children than in this study, where children were referred from the waiting list for assessment, and therefore could be of any level of involvement. Shriberg and Kwiatkowski based their classification categories on perceptual judgements by a large group of clinicians of conversational samples of speech from children with SSD. However the children they used in the study were aged 4;1 to 8;6, so it is likely that it may not apply to younger children, as seen in this sample. It may also not apply to PCC calculated from single word samples, although Masterson *et al.* (2005) showed that PCC calculated from a single word task were not significantly different from those derived from conversation samples.

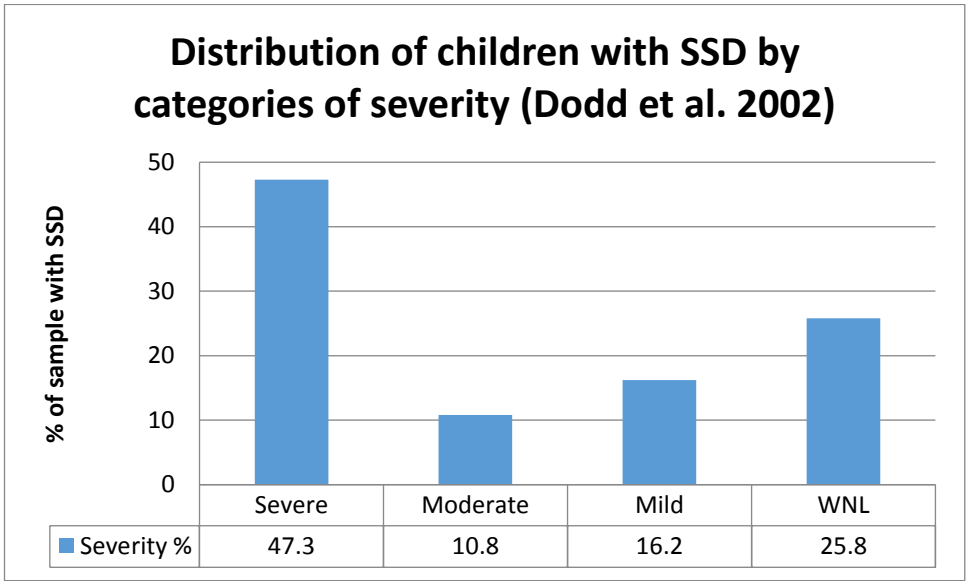
However, as PCC normally changes with age, with or without SSD, a scaled score of PCC adjusted for norms for age may be a more accurate indicator of severity than a PCC raw score. For example, at 3 years of age, Irish norms show that PCC scores between 77% and 92% are within one standard deviation of the mean. According to Shriberg and Kwiatkowski's classification, that would put some children with age appropriate speech into the mild to moderate range of SSD. Therefore, each participant was also categorised for severity of speech sound disorder according to the DEAP norms for children in Ireland, as in Table 3-4.

Figure 3-9 shows the distribution of the sample in this study by severity, calculated by the DEAP scaled score (Dodd *et al.* 2002). As the categorisation depends on norms for age, classes of severity in terms of PCC differ with age, i.e. a score of 70% might put a 3-year-old in the mild range, but a five-year-old in the severe range.

Table 3-4: Categorisation of severity on the DEAP

DEAP PCC Scaled Score	Percentile Rank	Severity grade	Description
7-13	16-99	WNL	Within Normal Limits: within one SD of the mean (SS 7-13) or above the normal range (SS 14-17)
6	9	Mild	>1SD <1.5 SD below the mean
5	5	Moderate	>1.5 SD <2 SD below the mean
4	2	Severe	>2 SD below the mean
3	1		

Figure 3-9: Distribution of children with SSD by categories of severity (Dodd et al. 2002)



It will be noted that a considerable number of children appear to have normal PCC for their age, according to Dodd's classification, despite a diagnosis of SSD. These are the proportion whose speech accuracy in terms of PCC is within normal limits for age, but who still demonstrate use of delayed or disordered processes in their speech. This demonstrates that quantitative measures of speech alone are not sufficient to identify or categorise SSD, as two samples with the same PCC might have quite different qualitative errors. For example, a high number of atypical omission errors would affect intelligibility much more than the same number of typical substitution errors. It also shows how clinical samples may differ from research samples, which often have

inclusion criteria that would exclude this section of quite mildly affected children, whose speech is nonetheless of concern.

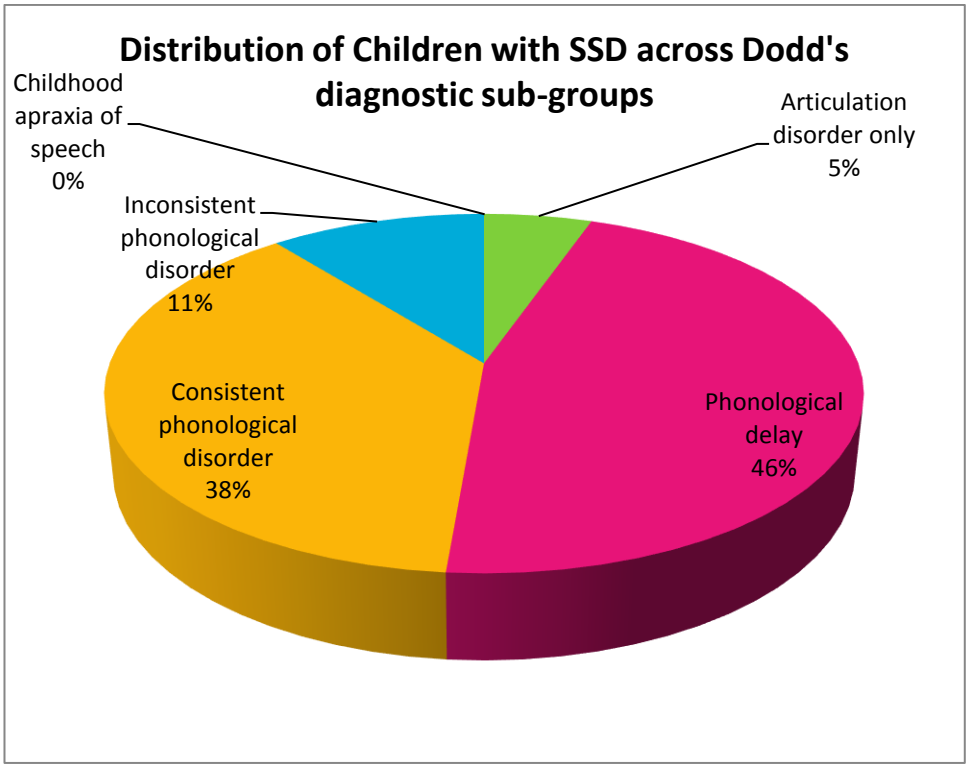
Severity ratings are important to show progress, either in therapy, or in the natural course of speech development over time. Normative ratings also help decisions about the presence of a SSD, although, as seen in the data above, they do not provide the only criterion. Severity ratings are also used to qualify children for services, or to decide on prioritisation for service. However, as discussed in section 1.5.8, at young ages PCC is not a predictor of improvement, i.e. children with low PCC at one point can catch up with normal peers.

There is little evidence regarding severity ratings in other clinical samples with which the distribution in this sample can be compared. In a study of 143 four to five year old children whose parents or teachers had reported concerns regarding their children's speech McLeod *et al.* (2013) found that, when classified using Shriberg's system, 9.1% were severe, 28.0% moderate-severe, 52.4% mild-moderate, and 10.5% were mild, which compares relatively well with the 20% severe, 15% moderate to severe, 50% mild to moderate and 15% mild in this study, using that classification. However, as there is no screening for speech difficulties in Australia, a large number of these had not been referred to SLT services. When only the 41 children who had attended SLT services were considered, they reported that 49% were mild to moderate, 40% moderate to severe and 14% were severe, which is difficult to interpret as it adds up to more than 100%. Children in Ireland are screened for speech and language development at the 2-year-old health check by public health nurses, and the vast majority of referrals come from that source, so it may be that the clinical sample in Ireland reflects the distribution in the community sample better than in the Australian clinical sample. Broomfield and Dodd (2004b)'s study classified severity using a different system, the Therapy Outcome Measures (TOM) impairment scale (Enderby and John 1997) which includes measures of the impact of impairment as well as the level of impairment. They found that 0.6% of children attending SLT services had impairment on this scale in the profound range, 11.3% in the severe range, 21.3% in the moderate to severe range, 44.4% in the moderate range and 22.5% in the mild range. This again shows the same trend of distribution of severity, with a smaller number at the severe and mild ends of the scale, and about half the sample in the moderate range.

3.4 Distribution across Dodd's diagnostic categories

Children were assigned a diagnosis according to Dodd's classification. Figure 3-10 shows the breakdown into diagnostic categories as a percentage of the sample of children with SSD.

Figure 3-10: Distribution of Children with SSD across Dodd's diagnostic sub-groups



As in Broomfield & Dodd's (2004) study, no child was found to fit the criteria for a diagnosis of Childhood Apraxia of Speech. However, these figures differ somewhat from those from their study with regard to the other diagnostic categories. In their sample of 320 children with speech sound disorders, they found the breakdown to be articulation disorder, 12.5%, Phonological Delay (PD) 57.5%, Consistent Phonological Disorder (CPD) 20.6% and Inconsistent Phonological Disorder (ICD), 9.4%. In Fox and Dodd's (2001) study of 100 German-speaking children with SSD, the breakdown was 20% articulation, 51% PD, 17% CPD and 12% IPD.

The low figure for articulation disorder in this sample can be accounted for by the fact that 57.5 per cent of the children with articulation disorder in Broomfield & Dodd's study were over six years of age and therefore outside the age range for this sample. In Fox and Dodd's sample, they account for the high occurrence of articulation disorder

with the fact that the data was collected in private practice, where more children with articulation disorder are referred because they are not prioritised in the public service. They also suggest that the majority of errors classed as articulation disorders, i.e. interdental productions of /s/ and /z/, may in fact be normal variation at that age, which would reduce the prevalence of articulation disorder to 5%. In Broomfield and Dodd's study, ten per cent of the children with phonological delay were also outside the age range, compared with only 7.6 per cent of those with consistent phonological disorder, accounting for a small amount of the difference in samples. Figures for IPD are roughly similar in all samples, but a bigger proportion of the sample in the current study present with CPD. However, the overall direction of size of groups as PD>CPD>IPD>articulation disorder is the same as that found by Broomfield and Dodd for English, Fox and Dodd (2001) for German, and other studies in different languages (So and Dodd 1994, Zhu and Dodd 2000). Table 3-5 shows the characteristics of children in each sub-group.

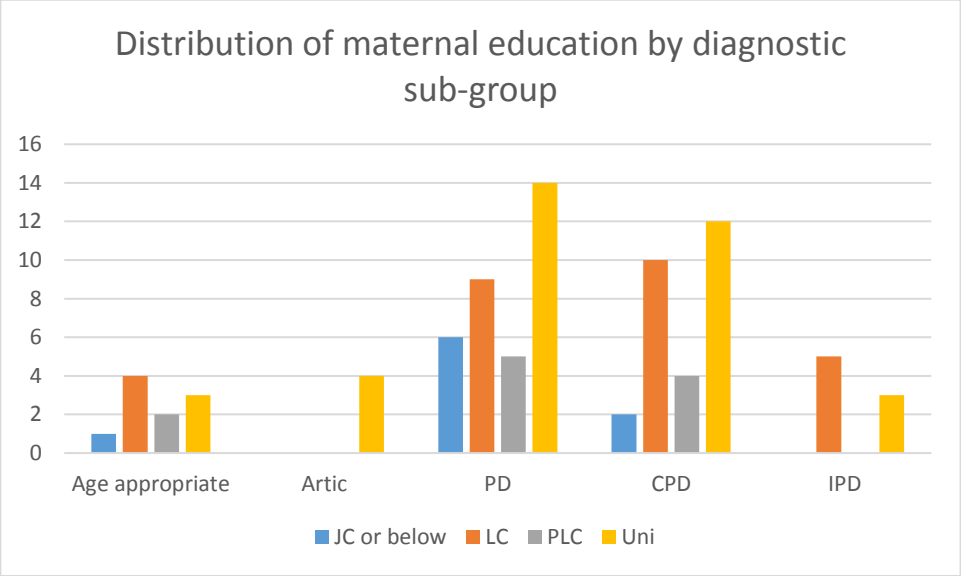
Table 3-5: Characteristics of children in each diagnostic sub-group

		Dodd classification										
		AA		**Artic**		**PD**		**CPD**		**IPD**		**Total**
		N	%	N	%	N	%	N	%	N	%	N
Gender	F	3	27	1	25	16	47	11	39	2	25	33
	M	8	73	3	75	18	53	17	61	6	75	52
Family History	No	4	36	4	100	17	50	16	57	2	25	43
	Yes	7	64	0	0	17	50	12	43	6	75	42
Motor Delay	No	10	91	4	100	30	88	26	93	8	100	78
	Yes	1	9	0	0	4	12	2	7	0	0	7
SL Delay	No	9	82	3	75	26	77	21	75	3	43	62
	Yes	2	18	1	25	8	23	7	25	4	57	22
Perinatal difficulties	No	11	100	3	75	29	85	25	89	6	75	74
	Yes	0	0	1	25	5	15	3	11	2	25	11
No. ear infections	None	4	36	3	75	13	38	14	50	4	50	38
	1 to 3	4	36	1	25	17	50	11	39	0	0	33
	4 to 6	3	27	0	0	2	6	1	4	2	25	8
	6+	0	0	0	0	2	6	2	7	2	25	6
Co-morbid language impairment	No	11	100	4	100	26	77	20	71	5	63	66
	Yes	0	0	0	0	8	23	8	29	3	37	19
Mean Age months (SD)		53 (8.2)		55 (11)		51 (9.3)		45 (6.7)		41 (4.4)		49 (8.9)
Mean PCC (SD)		91 (6.3)		82 (2.4)		76 (13.9)		62 (16)		41 (16.6)		70 (19.2)
Mean DEAP Standard Score (SD)		10 (2.1)		6.5 (2.6)		5.9 (2.2)		4.4 (2)		3.6 (1.2)		5.8 (2.7)

This shows that in each diagnostic sub-group, the ratio of boys to girls is between 2:1 and 3:1, except for the PD group, which was evenly balanced. None of the children with articulation disorder had a family history of speech, language or literacy disorders; as in Broomfield and Dodd's (2005) study, the IPD group had the highest ratio in favour of family history at 3:1, but the other groups were evenly balanced. Over half of the children with a history of delayed motor skills were in the PD group. The ratio of those with a history of speech and language delay was approximately 1:3 in each group except IPD, where more than half were reported to have delayed onset of speech. No children with age appropriate speech had a history of perinatal difficulties; the ratio of those with articulation disorder and IPD was highest at 1:3, and lower at 1:6 and 1:8 for those with PD and CPD respectively. Fox *et al.* (2002) also showed a higher ratio of perinatal difficulties in children with IPD. Children whose speech was age appropriate had the highest number of ear infections, followed by those with PD, whereas those with articulation disorder had the fewest. In the CPD and IPD groups, half had experienced at least one ear infection, with 50% of the IPD group experiencing recurrent infections. However, Chi-square tests showed none of the above differences between groups was significant.

Again, in accord with other studies showing breakdown into these subgroups, the severity of involvement in terms of number of errors was IPD>CPD>PD>Articulation disorder. The pattern for occurrence of co-morbid language impairment was the same. As in Broomfield and Dodd (2005b), no child with articulation disorder showed a co-morbid language impairment, and the IPD group showed the highest occurrence of language impairment, but a Chi-square test showed differences were not significant. Figure 3-11 shows the distribution of maternal education by diagnostic sub-group. As in Broomfield and Dodd's (2005) study, a Chi-square test showed that there was no significant trend in SES for sub-group.

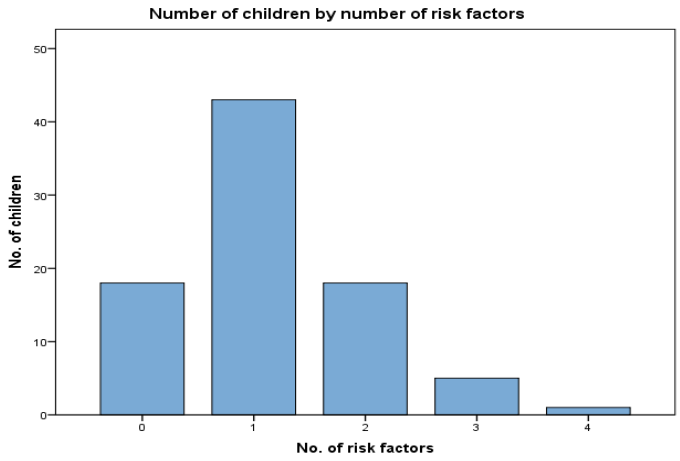
Figure 3-11: Distribution of maternal education by diagnostic sub-group



**JC=Junior Certificate =European Qualifications Framework (EQF) level 2, LC=Leaving Certificate = EQF levels 3&4, PLC=Post Leaving Certificate = EQF levels 4&5, Uni=University = EQF level 6. (National Qualifications Authority of Ireland, 2009)*

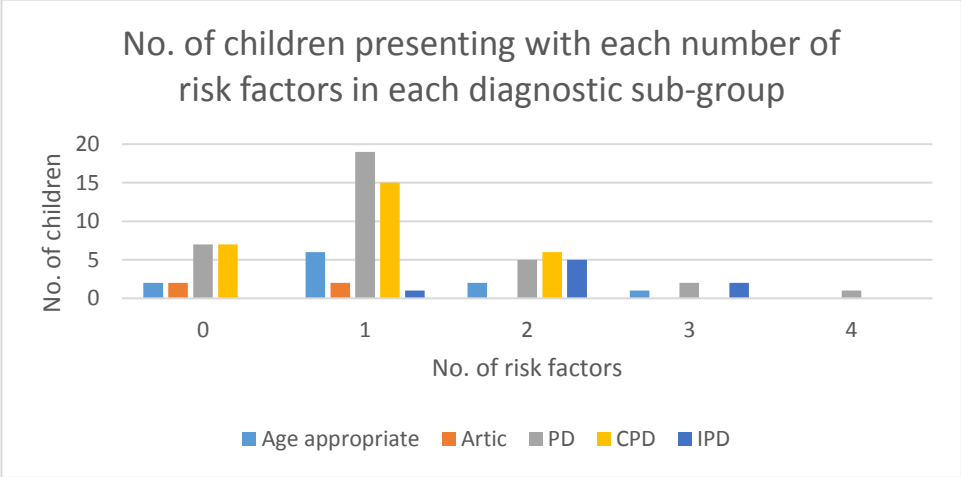
Overall, the IPD group showed the greatest incidence of risk factors: being male, having co-morbid LI, having a history of perinatal difficulties, delayed onset of speech, positive family history and highest number of ear infections, but the small number in the group precludes drawing any conclusions. Apart from male gender, no case history characteristic was demonstrated by more than half of the sample as a whole, limiting the value of such factors in identifying children at risk of SSD. However it may be that these factors interact, so that the presence of multiple risk factors, as in the group with IPD, or a certain combination of risk factors predict a greater likelihood or greater severity of SSD. To investigate whether having multiple risk factors was associated with a greater degree of severity, the number of risk factors from male gender, family history of SLCN, delayed onset of speech, delayed motor milestones, history of perinatal or hearing difficulties, and recurrent ear infections was calculated for each child. Figure 3-12 shows most children had only one risk factor, and very few had more than two. There was no correlation between number of risk factors and severity of SSD.

Figure 3-12: No. of risk factors



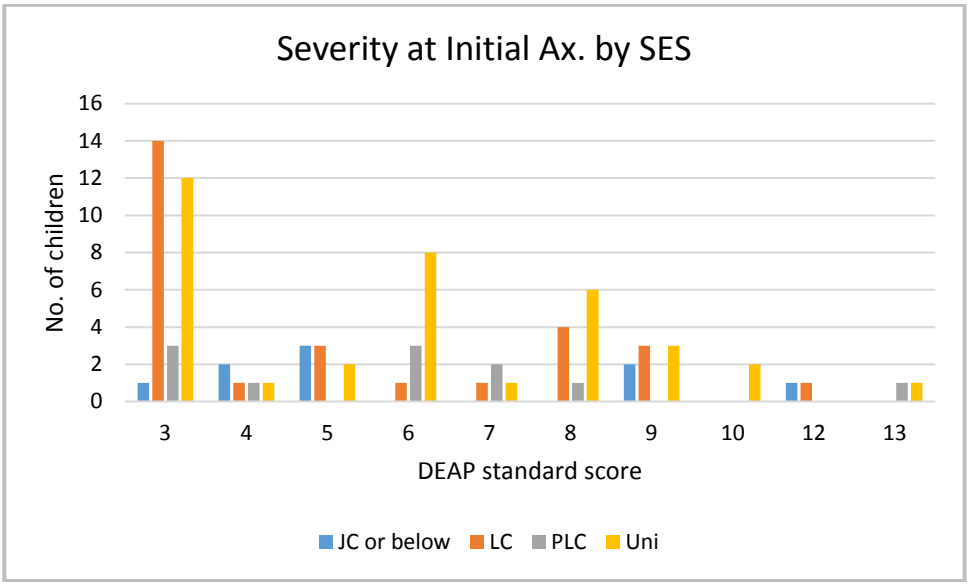
A Chi-square test showed no association between number of risk factors and diagnostic sub-group, as shown in Figure 3-13.

Figure 3-13: No. of risk factors by diagnostic sub-group



Interactions with SES might shed light on the role of environmental factors in SSD. Brroomfield and Dodd (2005) found an association of lower SES with greater severity of SSD. However in this sample, Figure 3-14 illustrates that the opposite association was found in this sample, with milder severity shown for children whose mothers left school at the age of 15/16 or below, χ^2 (27) = 41.08, p =.041.

Figure 3-14: Severity of SSD by maternal education



Broomfield and Dodd (2005b) also found an interaction between SES, gender, and family history, which might implicate a genetic disposition to SSD. They posited that intergenerational disadvantage might occur as families with a history of communication impairments gravitate to lower socio-economic strata, due to the consequences of SSD for social and academic attainment. In this sample, however, Chi-square tests showed no association between level of maternal education and gender or family history. Differences may be due to different ways of measuring SES: Broomfield and Dodd used the Index of Multiple Deprivation, which incorporates a large range of measures including employment, housing, health, car ownership and family status. Their study also took place in an area of high deprivation, which was not the case for this sample.

3.4.1 Underlying deficits for different diagnostic sub-groups

Studies by Dodd and colleagues have found that children with CPD perform more poorly on tests of rule abstraction and cognitive flexibility (Crosbie *et al.* 2009, Dodd 2011) than children with PD, suggesting that CPD group have cognitive-linguistic deficits at a central processing level which underlie their speech difficulties. Results on the Flexible Item Selection Test (Jacques and Zelazo 2001), the same test used by Dodd and colleagues to measure rule abstraction and cognitive flexibility, were analysed to establish if there were different patterns of performance for the different

diagnostic sub-groups in this sample. Mean scores on each measure for each diagnosic sub-group are shown in Table 3-6 below.

Table 3-6: Scores on the Flexible Item Selection Test, by Diagnostic sub-group

Dodd classification		**FIST score 1**	**FIST score 2**	**Difference FIST1 - FIST2**	**Age at Initial Ax (mths)**
Age appropriate n=9	**Mean**	**14.11**	**11.67**	**2.44**	**54**
	SD	1.691	3.041	2.78	8.602
	Range	10-15	6-15	0-7	41-65
Artic n= 4	**Mean**	**14.75**	**6.25**	**8.5**	54.5
	SD	0.5	6.131	5.92	10.97
	Range	14-15	0-12	3-15	46-69
PD n=34	**Mean**	**13.56**	**10.09**	**3.47**	**51**
	SD	2.765	3.098	2.74	9.267
	Range	0-15	0-15	0-9	36-69
CPD n=24	Mean	**11.58**	**7.79**	**3.79**	**45**
	SD	4.615	4.987	3.54	6.785
	Range	0-15	0-14	1-11	36-62
IPD n=7	**Mean**	**9**	**7.29**	**1.71**	**41**
	SD	5.657	6.102	1.98	4.271
	Range	2-15	0-14	-1-5	38-51

The means for each score for the PD group are strikingly similar to the findings of Dodd (2011), although the subjects' mean age is 4 months younger. However, the CPD group in this study were 6 months younger than Dodd's CPD group, and their scores on FIST1 were almost 2 points worse, and 1 point worse on FIST2. Therefore the difference between the FIST1 and FIST2 scores, which Dodd (2011) considers to be the measure of cognitive flexibility, was smaller in this CPD sample (3.79) than Dodd's CPD sample (5.44) and less than 1 point different from the PD group in both samples. In this sample, a Kruskal-Wallis test indicated that there was, in fact, no difference between diagnostic sub-groups on this measure, χ^2 (2) =2.29, p=.318. However, the measure itself may not be a good measure of cognitive flexibility: using this calculation, a child who scored 1 on both selections attained a better score for cognitive flexibility than a child who scored 15 on the first selection and 13 on the second. It is illogical to suggest that a child scoring 6 on the first selection and 4 on the second has better cognitive flexibility than one who scores the maximum 15 on the first selection and 12 on the second.

Dodd (2011) did not include children with IPD in the sample, but Crosbie *et al.* (2009) compared FIST scores for PD, CPD and IPD groups of older children (mean ages in months: PD, 62; CPD, 63; IPD, 65). They found that children with IPD scored similarly to children with PD, which was not the case in this sample, where the IPD children performed worst of all. However, the mean age of the IPD group in this sample was 42 months, almost 2 years younger than Crosbie et al's sample, so it is likely that age was a confounding factor, as pre-schoolers' performance on such tests of executive function have been shown to be age related (Carlson 2005). The small sample size of IPD children may also be an issue. Alternatively, deficits in children who still show inconsistency or atypical errors at age 5 may be different to those in children who show those error types at age 3-4, with only those who have specific deficits going on to have persisting SSD of such types.

Inspection of the range of scores, both in this and Dodd's (2011) sample, also indicated that there was overlap between diagnostic sub-groups, such that at least some of the CPD children scored as well as some of the PD children, with some attaining the maximum score, suggesting that not all children with CPD have deficits in rule abstraction and cognitive flexibility as measured by this test. In addition, some PD children scored at floor, indicating that poor rule abstraction and cognitive flexibility may not be characteristic of only one diagnostic sub-group.

In fact, it was the children with co-morbid language impairment who performed significantly worse than those with SSD alone on this assessment, as shown in Table 3-7. The two groups did not differ significantly in age. This accords with recent studies showing deficits in domain general executive function in children with language impairment (Kapa *et al.* 2017, Yang and Gray 2017). However, neither Dodd (2011) nor Crosbie et al. (2009) included children with LI in their samples, so this factor does not confound their results.

Table 3-7: Scores on the Flexible Item Selection Test, by language status

Mean Scores on the Flexible Item Selection Test, by language status				
Comorbidity Status		FIST score 1	FIST score 2	Age at Initial Ax
SSD only N=52	Mean (SD)	13.44 (2.6)	9.85 (3.7)	48.8 (9.1)
SSD+LI N=17	Mean (SD)	9.53 (5.9)	5.53 (5.0)	46.8 (7.4)
	Mann-Whitney U	U = 281, Z = -2.34, p = .019	U = 224, Z = -3.05, p = .002	U = 394, Z = -.66, p = .508

3.5 Conclusion: Characteristics of the sample

Large variation in findings across samples reporting the characteristics of children with SSD is apparent, but the demographic and diagnostic characteristics of the sample of children in this cross-sectional study appear to be broadly similar to at least some of those studies of children with SSD in this age range in each area examined, suggesting that they are broadly representative of children with SSD and do not violate any expectations of such samples as generated by the literature. At 26%, co-morbidity with LI was higher than the 15% reported for six year olds by Shriberg *et al.* (1999) but not as high as the 41% reported by Eadie *et al.* (2015) or 38.5% reported by Nathan *et al.* (2004b) for preschool children. However the possibility of selection bias, as discussed in section 3.2.5, and the small number of children with co-morbid LI must be kept in mind when interpreting results concerning children with LI in the following chapters.

The distribution of sub-groups of SSD, according to Dodd's Differential Diagnosis system is in line with cross-linguistic studies reported in the literature. The age distribution of the diagnostic sub-groups in the direction PD>CPD>IPD, also reflects that of similar studies that have examined this. The longitudinal study reported in Chapter 5 will allow for the investigation of the importance of age in variations in symptomatology, and the influence of age on predictors of persisting SSD.

An examination of the characteristics of the sample showed no risk factor that reliably identified children with SSD, and the presence of multiple risk factors was not associated with increased severity. Fewer than half of the children with SSD had a

positive family history of speech, language or literacy difficulties, and severity of SSD was not associated with either gender or positive family history, consistent with findings from Lewis and Freebairn (1997) that SSD in children with a putative genetic disposition to SSD do not present with a qualitatively different type of SSD. Consistent with much of the literature, SES, as measured by level of parental education was not associated with a higher risk of SSD. Other child factors such as motor delay or premature birth only applied to a very small proportion of the sample, limiting their usefulness for identification of children with SSD, and over half of the children presented with no risk factors at all, if being male is excepted.

When considering management of this population, it could be seen in this sample that over half of the children referred had mild to moderate impairments when measured by Shriberg's system, with 35% being classified as moderate to severe or severe. When using Dodd's system, which is adjusted for age, nearly half the children were classified as severe. However as this system depends on norms for age, a child at the older end of the sample where PCC for TD children is almost at ceiling, need only have a mild impairment in terms of raw PCC to fall into the severe category. Nevertheless, classification by either system shows that most children presenting to SLT services are not demonstrating severe impairments. Furthermore, increased age was associated with milder impairment, suggesting that children with SSD may continue to improve. In addition, 58% of the children in this sample were referred to SLT services before the age of 4, when the range of normal development in children tends to be much wider. As the normal course of development of speech finishes relatively early, unlike development of language, this allows the possibility that children with SSD may catch up with their peers in terms of speech development, in the relatively short term. Given the overwhelming demands on SLT services evidenced in section 1.1, such evidence would be very useful for planning best use of resources, and this possibility will be investigated in the longitudinal study reported in Chapter 5. It is a limitation of this study that a measure of activity and participation, such as the Therapy Outcome Measures (Enderby *et al.* 1997), was not used. This would have allowed for a better perspective of the impact of SSD than measures of speech accuracy alone.

Evidence that, unlike language, development of speech does not appear to be affected by SES, also contributes to evidence that aids in decision making about universal

services: if speech accuracy appears to be less vulnerable to environmental effects than language, it may require less resource to be dedicated to it at the universal level.

The next chapter is a cross-sectional analysis using exploratory factor analysis and cluster analysis to look for constructs that might underlie SSD, or suggest homogeneous sub-groups within the sample.

4 Chapter 4: Results 2: Factor analysis and subtyping

The goal of this section of the study was to attempt to classify children with speech sound disorders (SSD) empirically, using factor analytic techniques. Factor analyses are designed to identify latent variables or the cognitive processes underlying a construct by showing where performance on specific tasks factors together. This section therefore investigates the factor structure of a comprehensive battery of measures administered to a sample of pre-school and early school age children with SSD, to discover:

- Which components underlie children's speech abilities
- To what extent the underlying components allow homogeneous sub-types of children to be identified within the group of children with SSD of unknown cause
- Whether any profiles of deficit thus discovered are associated with risk factors, distinct surface symptomatology or previously defined sub-types of children with SSD

The validity of any factors and clusters defined will then be determined by examining their relation to outcomes in terms of persistence of SSD, phonological awareness and reading difficulties.

Previous studies using factor analyses have been carried out to determine associations between underlying deficits in speech and language and reading impairment (Lewis *et al.* 2006a, Anthony *et al.* 2010, Nathan *et al.* 2004a, Ramus *et al.* 2013); to investigate the relationship between phonological awareness abilities and performance on measures of speech perception, speech production, and naming tasks (Foy and Mann 2001), or to identify factors which differentiate sub-groups of children with hearing impairment from children with language deficits not due to hearing impairment (de Hoog *et al.* 2016). However, to the author's knowledge, no study has attempted to classify children with SSD empirically using factor analysis of a battery that includes tests of non-language cognitive ability. This study was motivated by Dodd's proposal that children who show deviant (i.e. use of atypical errors) rather than merely delayed phonological development do so because of an underlying deficit in the cognitive skills of pattern recognition, rule abstraction and cognitive flexibility.

Based on observation that a large number of children with SSD appear to present with normal language skills, and the contrary observation that children with language impairment often present with age appropriate speech, as well as previous research that distinguished separate factors of speech and language (Lewis *et al.* 2006a), it was hypothesised that measures of speech would load on a separate factor from measures of language abilities. However, given evidence regarding disassociation of speech input and output skills (Foy and Mann 2001, Nathan *et al.* 2004a), all speech skills might factor together, or separately. Considering there is evidence of motor deficits in children with SLI, but inconsistent findings regarding the involvement of motor abilities in SSD, it was not clear whether motor abilities might factor all together, or partially with speech or language, or separately. Regarding cognitive skills, it could be argued that language acquisition requires cognitive skills such as pattern finding and rule abstraction as much as the acquisition of speech (Tomasello 2005). However it was not known whether the cognitive abilities important to speech and language acquisition were the same as those which underlie non-verbal tasks such as early mathematical abilities. Therefore again it was not clear where these might factor.

Being an empirical method of analysis, that provides a data-driven method for investigating separable dimensions of abilities and disabilities, factor analysis, and exploratory factor analysis in particular, thus appeared to be an appropriate method to explore the interrelationships among the set of variables to enable the determination of the presence and nature of latent constructs, such that a deficit in one or more of these constructs could be causally related to deficits in the acquisition of speech. Exploratory factor analysis is an appropriate method in circumstances like those of this study, where the researcher does not know how many factors are necessary to explain the interrelationships among a set of items (Pett *et al.* 2003), as few restrictions are placed on the patterns of relations between the common factors and the measured variables, and no a priori number of common factors is specified.

As well as revealing how the different constructs measured by batteries of assessment measures are made up, and how they are associated with each other, factor analysis can be used to reduce larger numbers of test scores on different instruments to a smaller set of meaningful test composites, by allowing factor scale scores to be calculated for each research participant. These scores can then be used for further analysis with other external variables such as age or diagnosis, in order to determine whether children with

SSD have different profiles of strengths and deficits on the measured variables that might be used to classify them into meaningful sub-types.

4.1 Data analysis

All statistical analysis was carried out using SPSS Version 22 (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp). Please see Appendix 4-A for a full description of the procedure undertaken for the factor analysis.

4.1.1 Validity of the non-standardized tests.

As scores on these tests had not been established for a normative population, it was necessary to establish whether they discriminated well between children. The box plot in Figure 4-1 shows the range of scores on each test. This indicates that most scores were not at floor or ceiling, and discriminated adequately between children. However the results of the test of Phonological Legality showed that the majority of children's scores were not different to chance (chance level was 6/12, 29/52 children scored between 5 and 7, with no child scoring more than 9/12), so this test was not considered further.

4.1.2 Dealing with missing data

Where no data was collected for an individual variable e.g. due to attrition, or the child not completing a specific task, the result was treated as missing data. No imputations were made. Where multiple variables were required in the analysis, pairwise deletion was used to maximise the amount of data available for analysis and reduce bias associated with listwise deletion. Thus all observations that were non-missing on the variables in the analysis were included. Table 4-1 shows the number of children who completed each test, and the mean and range of scores.

Six tests had a high number (>10) of non-respondents; these were the test of Non-Word Discrimination, test of sequenced oro-motor movements, tests of production of multisyllabic words (spontaneous and in imitation) and syllable and alliteration awareness. Factors of age, gender and speech accuracy (measured by PCC) were investigated to establish whether these affected response status. Chi-square tests showed no difference on any test between responders and non-responders in terms of

gender. They also showed no difference in distribution of responders versus non-responders by six-monthly age group for the first four tests, but for the tests of syllable and alliteration awareness, there were significant differences, with the youngest children far less likely to respond than older children, as shown in Figure 4-2.

Figure 4-1: Range of scores on cognitive tests

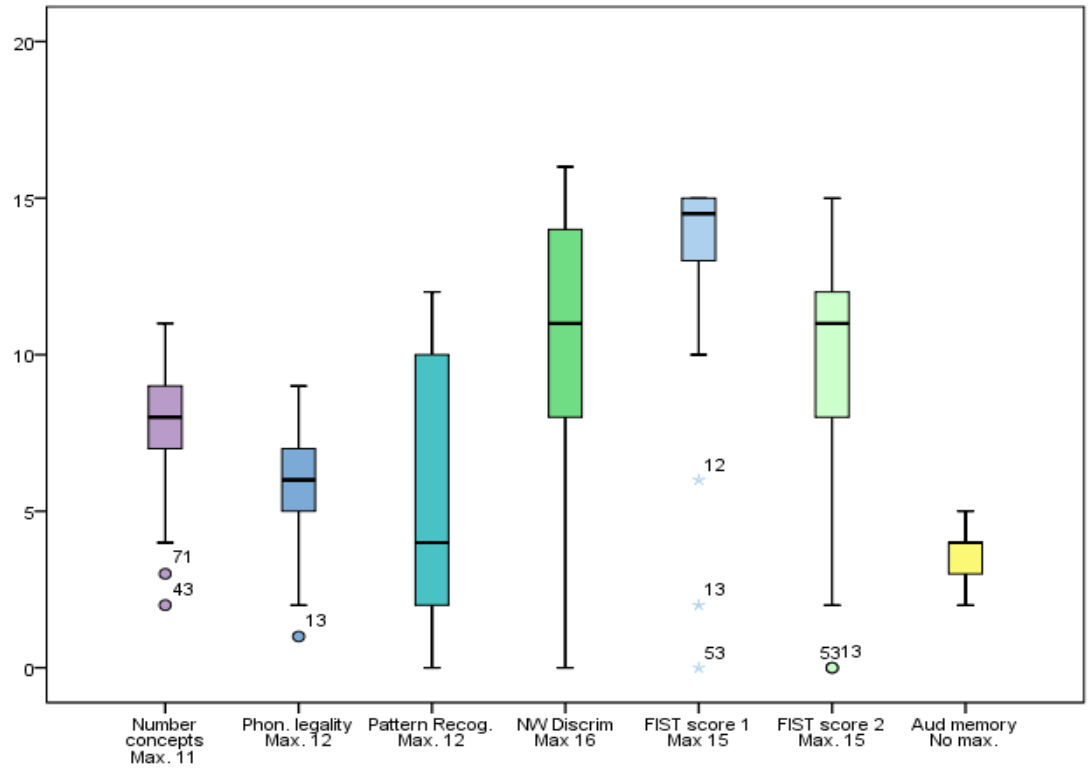
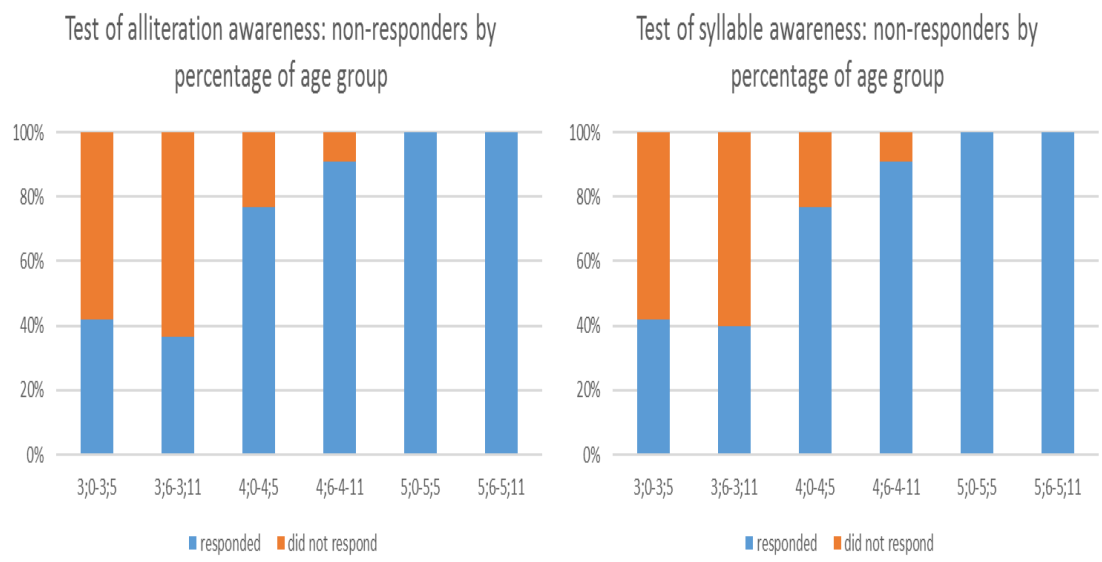


Table 4-1 Number of children who completed each assessment, mean scores and range

Number of children who completed each assessment and scores	N	Min	Max	Mean	SD
DEAP Percent Consonants Correct (PCC)	85	16	99	70.2	19.2
DEAP Standard Score (SS)	85	3	13	5.8	2.7
DDK Standard Score (SS)	76	3	14	10.1	1.9
Isolated Oral Movements Standard Score (SS)	77	3	13	9.0	1.9
Sequenced Oral Movements Standard Score (SS)	68	4	15	9.5	2.5
British Picture Vocabulary Scale (BPVS) Standard Score (SS)	83	70	132	98.9	12.0
CELF Sentence Structure Standard Score (SS)	75	2	15	9.1	2.7
CELF Word Structure Standard Score (SS)	75	2	17	9.4	3.7
CELF Expressive vocabulary Standard Score (SS)	75	4	16	10.5	3.0
CELF Core Standard Score (SS)	75	55	123	98.3	16.2
Number concepts	83	0	11	6.7	2.4
Pattern Recognition	73	0	12	5.1	3.9
Non-Word Discrimination Test	60	0	16	10.5	3.6
PCC Multisyllabic words: Naming	59	29	100	78.0	15.8
PCC Multisyllabic words: Imitation	57	35	100	83.2	17.4
Flexible Item Selection Test score 1	78	0	15	12.7	3.9
Flexible Item Selection Test score 2	78	0	15	9.1	4.4
Auditory memory word span	75	2	5	3.4	0.7
Beery Visual Motor Integration (BVMI) Standard Score (SS)	84	65	132	99.8	14.6
Phonological Awareness Syllable Structure Standard Score (SS)	52	4	14	8.9	2.2
Phonological Awareness Alliteration Standard Score (SS)	51	4	15	8.5	2.2

Figure 4-2: Percentage of non-responders by age on tests of PA



Mann-Whitney U tests showed no difference in speech accuracy, as measured by PCC, between responders and non-responders on tests of Non-Word Discrimination, production of Multisyllabic words or sequenced oro-motor movements. However, again there were significant differences for the tests of PA, with the non-responders in both tests having less accurate speech, both in terms of raw score PCC and age adjusted DEAP Scaled Score, as shown in Table 4- 2.

Table 4-2: Differences in PCC and DEAP Scaled Score of severity in responders vs non responders on Tests of PA

Test		Mean PCC	Significance	Mean DEAP Scaled Score	Significance
Syllable Awareness	**Responders N= 52**	76	*U*=424, *Z*=-3.91, *p*<.005	6.40	*U*=536, *Z*=-2.98, *p*=.003
	Non-Responders N=33	61		4.76	
Alliteration Awareness	**Responders N=51**	76	*U*=471, *Z*=-3.55, *p*<.005	6.27	*U*=587, *Z*=-2.58, *p*=.010
	Non-Responders N=34	62		5.00	

4.1.3 Analysis

Principal Axis Factoring with Direct Oblimin rotation was carried out on scores from the test battery for the complete sample of children, to reveal any latent variables that might factor together to show underlying cognitive processes associated with

identifiable sub-types in children with SSD. Factor scale scores for each child were then derived and were entered into a cluster analysis to cluster children in such a way that differences between clusters on these factor scores were as large as possible, and differences between children within the same cluster were as small as possible, in order to investigate whether sub-groups with homogenous profiles could be formed on the basis of the data. Box plots were used to visualize the profiles of the clusters, to reveal the different profiles based on the factor scores.

4.1.4 Selection of variables

The construct of interest is children's speech accuracy. The literature suggested five domains of ability that may contribute to this. Fabrigar *et al.'s* (1999) recommendations would suggest that between 15 and 25 variables should be selected for analysis, with more than one variable representing each domain. Eighteen variables with high face validity were selected as appropriate measures of aspects of these abilities, with more than one aspect being measured in each domain. The selected variables are shown in Table 4-3. It should be noted that the variables in bold appear in more than one potential constituent, as the literature suggests they may contribute to more than one domain.

Scores on the above test battery for all children, including the 11 with age appropriate speech at initial assessment, were entered into the analysis. All variables derived from standardised norm referenced tests were converted into Z-scores. For all the other variables, to remove age-related variance, raw scores were regressed along age, with standardised residuals used so that all scores are on the same scale and standardised to the same mean.

As reported above, not all children completed all tests, leading to missing data (see Table 4-1). It would have been desirable to add a measure of metaphonological ability to these items, as represented by the Preschool Inventory of Phonological Awareness sub-tests. However there was a relatively large amount of missing data on these items: 34 cases had no data on at least one of these measures, so including them would reduce the sample size and might limit the power of the model. Missing data was largely owing to the younger children's non-compliance with these tests, which may be an indicator that many children are not able to access metaphonological knowledge at this

age (Hesketh *et al.* 2007). Other studies have also reported having to abandon testing of alliteration and rhyme awareness in pre-schoolers, because too many children seemed unable or unwilling to attempt them (Snowling *et al.* 2000). Therefore, these variables were not included in the analysis, in order to allow for investigation of the factors from the largest amount of data possible.

Table 4-3: Measures entered into factor analysis

Potential Constituent	**Constituent item**	**Measured by:**	**Variable name**
Phonological ability	Accuracy of phonemes in single words	Percent consonants correct	DEAP PCC
	Severity of SSD according to norms for age	Standard Score on DEAP phonology sub test	DEAP SS
	Accuracy in multisyllable words, spontaneous production	Percent consonants correct	Mulitsyllsp
	Accuracy in multisyllable words, imitation	Percent consonants correct	Mulitsyllim
	Receptive Vocabulary	**British Picture Vocabulary Scale standard score**	**BPVS SS**
	Auditory Memory	**Animal name span test**	**Aud memory**
	Non-word Discrimination	Test of Non-word discrimination (single and multi-syllable non-words)	**NWD**
Motor ability	Oral-motor abilities	DEAP Oral-Motor test, Scaled Scores on: a) DDK rate b) isolated movements c) sequenced movements	 DDK SS SM SS IM SS
	Visual Motor Integration	**Beery Butenika test of Visual Motor Integration**	**BVMI SS**
Cognitive abilities	Pattern recognition ability	Test of pattern recognition	Pattern Recog.
	Rule derivation abilities	Flexible Item Selection Test, selection 1 & 2	FIST score 1 FIST score 2
	Mathematical competence	Test of Number concepts	Number Concepts
	Auditory Memory	**Animal name span test**	**Aud memory**
	Visual Motor Integration	**Beery Butenika test of Visual Motor Integration**	**BVMI SS**
Language ability	Language ability	CELF-PS2 (UK) sub-test scores a) Sentence structure b) Word Structure c) Expressive Vocabulary	 CELFSent CELFWord CELFEx
	Receptive Vocabulary	**British Picture Vocabulary Scale III standard score**	**BPVS SS**
	Non-word Discrimination	Test of Non-word discrimination (single and multi-syllable non-words)	**NWD**

4.2 Results of exploratory factor analysis

Principal Axis Factor analysis with Direct Oblimin rotation was conducted with age standardized scores for measures listed in Table 4-3. To identify distinct factors, variables with low primary loadings or weak and high cross loadings were excluded after the initial analysis. This resulted in the removal of 3 variables: DDK and Pattern Recognition did not correlate strongly with any other variable. The Test of Number Concepts loaded weakly (below .35,) and equally, with both the FIST and NWD scores on the one hand, and the CELF-4 (UK) sub-tests and BPVS on the other. Removing it increased the amount of variance explained from 65.4% to 67.4%.

Appendix 4-A describes the rationale for the sample size and the procedure undertaken for the analysis in more detail.

4.2.1 Factor structure

In the final analysis, four factors had eigenvalues greater than 1. Examination of the scree plot confirmed the presence of four distinct factors. Table 4-4 shows the rotated pattern matrix. To show the factor structure, the highest loading per variable is highlighted. The four factors revealed explained 67.4% of the variance. The factor loadings after Direct Oblimin rotation yielded relatively straightforward interpretations. The first factor included results of the FIST rule abstraction and cognitive flexibility tasks, as well as phonological processing abilities of non-word discrimination and auditory memory. These tasks seem to be related to cognitive-linguistic abilities, so it was named COGNITIVE. It explained 29.5% of the variance. The second factor comprised measures of speech output accuracy. It explained 20.1% of the variance and was named SPEECH. The third factor comprised measures of non-speech oral motor skills and visual motor integration, and was named MOTOR. It explained 10.3% of the variance. The last factor comprised measures of language ability, explained 7.5% of the variance and was named LANGUAGE.

After controlling for correlations among the factors, the pattern matrix showed only one variable that did not clearly load onto a single factor, Auditory Memory. It factored together positively with Non-Word Discrimination, another phonological processing ability, but surprisingly, negatively with speech output measures. As it is considered a core ability that might be expected to load on both input and output

factors, it was retained in the model. All other variables had an excellent to fair loading on one factor only, with no cross loadings over .4, showing a simple structure for the four factors, with each variable loading strongly on only one factor and each factor having multiple strong loadings.

Table 4-4: Principal Axis Factoring Pattern Matrix

Pattern Matrix

Assessment Variable	Factor			
	COGNITIVE	SPEECH	MOTOR	LANGUAGE
Flexible Item Selection Test FIST1	.909	.041	.084	-.116
Flexible Item Selection Test FIST2	.626	.028	.047	.041
Non Word Discrimination Test	.601	.011	-.083	.157
Auditory Memory Word Span	.392	-.369	.029	.291
DEAP Phonology sub-test PCC score	-.064	.935	.006	.112
DEAP Phonology sub-test Scale Score	-.103	.760	-.051	.172
Multisyllable words, picture naming	.253	.741	.091	-.055
Multisyllable words, imitation	.107	.615	.012	-.249
DEAP Sequenced Movements Test	.137	.014	.681	-.379
DEAP Isolated Movements Test	-.085	.030	.619	.087
Beery Test of Visual Motor Integration (BVMI)	.137	-.066	.455	.265
CELF PS2 (UK) Expressive vocabulary	.154	-.104	-.005	.745
CELF PS2 (UK) Sentence Structure	.306	.089	-.072	.599
British Picture Vocabulary Scale III	.044	.020	.383	.527
CELF PS2 (UK) Word Structure	.354	.198	.183	.405

Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization.

The cross loadings on COGNITIVE, of LANGUAGE variables (Sentence Structure and Word Structure) explain the moderate correlation shown in the correlation matrix (see Table 4-5) between those two factors. The moderate correlation between COGNITIVE and MOTOR is explained by the cross loading of BVMI, where visual

motor integration could be expected to reflect cognitive abilities, unlike sequenced and isolated movements which are more purely motor abilities.

Table 4-5: Factor Correlation Matrix

4.2.2 Robustness of the model

The resultant model showed a good simple structure, suggesting a strong model. Multiple high loadings on each factor with few cross loadings also suggest it is reliable at this sample size (Guadagnoli and Velicer 1988). Use of an orthogonal Varimax rotation was also trialled but the results remained the same, i.e. one factor representing cognitive skills, one speech accuracy, another non-speech motor skills and the last language skills, also suggesting a robust model.

4.2.3 Factor scores

Scores on each item that loaded strongly on the given factor were summed to obtain a score for each subject on that factor, for use in analysis of variance and regression. The data-driven approach had revealed four clearly interpretable sources of variance, but a more theory-driven approach was taken to calculating the factor scale scores for the purpose of further analysis. One consideration was the amount of missing data. Young children sometimes refuse to participate in testing (Dodd and McIntosh 2010), so although all children in the dataset had completed the DEAP, all other assessments had at least one participant with missing data. Factor scale scores could only be computed for children who had completed all assessments that made up that factor, so this reduced the number of participants for whom factor scale scores could be calculated. For this reason, the scores for the Multisyllable word tests were not included in the SPEECH factor scale, as their inclusion would have reduced the number of children for whom scores could be calculated. As they correlated very highly with PCC (Multisyllable words Spontaneous, r_s= .834, p <001, Multisyllable words Imitation, r_s= .834, p <001) anyway, their presence was not likely to add to the analysis. Table 6 shows the variables included in the scale and the number of participants for whom factor scale scores on each factor could be calculated. See Appendix 4-A for more details of how Factor scores were computed.

Table 4-6: Items on scales and number of children with SSD for whom factor scale scores could be computed

Number of children with SSD for whom factor scare scores could be computed				
	COGNITIVE	SPEECH	MOTOR	LANGUAGE
	Flexible Item Selection Test FIST1	DEAP Phonology Test PCC	DEAP Sequenced Movements Test	CELF PS2 (UK) Expressive Vocabulary
	Flexible Item Selection Test FIST2	DEAP Phonology Test Scaled Score	DEAP Isolated Movements Test	CELF PS2 (UK) Sentence Structure
	Non Word Discrimination Test		Beery Test of Visual Motor Integration	British Picture Vocabulary Scale III
	Auditory Memory Word Span			CELF PS2 (UK) Word Structure
N	57	85	68	73
% of sample	67%	100%	80%	86%

Scale scores for the COGNITVE factor could not be calculated for a large group of children (N=28) who had missing data, so that group was compared with children for whom scale scores could be generated, to check for any bias that may be introduced. Chi square tests showed that the 2 groups did not differ in gender composition, or diagnostic sub-group according to Dodd's classification. Mann-Whitney U tests showed that they did not differ significantly in age or in scores on the SPEECH or MOTOR factors. They did, however have significantly worse scores on the LANGUAGE factor, U=302, Z=-2.05, p=.040. However significance disappeared once the Bonferonni correction for multiple comparisons was applied. Thus despite the smaller number with scores on that factor, it seems that results may, with caution, be considered generally representative of the whole group.

4.2.4 Association between factors and membership of diagnostic sub-groups

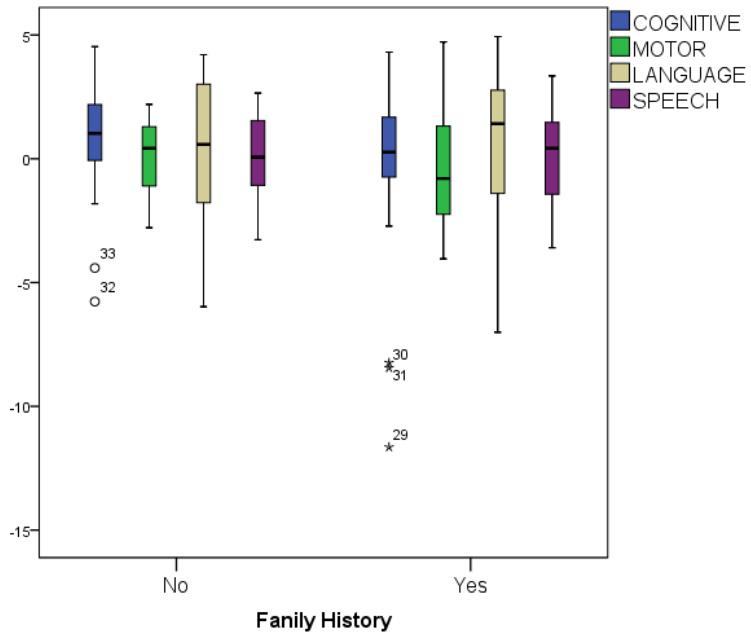
The distribution of factor scale scores for the children in previously identified diagnostic sub-groups was examined to investigate whether patterns of strengths and

deficits were associated with specific sub-groups identified by Shriberg's or Dodd's classification.

4.2.4.1 *Sub-groups defined by presence of Family History: Speech Delay-Genetic (Shriberg et al. 2010)*

To examine whether there was an association between family history and scores on any factor, factor scale scores were entered into a logistic regression with family history as the dependent variable. The model was not significant, indicating that a positive family history of speech, language or literacy disorders was not predicted by scores on any factor $\chi 2$ (4, N = 48) = 2.76, p =.598. The box plot in Figure 4-3 shows that children with and without a reported family history of speech, language and literacy difficulties performed similarly on each factor.

Figure 4-3: Performance on factors by presence of family history of speech, language or literacy difficulties



4.2.4.2 *Sub-groups defined by Number of ear infections (Speech Delay: Otitis Media with Effusion)*

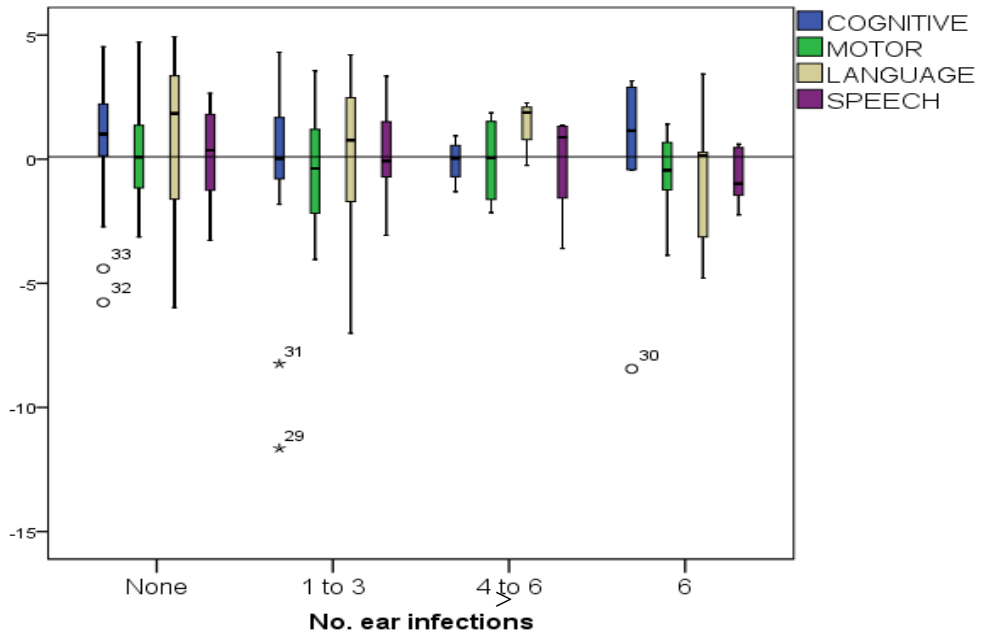
The box plots in Figure 4-4 below show that children with a history of more frequent ear infections did not perform any differently to children with a history of fewer ear infections on any factor. This was confirmed by a Kruskal-Wallis test (see Table 4-7)

that showed no statistically significant differences on any factor scale scores between groups.

Table 4-7: Association between factor scores and no. of ear infections

Results of Kruskal-Wallis test showing no difference on factor scores by number of ear infections				
Factor:	COGNITIVE	MOTOR	LANGUAGE	SPEECH
Chi-Square	3.195	2.063	.553	1.243
df	3	3	3	3
Asymp. Sig.	.363	.559	.907	.743
a. Kruskal Wallis Test				
b. Grouping Variable: No. ear infections				

Figure 4-4: Performance on factors by number of ear infections

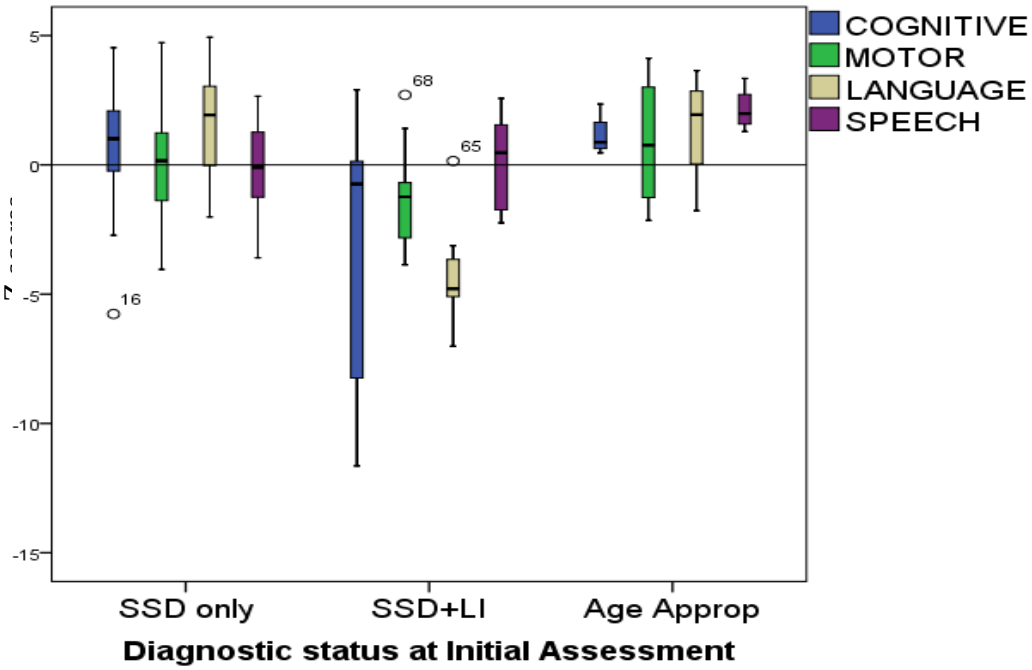


4.2.4.3 Sub-groups defined by presence of co-morbid language disorder.

Figure 4-5 shows the performance on the four factors for children with SSD alone, compared to those with SSD+LI, and those whose speech and language was age appropriate at Initial Assessment. Children with co-morbid language impairment showed worse performance on all factors except SPEECH, U=618, Z=-.09, p=.929, than those with SSD alone. Obviously, it would be expected that they score poorly on LANGUAGE, but a Mann-Whitney U test showed there was also a significant difference between children with SSD alone and those with SSD+LI for COGNITIVE, U=127, Z=-2.53, p=.011, and MOTOR, U=237, Z=-2.36, p=.018. These results

therefore indicate that children with SSD+LI, as well as having worse language skills than those with SSD alone, also performed worse on tests of more general cognitive-linguistic and non-speech motor abilities. In contrast, scores of the SSD only group did not differ significantly from those children with age appropriate speech and language at initial assessment on any factor except SPEECH, U=30, Z=-4.69, p=.000.

Figure 4-5: Performance of children with SSD alone, SSD+LI or age appropriate speech and language on factors



4.2.4.4 Sub-groups defined by Dodd's Differential Diagnosis Classification

A box plot was generated to examine whether the diagnostic sub-groups defined by Dodd (2005) were differentiated by different profiles of performance on the four factors. The performance of each of Dodd's diagnostic sub-groups on each factor is shown in Figure 4-6 below. A slight trend of AA>PD>CPD>IPD can be seen on all factors, but Kruskal-Wallis tests showed there was only a significant difference between diagnostic sub-groups for SPEECH, χ^2 (4) = 32.65, p<.001. In accordance with Dodd's findings (Dodd *et al.* 2005), the pattern of scores on SPEECH showed children with IPD were more impaired than those with CPD, U=60.5, Z=-1.96, p=.049, and those with CPD were more impaired than those with PD, U=304, Z=-2.43, p=.015. In contrast, children with CPD did not, as Dodd (2011) would predict, have worse scores on COGNITIVE than the other groups. However, COGNITIVE included the

results of a test of non-word discrimination, as well as the FIST results, introducing an element of auditory discrimination not considered by Dodd.

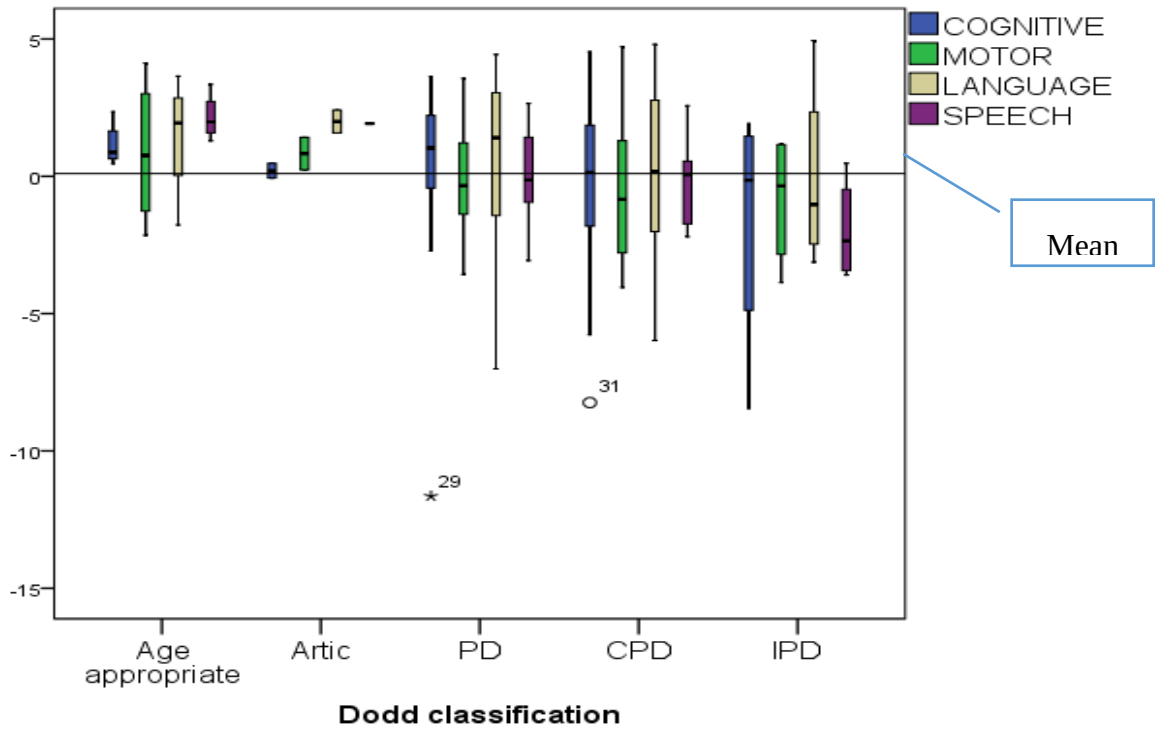
It must be taken into account that that group sizes for some sub-groups were very small, meaning results may not be generalizable to a bigger sample. Table 4-8 shows mean factor scale scores for each diagnostic sub-group, with the group size from which they were calculated.

Table 4-8: N and Mean factor scale scores for Dodd's diagnostic sub-groups

N and Mean factor scale scores for Dodd's diagnostic sub-groups								
Dodd Classification	COGNITIVE		SPEECH		MOTOR		LANGUAGE	
	N	Mean	N	Mean	N	Mean	N	Mean
Age appropriate	5	1.60	11	2.50	8	0.73	9	0.94
Artic	2	0.20	4	0.55	3	0.83	4	1.65
PD	28	0.42	34	0.21	28	0.08	33	0.22
CPD	17	-0.10	28	-0.74	22	-0.27	20	-0.41
IPD	5	-1.52	8	-2.03	7	-0.51	7	-1.84
Total	57		85		68		73	

When profiles for all sub-groups are viewed together, it can be seen that the PD and CPD groups have quite similar profiles. The sub-group with the worse COGNITIVE score is in fact the IPD group. However, it must be taken into account that some cell-sizes are very small, particularly so for children with IPD or articulation disorder.

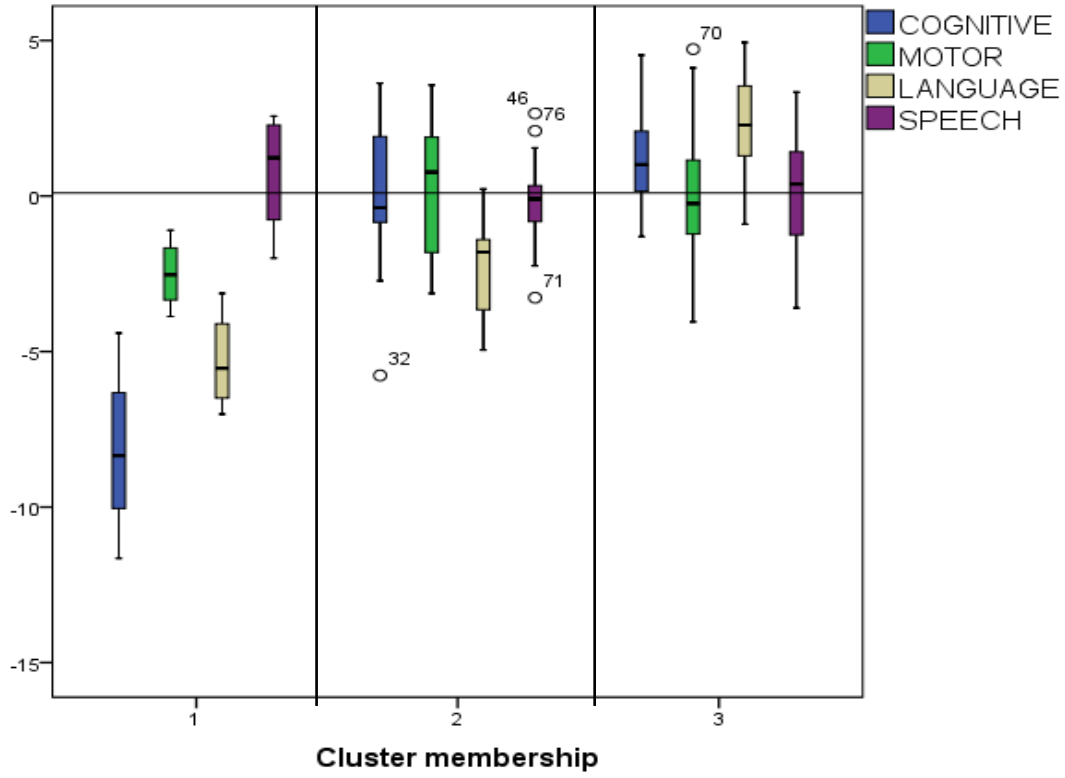
Figure 4-6: Performance on each factor by Dodd's diagnostic sub-groups



4.3 Identification of other sub-types from the data: cluster analysis

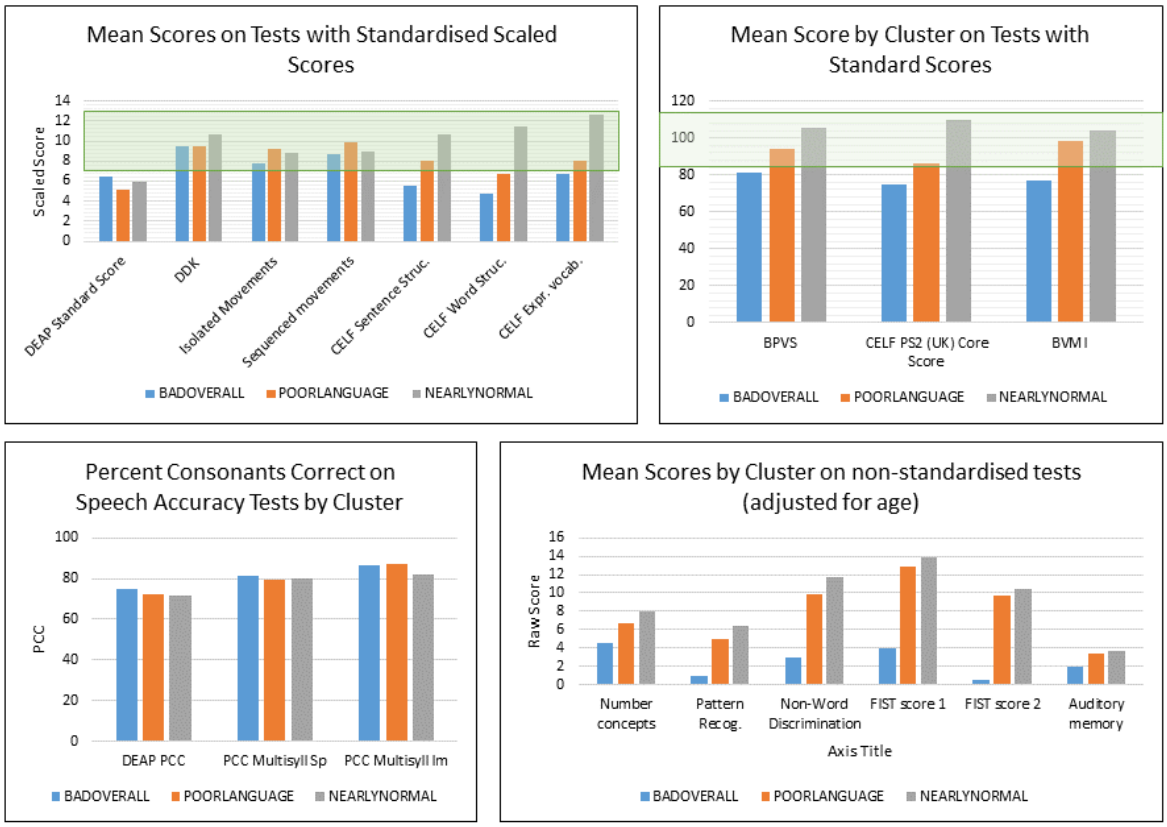
Given the fact that the data showed no strong association between factor scores and sub-groups previously identified in the literature, cluster analysis was used to explore the data to look for any previously unidentified sub-types based on factor scale scores. Cluster analysis is a statistical method of partitioning a sample into homogeneous classes to produce an operational classification on the factor scale scores for all factors, in order to identify sub-groups with homogeneous profiles of abilities. A hierarchical cluster analysis using Ward's method was used as it is the major statistical method for finding relatively homogeneous clusters of cases based on measured characteristics. This method of clustering is appropriate when there is no prior knowledge of how many clusters there may be or what they are characterized by (Burns and Burns 2008). Inspection of the dendrogram and the agglomeration schedule supported the extraction of three clusters.

Figure 4-7: Mean factor scale scores for each cluster



Due to combinations of missing data, only 48 (57%) of the children could be assigned cluster membership. Figure 4-7 shows mean scores on all factors for each cluster. Cluster 3 was predominant (N=31) and characterised by skills near to or above average for the group on all factors. It was labelled NEARLYNORMAL, as mean scores for this cluster on all normed tests except those for speech accuracy, fell within the normal range, as shown in Figure 4-8. Children in cluster 2 (N=13) showed poorer LANGUAGE performance than those in NEARLYNORMAL (although this was not necessarily in the impaired range), with skills in the normal range for motor abilities, and close to but slightly behind NEARLYNORMAL on the cognitive tests. It was labelled POORLANGUAGE. Cluster 1 (N=4) was very small and characterised by very poor COGNITIVE skills, poorer motor and poor LANGUAGE skills, but SPEECH skills similar to the other clusters. It was labelled BADOVERALL. Clusters were not differentiated by speech abilities. A Kruskal Wallis test showed no difference in age between clusters, $\chi 2$ (2, N = 48) = .478, p =.787, and a Chi-square test showed no difference in distribution of gender between clusters $\chi 2$ (2, N = 48) = .291, p =.865.

Figure 4-8: Mean scores on all tests for each cluster



Mean Scores on Tests by Cluster (green box shows scores within normal limits)

4.3.1 Association of Cluster Membership with other diagnostic sub-types or risk factors

Profiles of clusters of children that were formed based on the factor scores were visually compared with those of the clinical sub-groups defined by Dodd's Differential Diagnosis Classification system and Shriberg's SD-Gen and SD-OME sub-groups, to examine whether they aligned.

Figure 4-9 shows the distribution of Dodd's sub-types by cluster. Visual inspection indicated, as expected, that articulation cases and all but one age appropriate case fell into the NEARLYNORMAL cluster. Consistent with Dodd's findings that the PD group perform very similarly to TD children, most of the PD group fell into that cluster also. However, what is not consistent with Dodd's findings is that most of CPD group also fall into NEARLYNORMAL, where she would predict their membership of a cluster with the poorest cognitive abilities. In fact, all the sub-groups of phonological impairment are relatively evenly distributed among clusters, showing little association

between Dodd's diagnostic sub-groups and the groups defined by homogenous profiles of factor scale scores.

Figure 4-9: Distribution of Dodd's Diagnostic sub-types by Cluster Membership

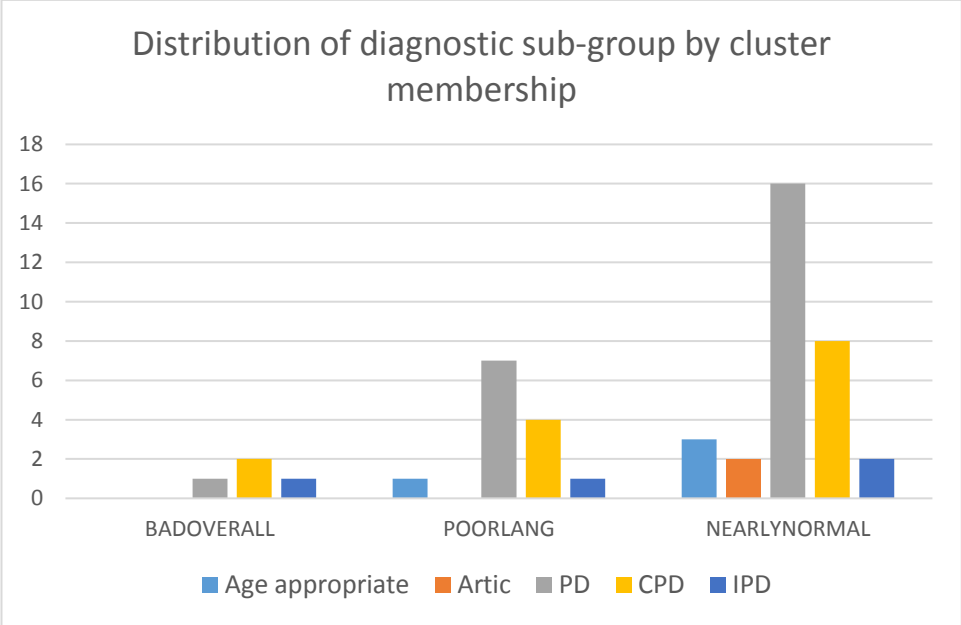


Figure 4-10 shows the affection status of family history for each cluster, indicating no obvious trend, although it is interesting that the cluster with the poorest performance overall also has the highest rate of affection status. Again the sample size is too small to show a trend, but it may be worthwhile to investigate whether this cluster also occurs in a larger sample and is also differentially affected.

Figure 4-10: Family history affection status by Cluster

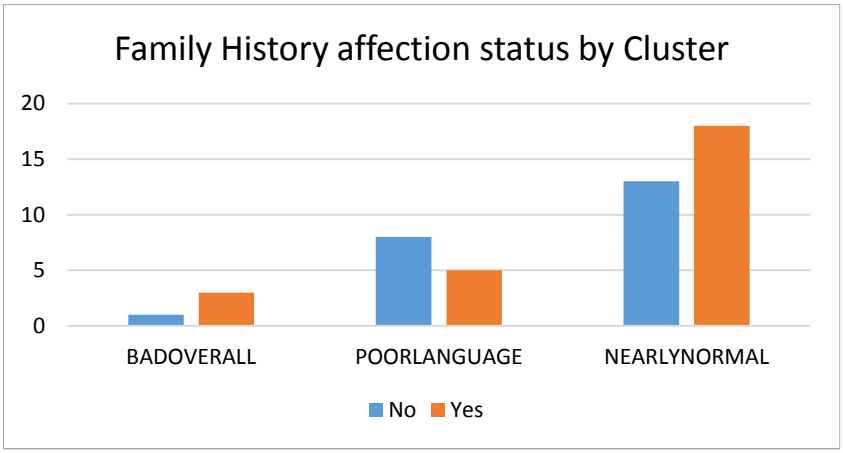
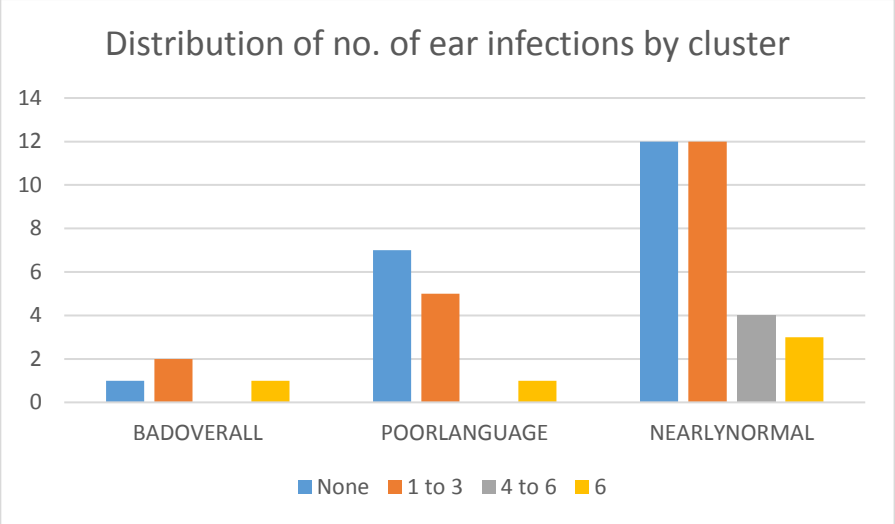


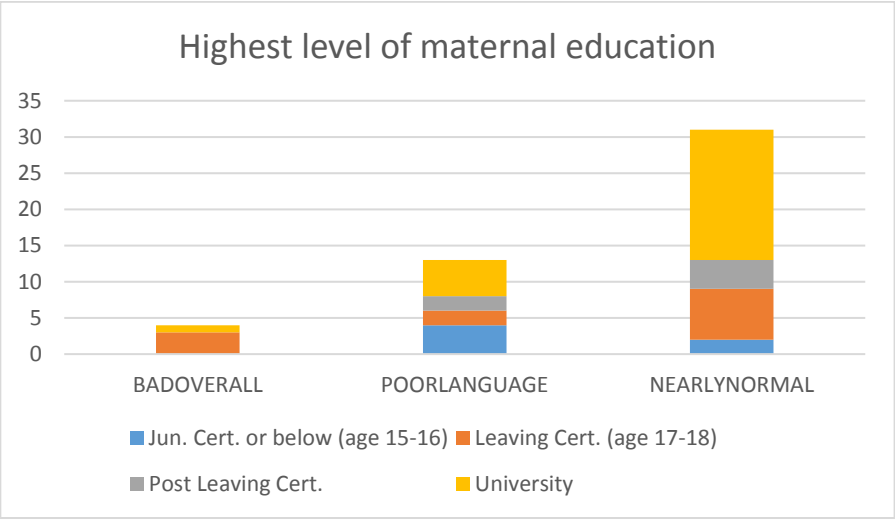
Figure 4-11 shows the distribution of the number of ear infections per cluster. Visual inspection shows no evident trend. Thus the clusters do not appear to align with Shriberg (2010)'s Speech-Delay: OME, or Speech-Delay: Genetic sub-groups.

Figure 4-11: No. of ear infections by cluster



Visual inspection of the distribution of cluster membership by highest level of maternal education, seen in Figure 4-12, showed no obvious effect of SES, and this was confirmed by a Chi-square test.

Figure 4-12: Distribution of cluster membership by highest level of maternal education



4.3.2 Predictive Value of Clusters

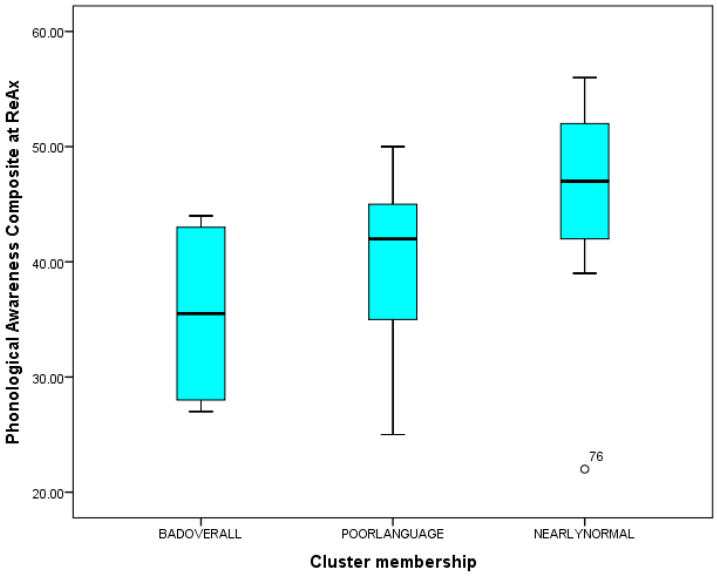
The outcomes of the children in this study after 2 years are properly described in the next chapter, but, as a preliminary attempt to validate the clusters, outcomes of the cluster members were briefly surveyed, to investigate whether cluster membership was predictive of resolved versus persistent SSD, or difficulties with reading. Forty-five of the 48 children who had been assigned to a cluster attended for reassessment, of whom four had presented with age appropriate speech at initial assessment. Of the remaining 41, 37 had resolved their SSD and eight had persistent SSD. As can be seen in Table 4-9, only 50% of the BADOVERALL cluster resolved their SSD, in contrast to 80% of the POORLANGUAGE cluster and 85% of NEARLYNORMAL. Due to small numbers, these differences were not significant.

Table 4-2: Distribution of children with persisting vs resolved SSD by cluster

	Cluster membership			Total
	BADOVERALL	POOR LANGUAGE	NEARLY NORMAL	
Persisting SSD	2	2	4	8
Resolved SSD	2	8	23	33
Total	4	10	27	41

Regarding the prediction of phonological awareness abilities (PA), a box-plot (see Figure 4-13) showed a clearly declining profile of performance on tests of phonological awareness abilities at reassessment in the direction NEARLYNORMAL>POORLANGUAGE>BADOVERALL. A Kruskal Wallis test showed this was significant $\chi^2(2) = 8.91$, $p = .012$. Mann-Whitney U tests initially showed the difference was significant between NEARLYNORMAL and POORLANGUAGE $U=74$, $Z=-2.31$, $p=.021$, and NEARLYNORMAL and BADOVERALL $U=18$, $Z=-2.22$, $p=.027$, but not between BADOVERALL AND POORLANGUAGE $U=13.5$, $Z=-.92$, $p=.374$. However, significance disappeared once the Bonferonni correction for multiple comparisons was applied.

Figure 4-13: Scores on the phonological awareness composite at Reassessment, by cluster membership



Overall just under 25% of children for whom cluster membership was defined had reported reading difficulties, as defined by receiving extra assistance specifically for reading at school. As expected for children with a history of SSD, this is higher than would be expected in the population as a whole, where, depending on the cut-off used, prevalence is approximately 9% (Pennington and Bishop 2009). It is however in line with figures reported for other studies of children with SSD (Johnson *et al.* 2011). Of the 41 children with cluster membership who had started reading, 50% of the BADOVERALL cluster had reported reading difficulties, compared to only 10% of POORLANGUAGE and 26% of NEARLYNORMAL, as shown in Table 10. The result that a higher proportion of children in the NEARLYNORMAL cluster, who had the best PA scores, had reading difficulties than those in the POORLANGUAGE cluster, was contrary to expectation. Although there is evidence that some children attain normal reading despite persistent phonological awareness difficulties (Pennington 2006), comorbid language impairment is a further risk factor for reading difficulty (Bishop and Adams 1990, Snowling *et al.* 2000, Nathan *et al.* 2004a), so the low number of children identified as poor readers in the group with poorer PA and language skills is surprising. It is possible that, as children were only beginning learning to read, and reading difficulties were dependent on report by parent rather than by direct assessment, these figures might change as children get older and formal assessments of reading are used. The small group size must also be taken into account.

Table 4-3: Distribution of children with reading difficulties by cluster

	Cluster membership			Total
	BADOVERALL	POOR LANGUAGE	NEARLY NORMAL	
Reading Ok	2	9	20	31
Reading Difficulty	2	1	7	10
Total	4	10	27	41

4.4 Discussion

4.4.1 Components which underlie children's speech abilities

The first aim of this analysis was to identify components underlying children's speech abilities. Fourteen of the 18 variables chosen to measure elements that have been posited to contribute to the development of speech, as outlined in Chapter 1, namely, speech perception, cognitive-linguistic abilities, vocabulary and language skills, motor abilities and production accuracy, were found, using exploratory factor analysis, to divide into 4 distinct factors, with a fifteenth being spread equally over 2 factors. Three variables, namely: scores on a DDK test which also contained a measure of consonant accuracy, the Test of Numerical Concepts, and Pattern Recognition, did not show a strong association with any factor.

Children with SSD thus showed four independent sources of variance: one for cognitive-linguistic ability, one for speech production accuracy, one for non-speech oral motor skills and one for language. The SPEECH factor, which loaded only on measures of output accuracy, showed no correlation with any other factor, indicating that speech accuracy, or inaccuracy in the case of pre-school children with SSD of unknown origin, is not associated with a single underlying deficit in either non-speech motor skills, language skills or broader underlying cognitive skills. In fact, three children with SSD scored above the mean of the age appropriate group on all of these factors, suggesting that a deficit in any one is not a necessary or sufficient condition for SSD.

The fact that measures of vocabulary factored together with syntactic/morphological abilities in LANGUAGE, which did not correlate with SPEECH prompts further consideration about the role of vocabulary in speech accuracy. If we follow the Lexical Restructuring Model (Walley *et al.* 2003), whereby phonological representations

become more refined and segmented as more words are acquired allowing for categorical representations of segments to be formed, the finding that speech accuracy does not correlate with either receptive or expressive vocabulary, suggests that there must either be a ceiling effect, whereby categorical representations can be adequately established once a critical mass of words is acquired, or that speech accuracy is, in at least a large number of cases, independent of such representations. Consistent with the first proposition is evidence cited by Stoel-Gammon (2011) from a number of studies showing an association between lexical and phonological development in under-threes, whereas studies of children older than this show that children with phonologically-based SSD can have vocabulary scores which are average or well above average according to normative data e.g. (Broomfield and Dodd 2004a, Waring *et al.* 2016), as well as evidence that children can have normal speech in the context of below average vocabulary (Broomfield and Dodd 2004a). Regarding the second proposition, Munson *et al.* (2005a) cites evidence that children with SSD do not find repetition of non-words with low as opposed to high frequency sequences relatively more difficult than TD children, to suggest that many children with phonologically based SSD do not have deficits in the higher-level phonological knowledge of categorical phonemic representations which is afforded by growth in the lexicon, but instead in building lower level perceptual and articulatory representations. Thus the association between vocabulary size and the phonological system may change with age or with the locus of impairment.

The cognitive-linguistic factor, COGNITIVE, had the strongest loading from the Flexible Item Selection Test (FIST), which is a measure of executive function in pre-schoolers (Jacques and Zelazo 2001). The first choice provides a direct measure of the child's ability to identify a single dimension (the abstraction component of the task), while the second score is a measure of participants' ability to switch flexibly between dimensions (the cognitive flexibility component), although it also requires abstraction. The other strong loading on this factor was the Non-Word Discrimination Test, with a weak loading from Auditory Memory. At first glance, it is surprising that the FIST and Non-Word Discrimination should factor together, and that the Test of Number Concepts and Pattern Recognition do not factor with the FIST. However the similarity between the FIST and Non-Word Discrimination tasks is that they both require categorisation abilities. The FIST requires children to identify two items that are the same on one dimension, while ignoring that they differ on two other dimensions. The

Non-Word Discrimination Test requires the child to categorise two acoustic signals as the same when they are made up of the same string of phonemes, despite being spoken by different voices, and as different when they differ by one phoneme. Both tasks require the identification of salient cues that are similar, while ignoring features that differ. While the Test of Number Concepts does include some items that involve classification of items into sets, it also involves matching and identifying one-to-one correspondence in dissimilar items, as well as the order and seriation abilities that are also required by the Pattern Recognition test. Thus the FIST and Non-Word Discrimination Test both require abilities that are used to abstract the phonological system of the ambient language from the input signal: namely, to detect sameness and difference, make distinctions, categorise on the basis of correct salient cues, while ignoring irrelevant features. This finding is thus consistent with Nittrouer et al.'s (2011) suggestion, introduced in section 1.3.2, that children with language and phonological processing deficits may have an underlying deficit in organizing sensory information to form coherent categories. Good auditory memory would obviously also aid this process, so that the Word Span test also factors together with the FIST and Non-Word Discrimination Test, although more weakly. However, the fact that scores on these two tests did not factor together with measures of speech accuracy shows that deficits in these skills are not associated with all cases of SSD.

Isolation of a separate factor of non-speech motor skills, MOTOR, that had no correlation with speech accuracy, converges with the findings of McGrath *et al.* (2007) in their study of 60 children aged 5;0 to 7;3 with SSD, 69 siblings and 41 controls. They found that the non-speech oral-motor sub-tests of the VMPAC were not associated with speech accuracy. This accords with recent literature concerning the lack of theoretical and empirical evidence for the efficacy of non-speech oral motor exercises (NSOME) in therapy for SSD (Lof & Watson, 2010; Powell, 2008; Lass & Pannbacker 2008). Lewis (2006a) also found that, in a group of 3 to 7 year-olds, oral-motor skills, including speech related oral-motor skills, factored with neither speech, language nor literacy. However, the fact that the results of the DDK task in this study did not factor with speech production accuracy does not accord with other studies: deficits in speech related oral-motor skills i.e. a DDK task, were found by Wren *et al.* (2012) in 8 year-olds with persistent SSD compared to controls, particularly in those with articulation as opposed to phonological deficits; however differences between groups were very small, with a relative risk ratio of less than 2, suggesting such

measures may not be useful clinically. Given the difference in age between this sample and Wren *et al.'s*, it may be that motor deficits are more characteristic of SSD that persists, than of SSD in pre-schoolers. Similar to Wren, McGrath *et al.'s* (2007) study cited above, found that speech related motor control measures loaded together with measures of PCC in their confirmatory factor analysis, although again only weakly (<.46), with non-speech oral-motor measures forming a separate factor. As well as differences in age of the samples, differing results may be due to the fact that different tools were used (Lewis: Oral and Speech Motor Control Protocol (Robbins and Klee 1987); McGrath: VMPAC (Hayden 1999); Wren: accuracy of repetition of /pətəkə/ /bədəgə/ over 10 seconds) and measures were purely observational. In this study, the DDK task used did not factor together with any other task. This is consistent with the findings of Brosseau-Lapré and Rvachew (2016) using a similar aged sample: in their study of 72 3 to 5 year-olds with SSD, measures of DDK were not associated with any other variable they measured, including receptive vocabulary, non-verbal intelligence, PCC and PA. However, the particular DDK test used from the DEAP, like Wren's, does in fact contain linguistic content in that children have to say "pat-a-cake" rather than repeating a single syllable, and are marked down for substitution errors (e.g. "pat-a-tate") which may have a linguistic rather than motor basis, meaning failure on this task might be associated with either motor or linguistic deficits, therefore not correlating strongly with either. It could also be due to the way speech accuracy was measured: Namasivayam *et al.* (2013) found that speech motor control as assessed by the VMPAC was significantly correlated with measures of speech intelligibility at word and sentence level, but not with articulatory proficiency as measured in PCC on the Goldman Fristoe single word test (Goldman & Fristoe 2000).

There was a weak correlation between MOTOR and COGNITIVE, which probably reflects the influence of the BVMI visual-motor integration test. This loaded on the MOTOR factor with weaker cross-loading on COGNITIVE; the fact that results of this test are stated to correlate with IQ scores (Beery and Beery 2010), may account for this crossloading.

As this study aimed to use measures that are commonly or feasibly used in clinical practice, it did not include measures of gross motor tasks to compare with Visscher *et al.* (2007) or Visscher *et al.* (2010), nor of any non-oral motor task to compare with Bishop (2002) or Bradford and Dodd (1996). This may account for the fact that there

was not a stronger correlation between MOTOR and LANGUAGE, given the reseach indicating co-morbidity of motor and language deficits in children with SLI (Hill 2001). However there was a moderate correlation between LANGUAGE and COGNITIVE. This may be explained by the presence of auditory discrimination skills in the make up of the COGNITIVE factor, as these are likely to contribute to the formation of robust phonological representations which are important in vocabularly learning (Bishop 2000).

Correlation between the SPEECH factor and COGNITIVE factor was low. It was interesting that phonological input skills, as represented by auditory non-word discrimination skills, and auditory memory did not factor with speech output skills, but formed a separate factor together with cognitive skills, which showed greater correlation with the language factor than the speech factor. The correlation of COGNITIVE with LANGUAGE suggests that cognitive pattern recognition and rule abstraction skills may be more important in early language learning than in the acquisition of phonology. Alternatively, the correlation may be due to the presence of variables measuring auditory memory and auditory discrimination in COGNITIVE, as these abilities are important in forming phonological representations (Anthony et al. 2010), and therefore also related to language learning. This study was not designed to test whether pattern recognition and rule abstraction skills are impaired in children with language deficits, but the finding in this sample that children with poor language skills also have poorer cognitive skills than those with SSD alone may be worthy of further investigation in future research, especially given the evidence that non-verbal IQ tends to decline with age in a large number of children with primary language impairment (Botting 2005).

The independence of the SPEECH factor from the other three factors suggests that deficits in speech accuracy, or speech output skills, are not directly associated with a single underlying deficit in either cognitve, linguistic or motor skills. It may suggest that deficits in speech accuracy can arise independently from deficits in any one or more of these areas, which would go some way to explain the heterogeneity in the natural history, course and response to intervention observed in children with SSD.

4.4.2 Identification of homogenous sub-types

The second aim of this analysis was to investigate whether the components extracted by the factor analysis allowed homogeneous sub-types of children to be identified within the group of children with SSD of unknown cause. Cluster analysis identified three sub-types: a large, predominant cluster, NEARLYNORMAL, a moderate size cluster: POORLANGUAGE, with poor language skills but average abilities in the other domains, and BADOVERALL, who showed low scores in all areas except speech accuracy. The fact that homogeneous clusters based on different profiles of strengths and weakness on factors identified by the factor analysis can be shown, indicates that identifiable sub-types, based on such profiles, rather than aetiology or speech symptomatology, might exist among the heterogeneous group of children with SSD of unknown cause.

In regard to the third aim, to establish whether any profiles of deficit thus discovered are associated with risk factors, distinct surface symptomatology or previously defined sub-types of children with SSD, what is shown in this small sample is that they do not closely align with sub-types identified by Dodd, in regard to surface symptomology, or Shriberg, in regard to the risk factors of family history or otitis media. Although there is a group of children with cognitive deficits as posited by Dodd, in that they performed poorly on the FIST, in this small sample they do not align with surface symptomatology. Neither Dodd's nor Shriberg's classifications take into account variation in language abilities, which appears to be a more important factor in this sample: as in Nathan *et al.'s.* (2004a) study, the children with SSD and poor language skills (BADOVERALL and POORLANGUAGE) tended to perform somewhat worse than the SSD group with normal language skills across all tasks except speech accuracy. Those children in the group with the best language skills performed near to, or above the normative mean in all tests where norms were available, and showed a trend to perform better than those with poor language on the non-standardised tests. This suggests that sub-typing children by presence of multiple deficits may be more useful for predicting ongoing difficulties than by presumed aetiology or by surface symptomatology of SSD. The suggestions that children who perform poorly overall may have worse outcomes is similar to the finding by Beitchman *et al.* (1996), whose low overall cluster had the worst outcomes, while children with isolated speech difficulties improved their speech and had low risk for other difficulties. It also aligns

with what was suggested by Lewis *et al.* (2000), who found that that children with SSD+LI at 4-6 years were at greater risk for reading, spelling, and language difficulties at age 8 to 9, than those with SSD alone. However, in this sample, although the children with poor language difficulties also had poorer PA, a bigger proportion (26%) of the group with good language skills were reported to have reading difficulties than those with poor language (21%). The fact that the children were at the early stages of reading may have played a part in this: Snowling *et al.* (2000) showed that phonological difficulties affect the acquisition of literacy most at the outset of reading, but that impairments of other language skills play a bigger part in later development. Reading status was defined by parental report, so reliance on this rather than direct assessment of reading may also have affected the findings, as well as the fact that very few children had severe or even moderate LI.

The fact that the three clusters do not differ in severity of speech impairment, nor in surface symptomatology, is not consistent with a severity model of SSD (Pennington 2006), whereby SSD is manifest at a particular point on a continuum of difficulties that encompasses language and other cognitive difficulties. The sample here is too small to show any clear trends, but a longitudinal study of a larger sample may show whether it is consistent with a multiple deficit model, where co-morbidities have an additive effect, or interact so that children are less able to use compensatory strategies or strengths in other areas to resolve the SSD, and therefore have more likelihood of persistent SSD and/or ongoing difficulties with phonological processing, leading to consequent difficulties with literacy, academic and social development. Wren *et al.* (2012), for example, found that 8 year olds who had persistent phonologically based SSD had lower IQ scores than a control group. The next chapter will examine whether deficits in particular skills are predictive of resolution of SSD.

It is also important to take into account that it cannot be determined from the results of this sample whether membership of a cluster is a stable characteristic, or whether children in different stages of development might move between clusters in a developmental fashion.

4.4.3 Comparisons with other studies

Most other factor analytic studies of children with SSD have looked at outcomes for reading rather than speech; however, the independence of speech output measures in

SPEECH from measures of language and vocabulary in LANGUAGE is consistent with the findings of those studies. Lewis *et al.* (2006a) in a study of 185 three to 7-year-olds tested on a battery of speech and language assessments, carried out a factor analysis which revealed distinct and independent constructs of speech and language. While it is frequently observed that children with pure articulation disorders rarely have co-morbid language impairment, phonology is considered to be a part of language, so SSD with a phonological basis is sometimes considered to be part of a continuum of language disorder, albeit at the less severe end (Bishop and Edmundson 1987). However the disassociation of speech and language disorder is also seen in over 50% of cases, indicating that co-morbid disorders might be dual deficits rather than a continuum of impairment. Results of this analysis would support that. In fact SPEECH, which included only speech output tasks, did not correlate with any other factor, which could mean that it is truly independent of other abilities, or alternatively that the group of children with SSD is so heterogeneous that it obscures associations with other constructs. A multi-deficit model might explain why severity of SSD is not a predictor of persistence, and why children with co-morbid conditions do not necessarily have worse SSD. Instead likelihood of persistence may depend on the presence of other risk and protective factors, which contribute to the heterogeneity of the presentation (Pennington and Bishop 2009).

Lewis *et al's* (2006a) assessment battery only included speech output skills, but another factor analytic study by Ramus *et al.* (2013) included phonological input as well as output tasks. This study comprised 64 children with SLI and/or dyslexia, aged 8-12 and 65 controls aged 5-12. As found by Lewis et al. (2006a) and by this study, they also found speech and language were independent constructs, but phonology measures were further broken down into 2 separate factors. The first, "phonological skills" showed high loadings on phonological awareness, a phonemic categorisation task and digit span tasks, and the second, which they named "phonological representations", showed strong loadings on non-word and real word production tasks, with weaker loadings for non-word and real word auditory discrimination tasks. Thus their findings differed from those in this sample by factoring input and output tasks together. However, Ramus et al. did not include any non-speech/language cognitive tasks such as those in this study; addition of these may have changed the loading, especially considering that the input tasks loaded only weakly on the same factor as output in their study.

Other factor analytic studies have found that speech/phonology measures divided into separate input and output factors. Nathan *et al.* (2004a) followed 38 children with SSD, and 19 controls from age 4 to 7, performing a principal components analysis at 3 points: at mean ages 4;6, 5;8 and 6;9. The output tasks included naming and repetition of words and non-words, and the input tasks were auditory discrimination tasks with both real and non-words, similar to the Non-Word Discrimination Test in this study. At each time, the analyses produced a two-factor solution with output and input tasks loading strongly on separate factors, as was found in this study. Foy and Mann (2001), in a study of 40 TD pre-schoolers aged 4-6, carried out a factor analysis to investigate the relation between children's phonological and phoneme awareness skills and their performance on a variety of spoken language tasks set to measure the strength of their phonological representations. As well as a factor they called Literacy, that comprised age, phoneme awareness, and letter knowledge, they also found 2 distinct factors for speech which they called Speech Production, which included articulation and non-word repetition measures, and Speech Perception, which included word discrimination and digit span task scores, similar to the non-word auditory discrimination and word span tasks that factored with the COGNITIVE factor in this study. These findings were similar to those of Anthony *et al.* (2010), who carried out a factor analytic study on a sample of 175 children aged 3 to 5 recruited from pre-schools; exclusion criteria comprised children with known sensory, physical, or cognitive impairments, abnormal oral structures or mechanism, but not children with speech problems or language delays. They found 3 discrete factors on a battery of tests aimed at measuring processing abilities that contribute to the formation and use of phonological representations, namely: Access to Phonological Representation, measured by rapid naming tasks; Precision of Expressive Phonological Representation, indexed by tasks tapping speech accuracy, and Precision of Receptive Phonological Representation, indexed by tasks where children had to judge the accuracy of real words produced by a puppet. Their analysis therefore showed the same division between speech output and input tasks as this study. They concluded that the quality of phonological representations affects both how accurately speech is produced and how precisely it is perceived, as well as how easily it can be accessed. However, it would seem likely that the relationship is reciprocal: the strengths of the input processing abilities must also affect the quality of the phonological representations formed.

The separation of auditory discrimination and auditory memory, which can be considered receptive phonological abilities, in the COGNITIVE factor, from production abilities in the SPEECH factor, together with evidence showing the independence of speech perception and production in the analyses of Nathan *et al.* (2004a), Anthony *et al.* (2010) and Foy and Mann (2001), converges with the findings of Rvachew (2006), that speech production skills can be disassociated from speech perception skills. The weak correlation between the factors supports Dodd and McIntosh's (2010) proposition, that deficient input processing is not sufficient to explain difficulties with phonological development, as some children had high scores at initial assessment on the factor representing receptive phonological abilities, with low scores on the one representing speech production. It is also consistent with Rvachew's (2007) suggestion that children with both speech perception and production difficulties form a separate sub-type to those with production difficulties alone.

4.5 Conclusion: Factor Analysis

Factor analysis indicated the independence of speech accuracy from semantic/syntactic deficits in pre-schoolers with SSD, as well as from cognitive-linguistic and motor abilities. This indicates that speech accuracy may be an encapsulated ability and SSD is not due solely to a deficit in any of these areas, nor is a deficit in one of these areas a necessary or sufficient cause. Profiles of deficit did not align with heritability or environmental risk factors of maternal education or otitis media, nor with previously defined sub-types of children with SSD. The results appear to better align with a multiple deficit model, where sub-types are defined by presence of co-morbid deficits, which may have an additive or interactive effect.

The next Chapter will report outcomes for the sample after two years, and investigate whether deficits in any particular skills are predictive of resolution or persistence of SSD.

5 Chapter 5: Results 3: Outcomes after 2 years.

This chapter reports the results of findings at Re-assessment of the sample, approximately 2 years after the initial data collection. The number of children who resolved their SSD by reassessment are reported, as well as changes in presentation in terms of increase in PCC and changes to diagnostic sub-group. The number and characteristics of children who did and did not receive therapy are reported, and outcomes compared. Following this descriptive analysis, univariable and multivariable logistic regression analyses were used to obtain odds ratios, 95% confidence intervals and likelihood ratio *p* values for associations between persistent SSD and demographic and other tested variables, to identify any that were predictors for resolved/persistent SSD, or for phonological awareness abilities at Reassessment.

5.1 Outcomes at Re-assessment

Seventy-five children attended for reassessment, of whom nine had had age appropriate speech at Initial Assessment, and 66 had been diagnosed with SSD. Mean age at reassessment was 6;7 (SD 12.6 months, range 5;0 to 7;9) for the group who had had age appropriate speech at Initial Assessment, and 6;1 (SD 9.4 months, range 4:8 to 8;1) for the 66 children with SSD at Initial Assessment. The average time between Initial and Reassessment was 24.5 months, (*SD*=3.5, Minimum =17, Maximum =35).

5.1.1 Number of children whose SSD had resolved by reassessment

Children were considered to have resolved their SSD if they presented at reassessment with age appropriate speech (as opposed to error-free speech). Age appropriate speech was attributed using the norms from the DEAP phonology sub-test. To be considered as having age appropriate speech, the child had to fulfil the following criteria:

- Have PCC in the 13th percentile or above (scale score 7 and above) for children in the same chronological age group according to the Irish standardisation norms of the DEAP, (where a dentalised production of /s/ or /z/ was counted as an age appropriate normal variation (see section 2.2.5.5 for justification).
- Use only phonological error patterns classified as typical for children in their chronological age group according to the DEAP norms
- Have no articulation disorder, as defined by the consistent distortion of a particular sound or sound group in all contexts (e.g. lateralisation of fricatives).

All nine children who had presented with age appropriate speech at initial assessment also had age appropriate speech at reassessment. Of the 66 children who had presented with an SSD at Initial Assessment, 52 (79%) presented with age appropriate speech at Reassessment. They had a mean age of 6;2 (SD 9.2, range 4;8 – 8;1), meaning they were significantly older than the 14 who had not resolved, (mean age 5;8, SD 8.6, range 4;10 – 7;1), U= 201, Z =-2.6, p=.010. They had also been older at Initial Assessment: resolved, mean age 4;1, SD 8.7, range 3;1 – 5;9, persisting, mean age 3;9, SD 8.5, range 3;0 – 5;2; U= 217, Z =-2.3, p=.021. The interval between assessments was somewhat greater for those who resolved (24.9 months, SD 3.2) than those who did not resolve (22.9 months, SD 3.4), but a Mann-Whitney U test showed this difference was not significant, U = 244.5, Z = -1.9, p = .058.

5.1.2 Change in PCC between Initial and Re-assessment: Children with SSD

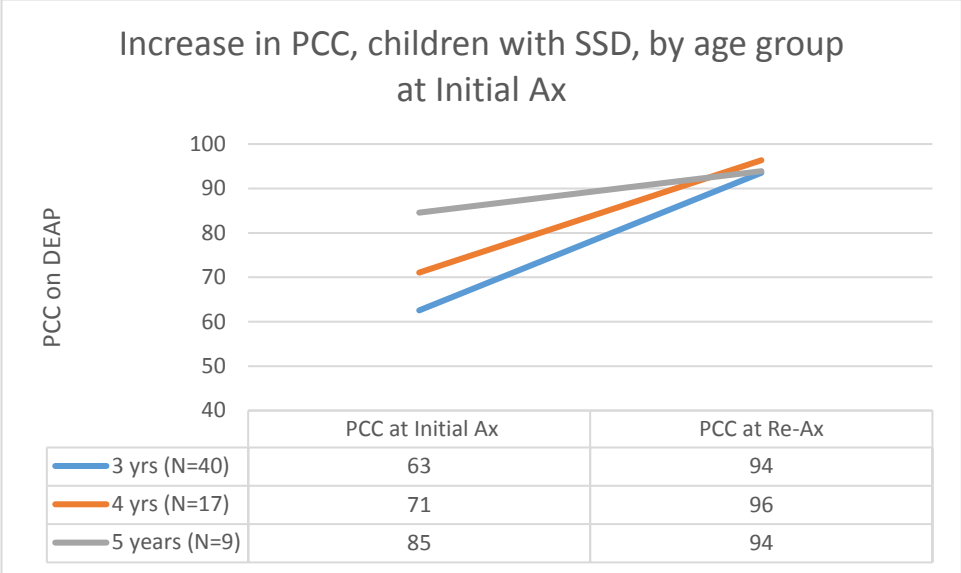
All children with SSD had increased PCC by Re-assessment except one who reduced PCC slightly (11503, PCC 91 at Initial Assessment and 89 at Re-assessment). He had an initial diagnosis of very mild delay at age 5;0, and by Reassessment at age 7;2 had only a dentalised production of /s/. He had received both a home programme and clinician directed therapy. The maximum increase in PCC was 70 points (11328, PCC 16 at Initial Assessment and 86 at Re-assessment). This child had a diagnosis of IPD at initial assessment, aged 3;2, and PD at Re-assessment, aged 5;3. He had received clinician directed therapy.

The mean increase in PCC was 27 percentage points, (range -2 to 70, SD 17, median 21). Boys made greater mean gains (30 points) than girls (22 points), but Mann-Whitney U tests showed that this difference was not significant, and both had similar PCC at Initial Assessment. Figure 5-1 shows that younger children made greater gains than older children: 3 year olds made mean gains of 31 points, 4 year olds 25 and 5 year olds 9 points, reflecting the fact that many children were at or near ceiling after 2 years. There was a moderate positive correlation between PCC at Time 1 and Time 2, $r_s(64) = .547$, $p < .001$).

Cross-sectional data from Initial and Re-assessment were examined to posit a trajectory in growth in mean PCC. To provide more observations, scores from measures at both times were combined where age groups overlapped between assessments taken at Initial and Re-assessment, i.e. age groups 4:6 to 4;11, 5 to 5;5 and

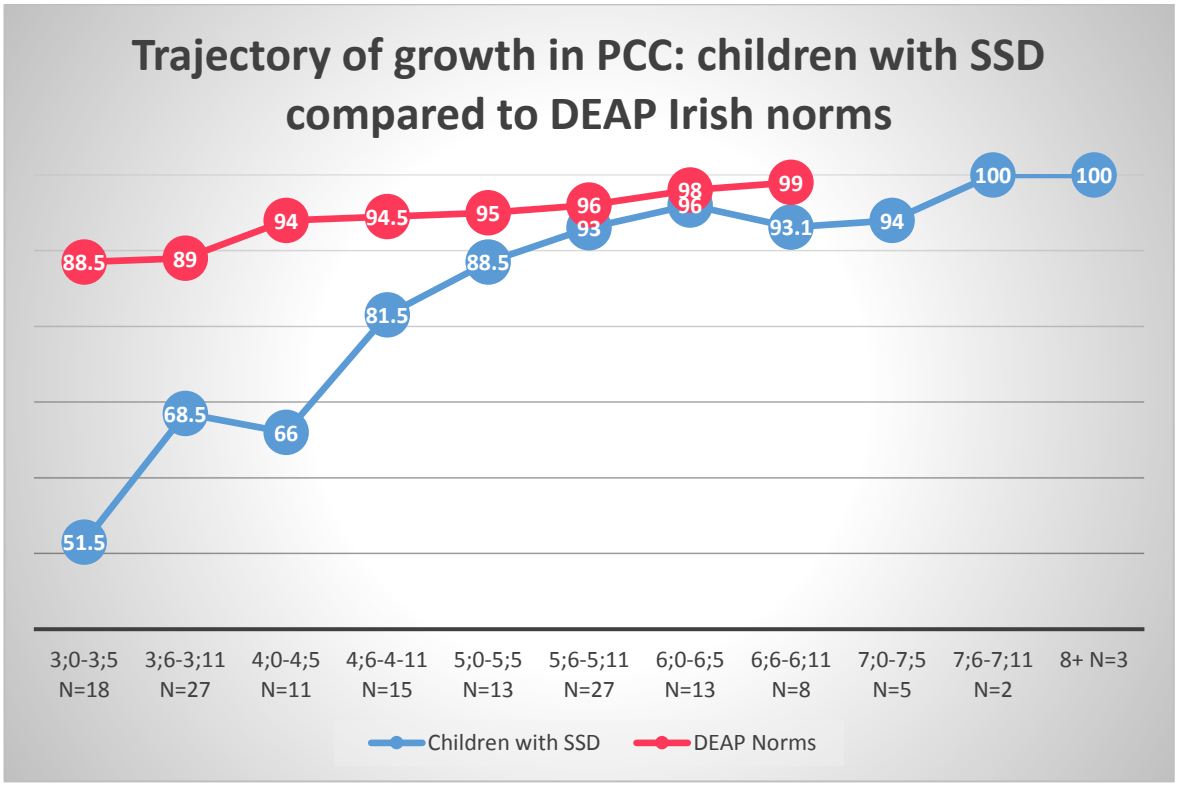
5;6 to 5;11. Figure 5-2 shows the trajectory of growth in mean PCC in the children with SSD in this sample compared to norms for the population from the Irish version of the DEAP. This shows a rapid growth in PCC between ages 3 and 5 for the children with SSD, tailing off to a slower increase and eventually convergence at ceiling, with the children with SSD reaching ceiling about one year later than norms for age.

Figure 5-1: Increase in PCC between Initial and Re-assessment, by age group



	PCC at Initial Ax	PCC at Re-Ax
3 yrs (N=40)	63	94
4 yrs (N=17)	71	96
5 years (N=9)	85	94

Figure 5-2: Trajectory of Growth in mean PCC by age group



5.1.2.1 Progress by Diagnostic Sub-group

It was expected that children with a diagnosis of CPD or IPD at initial assessment would make slower progress than those with PD, in accordance with the findings of Broomfield and Dodd (2005a). However, in fact they made greater mean gains than those with articulation disorder or PD, as shown in Table 5-1. A Kruskal Wallis test revealed a significant effect of diagnostic sub-group on increase in PCC, χ^2 (3) =14.58, *p*=.002. Mann-Whitney U tests showed the children with phonological delay made significantly less improvement than those with consistent phonological disorder, *U*= 224, *Z* =-2.45, *p*=.014, but the difference between the consistent and inconsistent disordered group was not significant (*U*= 43.5, *Z* =-1.82, *p*=.070). A Mann-Whitney U test showed no significant difference in the interval between Initial and Reassessment for the diagnostic sub-groups, *U*= 30.5, *Z* =-.11, *p*=.92, so this appears to indicate that the children with phonological disorder were increasing PCC faster than those with delay. However, there was a ceiling effect at reassessment, with 17 (26%) children having a PCC on the DEAP of 100%, and a further 13 (20%) having a PCC of 98% or more. These 30 children with PCC greater than 98% at reassessment comprised three with an original diagnosis of articulation disorder, 16 originally diagnosed with PD, 10 with CPD and 1 with IPD. Table 5-1 shows change in mean PCC for each sub-group between initial and reassessment.

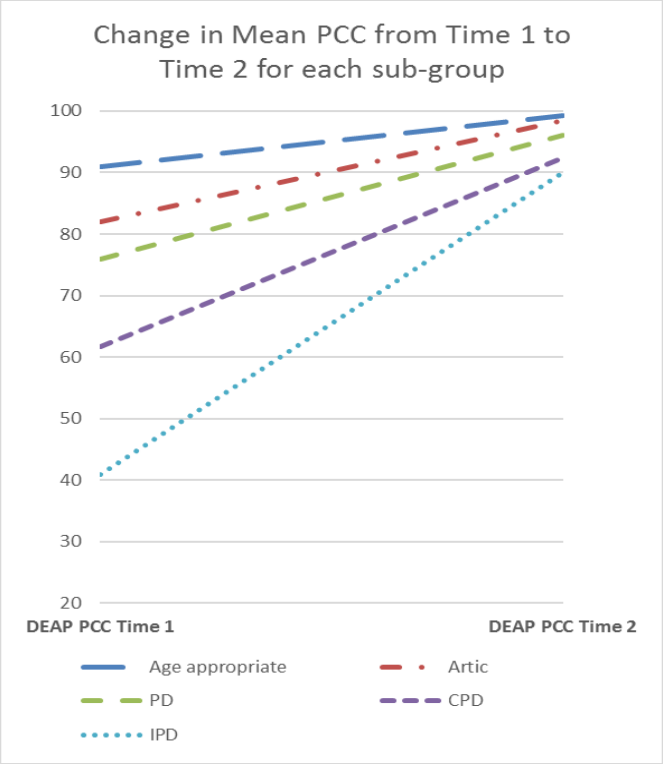
Table 5-1: Mean increase in PCC between initial and reassessment, by diagnostic sub-group at Initial Assessment

Diagnostic sub-group*	N	Mean PCC at Initial Ax (SD)	Mean PCC at Re-Ax (SD)	Mean increase (SD)	Min	Max
AA	9	90.7 (6.9)	99.3 (0.9)	**8.7 (6.5)**	0	20
Artic. Disorder	4	82 (2.4)	98.5 (3.0)	**16.5 (5.2)**	9	21
PD	32	76.0 (14.3)	96.01 (4.8)	**20.1 (12.6)**	0	60
CPD	23	61.3 (17.0)	92.4 (7.5)	**31.2 (17.3)**	3	62
IPD	7	43.1 (16.8)	90.0 (8.0)	**46.9 (19.8)**	18	70
*AA= Age Appropriate, Artic. Disorder = Articulation disorder, PD = Phonological Delay, CPD = Consistent Phonological Disorder, IPD = Inconsistent Phonological Disorder.						

Figure 5-3 shows an increase in PCC between Initial and Re-assessment for all diagnostic sub-groups, with a steeper incline for children with SSD than those with age appropriate speech, and a somewhat steeper gradient for those with disordered speech

than those with delay. However, given the ceiling effect and the relatively long period between assessments, it cannot be concluded that the rate of improvement was actually greater for the latter groups. Children with age appropriate or delayed speech may actually have hit ceiling much before the second assessment. At Reassessment there was still a significant difference in PCC between those originally classified as the PD and CPD groups, U= 248, Z = -.2.06, p=.039, and the PD and the IPD groups, U= 56.5, Z = -.2.06, p=.039, but this disappeared once the Bonferonni correction was applied.

Figure 5-3: Change in PCC between Initial & Reassessment by sub-group



As Age at Initial and Reassessment was shown to be significantly associated with chances of resolving the SSD, and the CPD and IPD groups were significantly younger than the PD group at Initial Assessment, it is likely that the gap in PCC between the PD and disordered groups at Reassessment was confounded by age, such that the disordered groups may have had a lower PCC because they were younger, rather than because they had a qualitatively different underlying condition.

5.1.3 Change in diagnostic subgroups between Initial and Reassessment.

To investigate whether children's membership of diagnostic sub-group remained stable, the children were re-classified into Dodd's diagnostic sub-groups at

Reassessment. Table 5-2 shows the number of children in each diagnostic sub group at Initial Assessment, and their subsequent diagnosis at Reassessment.

No child moved from a delay to a disorder between the two assessments. All children who had presented with age appropriate speech at initial assessment still presented so; however, parents of three children reported that they had received therapy, indicating either that development had plateaued at some point between Initial and Reassessment, or that therapy was being given for age appropriate speech processes. Counting children who presented with a dentalised production of alveolar fricatives as their only error as being in the resolved group (see rationale for this in Section 2.2.5.5), 3 out of 4 children (75%) originally diagnosed with an articulation disorder resolved, 29 out of 32 (90%) of those with PD resolved, 17 out of 23 (74%) of those with CPD and 3 out of 7 (43%) of those with IPD resolved. Thus a smaller proportion of those with disorder resolved than those with delay, but this may also be mediated by age: the disordered groups were significantly younger than the delayed.

Table 5-2 Distribution of diagnostic sub-groups at initial assessment and reassessment

Primary Sub-group Classification at Initial Assessment						
Sub-group	**AA**	**Artic**	**PD**	**CPD**	**IPD**	**Total N**
N at Initial Assessment	9	4	32	23	7	75
Primary Sub-group Classification at Reassessment						**Total N per sub-group at Reassessment**
AA (DD*)	9	3	29(7)	17(5)	3(2)	61
Artic	0	1	0	3	0	4
PD	0	0	3	2	4	9
CPD	0	0	0	1	0	1
IPD	0	0	0	0	0	0

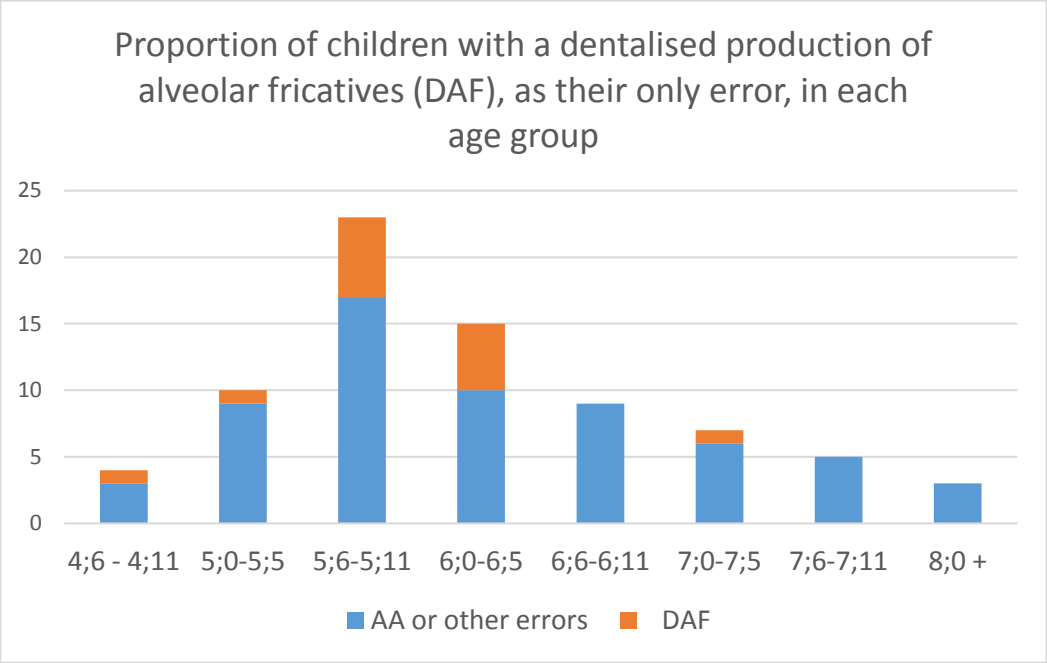
*Number included as AA who showed a dentalised production of alveolar fricatives as their only error are shown in brackets.

AA=age appropriate speech, Artic=articulation disorder, PD=phonological delay, CPD=consistent phonological disorder, IPD=inconsistent phonological disorder

Four children presented at Re-assessment with an articulation disorder only, in all cases lateralisation. Three of these had originally presented with a primary diagnosis of

CPD, and the fourth with articulation disorder. No child remained in the IPD sub-group, and only one child remained in the CPD group. The remaining children (n=61) either had age appropriate speech (n=52, including 14 with a dentalised production of alveolar fricatives), or presented with delayed error patterns only (n=9). Four children who had originally presented with IPD, and two who had presented with CPD, now presented with PD. Three children originally presenting with PD remained in that group. Dodd & Crosbie's (2005) classification does not specifically address a dentalised production of alveolar fricatives as a process, so the classification shown here is actually an amended one. It is assumed that Dodd would classify this production as fronting, so that these children would be classified in the delayed group. Therefore, for clarity, the number of children in the age appropriate group who presented with this dentalised production of alveolar fricatives is shown in brackets. Were these children assigned to the PD rather than the age appropriate group, 57.5% of children with SSD would have been classified as resolved, rather than the 79% so classified when an ungrooved dentalised production is counted as an age appropriate error in production of target alveolar fricatives, rather than a delayed error. However, Figure 5-4 below shows that this presentation diminishes rapidly after the age of 6;6, consistent with (Smit 1993)'s evidence that it is present in over 10% of TD children after the age of 6 and her suggestion that it diminishes naturally with age, consistent with an assumption that it is age appropriate in these younger children. (James 2001), in a study of 50 TD Australian children found a similar growth and decline in the use of dentalisation, but at somewhat younger ages, with a peak at age 4 and disappearance by age 6 and 7. However in her sample, dentalisation was not the only error shown.

Figure 5-4: Distribution of children with a dentalised production of alveolar fricatives as their only error, by age



5.1.4 Sub-groups within the diagnostic categories

As outlined in section 2.4.2, two patterns of impairment that did not fall into Dodd's diagnostic sub-groups as strictly defined in (Dodd and Crosbie 2005) were found in the sample. The first is shown by children who use only age appropriate error patterns but who still have a score of PCC in the 13th percentile or below (scaled score of 6 or below) according to the norms for the appropriate age group on the DEAP phonology sub-test. This group were assigned to the PD diagnostic sub-group. Within the PD group were also a group of children who had PCC in the normal range for age, but who still have at least one consistent speech error pattern that occurs in at least 10 percent of children, but which is typical only of children of a younger chronological age group (by at least 6 months).

The second pattern not described by Dodd & Crosbie (2005) is shown by children who consistently produce some words with atypical errors, although this does not reflect any consistent error pattern in their speech. These children were assigned to the CPD group. Outcomes for children in these different sub-groups within Dodd's categories were compared to investigate whether these sub-groups showed different patterns of resolution, as shown in Table 5-3.

Table 5-3: Outcomes for groups showing differing patterns of impairment within Dodd's PD and CPD diagnostic sub-groups

Outcomes for groups showing differing patterns of impairment within Dodd's PD and CPD diagnostic sub-groups				
Pattern at initial assessment		**Status at Re-assessment**		**Total**
		persisting SSD	**resolved SSD**	
PD	PCC<13th percentile and uses delayed error patterns	2	12	14
	PCC<13th percentile but only age appropriate error patterns	0	4	4
	PCC within normal limits but uses delayed error patterns	1	13	14
CPD	Uses at least one consistent atypical error pattern	3	11	14
	Makes atypical errors in at least 5 words consistently on the DEAP Phonology sub-tests, but not following a discernible pattern	3	6	9

A Chi-square test showed no significant differences between groups in the proportion who resolved, but some cells contained very small numbers. However, visual inspection also shows similar rates of resolution for groups within the PD group, and within the CPD group, supporting the allocation of children with these patterns of impairment to these diagnostic sub-groups. However there does not seem to be an association of any of these patterns with differential outcomes, within those groups.

5.2 Outcomes with and without therapy

As most children were no longer involved with SLT services by Re-assessment, files reporting whether therapy had been received, and type and amount, were not available. Therefore, classification of whether children had received therapy or not was designated by parent report.

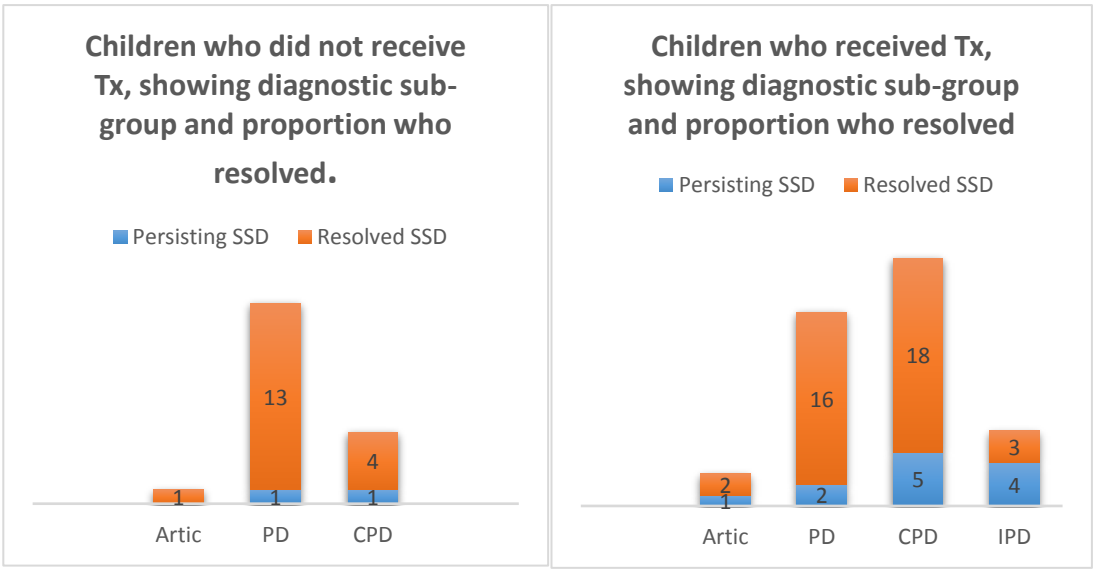
Children were considered to have been in receipt of therapy if they were reported to have received at least one session of clinician-directed therapy (n=22), or if they had received a home programme for parents to carry out (n=12), or both (n=12). Children were considered to have had no therapy if they had not been seen by a clinician since initial assessment (n=17) or if they had only had a review appointment (n=3). Thus 46

children with SSD had received therapy and 20 had received no therapy. Seventy per cent of boys and 69% of girls had received therapy, indicating no gender bias.

5.2.1 Resolution rate and changes in severity of SSD by therapy condition

Figure 5-5 shows that, of those who had received no therapy, 18 (90%) had resolved. Of those who received therapy, 34 (74%) had resolved. Given the unexpected result that a higher proportion of children had resolved without therapy than with therapy, the two groups were compared on age, presence of demographic risk factors or co-morbid language disorder, severity according to norms on the DEAP (i.e. adjusted for age) and diagnostic sub-group at initial assessment, to investigate whether those who received therapy were significantly different from those who had not.

Figure 5-5: Children who did and did not receive Tx, showing proportion who resolved.

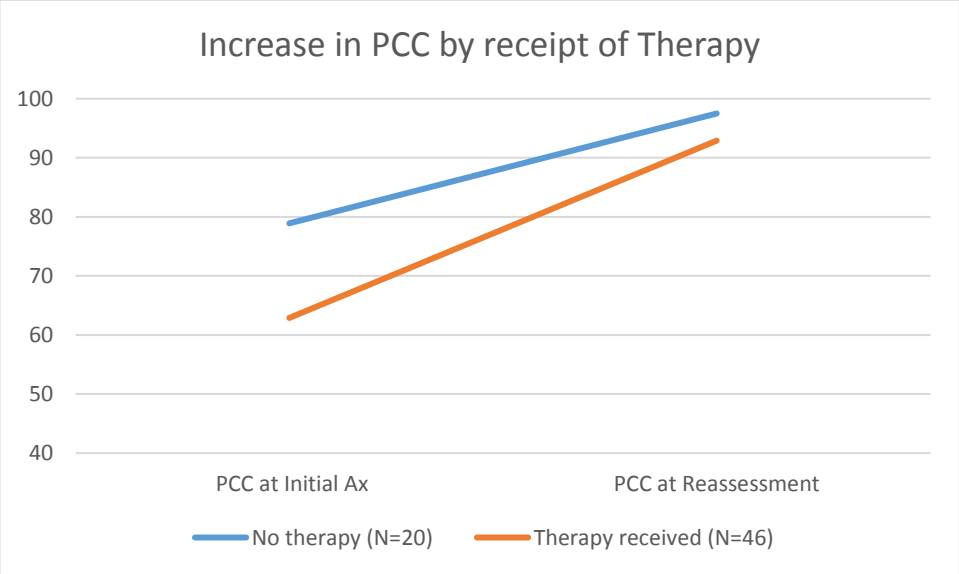


The mean amount of time between initial and reassessment did not differ between groups: those who had no therapy: Mean 24.7 months, SD 4.3, range 18-35; those who had therapy: Mean 24.3 months, SD 2.9, range 17-30, U= 425, Z = -.49, p=.63. A Mann Whitney U test showed no significant difference in age at initial assessment between those who had received therapy (mean 48 months, SD 8.6, range 37-69) and those who had not (mean 49 months, SD 9.6, range 36-69), U= 454, Z = -.08, p=.93, nor in age at reassessment (received therapy: mean 6;1, SD 8.9, range 56-96; no therapy: mean 6;2, SD 10.7, range 60-97), U= 457, Z = -.04, p=.97. Chi square tests were used to investigate whether the two groups differed in presence of risk factors.

These revealed no significant differences between the groups in terms of a positive family history of speech, language or literacy difficulties, history of birth difficulties or late onset of speech, history of persisting ear infections or level of maternal education. Ten (22%) of the 46 children who received therapy presented with comorbid language impairment at initial assessment, compared to five (25%) of the 20 who had no therapy. A Chi square test showed this difference was not significant.

However, regarding diagnostic factors, those who had received therapy had a significantly more severe SSD in terms of PCC (mean 63, SD 19.7) and DEAP scaled score (i.e. adjusted for age), (mean 4.7, SD 2.14, range 3-10) than those who had not (PCC: mean, 79, SD 9.03, DEAP Scaled Score: mean 6.7, SD 2.18, range 3-11), PCC: U= 237, Z=-.3.11 p=.002, DEAP Scaled Score: U= 232, Z=-.3.29 p=.001. Although a smaller proportion of children who received therapy resolved than those who did not, Figure 5-6 shows that receipt of therapy was associated with a significantly bigger mean gain in PCC across the sample as a whole, U=310, Z=-2.1, p=.037. Those who received therapy had a mean gain of 30 percentage points (SD 18.8, range 0-70) compared to a mean gain of 18.6 percentage points in those who did not (SD 8.5, range 4-34).

Figure 5-6: Increase in PCC by therapy condition



Again a ceiling effect must be taken into account, as the no therapy group was very near to ceiling (i.e. performance in the average range for age) at Re-assessment, as can be seen in Table 5-4.

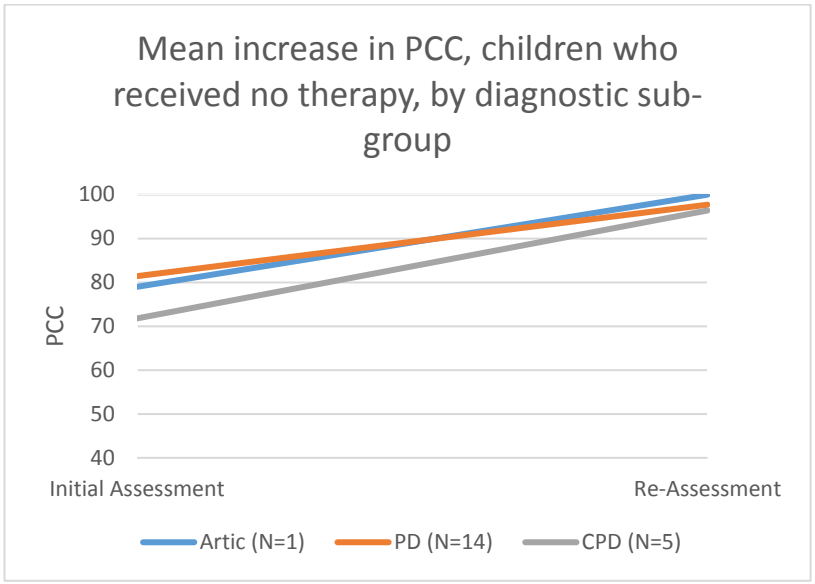
Table 5-4: Mean PCC and severity score at reassessment by therapy condition

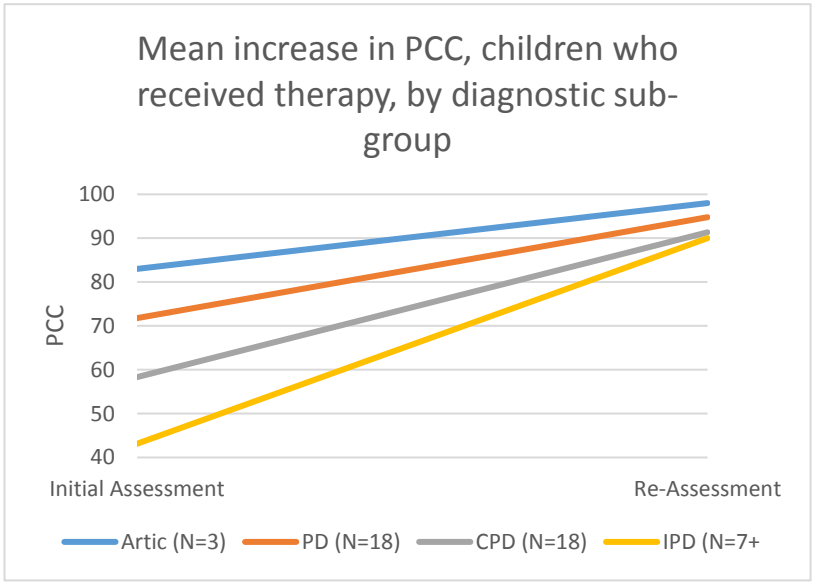
PCC and DEAP Scaled Score at Initial (In-Ax) and Re-assessment (Re-Ax) by receipt of therapy.								
	No therapy				Received therapy			
	PCC at Initial Ax	DEAP* Severity at In-Ax	PCC at Re-Ax	DEAP* Severity at Re-Ax	PCC at Initial Ax	DEAP* Severity at In-Ax	PCC at Re-Ax	DEAP* Severity at Re-Ax
Mean	78.9	6.6	97.5	10.3	62.9	4.7	92.91	8.09
SD	9.03	2.1	3.364	1.38	19.71	2.1	7.055	3.189
Min.	66	3	87	7	16	3	72	3
Max.	95	10	100	12	94	10	100	15
* Scaled Score, 10 = average. SD= Standard Deviation								

Only eight children made gains of less than 10 percentage points over the 2 years: three of these had a PCC of 100 and one of 99 at Re-assessment, so had hit ceiling. Two had a lower PCC due only to dentalised production of alveolar fricatives, one had a lateralised production of alveolar fricatives and the last was the child who remained with a diagnosis of CPD. Thus, apart from these last two, all the children made clinically significant gains over the two years, whether or not they received therapy.

Figure 5-7 shows mean increases in PCC for each diagnostic sub-group, comparing those who received therapy with those who did not. A larger proportion of children diagnosed with a phonological disorder (CPD or IPD) received therapy than those with a phonological delay (PD): all of the children with Inconsistent Phonological Disorder received therapy, and 78% of those with Consistent Phonological Disorder had therapy, compared to only 56% of those with a Phonological Delay. Receipt of therapy was associated with a trend for greater increase in PCC in both groups: children with phonological delay who received therapy improved by a mean of 23 percentage points, compared to a mean improvement of 16 points in those without therapy. However, a Mann-Whitney U test showed this difference was not significant, U=97, Z=-1.1, p=.157. Children with consistent phonological disorder who received therapy made a mean improvement of 33 points compared with 25 points in those who did not, but this was again shown to be non-significant by a Mann-Whitney U test, U=33.5, Z=-.86 p=.391. The children with inconsistent phonological disorder made the biggest gains with a mean improvement of 47 points; however they were both the youngest and the most severe at initial assessment. As all had received therapy, therapy condition could not be compared.

Figure 5-7: Mean increase in PCC between Initial and Reassessment, by therapy condition and diagnostic sub-group

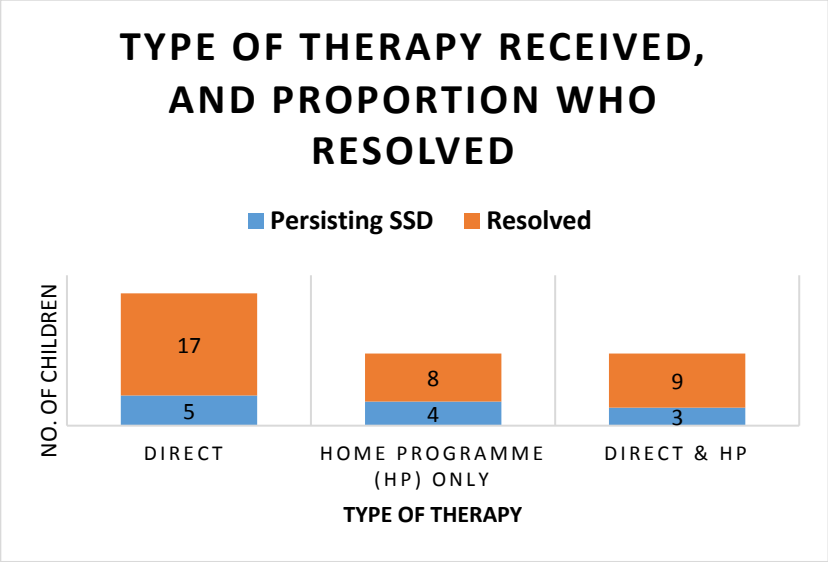




5.2.2 Type of therapy received

Figure 5-8 shows that 48% of children who received therapy had direct, clinician directed therapy (DIRECT), 26% received a home programme to be carried out by parents (HP), and 26% had both clinician directed therapy and a home programme (DIRECT & HP).

Figure 5-8: Type of therapy received & proportion who resolved



5.2.3 Characteristics of children who received no therapy

The two children in this group who had not resolved were the youngest in the group who did not receive therapy, and the fifth and sixth youngest in the sample as a whole. One was a girl who had also presented with language impairment at initial assessment, as well as an additional diagnosis of selective mutism. At initial assessment, she had presented with a mild delay, and still presented with a delayed error pattern at reassessment, although her PCC had increased by 20 percentage points. The second was a boy who had presented with severe CPD at initial assessment. He had resolved all his phonological difficulties at reassessment, but remained with a lateral lisp.

Of those who had resolved, one child had presented with an articulation disorder (substitution of a labiodental approximant for /w/, and a dental production of /s/ in all contexts, not stimulable for either sound), at initial assessment. Four children with mild CPD and 13 with PD (3 severe, 1 moderate and 9 mild, according to the DEAP severity scale, i.e. adjusted for age) resolved.

The disordered group made greater mean gains (25 percentage points, SD =9.2, Min = 11, Max = 34) than the delayed group (16 percentage points, SD =7.7, Min = 4, Max = 28) without therapy, but this difference was not statistically significant when tested with a Mann-Whitney U test, U=16.5, Z=-1.7, p=.086. In addition, there was a ceiling effect as the CPD group had been more severe to start, with a mean PCC of 72% compared to 81% for the delayed group.

5.2.4 Characteristics of children who had received therapy

Thirty-four of the 46 children with SSD who received therapy had age appropriate speech at reassessment: two had presented at Initial Assessment with an articulation disorder only, being non-stimulable for post alveolar fricatives/affricates. Three had originally presented with IPD (1 severe, 1 moderate, 1 mild according to the DEAP severity scale, i.e. adjusted for age), and 13 with CPD (8 severe, 2 moderate, 3 mild). Sixteen had originally presented with PD (6 severe, 1 moderate, 9 mild). Severity was classified according to scores on the DEAP Phonology sub-test, i.e. related to norms and adjusted for age (see section 3.3 for further details on classification of severity).

Of the 12 (26% of treated children) who had not resolved, one had received a diagnosis of articulation disorder at Initial Assessment and continued to present with a lateral lisp. Two children with original diagnoses of CPD had resolved their phonological difficulties and presented at Re-assessment with lateral lisps. Two with severe CPD and 4 with IPD (1 mild, 1 moderate, 2 severe) at Initial Assessment presented with PD at Reassessment. Two had presented with PD at Initial Assessment; one remained with a severe delay, the other moved from moderate to mild. The final child had presented with a severe CPD at Initial Assessment; her presentation had not changed, still showing atypical errors and having made an increase in PCC of only 3 percentage points. She also had a language impairment as well as reading and other academic difficulties.

5.2.5 Limitations regarding influence of therapy on outcomes

These findings suggest that SLTs were prioritising children with more severe deficits, and those with disorder rather than delay, for therapy. Details of the therapy received came from parent report, so no investigation could be made regarding the type of therapy offered. Some children received as little as one session, while others received only a home programme with no proven efficacy and little parent training. Some received therapy at weekly intervals, and others monthly. Thus this analysis was not designed to, and is not able to contribute to any discussion of the efficacy of therapy. What it does show is that children with SSD in this sample did resolve without therapy even, in a few cases, if their impairment was in the severe range. Both children with a diagnosis of disorder and delay resolved without therapy.

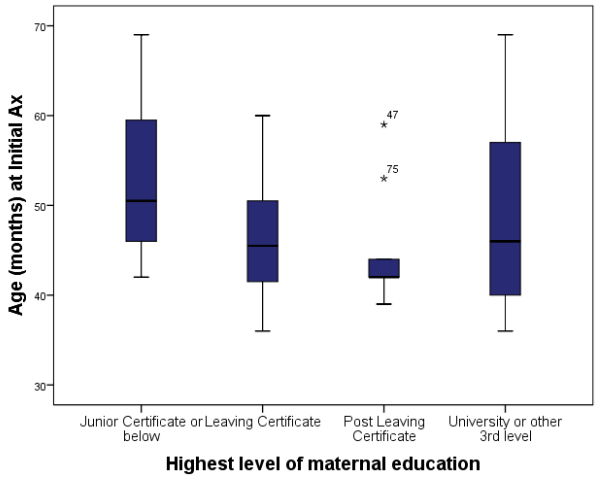
5.3 Predictors of outcome at reassessment

5.3.1 Demographic and case history factors associated with SSD

The distribution of demographic risk and concurrent factors associated with SSD was examined to investigate whether there was an influence on outcome.

The level of maternal education was found to have an effect almost the reverse of that suggested in the literature considering level of maternal education as a risk factor (Campbell *et al.* 2003): while 100% of the children whose mother's highest level of education was Junior Certificate, resolved (n=7), 80% of those with a Leaving Certificate (n=20), and 83% of those with further education (n=6), only 73% of those with mothers who had a university degree or equivalent (n=33), resolved. However a Chi Squared test showed this difference was not significant, $\chi 2$ (3) = 2.87, p = .413. The result in this sample may be an interaction with the presentation of the children of each group. Firstly, the children of mothers with the lowest level of education were older than the other groups (see Figure 5-9), possibly reflecting later referral, for instance referral by teachers rather than parents. As results for this sample showed that children who resolved were older, age might account for the fact that all resolved.

Figure 5-9: Age of children at Initial Assessment by mothers' education



There was no interaction between severity of SSD, as measured by the DEAP scaled score, and level of maternal education as revealed by a Kruskal-Wallace test, χ^2 (3) = 3.86, p=.277, so difference in severity levels at initial assessment does not explain results. Results may simply be due to the small sample size, or referral bias, as studies

reporting an effect of SES have been large epidemiological studies rather than clinical studies (Campbell *et al.* 2003, Eadie *et al.* 2015, Wren *et al.* 2016).

Parent reported motor delay was not associated with persistence of SSD in this sample: of the 7 children in the sample reported by parents to have a history of delayed motor skills, 6 had age appropriate speech by reassessment (one of whom had already had age appropriate speech at Initial Assessment) and one did not present for Re-assessment.

Table 5-5 below shows the distribution of other risk factors between the resolved and unresolved groups. Chi square tests showed there was no significant difference between the resolved and persisting group on any of these factors.

Table 5-5: Distribution of demographic factors in resolved and unresolved groups

	Gender				**Positive Family History**		**Delayed motor milestones**		**Delayed onset of speech**		**Birth difficulties**		**Insertion of grommets**	
	Male		**Female**											
Age appropriate at initial and reassessment, N=9	7		2		6		1		1		0		0	
Resolved SSD at reassessment N=52	32	62%	20	38%	23	44%	5	10%	14	27%	10	19%	1	2%
Persisting SSD at reassessment N=14	8	57%	6	43%	7	50%	0	0%	4	29%	1	7%	1	7%

Since both the group with motor delay and those whose mothers left education before Leaving Certificate were small (6 and 7 respectively), and results were not consistent with those of other larger studies, it is likely that the small sample size played a part, so these factors were therefore not included in further analysis.

A Chi square test showed number of ear infections did not differentiate the groups, with similar proportions reporting occurrence of ear infections in both the persisting SSD and resolved groups, as shown in Table 5-6.

Table 5-6: Number of ear infections by speech status at reassessment

No. ear infections by status at Re-assessment								
	None		1 to 3		4 to 6		6+	
Age appropriate at initial and Re-assessment N=9	4	44%	3	33%	2	22%	0	
Resolved SSD at reassessment N=52	23	44%	21	40%	4	8%	4	8%
Persisting SSD at reassessment N=14	7	50%	4	29%	1	7%	2	14%

5.3.2 Influence of age

Age and PCC were strongly correlated at both initial, (r_s (83) = .621, p < .001) and reassessment (r_s (73) = .419, p < .001), with a large effect size at initial assessment and a medium effect size at reassessment (Cohen, 1988). Squaring the correlation coefficients indicated that 38.6% of the variance in PCC was explained by age at initial, and 17.6% at reassessment, when a large number of cases reached ceiling. A Mann-Whitney U test was used to investigate whether children who had resolved their speech difficulties were different in age to those who had not. This showed that the children who had resolved were significantly older at Re-assessment (mean age 6;2) than those who had not (mean age 5;8), (U= 58, Z =-3.46, p=.002).

5.3.3 Predictors associated with SSD that resolved

For the 66 children with SSD, univariable and multivariable logistic regression analyses were used to obtain odds ratios (ORs), 95% confidence intervals and likelihood ratio p values for the associations between resolved SSD at Re-assessment and the demographic and risk variables, and results of assessment on the various speech, language and cognitive tasks. Both continuous and categorical variables were used. In the first stage, all variables were tested individually for their association with the outcome variable, that is, age appropriate speech at reassessment. Variables with a p value of <.10 in the univariable analyses were retained for use in the multivariable analysis. This tolerant level was used to ensure that any variable with potential to influence outcome was included. Given the difficulties in assessment of children of the age of those in this study, there were missing data for some of the variables collected. The number of children for whom data was available is shown in the N column in Table 7. In each analysis, the maximum data available for the variables under investigation were used. To quantify the predictive strength of the final model, the area

under the receiver operating characteristic curve (AUC) is reported. This quantifies the ability of the model to classify the children correctly as having resolved or persistent SSD. All statistical analysis was carried out using SPSS Version 22 (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp).

The variables for which the *p* value following univariable regression was less than .10 are shown in Table 5-7. These variables were taken forward to the multivariable analysis.

Table 5-7: Variables from univariable analysis showing association with resolved SSD at p < .10 level

Descriptive statistics and univariable regression model results for demographic, risk factors, and assessment results associated with resolved SSD, where *p* < .10

Variable	Category	N	Resolved SSD*	Persistent SSD*	Odds ratio	95% Confidence Intervals		*p* value
						Lower	Upper	
Age at Initial Ax (months)		66	49.4	44.6	1.082	0.989	1.184	0.085
Diagnostic sub-category	IPD	7	3	4	1			0.030
	PD	32	29	3	12.89	1.906	87.170	0.009
	CPD	23	17	6	3.778	0.648	22.017	0.139
PCC (DEAP Phonology sub-test) at Initial Ax		66	70.2	58.6	1.033	1.001	1.066	0.045
PCC Multisyllabic words: spontaneous at Initial Ax		47	79.8	69.2	1.047	0.996	1.1	0.07
Severity: DEAP Standard Score at Initial Ax		66	5.6	4.2	1.364	0.988	1.883	0.059
Non Standardised tests of cognitive skills**								
Non-word Discrimination at Initial Ax		50	11.2	7.3	1.322	1.058	1.651	0.014
FIST1 at Initial Ax		64	13.4	9.7	1.197	1.028	1.394	0.021

*Figures in yellow cells show N, in blue cells show mean score
**These scores are not adjusted for age, therefore, in the univariable analysis, a hierarchical regression was performed, with age entered in block 1 and test score in block 2

Descriptive statistics and results of the univariable analysis for all variables are shown in Appendix 5.A. Variables that did not show a significant association in the univariable analysis with resolved SSD at Re-assessment were: gender, family history of speech, language or literacy difficulties, delayed onset of speech, birth difficulties, number of ear infections, receipt of therapy, presence of more than 4 atypical errors in the DEAP Phonology sub-test, use of consistent atypical phonological processes, co-

morbid language impairment, PCC in multisyllabic words in imitation, scaled scores on the DEAP oro-motor subtests (DDK, isolated movements, sequenced movements), receptive vocabulary standard scores, scaled scores on the CELF P2 (UK) sentence structure, word structure or expressive vocabulary tests, or the overall CELF P2 (UK) core language score, standard scores of visual motor integration (BVMI), scaled scores for syllable awareness and alliteration awareness on the PIPA, and the scores for the following tests when controlled for age: Number Concepts, FIST2, Auditory Memory Wordspan and Pattern Recognition.

In the multivariable analysis $p<.05$ was the level used for significance. In addition, gender and family history were retained as important covariates given evidence in the literature that they might be potential confounders even though they were not significant at the .10 level, again to ensure that no potentially influential variables were excluded. Thus a hierarchical multiple regression was carried out, with gender and family history entered into the model in Block 1, and the remaining variables shown in Table 5-7 as Block 2. Predictors were included in a backward elimination multivariable logistic regression model, specifying the 10% level as the criterion for removing predictors.

The model with all univariable predictors entered was unable to converge on a solution. For predictors to be clinically most useful, they need to have predictive validity for the individual child, as well as show a tendency for the group. Therefore, scatter plots of the two novel test variables found to be predictors by the univariable analysis, Non-word Discrimination and the FIST1, were examined to discover whether they were actually useful in identifying individual children who would resolve by Re-assessment, and should be retained in the model. Scores on the FIST1 (see Figure 5-10) were nearly at ceiling for many participants, with a median score of 14 out of a possible 15, showing small discriminatory power. Four children who did not resolve attained maximum scores, while others with very low scores did resolve, showing no predictive value for the individual child. Scores on the test of Non-Word Discrimination (see Figure 5-11) looked more promising, as all cases that scored above the median score but one, had resolved. Case data revealed that this child had resolved his phonological difficulties, but retained a lateral lisp. Thus score above the median on the Non-Word Discrimination Test at Initial Assessment was a better predictor of resolved SSD at Re-assessment. As it had not shown good discriminatory power,

FIST1 was excluded from the model, and Non-word Discrimination was retained. The model was run again, including the remaining variables: Age at Initial Assessment, PCC and DEAP Standard Score, PCC in multisyllabic words in spontaneous production, Diagnostic sub-group and scores on the test of Non-word Discrimination.

Figure 5-10: Scatter plot showing scores on FIST1 for children with resolved/persisting SSD

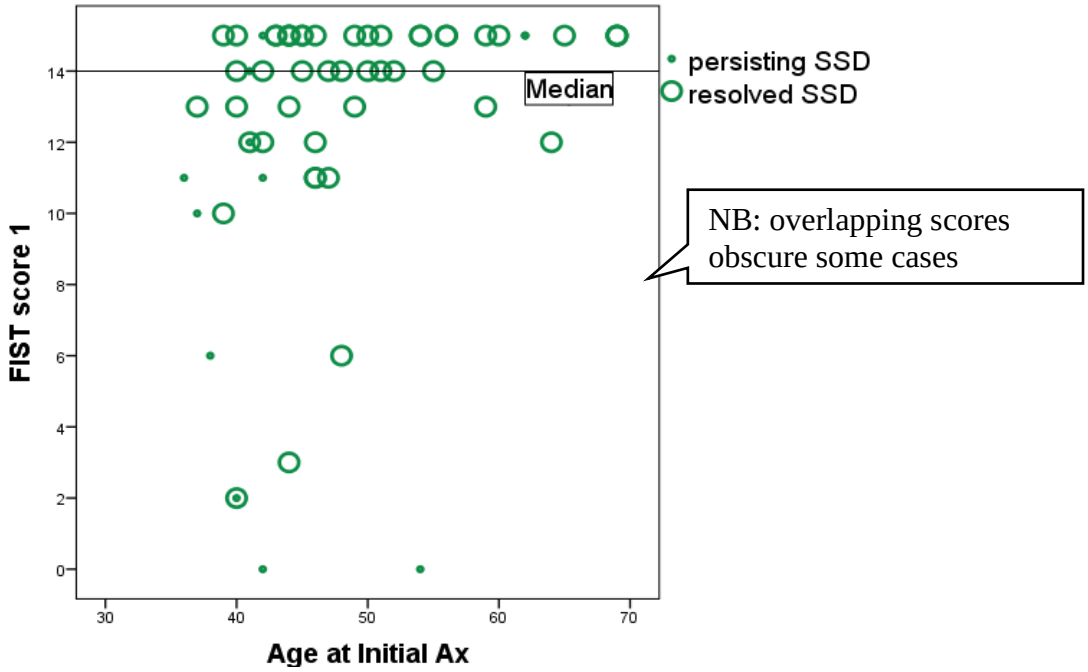
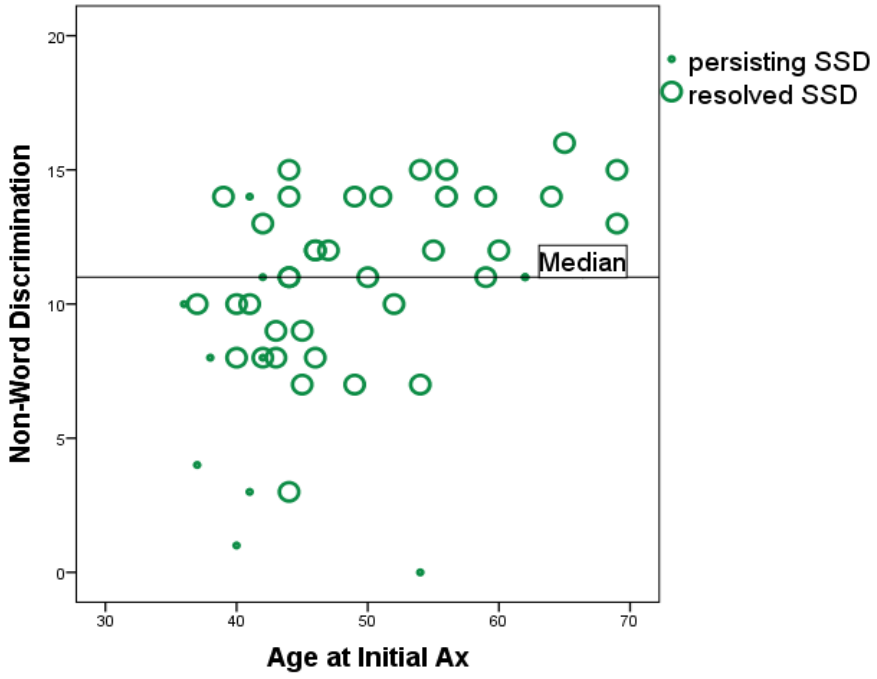


Figure 5-11: Scatter plot showing scores on Non-Word Discrimination for children with resolved/persisting SSD



The regression analysis was conducted using all but one of the children who had a full set of data: 34 with resolved SSD and 9 with persisting SSD. There was only one child with a full set of data with a diagnosis of articulation disorder, so this child and this category was omitted from the analysis. The final model, retaining Gender, Family History, Diagnostic sub-group, DEAP Standard Score, and scores on the test of Non-word Discrimination was statistically significant, χ^2 (6, N=43) = 23.67, p=.001. This model explained between 42% (Cox and Snell R square) and 66% (Nagelkerke R squared) of the variance and classified 91% of children correctly, with 91% sensitivity and 89% specificity. As shown in Table 5-8, only one of the independent variables, Non-word Discrimination, made a unique statistically significant contribution to the model. The odds ratio for this predictor was 1.56, indicating that for every point scored on the test, the participant was 1.56 times more likely to resolve their SSD. The Area under the Receiver Operating Characteristic Curve (AUC) for this model was .859, showing a good level of discrimination of resolved versus persisting SSD status.

Table 5-8: Logistic Regression of Resolved SSD status at Re-assessment

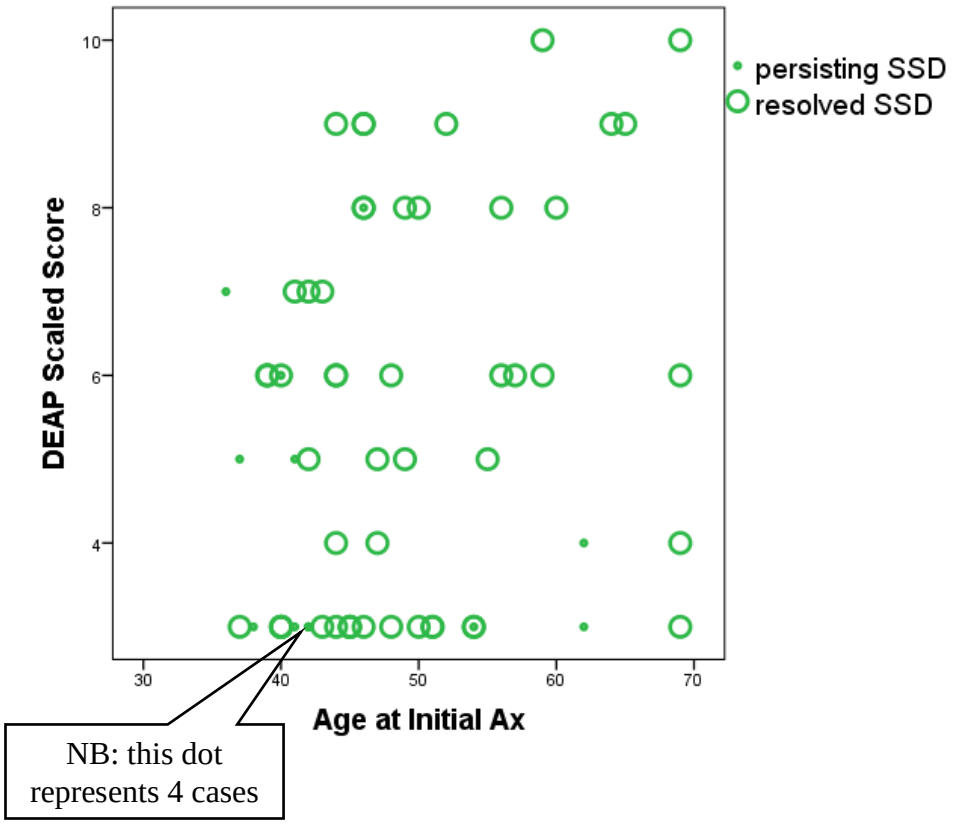
Logistic Regression of Resolved SSD status at Re-assessment					
Predictor	**Category**	**Odds ratio**	**95% Confidence Intervals**		
			Lower	**Upper**	***p* value**
Gender	**Female**				
	Male	NA			.745
Family History	**No**				
	Yes	NA			.292
Diagnostic sub-group	**IPD**	NA			.168
	PD	NA			.108
	CPD	NA			.560
Severity according to DEAP norms	**DEAPSS**	NA			.131
Non-word Discrimination		1.560	1.002	2.428	.049
Variable(s) entered on step 1: Age1, CollDiag, PCC, DEAPSS, NWD, MSS.					
Note: NA = not applicable, as the *p* value at this stage of the analysis was >.05					

This model was clinically useful at an individual level, in that a score above the median on the test of Non-Word Discrimination predicted certain resolution of phonological, if not articulation, difficulties in this sample. A Spearman's rank-order correlation was run to determine the strength of the relationship between Non-Word Discrimination scores at Initial Assessment and PCC, both concurrently and at Re-assessment. There was a significant, medium positive correlation between Non-Word

Discrimination scores at Initial Assessment and concurrent PCC, ($r_s(52) = .385$, p = .005), and a significant medium, bordering on strong, positive correlation between Non-Word Discrimination scores at Initial Assessment and PCC at Re-assessment, ($r_s(48) = .499$, p <.001).

Neither PCC on the Phonology sub-test of the DEAP nor the Standard Score at initial assessment were predictors of resolved SSD in this model, with the majority (21/30, 70%) of those classified as severe (DEAP Standard Score of 3 or 4) at Initial Assessment having resolved their SSD by Re-assessment, including three who had not received any therapy, and two who had only received a home programme. Seventeen of the 19 children (89%) who had presented at Initial Assessment with PCC within one standard deviation of norms for age, but who were classified as having SSD due to use of delayed speech error patterns or distortions, had resolved their SSD by Re-assessment. Of the 2 who had persisting SSD, one continued to use a delayed process and one had an articulation disorder. Figure 5-12 shows a scatter plot of severity (DEAP Scaled Score) at Initial Assessment by status at Re-assessment.

Figure 5-12: DEAP Scaled Score at Initial Ax by resolved/persistent SSD status at Re-assessment



Thus children who resolved their SSD two years after initial assessment in this study were more likely to have high scores on the test of Non-Word Discrimination, and to have PCC within normal limits relative to the normative sample on the DEAP, at initial assessment. Age was also a factor, with the resolved group being significantly older than the group with persisting SSD. However, in the logistic regression, only Non-word Discrimination ability was a significant predictor of resolution of SSD by Re-assessment.

At an individual level, no child who presented with a lateral lisp resolved this.

Table 9 shows the demographic characteristics and scores on selected variables of children who did not resolve their SSD by Reassessment.

Table 5-9: Demographic Characteristics and scores on predictor variables of children who did not resolve

Children who did not resolve their SSD: characteristics and scores on predictor variables												
ID	**Gender**	**Age at Initial Ax**	**Age at Re-Ax**	**Diagnosis Initial Ax**	**Secondary Artic**	**Co-morbidity**	**Family History**	**DEAP PCC**	**DEAP SS**	**Non-Word Discrim.**	**Diagnosis Re-Ax**	**Therapy**
12312	F	36	62	PD	No	SSD+LI	No	73	7	10	PD	No
11304	M	37	59	PD	No	SSD only	No	58	5	4	PD	Yes
11328	M	38	63	IPD	No	SSD only	Yes	16	3	8	PD	Yes
11322	M	40	65	IPD	No	SSD+LI	Yes	66	6	1	PD	Yes
12307	F	41	58	IPD	No	SSD only	No	57	5	3	PD	Yes
11310	M	41	59	CPD	No	SSD only	No	42	3	14	Artic/lat.*	Yes
1303	M	42	68	IPD	No	SSD only	Yes	39	3	--	PD	Yes
2302	F	42	68	CPD	No	SSD only	No	61	3	--	PD	Yes
11311	M	42	60	CPD	Lateralis-ation	SSD only	No	66	3	8	Artic/lat.	No
11329	M	42	68	CPD	No	SSD only	Yes	55	3	11	PD	Yes
12319	F	46	71	Artic/lateralisation		SSD only	No	85	8	7	Artic/lat.	Yes
11419	M	54	78	PD	No	SSD+LI	Yes	61	3	--	PD	Yes
12501	F	62	81	CPD	No	SSD+LI	Yes	69	4	11	CPD	Yes
12505	F	62	85	CPD	Lateralis-ation	SSD only	Yes	72	3	11	Artic/lat.	Yes
* this child was stopping fricatives at initial assessment, so lateralisation was not apparent at that point												

5.4 Differences in abilities between the resolved and persistent SSD group in DDK, vocabulary and auditory memory measured at Re-assessment

There may be skills related to the acquisition of speech, which are not predictive of persisting SSD at young ages, because of the large range of ability in very young children, but which nevertheless, differentiate groups of children with persistent rather than resolved speech at later ages. Older children with SSD have been shown in the literature to have poorer vocabulary, oral-motor skills and auditory memory than age-matched controls (Waring *et al.* 2016, Wren *et al.* 2016), so scores on measures of these were compared for the persistent and resolved groups.

5.4.1 Test of DDK rate

The DEAP Oro-motor DDK sub-test was re-administered at Re-assessment to measure speech motor skills. Table 5-10 shows that there was no significant difference at Initial Assessment between the scores of the children with persistent SSD and those whose SSD subsequently resolved. In fact, very few of the children in the study scored outside normal limits on oro-motor tasks: of the 76 children who completed the DDK task at initial assessment, only 4 (5%) had scores below 1SD, and one of those had age appropriate speech. Of the three with SSD, two did not resolve their SSD by Re-assessment, despite DDK scores within normal limits at that time, and one did not present for Re-assessment. The DDK skills of the children whose SSD resolved largely maintained similar scores at Re-assessment; however the scores of the children with persistent SSD had not maintained progress with age, so that by Re-assessment, they were significantly worse than those in the resolved group. For both groups, the range of scores had narrowed by Re-assessment, and all of the resolved group scores were within normal limits (WNL) according to norms on the DEAP. However, three in the persisting group had DDK scores below 1SD at Re-assessment, despite having been WNL at Initial Assessment. This may indicate that, for at least a sub-set of children whose SSD persists, there may be an association with slower development of oro-motor skills for speech. This would indicate the validity of Shriberg's classification of the SDCS subtype termed motor speech disorder—not otherwise specified.

Table 5-10: DDK scores at Initial & Re-assessment, resolved vs persistent SSD

Comparison of DDK scores at Initial & Re-assessment for the resolved vs persistent SSD groups					
		Age Approp.	**Persistent SSD (PSSD)**	**Resolved SSD (RSSD)**	**Significance test PSSD vs RSSD**
DEAP DDK standard score at Initial Ax:	Mean (SD) Min Max	10.50 6 13	9.31 (2.72) 3 13	10.22 (1.62) 6 14	*U*=279, *Z*=-.81 *p*=.42
DEAP DDK standard score at Re-Ax.	Mean (SD) Min Max	10.33 8 11	8.29 (1.54) 6 10	10.23 (1.2) 7 11	*U*=100, *Z*=-4.4 *p*<0001

5.4.2 Receptive Vocabulary

A comparison of scores of receptive vocabulary showed no significant differences between the resolved and persistent SSD group at either Initial or Re-assessment, as shown in Table 5-11. Both groups included some children above and below 1 standard deviation from the mean.

Table 5-11: Comparison of Receptive vocabulary scores at Initial & Re-Ax for the resolved vs. persistent SSD groups

Comparison of BPVS III scores at Initial & Re-assessment for the resolved vs persistent SSD groups					
		Age Approp.	**Persistent SSD (PSSD)**	**Resolved SSD (RSSD)**	**Significance test PSSD vs RSSD**
BPVS III standard score at Initial Ax:	Mean (SD) Min Max	100.0 (9.50) 86 111	96.7 (12.9) 70 112	99.8 (11.8) 73 132	*U*=329, *Z*=-.44 *p*=.66
BPVS III standard score at Re-Ax.	Mean (SD) Min Max	95.00 (11.38) 75 111	96.6 (11.4) 73 120	97.44 (7.9) 83 113	*U*=361, *Z*=-.05 *p*=.96

5.4.3 Auditory Memory

Auditory memory was tested using the same Word Span test at Initial and Re-assessment. Table 5-12 shows that both groups increased scores between Initial and Reassessment, but, although the persistent group had somewhat lower scores each time, a Mann-Whitney U test showed the difference between the persistent and

resolved groups at each time point was not significant. Nor was there a significant difference between the Age Appropriate and persistent SSD group scores.

Table 5-12: Comparison of Auditory Memory scores at Initial & Re-Ax for the resolved vs. persistent SSD groups

Comparison of Auditory Memory Wordspan test scores at Initial & Re-assessment for the resolved vs persistent SSD groups					
		Age Approp.	**Persistent SSD (PSSD)**	**Resolved SSD (RSSD)**	**Significance test PSSD vs RSSD**
Auditory Memory Wordspan score at Initial Ax:	Mean (SD) Min Max	3.67 (.50) 3 4	3.25 (.75) 2 4	3.43 (.79) 2 5	*U*=257 *Z*=-.73 *p*=.47
Auditory Memory Wordspan score at Re-Ax.	Mean (SD) Min Max	4.33 (.71) 3 5	3.79 (.80) 3 5	3.96 (.71) 2 5	*U*=311, *Z*=-.91 *p*=.36

5.4.4 Summary of differences in abilities between the resolved and persistent SSD group in DDK, vocabulary and auditory memory measured at Re-assessment

Across all three tasks, there was a slight trend for scores of the group who had age appropriate speech at both times to be higher than the two SSD groups, and for the group with resolved SSD to perform better than the persistent SSD group. However, this was significant only in the case of DDK. In all tasks, at both times there were some children in the persistent SSD group who scored better than some in the Age Appropriate and resolved SSD group, i.e. there was an overlap in scores between all groups on all tasks, such that none was a reliable indication of SSD status.

5.5 Phonological Awareness (PA) and Reading

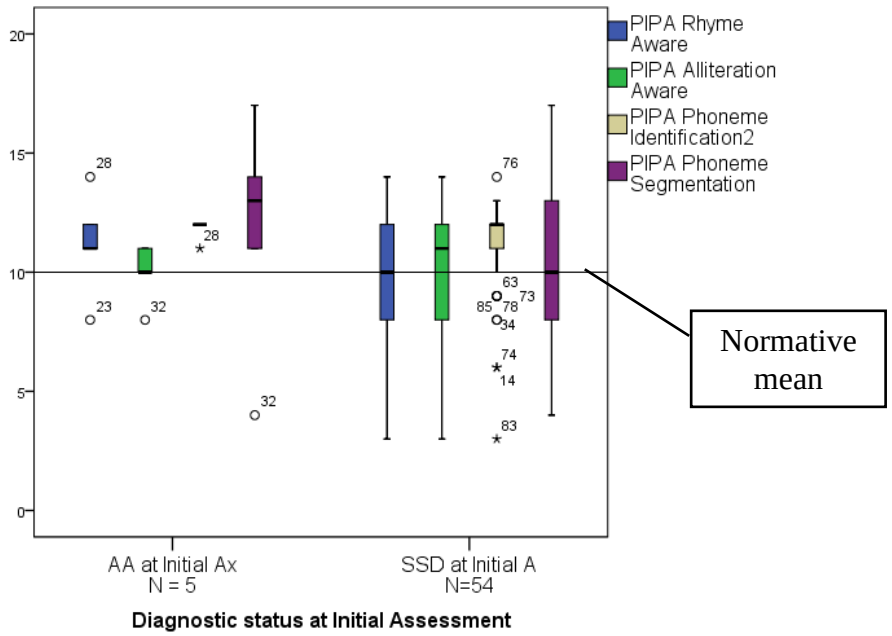
Children's phonological awareness ability at reassessment was examined to investigate the number who presented with poor phonological awareness abilities and/or reading difficulties, and whether there was an association between factors found significant in the literature and PA abilities at Re-assessment in this sample, i.e. error type, co-morbid language impairment, vocabulary, and non-word discrimination. To investigate whether those children who resolved their SSD without therapy may have been left

with underlying phonological deficits, which might otherwise have been addressed in therapy, PA abilities of the therapy group were compared with those who had not received therapy. Finally, the characteristics of those who showed PA abilities below the normal range are described.

5.5.1 Attainment on tests of PA at Re-assessment for the whole group

Data for 14 children were excluded as they fell outside the age range of the test used, and two had not completed the test, leaving 59 children in the sample for which results are reported. Scores reported are scaled scores, i.e. adjusted for age. The box plot in Figure 5-13 shows that, as expected, the children who had age appropriate speech at Initial Assessment scored better on sub-tests of phonological awareness abilities at Re-assessment than those who had an initial diagnosis of SSD. However, the eligible age appropriate group was very small, comprising only five children. It was unexpected to find that both groups scored at or above average for the normative sample on all sub-tests except rhyme awareness, where the SSD group scored just below average (9.77). This unexpected finding may have been related to the fact that local schools had introduced a well-known synthetic phonics programme, which all school-going children were using, and this may have had an effect on their PA abilities. It can be seen that the range of ability in the SSD group was very wide, with some children scoring very poorly, and others very well.

Figure 5-13: Comparison of PA abilities between AA and SSD group, and comparison to norms.

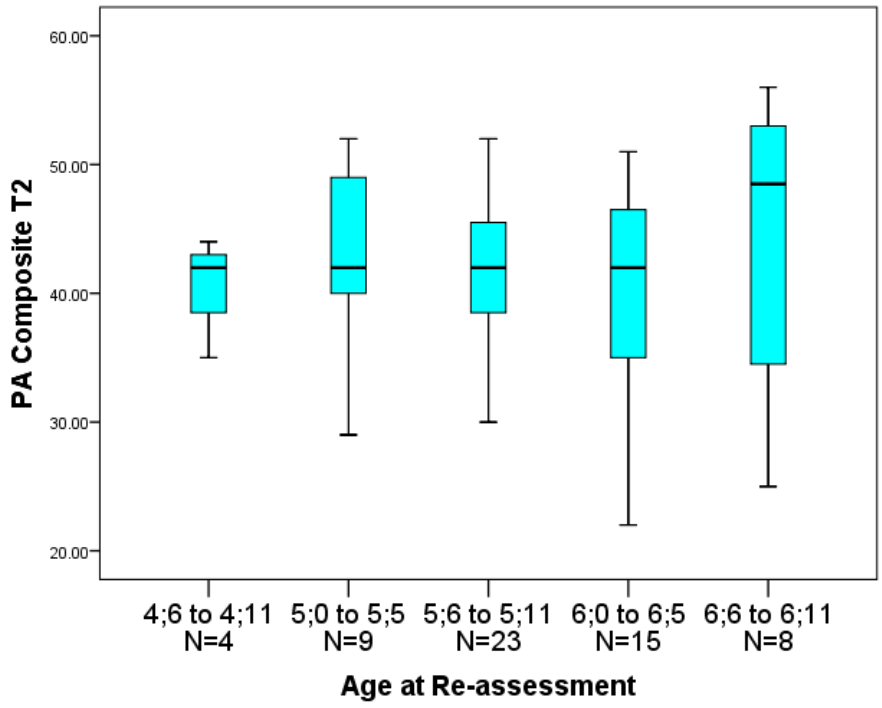


As it has been shown in several studies that PA tasks correlate highly with each other (Preston and Edwards 2010), a composite phonological awareness score was calculated for each child by adding together the scaled scores (adjusted for age) on the 4 sub-tests of the PIPA administered at Reassessment. This was then used to compare performance between different groups on PA abilities overall.

5.5.2 Effect of age

Inspection of a box plot (Figure 5-14) suggested no effect of age on attainment of PA scores in the whole sample. This was confirmed by a Kruskal-Wallis test, $\chi 2$ (4) =2.0 p=.735. As scores were scaled according to norms on the PIPA, this was an expected result.

Figure 5-14: PA Composite Scores by Age at Re-assessment



5.5.3 Attainment on PA by diagnostic sub-group

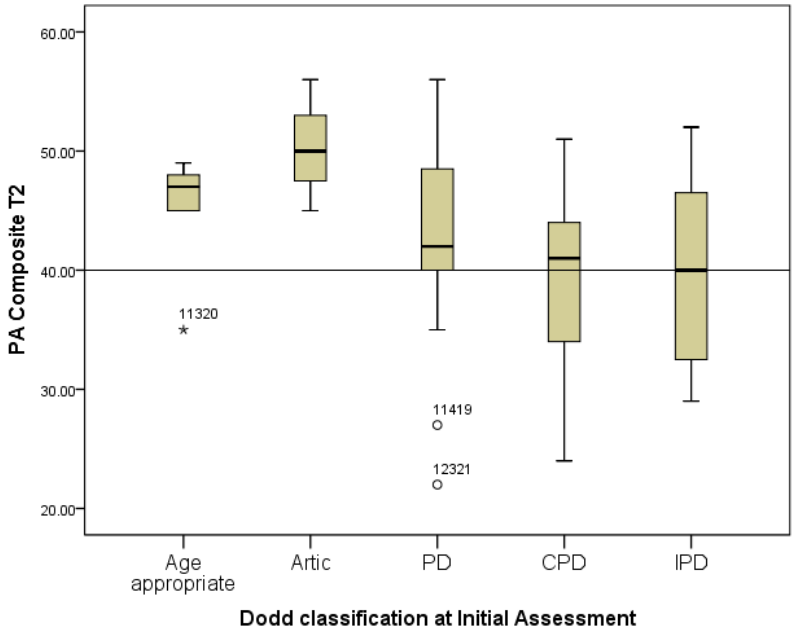
When broken down by diagnostic sub-group (see Figure 5-15), it can be seen that, in accordance with the literature, the articulation group performed as well as those with age appropriate speech, but these were both small groups of 3 and 5 children, respectively. The two disordered sub-groups, while still scoring an average at or above the test norms for age, showed a wider range of scores, and on average scored worse than the PD group, but a Kruskal Wallis test showed this difference was not

significant, $\chi 2$ (2) =2.7 p=.259. In addition, Table 5-13 shows that, in this sample, considerably more children with CPD were not yet at school, compared to the other groups, so it was likely that a greater number had not received any instruction in phonological awareness, which may have influenced this result.

Table 5-13: Number of children at school by diagnostic sub-group

Diagnostic sub-group	Mean Age (months)	Started school		% at school
		No	Yes	
Age appropriate	79	0	5	100%
Articulation disorder	81	0	3	100%
PD	76	3	20	87%
CPD	69	6	15	71%
IPD	66	0	7	100%

Figure 5-15: PA abilities by Dodd's diagnostic sub-groups



5.5.4 Children whose SSD had resolved

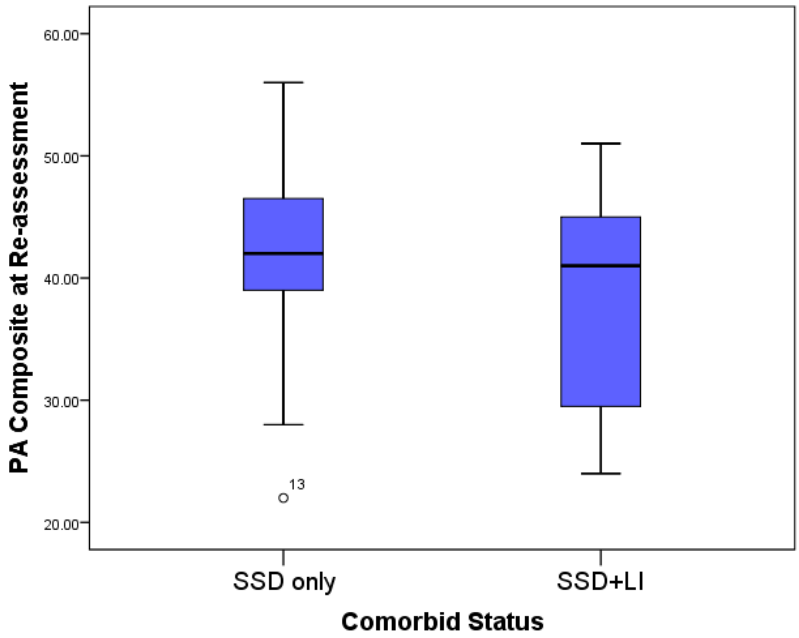
Children whose SSD had resolved had a better mean score (42) on the PA composite at Reassessment than those with persisting SSD (38), but a Mann-Witney U test showed this difference was not significant, *U*=190, *Z*=-1.285, *p*=.199.

5.5.5 Effect of language status

As shown in Figure 5-16 and expected from the literature, children with co-morbid language impairment (N=15, mean = 39) scored worse on the PA composite than those with SSD alone (N=39, mean = 42), but a Mann-Witney U test showed this difference was not significant, U=231, Z=-1.18 p=.237. This result may be due to the fact that most children referred to this study had only mild language impairments, compared to those in other studies described in the literature.

For the whole sample, including children who had had age appropriate speech at Initial Assessment, there was only a weak non-significant correlation between receptive vocabulary, as measured by the BPVS III, at Initial Assessment, and PA at Reassessment, r=.242, n=58, p=.067. There was however a moderate significant correlation, r=.324, n=59, p=.012, between concurrent measures of receptive vocabulary and PA at reassessment.

Figure 5-16: Comparison of PA abilities for children with SSD or SSD+LI



5.5.6 Influence of therapy on PA attainment

Children who had not received therapy had better PA composite scores (mean 44) than children who had received therapy (mean 40), which a Mann-Whitney U test showed to be significant, U = 236, Z = -2.0, p = .042. Of the 20 children with SSD who had received no therapy, one was outside the age range for the test, and the 17 who resolved their SSSD all had age appropriate PA. Only one had needed extra support for

reading. Of the two who had not resolved, one was not yet at school and had not completed the PA tests, and the other had PA in the normal range and no reading difficulty. This difference in scores may reflect the fact that the therapy group had more severe SSD overall, but it also indicates that, in this sample, those who resolved their SSD without help were not left with underlying PA deficits.

5.5.7 Predictors of phonological awareness at Reassessment

A linear regression was carried out to investigate whether phonological awareness at reassessment could be predicted by measures at Initial Assessment, in the whole sample (N=59). In the first stage, measures from variables related to those identified in previous studies to be predictors of PA, i.e. age (Preston and Edwards 2010), non-word discrimination skills and receptive vocabulary (Rvachew 2006), atypical errors (Preston *et al.* 2013), co-morbid language impairment (Sices *et al.* 2007), were tested individually for their association with the outcome variable, the PA composite score. See Appendix 5.B for full results of univariable analysis. Variables with a *p* value of <.10 in the univariable analyses (listed in Table 5-14 below) were then entered into the multivariable analysis. To control for different ages at initial assessment, Age at Initial Assessment was entered in the first block. A stepwise regression was used in order to find the best combination of independent (predictor) variables to predict PA abilities. Probability to enter was set at .05 and .10 to remove.

One outlier was removed. This resulted in a model which explained 30% (Adjusted R Square) to 35% (R Square) of the variance, F (3,39)= 6.91, p= .001, with three variables making a significant unique contribution: the Non-Word Discrimination test made the biggest contribution, *beta* =.48, p=001, explaining 22% of the variance, followed by Age at Initial Assessment explaining 13% of the variance, *beta* =-.37, p=009, indicating that older children had better scores. The third variable was Atypical Errors, so that absence of atypical errors explained 10% of the variance, *beta* =.48, p=001.

Table 5-14: Variables showing a significant association with PA composite score in univariable analysis

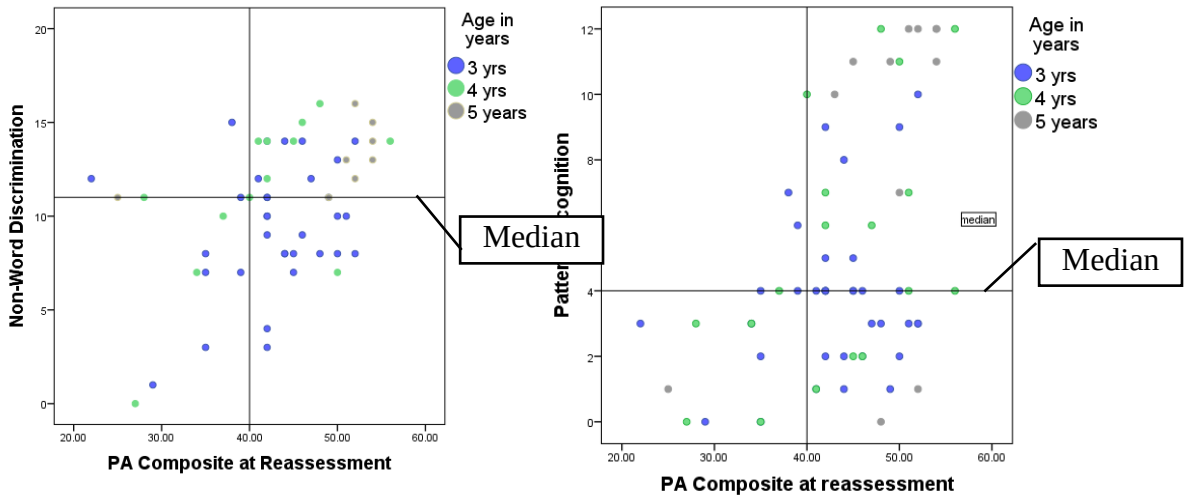
Variables showing a significant association with PA composite score in the univariable analysis			
	Sig.	**95.0% Confidence Interval for B**	
		Lower Bound	**Upper Bound**
BPVS SS	0.067	-0.011	0.317
CELF SS	0.009	0.041	0.282
Atypical errors	0.057	-8.281	0.123
DEAP PCC*	0.032	0.011	0.244
Pattern Recog.*	0.002	0.456	1.852
FIST score 1*	0.004	0.286	1.416
Number concepts*	0.035	0.078	2.118
Non-Word Discrimination *	0.013	0.163	1.333
*adjusted for age at Initial Assessment			

To determine whether more general underlying deficits in rule abstraction and pattern recognition might also contribute to deficits in phonological awareness, variables measuring such abilities were inspected to look for an association with PA abilities. Three variables showed an association with the PA Composite Score at a level of p <.10 in univariable analyses: Pattern Recognition, Number Concepts and FIST score 1. When these variables were added to the linear regression model, the amount of variance explained increased to between 32% (Adjusted R Square) and 37% (R Square), F (3,36)= 7.1, p= .001, again with 3 variables making a significant unique contribution. Age now made the greatest contribution to the variance, *beta* =-.38, p=.008, explaining 14% of the variance, followed by Pattern Recognition, *beta* =.36, p=.017 explaining 11%, and Non-Word Discrimination *beta* =.38, p=.025, explaining 9%.

To determine whether either variable of Non-Word Discrimination or Pattern Recognition was useful as a predictor of PA deficits at the level of the individual child, scatter plots were inspected (see Figure 5-17). This revealed that the Test of Pattern Recognition was a better predictor of good phonological awareness for the individual, with all but two children who scored above the median on Pattern Recognition attaining an average or above average score on the Phonological Awareness Composite according to the normative sample, and those two were quite close. However, like the

Test of Non Word Discrimination as a predictor for resolving SSD, it had poor specificity, in that many children with poor scores on Pattern Recognition at Initial Assessment nevertheless had good scores on the Phonological Awareness Composite at Reassessment. Once scores on tests of more general cognitive abilities were added to the model, production of atypical errors ceased to be a predictor of PA abilities.

Figure 5-17: Scatter plots of Age, Non-word Discrimination & Pattern Recognition scores at Initial Ax by PA scores at Re-assessment



Thus in the final model including tests of cognitive abilities, age at Initial Assessment was the biggest predictor of attainment on a PA Composite score at Re-assessment, with Non-word Discrimination and Pattern Recognition both making similar smaller contributions. At an individual level, scores above the median in Non-word Discrimination and Pattern Recognition both guaranteed PA scores above the normative mean in children who had been over three at Initial Assessment, but specificity was poor, with many children with poor scores on these tests still attaining good PA abilities 2 years later. Surprisingly, vocabulary and general language ability was not a significant predictor of PA abilities, and presence of atypical errors was not significant in the model containing tests of more general cognitive abilities.

5.5.8 Children with PA difficulties

A phonological awareness difficulty was defined as attainment of less than 32 points on the PA composite score; that is less than an average of eight (i.e. minus one SD) in scaled scores across the 4 sub-tests. Overall, only 8/59 (14%) children failed to attain a

composite score of 32: all had had therapy, four had persisting SSD and four had attained age appropriate speech.

Of those with an initial diagnosis of SSD alone, 3/44 (7%) presented with PA difficulties, as opposed to 5/15 (33%) with SSD+LI. Of those whose SSD had resolved by Re-assessment, 4/47 (9%) had phonological awareness difficulties, compared with 4/8 (50%) of those who had a persisting SSD; a Chi-Square test showed this was significant, χ^2 (1, N = 59) = 5.0, p =.025, so consistent with the "critical age" hypothesis (Nathan *et al.* 2004b). Of the seven with PA difficulties who had started reading, 4 were reported to have needed special help; all of these had concomitant language difficulties. Characteristics of children presenting with PA composite scores below the normal range at reassessment are shown in Table 14.

Table 5-15: Characteristics of children with PA deficits at Time 2

Characteristics of children with PA deficits at Time 2							
Gender	PA Standard Score composite	Age time 2	Family History	Comorbid	Diagnosis Initial assessment	Diagnosis reassessment	Reading Difficulty Reported
F	22	73	No	SSD only	PD	AA	No
M	27	78	Yes	SSD+LI	PD	PD	Yes
F	25	81	Yes	SSD+LI	CPD	CPD	Yes
F	31	68	No	SSD only	CPD	PD	No
M	28	72	No	SSD only	CPD	AA	No
M	24	72	No	SSD+LI	CPD	AA	Not started reading
M	29	65	Yes	SSD+LI	IPD	PD	Yes
M	30	67	Yes	SSD+LI	IPD	AA	Yes

The nine children who had not yet started school had a lower mean PA score (39) than those at school (43), but a Mann-Whitney U test showed no significance.

5.6 Reading difficulties

Sixty-five children were attending school at reassessment. Children were categorised as having a reading difficulty (RD) if parents reported they had received special help for reading at school. Fifteen children (23%) were reported to have RD, a figure similar to that reported by Peterson *et al.* (2009) who reported RD in 22% of 7-9 year olds with a history of SSD, but much lower than that in older studies (Bird *et al.* 1995, Larrivee and Catts 1999). This may be accounted for by the fact that RD was designated by parent report of a child receiving extra help for reading at school, rather

than by actual assessment of reading ability, as carried out in the studies named, or by the fact that many children in this sample had milder speech and/or language difficulties than those in many research studies.

Only 4 of the children with RD in this sample had scored below average on the PIPA, reflecting other studies that have shown that measures of phonological awareness are not always related to literacy outcome (Bird *et al.* 1995). Reading status was not reported for one child.

A much larger proportion of children with persisting SSD presented with reading difficulties (5/10 50%) than those who had age appropriate speech (10/54, 18.5%), but three of the five also had co-morbid language impairment, which is shown to be a larger risk for RD than persistence of SSD. Consistent with the literature, a larger proportion of those with co-morbid language impairment (5/14, 36%) presented with reading difficulties than those with SSD alone (10/50, 20%) (Leitão and Fletcher 2004, Nathan *et al.* 2004a).

None of those diagnosed with articulation disorder, but three of the nine children who presented as age appropriate at initial assessment presented with reading difficulties. The distribution of those with reading difficulties among the sub-groups of phonological impairment was 5/29 (17%) of those with PD, 3/15 (20%) of those with CPD and 4/7 (57%) of those with IPD. Table 5-16 shows the characteristics of children reported to have reading difficulties at reassessment.

Table 5-16: Characteristics of Children with Reading Difficulties at Reassessment

Characteristics of Children with Reading Difficulties at Reassessment							
Gender	**Age at reassessment**	**Family History**	**SSD Diagnosis at Initial Ax**	**SSD Diagnosis at reassessment**	**Comorbid**	**PA composite**	**Cluster membership**
M	75	No	CPD	AA	SSD only	39	--
M	67	Yes	IPD	AA	SSD+LI	30	--
M	71	No	CPD	AA	SSD+LI	41	--
F	58	No	IPD	PD	SSD only	35	--
F	90	Yes	AA	AA	AA	45	--
M	63	Yes	IPD	PD	SSD only	52	NEARLYNORMAL
M	92	Yes	AA	AA	AA	51	NEARLYNORMAL
M	91	Yes	PD	AA	SSD only	52	NEARLYNORMAL
M	93	Yes	AA	AA	AA	41	NEARLYNORMAL
F	96	Yes	PD	AA	SSD only	54	NEARLYNORMAL
F	96	Yes	PD	AA	SSD only	54	NEARLYNORMAL
M	70	No	PD	AA	SSD only	42	NEARLYNORMAL
F	81	Yes	CPD	CPD	SSD+LI	25	POORLANG
M	78	Yes	PD	PD	SSD+LI	27	BADOVERALL
M	65	Yes	IPD	PD	SSD+LI	29	BADOVERALL

5.7 Summary of findings regarding outcomes of children with SSD at Reassessment.

Seventy-five (88%) of the 85 children presented for Reassessment approximately 2 years later. All of the nine children who had presented with age appropriate speech at Initial Assessment also presented so at Reassessment. Of the 66 children who had presented with SSD at Initial Assessment, 52 (79%) presented with age appropriate speech at Re-assessment, where a dentalised production of alveolar fricatives was counted as age appropriate. Children who had resolved their SSD were significantly older than those with persisting SSD at Initial and Reassessment. All children in the SSD group but one, with a very mild delay, had increased their PCC. Both boys and girls made the same amount of progress in terms of PCC, but younger children improved more than older children did, possibly reflecting a ceiling effect as younger children had lower PCC initially. Children with IPD and CPD made greater gains over the two years than children with PD, so that they appeared to be catching up with both the PD group and the age appropriate group.

The distribution of children between diagnostic sub-groups of SSD changed between Initial and Re-assessment, but in a consistent manner, i.e. from disorder to delay and

delay to age appropriate speech. Those children classified at Initial Assessment as having age appropriate speech also had age appropriate speech at Re-assessment, but all the children in the phonologically disordered groups bar one changed diagnostic category, either to PD, articulation disorder or age appropriate speech. No child changed from a delay to a disorder, and only one child remained in the disordered category. No child changed from IPD to CPD, but given the long interval between Initial and Re-assessment, this is not evidence that this did not happen. All children classified as having an articulation disorder at Reassessment had lateralised fricatives, and no child with a lateralised production of fricatives resolved this.

A bigger proportion of those who did not receive therapy resolved than those who had received it, but the latter group were more severely impaired initially, and the amount of therapy provided was, in at least some cases, too small for there to be any expectation of an effect, so this study in no way contributes to evidence regarding efficacy of therapy. However, what this study did show was that some children with severe deficits did resolve their SSD even without therapy, and both children with phonological disorder and delay resolved without therapy. Over the two-year period, children who received no therapy made gains greater than those expected with maturation alone, so that they were catching up with their TD peers. Nevertheless, receipt of therapy was associated with a significantly bigger mean gain in PCC across the sample as a whole, but again, as the children in the therapy group were more severe to start with, they had a greater scope for improvement.

This study also sought to find factors predictive of resolving SSD within 2 years, using logistic regression. Neither gender nor family history were predictors, indicating that while these factors may increase the risk of having an SSD in preschool, in this sample they did not influence whether a child resolved this SSD within 2 years. Measures of oral-motor or gross motor skills, number of ear infections, history of birth difficulties, late talking and level of maternal education were similarly non-predictive. Unlike previous studies, presence of a co-morbid language impairment did not increase the risk of persistent SSD, but it must be taken into account that the language impairments in this sample were relatively mild compared to other studies. Measures of PCC at Initial Assessment, whether unadjusted or adjusted for age, were not associated with resolved SSD, with 70% of those classified as severe at initial assessment resolving their SSD by Re-assessment.

This study also aimed to discover whether more general cognitive deficits might underlie at least some types of phonological impairments. Thus it included measures not directly related to speech and language abilities, i.e. measures of number concepts, pattern recognition and cognitive flexibility. One measure of pattern recognition, the FIST 1, did show a univariate association with resolution, but it was not a good predictor at the individual level. The aim of this study is specifically to find clinically useful measures that can help in predicting whether an individual child will resolve their SSD. In the multivariable model, only one variable made a unique, statistically significant contribution, the Test of Non-Word Discrimination. This was a sensitive indicator, in that a score above the median predicted certain resolution of phonological difficulties in this sample, but it showed poor specificity, with many children who had poor scores on this test at Initial Assessment nevertheless resolving their SSD by reassessment.

A surprising finding was that the children in this sample, all referred for concerns regarding SSD, scored above the mean of the normative sample on most of the standardised PA tests; however it should be noted that the normative sample was a UK rather than an Irish sample. As predicted by the literature, children with CPD had worse PA abilities than children with PD, but this result may have been confounded by the fact that a smaller proportion of them had begun school. Eight children failed to attain a score on the phonological awareness composite that was within 1 SD of the mean of the normative sample, with a bigger proportion of those with co-morbid language impairment having PA difficulties than those with SSD alone, reflecting findings in the literature.

Multiple regression was used to find factors predictive of good PA abilities. When measures of Non-word discrimination, receptive vocabulary, presence of atypical errors, co-morbid language impairment and persisting SSD were entered into the model, age, Non-word discrimination and presence of atypical errors were significant predictors of poor PA. However, when measures of more general cognitive ability were added to the model, scores on a test of pattern recognition were a strong predictor of PA abilities and presence of atypical errors was no longer predictive.

Fifteen children (23%) had reported reading difficulties, consistent with other studies. Only 4 of these presented with impaired PA although all four had SSD+LI. Three

children with poor PA were reported to be progressing normally in reading; all of these had SSD only.

This chapter has reported results showing that the majority of children in a clinical sample of pre-schoolers with SSD resolved this by early school age. This study seeks to find indicators which would be useful at the level of the individual child to predict whether they will resolve their SSD. Only one factor was found to have good sensitivity and specificity in predicting SSD that would not resolve: this was presence of lateralised fricatives. Non-word discrimination ability also predicted resolution, with good sensitivity but poor specificity.

5.8 Discussion

This longitudinal study was motivated by the need for information about the outcomes of a cohort of children with SSD that is representative of the caseloads encountered in everyday clinical practice, so that, in a context of chronically overwhelmed SLT services, clinicians can decide where to target services to have the best effect. Botting *et al.* (2016) report discontent in the SLT profession with research findings from studies using highly selective participant groups that do not reflect actual SLT caseloads, or the resources available to clinicians. Much research into SSD focuses on specially selected criteria, such as a specified severity level, or exclusion of children with co-morbid language difficulties. However, as stated by Broomfield and Dodd (2005b), to reflect the 'real world of working clinicians' (p83), there is a need for studies that focus on groups of children like those who are referred to every SLT clinic every day, using assessment measures that are commonly available, in order to provide information that can be useful in the management of such caseloads.

The first question to be answered therefore, was how many of the children in such a caseload might resolve their SSD within Shriberg's age of short-term normalisation. The second purpose was to investigate whether predictors of outcome in terms of speech accuracy and phonological awareness abilities, could be identified. The third question sought to identify whether children's error profiles as categorised by Dodd's Differential Diagnosis Classification System remained stable, reflecting different underlying psycholinguistic processing deficits

5.8.1 Number of children who resolved their SSD

Of the 66 children who were diagnosed with SSD at Initial Assessment and who were re-assessed two years later, 52 (79%) presented with age appropriate speech at reassessment, when their mean age was 6;2. Only one child over the age of six had a persisting SSD, and this child had an articulation disorder. This figure is higher than many other longitudinal studies, but is in line with Shriberg's estimate that 75% of pre-school children with SSD normalise their speech by 6 years of age (Shriberg 1994c). This apparently high proportion of children who resolved their SSD was likely due to methodological factors. Firstly, it is a sample of children referred for SLT services, rather than a sample selected specifically for research purposes, so there was no set cut-off point for PCC, as used in many studies (e.g. Lewis *et al.* 2000), which often limit participants to those in the moderate to severe range of impairment. In this sample, 42% had only a mild impairment, including some with PCC in the normal range for age, but who still presented with immature speech in terms of delayed or disordered speech error patterns. There were no children with CAS, and only a small number with articulation disorder, which conditions are known to be persisting. Studies that report lower resolution rates have participants with more severe impairment: for example, Webster *et al.* (1997) report a resolution rate of 66% by age 6 for children with moderate to severe impairments.

Secondly, the assessment interval was relatively long. Studies with lower resolution rates have followed children over shorter intervals. Broomfield and Dodd (2011), in a study of 312 children referred to SLT services for SSD found that only 15% resolved, but the interval was only 6 months. A resolution rate of 66% was found in a clinical sample of children of the same age as those in this study by Roulstone *et al.* (2003), but they were followed up after only 12 months. When 83 children diagnosed with a speech and/or language impairment aged under 3;6 were followed up 4 to 7 years later by Roulstone *et al.* (2008), 78% no longer qualified as cases. Other studies with long intervals between initial and follow-up assessment have found rates of resolution similar to this study: Lewis *et al.* (2000) found a resolution rate of 86.5% in 4-6 year olds with SSD when followed up at age eight to nine; Peterson *et al.* (2009), found that 79 (88%) of 86 children with SSD before the age of 5 had resolved their difficulties when followed up at age 7 to 9; Hesketh (2004) found 28 (80%) of 35 children diagnosed with moderate to severe SSD at ages between 3;6 and 5;0 had resolved by

age 6;6 to 7;6. Bishop and Edmundson (1987) also found that 78% of children with pure phonological impairment at age 4 resolved by 5;6, although their sample was only 9 children. A high rate of resolution is also supported by an older cohort study of speech development in 944 children born in Newcastle-upon-Tyne, England, which found that at 4 years of age, 17% had "defective articulation" not attributable to organic disease. On follow-up at 4;8 years, this figure had dropped to 14%, and by age 6 years to 3%, indicating a resolution rate of over 80% (Morley 1972).

A third factor is the age of the sample at initial assessment. Children are usually referred for assessment of speech difficulty at around age 3 (Broomfield and Dodd 2005b), so it is at that age when clinicians need to make decisions about provision of therapy. The majority (58%) of this sample were under the age of 4 years at initial assessment and 22% were under 3;06. Some studies reporting lower resolution rates include only participants aged 4 or over, when speech sound systems are known to be relatively more stable (Dodd *et al.* 2003). For example, Morgan *et al.* (2017) report a resolution rate of just under 60% for 4 year olds reassessed at age 7. In Broomfield and Dodd's (2011) study cited above, 15% of the sample were over six at referral, meaning they were already classified as having persistent SSD.

A final factor is the definition of resolution used. Law *et al.* (2000) suggested that in many studies it would appear that the children's' speech is reported in absolute rather than developmental terms. This study reported as resolved children whose speech became age appropriate rather than error free. Thus at 6 years many still had dentalised /s/ and /z/ productions, or had not yet acquired an adult like /ɹ/, but these are not considered to be markers of SSD until age 8 or over. Dodd *et al.* (2017) also reports that 9 children in the ELVS sample, who had shown no distortions of /s/ or /z/ at age 4, had an interdental lisp at age 7, although, like the children in this sample, all produced these consonants correctly in some contexts. Conversely, nine of 14 children in her sample who had interdental lisps at age 4 presented with no speech errors at age 7. In this study, children whose only error was a dentalised production of alveolar fricatives, were classified as having age appropriate speech, but this is not the case in all studies, (e.g. Morgan *et al.* 2017). However, Bowen (2015) also suggests that interdental lisps might be considered a developmental error since they are so common. It is possible that some of the children still making age appropriate errors at age 6 will fall back into the category of SSD later, if those productions of /s/ and /ɹ/ do not continue to develop.

However, the number counted as resolved here reflects the number who would be discharged from a clinical caseload as having no current problems.

The results of this study regarding resolution rate therefore show that a large majority of children will resolve their SSD before age 7. A limitation of this study is that there were only two measurement points, so it is not possible to know how soon within the 2 years before Re-assessment it was that children made significant improvement or caught up with their peers. To attempt to account for this, a trajectory curve (see Figure 5-2) was posited from cross sectional observations of children in the sample at different ages. This suggested the most rapid increase in speech accuracy would take place in the youngest age groups, with improvement levelling off by age 4;6. This converges with the results of Roulstone *et al.* (2003), who found that children aged between 2 and 3;6 showed more improvement in phonology in the first six months of their study, than in the second six months.

5.8.2 Influence of therapy

It was fortuitous that this study allowed follow-up of a group of 20 children who did not receive therapy, which enabled spontaneous resolution in the longer term to be seen. This was due to the political and economic context, where austerity led to vacant posts in health services not being filled, and a consequent undersupply of services. Thus children were not randomly assigned to a waitlist group, but received no therapy either because there was no therapist in post in their local service, or because they did not reach a high enough priority on the waiting list.

In contrast to the findings of Broomfield and Dodd (2011), children who received no therapy still made gains greater than those expected with maturation alone, so that they were catching up with their typically developing peers, as can be seen in Table 5-4. However, the period between Initial and Reassessment was 2 years as compared to 6 months in Broomfield and Dodd's study, so it is possible that children with SSD needed longer to catch up, and children with age appropriate speech hit ceiling much more quickly.

The finding that 18 out of 20 had achieved age appropriate speech at follow up 2 years later, without therapy, is greater than the degree of spontaneous resolution (i.e. without therapy) found in the studies noted in the introduction (e.g. Roulstone *et al.* 2003,

Williams and Elbert 2003), but again may be accounted for by the fact that most children had only a mild impairment, and the follow up interval was twice as long at 2 years than the one year reported in most of those studies. Nevertheless, disordered as well as delayed children resolved without therapy, and four of the children resolved without therapy despite having SSD in the moderate to severe range. It was evident that local SLTs were prioritising children with disorder, and children who were more severe, for therapy. For instance, all of the children with IPD received therapy, so the study shed no light on this group's progress without therapy.

In fact, many of the children assigned to the "therapy received" group did not receive an amount or type of therapy that could be expected to be clinically effective. For example 12 children received only a home programme, for which there is little proven efficacy for SSD (Sugden *et al.* 2016), while of those who received clinician directed therapy, some received as little as one session. Thus this study cannot contribute to any evaluation of therapy. Despite this, the majority of children resolved, so it may be that some of these would also have resolved without any therapy.

Other longitudinal studies of outcomes have also reported lack of evident therapy effect for children with SSD. In a study of 93 4 year-olds with SSD, Morgan *et al.* (2017) found that only five of the 11 (45%) who had received therapy went on to resolve their SSD by age 7, compared to 58% of the sample as a whole. Shriberg *et al.* (1994b) found that 78% of the children with persisting SSD had received some type of direct intervention by follow up, compared to only 60% of the children who normalized; however, this difference was not statistically significant. Paul (1996), in a small study of 31 children with expressive speech and language delay, found 39% of those who had not resolved had received intervention, compared to only 29% of those who resolved by school age. Bishop and Edmundson (1987) found no effect of therapy on outcome in their study of 87 language impaired 4 year olds after one year, but as they note, therapy was not randomly assigned: children with more persistent deficits tended to obtain more therapy, as in this study, where the child who received most therapy did not resolve, although her diagnosis changed from consistent phonological disorder to articulation disorder. Thus amount of therapy received is unlikely to be a good predictor of outcome, as those children who have the most severe and intransigent deficits are both more likely to receive therapy and less likely to resolve, and therefore likely to receive more therapy without resolving. As Shriberg *et al.*

(1994b) observed, in studies such as this where there is a wide range of severity of disorder and therapy is not randomly allocated, any effect of therapy on outcome is confounded by differences due to initial status.

5.8.3 Predictors of resolving SSD.

Evidence that most pre-schoolers will resolve their SSD, is, however, not enough for clinicians who need to make decisions regarding therapy for the individual child. Another aim of this study was to examine a range of factors to look for predictors of outcome at the level of the individual, as well as at the group level. However, the findings of this study, because such a large proportion of children resolved their SSD, need to be interpreted with caution. Because the number of children who did not resolve their SSD by Re-assessment is small, conclusions regarding predictors can only be tentative.

5.8.3.1 Demographic and case history factors

Case history factors such as gender, family history, SES, birth history and history of ear infections were not a predictor of normalisation of SSD, so it may be that while they predict membership of the group of children with SSD, they do not predict persistence of SSD. In this study there were more boys in the sample with SSD at Initial Assessment (60%) than girls, which is consistent with a large proportion of studies of children with SSD, e.g. (Fox *et al.* 2002, Campbell *et al.* 2003, Broomfield and Dodd 2004b). At Re-assessment they formed a similar proportion in the group of those whose SSD did not resolve (57%), suggesting that boys and girls resolve at the same rate. If this is the case, the proportion of boys in the population with SSD will usually be greater than the proportion of girls at all ages, as found in Wren *et al.'s* (2016) study, where 64% of the 8 year olds with persisting SSD were boys. However the likelihood of having persistent as opposed to transient difficulties was not higher for boys than girls in this sample, in accordance with studies by Shriberg *et al.* (1994b), Bernhardt and Major (2005), Young *et al.* (2002), Lewis *et al.* (2015) and Morgan *et al.* (2017).

In the same way, 45% of children in the whole sample of children with SSD had a positive family history of speech, language and literacy disorders, compared to 50% of those with persisting SSD, a non-significant difference, suggesting family history does

not affect normalisation. This is consistent with findings from Morgan *et al.* (2017), Roulstone *et al.* (2003) and Shriberg *et al.* (1994b), who similarly found no effect of family history on outcome. Broomfield and Dodd (2005a) report poorer outcomes for children with a positive family history over 6 months, but no figures are given.

Like other longitudinal studies of clinical samples, (e.g. Shriberg et al., 1994b; Roulstone *et al*., 2003), this study found no effect of level of maternal education on persistence of SSD. Studies using other measures of SES in clinical samples, as reported in section 1.5.6.3 (e.g. Lewis *et al.*, 2015), have similarly found no association. It is in population samples where measures of SES have been found to be associated with prevalence of SSD at both younger and older ages, (e.g. Campbell *et al.* 2003, Eadie *et al.* 2015, Wren *et al.* 2016), so this difference might reflect that population samples capture a section of the population who do not access clinical services, or who do not take part in clinical research projects. Alternatively, it might be that this factor only becomes significant in large sample sizes. However, one longitudinal population study, Morgan *et al.* (2017), also found no effect of SES on persistence of SSD. This evidence suggests that SES is not a good predictor of transient versus persistent SSD in children who present to clinical services.

The distribution of diagnostic sub-groups by level of maternal education shown in this sample is in contrast to Broomfield and Dodd's (2004) study, which found that CPD was most strongly associated with mild deprivation, and IPD with severe deprivation, while PD and articulation disorder were more strongly associated with children from affluent backgrounds. This might also indicate that highest level of maternal education was not a good measure of disadvantage for this particular sample, a good number of whom came from rural backgrounds, rather than the urban areas in Broomfield and Dodd's study. Alternatively, it may be that maternal education and other measures of socio-economic status are risk factors for a diagnosis of SSD, but not predictive of its course, once diagnosed. However, evidence against this argument is provided by Wren (2015), who, in the ALSPAC large cohort study, found that children with persistent SSD at age 8 were more likely to be from families who did not own their own homes. There may also be an interaction with maternal age, which was not measured in this study. Roulstone et al. (2009) found an association between younger mothers and likelihood of persisting SSD. Older mothers in Ireland are less likely to have higher

educational qualifications (Dept. of Education and Science 2007), so if older maternal age is a protective factor it may confound educational qualifications.

Parent report of delayed motor abilities was not predictive of persistence of SSD in this sample. This result was consistent with findings from Wren *et al.* (2016), who found that measures of gross and fine motor skills at 42 months did not predict SSD status at age 8. This may be mediated by age: Eadie *et al.* (2015) found delayed motor abilities at 2 predicted SSD at 4, but given the age of their sample, it may again be that this is a risk factor for SSD in pre-schoolers, but not for persistence of SSD past the age of normalisation. Alternatively, it may be that the association only becomes significant in large samples, or that differences in motor development only become clearly apparent at later ages, and in samples that include children with known motor deficits like CAS: Visscher *et al.*'s (2007) study showing impaired motor skills in children with SSD looked at children aged 6-9. It did not exclude children with CAS, which is acknowledged to have a motor component and to persist into later ages (Murray *et al.* 2015), and only included children at special schools, indicating more severe impairment.

The fact that number of ear infections or insertion of grommets did not predict speech status at Re-assessment is consistent with other studies, including a meta-analysis (Roberts *et al.* 2004) and a large scale longitudinal study (Paradise *et al.* 2007) that show little to no effect of otitis media on development of speech and language. As seen here, almost half the sample did not have any ear infections, while it is also known that many children who have ear infections do not have resulting speech and language difficulties. Thus, while double dissociation of otitis media and SSD does not prove that it is not an aetiological factor in SSD for some children, it is not useful as a predictor for persistence of SSD.

Report of birth difficulties, shown to be a risk factor for SSD in Fox *et al.* (2002), was not a predictor of persistent SSD in this sample, consistent with results from Eadie *et al's.* (2015) large population study of younger children. Conversely Eadie *et al.* found report of delayed onset of speech at 2 to be a predictor of SSD at 4, whereas this study did not find report at 3-5 to be a predictor of SSD at 5-7, so it may be that it is only a predictor of delay at young ages, given that delayed language is often transient (Hayiou-Thomas *et al.* 2014). Similar to the larger study by Broomfield and Dodd

(2004b), a history of delayed onset of speech and language was less frequent in children with SSD than in children with language impairment, possibly indicating that children with SSD alone can catch up quickly from initial delay.

5.8.3.2 Severity of speech impairment

As seen in other studies where children were under five years of age at initial assessment, number of errors was not a good predictor of ongoing speech errors (Bishop and Edmundson 1987, Shriberg *et al.* 1994b, Bernhardt and Major 2005, Rvachew 2006, Roulstone *et al.* 2009, Morgan *et al.* 2017). While severity of impairment adjusted for age appeared in the final model of predictors, it did not make a unique contribution to likelihood of resolving the SSD. Although most children with mild impairments resolved, so did many children with the most severe deficits. Inspection of the scatter plot (see Figure 5-12) showed it was not a good predictor for the individual child, as some with milder deficits did not overcome these by Re-assessment, while other children with severe impairments made very large improvements and caught up with TD peers.

5.8.3.3 Error type

One particular error type was shown to remain stable: the four children who presented with lateral fricatives, either as a primary or secondary diagnosis, did not resolve this, even with therapy. This accords with evidence cited in Shriberg et al. (1994b) that lateralisation does not correct spontaneously and with Preston *et al.*'s. (2013) findings that a history of distortions at age 4 was a better predictor of speech outcome at age 8 than a history of substitutions or omissions. Dodd *et al.* (2017) also reports that the one child in the ELVS sample with a lateral lisp at age 4 had not improved by age 7. Conversely, the 3 children who presented at initial assessment as non-stimulable for other sounds, two for post-alveolar fricatives/affricates and one for /w/, all resolved, the latter without therapy and the former with. Shriberg *et al.* (1994b) found that stimulability measures produce high rates of false positives and negatives at the level of predicting individual improvement, and recommended that clinicians should hold back from diagnosing articulation difficulties based on lack of stimulability until well past the normal age of acquisition of a sound, except where the sound is lateralised.

Morgan *et al.* (2017), in their longitudinal study of 4 year-olds followed up at age 7, found that only type of error (delay vs disorder) was a statistically significant predictor of persisting versus resolved SSD, with children who made disordered errors (i.e. in diagnostic sub-groups CPD and IPD) twice as likely to show persisting errors at age 7 as those with delayed errors. Again the delayed group had less severe impairment at age 4, but a comparison of groups by age is not reported (i.e. whether the children with delay were older than those with disorder). In contrast to that finding, in the study reported here, neither diagnostic status nor production of atypical errors was a predictor of persisting SSD, but again there were methodological differences. Presence of disordered errors was defined using Dodd's classification in both studies, but in this study the children who presented with disordered errors were much younger: 23 out of 30 (76%) were below the age of 4, and 28/30 (93%) were below the age of 4;6. At Re-assessment, only 1 child, who was aged 5 at initial assessment, still presented with disordered errors. It may be then that disordered errors are a sign of a more poorly developed phonological system, but children who present with them at age 3 have a lower risk of persistent SSD than children who are still producing them after the age of 4. Unlike Broomfield and Dodd's (2005) findings over a shorter 6 month follow up, children classified as having CPD or IPD in this study made more progress in terms of PCC between Initial and Re-assessment than PD children, and appeared to be catching up with TD peers. In fact, a smaller proportion of the children with atypical errors (74%) or inconsistent errors (43%) resolved than those with PD (91%), but the children with disordered errors were also more severe at initial assessment, and significantly younger (5:8) than the PD group (6;4) at Re-assessment, allowing that they might still resolve by the age of 6. This concurs with Preston *et al.* (2013)'s findings, that a history of atypical errors did not predict persisting SSD at age 8. Given Claessen *et al.*'*s* (2016) findings in a sample of under threes, that younger children made more atypical errors than older children, and both groups continued to decrease use of atypical errors as they approached the age of 4, it may be that atypical errors are only a sign of more pervasive SSD if they continue to be produced after that age.

Dodd (2011) explains atypical errors as a surface symptom of underlying cognitive deficits, but it is possible that not all atypical error patterns reflect difficulties with rule abstraction and cognitive flexibility as posited by Dodd. The reason why not all children who show atypical errors have the same deficits or outcomes may be due to difficulty in identifying whether use of an atypical phonological process actually

reflects deficient knowledge of the phonological system, or when it is simply a strategy for dealing with a current constraint (Gibbon and Lee 2016). While some processes, such as omission of all initial consonants, or use of /h/ to substitute for all consonants, could point to mislearning about the rules of the English phonological system, other processes seem more likely to be devised as strategies to deal with current maturational constraints.

For example, cluster reduction is a process which in non-linear phonology can be viewed as a response to a constraint at the syllable level on what may be contained in an onset. However, some TD children who use cluster reduction, while not producing the /s/ in clusters which contain /s/ plus another voiceless plosive (e.g. st, sk,), still signal knowledge of its presence by producing an unaspirated voiceless consonant initially (voiceless plosives are usually aspirated in word initial position before vowels in English, but not when forming a cluster with /s/). Another way to signal knowledge of a segment missing in production is to coalesce features of both into the one segment allowed. For example, the use of a bilabial fricative to substitute for clusters such as /sw/ and /sp/ (classified as an atypical error by Dodd) shows the coalescence of the fricative feature of /s/ with the labial feature of /w/ and /p/, allowing the child to mark that there are two segments in the cluster while being constrained to produce only one. Both the typical and atypical error production signal that children in fact have accurate knowledge of the structure of the target word that they cannot yet produce. Similarly, So and Dodd (1995), in a study of the acquisition of phonology by Cantonese speaking children, found that children reduced the clusters /kw/ and /k^{h}w/ by deleting /w/. However, although singleton /k/ was usually fronted to /t/, they signalled knowledge of the /w/ by substituting /kw/ with /p/. Older children still using cluster reduction often marked the contrast in aspiration by substituting /k^{h}w/ with /p(^{h})/, whereas /kw/ was substituted with /f/.

From another perspective, Bernhardt and Stemberger have proposed that what appears to be atypical development can result from "randomness in the resolution of multiple constraints" (Bernhardt and Stemberger 2007, p584). Bernhardt and Stemberger have applied Optimality Theory to the study of phonological acquisition and SSD (Bernhardt and Stemberger 1998). This theory posits a set of universal constraints that act on underlying representations, the input, to produce surface realisations, the output. It relies on the theory of markedness referred to above, with the observed language

resulting from the optimal satisfaction of conflicting constraints of markedness and faithfulness. Markedness constraints favour output that contains only unmarked properties of the input, whereas faithfulness constraints favour output that is most similar to the input, i.e. the adult target (Rvachew and Brosseau-Lapré 2018). Phonological acquisition is thus seen as the process of adjusting the relative ranking of these constraints until the output matches the input. In a discussion of the application of Optimality Theory to children with phonological impairment, they note that the data from children with atypical phonological development still fit the observation that in phonological development, unmarked structures tend to emerge before marked structures. They argue that where children need to overcome multiple markedness constraints, the data may appear to be atypical when children have followed a less common option to deal with conflicting constraints. In fact the same constraints are relevant and the developmental process is the same: i.e. faithfulness constraints gradually overcome markedness constraints, but the data may appear atypical when the child chooses a less common option where multiple valid options are available (Bernhardt and Stemberger 2007). Barlow and Geirut (1999) show an example of how different patterns of cluster reduction for the same word can be produced by different rankings of the same set of constraints. Thus it is not that the children with atypical development do not follow the same constraints as typically developing children; they merely use a less common strategy in resolving conflicts. Optimality theory thus preserves continuity in the grammar, and captures common error patterns while still allowing for individual variation in production by way of alternative ranking and re-ranking of constraints (Barlow and Geirut 1999). The theory is vulnerable to the same cross-linguistic evidence showing the lack of a universal hierarchy of markedness referred to above, as well as its reliance on the child's perception of the input as being perfect, so that mismatches between input and output can be detected. It also shares with natural phonology the negative conception of phonological development as a process of demotion of prohibitive constraints rather than the acquisition of new skills or knowledge (Rvachew and Brosseau-Lapré 2018). Nevertheless, it also allows for the conception of atypical speech production as a creative response to constraints, rather than a lack of knowledge or understanding of the ambient phonological system.

Atypical errors may thus result from sources other than mislearning of the rules of the phonological system. For example, one of the most frequently identified atypical processes in this sample was backing of alveolar plosives to velars. However several

studies (Edwards *et al.* 1997, Gibbon 1999, Goozée *et al.* 2007) have shown that some children who appear to be using the atypical phonological process of backing, or the developmental process of fronting, do in fact produce covert contrasts between alveolar and velar plosives, illustrating that they do have phonological knowledge of the contrast, but this can only be seen when instrumental measurements are used. The difficulty they have is in moving parts of the tongue independently of each other, so what is heard as either fronting or backing in such a case is not a phonological process nor a sign of lack of knowledge of the phonological system. Both Tyler *et al.* (1993) and Forrest *et al.* (1990) found that children who made such a covert contrast for /k/ versus /t/ initially, made progress in therapy more quickly than children who made no contrast, indicating that possession of phonological knowledge in a child who cannot yet make an articulatory contrast may predict faster resolution.

One example of a child in the current study who fell into the CPD sub-type who may have been making atypical errors as a response to other constraints was 12505, who also had a secondary diagnosis of articulation disorder. She used atypical processes of backing alveolar stops to velars, and substituting /l/ with /n/. She also lateralised alveolar fricatives, thus showing difficulty with producing any oral consonants at the alveolar ridge. A possible hypothesis is that this child had difficulty with differential movement of parts of the tongue, so that what was heard as backing was in fact release of the plosive at the alveolar ridge before the dorsal release, thus resulting in perception of /k,g/ for /t,d/. However /n/ does not require oral release so can be produced with an undifferentiated tongue gesture. Thus producing /l/ as /n/ allowed this child to preserve the sonorant feature of /l/ as well as the alveolar place of articulation. This child had high scores on both parts of the FIST (15 and 12 respectively), but scored more than 1SD below the mean for the sample in DDK rate, consistent with the hypothesis that motor rather than phonological constraints were the basis of her SSD. By Re-assessment she had resolved all errors except lateralisation of fricatives. In fact, 3 of the 4 children who presented with a diagnosis of articulation disorder (lateral lisp) at Reassessment, originally presented with CPD at Initial Assessment, possibly indicating difficulties with tongue movement in addition to phonological difficulties. Thus it may be that some children will produce atypical errors due to inability to abstract the correct rules of the phonological system from the input, whereas others produce them due to motor constraints, such that the two groups

have different underlying deficits, and potentially different outcomes, despite similar surface symptoms.

5.8.3.4 Oro-motor skills

As discussed in section 1.3.5, there is evidence of motor deficits in older children with persistent SSD. However, the results of the current study show that oro-motor skills in pre-schoolers were not a predictor of resolved SSD in this sample, consistent with findings from Brosseau-Lapré and Rvachew (2016) that DDK scores did not contribute to articulation accuracy in a similar aged sample. Only 4 (5%) children in this sample scored below normal limits in the DDK task at initial assessment, similar to the figure of 4.9% of children who had difficulty with the repetition of both single- and multi-syllable sequences found in Brosseau-Lapré and Rvachew (2016)'s study of 72 children with SSD. Neither of the two who presented for Re-assessment resolved their SSD, so it may indicate some potential for an indicator for the individual child, but this would need to be confirmed by further research. However, as an indicator it has very poor sensitivity, as poor oro-motor skills seem to affect only a small sub-set of children with SSD, consistent with findings from Vick *et al.* (2014) as outlined in section 1.3.5. The data collected here suggests that differences in these skills may not become apparent until children pass the age of short-term normalisation of SSD: oro-motor skills measured at a mean age of 4;1 did not differentiate the resolved SSD from persistent SSD group, but the persistent SSD had a significantly poorer mean score than the resolved group on the same task when measured at 6;2. Additionally, no child in the resolved group showed DDK scores outside normal limits at Re-assessment, compared to three children in the persistent group who did. Several children in the persistent group appeared to lose pace with development of oro-motor skills, showing worse scaled scores at Re-assessment than at Initial Assessment. Differentiation at older ages is consistent with the findings of Wren *et al.* (2012), who found children identified with SSD at age 8 had poorer scores on measures of DDK than TD children, when all types of SSD were included. However, once children with distortions, largely acknowledged to have a motoric basis, were excluded from the sample, in Wren *et al.* (2016) concurrent measurements of DDK were not significant predictors of persisting SSD.

It may be that more fine-grained measures of motor ability are needed to identify oral motor deficits than those widely available to clinicians in Ireland and the UK: other studies which have found motor deficits in children with SSD have used instrumental measures (Edwards 1992, Vick *et al.* 2014, Gibbon and Lee 2016). Alternatively, it may be that the range of motor abilities in very young children is so wide as to mask those with true motor difficulties, so that those who go on to have persisting SSD from this cause are not those that stand out at the start, but those who fail to develop at the same rate as their peers who attain age appropriate speech.

5.8.3.5 Co-morbid language impairment

Whilst numerous studies have shown worse outcomes in regard to literacy and academic achievement for children with SSD+LI compared to those with SSD alone, few studies have investigated whether presence of a co-morbid language impairment affects normalisation of the SSD itself. The finding of this study that language abilities did not predict normalisation of speech and that there was no significant difference in resolution rate between those with and without co-morbid language impairment, is consistent with findings from Morgan *et al.* (2017), Shriberg *et al.* (1994b) and Lewis *et al.* (2000), who all found that language status was not a predictor of normalisation of SSD. Similarly, Roulstone *et al.* (2003) found similar rates of improvement in speech accuracy for 2 to 3 year olds with and without LI over 12 months, and Broomfield and Dodd (2005a) found no broad effect of co-morbid LI on rates of improvement in SSD over 6 months. However, again there are conflicting findings: Roulstone *et al.* (2009) found that expressive language skills measured at age 2 and age 5 predicted SSD status at age 8, and in the longitudinal study reported by Stackhouse *et al.* (2007) and Nathan *et al.* (2004a), fewer children with SSD+LI (7/19) at a mean age of 4;6 resolved their SSD by age 6;9 than children with SSD alone (12/19). However children with SSD+LI had much more severe speech difficulties at 4;6 than those with SSD+LI, and the latter group had actually improved more in terms of mean PCC by 6;9, indicating that they were catching up.

In the study reported here, the fact that co-morbid language impairment appears not to predict a smaller likelihood of resolving the SSD, is consistent with findings from Chapter 4 of this study, which showed speech accuracy and language abilities were not associated, suggesting that speech output is an encapsulated ability. However, it must

also be taken into account that most of the SSD+LI children in this sample had only mild language impairment.

5.8.3.6 Non-word discrimination

The one factor that did significantly predict normalisation of phonological deficits was the score on a test of Non-Word Discrimination. However, while the measure used in this study had high specificity, it had low sensitivity; that is, in this sample, a score above the median at initial assessment guaranteed that phonological deficits would resolve by reassessment, but a poor score did not predict that a child would not resolve. Stackhouse *et al.* (2007), in their longitudinal study of 47 children with SSD aged 4-5 over 2 years, also report that children whose speech problems persisted performed less well on the auditory discrimination task. Thus they concluded that children with persisting speech difficulties are more likely to have both input and output difficulties. Rvachew (2006) also found a significant association in a preschool longitudinal cohort between speech perception skills at initial assessment and speech accuracy 14 months later. Her results suggest that speech perception leads growth in articulation skills as development occurs. As results from categorical perception tasks and word-recognition experiments suggest that speech perception abilities are still developing in preschool and early school age (Nittrouer 2002b, Munson *et al.* 2005a), it may be that some of those children who presented at initial assessment with poor speech perception developed better speech accuracy as their speech perception abilities improved. For example, they may have formed better phonological representations due to better speech perception, or they may have developed categorical perception skills allowing them to focus on the features of the phonological system that are more salient. In contrast, those who did not improve in speech perception may not have improved in production either, while those who already had good speech perception and auditory discrimination at initial assessment went on to resolve their production-only deficits more quickly. There is a difficulty with the task used in this study in identifying exactly what it is testing. As well as speech perception and discrimination abilities, it requires phonological encoding and memory, although the fact that results did not correlate highly with the auditory memory task may mean that it is not testing memory alone. Wren *et al.* (2016) found that a related task, non-word repetition (NWR) was associated with persisting SSD, but NWR also requires skills of phonological assembly and motor planning and execution. The auditory Non-Word

Discrimination task might be better able to show deficits in receptive phonology, as results are not confounded by articulatory abilities. The unique benefit of this particular task was its ability to discriminate individual children who resolved (if a child scored above the median, his/her phonological problems resolved), rather than note a characteristic for a group as a whole, which does not allow a judgement about an individual. However, the fact that so few children in this sample did not resolve may mean that this does not apply in different samples, or over shorter time-spans. Further research is needed to test the utility of this task.

5.8.3.7 The importance of age

For clinical decision-making, it is important to know whether factors posited as being predictive of persisting SSD are predictive for children at specific ages. As in the study being reported here, age itself has not been found to be a predictor of normalisation of SSD; however some of the conflicting findings reported by studies looking for predictors of persisting SSD may be due to those factors, or the outcome, being measured at different ages. Samples measured at more than two time-points allow the investigation of this. One such study is the ALSPAC study, where Roulstone *et al.* (2009) found that the proportion of speech errors produced by children at age 2 did not predict SSD status at age 8, but the proportion at age 5 did predict this. Results from another longitudinal study, the ELVS study of 1,494 children, are reported by Eadie *et al.* (2015) when the children were aged 4, and by Morgan *et al.* (2017) when they were aged 7. The demographic factors of male gender, positive family history of speech and language difficulties, and SES, which predicted SSD at age 4, were no longer significant predictors for SSD at age 7. In this study, although most children with SSD were male, gender did not distinguish those who resolved from those whose SSD persisted: thus male gender is a concurrent but not a prospective predictor. Similarly, oro-motor skills measured at preschool did not differentiate the persisting from resolved SSD groups 2 years later, but oro-motor skills measured concurrently were significantly poorer for the persisting SSD group. Thus it is important to distinguish factors that are a risk for having an SSD at pre-school school ages from factors that predict recovery or persistence of the SSD beyond the age of normalisation, and again to distinguish deficits that differentiate children with persisting SSD from those with resolved SSD concurrently, from true predictors that predict a future state from an earlier circumstance. It is the latter type of predictor that will aid in clinical decision-

making about provision of therapy for children with SSD, at the time of referral. Most specifically, what would be most useful is a factor that can be measured at the age of referral, and that will predict the likelihood of the SSD being resolved by the age of normalisation. In this study, only Non-word Discrimination was a useful predictor for that.

5.8.4 Change in diagnostic sub-group and characteristics

A second aim of this study was to investigate whether children's error profiles, as categorised by Dodd's Differential Diagnosis system, remained stable. In common with many cross-linguistic studies, it proved possible to classify the children's error profiles at initial assessment using this framework (e.g. Fox and Dodd 2001, Topbas 2006). However, the distribution of children in the different sub-groups was not maintained, as the IPD sub-group disappeared and the CPD sub-group retained just one member by Re-assessment. Children who were age appropriate remained age appropriate, and children moved in and out of the articulation disorder sub-group. The remaining changes were systematic: children with PD nearly all became age appropriate; none moved into the disordered groups, although some remained delayed. All but one child with a diagnosis of phonological disorder (i.e. CPD or IPD), either became age appropriate or moved into the delayed sub-group, or articulation disorder group. The exception was a child with CPD at age 5;2, who remained with CPD at Re-assessment. Only two children over the age of 4;5 at Initial assessment in this sample were diagnosed with CPD, and neither resolved: the second remained with an articulation disorder. Like the findings in Broomfield and Dodd (2005b), the diagnostic sub-groups exhibited a continuum of severity and age: as mean age increased, mean severity as reflected both in PCC and age adjusted DEAP Scaled Score, decreased, so that the IPD sub-group were the youngest and most severe, the CPD group were second youngest and second most severe, and the PD group were oldest and mildest. However, children in the two disordered sub-groups made bigger gains in PCC than those in the delayed group over the 2 years, suggesting greater progress toward normalisation – "catching up", unlike the findings of Broomfield and Dodd (2005a), where children with disorder made little progress without therapy. The interval measured in their study was much shorter, at six months, and the children spanned a greater age range, which may account for some of the difference.

The fact that there is systematic change from disordered to delayed or age appropriate errors was also supported by data on speech error stability from the ELVS study reported by Dodd *et al*. (2017). This showed that all children with PD had either age appropriate speech, or continued in the PD group at age 7, whereas only one of the 20 children with CPD at age 4 still showed atypical errors at age 7, with the remainder having either age appropriate speech, delayed errors, or an articulation disorder only. The 4 children in the IPD group had increased mean PCC from 56 to 86 over the 3 years, 2 now showing only delayed errors, and 2 atypical errors, but the rate of inconsistency at age 7 was not reported. However the high mean PCC for the group at age 7 makes it unlikely that many of them could have still been producing a rate of inconsistency of over 40%.

Regarding the group with articulation disorder, those with lateralisation did not resolve. Some children acquired dentalised lisps after the age of referral, which were not articulation disorders, as all children produced the sound correctly in some contexts. It is not clear whether this could be a phonological phenomenon, or may be related to demands of producing the fine motor movements needed for sibilants as linguistic contexts become more complex, as it has been found that increased movement variability in children is exacerbated by increased utterance complexity, as a result of the cognitive processing demands on speech motor control (Vick *et al.* 2012). More precise tracking of this phenomenon using instrumental measures, might serve to clarify this. For example, Baum and McNutt (1990) found that children with a frontal production of /s/, heard as /θ/, in fact maintained temporal, amplitude, and spectral differences between "misarticulated" /s/ and /θ/ when instrumental measures were used. However, whether phonological or motoric in nature, in the sample reported in this study, and as reported by Fox and Dodd (2001) and Dodd *et al.* (2017), dentalised productions of /s/ appear to be common in this age group. Smit (1993), in the Iowa and Nebraska norms study also noted that dental distortions for /s/ remained common throughout age groups, finally plateauing at about 10% in the 6 to 9 year olds. Figure 5-4 in section 5.1.3 shows a decline in the presence of dental distortions after age 6;06 in this sample. Also consistent with the results of this study, Smit found that lateral distortions generally remained, at a low level, throughout.

Three children with CPD resolved their phonological errors but had persistent articulation disorders, in all cases lateralisation of alveolar fricatives, whereas no child

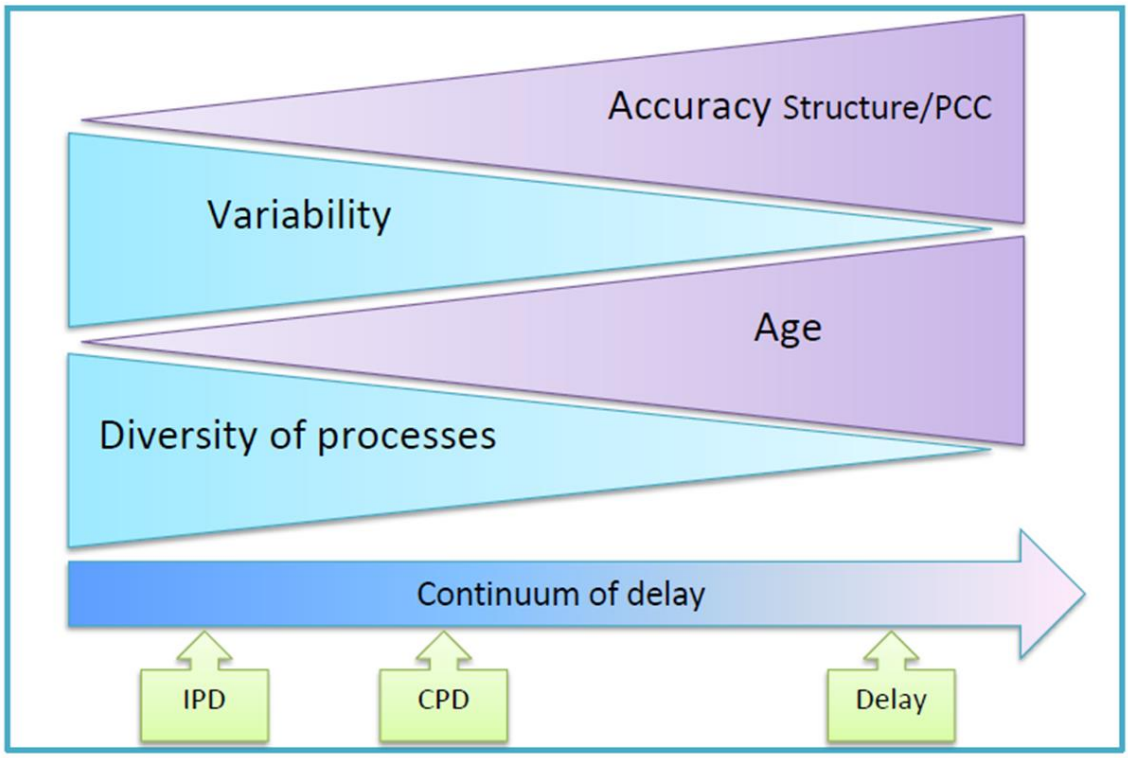
from another diagnostic sub-group showed a residual articulation disorder in the study reported here. An association between articulation disorder and the CPD group was also shown in a smaller study of 17 Cantonese-speaking children with SSD, where the two children with additional articulation disorders (i.e. difficulty with the articulation of /s/ and /l/) were both in the CPD group (So and Dodd 1995). Dodd et al's (2017) study of the ELVS sample also showed that two children with CPD changed diagnosis to an articulation disorder only after 3 years, with a further one moving to PD plus articulation disorder. In this sample, the children who had CPD at initial assessment had the worse mean scores on DDK rate at Re-assessment (Standard Scores: PD: 10.1; IPD: 9.7; CPD: 9.3). It may be therefore, that the CPD sub-group is complicated by a minority of children who have motoric deficits, so that some of the errors classified as atypical have a motoric basis, as posited in section 5.8.3.3, which may be part of the reason why fewer are reported to resolve their SSD (e.g. (Morgan *et al.* 2017). It is possible that some of them, despite resolving their phonological deficits, are left with distortion errors, so it is actually distortion errors, as found by Preston *et al.* (2013), that predict persisting SSD rather than atypical errors. This accords also with the findings of Wren *et al.* (2012), that children who demonstrate both phonological and articulatory errors are more likely to have SSD that persists. Whether children who make atypical errors are more likely to also have motoric deficits would be worthy of investigation in a larger sample with more rigorous measures of motor ability.

5.8.4.1 SSD as a continuum of delay

The fact that children appeared to move in one direction only regarding change in sub-groups between initial and reassessment, i.e. either to PD or age appropriate speech, allows the suggestion that SSD of unknown origin is a continuum of delay rather than representing different sub-groups of deficits. Support for this is suggested by the fact that the IPD were youngest and most severely affected, the CPD group were second youngest and second most severely affected, the PD group were oldest and mildest. In this study, one child maintained her diagnosis of CPD, but all of the others progressed either to PD or age appropriate speech. No children were observed to pass from IPD to CPD, but this could be due to the long time between initial and reassessment, and some support for this trajectory was in fact observed for two children in ELVS sample (Dodd *et al.* 2017). It is therefore possible that the different surface symptomatology reflects different stages in the process of normal phonological acquisition, with high

token-to-token variability and diversity of processes at young ages gradually reducing with age, to result in speech that is more accurate. We can then conceptualise children still at the beginning of the continuum when variability is high as those with IPD, children who have reduced variability but who still have a limited range of possible structural templates (which has a larger impact on PCC) and who use a wider range of error patterns, (some uncommon) as those with CPD, and children who are purely having difficulties at the segmental level and with the most complex structures (i.e. CC and CCC, especially where these do not fit the sonority curve), as those with phonological delay. This model is depicted in Figure 5-18 below.

Figure 5-18: Model of continuum of delay



5.8.4.1.1 Severity of impairment

To support this model, it would be necessary to show that severity of impairment and diagnosis changed as children got older, and that diagnoses conceptualised as being further up the continuum contained children with less severe impairment. This was true for this sample, as well as in the sample reported by Dodd *et al.* (2017). Broomfield and Dodd (2004b) also noted that children with PD rarely presented with a delay of more than 2 years, or with severe or profound impairment. In this sample just 6% of

children with PD were classified as severe in Shriberg's system, compared to 29% of the CPD and 63% of those with IPD.

5.8.4.1.2 Variability

It would also need to be shown that in typical development, variability decreases with increasing age. This has been shown by several studies (Burt *et al.* 1999, Holm *et al.* 2007, Macrae 2013). If the children with IPD are merely at the most severe end of a continuum of delay, it would be expected that the degree of inconsistency with which they present is the same as that seen in typically developing children at a younger age, and that all children who are delayed will show more variability than typically developing children of the same age, even if they do not meet the diagnostic criterion for IPD. There are conflicting findings in the literature regarding the degree of inconsistency in young, typically developing children's speech. Holm *et al.* (2007), found mean rates of variability of under 13% in 20 children aged 3:0 to 3:5. However, more recent studies have found much higher variability rates. Sosa (2015) in a study of 33 children aged 2;6 to 3;11, found a mean variability rate of 57% when only consonant differences were counted, as in this study, over a similar sample of words. Similarly high rates of inconsistency were found by Macrae (2013) and McLeod and Hewett (2008) with 2 to 3 year olds, using different sampling tools. Thus, these findings suggest that the mean inconsistency rate of 54% in the IPD children in this study is not atypical of normal development at an earlier stage, and the fact that 7 of the 8 children identified with IPD were in the youngest age group, 3;6 or below, also suggests that high variability may be a typical phase in development of very young children. All children with IPD progressed to normal speech or PD, including those who did not receive the appropriate type of treatment or a significant amount of treatment. There was no child with IPD in this sample over the age of 51 months, whereas children with consistent disorder and phonological delay were present up to 81 and 78 months respectively. Broomfield and Dodd (2004a) also reported that, in their sample of all children referred between the ages of 3:0 and 5:11, the majority of children with inconsistent deviant phonological disorder were referred between three and four years. Fox and Dodd (2001), in a sample of 110 German speaking children, also found the IPD group to be the youngest, with a mean age of 4;2. Preston and Koenig (2011), in a study of 20 children aged 9-15 with residual SSD, found no data to support the existence of a sub-group of children with a high level of variability in this

age group, suggesting it is not a qualitatively different condition. Children with CPD and PD have also been reported to have higher inconsistency scores than normative data for children with TD speech (Dodd 2011), supporting the view of inconsistency being an early stage on the continuum of development rather than a qualitatively different condition.

5.8.4.1.3 Diversity of Processes

Children with CPD are classified as such on the basis of use of uncommon processes (used by <10% of the population). However they also use common processes such as those used in normal development. Some researchers have indicated that these idiosyncratic or uncommon processes may also be found to some degree in children whose phonological development is considered normal (Hodson and Paden, 1981; Stoel-Gammon and Dunn, 1985). Roberts et al. (1990), suggest two alternative explanations: it could be that uncommon processes are used often by children whose phonological development is classified as disordered but are used rarely by children developing a normal phonological system. Alternatively, it may be that all children use these processes at some point in their development, but that children with speech disorders persist in their use longer than children exhibiting normal phonological development, which would again show a type of delay. In a study of 147 typically developing children between the ages of 24 months and 8 years, they found that the TD children infrequently used processes usually considered as disordered, and that in fact, some processes defined as uncommon in normally developing speech (e.g., deletion of medial consonants) occurred with a higher frequency than other processes classified as common. James (2001), in a study of 50 children aged 2 to 7 who had not been referred to SLT services and were considered to be typically developing, found that over 50% of the 2 to 4 year-olds demonstrated the phonological processes of backing and initial consonant deletion, and over 40% showed affrication, all of which are classified by Dodd *et al.* (2002) as atypical errors. Preston and Edwards (2010) also found that, in a sample of 43 children diagnosed with SSD of unknown origin, aged 4 to 5, all made some atypical speech errors. There is also some evidence that the number of atypical errors reduces with age and severity. Rvachew and Brosseau-Lapré (2012), in a study of 58 preschool children with SSD followed up after 1 year, found more atypical errors in the more severe groups at both ages, and an overall reduction in numbers of atypical errors after 1 year. Preston *et al.* (2013) in a study of 25 children with SSD at 4;6, found only one occurrence of an atypical error at follow-up at age 8,

indicating that atypical errors had resolved by this age. James (2001) found that use of the atypical processes listed above declined rapidly between the ages of 3 and 4. It is therefore possible that the persisting or frequent use of uncommon processes is a symptom of delayed development, rather than a qualitative difference.

There is therefore some support for the model above, to show that inconsistency and use of atypical processes reflect a gradient of severity and age along a single continuum of delay. Further support for the continuum of delay model is found from longitudinal studies by Roulstone *et al.* (2009) and Shriberg *et al.* (1994a), who both found that the order of acquisition of sounds was the same for children with or without SSD. Bernhardt and Stemberger (1998) have also noted that similarities in phonological development outweigh the differences between children with SSD and children with TD speech, with most developmental patterns being the same for both. They also indicate that, for both children with TD speech and children with SSD, children's productions change with time, with most change being positive, i.e. in the direction of the adult form. They suggest that in TD children this change to the phonological system is rapid with frequent developmental changes, while in children with SSD these changes are more protracted with the possibility of plateaus in learning occurring. Thus where the children with SSD differ from those with TD speech is in degree of prevalence of developmental patterns and constraints, and age of acquisition for the components of the phonological system. They therefore coined the term protracted phonological development, rather than phonological disorder or speech sound disorder for those children, which is both a more neutral and positive term (Chávez-Peón et al. 2012) and is in accord with the argument presented here that SSD is a continuum of delay in acquisition, rather than a set of qualitatively different disorders, some of which arise from lack of knowledge of the phonological structure of the language. Cross-linguistics studies of children with protracted phonological development by Bernhardt et al. (2015a), Dubasik and Ingram (2013), and Chávez-Peón et al. (2012) showed a bias towards within-inventory substitutions (Bernhardt et al. 2015b), and language specific mismatches, showing that differences in phonological systems appeared to affect developmental patterns for children with protracted development in the same way as those with TD speech, suggesting that the former have similar knowledge of the ambient phonological system to the latter. Thus protracted phonological development may be a better term than SSD for children whose SSD is of a phonological rather than structural or organic nature, reflecting the

fact that most of these children will eventually attain age appropriate phonology in a process of development similar to, but slower than, their TD peers.

5.8.5 Predictors of phonological awareness (PA) and reading abilities

It is known that children with SSD are at risk of later problems with acquisition of literacy even after their SSD resolve. Therefore a fourth aim of this study was to investigate outcomes in regard to phonological awareness abilities and reading.

Findings with regard to PA show fewer children went on to have PA difficulties than recorded in other studies. This may be due to several factors, including different levels of severity in the samples and different assessment measures. For example, Rvachew *et al.* (2007) found 47% of 5-6 year old children failed the test of phonological awareness, compared to 12% in this study, but their participants did not include any in the mild range of speech impairment, whereas this study included 42% in the mild range. Another possible confounding factor is that local schools and preschools had adopted a synthetic phonics programme for literacy just before this study took place, and many parents reported success both for beginning reading and improvements in speech when their children started this programme.

Although there was a trend toward lower mean PA scores for the groups who had had disordered rather than delayed phonology at Initial Assessment, the finding was not significant. When entered into a multiple regression model that contained results of Pattern Recognition and Non-word Discrimination tests, atypical errors no longer made a significant contribution to the variance in scores of PA at Re-assessment. Thus, unlike Preston *et al.* (2013)'s study, there was no clear finding that atypical errors at preschool age predicted low phonological awareness skills at school age, a similar finding to Rvachew *et al.* (2007)'s study. However, there may be an effect of age, or time between follow up: the children in Preston *et al's* study were followed up at age 8, after 4 years, rather than after 2 years, at a mean age of 6;2 in this study. The children in Rvachew *et al's* study, which found atypical errors were not significant predictors of PA difficulties, were a more similar age to this study, and follow up was after one year. Hayiou-Thomas *et al.* (2016) also found no strong evidence that disordered rather than delayed errors in a sample of 59 children indicated more serious deficits in PA or literacy, when children's speech was assessed at a mean age of 3;8, and reading at age 5;6 and 8. It may be that clear differences do not emerge until the

variability and diversity of errors in the younger children's speech profiles reduces, which would also explain why findings differed from those of Leitão and Fletcher (2004), as their finding of an association between atypical errors and poorer PA and greater difficulties with literacy at age 7 and again at 12-13, involved children who had used atypical error patterns at age 5-6, which is much less common. Only two children in the study reported here made atypical errors after the age of five.

In fact, it was results on tests of cognitive abilities, i.e. pattern recognition, at Initial Assessment that were more reliable predictors of PA abilities at Re-assessment. This aligns with findings reported by Peterson *et al.* (2009) that there is growing evidence for the importance of nonverbal IQ in literacy development among children with early speech and/or language impairments. Several studies have reported that children with speech and/or language impairment and low nonverbal IQ have poorer literacy outcomes than children with speech and language impairment and normal nonverbal IQ (Bishop and Adams 1990, Bird *et al.* 1995, Catts *et al.* 2002).

In the sample reported here, NWD ability also made a significant unique contribution to the variance in PA scores at Re-assessment. This reflects findings by Rvachew (2006), in a longitudinal study of a sample of 47 children, initially aged between 4 and 5;7 (mean 4;10) and reassessed one year later. She found that speech perception ability, as tested on a Word Identification task, explained 11% of the variance in PA skills at re-assessment, very similar to the 9% in this study. However, she found that receptive vocabulary explained a further 10%, whereas that was not a significant predictor in this study. It may be that the standard score used in the analysis in the study reported here was not sensitive enough: Rvachew's study used raw scores. In addition, she did not include any measures of cognitive ability in her analysis.

The figure of 22% of children with SSD going on to have reading difficulties is the same as that reported by Peterson *et al.* (2009), although the distribution between those with and without concomitant language impairment is somewhat different: Peterson found 67% of those with SSD+LI went on to develop reading difficulties, compared with 10% of those with SSD alone, while the respective figures in this study were 38% and 14%. The lower figure in the current study for children with co-morbid LI may be explained by the generally mild nature of language deficits in the sample reported here. Nevertheless, the present results are in line with the growing body of research

indicating that a substantial portion of the overlap between SSD and reading difficulties is due to the third variable of LI.

5.9 Limitations

One limitation of this study is that it is not possible to know how soon the children with good outcomes resolved. Two years is a long interval and it cannot be known whether those children who resolved their SSD without therapy would have benefited from therapy to resolve sooner. However, no parent of a child who did not receive therapy expressed dissatisfaction with this. Additionally, it is not possible to tell whether those children who resolved did so by making steady progress, or whether there were spurts and plateaus of development.

Another limitation is the sample size, which meant, when children were split into sub-groups, some were very small. This was exacerbated by the large amount of missing data, which reflects the difficulty of assessing children in this age range. These two factors meant some categories of children were very small, so that all results must be treated with caution. However, attempts to compensate for this were made by looking at frequencies and the distribution of results in scatter plots: these were in accordance with the findings from statistical analysis, showing usefulness at a clinical level.

Many studies of the natural history of SSD in children are confounded by therapy, and it was an uncommon feature of this study that circumstance provided a group of children who did not receive therapy. However, assignment to the therapy/no therapy group was not random and it was evident that clinicians were prioritising children with more severe deficits for therapy, so the outcomes for those who had no therapy are not generalizable to the wider population. As many of the children had been discharged some time previous to the follow-up assessment, their files had been archived so details of therapy received only came from parent report. Nevertheless, it was evident that many children had received an amount of therapy far smaller than anything for which efficacy has been proven, so it is possible that many in the therapy group would also have resolved without therapy.

5.10 Conclusion and clinical implications

This study was designed so that its findings would be replicable by clinicians who treat children with SSD. To do that, it used a sample of children referred to local SLT services, thus reflecting the caseload they are dealing with, case history and demographic information usually collected during assessment, and assessments which are widely available or at least feasible for clinicians to use in practice in the Irish and UK health services.

The high resolution rate in this study reflects the difficulty in ascertaining at a young age which children are truly cases, in the sense that without intervention they would necessarily go on to experience problems associated with early SSD in school and beyond (Law *et al.* 2000). Given the wide variation in trajectories of development at young ages, the younger children are, the harder it is to ascertain which are actual cases. The importance of early intervention for children with defined disabilities, and the effects of early deprivation in the home are well evidenced and have led to a "type of conventional wisdom" (Casto and Mastropieri 1986) that the earlier intervention can be provided, the better the outcomes will be. SLT services are often focused on the needs of children in the early years (Wren 2015). This emphasis on early intervention, in the context of limited resources, necessarily leads to the curtailment of provision, evidenced by Dockrell *et al.* (2006) and Gross (2011), for older children, amongst whom definite cases who will not improve without therapy are easier to identify. Early intervention has been conflated with intervention at an early age, but what it should mean is intervention as soon as a problem becomes evident. The evidence from this study, that the parents of children who did not receive therapy did not express dissatisfaction with this, and from studies by Roulstone et al. (2003), Paul (1996) and Dodd *et al.* (2017), that the parents of a large number of pre-school children with speech and language disorders did not seek to be referred for intervention, even when told that their children's speech and/or language was not developing at the normal rate, suggests that a clinical diagnosis of SSD or language delay in pre-schoolers is not always viewed by parents as a problem big enough to warrant intervention. As Bishop and Leonard (2001, p.xii) wrote: "We need to balance the benefits of starting intervention early …, against wasting intervention resources on children likely to resolve spontaneously." We also need to consider the costs to families and children of taking up intervention that may not be necessary. Therapy given to children who will

resolve without it represents an opportunity cost to them and their families in terms of time and other activities, and may represent significant financial costs for parents who have to take time off work. Where services are rationed, it also represents a cost to those children who are definite cases and who cannot access enough, or indeed any, therapy. Rationing may even result in no one getting enough therapy to make a difference. In a randomised controlled trial of therapy for preschool children which found few differences in outcome between those who did and did not receive therapy, Glogowska *et al.* (2000) suggested that for this group of children, the appropriateness, timing, nature and intensity of therapy should be reconsidered. In interviews with parents of children assigned to a one-year waiting period in the randomised controlled trial, Glogowska and Campbell (2000), found that the identification of early difficulties with speech and language did not inevitably mean that parents wanted therapy to begin. Some parents were happy with the opportunity to let the children develop on their own at their own pace, while others felt that after the waiting period the children were better able to participate in therapy, and to derive more benefit from it than they would have at an earlier age.

In the sample of children in this study, neither severity nor error type, both factors commonly used in decisions about provision of therapy by clinicians, were specific predictors of resolution. For children with SSD of the phonological type, it may be that the rate of progress is more important than children's presentation at a single point in time. In the literature, the youngest age for which it was shown that error status at that point predicted error status later, was 4;6 (Nathan, 2004). It may therefore be appropriate for clinicians to operate a monitoring approach for children younger than 4;6, with therapy offered to those who appear to have plateaued in development, or to those who are experiencing significant limitations to activity and participation due to their SSD.

The only factor that was both specific and sensitive to case status in this study was lateralisation, but while there is evidence that lateralisation can be successfully treated in older children (Gibbon and Hardcastle 1987, Wright *et al* 2017) this is lacking for preschool children.

In this study, children who had good non-word discrimination abilities were more likely to resolve than those who do not. This is in line with findings from Rvachew

(2006) that speech perception abilities are important to the acquisition of speech accuracy. However currently there are no standardised or widely available assessment tools for auditory discrimination or speech perception for use with pre-schoolers with SSD. The one training programme that has shown efficacy in improving speech perception abilities in children with SSD of unknown origin (SAILS; Rvachew *et al.* 2004) has to be tailored to the variety of English spoken, so is not currently available outside of Canada. The need for assessment tools to identify children who have deficits in auditory discrimination and speech perception, and effective interventions for those who could benefit from them has already been raised (Preston *et al.* 2015, Rvachew 2016).

This study was carried out with the aim of aiding clinicians in making difficult day-to-day decisions in practice about who should receive intervention and who can be safely monitored. It was carried out with a sample who would qualify for services according to service criteria, rather than the more exclusionary samples seen in many research studies; the assessments carried out are generally widely available and feasible to use in health services, and resolved status was judged on criteria used for discharge in such services.

Replication of these findings with bigger samples is needed, as well as ethically sound randomised control trials with a monitoring only arm, such as that carried out by Roulstone *et al.* (2003), to be able to answer the question posed above with any confidence.

The next chapter is the last in this thesis and is a concluding discussion.

6 Chapter 6: Concluding Discussion

This chapter attempts to bring together results from the cross-sectional and longitudinal study, putting them into the context of the existing literature. Findings regarding the abilities which determine speech accuracy and a theory of how delay might occur are discussed, as well as the utility of sub-types in SSD. This is followed by an outline of what was discovered about the natural history of SSD in a clinical sample that reflects the caseload of SLT services in Ireland and the UK, and what might predict transient as opposed to persisting SSD. Implications regarding management of SSD in clinical services are set out. Finally the limitations of the study are discussed, together with ideas for further research.

6.1 Which abilities determine speech accuracy?

As introduced in section 1.3, Munson *et al.* (2005a) suggests that perceptual, articulatory, and higher-level phonological knowledge (i.e. how words are segmented into sounds, and how sounds can be combined into sequences in words), and social-indexical knowledge (not considered here) are needed for the successful acquisition of speech. In addition, Dodd (2011) posits that SSD in some children is caused by an underlying cognitive difficulty in abstracting the rules of the phonological system, thus resulting in a deficit in higher-level phonological knowledge. A factor analysis was carried out to investigate the association between these skills and speech accuracy in children with SSD. This was the first factor analysis of the speech domain to include measures of general cognitive ability, such as number concepts and pattern recognition, as well as measures of rule abstraction and cognitive flexibility, along with measures of language and motor skills.

The results of the cross sectional factor analysis in this sample suggested that speech, as measured by output accuracy, is not determined by any one factor of perceptual, motor, cognitive or linguistic abilities. The results factored very clearly into four discrete and easily interpretable domains, with separate factors for perceptual/cognitive, language, and non-speech motor abilities that were not associated with the factor representing speech accuracy. The non-speech cognitive skills represented by tests of number concepts and pattern recognition were not associated strongly with any factor, suggesting that, in children of this age, speech accuracy is not associated with general intelligence. The finding that no association was seen with

language skills gives further support to the view that SSD is not simply the top end of a continuum of language impairment (Lewis *et al.* 2006a), and the lack of association with non-speech motor skills is consistent with the notion that it is not part of a pervasive motor disorder, in accord with the absence of evidence that SSD is amenable to intervention targeting non-speech motor skills (Lof and Watson 2010). However DDK rate also failed to factor with speech, and there was no association between the perceptual/cognitive factor and speech. It seems counterintuitive, and at odds with Munson's contention above, that speech accuracy does not have some relation to perceptual/cognitive abilities, so what it brings us back to is the well-recognised heterogeneity of children with SSD: if SSD in different children arises from different causes, and if children show similar presentations of delayed speech for different reasons, speech accuracy appears to be independent of any one factor, because for different children poor speech accuracy is due to a different factor or factors. The same phenotype, then, may have different underlying causes or deficits.

6.2 How does SSD occur?

If all SSD is delay, as suggested in section 5.8.4.1, how does this delay occur? Three possible alternative explanations were suggested in section 1.3.3.: firstly that SSD arises from a multiplicity of sources, perceptual, cognitive, or motor, with different children having different deficits; secondly that domain general cognitive deficits originally contributed to difficulties with the acquisition of phonology, but have since been outgrown, leaving the phonological structure, however, fixed at an immature stage; and thirdly that in some children, cognitive deficits accompany, rather than underlie, SSD, causing SSD to be more resistant to normalization or treatment, as children are unable to use adaptive or compensatory strategies to make up for deficits in other areas, or to benefit from treatment.

Because speech develops while children are still very young, it is developing contemporaneously with perceptual, motor and cognitive abilities, as outlined in section 1.2. In neuroconstructivist theory, cognitive development is conceptualized as a process that emerges out of multidirectional interactions between cognitive, genetic, biological, behavioural, and environmental systems (Karmiloff-Smith 2009). In a more specific way, it is possible to conceptualise the development of speech as emerging out of multidirectional interactions between perceptual, cognitive and motor skills, for

example, as envisaged by Velleman and Vihman (2002), where sound patterns a child has learned to produce in babble are more salient to her/him when they are heard in the input. As abilities in these domains are still developing, a mismatch in maturation between any of them could result in an overall delay. The different parts of the developing system may interact with other parts differently at different times, so that a process that is vital at an earlier time may no longer play a role at a later time, but its delay at the earlier time may have affected the developmental trajectory (Thomas *et al.* 2009). For example a delay in the perceptual system may result in underspecified or wrongly specified motor programs for production, while a delay in domain general cognitive processes that support statistical learning might result in difficulties in the transition from holistic to segmented phonological representations. Processes which are upstream fail to feed forward to processes downstream, while delay in downstream processes both affects output and fails to provide necessary feedback to upstream processes. However, if the delayed process then matures, so that the necessary feed forward/feedback recommences, speech can progress in development again, and because the other systems have matured in the meantime, this progress can be rapid. However, where delays occur in the maturation of more than one process, possibly at different times, the system as a whole does not experience a time when all systems are mature at the same time, so this rapid progress does not happen, or becomes stalled again at a different point. We can posit then, that there is a wide variability in the rate of development of underlying skills in younger children, but, if as described in McKean *et al.* (2013), the phonological system fixes on a pattern of processing before all contributory processes have matured, a persisting SSD may occur. This may be more likely to happen if a child has a delay in more than one of the contributory processes, hence the finding that persisting SSD is likely to be multifactorial in nature.

As the majority of children in this sample were reported to have onset of speech at the normal time, and the majority also attained age appropriate speech, we can surmise a non-linear trajectory of speech development (Thomas *et al.* 2009) for children with SSD, with normal onset for most and delayed onset for a minority, slow rate in early development, followed by an accelerated rate (catching up) for most children with SSD of unknown cause, with a premature asymptote for the minority of children who do not normalise. This rapid rate of improvement from a low beginning, followed by a period of slower but continuing development was in fact what the trajectory constructed from cross sectional observations at both times of assessment showed (see section 5.1.2).

Because children were not referred until age 3, the onset and slow rate in early development could not be constructed, but must be assumed to have occurred before the age of referral, for SSD to be suspected.

6.3 Existence of sub-groups inside the group of children with SSD

If it is the case that deficits in one or more of these underlying processes leads to SSD, it may still be possible to find sub-groups of children within the population of children with SSD who share the same deficit or causal factor. Two systems of classification of children with SSD of unknown origin were introduced in section 1.4: Shriberg's system, (Shriberg *et al.* 2005), based on aetiology, and Dodd's system (Dodd *et al.* 2005), based on surface symptomatology. Neither was strongly supported by the data from the sample in the study reported here. Shriberg's suggested sub-groups did not align with a particular presentation, and neither Shriberg's nor Dodd's aligned with specific deficits in the tasks measured in this sample, nor did they predict outcome, once the analysis was controlled for age.

Shriberg's system did not account for all children with SSD, with 22% not falling into any of his suggested sub-groups, (although psychosocial involvement was not measured). The Speech Errors-Sibilants sub-group was partially supported by the data, with a considerable number of children presenting at Re-assessment with distortions of sibilants as their only error. However, only some of these fit Shriberg's explanation for this error, i.e. that they adopt a production pattern for the sounds when very young which reflects an immature motor ability, and retain this immature pattern despite subsequent maturation of production abilities. Those that displayed lateralisation appeared, as Shriberg would suggest, to have acquired an inaccurate pattern for production from the start of producing these sounds, but those who displayed a dentalised production of sibilants were frequently observed to produce the sounds correctly in some contexts, and in some cases had shown a normal production when younger, suggesting that this production is not always due to initial motor mislearning. Evidence from the literature (Smit 1993) suggests that the latter group only show a variation in normal acquisition, while the former group will go on to have persisting difficulties: this conclusion was supported by the results of this study. There were also children with distortions of /ɹ/, but none was outside of normal limits for acquisition of this sound, so were not yet classified as having persisting SSD if this was their only

error. In regard to the genetic sub-group, children who might be presumed to have a genetic disposition to SSD due to being male and/or having a family history of SSD did not show a qualitatively different symptomatology, and were not more likely to have specific deficits on any factor, nor to have persisting as opposed to transient SSD. Similarly, number of ear infections was not associated with a particular symptomatology, deficit or outcome. There was evidence that the children with persisting SSD at Re-assessment showed poorer scores on a DDK measure than those who resolved, providing some evidence supporting Shriberg's SDCS subtype of motor speech disorder—not otherwise specified (MSD-NOS), but this group could not be identified at Initial Assessment. Psychosocial involvement (PSI) was not measured in the longitudinal sample, so no conclusions can be drawn as to that sub-group.

Dodd *et al.'s* (2005) classification was able, with small amendments, to classify all children with SSD in terms of their symptomatology, and the distribution and characteristics of sub-groups was similar to other studies in the literature. However, contrary to Dodd's hypotheses, the sub-groups in this sample were not associated with specific deficits in any factor, nor with outcomes when age was controlled for. Instead, the sub-groups outlined in Dodd's system appeared to be associated with age and severity, illustrating, then, a continuum of development where variability and diversity of errors reduces as age increases. Where progress along this continuum stalled, SSD resulted, with differing symptomatology resulting from the stage at which the stall occurs. For example, large amounts of variability have been associated with transitions between stages and are said to occur at the beginning of the emergence of an organized system (Vihman 1996). Thus Sosa and Stoel-Gammon (2006) suggest that variability may occur as children move from holistic to segmented phonological representations, i.e. at the transition from Stackhouse and Wells' (1997) whole word phase to the systematic simplification phase. In children whose progress is arrested at this transition, this could result in the continued inconsistency in word production that is characteristic of Dodd's IPD sub-group.

The contention that children who make consistent, atypical errors do so because they are not able to abstract the rules of the ambient system is not supported by this study, which did not find that these children had specific difficulties with rule abstraction or cognitive flexibility. The contention is also difficult to defend from the point of view of the theory of acquisition. Firstly, it is not clear that they abstract the "wrong" rules:

such children do not violate the phonotactics of the language by using phonemes in illegal positions or combinations, nor do they generally introduce sounds that are not present in the native phonological system, except where coalescence occurs. They also usually make typical as well as atypical errors, showing that they have abstracted at least some of the same rules as typically developing children. Secondly, errors categorised as atypical comprise both structural and substitution errors, whereas substitution errors imply a better knowledge of the phonological system than structural errors. Consistent patterns of substitution errors imply that the child has progressed to the stage of a segmented phonological representation with at least partial knowledge of the target phoneme, whereas structural errors may indicate that the phonological representation is incomplete (Shriberg et al. 2005), or that the child's output is still constrained by a holistic representation (Stackhouse and Wells 1997) or a gestalt template, as described for an earlier stage of development than the segmented representation (Velleman and Vihman 2002). The child who makes atypical substitution errors must therefore have construed enough correct rules to progress to the stage of more finely specified segmented phonological representations, and it would thus seem counterintuitive to suggest that he or she makes atypical errors due to a lack of rule abstraction ability. If we consider the process of acquisition as described by usage-based theorists, it would seem more likely that errors that affect the structure of words identify the child who is having difficulty in passing from the early stage of holistic representations that may not be fully specified, to the stage of more detailed segmental representations, especially where structural omission errors create a large number of homophonous words, such as when the processes of initial and final consonant deletion are used. Although TD children also make structural errors by omitting elements in clusters later in development, there is evidence that children use strategies such as coalescence of features, and other contrasts (e.g. lack of aspiration of voiceless consonants following /s/, even where the [s] is not produced) to signal knowledge of the structure and reduce homophony. Thus these later constraints may be motoric rather than representational in nature. The theory of acquisition outlined in the introduction to this thesis suggests that holistic, unsegmented representations are a stage before segmented representations, so that those children, who make a large number of structural errors, especially in less complex structures, may be still at an earlier stage of development than those who make substitution errors. Thus omissions of structure rather than substitution errors might, as suggested by Shriberg *et al.*

(1994b), be more indicative of persisting SSD and difficulties with PA, which also requires accurately segmented phonological representations (Nathan *et al.* 2004a), than atypical errors per se. Previous research findings that children who use atypical errors have worse outcomes for speech and PA may have been confounded by the fact that errors in basic structures, such as initial and medial consonant deletion, are classed as atypical, so children using a large amount of this type of error would have confounded the results for the group as a whole.

The model where SSD of unknown cause is exhibited as gradations of delay rather than qualitatively different disorders is supported by studies by Shriberg *et al.* (1994a) and Roulstone *et al.* (2009), which showed that the order of acquisition of sounds in children with SSD is essentially the same as that seen in normal development.

Thus it seems that the deficit underlying SSD in a particular child cannot be detected from the aetiology, nor from the symptomatology. It was an aim of this study to look for evidence supporting the existence of newly defined sub-groups in the population of children with SSD. As in the very large international CATALISE study of developmental language disorder (Bishop *et al.* 2017), it was hoped that a pattern of verbal and nonverbal skills could be found in this study, that would point to natural subtypes with distinct causes or deficits, so that a profile of cognitive, linguistic and motor skills could be used to distinguish groups with SSD of different origins, thus suggesting the most appropriate interventions and likely outcomes. However, as in that study, no such sub-groups were found. Although cluster analysis of this sample did support three relatively clear sub-groups among the children with SSD, the small number of children with enough data to be included in the analysis, and the small number who went on to have persisting SSD, mean that results are not generalizable. However they did show some trends that may be worthy of further investigation. The profiles of cluster membership and likelihood to resolve support the contention of Dodd (2011) and Rvachew (2006), that it is underlying deficits that predict which children will go onto have persisting SSD, in that the children in the cluster with worst performance on all three factors other than SPEECH were least likely to resolve their SSD and most likely to have PA difficulties, whereas those with performance most like age appropriate children on all factors except SPEECH were most likely to resolve their SSD and have good PA. The fact that children least likely to resolve their SSD showed poor performance in more than one area also supports Stackhouse *et al.*'s. (2007)

conclusion that there is no one task that will predict a child's speech and literacy outcomes, and that persistent SSD are multifactorial in nature. It is also consistent with Wren *et al.*'s (2016) contention that there is involvement across more than one domain of skills in children whose SSD persist. In individual cases, there may be more than one locus of impairment: In Wright *et al.'s* (2017) case study of an adolescent with a severe persistent SSD, Liam's SSD also appeared initially to be a case of articulatory mislearning, but further investigation revealed how this initial articulatory mislearning appeared to have become compounded into deficits in phonological discrimination and inaccurate phonological representations. Impairment at multiple levels may in part explain why SSD in older children is resistant to therapy. Wren *et al.* (2012) similarly found in a sample of children with persisting SSD, that there were a group who demonstrated more frequent errors affecting a larger range of speech sounds. She suggested that these children also appeared to have a combined phonological and articulation basis to their SSD, as was the case for Liam in Wright *et al.* (2017). Two children in the study reported here had a primary diagnosis of CPD, and also a secondary diagnosis of articulation disorder at Initial Assessment, while only one presented with lateralised fricatives as his only problem, suggesting further research into whether SSD with combined phonological and articulation deficits predicts poorer outcome, might be worthwhile.

A single cognitive marker or endophenotype, therefore, may not be sufficient to identify homogeneous groups of children (Lewis *et al.* 2006b).

6.4 The natural history of SSD

Children with SSD of unknown origin, and not including CAS, made considerable improvements in PCC over two years, both with and without therapy. Almost all children improved, and the most severe improved the most. Bishop and Edmundson (1987) noted that the areas of speech and language that show most propensity for "recovery" are those that are generally mastered at a young age. As speech is largely adult-like by the age of 5 or 6 in TD children (although there is evidence for continuing development of accuracy in polysyllabic words, and further development of perceptual skills), the fact that the normal child does not continue to make further progress in speech skills, unlike for example continuing development of language, allows the majority of children with SSD to catch up with their peers within a

relatively short period. Only one child in the longitudinal study reported here showed very little change in presentation by re-assessment, indicating that in this clinical sample, it was unusual for a child to become permanently arrested at an earlier stage of development, or to remain with a disordered phonological system. In this sample, children with IPD all became consistent (although some remained with typical delayed errors), indicating that their phonological or motor planning had continued to improve, and/or that their phonological representations had become better specified. Children who made atypical errors at initial assessment no longer showed atypical errors at reassessment in all but one case. The fact that many children received therapy is, of course, a confound here and it is possible that some children would not have continued to improve without therapy. However, evidence from parent report suggested that few children received an amount and type of therapy that would be expected to make a significant difference to an intransigent SSD. For example, 27% of those who received therapy only received a home programme, with very little parent training.

Although evidence from studies where children received no therapy is sparse, findings from the two studies reported in section 1.5.5 where children did not receive therapy, as well as findings from the longitudinal study reported here, still showed high resolution rates within the period of normalisation, even without therapy. Therefore it must also be taken into consideration that, as suggested by Wake *et al.* (2015) for pre-schoolers with language impairment, much of the change ascribed to intervention may actually be due to natural improvement. For efficacy studies, this shows the importance of multiple base line designs, to track whether any improvement shown is more than that occurring naturally in the child.

6.5 What does and does not predict outcome?

Given the undersupply of SLT services reported in the UK and Ireland, as well as other countries, it is important that services be directed to where they can be of most effect. This section discusses the utility of measures used to decide prognosis.

6.5.1 Risk factors

It is important to separate risk factors for SSD in early childhood from factors that predict persistence of SSD: it is the latter that is important to clinicians who are making decisions on service provision for a clinical caseload. By definition, the

children in the caseload are already showing delayed speech development. There is a preponderance of evidence that boys are more likely to be identified with SSD than girls, but similar to findings from Dale *et al.* (2003) and Armstrong *et al.* (2017) for children with early language delay, there is no evidence that boys are significantly more likely to have persisting SSD. Evidence that older children with SSD are more likely to be male is to be expected if more males are identified as having SSD initially, and rates of resolution are the same for both genders. When considering factors such as family history as risks, it is also important to take into account how many children display the risk factor but do not have an impairment: with the case of family history, some population studies show that as many unaffected children with TD speech have this risk factor, as children in clinical samples where all children have SSD, meaning it is of small use in identifying children at risk of persistent SSD: for example Campbell *et al.* (2003) showed 25% of children with TD speech had a positive family history of SSD, compared to only 22% of children with SSD in Broomfield and Dodd's (2005b) sample. Evidence regarding SES as a risk factor for SSD is equivocal; associations seem to be seen in population but not clinical samples. Again we need to differentiate whether low SES is just a risk factor for SSD at all, or whether it is a risk factor for persistent SSD. Wren *et al.* (2016) showed that children with persisting SSD at age 8 came from families less likely to be home owners, but the odd ratio was less than 2, i.e. unlikely to be clinically significant (Sackett *et al.* 1991), and it is not known whether children with transient SSD at younger ages also had parents less likely to be home owners. None of the factors of speech and language delay, delayed motor development, birth difficulties, incidence of otitis media or family history were predictors of persistent SSD in this study. Therefore Shriberg's statement in 1994, that risk factors cannot be "used to indicate the probable need for speech services" p 1129, remains accurate, and demographic and early developmental variables do not seem to be useful as guides to improve early identification practices (Armstrong *et al.* 2017).

In fact it may be that SSD below the age of 4 is better considered a risk factor for SSD rather than a disorder per se, as suggested by Whitehurst and Fischel (1994) for early language delay, since the majority of children both in this study and most of the literature considered in section 1.5.3 will go on to have speech accuracy in the normal range at age 6.

6.5.2 Delay or disorder: Error types and severity

Dodd's classification of delay and disorder is used by many clinicians to decide on priorities for service (Joffe and Pring 2008), where children who make atypical errors are considered to be less likely to resolve without therapy. The study reported in this thesis is one of the first studies to track the development of children classified by Dodd's system of differential diagnosis longitudinally up to the age of normalisation, and it showed that children from Dodd *et al*.'s disordered categories (Dodd *et al.* 2005) did not plateau in development or get worse without therapy, (with the exception of one case), as predicted by Broomfield and Dodd (2005a), but in fact made greater improvements in PCC than children with delay, and either changed presentation to fit Dodd's category of PD, or showed only transitory SSD. Thus, as discussed in section 5.8.4, it appears that inconsistency and use of atypical errors is indicative of a more severe delay rather than a qualitatively different disorder, so that these surface symptoms need only raise concern if they persist or present in the eldest pre-schoolers, where they might indicate that development has plateaued, or delay is still very severe. Findings from two children in this study support this interpretation. They were the only 2 children over the age of 4;6 to present at Initial Assessment with a diagnosis of disorder (CPD), both aged 5;2. One (12505) received over 30 sessions of clinician-directed therapy, which revealed significant difficulty with motor gestures at the alveolar ridge, (although she demonstrated none of the inconsistency, prosodic or sequencing difficulties associated with CAS), so it appeared that her atypical substitution errors (backing and substitution of /l/ with /n/), had a motoric as well as a phonological component. At Re-assessment she was still being treated for lateralisation of alveolar fricatives, but showed no phonological errors. She was achieving well academically. The other child (12501) remained with a diagnosis of CPD at Re-assessment. She had received only reviews and home programmes, and it appeared that her development had plateaued as she made very little change in presentation by Re-assessment. She also had co-morbid language impairment and difficulties with reading.

Following the children with SSD longitudinally showed that use of atypical errors reduced with age, and it was not a significant predictor of persisting SSD in this sample. Although a somewhat smaller proportion of children with diagnoses of disorder resolved their SSD than those with PD, they were significantly younger than

children with PD, and making on average more progress in PCC, so it might be expected that they would still resolve within the period of normalisation. Therefore the classification of delay or disorder does not appear to be useful to predict prognosis, at least in this sample. Neither can severity, by whatever classification, be used as a good predictor. While all of the children in Shriberg's mild category resolved, so did 77% of the severe category, compared to 82% of the mild to moderate category. Rate of progress therefore appears to be a better measure of prognosis than severity level or error type, with a slow rate or cessation in progress indicating cause for concern.

There was however one type of case that could be identified with complete sensitivity and specificity at an early age: in this sample none of the children who showed a lateralised distortion of fricatives recovered from this. This study provides no evidence for how early intervention for this is effective, but intervention appears to be necessary to correct this.

6.5.3 Non-word discrimination

Similar to results of the study by Rvachew (2006), where speech perception skills in pre-schoolers explained growth in speech accuracy after 12 months, in the study reported here, a multivariable logistic regression showed that score on a test of Non-Word Discrimination at Initial Assessment was the only unique predictor of resolution of SSD two years later. This result was also useful at the level of the individual child, in that all children who scored above the median resolved their phonological deficits. However, it had poor sensitivity, as a large number of children who scored below the median also resolved their SSD. Nevertheless, there was a significant medium correlation between NWD scores and concurrent PCC at Initial Assessment, and a significant medium, bordering on strong correlation with PCC measured at Re-assessment in the sample diagnosed with SSD at initial assessment. This suggests that speech production accuracy is dependent on the perceptual and cognitive abilities needed for auditory discrimination, so that if these are well developed, even in children with SSD, speech accuracy will improve. However, perceptual and categorisation abilities are still developing during pre-school and school age (Hazan and Barrett 2000), so those who have poor perceptual/discrimination abilities at an early age may still show development of these, which allows speech accuracy also to progress. It may be then, that for those whose auditory discrimination abilities do not develop as

quickly as their peers' do, speech accuracy will remain poor. Unfortunately, NWD skills were not measured again at Re-assessment, so this could not be established for this sample. However, there is evidence to support this hypothesis from other studies, as reported in section 1.3.2. Stackhouse *et al.* (2007) found that children whose SSD persisted, consistently performed less well on measures of auditory discrimination than those who resolved, at all assessment times: ages 4;6, 5;8 and 6;8. Gósy and Horváth (2014) found NWD abilities in 7 to 8 year-old children with SSD were significantly poorer than in age-matched controls, with no overlap between groups, and Kenney *et al*. (2006) found all participants in their sample of adults with persisting SSD had deficits in discrimination compared to controls.

These findings then suggest that good perception and discrimination skills are necessary but not sufficient for good speech accuracy. It may be that poor discrimination skills lead to phonological representations inadequately well specified to form a basis for accurate motor programmes for production. Conversely, findings from the single case study described by Wright *et al*. (2017) that Liam could not discriminate /s/ and /ʃ/ after many years of substituting both with the same production as a voiceless nasal, shows how possible deficits downstream at the output level might influence cognitive/linguistic representations upstream, supporting the reciprocal development of motor and perception abilities suggested by Velleman and Vihman (2002) and Byun (2012).

It may be significant that auditory discrimination in this study was tested using non-words rather than real words. Findings from Rvachew and Jamieson (1989) showed that children can sometimes use cues from coarticulation to discriminate minimal word pairs, despite not being able to discriminate phonemes at sound level. Non-word discrimination eliminates use of such strategies and may indicate a general deficit not identifiable with real words. Conversely, when planning for therapy, it is important that discrimination is tested for the sounds that the child is having difficulty with, to establish whether the locus of difficulty is at the input and output level, or at the output level only. Findings from the case study reported by Wright *et al*. (2017) showed that, as indicated above, children may be able to discriminate words using strategies that do not implicate a complete knowledge of the salient acoustic features of a phoneme needed to discriminate it at sound level. The development of an Irish version of an application such as the SAILS (Speech Assessment and Interactive Learning System)

tool for speech perception assessment and training (Innovations 1995), would be useful here.

6.5.4 The effect of age of assessment

A problem in the identification of children with persisting SSD is the difficulty of testing abilities in children under the age of 4. Performance at chance in some tasks is reported by several studies (Brosseau-Lapré and Rvachew 2016), and the report by Bird *et al.* (1995 p. 453), on attempting to establish rhyming abilities, that "the impression of the experimenter was that many of these children had no idea of what was meant by the concept of rhyme, and they were unable to work out what was required, despite being given examples with practice and feedback" also described the experience in this study. It may be unwise to ascribe a deficit to a skill that has not yet emerged, just as a 12 month-old child who is not yet walking cannot be described to have a walking deficit, even though some younger children are already walking.

Thus the large amount of variability in the development of young children may make it difficult to identify prognostic factors in young pre-schoolers. This is exemplified by the finding that performance on a test of DDK rate differentiated, at Re-assessment, children who had persisting SSD from those who had resolved, but did not show any difference two years earlier at Initial Assessment. Vick *et al.* (2014)'s findings that a sub-group of pre-schoolers with SSD who showed poor motor control could be identified using instrumental measures, requires a follow-up to see if this is prognostic of persisting SSD, and their finding that such children are poor at imitating words with iambic stress holds out hope that a non-instrumental measure might be developed to identify them.

6.5.5 Identification of children who will go on to resolve their SSD

From the findings reported here, it seems that subtyping children with SSD may not be useful in indicating prognosis or suggesting novel therapies. Neither sub-types derived from supposed aetiology nor on observed symptomatology, informed likely underlying deficits or prognosis, in the latter case because children's symptoms changed with age and development. However, symptoms are still useful for revealing the developmental stage children have reached and for indicating effective therapies, i.e. Core Vocabulary therapy for children who show high inconsistency (Dodd *et al.* 2006), and

phonological contrast therapy (Dodd and Bradford 2000) for children with structural and substitution errors which appear to be due to difficulties with phonological organisation. But we must avoid the danger of sub-typing possibly leading to blanket 'one size fits this group' approaches which may not address individual profiles of strengths and weaknesses. For instance, output based phonological contrast therapy will not work for the child who has difficulties in discriminating correct from error productions.

It might be expected that upstream errors will have a large effect on speech output, and it was interesting that no child with good auditory discrimination of non-words had a measurement of DDK below normal limits at either Initial or Reassessment (although very few had a poor measure of DDK at either time). Although it seems unlikely that discrimination abilities could drive motor development, this finding allows the possibility, as suggested by Velleman and Vihman (2002) and Byun (2012) that development of motor and perception abilities is a reciprocal process.

This study did not find a certain indicator of persisting SSD in children, but found that most children would resolve their SSD, with good auditory discrimination abilities predicting certain resolution. Results of the cluster analysis also suggested that children with poor performance over multiple factors, including perceptual/cognitive, language and motor skills, would be less likely to resolve, but with the caveat that at this point the sample size was very small, so no generalisations could be made.

6.6 Clinical implications

This study sought to provide evidence useful to SLTs to make decisions about service delivery in their everyday practice, using materials that are widely available to them, and participants similar to those who make up the caseloads of SLT services. The findings of this study and supporting evidence suggest considerations for decisions about likely prognosis, prioritisation of services and assessment that will be discussed in the following sub-sections.

The benefit of carrying out research with a sample that reflects a real caseload is that results can be applied to real life clinical practice. Much research is carried out with children who have moderate to severe deficits, whereas in this clinical sample and those cited in section 3.3, the majority of children had only mild to moderate

difficulties. This, in combination with the long period of the study, showed the proportion of children who resolved their SSD, i.e. 79%. was higher than frequently reported. As the sample used here reflects clinical rather than research samples, this result is more pertinent for decision making in clinical practice. The large number of children who resolved, and the predictors, or non-predictors of that, have implications for assessment and service delivery to this population.

6.6.1 Assessment

As neither of the sub-groups suggested by Shriberg *et al*. (2010) nor Dodd *et al.* (2005) proved useful in indicating likely prognosis or locus of difficulty, it appears that the best way currently available to deal with the heterogeneity shown by children with SSD may be to use psycholinguistic assessment, as described by Stackhouse and Wells (1997), to identify specific deficits in the individual child. In the case study reported by Wright et al. (2017), Liam's case illustrates an SSD that involved difficulties at multiple levels of representation: it was initially thought to be a problem on the output side of speech processing, but further investigation showed deficits in perception and in phonological representations. Coupled with findings that good non-word discrimination predicted certain resolution of SSD in this study, this illustrates the importance of assessment of perceptual abilities.

Surveys of clinicians' practice for children with SSD carried out in the UK, (Joffe and Pring 2008), USA, (Brumbaugh and Smit 2013, Skahan et al. 2007), the Netherlands (Priester et al. 2009) and Australia (Mcleod and Baker 2014) show that targeting auditory discrimination is a frequently used method of intervention, with McLeod and Baker's study showing that 33.5% of clinicians used it always and 55.5% used it sometimes, and only 3.5% never used it. However, in the same study, far fewer reported they always assessed perception and auditory discrimination skills. The fact that such skills are targeted in therapy more often than they are assessed adds strength to Joffe and Pring's (2008) conclusion that work on auditory discrimination may be unwarranted for some children, who do not have deficits in this area.

The finding that auditory discrimination ability was the only significant predictor of likelihood to resolve SSD in this study strengthens the case for assessment of speech perception/discrimination abilities to be an essential item in the assessment battery for SSD. The fact that the measure lacked sensitivity is compensated for by the fact that it

appeared, unlike measures that indicate group trends only, to work at the level of the individual child, with good discrimination skills ensuring a good outcome for phonologically based SSD. Further research is of course needed to replicate this result in a bigger sample, and to develop assessment tools that have good validity and reliability, and are widely available. It may be that the lack of a commercial assessment tool suitable for use with pre-schoolers partially explains failure to assess these skills. However Stackhouse and colleagues include tasks assessing Non-Word Discrimination in their Compendium (Stackhouse et al. 2007), which is widely available and easy to use. Use of such assessments may at least indicate to SLTs which children are likely to resolve without treatment, and are therefore not a priority for treatment, as well as identify those for whom intervention for auditory discrimination is necessary, and those for whom it is not warranted.

6.6.2 Intervention: when is early too early?

Given the pressure on SLT services described in section 1.1, SLTs cannot treat every child with an SSD and so must prioritise children for intervention. As in the study by Roulstone *et al.* (2008), SLTs in this study appeared to be selecting children for therapy who seemed least likely to resolve. Despite this, in the study reported here, 50% of young pre-schoolers in the mild category using Dodd's classification, and 55% using Shriberg's classification, received some form of therapy, and similar practice is reported in the literature: Emanuel *et al.* (2007) reported that 76.5% of two to three year-olds with mild speech or language difficulties were offered therapy in a community paediatric UK SLT service. Given the evidence reported here that the majority of children will resolve their SSD before they are seven years old, even without therapy, this may not be the best use of scarce resources. While children who have severe deficits may also resolve them without therapy, significant, even if temporary, impairment of the ability to communicate justifies intervention to accelerate the process of development. However, for children with only a mild impairment, this is not likely to be true for the majority of cases. Botting et al. (2016) has shown that children with language impairment who are nevertheless improving without therapy, do not benefit from therapy, but continue to improve at the same rate with or without therapy. If this is also true for children with SSD, the best use of resources with pre-schoolers might be a "watch and measure" approach (rather than merely "watch and wait") so that children who seem not to be "catching up" in speech

development can be offered intervention at that point. Such a monitoring approach has been recommended by Reilly *et al*. (2014) for pre-schoolers with language impairment. Like children with SSD, many of these children show only a transient impairment, so Reilly et al. recommend that a child should only be considered a case when they show language impairment at two time points, 12 months apart. Given the shorter course of development in speech than in language, a shorter time period might be more appropriate, but it would seem an efficient use of resources to establish whether natural improvement is occurring before offering therapy. This does not mean that pre-school children should be denied intervention altogether in the short term: many childhood illnesses are self-limiting, but we still intervene to alleviate the symptoms. It may be that measures showing that a child's speech development is failing to progress, or a quality of life measure that indicates that a child's SSD is affecting their participation and activities of daily life, are better indicators of children who need intervention, rather than number or type of errors at this age. In support of this, Roulstone et al. (2003) found that the only predictor of outcome in their sample of young pre-schoolers was the disability scale of the Therapy Outcome Measures (TOMs), (Enderby and John 1997), suggesting that SLTs are well able to make such judgements.

Thus, early intervention must be construed as intervention that takes place as soon as the need for it is established, not just intervention at an early age. Identification of a true SSD may be difficult in younger children. Evidence from a number of studies shows that the peak age of decline of use of phonological processes is between the ages of 2 and 4 (Roberts et al. 1990, Haelsig and Madison 1986, Khan and Lewis 1986), with James (2001) showing the decline in the use of atypical processes as greatest between 3 and 4, so frequency and types and errors at ages under 4 may not be a reliable indicator of SSD. Sosa (2015) found that 3 year-olds with TD speech showed rates of inconsistency high enough to be classified as having IPD in Dodd's classification, and concluded that such intra-word variability is not a marker of SSD at that age. McLeod and Hewett (2008) drew the same conclusion, finding inconsistency rates between 42% and 79% in sixteen 2 to 3 year-olds' production of words containing consonant clusters. However, it should be noted that findings regarding consistency in the speech of TD children vary markedly: Stackhouse *et al.* (2007) found a mean inconsistency rate in 7 TD children aged between 3;7 and 4;0 of 26%, with the highest rate at 39%, just below Dodd's threshold for IPD, whereas Holm *et al.* (2007) found it to be as low as 13% in 20 children under 3;6.

Nevertheless, it may be that a large amount of variability, and use of both typical and atypical errors, is not predictive of case status in children under the age of 4, and that reliable identification of case status can only be made after this age. Stackhouse *et al.* (2007), from findings from their 3 year longitudinal study of children with SSD at a mean age of 4;6, also suggest that 4 years is the age at which children likely to have persisting difficulties can be identified. Roulstone *et al.'s* (2009) study showing that persisting SSD at age 8 was associated with proportion of errors at age 5, but not at age 2, also supports this contention. Even at age 4, there is evidence that children continue to improve productions of more complex words, with Masso et *al.'s* (2017) study of 80 children with SSD showing an increase in PCC in polysyllabic words of 8 points in a period as short as 18 weeks, without intervention. Half the children were classed as non-improvers by the end of the study, meaning the growth in PCC of the improvers must have been even greater. This finding also holds promise that a "watch and measure" strategy could identify for intervention those children who are not continuing to catch up without help.

It is also necessary to take into account the appropriate age for intervention. Although early intervention is recommended as it may be easier to shape a developing system (Broomfield and Dodd 2005a), if speech development is dependent on abilities that are developing slowly in children with SSD, intervention offered too soon could be counterproductive. Priester *et al.* (2009) cite cases where intervention undertaken before a child's motor control is sufficiently developed for it, has had a negative rather than a positive effect on the child's social-emotional development. Broomfield and Dodd (2005a) report evidence that intervention for children showing atypical errors is most effective at age 4, for delayed speech at age 5, and for articulation disorders at age 7, which ages are in accord with identification of case status not taking place before age 4.

The importance of avoiding difficulties with literacy is also a large concern for SLTs. Age 4 also appears to be an appropriate age to measure and start intervention for PA. Evidence from this study and others such as Bird and Bishop (1992), as well as norms from the PIPA (Dodd *et al.* 2000), (which show a raw score of zero on some sub-tests indicates normal development for children younger than 4), suggests that many children do not have sufficient meta-phonological skills for effective intervention for PA before this age.

6.6.3 Implications for service delivery

The findings of this study, and the supporting evidence presented, show that focusing SLT services on early intervention for children with SSD may not be the best use of severely limited resources. The majority of children referred for SSD are aged under four, an age when, for the majority, it is only a transitory delay, and most children who present have only mild involvement, which does not cause any significant current problems. Nevertheless, there is evidence that over half of these are receiving therapy, which for the majority who are naturally improving, is possibly not going to have any effect (Botting *et al.* 2016). Given that children with SSD make up the largest part of caseloads for paediatric community SLTs (Joffe and Pring 2008), moving to a monitoring approach for these under fours would allow a large amount of resource to be diverted to children who need specialised, individualised treatment for clearly identified persisting speech and language needs, while allowing for rapid intervention if monitored children were found to have arrested in speech development, or to be suffering negative social and personal consequences.

Table 6-1 provides a summary of findings and recommendations regarding clinical practice.

Table 6-1: Summary of findings and recommendations regarding clinical practice

Finding	Recommendation
79% of children aged 3 to 511 in this sample resolved their SSD within 2 years. Many did not receive therapy. Over half the children in this sample had only mild deficits, but over half of these received therapy.	SLTs should consider offering therapy to children under 4 only if their SSD is having social or personal consequences, or if development appears to have stalled. Monitoring of progress may be more appropriate for most children of this age than routinely offering therapy.
Pre-schoolers in this sample who showed good performance on a test of non-word discrimination all resolved their SSD. Mean scores of children who did not resolve their SSD on this test was 7, whereas the mean score of those who resolved was 11. All children with a score over 11 resolved.	Auditory discrimination of non-words should be an essential part of assessment of children referred for SSD.
Continued on next page	

Table 6-1: Summary of findings and recommendations regarding clinical practice cont.

Finding	Recommendation
Severity of involvement is not a good indicator of SSD that will persist. 77% of children under 4;6 who presented with severe SSD according to Shriberg's classification, resolved their SSD within two years, some without therapy. No child over 4;6 presented with a severe SSD in this study.	Decisions to offer intervention for young pre-schoolers may be better based on a quality of life measure, or signs that natural development has stalled, than a measure of PCC.
Inconsistency and atypical errors in children under 4;6 do not appear to indicate that SSD will persist.	Inconsistency and atypical errors are a cause for concern only when they persist past the age of 4;6.
Almost all children made significant increases in PCC over 2 years. Children with disordered speech and severe SSD increased PCC more than children with delay over two years.	Rate of progress is a better measure of likelihood to resolve than severity level. Children who are not showing natural improvement in speech development are a cause for concern. Improvement ascribed to therapy effect may in fact be natural improvement. A multiple baseline measure before therapy begins would allow clinicians to control for this.
Lateralisation of fricatives did not resolve without help, but all children who had other non-stimulable sounds at Initial Assessment acquired them by Re-assessment, with the exception of /ɹ/ as this was still within normal limits for acquisition in the time period of this study.	Research is needed to establish the optimal age for therapy in cases of lateralisation of fricatives. Therapy for lateralisation should always be offered, as soon as ability to engage in therapy is shown. Lack of stimulability for a sound is not an indicator for concern until well after the normal acquisition of such sound.
Many children acquired a dentalised production of alveolar fricatives in some contexts, with many presenting with this as their only error between the ages of 4;6 and 6;6, at which age it effectively disappeared.	A dentalised production of alveolar fricatives is not cause for concern in children under 6;6

6.7 Limitations of the study

A major limitation of this study was that measures were taken at only two time points, a relatively long interval apart, so we have no indication of how quickly children improved, or of how measures of individual abilities changed over time, which might give an indication as to what stimulated such improvement. A related factor is that only a subset of measures were repeated at both times, so it cannot be shown whether, for example, children who originally had poor auditory discrimination skills but resolved nevertheless, did so because those skills improved. Furthermore, the novel tests were not standardised, so no comparison of abilities can be made with children whose speech is typically developing. Receipt of therapy, is, of course, also an important confounding factor, but this is true in most studies of natural history, where large effects of therapy are not noted e.g. Shriberg *et al.* (1994b), Roulstone *et al.* (2009), and indications from parents suggest that like some of these studies, a very low dosage of therapy was offered to those who received it, and did not appear to account in any large measure for the degree of recovery reported (Paul 1996, Wren *et al.* 2016).

The sample size of 85 children is small for predictive analysis, especially given the fact that so few remained with persisting SSD. Although the 12% attrition rate was relatively small, the difficulties of assessment in such a young population meant that there was a large amount of missing data, further reducing the sample size for analysis. It is also possible that there was a referral bias away from children with significant language deficits, as discussed in section 3.2.5. Nevertheless, the sample is comparable with those used in much clinical research with children with SSD.

The measure used for designation of SSD status, and for measuring development of speech over time, was PCC on a standardised single word naming elicitation test. A major criticism of single-word articulation tests is that they may not accurately reflect speech sound production abilities or intelligibility in spontaneous speech, nor an evaluation of the prosodic features of speech. However, standardized single-word articulation tests provide an efficient means of obtaining a speech sample that includes all the sounds in the language across different word positions, and allow comparison to a normative sample to determine if a SSD is present, and the severity of involvement. Multiple scores obtained from the same probe allow for comparison between children and for measurement of progress over time, essential in a longitudinal study. In

addition, control of the stimuli is necessary when transcribing productions from children with much reduced intelligibility. The major justification for use as a measure in this study, however, is that they are the most common method used to evaluate speech in clinical practice (Skahan *et al.* 2007), as well as to describe inclusion/exclusion criteria, and to measure baselines and change in research. There is also evidence that children continue to have difficulties with word production that are not captured by single word tests, but which may only be captured by performance on batteries of polysyllabic words (James *et al.* 2008) or non-word repetition, e.g. (Stackhouse *et al.* 2007b). However, where services are already overstretched, the priority is to bring children to a level of output that presents as normal speech without overt errors. Although errors on polysyllabic words and non-words are suggested to implicate development of phonological awareness, this may be equally as well or better tested on an assessment of PA abilities.

Use of phonetic transcription to record children's speech output is also a limitation in that it forces categorical representations of continuous data, and must be coloured by the perception of the transcriber. For instance, by definition, covert contrasts are frequently not perceived by the listener. Even if a contrast is noted, the transcriber has to ascribe the signal to one phonetic category or another, and may not be able to record the difference by means of diacritic, due to the essential flaw of trying to record the characteristics of an acoustic signal by means of a description of an articulatory gesture. Similarly, it is often difficult to record coarticulatory information. However, transcription is the only method widely available to SLTs, and is probably the most feasible when working with children who produce a wide range of errors.

A factor analysis can only factor whatever is put into it. In order to be relevant to the realities of clinical practice, the assessments used were those widely available or feasible to use in the constraints of everyday practice. Therefore results of this study may be limited by the sensitivity of measures, so that these results may only show that associations were not revealed by measurements from the sorts of tasks widely used by SLTs in the UK and Ireland to assess a pre-school population of children with SSD. The inclusion of instrumental measures to measure perceptual and motor abilities, or a different selection of tasks for assessment, may have revealed other associations.

Despite these limitations, findings were relevant to clinical practice and had some usefulness for considerations regarding service delivery, and some indications can be given for future research.

6.8 Future research

A replication of this study would of course be necessary to establish how far results from this sample can be generalised to other clinical samples. To increase usefulness of the replication, repeated reassessment of all tasks at intervals of say, six months, would provide much more useful information that would allow investigation of, for example, how quickly children catch up, whether improvement is incremental or whether it proceeds in stages, and whether progress in underlying skills stimulates progress in speech accuracy. As it is unusual for a child in Irish SLT services to receive more than one block of therapy a year, such regular monitoring would also allow for observation of how far therapy influences rate of improvement, and for comparison of the effects of therapy on children who are improving naturally as opposed to those making little change.

To make such a replication feasible and not so onerous that attrition rates would rise, it would need to be carried out in collaboration with local services, such that regular review assessments would be carried out by the researchers, in place of rather than in addition to the normal service provider, and the review results shared with them to direct further management. Children who resolved within the course of the study could be discharged at that point, with a single follow up to investigate PA and reading ability after they began school. Scheduling this final assessment at a fixed age rather than after a set period would allow all children to be at the stage of beginning reading, and testing of reading ability, rather than reliance on parent report.

Experience from this study would inform the selection of tasks and configuration of the assessment battery at each point. A quality of life measure should be added to investigate predictive validity of such a measure, as well as its utility in decision making regarding need for/effect of intervention. A somewhat more extensive measure of speech motor ability, (although still one feasible for clinical use) such as the VMPAC (Hayden 1999) might provide more sensitivity, as few children scored below normal limits on the oro-motor section of the DEAP. It would also provide for better comparability with other studies. The inclusion of a test of repetition of iambic words

could also be incorporated to evaluate the clinical utility of a test suggested by Vick *et al.* (2014). The test of Number Concepts appeared to have no usefulness so would not be included in the assessment battery. Replacing the novel assessments used in this study with norm-referenced tests of auditory memory, non-word discrimination and pattern recognition would also add to comparability of results. Tests of PA abilities or pattern recognition would not be included in the battery for children under four, due to the large number of children of this age who seem to have no understanding of the tasks, so that assessment for these children would be less onerous.

A further study that logically arises from the recommendations presented here could also be carried out in collaboration with local services, to monitor the effect and feasibility of moving to a 'wait and measure' system of management of SSD in pre-schoolers, where intervention is not routinely offered to children aged under 4 unless they are suffering social or emotional consequences from their SSD, or unless parents desire it. Experience from other studies shows that many parents do not seek therapy, even when offered, but are happy to wait until children are older to see if natural improvement occurs, or for children to be more ready for therapy (Paul 1996, Roulstone *et al.* 2003, Dodd *et al.* 2017). To evaluate outcomes, the system could be implemented in one local clinic, and outcomes in terms of improvements in speech accuracy and discharge rate, as well as amount of resource required, could be compared with another clinic operating service as usual.

A third study is motivated by the theoretical implications of the usage-based theory of speech acquisition described here. Atypical errors, did not, as shown in other studies, predict persisting SSD. Theory regarding the course of acquisition as described by Stackhouse and Wells (1997) sees children passing from a stage of holistic representations to finely specified segmented representations. It would therefore seem probable that errors involving systematic substitutions of individual segments would arise at the second stage, and structural errors at an earlier stage of development. Therefore persistent use of structural errors that implicate delayed transition to the stage of systematic simplification might better predict likelihood of persisting SSD than atypical errors per se. The data from this study will be reanalysed to investigate whether the distribution of structural versus substitution errors changes with age and/or severity of SSD, and whether structural omission errors are a better predictor of persisting SSD than substitution errors, as suggested by Shriberg (1994).

7 References

Adler, N. E. and Newman, K. (2002) 'Socioeconomic disparities in health: pathways and policies', *Health affairs*, 21(2), 60-76.

Alt, M. and Plante, E. (2006) 'Factors that influence lexical and semantic fast mapping of young children with specific language impairment', *Journal of Speech, Language, and Hearing Research*, 49(5), 941-954.

Altvater-Mackensen, N. and Fikkert, P. (2010) 'The acquisition of the stop-fricative contrast in perception and production', *Lingua*, 120(8), 1898-1909.

Ambridge, B. and Lieven, E. V. (2011) *Child language acquisition: Contrasting theoretical approaches*, Cambridge: Cambridge University Press.

American Speech-Language-Hearing Association (2008) *2008 Schools Survey report: Caseload characteristics.*, Rockville, MD: ASHA.

Anthony, J. L., Aghara, R. G., Dunkelberger, M. J., Anthony, T. I., Williams, J. M. and Zhang, Z. (2011) 'What Factors Place Children With Speech Sound Disorders at Risk for Reading Problems?', *American Journal of Speech-Language Pathology*, 20(2), 146.

Anthony, J. L., Williams, J. M., Aghara, R. G., Dunkelberger, M., Novak, B. and Mukherjee, A. D. (2010) 'Assessment of individual differences in phonological representation', *Reading and Writing*, 23(8), 969-994.

Armstrong, R., Scott, J. G., Whitehouse, A. J., Copland, D. A., Mcmahon, K. L. and Arnott, W. (2017) 'Late talkers and later language outcomes: Predicting the different language trajectories', *International Journal of Speech-Language Pathology*, 19(3), 237-250.

Armstrong, S. and Ainley, M. (1992) *South Tyneside assessment of phonology*, St Mabyn: STASS Publications.

ASHA (2007) *Childhood apraxia of speech. Technical report.* Rockville, MD: ASHA

Auinger, P., Lanphear, B. P., Kalkwarf, H. J. and Mansour, M. E. (2003) 'Trends in Otitis Media Among Children in the United States', *Pediatrics*, 112(3), 514-520.

Bacon, P. and Associates (2001) *Current and future supply and demand conditions in the labour market for certain professional therapists*, Dublin, Ireland: Department of Health and Children.

Baker, E. (2006) 'Management of speech impairment in children: The journey so far and the road ahead', *Advances in Speech Language Pathology*, 8(3), 156-163.

Ball, M.J. (2016). Principles of clinical phonology: theoretical approaches. Abingdon , OX: Psychology Press.

Ball, M.J. and Müller, N. (2002). 'The use of the terms phonetics and phonology in the description of disordered speech'. *Advances in Speech Language Pathology*, 4(2), 95-108.

Ball, M., Manuel, R. and Muller, N. (2004) 'An atypical articulatory setting as learned behaviour: a videofluorographic study', *Child Language Teaching and Therapy*, 20(2), 153-162.

Ball, M. J., Müller, N. and Munro, S. (2001) 'The acquisition of the lateral fricative in Welsh-English bilinguals'. *Multilingua*, 20, 269-284.

Barre, N., Morgan, A., Doyle, L. W. and Anderson, P. J. (2011) 'Language Abilities in Children Who Were Very Preterm and/or Very Low Birth Weight: A Meta-Analysis', *The Journal of Pediatrics*, 158(5), 766-774.e1.

Baum, S. R. and McNutt, J. C. (1990) 'An acoustic analysis of frontal misarticulation of /s/ in children', *Journal of phonetics*, 18(1), 51-63.

Beery, K. E. and Beery, N. A. (2010) *The Beery-Buktenica developmental test of visual-motor integration (6th ed.).*, Bloomington: Pearson.

Beitchman, J. H., Wilson, B., Brownlie, E., Walters, H. and Lancee, W. (1996) 'Long-term consistency in speech/language profiles: I. Developmental and academic outcomes', *Journal of the American Academy of Child and Adolescent Psychiatry*, 35(6), 804-814.

Benasich, A. A. and Brooks-Gunn, J. (1996) 'Maternal attitudes and knowledge of child-rearing: Associations with family and child outcomes', *Child development*, 67(3), 1186-1205.

Bernhardt, B. and Major, E. (2005) 'Speech, language and literacy skills 3 years later: a follow-up study of early phonological and metaphonological intervention', *International Journal of Language and Communication Disorders*, 40(1), 1-27.

Bernhardt, B.M., Másdóttir, T., Stemberger, J.P., Leonhardt, L. and Hansson, G.Ó., 2015a. Fricative acquisition in English-and Icelandic-speaking preschoolers with protracted phonological development. *Clinical linguistics and phonetics, 29(8-10), 642-665.*

Bernhardt, B.M. and Stemberger, J.P., 2015b, Same challenges, diverse solutions: Outcomes of a crosslinguistic project in phonological development. In Babatsouli & Ingram (eds.) *Proceedings of the International Symposium on Monolingual and Bilingual Speec*h, 47-56.

Bernhardt, B. and Stemberger, J. P. (2007) 'Phonological impairment in children and adults' in De Lacy, P., ed., *The Cambridge handbook of phonology*, Cambridge: Cambridge University Press, 575-595.

Bernthal, N., Bankson, J. and Flipson, P. (2009) *'Articulation and Phonological Disorders: Speech Sound Disorders in Children'*. 7th Edition. Boston: Allyn and Bacon.

Bhargava, A. and Kirova, A. (2006) 'Assessing the development of mathematical concepts in preschool children: checklists for teachers', *Journal of Teaching and Learning*, 2(1).

Bird, J. and Bishop, D. (1992) 'Perception and awareness of phonemes in phonologically impaired children', *International Journal of Language and Communication Disorders*, 27(4), 289-311.

Bird, J., Bishop, D. and Freeman, N. (1995) 'Phonological awareness and literacy development in children with expressive phonological impairments', *Journal of Speech, Language, and Hearing Research*, 38(2), 446-462.

Bishop, D. and Edmundson, A. (1987) 'Language-Impaired 4-Year-OldsDistinguishing Transient from Persistent Impairment', *Journal of Speech and Hearing Disorders*, 52(2), 156-173.

Bishop, D. V. (2002) 'Motor immaturity and specific speech and language impairment: Evidence for a common genetic basis', *American Journal of Medical Genetics*, 114(1), 56-63.

Bishop, D. V. M. (2000) 'How does the brain learn language? Insights from the study of children with and without language impairment', *Developmental Medicine & Child Neurology*, 42(2), 133-142.

Bishop, D. V. M. and Adams, C. (1990) 'A Prospective Study of the Relationship between Specific Language Impairment, Phonological Disorders and Reading Retardation', *Journal of Child Psychology and Psychiatry*, 31(7), 1027-1050.

Bishop, D. V. M. and Edmundson, A. (1986) 'Is otitis media a major cause of specific developmental language disorders?', *International Journal of Language and Communication Disorders*, 21(3), 321-338.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T. and Catalise-consortium (2017) 'Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology', *Journal of Child Psychology and Psychiatry*, early online.

Boada, R. and Pennington, B. F. (2006) 'Deficient implicit phonological representations in children with dyslexia', *Journal of Experimental Child Psychology*, 95(3), 153-193.

Bornstein, M. H., Hahn, C.-S., Suwalsky, J. T. and Haynes, O. M. (2003) 'The Hollingshead four-factor index of social status and the socioeconomic index of occupations', In Bornstein, M. and Bradley, H., (eds) *Socioeconomic status, parenting, and child development*, 29-81., Mahwah, NJ: Lawrence Erlbaum

Botting, N., Gaynor, M., Tucker, K. and Orchard-Lisle, G. (2016) 'The importance of natural change in planning school-based intervention for children with Developmental Language Impairment (DLI)', *Child Language Teaching and Therapy*, 32(2), 159-177.

Bowen, C. (2009) *Children's speech sound disorders*, 1st ed. Chichester:John Wiley and Sons.

Bowen, C. (2014) *Children's speech sound disorders*, 2nd ed. Chichester:John Wiley and Sons.

Bowen, C. (2015) *When /s/ and /z/ are hard to say.* [online], available: http://www.speech-language-therapy.com/index.php?option=com_content&view=article&id=86:lisp [accessed 27/06/16].

Bradford, A. and Dodd, B. (1996) 'Do all speech-disordered children have motor deficits', *Clinical Linguistics and Phonetics*, 10(2), 77-101.

Bradley, R. H. and Corwyn, R. F. (2002) 'Socioeconomic status and child development', *Annual Review of Psychology*, 53(1), 371-399.

Braveman, P. A., Cubbin, C., Egerter, S. and et al. (2005) 'Socioeconomic status in health research: One size does not fit all', *Journal of the American Medical Association,* 294(22), 2879-2888.

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

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

Broomfield, J. and Dodd, B. (2005a) 'Clinical effectiveness' in *in Differential diagnosis and treatment of children with speech disorder*, 2nd ed., London: Whurr, 211-229.

Broomfield, J. and Dodd, B. (2005b) 'Epidemiology of speech disorders', *Differential Diagnosis and Treatment of Children with Speech Disorder*, 2nd ed., London: Whurr, 83-99.

Broomfield, J. and Dodd, B. (2011) 'Is speech and language therapy effective for children with primary speech and language impairment? Report of a randomized control trial', *International Journal of Language and Communication Disorders*, 46(6), 628-640.

Brosseau-Lapré, F. and Rvachew, S. (2016) 'Underlying manifestations of developmental phonological disorders in French-speaking pre-schoolers', *Journal of Child Language*, 11/17, 1-25.

Burns, R. and Burns, R. (2008) 'Cluster analysis', *Business research methods and statistics using SPSS*, 552-567. London: Sage

Burt, L., Holm, A. and Dodd, B. (1999) 'Phonological awareness skills of 4-year-old British children: an assessment and developmental data', *International Journal of Language and Communication Disorders*, 34(3), 311-335.

Byers Brown, B., Bendersky, M. and Chapman, T. (1986) 'The early utterances of preterm infants', *British Journal of Disorders of Communication*, 21(3), 307-319.

Byun, T. M. (2011) 'A gestural account of a child-specific neutralisation in strong position', *Phonology*, 28(3), 371-412.

Byun, T. M. (2012) 'Bidirectional perception-production relations in phonological development: evidence from positional neutralization', *Clinical Linguistics and Phonetics*, 26(5), 397-413.

Byun, T. M. A. and Tessier, A. M. (2016) 'Motor influences on grammar in an emergentist model of phonology', *Language and Linguistics Compass*, 10(9), 431-452.

Campbell, T. F., Dollaghan, C. A., Rockette, H. E., Paradise, J. L., Feldman, H. M., Shriberg, L. D., Sabo, D. L. and Kurs-Lasky, M. (2003) 'Risk Factors for Speech Delay of Unknown Origin in 3-Year-Old Children', *Child development*, 74(2), 346-357.

Carlson, S. M. (2005) 'Developmentally Sensitive Measures of Executive Function in Preschool Children', *Developmental Neuropsychology*, 28(2), 595-616.

Carroll, J. (2008) *The nature of phonological deficits in children with speech and literacy difficulties: Full Research Report. ESRC End of Award Report, RES-062-23-0195*, Swindon: ESRC.

Carroll, J. M. and Snowling, M. J. (2004) 'Language and phonological skills in children at high risk of reading difficulties', *Journal of Child Psychology and Psychiatry*, 45(3), 631-640.

Carroll, J. M., Snowling, M. J., Stevenson, J. and Hulme, C. (2003) 'The development of phonological awareness in preschool children', *Developmental psychology*, 39(5), 913-923.

Casby, M. W. (2001) 'Otitis media and language development: a meta-analysis', *American Journal of Speech-Language Pathology*, 10(1), 65-80.

Casto, G. and Mastropieri, M. A. (1986) 'The efficacy of early intervention programs: A meta-analysis', *Exceptional Children*, 52(5), 417-424.

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

Chávez-Peón, M.E., Bernhardt, B.M., Adler-Bock, M., Ávila, C., Carballo, G., Fresneda, D., Lleó, C., Mendoza, E., Perez, D. and Stemberger, J.P., 2012. 'A Spanish pilot investigation for a crosslinguistic study in protracted phonological development.' *Clinical linguistics and phonetics*, 26(3), 255-272.

Choudhury, N. and Benasich, A. A. (2003) 'A Family Aggregation StudyThe Influence of Family History and Other Risk Factors on Language Development', *Journal of Speech, Language, and Hearing Research*, 46(2), 261-272.

Cleland, J., Scobbie, J. M., Heyde, C., Roxburgh, Z. and Wrench, A. A. (2016) 'Covert contrast and covert errors in persistent velar fronting', *Clinical Linguistics and Phonetics*, 31(1), 35-55.

Crosbie, S., Holm, A. and Dodd, B. (2005) 'Intervention for children with severe speech disorder: A comparison of two approaches', *International Journal of Language and Communication Disorders*, 40(4), 467-491.

Crosbie, S., Holm, A. and Dodd, B. (2009a) 'Cognitive flexibility in children with and without speech disorder', *Child Language Teaching and Therapy*, 25(2), 250-270.

Crowe Hall, B. J. (1991) 'Attitudes of Fourth and Sixth Graders Toward Peers With Mild Articulation Disorders', *Language, Speech and Hearing Services in Schools*, 22(1), 334-340.

Dale, P. S., Price, T. S., Bishop, D. V. and Plomin, R. (2003) 'Outcomes of early language DelayI. Predicting persistent and transient language difficulties at 3 and 4 Years', *Journal of Speech, Language, and Hearing Research*, 46(3), 544-560.

Davis, B. (2005) 'Clinical diagnosis of developmental speech disorders' in Kamhi, A. and Pollock, K., eds., *Phonological disorders in children*, Baltimore: Brookes, 3-21.

de Boysson-Bardies, B. and Vihman, M. M. (1991) 'Adaptation to language: Evidence from babbling and first words in four languages', *Language*, 67(2), 297-319.

de Hoog, B. E., Langereis, M. C., van Weerdenburg, M., Knoors, H. E. T. and Verhoeven, L. (2016) 'Linguistic profiles of children with CI as compared with children with hearing or specific language impairment', *International Journal of Language and Communication Disorders*, n/a-n/a.

Delgado, C. E., Vagi, S. J. and Scott, K. G. (2005) 'Early risk factors for speech and language impairments', *Exceptionality*, 13(3), 173-191.

Department of Education and Science (2007) *Se Si: gender in Irish Education*, Dublin: Department of Education and Science [Online]. Available at: https://www.education.ie/en/Publications/Statistics/Se-Si-Gender-in-Irish-Education-Introduction-to-Chapter-9.pdf [accessed 14/04/2017].

DES (2015) *Retention Rates of Pupils in Second Level Schools: 2008 Cohort*, Dublin: Department of Education and Science [Online]. Available at: Skillshttp://www.education.ie/en/Publications/Statistics/Statistical-Reports/Retention-Rates-of-Pupils-in-Second-Level-Schools-2008-Cohort.pdf [accessed 14/04/2017].

Dockrell, J. E., Lindsay, G., Letchford, B. and Mackie, C. (2006) 'Educational provision for children with specific speech and language difficulties: perspectives of speech and language therapy service managers', *International Journal of Language and Communication Disorders*, 41(4), 423-440.

Dockrell, J. E. and Marshall, C. R. (2015) 'Measurement issues: Assessing language skills in young children', *Child and Adolescent Mental Health*, 20(2), 116-125.

Dodd, B. (1995a) *Differential diagnosis and treatment of children with speech disorder*, London: Whurr.

Dodd, B. (1995b) 'Procedures for classification of sub-groups of speech disorder'. In Dodd, B., ed., *Differential diagnosis and treatment of children with speech disorder*, London: Whurr.

Dodd, B. (2005a) Differential diagnosis and treatment of children with speech disorder, 2nd Ed. London: Whurr.

Dodd, B. (2005b) 'Children with speech disorder: Defining the problem'. In Dodd, B., ed., *Differential diagnosis and treatment of children with speech disorder*, London: Whurr, 3-23.

Dodd, B. and Crosbie, S., (2005c). A procedure for classification of speech disorders. In Dodd, B. (ed.), *Differential diagnosis and treatment of children with speech disorder*, London: Whurr, 140-162.

Dodd, B. (2011) 'Differentiating Speech Delay From Disorder: Does it Matter?', *Topics in Language Disorders*, 31(2), 96-111

Dodd, B. (2014) 'Differential Diagnosis of Pediatric Speech Sound Disorder', *Current Developmental Disorders Reports*, 1(3), 189-196.

Dodd, B. and Bradford, A. (2000) 'A comparison of three therapy methods for children with different types of developmental phonological disorder', *International Journal of Language and Communication Disorders*, 35(2), 189-209.

Dodd, B. and Crosbie, S. (2005) 'A procedure for classification of speech disorders' in Dodd, B., ed., *Differential diagnosis and treatment of children with speech disorder*, London: Whurr, 3-23.

Dodd, B., Holm, A., Crosbie, S. and McCormack, P. (2005) 'Differential diagnosis of phonological disorders' in Dodd, B., ed., *Differential Diagnosis and Treatment of Children with Speech Disorder,* 2nd Edition, London: Whurr.

Dodd, B., Holm, A., Crosbie, S. and McIntosh, B. (2006) 'A core vocabulary approach for management of inconsistentspeech disorder', *International Journal of Speech-Language Pathology*, 8(3), 220-230.

Dodd, B., Holm, A., Hua, Z. and Crosbie, S. (2003) 'Phonological development: a normative study of British English-speaking children', *Clinical Linguistics and Phonetics*, 17(8), 617-643.

Dodd, B., Leahy, J. and Hambly, G. (1989) 'Phonological disorders in children: Underlying cognitive deficits', *British Journal of Developmental Psychology*, 7(1), 55-71.

Dodd, B. and McIntosh, B. (2008) 'The input processing, cognitive linguistic and oro-motor skills of children with speech difficulty', *International Journal of Speech-Language Pathology*, 10(3), 169-178.

Dodd, B. and McIntosh, B. (2010) 'Two-year-old phonology: impact of input, motor and cognitive abilities on development', *Journal of Child Language*, 37(05), 1027-1046.

Dodd, B., Ozanne, A., McIntosh, B., Crosbie, S. and Teitzel, T. (2000) *Preschool and Primary Inventory of Phonological Awareness.* London*:*Pearson Education.

Dodd, B., Ttofari-Eecen, K., Brommeyer, K., Ng, K., Reilly, S. and Morgan, A. (2017) 'Delayed and Disordered Development of Articulation and Phonology between Four and Seven Years', *submitted*.

Dodd, B., Zhu Hua, Crosbie, S., Holm, A. and Ozanne, A. eds. (2002) *Diagnostic Evaluation of Articulation and Phonology,* London: Psychological Corporation.

Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn, D. N. and Kurs-Lasky, M. (1999) 'Maternal Education and Measures of Early Speech and Language', *Journal of Speech, Language and Hearing Research*, 42(6), 1432-1443.

Donegan, P. (2013) 'Normal vowel development' in Ball, M. J. and Gibbon, F., eds., Handbook of vowels and vowel disorders, Hove: Psychology Press, 24-60.

Dubasik, V.L. and Ingram, D., 2013. 'Comparing phonology of dyads of children with typical development and protracted development'. *Clinical linguistics and phonetics*, 27(9), 705-719.

Claessen, M., Beattie, T., Roberts, R., Leitão, S., Whitworth, A. and Dodd, B. (2016) 'Is two too early? Assessing toddlers' phonology', *Speech, Language and Hearing*, 1-11.

Dunn, L. M., Dunn, D. M., Styles, B. and Sewell, J. (2009) *The British Picture Vocabulary Scale III - 3rdEdition*, London: GL Assessment.

Eadie, P., Morgan, A., Ukoumunne, O. C., Ttofari Eecen, K., Wake, M. and Reilly, S. (2015) 'Speech sound disorder at 4 years: prevalence, comorbidities, and predictors in a community cohort of children', *Developmental Medicine and Child Neurology*, 57(6), 578-584.

Edwards, J. (1992) 'Compensatory speech motor abilities in normal and phonologically disordered children', *Journal of phonetics*, 20(2), 189-207.

Edwards, J., Fourakis, M., Beckman, M. E. and Fox, R. A. (1999) 'Characterizing knowledge deficits in phonological disorders', *Journal of Speech, Language, and Hearing Research*, 42(1), 169-186.

Edwards, J., Fox, R. A. and Rogers, C. L. (2002) 'Final Consonant Discrimination in Children: Effects of Phonological Disorder, Vocabulary Size, and Articulatory Accuracy', *Journal of Speech, Language and Hearing Research*, 45(2), 231.

Edwards, J., Gibbon, F. and Fourakis, M. (1997) 'On Discrete Changes in the Aquisition of the Alveolar/Velar Stop Consonant Contrast', *Language and Speech*, 40(2), 203.

Eilers, R. E., Gavin, W. and Wilson, W. R. (1979) 'Linguistic Experience and Phonemic Perception in Infancy: A Crosslinguistic Study', *Child development*, 50(1), 14-18.

Eimas, P. D., Siqueland, E. R., Juscyk, P. and Vigorito, J. (1971) 'Speech perception in infants', *Science*, 171(3968), 303-306.

Ellis, E. M. and Thal, D. J. (2008) 'Early language delay and risk for language impairment', *SIG 1 Perspectives on Language Learning and Education*, 15(3), 93-100.

Emanuel, R., Chiat, S. and Roy, P. (2007) 'Evaluation of the clinical decisions made for 2-year-olds referred for speech and language therapy: a follow-up study' *International Journal of Language and Communication Disorders*, 42, 1-15.

Enderby, P. and John, A. (1997) *Therapy outcome measures (speech and language therapy)*, London: Singular.

Entwisle, D. R. and Astone, N. M. (1994) 'Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status', *Child development*, 1521-1540.

European Commission (2014) *European Qualifications Framework* [online], available: http://ec.europa.eu/ploteus/search/site?f[0]=im_field_entity_type%3A97# [accessed 14/04/2017].

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. and Strahan, E. J. (1999) 'Evaluating the use of exploratory factor analysis in psychological research', *Psychological methods*, 4(3), 272.

Felsenfeld, S., Broen, P. A. and McGue, M. (1992) 'A 28-Year Follow-Up of Adults With a History of Moderate Phonological Disorder: Linguistic and Personality Results', *Journal of Speech, Language and Hearing Research,*, 35(5), 1114-1125.

Felsenfeld, S., McGue, M. and Broen, P. A. (1995) 'Familial Aggregation of Phonological DisordersResults From a 28-Year Follow-Up', *Journal of Speech, Language, and Hearing Research*, 38(5), 1091-1107.

Felsenfeld, S. and Plomin, R. (1997) 'Epidemiological and Offspring Analyses of Developmental Speech Disorders Using Data From the Colorado Adoption Project', *Journal of Speech, Language, and Hearing Research*, 40(4), 778-791.

Ferguson, C. A. and Farwell, C. B. (1975) 'Words and sounds in early language acquisition: English initial consonants in the first fifty words', *Language*, 51, 419-439.

Fey, M. E. (1992) 'Articulation and PhonologyInextricable Constructs in Speech Pathology', *Language, Speech, and Hearing Services in Schools*, 23(3), 225-232.

Fikkert, P. (2010) 'Developing representations and the emergence of phonology: Evidence from perception and production', *Laboratory phonology*, 10(4), 227-255.

Flipsen, P. (2015) 'Emergence and Prevalence of Persistent and Residual Speech Errors', *Seminars in Speech and Language*, 36(04), 217-223.

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A. and Elbert, M. (1990) 'Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children', *Clinical Linguistics and Phonetics*, 4(4), 327-340.

Fox, A., Dodd, B. and Howard, D. (2002) 'Risk factors for speech disorders in children', *International Journal of Language and Communication Disorders*, 37(2), 117-131.

Fox, A. V. and Dodd, B. (2001) 'Phonologically Disordered German-Speaking Children', *American Journal of Speech-Language Pathology*, 10(3), 291.

Fox, A. V. and Dodd, B. J. (1999) 'Der Erwerb des phonologischen Systems in der deutschen Sprache', *Sprache-Stimme-Gehör*, 23(4), 183.

Foy, J. G. and Mann, V. (2001) 'Does strength of phonological representations predict phonological awareness in preschool children?', *Applied Psycholinguistics*, 22(03), 301-325.

Freeby, N. and Madison, C. L. (1989) 'Children's Perceptions of Peers with Articulation Disorders', *Child Study Journal*, 19(2), 133-44.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C. and Fletcher, J. M. (2006) 'The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems', *Journal of Educational Psychology*, 98(1), 29.

Gathercole, S. E. (2006) 'Nonword repetition and word learning: The nature of the relationship', *Applied Psycholinguistics*, 27(04), 513-543.

Gathercole, S. E. and Adams, A.-M. (1994) 'Children′ s phonological working memory: Contributions of long-term knowledge and rehearsal', *Journal of Memory and Language*, 33(5), 672-688.

Gathercole, S. E. and Baddeley, A. D. (1993) *Working memory and language*, Hove: Psychology Press.

Gernand, K. L. and Moran, M. J. (2007) 'Phonological Awareness Abilities of 6-Year-Old Children with Mild to Moderate Phonological Impairments', *Communication Disorders Quarterly*, 28(4), 206-215.

Gibbon F.E. (2007). Research and practice in developmental phonological disorders. In: Pennington M.C. (ed.) *Phonology in Context*. London: Palgrave Macmillan, 245-273.

Gibbon, F. and Hardcastle, W. (1987) 'Articulatory description and treatment of "lateral /s/" using electropalatography: A case study', *British Journal of Disorders of Communication*, 22(3), 203-217.

Gibbon, F. E. (1999) 'Undifferentiated lingual gestures in children with articulation/phonological disorders', *Journal of Speech, Language, and Hearing Research*, 42(2), 382-397.

Gibbon, F. E. and Lee, A. (2016) 'Electropalatographic (EPG) evidence of covert contrasts in disordered speech', *Clinical Linguistics and Phonetics*, 31(1), 4-20.

Gierut, J. A. (1998) 'Treatment Efficacy: Functional Phonological Disorders in Children', *Journal of Speech Language and Hearing Research*, 41(1), S85-100.

Gierut, J. A., Morrisette, M. L. and Ziemer, S. M. (2010) 'Nonwords and generalization in children with phonological disorders', *American Journal of Speech-Language Pathology*, 19(2), 167-177.

Glascoe, F. P. (1997) *Parents' evaluation of developmental status (PEDS)*. Nashville, TN: Ellsworth and Vandermeer Press Ltd.

Glogowska, M. and Campbell, R. (2000) 'Investigating parental views of involvement in pre-school speech and language therapy', *International Journal of Language and Communication Disorders*, 35(3), 391-405.

Glogowska, M., Roulstone, S., Enderby, P. and Peters, T. J. (2000) 'Randomised controlled trial of community based speech and language therapy in preschool children', *British Medical Journal*, 321(7266), 923.

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

Goldstein, M. H., King, A. P. and West, M. J. (2003) 'Social interaction shapes babbling: Testing parallels between birdsong and speech', *Proceedings of the National Academy of Sciences*, 100(13), 8030-8035.

Goldstein, M. H. and Schwade, J. A. (2008) 'Social feedback to infants' babbling facilitates rapid phonological learning', *Psychological Science*, 19(5), 515-523.

Goozée, J., Murdoch, B., Ozanne, A., Cheng, Y., Hill, A. and Gibbon, F. (2007) 'Lingual kinematics and coordination in speech-disordered children exhibiting differentiated versus undifferentiated lingual gestures', *International Journal of Language and Communication Disorders*, 42(6), 703-724 22p.

Gósy, M. and Horváth, V. (2014) 'Speech processing in children with functional articulation disorders', *Clinical Linguistics and Phonetics*, 29(3), 185-200.

Gross, J. (2011) *Two Years On: final report of hte communication champion for children*, [online]. London: Office of the Communication. Available: Championhttps://www.thecommunicationtrust.org.uk/media/9683/nwm_final_jean_gross_two_years_on_report.pdf [accessed 23/04/2016].

Grunwell, P. (1981) *The nature of phonological disability in children*, Cambridge MA:Academic Press.

Guadagnoli, E. and Velicer, W. F. (1988) 'Relation of sample size to the stability of component patterns', *Psychological bulletin*, 103(2), 265.

Harris, J., Botting, N., Myers, L. and Dodd, B. (2011) 'The relationship between speech impairment, phonological awareness and early literacy development', *Australian Journal of Learning Difficulties*, 16(2), 111-125.

Harrison, L. J. and McLeod, S. (2010) 'Risk and protective factors associated with speech and language impairment in a nationally representative sample of 4-to 5-year-old children', *Journal of Speech, Language, and Hearing Research*, 53(2), 508-529.

Harrison, L. J., McLeod, S., Berthelsen, D. and Walker, S. (2009) 'Literacy, numeracy, and learning in school-aged children identified as having speech and language impairment in early childhood', *International Journal of Speech-Language Pathology*, 11(5), 392-403.

Hauner, K. K., Shriberg, L. D., Kwiatkowski, J. and Allen, C. T. (2005) 'A subtype of speech delay associated with developmental psychosocial involvement', *Journal of Speech, Language, and Hearing Research*, 48(3), 635-650.

Hawa, V. V. and Spanoudis, G. (2014) 'Toddlers with delayed expressive language: An overview of the characteristics, risk factors and language outcomes', *Research in Developmental Disabilities*, 35(2), 400-407.

Hayden, D. (1999) *VMPAC: Verbal motor production assessment for children*, London: Psychological Corporation.

Hayes, N., Keegan, S. and Goulding, E. (2012) *Evaluation of the Speech and Language Therapy Service of Tallaght West Childhood Development Initiative*, Dublin: Childhood Development Initiative (CDI).

Hayiou-Thomas, M. E., Dale, P. S. and Plomin, R. (2014) 'Language impairment from 4 to 12 years: Prediction and etiology', *Journal of Speech, Language, and Hearing Research*, 57(3), 850-864.

Hayiou-Thomas, M. E., Carroll, J. M., Leavett, R., Hulme, C. and Snowling, M. J. (2016) 'When does speech sound disorder matter for literacy? The role of disordered speech errors, co-occurring language impairment and family risk of dyslexia', *Journal of Child Psychology and Psychiatry*, 58(2), 197-205.

Hazan, V. and Barrett, S. (2000) 'The development of phonemic categorization in children aged 6–12', *Journal of phonetics*, 28(4), 377-396.

Healthcare Pricing Office (2016) *Perinatal Statistics Report 2013*. Dublin:Health Service Executive

Heaton, R. K. (1993) *Wisconsin Card Sorting Test (WCST)*, Revised and expanded. Odessa, Fla: Psychological Assessment Resources.

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

Hesketh, A., Dima, E. and Nelson, V. (2007) 'Teaching phoneme awareness to pre-literate children with speech disorder: a randomized controlled trial', *International Journal of Language and Communication Disorders*, 42(3), 251-271.

Hill, E. L. (2001) 'Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments', *International Journal of Language and Communication Disorders*, 36(2), 149-171.

Hickey, R., 2004. 'The phonology of Irish English.' In Kortmann, B., Schneider, E.W., Burridge, K., Mesthrie, R. and Upton, C., (eds.), *Handbook of varieties of English. Volume 1: Phonology*. Berlin: Mouton de Gruyter, 68-97.

Hitchcock, E. R., Harel, D. and Byun, T. (2015) 'Social, Emotional, and Academic Impact of Residual Speech Errors in School-Aged Children: A Survey Study', *Seminars in Speech and Language*, 36(4), 283-294.

Hoff, E. (2003) 'The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech', *Child development*, 74(5), 1368-1378.

Hoff, E., Laursen, B., Tardif, T. and Bornstein, M. (2002) 'Socioeconomic status and parenting', *Handbook of parenting Volume 2: Biology and ecology of parenting*, 8(2), 231-52.

Holm, A., Crosbie, S. and Dodd, B. (2007) 'Differentiating normal variability from inconsistency in children's speech: normative data', *International Journal of Language and Communication Disorders*, 42(4), 467-486.

Holm, A. and Dodd, B. (2010) 'A novel intervention for non-developmental speech error patterns.', *Poster presented at Child Language Seminar*, London, June 2010.

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

Holm, A., Farrier, F. and Dodd, B. (2008) 'Phonological awareness, reading accuracy and spelling ability of children with inconsistent phonological disorder', *International Journal of Language and Communication Disorders*, 43(3), 300-322.

Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P. and Serniclaes, W. (2009) 'French native speakers in the making: From language-general to language-specific voicing boundaries', *Journal of Experimental Child Psychology*, 104(4), 353-366.

Hough (2010) '23,000 on waiting list for speech therapy', *Irish Examiner*, 3rd May 2010, [online], available at: http://www.irishexaminer.com/ireland/health/23000-on-waiting-list-for-speech-therapy-118806.html [accessed 07/07/2015].

Hresko, WP.; Herron, SR.; Peak, PK. (1996) *Test of Early Written Language: A Method of Evaluating the Products and Processes of Emerging Writing Ability.* Vol. 2nd ed. Austin, TX: Pro-ed.

HSE (2016a) *Health Service Performance Report December 2016* [online], available: http://www.hse.ie/eng/services/publications/performancereports/ [accessed 23/03/17].

HSE (2016b) *HSE Census Report, July 2016.* [online], available at: http://www.hse.ie/eng/staff/Resources/Employment_Reports/Census-Report-July-2016.pdf. Dublin: Health Service Executive [accessed 23/03/17]

Hua, Z. and Dodd, B. (2000) 'Putonghua (modern standard Chinese)-speaking children with speech disorder', *Clinical Linguistics and Phonetics*, 14(3), 165-191.

Ingram, D. (1976) 'Aspects of phonological acquisition', in Ingram D. ed. *Phonological disability in children.* London: Edward Arnold, 10-49.

Innovations, A. (1995) SAILS: *Speech assessment and interactive learning system* [Computer software], London, Ontario, Canada: Author.

Jacques, S. and Zelazo, P. D. (2001) 'The Flexible Item Selection Task (FIST): A Measure of Executive Function in Preschoolers', *Developmental Neuropsychology*, 20(3), 573-591.

Jacques, S. and Zelazo, P. D. (2005) 'Language and the Development of Cognitive Flexibility: Implications for Theory of Mind', in Astington, J.W. and Baird, J.A. eds., 2005. *Why language matters for theory of mind*. Oxford: Oxford University Press.

Jakobson, R. (1968) *Child language, aphasia and phonological universals*, The Hague: Mouton Publishers.

James, D. G. (2001) 'Use of phonological processes in Australian children ages 2 to 7; 11 years', *Advances in Speech Language Pathology*, 3(2), 109-127.

James, D. G. H., van Doorn, J. and McLeod, S. (2008) 'The contribution of polysyllabic words in clinical decision making about children's speech', *Clinical Linguistics and Phonetics*, 22(4-5), 345-353.

Jamieson, D. G. and Rvachew, S. (1992) 'Remediating speech production errors with sound identification training', *Journal of Speech-Language Pathology and Audiology*. 16(3), 201-210.

Jenni, O. G., Chaouch, A., Caflisch, J. and Rousson, V. (2013) 'Infant motor milestones: poor predictive value for outcome of healthy children', *Acta Paediatrica*, 102(4), 181-184.

Jessup, B., Ward, E., Cahill, L. and Keating, D. (2008) 'Prevalence of speech and/or language impairment in preparatory students in northern Tasmania', *International Journal of Speech-Language Pathology*, 10(5), 364-377.

Joffe, V. and Pring, T. (2008) 'Children with phonological problems: a survey of clinical practice', *International Journal of Language and Communication Disorders*, 43(2), 154-164.

Johnson, C. J., Beitchman, J. H. and 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), 51-65.

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., Brownlie, E. B., Douglas, L., Taback, N., Lam, I. and Wang, M. (1999) 'Fourteen-Year Follow-Up of Children With and Without Speech/Language ImpairmentsSpeech/Language Stability and Outcomes', *Journal of Speech, Language, and Hearing Research*, 42(3), 744-760.

Johnson, E. P., Pennington, B. F., Lowenstein, J. H. and Nittrouer, S. (2011) 'Sensitivity to Structure in the Speech Signal by Children with Speech Sound Disorder and Reading Disability', *Journal of Communication Disorders*, 44(3), 294-314.

Kapa, L. L., Plante, E. and Doubleday, K. (2017) 'Applying an Integrative Framework of Executive Function to Preschoolers With Specific Language Impairment', *Journal of Speech, Language, and Hearing Research*, 1-15.

Kaplan, W., Wirtz, V. J., Mantel-Teeuwisse, A. K., Stolk, P., Duthey, B. and Laing, R. (2013) 'Priority Medicines for Europe and the World 2013 Update.World Health Organization [online]. Available at: http://www. who. int/medicines/areas/priority_medicines/MasterDocJune28_FINAL_Web. pdf. [accessed 12/12/2014].

Karlsson, H. B., Shriberg, L. D., Jr, P. F. and McSweeny, J. L. (2002) 'Acoustic phenotypes for speech-genetics studies: Toward an acoustic marker for residual/s/distortions', *Clinical Linguistics and Phonetics*, 16(6), 403-424.

Karmiloff-Smith, A. (1998) 'Development itself is the key to understanding developmental disorders', *Trends in Cognitive Science*, 2(10), 389-398.

Karmiloff-Smith, A. (2009) 'Nativism versus neuroconstructivism: rethinking the study of developmental disorders', *Developmental psychology*, 45(1), 56.

Keating, D., Turrell, G. and Ozanne, A. (2001) 'Childhood speech disorders: Reported prevalence, comorbidity and socioeconomic profile', *Journal of paediatrics and child health*, 37(5), 431-436.

Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S. and Diliberti, M. (2016) 'The Condition of Education 2016. NCES 2016-144', *National Center for Education Statistics*.

Kenney, M. K., Barac-Cikoja, D., Finnegan, K., Jeffries, N. and Ludlow, C. L. (2006) 'Speech perception and short-term memory deficits in persistent developmental speech disorder', *Brain and language*, 96(2), 178-190.

Kent, R. D. (1992) 'The biology of phonological development' in Ferguson, C. A., Menn, L. and Stoel-Gammon, C., eds., *Phonological development: Models, research, implications*, York: York Press, 65-90.

Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C. J. and Williams, N. M. (2010) 'The role of vocal practice in constructing phonological working memory', *Journal of speech, language, and hearing research*, 53(5), 1280-1293.

Kerstjens, J. M., de Winter, A. F., Bocca-Tjeertes, I. F., ten Vergert, E. M. J., Reijneveld, S. A. and Bos, A. F. (2011) 'Developmental Delay in Moderately Preterm-Born Children at School Entry', *The Journal of Pediatrics*, 159(1), 92-98.

Kraemer, S. (2000) 'The fragile male', *British Medical Journal*, 321(7276), 1609-1612.

Kuhl, P. K. (2004) 'Early language acquisition: cracking the speech code', *National Review of Neuroscience*, 5(11), 831-843.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T. and Pruitt, J. (2005) 'Early Speech Perception and Later Language Development: Implications for the "Critical Period"', *Language Learning and Development*, 1(3-4), 237-264.

Larrivee, L. S. and Catts, H. W. (1999) 'Early reading achievement in children with expressive phonological disorders', *American Journal of Speech-Language Pathology*, 8(2), 118-128.

Lasky, R. E., Syrdal-Lasky, A. and Klein, R. E. (1975) 'VOT discrimination by four to six and a half month old infants from Spanish environments', *Journal of Experimental Child Psychology*, 20(2), 215-225.

Law, J., Boyle, J., Harris, F., Harkness, A. and Nye, C. (2000) 'Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature', *International Journal of Langiage and Communication Disorders*, 35(2), 165-88.

Leather, C. V. and Henry, L. A. (1994) 'Working Memory Span and Phonological Awareness Tasks as Predictors of Early Reading Ability', *Journal of Experimental Child Psychology*, 58(1), 88-111.

Leitão, S. and Fletcher, J. (2004) 'Literacy outcomes for students with speech impairment: long-term follow-up', *International Journal of Language and Communication Disorders*, 39(2), 245-256.

Leitão, S., Hogben, J. and Fletcher, J. (1997) 'Phonological processing skills in speech and language impaired children', *International Journal of Language and Communication Disorders*, 32(2s), 91-111.

Lewis, B. A., Avrich, A. A., Freebairn, L. A., Taylor, H. G., Iyengar, S. K. and Stein, C. M. (2011) 'Subtyping Children With Speech Sound Disorders by Endophenotypes', *Topics in Language Disorders*, 31(2), 112–127.

Lewis, B. A. and Freebairn, L. (1997) 'Subgrouping children with familial phonologic disorders', *Journal of Communication Disorders*, 30(5), 385-402.

Lewis, B. A., Freebairn, L., Tag, J., Ciesla, A. A., Iyengar, S. K., Stein, C. M. and Taylor, H. G. (2015) 'Adolescent Outcomes of Children with Early Speech Sound Disorders with and without Language Impairment', *American Journal of Speech-Language Pathology*. 24(2),150-163

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Miscimarra, L., Iyengar, S. K. and Taylor, H. G. (2007) 'Speech and Language Skills of Parents of Children With Speech Sound Disorders', *American Journal of Speech-Language Pathology*, 16(2), 108-118.

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Stein, C. M., Shriberg, L. D., Iyengar, S. K. and Taylor, H. G. (2006a) 'Dimensions of early speech sound disorders: A factor analytic study', *Journal of Communication Disorders*, 39(2), 139-157.

Lewis, B. A., Freebairn, L. A. and Taylor, H. G. (2000a) 'Academic outcomes in children with histories of speech sound disorders', *Journal of Communication Disorders*, 33(1), 11-30.

Lewis, B. A., Freebairn, L. A. and Taylor, H. G. (2000b) 'Follow-Up of Children with Early Expressive Phonology Disorders', *Journal of learning disabilities*, 33(5), 433-444.

Lewis, B. A., Shriberg, L. D., Freebairn, L. A., Hansen, A. J., Stein, C. M., Taylor, H. G. and Iyengar, S. K. (2006b) 'The genetic bases of speech sound disorders: Evidence from spoken and written language', *Journal of Speech, Language, and Hearing Research*, 49(6), 1294-1312.

Li, F., Edwards, J. and Beckman, M. E. (2009) 'Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers', *Journal of phonetics*, 37(1), 111-124.

Lipetz, H. M. and Bernhardt, B. M. (2013) 'A multi-modal approach to intervention for one adolescent's frontal lisp', *Clinical Linguistics and Phonetics*, 27(1), 1-17.

Locke, J. L. (1980) 'The Inference of Speech Perception in the Phonologically Disordered Child. Part IISome Clinically Novel Procedures, Their Use, Some Findings', *Journal of Speech and Hearing Disorders*, 45(4), 445-468.

Lof, G. L. and Watson, M. (2010) 'Five reasons why nonspeech oral motor exercises (NSOME) do not work', *Perspectives on School-Based Issues*, 11(4), 109-117.

Loucas, T., Baird, G., Simonoff, E. and Slonims, V. (2016) 'Phonological processing in children with specific language impairment with and without reading difficulties', *International Journal of Language and Communication Disorders*, 51(5), 581-588.

Luoma, L., Herrgård, E., Martikainen, A. and Ahonen, T. (1998) 'Speech and language development of children born at 32 weeks gestation: a 5-year prospective follow-up study', *Developmental Medicine and Child Neurology*, 40(6), 380-387.

Macken, M. A. (1995) 'Phonological acquisition', in Selkirk, E., Goldsmith, J.A., Riggle, J. and Alan, C.L. eds., *The handbook of phonological theory*. Hoboken, NJ:John Wiley and Sons. 671-696

Macrae, T. (2013) 'Lexical and child-related factors in word variability and accuracy in infants', *Clinical Linguistics and Phonetics*, 27(6-7), 497-507.

Macrae, T., Tyler, A. A. and Lewis, K. E. (2014) 'Lexical and Phonological Variability in Preschool Children With Speech Sound Disorder', *American Journal of Speech-Language Pathology*, 23(1), 27-35.

Masterson, J. J., Bernhardt, B. H. and Hofheinz, M. K. (2005) 'A comparison of single words and conversational speech in phonological evaluation', *American Journal of Speech-Language Pathology*, 14(3), 229-241.

McGrath, L. M., Pennington, B. F., Willcutt, E. G., Boada, R., Shriberg, L. D. and Smith, S. D. (2007) 'Gene × Environment interactions in speech sound disorder predict language and preliteracy outcomes', *Development and Psychopathology*, 19(04), 1047-1072.

McIntosh, B. and Dodd, B. (2011) *Toddler Phonology Test*. London:Pearson

McKean, C., Letts, C. and Howard, D. (2013) 'Developmental change is key to understanding primary language impairment: The case of phonotactic probability and nonword repetition', *Journal of Speech, Language, and Hearing Research*, 56(5), 1579-1594.

McKinnon, D. H., McLeod, S. and Reilly, S. (2007) 'The prevalence of stuttering, voice, and speech-sound disorders in primary school students in Australia', *Language, Speech, and Hearing Services in Schools*, 38(1), 5-15.

McLeod, S., Baker, E., Murdoch, B., Goozee, J., Whelan, B. and Docking, K. (2004) 'Current clinical practice for children with speech impairment', in *Proceedings of the 26th world congress of the International Association of Logopedics and Phoniatrics. Brisbane: University of Queensland*, Brisbane, Australia: The University of Queensland.

McLeod, S., Crowe, K., Masso, S., Baker, E., McCormack, J., Wren, Y., Roulstone, S. and Howland, C. (2017) 'Profile of Australian preschool children with speech sound disorders at risk for literacy difficulties', *Australian Journal of Learning Difficulties*, 22(1), 15-33.

McLeod, S., Harrison, L. J., McAllister, L. and McCormack, J. (2013) 'Speech Sound Disorders in a Community Study of Preschool Children', *American Journal of Speech-Language Pathology*, 22(3), 503-522.

McLeod, S. and Hewett, S. R. (2008) 'Variability in the production of words containing consonant clusters by typical 2-and 3-year-old children', *Folia Phoniatrica et Logopaedica*, 60(4), 163-172.

McLeod, S. and McKinnon, D. H. (2007) 'Prevalence of communication disorders compared with other learning needs in 14,500 primary and secondary school students', *International Journal of Language and Communication Disorders*, 42(s1), 37-59.

McNally, S. and Quigley, J. (2014) 'An Irish cohort study of risk and protective factors for infant language development at 9 months', *Infant and Child Development*, 23(6), 634-649.

McSweeny, J.L. and Shriberg, L.D., 1995. Segmental and suprasegmental transcription reliability (Tech. Rep. No. 2). *Phonology Project, Waisman Center on Mental Retardation and Human Development*, University of Wisconsin-Madison.

Morgan, A., Eecen, K. T., Pezic, A., Brommeyer, K., Mei, C., Eadie, P., Reilly, S. and Dodd, B. (2017) 'Who to Refer for Speech Therapy at 4 Years of Age Versus Who to "Watch and Wait"?', *The Journal of Pediatrics*. 185, 200-204

Morley, M. E. (1972) *The development and disorders of speech in childhood*, London: Churchill Livingstone.

Morrison, C. M., Chappell, T. D. and Ellis, A. W. (1997) 'Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables', *The Quarterly Journal of Experimental Psychology: Section A*, 50(3), 528-559.

Mulligan, J. and Mitchelmore, M. (2009) 'Awareness of pattern and structure in early mathematical development', *Mathematics Education Research Journal*, 21(2), 33-49.

Mulligan, J., Prescott, A. and Mitchelmore, M. (2004) 'Children's Development of Structure in Early Mathematics', in M Høines and A Fuglestad (Eds) *Proceedings of the 28th annual conference of the International Group for the Psychology of Mathematics Education.* Vo.l 3, 393-401, Bergen, Norway: PME,

Munson, B., Edwards, J. and Beckman, M. E. (2005a) 'Phonological knowledge in typical and atypical speech–sound development', *Topics in Language Disorders*, 25(3), 190.

Munson, B., Edwards, J. and Beckman, M. E. (2005b) 'Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders', *Journal of Speech, Language, and Hearing Research*, 48(1), 61-78.

Munson, B., Edwards, J., Beckman, M. E., (2012) 'Phonological representations in language acquisition: Climbing the ladder of abstraction', in Cohn, A., Fougeron, C. and Huffman, M. eds. *The Oxford Handbook of laboratory phonology*, Oxford: Oxford Univeristy Press, 288-309.

Munson, B., Edwards, J., Schellinger, S. K., Beckman, M. E. and Meyer, M. K. (2010) 'Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of Vox Humana', *Clinical Linguistics and Phonetics*, 24(4-5), 245-260.

Munson, B. and Krause, M. O. P. (2016) 'Phonological encoding in speech-sound disorder: evidence from a cross-modal priming experiment: Picture-word interference', *International Journal of Language and Communication Disorders*.

Murphy, C. F., Pagan-Neves, L. O., Wertzner, H. F. and Schochat, E. (2014) 'Auditory and visual sustained attention in children with speech sound disorder', *PloS one*, 9(3), e93091.

Murray, A., McCrory, C., Thornton, M., Williams, J., Quail, A., Swords, L., Doyle, E. and Harris, E. (2011) 'Growing Up in Ireland: Design, instrumentation and procedures for the child cohort', Ireland, Dublin: Department of Health and Children.

Murray, E., McCabe, P. and Ballard, K. J. (2015) 'A Randomized Controlled Trial for Children With Childhood Apraxia of Speech Comparing Rapid Syllable Transition Treatment and the Nuffield Dyspraxia Programme–Third Edition', *Journal of Speech, Language, and Hearing Research*, 58(3), 669-686.

Namasivayam, A. K., Pukonen, M., Goshulak, D., Yu, V. Y., Kadis, D. S., Kroll, R., Pang, E. W. and De Nil, L. F. (2013) 'Relationship between speech motor control and speech intelligibility in children with speech sound disorders', *Journal of Communication Disorders*, 46(3), 264-280.

Nathan, L., Stackhouse, J., Goulandris, N. and Snowling, M. J. (2004a) 'The Development of Early Literacy Skills Among Children With Speech DifficultiesA Test of the Critical Age Hypothesis', *Journal of Speech, Language, and Hearing Research*, 47(2), 377-391.

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

Nathan, L. and Wells, B. (2001) 'Can children with speech difficulties process an unfamiliar accent?', *Applied Psycholinguistics*, 22(03), 343-361.

Newcomer, P. L. and Hammill, D. D. (1988) *Test of language development-2: primary*, Austin, TX: Pro-ed.

Niclasen, J., Obel, C., Homøe, P., Kørvel-Hanquist, A. and Dammeyer, J. (2016) 'Associations between otitis media and child behavioural and learning difficulties: Results from a Danish cohort', *International journal of pediatric otorhinolaryngology*, 84, 12-20.

Nip, I. S. B., Green, J. R. and Marx, D. B. (2009) 'Early speech motor development: Cognitive and linguistic considerations', *Journal of Communication Disorders*, 42(4), 286-298.

Nittrouer, S. (1992) 'Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries', *Journal of phonetics*, 20, 351-382.

Nittrouer, S. (2001) 'Challenging the notion of innate phonetic boundaries', *The Journal of the Acoustical Society of America*, 110(3), 1598-1605.

Nittrouer, S. (2002a) 'From Ear to CortexA Perspective on What Clinicians Need to Understand About Speech Perception and Language Processing', *Language, Speech, and Hearing Services in Schools*, 33(4), 237-252.

Nittrouer, S. (2002b) 'Learning to perceive speech: How fricative perception changes, and how it stays the same', *The Journal of the Acoustical Society of America*, 112(2), 711-719.

Nittrouer, S. and Miller, M. E. (1997) 'Developmental weighting shifts for noise components of fricative-vowel syllables', *The Journal of the Acoustical Society of America*, 102(1), 572-580.

Nittrouer, S., Shune, S. and Lowenstein, J. H. (2011) 'What is the deficit in phonological processing deficits: Auditory sensitivity, masking, or category formation?', *Journal of Experimental Child Psychology*, 108(4), 762-785.

Nittrouer, S., Studdert-Kennedy, M. and McGowan, R. S. (1989) 'The Emergence of Phonetic SegmentsEvidence from the Spectral Structure of Fricative-Vowel Syllables Spoken by Children and Adults', *Journal of Speech, Language, and Hearing Research*, 32(1), 120-132.

Nys, J. and Leybaert, J. (2013) 'Impact of language abilities on exact and approximate number skills development: evidence from children with specific language impairment', *Journal of Speech, Language, and Hearing Research*, 56(3), 956-970.

OECD (2008) *Growing Unequal?* [online], available: http://dx.doi.org/10.1787/9789264044197-en [accessed 04/05/2017].

Overby, M., Carrell, T. and Bernthal, J. (2007) 'Teachers' Perceptions of Students With Speech Sound Disorders: A Quantitative and Qualitative Analysis', *Language, Speech and Hearing Services in Schools*, 38(4), 327-341.

Papic, M. and Mulligan, J. (2005) 'Pre-schoolers' mathematical patterning', in P. Clarkson, A. Downton, D.Gronn, M. Horne, A. McDonough, R. Pierce, & A. Roche (Eds.), *Building connections: Research,theory and practice. (Proceedings of the 28th annual conference of the Mathematics Education Research Group of Australasia)*, Melbourne, Sydney: MERGA., 609-616.

Papic, M. and Mulligan, J. (2007) 'The Growth of Early Mathematical Patterning: An Intervention Study, ' in Watson, J., and Beswick, K., Eds. *Mathematics: Essential Research, Essential Practice — Volume 2, Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia* MERGA Inc.

Paradise, J. L., Dollaghan, C. A., Campbell, T. F., Feldman, H. M., Bernard, B. S., Colborn, D. K., Rockette, H. E., Janosky, J. E., Pitcairn, D. L., Sabo, D. L., Kurs-Lasky, M. and Smith, C. G. (2000) 'Language, Speech Sound Production, and Cognition in Three-Year-Old Children in Relation to Otitis Media in Their First Three Years of Life', *Pediatrics*, 105(5), 1119-1130.

Paradise, J. L., Campbell, T. F., Dollaghan, C. A., Feldman, H. M., Bernard, B. S., Colborn, D. K., Rockette, H. E., Janosky, J. E., Pitcairn, D. L. and Kurs-Lasky, M. (2005) 'Developmental outcomes after early or delayed insertion of tympanostomy tubes', *New England Journal of Medicine*, 353(6), 576-586.

Paradise , J. L., Feldman , H. M., Campbell , T. F., Dollaghan , C. A., Rockette , H. E., Pitcairn , D. L., Smith , C. G., Colborn , D. K., Bernard , B. S., Kurs-Lasky , M., Janosky , J. E., Sabo , D. L., O'Connor , R. E. and Pelham , W. E. J. (2007) 'Tympanostomy Tubes and Developmental Outcomes at 9 to 11 Years of Age', *New England Journal of Medicine*, 356(3), 248-261.

Passolunghi, M. C., Lanfranchi, S., Altoè, G. and Sollazzo, N. (2015) 'Early numerical abilities and cognitive skills in kindergarten children', *Journal of Experimental Child Psychology*, 135, 25-42.

Passolunghi, M. C., Vercelloni, B. and Schadee, H. (2007) 'The precursors of mathematics learning: Working memory, phonological ability and numerical competence', *Cognitive Development*, 22(2), 165-184.

Paul, R. (1993) 'Outcomes of early expressive language delay', *Journal of Childhood Communication Disorders*, 15(15), 7-14.

Paul, R. (1996) 'Clinical implications of the natural history of slow expressive language development', *American Journal of Speech-Language Pathology*, 5(2), 5-21

Paul, R. (2000) 'Predicting outcomes of early expressive language delay: Ethical implications.' in Bishop, D. V. M. and Leonard, L. B., eds., *Speech and language impairments in children: Causes, characteristics, intervention and outcome*, Hove, U.K: Psychology Press, 195-209.

Paul, R. and Jennings, P. (1992) 'Phonological behavior in toddlers with slow expressive language development', *Journal of Speech, Language, and Hearing Research*, 35(1), 99-107.

Pennington, B. F. (2006) 'From Single to Multiple Deficit Models of Developmental Disorders', *Cognition*, 101(2), 385-413.

Pennington, B. F. and Bishop, D. V. M. (2009) 'Relations among speech, language, and reading disorders', *Annual review of psychology*, 60(1), 283-306.

Peterson, R. L., Pennington, B. F., Shriberg, L. D. and Boada, R. (2009) 'What Influences Literacy Outcome in Children With Speech Sound Disorder?', *Journal of speech, language, and hearing research : Journal of Speech, Language and Hearing Research*, 52(5), 1175-1188.

Pett, M. A., Lackey, N. R. and Sullivan, J. J. (2003) *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*, London: Sage.

Pharr, A. B., Ratner, N. B. and Rescorla, L. (2000) 'Syllable structure development of toddlers with expressive specific language impairment', *Applied Psycholinguistics*, 21(04), 429-449.

Pollock, K. E. and Berni, M. C. (2003) 'Incidence of non-rhotic vowel errors in children: data from the Memphis Vowel Project', *Clinical Linguistics and Phonetics*, 17(4-5), 393-401.

Poteat, G. M. (2014) *The Beery-Buktenica Developmental Test of Visual-Motor Integration, Sixth Edition (Beery VMI).* London:Pearson

Powell, T. W. (2008) 'The use of nonspeech oral motor treatments for developmental speech sound production disorders: Interventions and interactions', *Language, Speech, and Hearing Services in Schools*, 39(3), 374-379.

Preston, J. and Edwards, M. L. (2010) 'Phonological awareness and types of sound errors in preschoolers with speech sound disorders', *Journal of Speech, Language, and Hearing Research*, 53(1), 44-60.

Preston, J. L., Hull, M. and Edwards, M. L. (2013) 'Preschool Speech Error Patterns Predict Articulation and Phonological Awareness Outcomes in Children With Histories of Speech Sound Disorders', *American Journal of Speech-Language Pathology*, 22(2), 173-184.

Preston, J. L., Irwin, J. R. and Turcios, J. (2015) 'Perception of Speech Sounds in School-Aged Children with Speech Sound Disorders', in *Seminars in Speech and Language*, 36(4), 224-233.

Preston, J. L. and Koenig, L. L. (2011) 'Phonetic variability in residual speech sound disorders: Exploration of subtypes', *Topics in Language Disorders*, 31(2), 168-184.

Preston, J. L., Leece, M. C. and Maas, E. (2016) 'Motor-based treatment with and without ultrasound feedback for residual speech-sound errors', *International Journal of Language and Communication Disorders*, 52(1), 80-94.

Priester, G., Post, W. and Goorhuis-Brouwer, S. (2009) 'Problems in speech sound production in young children. An inventory study of the opinions of speech therapists', *International journal of pediatric otorhinolaryngology*, 73(8), 1100-1104.

Raitano, N. A., Pennington, B. F., Tunick, R. A., Boada, R. and Shriberg, L. D. (2004) 'Pre-literacy skills of subgroups of children with speech sound disorders', *Journal of Child Psychology and Psychiatry*, 45(4), 821-835.

Ramus, F., Marshall, C. R., Rosen, S. and van der Lely, H. K. J. (2013) 'Phonological deficits in specific language impairment and developmental dyslexia: towards a multidimensional model', *Brain*, 136(2), 630-645.

Ramus, F., Peperkamp, S., Christophe, A., Jacquemot, C., Kouider, S. and Dupoux, E. (2010) 'A psycholinguistic perspective on the acquisition of phonology', *Laboratory phonology*, 10(3), 311-340.

RCSLT (2017) *The SEND reforms in England: Supporting children and young people with speech, language and communication needs*, London: Royal College of Speech and Language Therapists.

Reid, D. K., Hresko, W. P. and Hammill, D. D. (1989) *Test of early reading ability-2* (TERA-2), Austin, TX: Pro-ed.

Rescorla, L. (1989) 'The language development survey: a screening tool for delayed language in toddlers', *Journal of Speech and Hearing Disorders*, 54(4), 587-599.

Rescorla, L. and Schwartz, E. (1990) 'Outcome of toddlers with specific expressive language delay', *Applied Psycholinguistics*, 11(04), 393-407.

Reuterskiöld-Wagner, C., Sahlén, B. and Nyman, A. (2005) 'Non-word repetition and non-word discrimination in Swedish preschool children', *Clinical Linguistics and Phonetics*, 19(8), 681-699.

Richtsmeier, P. T. (2010) 'Child phoneme errors are not substitutions', *Toronto Working Papers in Linguistics*, 33(1).

Robbins, J. and Klee, T. (1987) 'Clinical assessment of oropharyngeal motor development in young children', *Journal of Speech and Hearing Disorders*, 52(3), 271-277.

Roberts, J., Hunter, L., Gravel, J., Rosenfeld, R., Berman, S., Haggard, M., Hall, J., Lannon, C., Moore, D., Vernon-Feagans, L. And Wallace, I. (2004) 'Otitis Media, Hearing Loss, and Language Learning: Controversies and Current Research', *Journal of Developmental and Behavioral Pediatrics*, 25(2), 110-122.

Roberts, J. E., Burchinal, M. and Footo, M. M. (1990) 'Phonological process decline from 2½ to 8 years', *Journal of Communication Disorders*, 23(3), 205-217.

Robertson, C. and Salter, W. (1997) *The phonological awareness test*, East Moline, IL: LinguiSystems.

Rooke, Z., O'Toole, C. and Gibbon, F. (2013) 'The Prevalence of Speech and Language Impairments among a Nationally Representative Sample of Irish Nine-Year-Olds.', *paper presented to the Irish Association of Speech and Language Therapists' Biennial Conference*, Dublin, 23rd and 24th April 2013,

Rose, Y. (2014) 'The emergence of first language phonology: Perception, articulation and representation', *New directions in the acquisition of Romance languages: Selected proceedings of The Romance Turn V*, 35-61.

Roulstone, S. (2001) 'Consensus and variation between speech and language therapists in the assessment and selection of preschool children for intervention: a body of knowledge or idiosyncratic decisions?', *International Journal of Language and Communication Disorders*, 36(3), 329-348.

Roulstone, S., Loader, S., Northstone, K. and Beveridge, M. (2002) 'The Speech and Language of Children Aged 25 Months: Descriptive Data from the Avon Longitudinal Study of Parents and Children', *Early Child Development and Care*, 172(3), 259-268.

Roulstone, S., Miller, L. L., Wren, Y. and Peters, T. J. (2009) 'The natural history of speech impairment of 8-year-old children in the Avon Longitudinal Study of Parents and Children: Error rates at 2 and 5 years', *International Journal of Speech and Language Pathology*, 11(5), 381-391.

Roulstone, S., Peters, T. J., Glogowska, M. and Enderby, P. (2003) 'A 12-month follow-up of preschool children investigating the natural history of speech and language delay', *Child: Care, Health and Development*, 29(4), 245-255.

Roulstone, S., Peters, T. J., Glogowska, M. and Enderby, P. (2008) 'Predictors and outcomes of speech and language therapists' treatment decisions', *International Journal of Speech-Language Pathology*, 10(3), 146-155.

Rowe, M. L. (2008) 'Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill', *Journal of Child Language*, 35(1), 185-205.

Ruscello, D. M. (1995) 'Visual feedback in treatment of residual phonological disorders', *Journal of Communication Disorders*, 28(4), 279-302.

Ruscello, D.M. (2003). 'Residual phonological errors'. In Kent, R. D., ed. *The MIT encyclopedia of communication disorders*. Cambridge, MA: MIT Press. 156-158

Rvachew, S. (2006) 'Longitudinal predictors of implicit phonological awareness skills', *American Journal of Speech-Language Pathology*, 15(2), 165-176.

Rvachew, S. (2007) 'Phonological Processing and Reading in Children With Speech Sound Disorders', *American Journal of Speech-Language Pathology*, 16(3), 260-270.

Rvachew, S. (2016) 'Speech Perception and Persistent Speech Errors', *Developmental Phonological Disorders* [online], 02/05/16, available: https://developmentalphonologicaldisorders.wordpress.com/2016/03/27/speech-perception-and-persistent-speech-errors-2/ [accessed 2/5/2016].

Rvachew, S. and Brosseau-Lapré, F. (2012) *Developmental phonological disorders: Foundations of clinical practice*, San Diego, CA: Plural Pub.

Rvachew, S. and Jamieson, D. G. (1989) 'Perception of voiceless fricatives by children with a functional articulation disorder', *Journal of Speech and Hearing Disorders*, 54(2), 193-208.

Rvachew, S., Mattock, K., Polka, L. and Ménard, L. (2006) 'Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and

Canadian French', *The Journal of the Acoustical Society of America*, 120(4), 2250-2259.

Rvachew, S. and Nowak, M. (2001) 'The Effect of Target-Selection Strategy on Phonological Learning', *Journal of Speech, Language, and Hearing Research*, 44(3), 610-623.

Rvachew, S., Nowak, M. and Cloutier, G. (2004) 'Effect of phonemic perception training on the speech production and phonological awareness skills of children with expressive phonological delay', *American Journal of Speech-Language Pathology*, 13(3), 250-263.

Rvachew, S., Ohberg, A., Grawburg, M. and Heyding, J. (2003) 'Phonological Awareness and Phonemic Perception in 4-Year-Old Children With Delayed Expressive Phonology Skills', *American Journal of Speech-Language Pathology*, 12(4), 463.

Rvachew, S., Pi-Yu, C. and Evans, N. (2007) 'Characteristics of Speech Errors Produced by Children With and Without Delayed Phonological Awareness Skills', *Language, Speech and Hearing Services in Schools*, 38(1), 60-71.

Sackett, D., Haynes, R., Guyatt, G. and Tugwell, P. (1991) *Clinical Epidemiology: a Basic Science for Clinical Medicine (2nd ed.)*, Boston: Little Brown and Company.

Sansavini, A., Guarini, A., Justice, L. M., Savini, S., Broccoli, S., Alessandroni, R. and Faldella, G. (2010) 'Does preterm birth increase a child's risk for language impairment?', *Early human development*, 86(12), 765-772.

Selby, J. C., Robb, M. P. and Gilbert, H. R. (2000) 'Normal vowel articulations between 15 and 36 months of age', *Clinical Linguistics and Phonetics*, 14(4), 255-265.

Semel, E., Wiig, E. and Secord, W. (2004) *Clinical evaluation of language fundamentals—4 (UK)* San Antonio, TX: The Psychological Corporation.

Semel, E., Wiig, E. and Secord, W. (2006) *Clinical evaluation of language fundamentals-Preschool 2 (UK)*, San Antonio, TX: The Psychological Corporation.

Shavers, V. L. (2007) 'Measurement of socioeconomic status in health disparities research', *Journal of the national medical association*, 99(9), 1013.

Shiller, D. M., Rvachew, S. and Brosseau-Lapré, F. (2010) 'Importance of the auditory perceptual target to the achievement of speech production accuracy. Importance de la cible perceptive auditive dans l'atteinte d'une production adéquate de la parole', *Revue canadienne d'orthophonie et d'audiologie*, 34(3), 181.

Shriberg, L.D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech, Language, and Hearing Research*, 36(1), 105-140.

Shriberg, L. (1994) 'Five subtypes of developmental phonological disorders', *Clinics in Communication Disorders*, 4(1), 38-53.

Shriberg, L. and Austin, D. (1998) 'Comorbidity of speech-language disorder: Implications for a phenotype marker for speech delay' in Paul, R., ed., *The speech-language connection*, Baltimore: Brookes, 73-117.

Shriberg, L. D. (2003a) 'Diagnostic markers for child speech-sound disorders: introductory comments', *Clinical Linguistics and Phonetics*, 17(7), 501-505.

Shriberg, L. D. and Austin, D. (1997) 'The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data', *Journal of Speech, Language and Hearing Research*, 40(4), 723.

Shriberg, L. D., Flipsen Jr, P., Kwiatkowski, J. and Mcsweeny, J. L. (2003a) 'A diagnostic marker for speech delay associated with otitis media with effusion: the intelligibility-speech gap', *Clinical Linguistics and Phonetics*, 17(7), 507-528.

Shriberg, L. D., Fourakis, M., Hall, S. D., Karlsson, H. B., Lohmeier, H. L., McSweeny, J. L., Potter, N. L., Scheer-Cohen, A. R., Strand, E. A., Tilkens, C. M. and Wilson, D. L. (2010) 'Extensions to the Speech Disorders Classification System (SDCS)', *Clinical Linguistics and Phonetics*, 24(10), 795-824.

Shriberg, L. D., Friel-Patti, S., Flipsen Jr, P. and Brown, R. L. (2000) 'Otitis Media, Fluctuant Hearing Loss, and Speech-Language Outcomes: A Preliminary', *Journal of Speech, Language and Hearing Research*, 43(1), 100.

Shriberg, L. D., Gruber, F. A. and Kwiatkowski, J. (1994a) 'Developmental Phonological Disorders III Long-Term Speech-Sound Normalization', *Journal of Speech, Language, and Hearing Research*, 37(5), 1151-1177.

Shriberg, L. D., Kent, R. D., Karlsson, H. B., Mcsweeny, J. L., Nadler, C. J. and Brown, R. L. (2003b) 'A diagnostic marker for speech delay associated with otitis media with effusion: backing of obstruents', *Clinical Linguistics and Phonetics*, 17(7), 529-547.

Shriberg, L. D. and Kwiatkowski, J. (1982) 'Phonological disorders III: a procedure for assessing severity of involvement', *Journal of Speech and Hearing Disorders*, 47(3), 256-70.

Shriberg, L. D. and Kwiatkowski, J. (1994a) 'Developmental Phonological Disorders IA Clinical Profile', *Journal of Speech, Language, and Hearing Research*, 37(5), 1100-1126.

Shriberg, L. D., Kwiatkowski, J. and Gruber, F. A. (1994b) 'Developmental Phonological Disorders II Short-Term Speech-Sound Normalization', *Journal of Speech, Language, and Hearing Research*, 37(5), 1127-1150.

Shriberg, L. D., Lewis, B. A., Tomblin, J. B., McSweeny, J. L., Karlsson, H. B. and Scheer, A. R. (2005) 'Toward Diagnostic and Phenotype Markers for Genetically Transmitted Speech Delay', *Journal of Speech, Language, and Hearing Research*, 48(4), 834-852.

Shriberg, L. D. and Lof, G. L. (1991) 'Reliability studies in broad and narrow phonetic transcription', *Clinical Linguistics and Phonetics*, 5(3), 225-279.

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R. and Moore, C. A. (2009) 'A nonword repetition task for speakers with misarticulations: The Syllable Repetition Task (SRT)', *Journal of Speech, Language, and Hearing Research*, 52(5), 1189-1212.

Shriberg, L. D., Tomblin, B. and McSweeny, J. L. (1999) 'Prevalence of Speech Delay in 6-Year-Old Children and Comorbidity With Language Impairment', *Journal of Speech, Language, and Hearing Research*, 42, 1461-1481.

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

Shuster, L. I. (1998) 'The perception of correctly and incorrectly produced /r/', *Journal of Speech, Language, and Hearing Research*, 41(4), 941-950.

Sices, L., Taylor, H. G., Freebairn, L., Hansen, A. and Lewis, B. (2007) 'Relationship Between Speech-Sound Disorders and Early Literacy Skills in Preschool-Age Children: Impact of Comorbid Language Impairment', *Journal of developmental and behavioral pediatrics*, 28(6), 438-447.

Silverman, F. H. and Paulus, P. G. (1989) 'Peer Reactions to Teenagers Who Substitute /W/ for /R/', *Language, Speech and Hearing Services in Schools*, 20(2), 219-221.

Skahan, S. M., Watson, M. and Lof, G. L. (2007) 'Speech-language pathologists' assessment practices for children with suspected speech sound disorders: Results of a national survey', *American Journal of Speech-Language Pathology*, 16(3), 246-259.

Skebo, C. M., Lewis, B. A., Freebairn, L. A., Tag, J., Ciesla, A. A. and Stein, C. M. (2013) 'Reading Skills of Students With Speech Sound Disorders at Three Stages of Literacy Development', *Language, Speech, and Hearing Services in Schools*, 44(4), 360-373.

Smit, A. B. (1993) 'Phonologic error distributions in the Iowa-Nebraska articulation norms project', *Journal of Speech and Hearing Research*, 36(3), 533.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E. and Bird, A. (1990) 'The Iowa articulation norms project and its Nebraska replication', *Journal of Speech and Hearing Disorders*, 55(4), 779-798.

Smith, A. (2006) 'Speech motor development: Integrating muscles, movements, and linguistic units', *Journal of Communication Disorders*, 39(5), 331-349.

Smith, B. (2006) 'Precautions regarding nonword repetition tasks', *Applied Psycholinguistics*, 27(04), 584-587.

Smith, N. V. (1973) *The acquisition of phonology: A case study*, Cambridge: Cambridge University Press.

Snowling, M., Bishop, D. V. M. and Stothard, S. E. (2000) 'Is Preschool Language Impairment a Risk Factor for Dyslexia in Adolescence?', *Journal of Child Psychology and Psychiatry*, 41(5), 587-600.

Snowling, M. and Hulme, C. (1994) 'The Development of Phonological Skills', *Philosophical Transactions: Biological Sciences*, 346(1315), 21-27.

So, L. K. H. and Dodd, B. J. (1994) 'Phonologically Disordered Cantonese-Speaking Children', *Clinical Linguistics and Phonetics*, 8(3), 235-255.

So, L. K. H. and Dodd, B. J. (1995) 'The acquisition of phonology by Cantonese-speaking children', *Journal of Child Language*, 22(3), 473-495.

Sosa, A. V. (2015) 'Intra-word Variability in Typical Speech Development', *American Journal of Speech-Language Pathology*, Vol. 24(1), 24-35.

Sosa, A. V. and Stoel-Gammon, C. (2006) 'Patterns of intra-word phonological variability during the second year of life', *Journal of Child Language*, 33(01), 31-50.

Stackhouse, J., Vance, M., Pascoe, M. and Wells, B. (2007) *Compendium of auditory and speech tasks*, Chichester: John Wiley and Sons.

Stackhouse, J. and Wells, B. (1997) *Children's Speech and Literacy Difficulties 1: A Psycholinguistic Framework*, London: Whurr.

Stager, C. L. and Werker, J. F. (1997) 'Infants listen for more phonetic detail in speech perception than in word-learning tasks', *Nature*, 388(6640), 381-382.

Stampe, D. (1979) *Dissertation on natural phonology*, Abingdon: Taylor and Francis.

Stoel-Gammon, C. (2011) 'Relationships between lexical and phonological development in young children', *Journal of Child Language*, 38(01), 1-34.

Studdert-Kennedy, M. (1987) 'The phoneme as a perceptuomotor structure', *Haskins Laboratories: Status Report on Speech Research, SR*, 91, 45-57.

Sugden, E., Baker, E., Munro, N. and Williams, A. L. (2016) 'Involvement of parents in intervention for childhood speech sound disorders: a review of the evidence', *International Journal of Language and Communication Disorders*, 51(6), 597-625.

Sutherland, D. and Gillon, G. T. (2007) 'Development of phonological representations and phonological awareness in children with speech impairment', *International Journal of Language and Communication Disorders*, 42(2), 229-250.

Thomas, M. and Karmiloff-Smith, A. (2005) 'Can developmental disorders reveal the component parts of the human language faculty?', *Language Learning and Development*, 1(1), 65-92.

Thomas, M. S., Annaz, D., Ansari, D., Scerif, G., Jarrold, C. and Karmiloff-Smith, A. (2009) 'Using developmental trajectories to understand developmental disorders', *Journal of Speech, Language, and Hearing Research*, 52(2), 336-358.

Thomas, M. S. and Knowland, V. (2009) 'Sensitive periods in brain development: Implications for education policy', *European Psychiatric Review*, 2(1), 17-20.

To, C. K., Cheung, P. S. and Mcleod, S. (2013) 'The impact of extrinsic demographic factors on Cantonese speech acquisition', *Clinical Linguistics and Phonetics*, 27(5), 323-338.

Tomasello, M. (2005) 'Beyond formalities: The case of language acquisition', *The Linguistic Review*, 22(2-4), 183-197.

Topbas, S. (2006) 'Does the speech of Turkish-speaking phonologically disordered children differ from that of children speaking other languages?', *Clinical Linguistics and Phonetics*, 20(7-8), 509-522.

Torrington Eaton, C. and Ratner, N. B. (2016) 'An exploration of the role of executive functions in preschoolers' phonological development', *Clinical Linguistics and Phonetics*, 1-17.

Tyler, A. (2010) 'Subgroups, comorbidity, and treatment implications' in Paul, R. and Flipsen, P., eds., *Speech sound disorders in children: In honor of Lawrence D. Shriberg*, Plural: San Diego, 71-92.

Tyler, A. (2011) 'Foreword to speech sound disorders in children', *Topics in Language Disorders*, 31(2), 93-95.

Tyler, A. A., Figurski, G. R. and Langsdale, T. (1993) 'Relationships between acoustically determined knowledge of stop place and voicing contrasts and phonological treatment progress', *Journal of Speech and Hearing Research*, 36(4), 746-59.

Velleman, S. L. and Vihman, M. (2002) 'Whole-Word Phonology and Templates: Trap, Bootstrap, or Some of Each?', *Language, Speech and Hearing Services in Schools*, 33(1), 9-23.

Veríssimo, A., van Borsel, J. and de Britto Pereira, M. (2011) 'Residual /s/ and /r/ distortions: The perspective of the speaker', *International Journal of Speech-Language Pathology*, 14(2), 183-186.

Vick, J. C., Campbell, T. F., Shriberg, L. D., Green, J. R., Abdi, H., Rusiewicz, H. L., Venkatesh, L. and Moore, C. A. (2012) 'Distinct developmental profiles in typical speech acquisition', *Journal of neurophysiology*, 107(10), 2885-2900.

Vick, J. C., Campbell, T. F., Shriberg, L. D., Green, J. R., Truemper, K., Rusiewicz, H. L. and Moore, C. A. (2014) 'Data-Driven Subclassification of Speech Sound Disorders in Preschool Children', *Journal of Speech, Language and Hearing Research*, 57(6), 2033-2050.

Vihman, M. and Croft, W. (2007) 'Phonological development: toward a "radical" templatic phonology', *Linguistics*, 45(4), 683-725.

Vihman, M. M. (1996) *Phonological development: The origins of language in the child*, Hoboken, NJ: Blackwell Publishing.

Visscher, C., Houwen, S., Moolenaar, B. E. N., Lyons, J. I. M., Scherder, E. J. A. and Hartman, E. (2010) 'Motor proficiency of 6- to 9-year-old children with speech and language problems', *Developmental Medicine and Child Neurology*, 52(11), e254-e258.

Visscher, C., Houwen, S., Scherder, E. J., Moolenaar, B. and Hartman, E. (2007) 'Motor profile of children with developmental speech and language disorders', *Pediatrics*, 120(1), e158-e163.

Vohr, B. (2014) 'Speech and language outcomes of very preterm infants', in *Seminars in Fetal and Neonatal Medicine*, 19, 78-83.

Vought, J. R. and Dean, R. S. (2011) 'Woodcock-Johnson III Tests of Cognitive Abilities' in *Encyclopedia of Child Behavior and Development*, New York: Springer, 1577-1578.

Wake, M., Levickis, P., Tobin, S., Gold, L., Ukoumunne, O. C., Goldfeld, S., Zens, N., Le, H. N., Law, J. and Reilly, S. (2015) 'Two-year outcomes of a population-based intervention for preschool language delay: an RCT', *Pediatrics*, 136(4) e838-e847.

Walley, A., Metsala, J. and Garlock, V. (2003) 'Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability', *Reading and Writing*, 16(1-2), 5-20.

Walsh, B. and Smith, A. (2002) 'Articulatory Movements in Adolescents: Evidence for Protracted Development of Speech Motor Control Processes', *Journal of Speech, Language and Hearing Research*, 45(6), 1119.

Waring, R., Eadie, P., Rickard Liow, S. and Dodd, B. (2016) 'Do children with phonological delay have phonological short-term and phonological working memory deficits?', *Child Language Teaching and Therapy*. 33(1), 33-46.

Waring, R. and Knight, R. (2013) 'How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems', *International Journal of Language and Communication Disorders*, 48(1), 25-40.

Waterson, N. (1976) 'Perception and production in the acquisition of phonology', *Baby talk and infant speech.* Amsterdam: Swets and Zeitlinger, 294-322.

Webster, P. E., Plante, A. M. and Couvillion, L. M. (1997) 'Phonologic impairment and prereading: update on a longitudinal study', *Journal of learning disabilities*, 30(4), 365-375.

Werker, J. F. and Tees, R. C. (1999) 'Influences on infant speech processing: Toward a new synthesis', *Annual review of psychology*, 50(1), 509-535.

Whitehouse, A. J., Robinson, M. and Zubrick, S. R. (2011) 'Late talking and the risk for psychosocial problems during childhood and adolescence', *Pediatrics*, 128(2), e324-e332.

Whitehurst, G. J. and Fischel, J. E. (1994) 'Practitioner Review: Early Developmental language Delay: What. If Anything. Should the Clinician Do About It?', *Journal of Child Psychology and Psychiatry*, 35(4), 613-648.

Wiig, E., Secord, W., Semel, E. (1992) *Clinical Evaluation of Language Fundamentals-Preschool*. San Antonio, TX: Psychological Corp.

Wiig, E., Semel, E. and Secord, W. (2000) *Clinical evaluation of language fundamentals—Preschool (UK)*. San Antonio, TX: The Psychological Corporation'.

Williams, A. L. and Elbert, M. (2003) 'A prospective longitudinal study of phonological development in late talkers', *Language, Speech, and Hearing Services in Schools*, 34(2), 138-153.

Wolke, D., Samara, M., Bracewell, M. and Marlow, N. (2008) 'Specific Language Difficulties and School Achievement in Children Born at 25 Weeks of Gestation or Less', *The Journal of Pediatrics*, 152(2), 256-262.e1.

Wren, Y. (2015) "“He'll Grow out of it Soon—Won't He?”—The Characteristics of Older Children's Speech When They Do—and Don't—Grow out of it', *SIG 16 Perspectives on School-Based Issues*, 16(2), 25-36.

Wren, Y., McLeod, S., White, P., Miller, L. L. and Roulstone, S. (2013) 'Speech characteristics of 8-year-old children: findings from a prospective population study', *Journal of Communication Disorders*, 46(1), 53-69.

Wren, Y., Miller, L. L., Peters, T. J., Emond, A. and Roulstone, S. (2016) 'Prevalence and Predictors of Persistent Speech Sound Disorder at Eight Years Old: Findings From a Population Cohort Study', *Journal of Speech, Language, and Hearing Research*, 1-27.

Wren, Y. E., Roulstone, S. E. and Miller, L. L. (2012) 'Distinguishing groups of children with persistent speech disorder: Findings from a prospective population study', *Logopedics Phoniatrics Vocology*, 37(1), 1-10.

Wright, A., James, E., and Franklin, S. (2017) 'Persistent speech disorder in adolescence: a treatment case study.' in Dodd, B. and Morgan, A., eds., *Intervention Case Studies of Child Speech Impairment*, Guildford: J&R Press. 337-357.

Yang, H.-C. and Gray, S. (2017) 'Executive function in preschoolers with primary language impairment', *Journal of Speech, Language, and Hearing Research*, 60(2), 379-392.

Young, A. R., Beitchman, J. H., Johnson, C., Douglas, L., Atkinson, L., Escobar, M. and Wilson, B. (2002) 'Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children', *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43(5), 635-645.

Zamuner, T. (2004) *Input-based phonological acquisition*, London: Routledge.

Zhu, H. and Dodd, B. (2000) 'Putonghua (modern standard Chinese)-speaking children with speech disorder', *Clinical Linguistics and Phonetics*, 14(3), 165-191.

Ziegler, W. (2008) 'Neurophonetics', In Ball, M.J., Perkins, M.R., Müller, N. and Howard, S. eds., *The handbook of clinical linguistics*. Oxford: Blackwell, 491-505

8 Appendices

8.1 Appendix 1-A Table of longitudinal studies: characteristics of the samples

Appendix 1-A Table of longitudinal studies 1: Characteristics of the samples									
Study	**Type of sample**	**Aim**	**Age 1**	**Age 2**	**Inc/ex criteria**	**Number/ Gender**	**Interval**	**Initial Severity**	**Speech measure**
Bishop & Edmundson 1987	Clinical	Investigate the range of severity of S&L difficulties at age 4, and outcomes at 5;6	3;9-4;2	5;6	LI of unknown origin referred to SLT services	88, 72M 16F, 9 with SSD only.	18 mths	mixed	PCC single word
Bishop & Adams 1990	Clinical	Investigate language and literacy outcomes at 8;6 in children with LI at 4 years	5;6	8;6	children from Bishop & Edmundson 1987 with only phonological difficulties by 5;6	6, gender not reported	2.5 yrs	mixed	PCC single word
Shriberg, Gruber, & Kwiatkowski, 1994: Long term	Clinical	Characterize sequence, rates, and error patterns of long-term speech-sound normalization	3;8-5;4	8;11-11;4	not resolved by 6	9M 1F	5-6 yrs	mild-moderate to severe	PCC single word
Shriberg, Kwiatkowski, & Gruber, 1994: Short term	Clinical	Investigate predictors of resolution after one year	3;8-5;4, mean 4;3	mean 5;3	children with SSD severe enough to interfere with intelligibility	64: 64% M	1 year	mild-moderate to severe	PCC single word
Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996	Population	Examine developmental and academic outcomes of speech/language-impaired children	5 yrs	12;6	random sample, speech impaired category included children with voice & fluency issues	85 M 39 F	7 yrs	mixed speech and language impairments	Raw error totals on Photo Articulation Test

Study	Type of sample	Aim	Age 1	Age 2	Inc/ex criteria	Number/ Gender	Interval	Initial Severity	Speech measure
Webster, Plante, & Couvillion, 1997	Population	Investigate relationship between phonological ability, syntax, PA and reading	3;6	6;0	children with SSD, normal receptive language	29, gender not reported	2.5 years	moderate-severe	PCC single word and processes
Lewis, Freebairn & Taylor, 2000	Clinical	Compare language/academic outcomes for SSD alone compared to SSD+LI	4 to 6 yrs	8 to 10 yrs	>-1SD on Goldman Fristoe + 3 error patterns. PIQ>80, LI = <8 on 2 subtests of TOLD-P	28x SSD: 64% M, 24x SSD+LI: 63% M	4 yrs	mixed. SSD+LI significantly more severe SSD than SSD alone	PCC single word
Glogowska, Roulstone, Enderby & Peters, 2000	Clinical	to investigate the efficacy of therapy for pre-schoolers with SSD and/or LI	<3;6	1 yr later	children referred to SLT services. SSD = 40%+ errors on velars, fricatives or postvocalic singletons + therapists judgement LI = <10 percentile on expressive or receptive subtests of PLS	71 in therapy (77% M), 84 no therapy (74% M)	I year		Error rate single word + therapists judgement
Roulstone, Peters, Glogowska, & Enderby, 2003	Clinical	Investigate natural history of pre-schoolers with S&L delay without Tx	<3;6	1 year later	as Glogowska above	54M, 15F, 12 SSD only	1 year	mixed	Error rate single word + therapists judgement

Study	Type of sample	Aim	Age 1	Age 2	Inc/ex criteria	Number/ Gender	Interval	Initial Severity	Speech measure
Williams and Elbert, 2003	Clinical	Identify possible predictors of delayed versus deviant development.	2x 22 mths, 3x 30-31 mths		late talkers in absence of other conditions.	3M 2F	10-12 mths	< 12 sounds at 22 mths, <15 sounds at 31 mths. Lexicon<50 words	PCC, inventory, variability, error patterns, continuous speech & single word
Nathan, Stackhouse, Goulandris, & Snowling, 2004a/ Stackhouse, Vance, Pascoe, & Wells, 2007	Clinical	Identify predictors of persisting speech difficulties and associated literacy problems	4;6	6;8	SSD of unknown cause, 1 SD below mean, normal NVIQ. LI: < 10th centile on 2 receptive or expressive language measures.	47: 28 SSD only, 19 SSD+LI. 31 M 16 F	2;2 years	mixed	PCC single word; word repetition and NWR
Hesketh, 2004	Clinical	Assess literacy and PA outcomes in children with a history of SSD, and to identify predictors	3;6 to 5;0	6;6 to 7;6	SSD, no severe receptive language deficit, normal PIQ	24 M, 11F	2;6 to 3 yrs.	moderate to severe	PCC single word
Leitão & Fletcher, 2004	Clinical	Examine phonological processing and literacy skills of children with SSD only, comparing those with typical and atypical errors	5;4 to 6;7	12 to 13 yrs	SSD. Language within -1.5 SD. Normal PIQ. 10% or more of atypical process use = atypical group	14, gender not reported	7 yrs.	mixed	PCC on GFTA, typical and atypical processes

Study	Type of sample	Aim	Age 1	Age 2	Inc/ex criteria	Number/ Gender	Interval	Initial Severity	Speech measure
Nathan, Stackhouse, Goulandris, & Snowling, 2004b	Clinical	To examine academic outcomes in children with a history of SSD, contrasting those with resolved and persisting SSD	4;6	6;7	SSD, = 1 SD below mean, normal NVIQ. 38.5% had co-morbid LI. 23.4% had special educational needs.	26M 13F	2 yrs.	Mixed	PCC single word; word repetition and NWR
Bernhardt & Major 2005	Clinical	Investigate predictors and outcomes in speech, language and literacy	3;6 to 4;11	6;1 to 8;5	SSD, moderate to severe, normal hearing and oro-motor function	8 M 4 F (11 SSD+LI)	3 yrs.	11 severe, 1 moderate	PCC single word
Broomfield & Dodd 2005/ Broomfield 2011	Clinical	Investigate efficacy of therapy in a clinical sample	272 aged 0-6 yrs. 48 > 6 yrs.	6 mths older	1SD below mean, or profile of clinical symptoms. LI not excluded	211 in therapy, 101 no therapy, 63% M	6 mths	Mixed	PCC single word
Lewis, Freebairn, Hansen, Stein, Shriberg, Iyengar, & Taylor, 2006	Clinical	Factor analysis of S&L measures	3-7 yrs., mean 5.2 yrs.	9.1	SSD, normal IQ, hearing, speech mechanism, no neurological or developmental difficulties	26 M 12 F	4 yrs.	Moderate to severe SSD	PCC single word and conversation
Rvachew 2006	Clinical	To find predictors of poor PA skills in pre-schoolers with SSD	4;0-5;7, mean 4;10	5;1 to 6;7, mean 5;10	SSD < 16th percentile on a standardized assessment. LI & CAS not excluded)	28 M, 19 F, 37 SSD at IAx, 10 already resolved	1 yr.		PCC single word

Study	Type of sample	Aim	Age 1	Age 2	Inc/ex criteria	Number/ Gender	Interval	Initial Severity	Speech measure
Paul, 2006	Clinical	Investigate the outcomes of toddlers with slow expressive language development	1;8 to 2;10	Kinder garden	< 10th percentile on Rescorla's 1989 Language Development Survey	31, gender not reported	3 to 4 years	mixed	PCC connected speech
Rvachew, Pi-Yu, & Evans, 2007	Clinical	Examine relationship between typical versus atypical speech errors and PA skills	4;6	5;6	SSD, LI and CAS not excluded. normal hearing and oral-motor function	58, gender not reported	1 year	moderate	PCC single word
Rvachew, 2007	Clinical	Examine the predictive relationship between speech perception and PA at 4-5 years and reading at 6-7 years	4;5 to 5;7	6;2 to 7;9	SSD, LI and CAS not excluded. normal hearing and oral-motor function	20 M, 13 F in SSD group, and 35 TD controls	1.5 to 2 years	mixed	PCC single word
Roulstone, Peters, Glogowska & Enderby, 2008	Clinical	Predictors and outcomes of SLTs' decisions to offer intervention	<3;6	7 to 9 yrs	referred to local SLT services for speech or language concerns	191 , gender not reported	4 to 5 years	mixed	PCC single word
McIntosh & Dodd 2008	Clinical	Investigate whether errors at 2 predicted SSD at 3	2;1 to 2;11	3	SSD > 1SD below mean on normative sample	2F, 8M 5 with SSD 5 TD controls	6-11 months	mixed	PCC single word

Study	Type of sample	Aim	Age 1	Age 2	Inc/ex criteria	Number/ Gender	Interval	Initial Severity	Speech measure
Roustone, Miller, Wren, & Peters, 2009	Population	To determine whether error patterns at 2 and 5 years predict SSD at 8 years	2 yrs.	8 yrs.	SSD defined as >1 perceptible error in a connected speech sample: gave prevalence of 18%	79 M 53 F	2 intervals of 3 years each	mixed	At 2 and 5 yrs., no.of errors in single words. At 8 years no. of errors in connected speech
Young, Beitchman, Johnson, Douglas, Atkinson, Escobar, & Wilson,2002	Population	Investigate outcomes of children at 5 with SSD and/or LI at age 19	5 years	19 years		35 SSD only (74% M), 25 SD+LI (44% M), 49 LI only (65% M), 120 controls	14 years	mixed	errors on single word test
Preston et al. 2013	Clinical	Determine whether speech error patterns in pre-schoolers with SSD predict persisting SSD and PA.	4;6	8;3	<89 on GFTA2, average receptive language, average PIQ	18 M, 7 F	3;9	range 50-87 on GFTA2	PCC single word

Study	Type of sample	Aim	Age 1	Age 2	Inc/ex criteria	Number/ Gender	Interval	Initial Severity	Speech measure
Eadie, Morgan, Ukoumunne, Ttofari Eecen, Wake, & Reilly, 2015	Population	Determine prevalence of idiopathic SSD, comorbidity of SSD with language and pre-literacy difficulties; and predictors of speech outcome at 4 years.	birth	4 years	children with disabilities known to affect language excluded. SSD defined as <80 on GFTA2. LI=>1.25SD below mean on expressive/receptive composite	1494. 51 with SSD: 25M 26F	up to 4 years	mixed	PCC single word
Lewis Freebairn, Tag, J., Ciesla, Iyengar, Stein, & Taylor, 2015	Clinical	Determine predictors and outcomes for speech, language, and literacy for individuals with histories of SSD and SSD+LI	4 to 6	11 to 18	SSSD < 10th percentile on GFTA +3 phonological process error types on Khan-Lewis Phonological Analysis. Normal PIQ, < 6 episodes of OME. LI=scaled score of 8 or under on 2 subtests of CELF-PS or TOLD-P	170, 77 SSD (50 M), 93 SSD+LI (67 M)	7+ years	moderate to severe	PCC single word
Hayiou-Thomas, Carroll, Leavett, Hulme, & Snowling 2016		Investigate the role of early SSD in literacy development, in the context of additional risk factors	3;6	8;6	>-2SD on DEAP excluded: MZ twinning, chronic illness, deafness, English as a second language, in care, psychiatric disorder	245	2 and 5 years	severe	PCC single word

Study	Type of sample	Aim	Age 1	Age 2	Inc/ex criteria	Number/ Gender	Interval	Initial Severity	Speech measure
Morgan, Eecen, Pezic, Brommeyer, Mei, Eadie, Reilly & Dodd, 2017	Population	Determine predictors of speech outcome at 7 years, for children with speech errors at age 4.	4	7	Recruited from ELVS population sample	93 with SSD, 54 phonological delay, 39 disorder. Gender not reported .	3 years	mixed	PCC single word

8.2 Appendix 1-B Table of Predictors

Appendix 1-B Table of longitudinal studies 2: Predictors of transient vs persistent SSD

(where factors are predictors, a Y+ means more likely to resolve the SSD, Y- means less likely to resolve)

Study	Male Gender	Positive Family History	SES	Motor Skills Delay	Birth Difficulties	Feeding / sucking/ drooling issues	OME History /hearing issues	Age	Poor language/ vocab	Severity of SSD	Receipt of therapy	Error type	Poor input phonology	Low DDK	Late talker	Low PIQ
Bishop & Edmundson 1987									Y-	N	N					
Bishop & Adams 1990	isolated phonological problems in pre-literate children are not predictive of reading difficulties															
Shriberg, Gruber, & Kwiatkowski, 1994: Long term						normalisation not likely after 8;6		Y		Y	moderate to severe SSD predicts residual distortion errors					
Shriberg, Kwiatkowski, & Gruber, 1994: Short term	N	N	N[1]	N	N	N	N	N	N	N	N	Y- omission errors		N	Y+	
Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996			N[2]	low ability overall on a set of speech, language and cognitive measures predicted worse outcomes for later linguistic, cognitive and academic performance. Isolated speech difficulty had low risk for persisting difficulties.												
Webster, Plante, & Couvillion, 1997	Speech accuracy at 3;6 predicted working memory at 5;6. Concurrent PCC predicted letter knowledge at 5;6 better than PCC at 3;6 did.															
Lewis, Freebairn & Taylor, 2000	SSD+LI had worse academic outcomes than SSD alone.															

Study	Male Gender	Positive Family History	SES	Motor Skills Delay	Birth Difficulties	Feeding / sucking/ drooling issues	OME History /hearing issues	Age	Poor language /vocab	Severity of SSD	Receipt of therapy	Error type	Poor Input phonology	Low DDK	Late talker	Low PIQ
Glogowska, Roulstone, Enderby & Peters, 2000											Y+					
Roulstone, Peters, Glogowska, & Enderby, 2003		N	N[1]							Y-	Measured by TOMS					
Williams and Elbert, 2003							Y-			Y-		Y-	atypical errors			
Nathan, Stackhouse, Goulandris, & Snowling, 2004a/ Stackhouse, Vance, Pascoe, & Wells, 2007									Y-				Y-			Y- Con-current
Hesketh, 2004	As a group, children with SSD alone were not at increased risk of a delay in literacy or PA. However, within the group there was a wide range of performance. PA skills were the best predictor of single-word literacy outcome.															
Leitão & Fletcher, 2004												Y-	atypical errors			
Nathan, Stackhouse, Goulandris, & Snowling, 2004b									Y-							
Bernhardt & Major 2005										N						

Study	Male Gender	Positive Family History	SES	Motor Skills Delay	Birth Difficulties	Feeding / sucking/ drooling issues	OME History /hearing issues	Age	Poor language/ vocab	Severity of SSD	Receipt of therapy	Error type	Poor input phonology	Low DDK	Late talker	Con-current PIQ
Broomfield & Dodd 2005/ Broomfield 2011		Y-			Y+		Y+		N		Y+	Y- atypical errors			Y-	
Lewis, Freebairn, Hansen, Stein, Shriberg, Iyengar, & Taylor, 2006										Y-						
Rvachew 2006										N			Y-			
Paul, 2006											N					
Rvachew, Pi-Yu, & Evans, 2007												N				
Rvachew, 2007												N				
Roulstone, Peters, Glogowska & Enderby, 2008												Error type did not predict PA				
McIntosh & Dodd 2008												Y- atypical errors				
Roustone, Miller, Wren, & Peters, 2009								Y		Y- at 5 yrs. but not at 2 yrs.						

Study	Male Gender	Positive Family History	Low SES	Motor Skills Delay	Adverse Birth Circumstances	Feeding / sucking/ drooling issues	OME History /hearing issues	Age	Language skills/ vocab	Severity of SSD	Receipt of therapy	Error type	Poor input phonology	Low DDK	Late talker	Low PIQ
Young, Beitchman, Johnson, Douglas, Atkinson, Escobar, & Wilson, 2002	At 5, there no differences between SSD only and controls on any demographic variable, and no differences between SSD+LI and LI only. LI groups had lower SES, VIQ and PIQ. SSD only and control groups were generally quite similar by age 19, clear and pervasive differences were evident between LI and non-LI groups.															
Preston et al. 2013												Y- distortion but not atypical errors predicted PSSD				
Eadie, Morgan, Ukoumunne, Ttofari Eecen, Wake, & Reilly, 2015	Y-	Y-	Y-	Y-	N										Y+	
Lewis Freebairn, Tag, J., Ciesla, Iyengar, Stein, & Taylor, 2015			Y-						Y-				Y- NWR	Y-		Y-
Hayiou-Thomas, Carroll, Leavett, Hulme, & Snowling 2016								Y				Y/N atypical errors are associated with poorer word reading at age 5½, but not at age 8				
Morgan, Eecen, Pezic, Brommeyer, Mei, Eadie, Reilly & Dodd, 2017	N	N	N[3]						N	N	N	Y- atypical errors				N

[1] maternal/parental education, [2] parental occupation, [3] Socioeconomic Indexes for Areas (SEIFA, Australia), [4] Hollingshead Four Factor Index of Social Class, [5] parents not homeowners.

Concurrent predictors of persisting SSD at age 8, compared to TD controls

Study	Male Gender	Positive Family History	Low SES	Motor Skills Delay	Birth Difficulties	Feeding / sucking/ drooling issues	OME History /hearing issues	Age	Poor language/ vocab	Severity of SSD	Receipt of therapy	Error type	Poor input phonology	Low DDK	Late talker	PIQ
Wren, Roulstone & Miller, 2012	Y-		Y[1-]										poor NWR	Y-		Y-
Wren, Miller, Peters, Emond, & Roulstone, 2016	Y-	N	Y[5-]	Y-		Y-	Y-						poor NWR		Y-	

[1] maternal/parental education, [2] parental occupation, [3] Socioeconomic Indexes for Areas (SEIFA, Australia), [4] Hollingshead Four Factor Index of Social Class, [5] parents not homeowners.

8.3 Appendix 2-A Novel Tests

8.3.1 Test of Non-Word Discrimination

Total Correct: 1 syll. _____ 3 syll:_____total_____

	Production 1	Production 2	Feature diff.	Score
1.	mɛp	mɛp		S D
2.	tɪfəlɒn	tɪsəlɒn	P	S D
3.	tʌz	tʌd	M	S D
4.	gʌməlʌm	gʌməlʌm		S D
5.	nɪg	nɪd	P	S D
6.	taz	taz		S D
7.	sɪləbɔn	sɪləbɔn		S D
8.	vɔzəlm	vɔdəlɪn	M	S D
9.	tɛməti	tɛməti		S D
10.	dat	dat		S D
11.	dɪzalʌp	dɪzalʌp		S D
12.	fas	faʃ	P	S D
13.	sɪdʒali	sɪdali	Aff	S D
14.	bɛnɪsən	bɛnɪsən		S D
15.	tɛp	tɛb	V	S D
16.	mɛtiman	mɛkiman	P	S D

Administration:

Each slide contains a video clip of Girl 1 and Girl 2, side by side

1. Administrator clicks to play clip: Girl 1 says production 1, then Girl 2 says production 2
2. Say: "look at these two girls. This one (indicate left hand girl) is going to say a funny word, and the other girl has to say it after her. She has to say it just the same. I want you to tell me if she gets it right."
3. Play clip. Say: "did she get it right?"
4. Ring S if the child thought she said it the same, D if he thought she said it differently. Answers in bold are correct.
5. Final score is number of correct answers.

8.3.2 Test of phonological legality: stimulus words

1.	slaʃi	svaʃi
2.	gjʌpoʊ	kjʌpoʊ
3.	θrɪpi	θlɪpi
4.	frɪŋi	fnɪŋi
5.	ʃtʌli	stʌli
6.	ʃroʊpi	ʃloʊpi
7.	frʌdi	vrʌdi
8.	zlɛti	slɛti
9.	stʌtoʊ	srʌtoʊ
10.	vlɛti	flɛti
11.	ʒroʊpi	ʃroʊpi
12.	zmɛbi	smɛbi

8.3.3 Test of pattern recognition

Test of pattern recognition. Instructions:

The test is presented on computer. Each page is one test item. The task is to select the shape from the bottom line that will complete the pattern on the top line.

Administration:

1. Go to the Trial Item page. Say "Look at these circles. They make a pattern. Blue, yellow, blue, yellow, blue, yellow (run your finger below shapes as you talk). Now we need to finish the pattern. What comes next? Which of these (point to circles below) do we need to put here (indicate empty box) to make the pattern?"
2. Child should point to a circle. Grab the chosen circle with the mouse and put it in the empty box. Ask "is that right?" If child indicates it is not right, let him have another go. Give appropriate feedback till he gets it right.
3. Repeat with next trial item. Say "Look at these squares. They make a pattern. Big, small, big, small, big, small. Now we need to finish the pattern. What comes next? Which of these (point to shapes below) do we need to put here (indicate empty box) to make the pattern?"
4. Repeat as in 2 above
5. Scroll down to Test items. Say "Now we need to make this pattern. Which of these shapes (indicate) do we need to put in this box to make the pattern right?"
6. Child points to a shape. Grab the chosen shape with the mouse and put it in the empty box. Ask "is that right?" If child indicates it is not right, let him have another go until he indicates it is right (whether it is or not).
7. Move to next test item
8. Final score is the number of items he got correct.

8.3.4 Test of Numerical Concepts

Case Number		Date		
Ability	**Equipment**	**Task**		**Result**
Matching, 1 to 1 correspondence	5 knives, 5 forks, thrown into a heap on the table	***Can you do this?*** Demo putting knife and fork together Child must put all knives and forks together		1 0
Matching similar items even set = <5 items	2 of each, cow, dog, horse, pig, sheep, thrown into a heap on the table	***Can you match these? I'll do the first one*** Put one pair of animals together on the table. Child must pair all animals correctly		1 0
Matching similar items odd set = <5 items	2 of each, horse, sheep, 3 each of dog, pig, cow, thrown into a heap on the table	***Now, can you match these?*** Child has to sort animals into pairs or groups of the same species		1 0
Class inclusion	4 green card fields, with 2 animals in each field, e.g. 2 cows, 2 pigs, 2 horses, 2 sheep.	Offer 2 more of each animal ***Put these where they go:*** Child has to put them in the field with the same species		1 0
Class exclusion	As above with all animals now in the correct fields. One dog.	Tell child to close their eyes. Put the dog into one of the fields. ***Now look, which one doesn't belong?*** Child must indicate/remove the dog		1 0
Ordering	large medium and small cows & horses to demo	***Look, I'm putting my horses in a row.*** Put horses in a row according to size, big to small, one behind the other. ***Can you put these cows in a row like that?*** Child must order cows big to small	orders 3 objects in random order orders 3 objects by trial and error orders 3 objects in a systematic manner	1 2 3
Conservation task 1	2 green card fields, 10 identical animals, 4 in one field, 6 in other, lined up with equal spacing	***Which field has more?*** Child must indicate field with six cows		1 0
Conservation task 2	same 4 animals in one field spread out, 6 in the other squashed together	Tell child to close their eyes. Change spacing so the four cows take up more of their field than the six cows. ***Now which field has more?*** Child must indicate field with six cows		1 0
Total				

8.4 Appendix 4-A: Description of the procedure undertaken in the factor analysis.

Factor analysis is a means to undertake structure analysis of a construct, including a complex array of structure analysing procedures that are used to identify the interrelationships among a large set of observed variables. Through data reduction a smaller set of these variables are then grouped into factors that have common characteristics (Nunnally and Bernstein 1994). Each factor is a linear combination or cluster of related observed variables that represents a specific underlying dimension of a construct, which is as distinct as possible from the other factors included in the solution (Tabachnick and Fidell 2001). Where clusters of large correlation coefficients between subsets of variables occur, it suggests that those variables could be measuring aspects of the same underlying dimension, a latent variant, or factor. Reducing a data set from a group of interrelated variables to a smaller set of factors enables the maximum amount of common variance in a correlation matrix using the smallest number of explanatory constructs (Field 2009), thus revealing the structure of the construct, here speech accuracy.

Once the internal structure of a construct has been established, factor analysis may also be used to identify external variables, e.g. gender, SES, that appear to relate to the various dimensions of the construct of interest (Nunnally and Bernstein 1994).

8.4.1 Determining the factor extraction method

The term "factor analysis" encompasses a variety of different techniques (Pallant 2010), so it is necessary to determine the extraction method most appropriate to the aims of the analysis. The main divide is between Principal Components Analysis (PCA) and "true" or *Common factor analysis* (CFA) (Pett *et al.* 2003). Both procedures attempt to produce a smaller number of linear combinations of the original variables that can account for most of the variability in the pattern of correlations (Pallant, 2010).

Factor analysis is based on the belief that measurable and observable variables can be reduced to fewer latent variables that share a common variance but are unobservable. (Bartholomew *et al.* 2011). As an empirical method of analysis, that provides a data-driven method for investigating separable dimensions of abilities and disabilities, factor analysis, and exploratory factor analysis in particular, thus appeared to be an appropriate method to explore the interrelationships among the set of variables to enable

the determination of the presence and nature of latent constructs, such that a deficit in one or more of these constructs could be causally related to deficits in the acquisition speech. Exploratory factor analysis is an appropriate method in circumstances like those of this study, where the researcher does not know how many factors are necessary to explain the interrelationships among a set of items (Pett *et al.* 2003), as few restrictions are placed on the patterns of relations between the common factors and the measured variables, and no a priori number of common factors is specified.

In this analysis, CFA is used for the exploratory factor analysis in order to reveal the structure of the construct of speech disorder by identifying underlying factors in a battery of assessments that may contribute to it. CFA assumes an underlying causal structure, where the covariation in the observed variables is due to the presence of one or more latent variables (factors) that exert causal influence on these observed variables, and is thus used to identify the number and nature of these latent factors.

8.4.2 Selection of Common Factor Analysis extraction procedure

Principal axis factoring (PAF) was chosen as the extraction procedure for the CFA analysis, as it is the method most commonly used for exploratory, as opposed to confirmatory, factor analysis. Principal factors methods have the advantage of entailing no distributional assumptions and are less likely than other procedures, such as Maximum Likelihood, to produce Heywood cases or a solution that fails to converge on a final set of parameter estimates (Fabrigar et al, 1999). Principal axis factoring, a type of EFA, is superior to principal components analysis as it analyses common variance only, which is a key requirement for theory development. In addition to this, it is a useful technique for identifying items that do not measure an intended factor or that simultaneously measure multiple factors (Worthington & Whittaker, 2006).

8.4.3 Suitability of the data for factor analysis.

The two issues to consider in determining whether a data set is suitable for factor analysis are the strength of the relationship between the variables, and sample size.

Following selection of the variables, an initial Correlation Matrix was generated to inspect the relationships between the variables, in order to determine whether the matrix was factorable. Inspection of the Matrix showed that all variables had an adequate correlation (Pearson's r >.30) with at least one other item (too little shared common variance can lead to the isolation of too many factors with poor loadings). There were

numerous coefficients greater than .30, as recommended by Tabachnik & Fidell (2007), indicating that the matrix was factorable.

There is wide variation in the literature about the necessary sample size for factor analysis. Field recommends 10-15 subjects per variable (2009 p647). Gorsuch (1983) p.332) and Hatcher (1994) p. 73) recommend a minimum subject to item ratio of at least 5:1. Monte Carlo simulations have been used to model the effect of different participant to variable ratios using simulated data. Guadagnoli and Velicer (1988) concluded that the most important issues in determining reliable factor solutions were the absolute sample size and the absolute magnitude of factor loadings: the more frequent and higher the loadings are on a factor, the smaller the sample needed. They argue that if a factor has 4 or more loadings greater than 0.6 it is reliable regardless of sample size. The stronger (i.e. uniformly high communalities, no cross loadings, several variables loading strongly on each factor) the data, the smaller the sample can be for an accurate analysis.

Two further statistical measures of sample adequacy are offered by SPSS. The Kaiser-Meyer-Olkin value was .754, thus in the 'good' range (Field 2009 p647). Bartlett's test of sphericity reached significance, showing this was not an identity matrix. Both tests thus support the factorability of the correlation matrix.

8.4.4 Selecting the rotation

The goal of rotation is to simplify and clarify the data structure. Orthogonal rotations produce factors that are uncorrelated; oblique methods allow the factors to correlate. Although orthogonal rotations seem to be those most commonly used, many reviewers support the use of oblique rotations (Gorsuch 1983, Fabrigar *et al.* 1999, Conway and Huffcutt 2003). Conway & Huffcutt (2003) state that if factors are actually correlated, then orthogonal rotation forces an unrealistic solution that may distort loadings away from simple structure, whereas an oblique rotation will better represent reality and produce better simple structure. In cases where factors are poorly correlated or uncorrelated, an orthogonal rotation would be appropriate, but this will not indicate whether factors are actually correlated. They hold that an oblique rotation will still give loadings that are very similar to those from an orthogonal rotation, plus a factor correlation of about zero. Therefore, they concluded that there is no reason to use an orthogonal rotation as oblique rotations sometimes produce better solutions (i.e., when

the constructs are in reality correlated) and the rest of the time essentially equivalent solutions (i.e., when constructs are really uncorrelated or nearly so).

Fabrigar et al. (1999) suggest first examining solutions produced by one or more of the common methods of oblique rotation; if this indicates that the factors are uncorrelated, a varimax rotation can then be conducted. However, if at least some of the factors are found to be correlated with one another, they suggest it is most defensible to use the oblique solution as the basis for interpretation. Fabrigar's suggested method will be used in this study.

8.4.5 Selecting the number of factors to retain

Fabrigar et al., (1999) state that the objective when deciding the number of factors to retain, is "to select a model that explains the data substantially better than simpler alternative models (i.e., models with fewer factors) but does as well or nearly as well as more complex alternative models (i.e., models with more factors)".

However, no single technique has been shown to be consistently highly accurate in pinpointing the most accurate number of factors to retain (Conway & Huffcutt, 2003). Ford *et al.* (1986) and Fabrigar et al. (1999) both recommend using a combination of techniques to make this decision. Fabrigar et al. also noted that choice of number of factors is a substantive issue as well as a statistical one, because a solution that is not able to be interpreted is not helpful. Therefore, Conway & Huffcutt (2003) recommend that the combination of techniques should include examination of multiple solutions with different numbers of factors, for interpretability.

8.4.6 Principal Axis Factoring

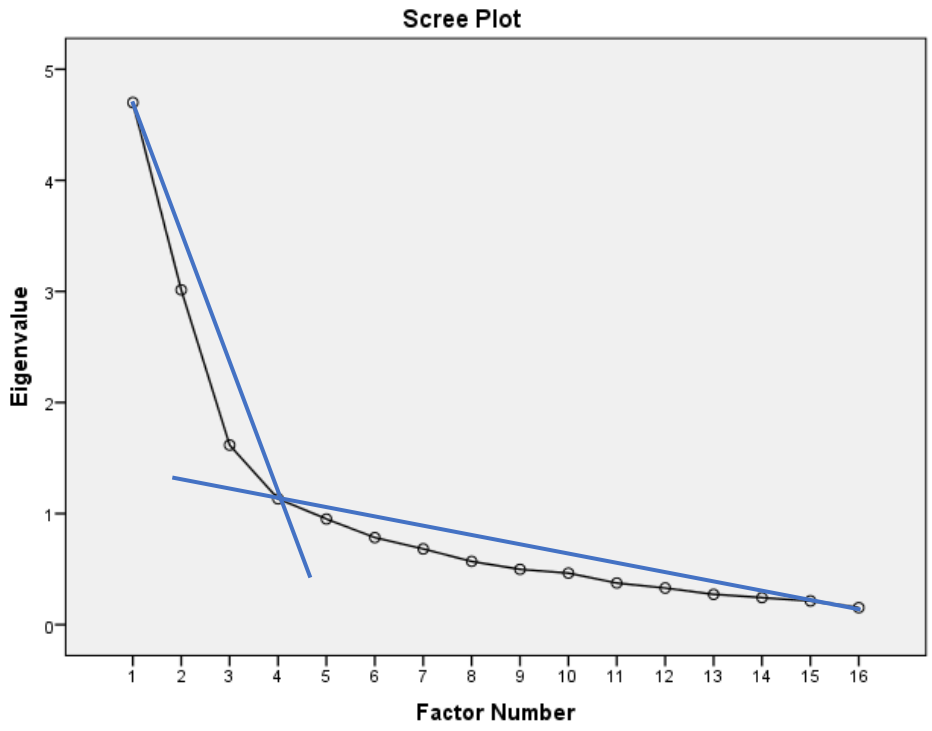
Principal Axis Factoring (PAF) was used to confirm the presence and nature of any latent constructs underlying these variables that could be causally related to deficits in speech.

Scores on the study test battery for all children, including the 11 with age appropriate speech at initial assessment, were entered into the analysis. All variables derived from standardised norm referenced tests were converted into Z-scores. For all the other variables, to remove age-related variance, raw scores were regressed along age, with standardised residuals used so that all scores were on the same scale and standardised to the same mean.

Principal axis factoring initially revealed the presence of 5 components with Eigenvalues greater than 1, explaining 27.5%, 17.4%, 9.0%, 6.6% and 5.9% of the variance respectively. However, inspection of the Structure Matrix revealed that 2 variables, DDK and Pattern recognition, did not load strongly on any factor, so they were removed and the analysis repeated. This then revealed the presence of four components with Eigenvalues greater than 1, explaining 29.4%, 18.8%, 10.1%, 7.1% of the variance respectively. These components explained 65.4% of the variance altogether.

Inspection of the scree plot in Figure 8-1 confirmed a clear break after the fourth component, confirming the selection of 4 factors.

Figure 8-1: Scree plot for PAF, showing break after 4th component:



To aid in the interpretation of these components, an Oblimin rotation was performed. An oblique rotation generates two different matrices: the factor structure matrix and the factor pattern matrix. The pattern factor matrix contains the loadings that represent the unique relationship of each item to a factor, having first controlled for the correlation among the factors; thus the magnitudes of their loadings is influenced by the strength of the correlations among the factors. In the factor structure matrix, in contrast, the loadings are simple zero order correlations of the items with the factors. The magnitudes

of these loadings are not influenced by the strength of the correlations among factors (Pett et al., 2003).

As recommended by Pett et al., the analysis here will focus on the factor structure matrix first to decide interpretability, and then compare the decisions made with the factor pattern matrix.

Observation of the structure matrix showed two variables loading weakly (below .5) on any factor: Auditory Memory and Number Concepts. The weakest, Number Concepts, was removed. After reanalysis this showed a stronger model explaining 67.4% of the variance (see Figure 8-2). The Oblimin rotation (see Figure 8-3) then revealed fewer cross loadings on the structure matrix with all variables loading at least .4 on one factor

This four factor solution explained 67.4% of the variance. The first factor included pattern recognition and rule abstraction tasks, as well as phonological processing abilities of non-word discrimination and auditory memory. It explained 29.5% of the variance, and was named COGNITIVE. The second factor comprised measures of speech output accuracy. It explained 20.1% of the variance and was named SPEECH. The third factor comprised measures of non-speech oral motor skills and visual motor integration, and was named MOTOR. It explained 10.3% of the variance. The last factor comprised measures of language ability, explained 7.5% of the variance and was named LANGUAGE.

Figure 8-2: Structure matrix

Structure Matrix

	Factor			
	1	2	3	4
R_FIST1	.894	.133	.397	.263
R_FIST2	.662	.079	.278	.301
R_NWD	.637	.037	.153	.392
Zscore: CELF Word	.601	.206	.380	.554
R_Auditory Memory	.492	-.361	.170	.490
R_PCC1	.058	.920	.089	-.001
R_MultisyllSp	.322	.775	.247	-.009
Zscore: DEAP SS	.010	.731	.008	.053
R_MultisyllIm	.058	.647	.079	-.261
Zscore: SM SS	.226	.127	.683	-.237
Zscore: IM SS	.174	.076	.603	.128
Zscore: BVMI SS	.402	-.035	.530	.385
Zscore: CELF Ex vocab	.449	-.162	.134	.817
Zscore: CELF SenStr	.532	.051	.121	.706
Zscore: BPVS SS	.398	.012	.468	.592

Extraction Method: Principal Axis Factoring.

Rotation Method: Oblimin with Kaiser Normalization.

After controlling for correlations among the factors, the Pattern Matrix (Figure 8-3) showed only one variable that did not clearly load onto a single factor, Auditory Memory. It factored together positively with Non-Word Discrimination, another phonological processing ability, but surprisingly negatively with speech output measures. As it is considered a core ability that might be expected to load on both input and output factors, it was retained in the model. All other variables had an excellent to fair loading on one factor only, with no cross loadings over .4, showing a simple structure for the four factors, with each variable loading strongly on only one factor and each factor having multiple strong loadings.

Figure 8-3: Pattern Matrix

Pattern Matrix[a]

	Factor			
	1	2	3	4
R_FIST1	.909	.041	.084	-.116
R_FIST2	.626	.028	.047	.041
R_NWD	.601	.011	-.083	.157
R_Auditory Memory	.392	-.369	.029	.291
R_PCC1	-.064	.935	.006	.112
Zscore: DEAP SS	-.103	.760	-.051	.172
R_MultisyllSp	.253	.741	.091	-.055
R_MultisyllIm	.107	.615	.012	-.249
Zscore: SM SS	.137	.014	.681	-.379
Zscore: IM SS	-.085	.030	.619	.087
Zscore: BVMI SS	.137	-.066	.455	.265
Zscore: CELF Ex vocab	.154	-.104	-.005	.745
Zscore: CELF SenStr	.306	.089	-.072	.599
Zscore: BPVS SS	.044	.020	.383	.527
Zscore: CELF Word	.354	.198	.183	.405

Extraction Method: Principal Axis Factoring.

Rotation Method: Oblimin with Kaiser Normalization.

8.4.7 Robustness of the model

Despite the small sample size, the model had good simple structure, and deletion of less relevant variables did not affect the structure, suggesting it was robust. Multiple high loadings on each factor with few cross loadings also suggest it is reliable at this sample size (Guadagnoli and Velicer 1988).

Use of an orthogonal Varimax rotation was also trialled but the results remained the same, i.e. one factor representing cognitive skills, one speech accuracy, another non-speech motor skills and the last language skills, also suggesting a robust model.

8.4.8 Factor scores

Factor analysis is also useful for reducing items into discrete dimensions that can be summed or aggregated and subsequently used as input for further analysis. Factor scores are composite variables which provide information about a subject's placement on the factors (DiStefano *et al.* 2009).

Factor scores were computed using a non-refined method as non-refined methods are thought to be more stable across samples than refined methods, meaning the obtained

results do not depend as heavily on the particular sample used, and they are easy to interpret (DiStefano *et al.* 2009). Non-refined methods are relatively simple, cumulative procedures to provide information about individuals' placement on the factor distribution. The method chosen was to sum scores by factor, i.e. to add together the scores on the items loading on that factor for each subject. As the structure generated was a simple one, identifying these items was straightforward (highlighted in Figure 8-3 above). Tabeachinck and Fidell (2001) noted that this approach is generally acceptable for most exploratory research situations.

8.4.9 References

Bartholomew, D. J., Knott, M. and Moustaki, I. (2011) *Latent variable models and factor analysis: A unified approach*, Hoboken, NY: John Wiley & Sons.

Conway, J. M. and Huffcutt, A. I. (2003) 'A review and evaluation of exploratory factor analysis practices in organizational research', *Organizational research methods*, 6(2), 147-168.

DiStefano, C., Zhu, M. and Mindrila, D. (2009) 'Understanding and using factor scores: Considerations for the applied researcher', *Practical Assessment, Research & Evaluation*, 14(20), 1-11.

Dodd, B. and McIntosh, B. (2010) 'Two-year-old phonology: impact of input, motor and cognitive abilities on development', *Journal of Child Language*, 37(05), 1027-1046.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. and Strahan, E. J. (1999) 'Evaluating the use of exploratory factor analysis in psychological research', *Psychological methods*, 4(3), 272.

Field, A. (2009) 'Exploratory factor analysis', *Discovering statistics using SPSS*, 3, 627-685. London: Sage

Ford, J. K., MacCallum, R. C. and Tait, M. (1986) 'The application of exploratory factor analysis in applied psychology: A critical review and analysis', *Personnel psychology*, 39(2), 291-314.

Gorsuch, R. L. (1983) *Factor Analysis*, 2nd ed., Hillsdale, NJ: Lawrence Erlbaum Associates.

Guadagnoli, E. and Velicer, W. F. (1988) 'Relation to sample size to the stability of component patterns', *Psychological bulletin*, 103(2), 265.

Hatcher, L. (1994) *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*, Cary, North Carolina: SAS Institute Inc.

Nunnally, J. C. and Bernstein, I. (1994) 'The assessment of reliability', *Psychometric theory*, 3(1), 248-292.

Pallant, J. (2010) 'SPSS survival manual: *A step by step guide to data analysis using SPSS,' 4th ed.*, Maidenhead: Open University Press.

Pett, M. A., Lackey, N. R. and Sullivan, J. J. (2003) *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*, London: Sage.

Suhr, D. (2005) 'Principal Component Analysis versus Exploratory Factor Analysis' SUGI 30 Proceedings, Philadelphia, Pennsylvania April 10-13,. online available at http://www2.sas.com/proceedings/sugi30/203-30.pdf accessed 1-10-13.)

Tabachnick, B. G. and Fidell, L. S. (2001) 'Principal components and factor analysis', *Using multivariate statistics*, 4, 582-633. New York: Harper & Row.

8.5 Appendix 5.A: Descriptive statistics and univariable regression model results for demographic, risk factors, and assessment results associated with resolved SSD.

NB: Resolved versus Persistent: Figures in yellow cells show N, in blue cells show mean score for each group.

Highlighted variables with a *p* value <.10 were retained for use in the multivariable analysis.

Variable	Category	N	Resolved SSD	Persistent SSD	Odds ratio	95% Confidence Intervals		*p* value
						Lower	Upper	
Gender	Female	26	20	6	1			
	Male	40	32	8	1.2	0.363	3.972	0.765
Age at Initial Ax (months)		66	49.4	44.6	1.082	0.989	1.184	0.085
Family history of speech, language or literacy difficulties	No	36	29	7	1			
	Yes	30	23	7	0.793	0.243	2.586	0.701
Delayed onset of speech	No	48	39	9	1			
	Yes	17	13	4	0.75	0.197	2.849	0.673
Birth difficulties	No	55	42	13	1			
	Yes	11	10	1	3.095	0.361	26.511	0.302
No. of ear infections	None	30	23	7	1			0.803
	1 to 3	25	21	4	1.598	0.409	6.247	0.501
	4 to 6	5	4	1	1.217	0.116	12.752	0.87
	6 or more	6	4	2	0.609	0.091	4.056	0.608
Received therapy	No	20	18	2	1			
	Yes	46	34	12	0.315	0.063	1.563	0.157
Diagnostic sub-category	IPD	7	3	4	1			0.030
	PD	32	29	3	12.89	1.906	87.170	0.009
	CPD	23	17	6	3.778	0.648	22.017	0.139
Atypical errors (>4 words, consistently)	No	43	35	8	1			
	Yes	23	17	6	0.648	0.194	2.165	0.480
Atypical process (>4 occurrences)	No	52	41	11	1			
	Yes	14	11	3	.984	.233	4.151	0.982
Comorbid language impairment	No	51	41	10	1			
	Yes	15	11	4	0.671	0.176	2.554	0.558

Table continued overleaf

Appendix 5-A: Descriptive statistics and univariable regression model results for demographic, risk factors, and assessment results associated with resolved SSD, *continued*.

Area	Variable	N	Resolved SSD	Persistent SSD	Odds ratio	95% Confidence Intervals		*p* value
						Lower	Upper	
Speech accuracy at Initial Ax	PCC (DEAP Phonology sub-test)	66	70.2	58.6	1.033	1.001	1.066	0.045
	PCC Multisyllabic words: spontaneous	47	79.8	69.2	1.047	0.996	1.1	0.07
	PCC Multisyllabic words: imitation	47	82.8	75.8	1.019	0.983	1.056	0.305
	DEAP Standard Score	66	5.6	4.2	1.364	0.988	1.883	0.059
DEAP Oral Motor	DDK	63	10.2	9.3	1.274	0.927	1.75	0.136
DEAP Oral Motor	Isolated movements	63	9.1	8.9	1.075	0.763	1.515	0.678
DEAP Oral Motor	Sequenced movements	55	9.4	8.3	1.233	0.9	1.687	0.192
Receptive vocabulary	BPVS	65	99.8	96.7	1.022	0.972	1.075	0.388
Language, receptive	CELF P2(UK) Sentence structure sub-test	60	9.3	8.5	1.115	0.888	1.4	0.347
Language, morphology	CELF P2(UK) Word structure sub-test	60	9.8	7.7	1.146	0.972	1.35	0.105
Expressive vocabulary	CELF P2(UK) Expressive vocabulary	60	10.8	10.4	1.045	0.847	1.289	0.68
Language overall	CELF Core Language Score	60	100	93.3	1.023	0.986	1.061	0.226
Non-speech motor	Beery Butenika test of Visual Motor Integration	66	99.5	93.9	1.026	0.986	1.069	0.207
Phonological awareness	PIPA Syllable segmentation	39	10	8.7	0.775	0.5	1.2	0.254
	PIPA Alliteration Awareness	39	8.9	8	1.237	0.741	2.066	0.415
Cognitive tests*	Number Concepts	65	6.9	6.4	0.95	0.691	1.306	0.75
	Pattern Recognition	58	5.4	3.2	1.229	0.92	1.642	0.162
	Non-word Discrimination	50	11.2	7.3	1.322	1.058	1.651	0.014
	FIST1	64	13.4	9.7	1.197	1.028	1.394	0.021
	FIST2	64	8.9	8.3	1.007	0.872	1.163	0.924
	Auditory Memory	61	3.4	3.3	1.093	0.457	2.616	0.842

*These scores are not adjusted for age, therefore, in the univariable analysis, a hierarchical regression was performed, with age entered in block 1 and test score in block 2.