Language and memory in Williams syndrome

Jonathan Peter Brock

Thesis submitted for the degree of Doctor of Philosophy

Department of Psychology, University of Warwick

July 2002

Abstract

This thesis investigated the relationship between language, memory, and temporal processing abilities in Williams syndrome (WS). Despite mild to moderate mental retardation, there is evidence that individuals with WS have relatively spared language abilities and, in particular, that their receptive vocabulary increases at a rate faster than that predicted by overall abilities. However, it has also been suggested that semantic processing is impaired in WS and that individuals with WS have relative difficulty learning the meanings of words.

In Experiments 1 and 2, the performance of children with WS on temporal discrimination tasks was comparable with that of typically developing (TD) children and children with moderate learning disabilties (MLD) matched on vocabulary. However, there was little evidence to support the idea that temporal discrimination mechanisms contribute to language abilities.

Experiments 3, 4, 5, 6, and 7 investigated short-term memory abilities. Contrary to previous findings, there was no evidence for a reduced influence of lexical-semantic knowledge on STM in WS. Children with WS and TD and MLD controls matched on vocabulary showed equivalent effects of lexicality, word-frequency, concreteness, and wordlikeness, and similar tendencies to lexicalize nonwords. Overall performance in the WS group was generally worse than that of TD controls, but was comparable with that of MLD children matched on chronological age. This group difference was accounted for by differences in order memory, suggesting that serial order mechanisms may influence the rate of vocabulary acquisition.

Experiments 8 and 9 investigated free recall. Contrary to previous findings, children with WS and TD controls showed no evidence of a primacy effect in a standard free recall task, but both groups showed significant primacy effects when taught to rehearse.

Overall, therefore, children with WS showed a normal relationship between language and memory abilities. Implications for theoretical accounts of language in WS are discussed.

Table of Contents

List of Abbreviations

BAS	British Abilities Scale
BPVS	British Picture Vocabulary Scale
CDI	Communicative Development Inventory
CNRep	Childrens test of Nonword Repetition
DAS	Differential Abilities Scale
ERP	event-related potential
IPSyn	Index of Productive Syntax
K-BIT	Kaufman Brief Intelligence Test
LTM	long-term memory
MLD	moderate learning disability
OSCAR	OSCillator-based Associative Recall
PPVT-R	Peabody Picture Vocabulary Test
SLI	specific language impairment
STM	short-term memory
SVAS	supravalvular aortic stenosis
TD	typically developing
TROG	Test for Reception of Grammar
WAIS	Wechsler Adult Intelligence Scale
WISC	Wechsler Intelligence Scale for Children
WS	Williams syndrome

Part I: Introduction

Chapter 1: Williams syndrome

1.1: Overview of the thesis

In recent years, there has been considerable interest in the linguistic abilities of individuals with Williams syndrome (WS). Despite mild to moderate mental retardation, language abilities in WS appear to be relatively preserved, although there is growing evidence that there are strengths and weaknesses within the language profile. This thesis investigates possible constraints on language acquisition in WS by looking at the verbal memory and temporal processing abilities of a group of 14 children with WS in a series of nine experiments.

The thesis is divided into five parts. Part I (chapters 1 and 2) provides a review of previous research in WS. Chapter 1 is a general review of the medical, genetic, neurological and cognitive features of the syndrome, while chapter 2 provides a more in-depth review of studies of language in WS. These studies suggest that, compared with overall cognitive abilities, receptive vocabulary and verbal short-term memory (STM) are relative strengths, but there is little evidence for relative preservation of other linguistic abilities. Studies looking at possible factors which might contribute to the language profile in WS are then presented in the remainder of the thesis.

Part II of the thesis (chapters 3, 4, and 5) looks at the possibility that language abilities are relatively strong as a result of preserved temporal processing abilities. Parts III (chapters 6, 7, 8, 9, and 10) and IV (chapter 11) follow up existing research looking at the causal relationship between language and memory skills in WS. Part V (chapter 12) provides a summary of the main findings in the thesis and discusses their implications.

1.2: Overview of Williams syndrome

Williams syndrome is a genetic disorder with an estimated incidence of around 1 in 20,000 live births (Morris, Demsey, Leonard, Dilts, & Blackburn, 1988). Early studies noted the sporadic (as opposed to hereditary) occurrence of supravalvular aortic stenosis (SVAS, a narrowing of the aorta), and hypercalcaemia (excessive blood calcium levels) in association with developmental delay and unusual facies (Beuren, Schulze, Eberle, Harmajanz, & Apitz, 1964; Bongiovanni, Eberlein, & Jones, 1957; Garcia, Friedman, Kaback, & Rowe, 1964; Joseph & Parrott, 1958; Williams, Barratt-Boyes, & Lowe, 1961). Studies of the hereditary form of SVAS (Ewart, Morris, Ensing et al., 1993) led to the discovery of submicroscopic deletions of about 20 genes in the 7q11.23 region of chromosome seven in WS (Ewart, Morris, Atkinson et al., 1993). Subsequently it was shown that more than 98% of individuals with WS have deletions of the gene ELN (Lowery et al., 1995; Mari et al., 1995) and, based on these findings, a FISH test (fluorescent in situ hybridization) was developed to probe for the ELN deletion and provide a simple reliable genetic test for WS (Lowery et al., 1995).

Individuals with WS tend to have a characteristic 'elfin' facial profile with full prominent lips, a wide mouth, flat nasal bridge and short nose (Joseph & Parrot, 1958). Other physical characteristics include a stellate iris pattern, small widely-spaced teeth and short stature. As well as SVAS, and hypercalcaemia, high blood pressure and hypertension are also common in WS, together with eating and feeding problems and a failure to thrive during infancy (Martin, Snodgrass, & Cohen, 1984; Morris et al., 1988). WS is associated with mild to moderate mental retardation although there is wide individual variation which is reflected in the fact that some adults with WS are able to live independently or semi-independently whereas others require significant support (Udwin, 1990). Personality characteristics include hyperactivity, anxiety and a tendency to be extremely sociable and affectionate, especially towards adults (Bellugi, Adolphs, Cassady, & Chiles, 1999; Udwin, Howlin, & Davies, 1996). Perhaps the most common feature of WS, however, is an abnormal sensitivity to certain sounds, known as auditory hyperacusis (Marriage, 1996; van Borsel, Curfs, & Fryns, 1997).

Much of the interest in WS stems from the unevenness of the cognitive profile. Although studies typically report mean IQs in the mid 50s to low 60s range (e.g. Arnold, Yule, & Martin, 1985; Bellugi, Lichtenberger, Jones, Lai, & St George, 2000; Howlin, Davies, & Udwin, 1998; Mervis, Morris, Bertrand, & Robinson, 1999; Udwin & Yule, 1990; Udwin, Yule, & Martin, 1987), several 'islets of ability' have been identified in language, social cognition and face processing. Moreover, increasing knowledge of the neurophysiology of WS and the recent discovery of its genetic basis has offered the promise of linking the genotype with the neurological and cognitive phenotype (e.g. Bellugi, Lichtenberger, Mills, Galaburda, & Korenberg, 1999). The following sections provide a brief review of spatial, numerical, face-processing, social and language abilities in WS and their hypothesised links with neurological and genetic abnormalities.

1.3: The cognitive profile in WS

1.3.1: Spatial skills

Spatial tasks represent a particular area of difficulty in WS, although performance across tasks is uneven (see Farran & Jarrold, 2002, for review). For example, numerous studies (Bellugi, Marks, Bihrle, & Sabo, 1988; Bellugi, Bihrle, Neville, Doherty, & Jernigan, 1992; Dall'oglio & Milani, 1995; Howlin et al., 1998; Mervis et al., 1999; Frangiskakis et al., 1996; Udwin et al., 1987) have reported poor performance on the block design subtests of the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1974) and the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1986), and on the Pattern Construction subtest of the Differential Abilities Scale (DAS; Elliot, 1990). Indeed, poor Pattern Construction performance is one of the criteria for the WS Cognitive Profile (Mervis et al., 2000). Similarly, children with WS have been found to perform worse than language-matched controls on copying the Rey figure (Rey, 1968; Vicari, Brizzolara, Carlesimo, Pezzini, & Volterra, 1996), and drawing abilities in general are severely delayed, although they seem to follow a normal developmental course (Bellugi et al., 1988; Bertrand, Mervis, & Eisenberg, 1997; Bertrand & Mervis, 1996).

Several studies have reported poor performance on orientation matching (Bellugi et al., 1988; Rossen, Klima, Bellugi, Bihrle, & Jones, 1996; Wang, Doherty, Rourke, & Bellugi, 1995) but these have suffered from floor effects (cf. Farran & Jarrold, 2002) and other studies have found that performance is commensurate with non-verbal ability (Atkinson et al., 1997; Stiers, Willekens, Borghgraef, Fryns, & Vandenbussche, 2000). Performance on the Corsi blocks, a test of visuo-spatial STM, is worse than auditory-verbal STM (Jarrold, Baddeley, & Hewes, 1999; Wang & Bellugi, 1994). Finally, object recognition has been shown to better than in Down syndrome (Wang et al., 1995) but at a similar level to that in TD controls matched on overall mental age (Hoffman & Landau, 2000).

There has been considerable debate regarding the precise nature of the spatial deficit in WS. Bellugi and colleagues (e.g. Bellugi, Wang, & Jernigan, 1994) have proposed that individuals with WS have a deficit in global visuo-spatial perception. Supporting evidence comes from the poor global organization shown in free drawing, block design and copying of hierarchical figures (Bihrle, Bellugi, Delis, & Marks, 1989; Rossen et al., 1996; see Stevens, 1996 for limited support for this view). However, unlike individuals with autism who are also thought to have a deficit in global perception (Shah & Frith, 1993), people with WS demonstrate normal facilitation of block design performance when the target design is segmented (Farran, Jarrold, & Gathercole, 2001; Mervis et al., 1999).

There is also no evidence for a local bias in perception of hierarchical figures. Farran, Jarrold, and Gathercole (2002) reported that individuals with WS show normal global precedence and interference effects in responding to elements of these figures, while Stevens (1996) found that individuals with WS were able to give accurate descriptions of hierarchical figures. People with WS have also been found to show normal global processing in a visual search task (Pani, Mervis, & Robinson, 1999). Farran and Jarrold (2002) have therefore argued that individuals with WS can perceive information at local and global levels but have difficulty using this information in visuo-spatial construction at the global level. Similarly, Pani et al. have suggested that visuo-spatial constructive deficits result from a general weakness in planning and organising information in working memory, while Atkinson et al. (1997) have suggested that there may be a deficit in frontal function linked to the selection and control of spatial behaviour.

Atkinson and colleagues have also proposed that spatial deficits in WS are associated with deficits in visual information processing in the dorsal stream, but an intact ventral stream. This view is supported by evidence of abnormal motion coherence, a test of dorsal stream function, but normal form coherence which involves the ventral stream (Braddick, Atkinson, Wattam-Bell, & Anker, 1997). Furthermore, individuals with WS showed severe impairment in posting a card through an oriented slot but only modest impairment in matching the orientation of the card to the slot, the opposite pattern of results to patients with selective damage to ventral visual processing stream (King et al., 1997).

1.3.2: Number

Another area of difficulty for individuals with WS is mathematics and its everyday application in dealing with, for example, money, recipes, and estimation of quantities, and this provides a serious obstacle to independent living (Bellugi, Lichtenberger et al., 1999). Some people with WS are able to master addition and a small number can also perform subtraction and division (Bellugi et al., 2000). However, Udwin et al. (1996) tested 23 young adults on the Arithmetic subtest of the WISC and found that only six scored above basal level. Johnson (2000), meanwhile, reported that young adults with WS had poor understanding of the concepts of zero, infinity, and fractions. She suggested that their difficulty lay in the reconceptualization that numbers do not necessarily refer to whole objects.

1.3.3: Social cognition

In contrast to their deficts in spatial cognition and number, social cognition has been characterised as a relative strength in WS, prompting comparisons with social deficits in autism (see e.g. Karmiloff-Smith, Klima, Bellugi, Grant, & Baron-Cohen, 1995; Tager-Flusberg & Sullivan, 2000). Individuals with WS are characterised as having sociable, empathetic, affectionate and outgoing personalities (Einfeld, Tonge, & Florio, 1997; Gosch & Pankau, 1997; Sarimski, 1997; Tomc, Williamson, & Pauli, 1990; Udwin & Yule, 1991). Moreover, Tager-Flusberg and colleagues have reported that, compared with controls with Prader-Willi syndrome, young children with WS show greater empathy (Tager-Flusberg & Sullivan, 1999) and perform better at matching complex emotions to pictures of faces (Tager-Flusberg, Boshart, & Baron-Cohen, 1998; but see Tager-Flusberg & Sullivan, 2000). Similarly, Karmiloff-Smith et al. (1995) reported that individuals with WS were able to use eye direction to infer mental states and intentions.

However, people with WS often have difficulty understanding unwritten social rules and have more difficulty making social relationships with peers than with adults (Einfeld et al., 1997; Udwin & Yule, 1998). Tager-Flusberg and Sullivan (2000) suggested that this reflects a dissociation between relatively spared social perception and deficits in social cognition. Consistent with this, they found that on traditional tests of theory of mind, and tests in which they were required to interpret behaviour in terms of mental states, children with WS were no better than age-, language-, and IQ-matched controls with Prader-Willi syndrome or mental retardation of mixed aetiology.[1]

1.3.4: Face processing in WS

Related to the relatively good social perception of individuals with WS, a second area of relative strength is in face-processing, with several studies reporting performance in the normal range on tests such as Benton Test of Facial Recognition (Benton, Hamsher, Varney, & Spreen, 1983; see Bellugi et al., 1992; Rossen et al., 1996; Wang et al., 1995). However, because of ceiling effects, it is not clear from these studies whether performance really is age-appropriate (cf. Farran & Jarrold, 2002). Moreover, Deruelle, Mancini, Livet, Casse-Perot, and de Schonen (1999) reported that individuals with WS performed at the same level as TD children matched on mental age on tasks requiring matching of faces by gender, age, gaze direction, emotional expression and identity. The WS group in this study were found to be at the same level as chronological age-matched controls on lip-reading, but this may again be due to ceiling effects (cf. Farran & Jarrold, 2002).

However, even if face-processing is a relative strength in WS, there is considerable evidence to show that the underlying processes are different to those in typical development. In an event-related potential (ERP) study of face-processing, Mills (1998) reported that the N200 component was relatively large when processing faces but not other objects, and the anterior N320 component did not show the normal right hemisphere asymmetry. In a subsequent study, Mills et al. (2000) found abnormal early ERP components which were not found at any age in TD and brain-injured individuals and people with other learning disabilities. One suggestion is that individuals with WS have a relatively local approach to face-processing (e.g. Deruelle et al., 1999). Consistent with this view, a number of studies have found that people with WS are unusually good at recognising upside-down faces which is indicative of a local approach (Karmiloff-Smith, 1997; Wang et al., 1995; but see Mills et al., 2000).

1.3.5: Language in WS

The third and most widely researched area of relative strength in WS is language. Recent interest in WS began with a study by Bellugi et al. (1988) of three children with WS aged between 11 and 16 years. Despite IQs around 50, visuo-spatial construction skills at the five-year-old level, and the inability to master Piagetian conservation tasks normally passed by seven-year-olds, these children all produced language described as 'complex in terms of morphological and syntactic structures' (p. 183; see Clahsen & Almazan, 1998 for similar findings). Later studies noted an ability to tell sophisticated stories with a large variety of narrative enrichment devices (Bellugi et al., 1994), and the extensive use of affective socially-oriented language (Reilly, Klima, & Bellugi, 1990; see also Stojanovik, Perkins, & Howard, 2001). Results such as these led Bellugi and her colleagues to conclude that in WS, 'linguistic functioning is selectively preserved in the face of severe general cognitive deficits' (Bellugi et al., 1992, p. 201).

However, there are two important objections to this claim. First of all, language abilities are not preserved in the strict sense that they are at the level predicted by chronological age. Performance on standardized tests is rarely if ever at age-appropriate levels (Bellugi et al., 1988, 1992, 1994; Bellugi & Wang, 1998; Rossen et al., 1996; Jarrold, Baddeley, & Hewes, 1998; Mervis et al., 1999), and early language is severely delayed (e.g. Singer–Harris, Bellugi, Bates, Jones, & Rossen, 1997). A more conservative claim is that language abilities are not absolutely preserved but are relatively preserved compared with general cognitive abilities (e.g. Bellugi, Lichtenberger et al., 1999; Mervis et al., 1999).

The second objection is that, even within the language profile, there appear to be strengths and weaknesses. For example, studies of older children and adults with WS have consistently found that receptive vocabulary is at a higher level than other language abilities such as syntax and morphology (Clahsen & Almazan, 1998; Grant et al., 1997; Karmiloff-Smith et al., 1997; Klein, Mervis, Hutchins & Bertrand, 1996, cited in Mervis et al., 1999; Volterra, Capirci, Pezzini, Sabbadini, & Vicari, 1996; Zukowski, 2001). Other studies have suggested a dissociation between intact regular and deficient irregular morphology (e.g. Clahsen & Almazan, 1998), and a relative advantage for learning the sound patterns of words over their meanings (e.g. Paterson, 2000). These findings suggest that it is inappropriate to talk of language as an entity as being 'preserved'.

1.4: Neurological profile

Post-mortem and magnetic resonance imaging studies of neuroanatomy in WS (e.g. Bellugi, Mills, Jernigan, Hickock, & Galaburda, 1999; Galaburda & Bellugi, 2000; Galaburda, Wang, Bellugi, & Rossen, 1994; Hickock et al., 1995; Jernigan, Bellugi, Sowell, Doherty, & Hesselink, 1993; Reiss et al., 2000; Wang, Hesselink, Jernigan, Doherty, & Bellugi, 1992) have shown that, although overall brain size is reduced in WS, several different regions are disproportionately preserved or reduced in volume. Based on the known role of these regions in TD individuals, these findings have been tentatively linked to features of the WS cognitive profile. For example, disproportionate reduction in posterior (parietal-occipital) volume has been related to visuo-spatial deficits, while relative preservation of frontal cortex and the cerebellum have been linked with relatively preserved language abilities. Several studies have noted that the superior temporal region is preserved in WS, a finding which has been related to strengths in music and auditory processing. Finally, limbic structures such as the amygdala, hippocampus and parahippocampal gyrus are smaller in WS than in normally developing brains but, in contrast to Down syndrome, they are in proportion to overall brain size. This finding has been related to the relatively normal socio-affective behaviour of individuals with WS. These links, however, remain speculative at present as most studies so far conducted have not been large enough to look at correlations between neuroanatomy and cognitive functioning.

1.5: Genetics

Most individuals with WS have a deletion of about 20 genes although some rare individuals have smaller deletions (Robinson et al., 1996). It has been argued that the study of these individuals provides the opportunity to determine the relevance of particular genes and determine the relationship between genotype and phenotype (e.g. Bellugi, Lichtenberger et al., 1999). However, the only unequivocal link so far is between the ELN gene and SVAS (e.g. Tassabehji et al., 1999). Several authors (e.g. Frangiskakis et al., 1996; Mervis et al., 1999) have suggested that the spatial deficit in WS may be related to the deletion of the gene LIMK1 which encodes the protein tyrosine kinase and is expressed in the developing brain (Proschel, Blouin, Gutowski, Ludwig, & Noble, 1995) where it seems to be indirectly involved in axonal growth (Arber et al., 1998). However, individuals have recently been identified who lack this gene but do not have spatial problems (Tassebehji et al., 1999). It therefore seems likely that the relationship between genotype and cognitive phenotype in WS is more complicated than at first assumed, with the uneven cognitive profile resulting from a complex interaction between multiple genes in the deleted region (Donnai & Karmiloff-Smith, 2000; Karmiloff-Smith, Scerif, & Thomas, 2002; Osborne, 1999).

1.6: Modular and neuroconstructivist accounts of WS

Several authors have taken WS as evidence for the existence of innately specified cognitive modules. Thus, for example, it has been suggested that modules for language, social cognition and face-processing are selectively spared in WS, in contrast to their selective impairment in other developmental disorders such as specific language impairment (SLI) and autism (e.g. Anderson, 1998; Karmiloff-Smith et al., 1995; Pinker, 1994; Rossen et al., 1996). The fact that these developmental disorders have a genetic basis has then been assumed to show that the relevant modules are innately specified (e.g. Pinker, 1999). However, at least as far as WS is concerned, this view seems to be oversimplistic. First, there is no convincing evidence at present for links between individual genes in the deleted region and specific higher-level cognitive impairments. Second, it seems inappropriate to characterise language and face-processing abilities as reflecting 'spared' modules as there is increasing evidence that the cognitive processes underlying performance are not the same as in typical development.

In contrast to this modular account, the 'neuroconstructivist' account (e.g. Elman et al., 1996; Karmiloff-Smith, 1992, 1998) assumes that modules are not innately specified but emerge during cognitive development, and there is no direct mapping from genes to modules. Rather, in developmental disorders such as WS and SLI, genetic factors influence the constraints under which cognitive development takes place (cf. Karmiloff-Smith et al., 2002). These constraints are non-specific but may affect certain cognitive processes more than others, leading to an uneven cognitive profile. In WS, these altered constraints lead to a pattern of relatively good language, face-processing and social perception in older children and adults (i.e. in the 'end-state'), but it does not necessarily follow that the same pattern will be observed early in development (cf. Paterson, Brown, Gsödl, Johnson, & Karmiloff-Smith, 1999), nor is it inevitable that good overall performance in a particular domain reflects the same underlying cognitive or neural mechanisms that are operating in typical development.

The problem facing this neuroconstructivist account of WS is that the relevant constraints on cognitive development are currently unspecified. One of the aims of this thesis is therefore to investigate some of the possible constraints upon language development in WS. The studies in Part II investigate the novel hypothesis that linguistic strengths in WS are a consequence of relatively preserved temporal processing abilities, while the studies in Parts III and IV look at the suggestion that language abilities may be related to relative strengths in phonological short-term memory and deficits in semantic memory. However, before these studies are reported, chapter 2 provides a more detailed review of language abilities in WS.

[1] In an earlier study, Karmiloff-Smith et al (1995) reported that older children and young adults with WS performed well on both first and second order theory of mind tests and half of those tested showed appreciation of both sarcasm and metaphor. However, there was no control group in this study and the participants were much older than the normal age at which these tests are passed.

Chapter 2: Language in Williams syndrome

2.1: Introduction

Chapter 1 provided a brief introduction to WS. This chapter takes a more detailed look at language abilities in WS and the twin claims that (a) language in WS is relatively preserved and (b) language in WS is atypical in that there are certain strengths and weaknesses within the language profile.

As noted in chapter 1, early studies led to claims that language in WS is selectively preserved in the face of severe cognitive deficits (e.g. Bellugi et al., 1992), but the fact that language is often severely delayed and is rarely at age-appropriate levels has led to the current consensus among researchers that language abilities are not absolutely preserved but are relatively preserved compared with general cognitive abilities (e.g. Bellugi, Lichtenberger et al., 1999; Mervis et al., 1999). Nevertheless, even within the language profile there seem to be certain strengths and weaknesses and this has led to the claim that language development does not follow the normal developmental course. However, there are methodological concerns with many of the studies upon which these claims are based. Consequently, it may be premature to make any major theoretical claims about language in WS or to suggest specific therapeutic interventions.

The starting point for this chapter is the null hypothesis that language abilities in WS are exactly what should be expected given overall cognitive abilities (i.e. full-scale IQ). If this is true then WS could not be used to support theoretical claims about the development of language, and it would suggest that individuals with WS would benefit from similar language interventions to other groups with the same level of mental retardation. The null hypothesis predicts that overall performance on tests of, for example, vocabulary, syntax, and morphology should be in line with overall mental age. In contrast, if language as a whole is relatively spared then performance on all linguistic tasks should be superior to that predicted by overall mental age, and performance across language tasks should be at similar levels. Evidence for relative strengths and weaknesses in the language profile would suggest that language development is not only delayed but also atypical. If language development followed a normal trajectory, then at any point in time, an individual's age-equivalent scores across a range of linguistic tests should be approximately equivalent.

The following sections consider support for these predictions in relation to studies of syntax, morphology and vocabulary, before looking at evidence for atypical lexical-semantic processing, vocabulary acquisition, preverbal communication and verbal memory. However, before reviewing these studies it is necessary to consider some methodological issues.

2.2: Methodological issues

Bishop (1997) and Burack, Iarocci, Bowler, and Mottron (2002) have provided comprehensive reviews of many of the methodological issues involved in studies of developmental disorders, while Farran and Jarrold (2002) and Thomas et al. (2001) have discussed methodological issues in relation to studies of visuo-spatial skills and past-tense morphology in WS. Many of the issues raised in these papers are relevant here.

2.2.1: Use of standardized scores

Early studies of WS were necessarily exploratory and often adopted the approach of comparing standard or age-equivalent scores on different tasks, for example comparing language scores with measures of overall mental age. The null hypothesis would predict that there should be no significant discrepancies across different tasks. However, if the two tests were standardized on different samples of TD children (as is usually the case unless performance is being compared on different subtests of the same battery), age-equivalent scores may vary substantially so apparent dissociations may not necessarily be genuine. Ideally, therefore, the performance of individuals with WS on one measure should be compared with that of TD controls matched on the second measure. Clearly, if novel unstandardized tasks are being used, this is a necessary procedure in any case.

2.2.2: Choice of control group

One problem with comparing individuals with WS to TD children is that matching on a measure such as overall mental age will inevitably mean that groups are mismatched in terms of chronological age. Potential group differences may be a consequence of the fact that the WS group have much greater language experience or world knowledge. This problem can be overcome by comparing the performance of individuals with WS to that of individuals with other forms of learning disability who are also matched on chronological age. Using a learning disabled control group also allows the specificity of the findings to be addressed. Obviously, showing that a particular pattern of performance is only demonstrated by individuals with WS would be an exhaustive process. Nevertheless it may be that group dissociations result from factors such as poor attention or working memory capabilities that are common to many learning disabilities, so including a learning disabled control group is an important consideration.

Many early studies used individuals with Down syndrome as a control group. However, verbal abilities in Down syndrome are typically inferior to nonverbal abilities (e.g. Fowler, 1990; Klein & Mervis, 1999). Results may therefore be ambiguous as, for example, superior performance on language tasks in WS could reflect language deficits in Down syndrome rather than spared language in WS. An alternative approach is to compare individuals with WS with learning-disabled controls of mixed aetiology, often referred to as an MLD (moderate learning disability) group. The problem here is that because the group composition may be variable, comparisons between different studies may not be particularly straightforward.

2.2.3: Criteria for matching controls

A crucial factor of course is the measure on which control groups are matched. This is a particularly thorny issue in WS where the cognitive profile is so uneven and matching on one measure inevitably entails that groups will be mismatched in terms of several other abilities. As such, it is important to choose the matching measure depending on the specific research question under consideration. For example, to test whether a language skill is better than predicted by overall cognition, groups should be matched on a measure of overall mental age. In contrast, if the aim is to show that language profiles are uneven, then groups should be matched on some measure of language ability.

2.2.4: Confounding task demands

A further important consideration is that group differences in performance on two tasks may result from differences in task demands such as attention, memory load or metacognitive awareness rather than from meaningful differences in the domain of specific interest. The ideal situation is to compare performance on closely matched tasks but when comparing performance on verbal and nonverbal tasks this may not always be possible. To be convincing, such studies need to show that dissociations hold across a range of tasks.

2.2.5: Floor and ceiling effects

Another potential problem is that of floor and ceiling effects which can mask group differences. For example, individuals with WS and controls may both be at ceiling but this does not mean that there would be no group differences if the test was more difficult or the participants had been tested at an earlier age so they were below ceiling. However, it is often the case that ceiling performance by WS participants is taken as indicating a 'spared' ability. Moreover, ceiling effects can lead to spurious group by condition interactions. For example, if there are ceiling effects in test A but not in test B then a general deficit across tasks will appear as a specific deficit in test B (see for example the studies of regular and irregular past tense in Section 2.3.3).

2.2.6: Group sizes

The final issue concerns the group sizes. Given the rarity of the syndrome it may be difficult to recruit large groups, especially if the study is focusing on a relatively narrow age range. However, it is important to employ reasonable sized groups to disentangle systematic variation between groups from random noise, particularly if the tests involved only produce a small range of scores. Thus while case studies and studies of small groups provide useful starting points for further investigation, caution should be exercised when generalizing results to the syndrome as a whole. On the other hand, it is also important to realise that the average WS profile may not be representative of all individuals.

With these considerations in mind, the following sections review studies of grammatical abilities and vocabulary knowledge in WS.

2.3: Grammatical abilities

2.3.1: Performance on the TROG

Early studies (e.g. Bellugi, Bihrle, Jernigan, Trauner, & Doherty, 1990; Bellugi et al., 1994) demonstrated that individuals with WS performed well on selected blocks of the Test for Reception of Grammar (TROG; Bishop, 1989). However, later studies using the full-scale TROG have consistently failed to find evidence that grammatical abilities are above the level predicted by overall mental age or nonverbal mental age (Clahsen & Almazan, 1998; Grant et al., 1997; Karmiloff-Smith et al., 1997; Volterra et al., 1996; Zukowski, 2001).

2.3.2: Syntax

Klein et al. (1996; cited in Mervis et al., 1999) reported that the performance on the Index of Productive Syntax (IPSyn; Scarborough, 1990) of 39 children with WS (mean age seven years) was in line with their overall mental age. Moreover, the relationship between productive syntax and mean length of utterance was the same as that found in TD children. Meanwhile, Singer-Harris et al. (1997; see also Bates & Goodman, 1999) found that parental reports of grammatical complexity in WS were in accordance with their productive vocabulary and mean length of utterance. Thus, although syntactical development is delayed, there is little evidence from these studies that it is any better than would be predicted from overall cognitive abilities or that it follows an atypical trajectory.

However, detailed analysis of performance on the TROG and similar tests has revealed a pattern of relative strengths and weaknesses in syntactical awareness in WS. People with WS appear to have little trouble with negatives or passives (Bellugi et al., 1990, 1992; Clahsen & Almazan, 1998; but see Stojanovik et al., 2001), but difficulties have consistently been found with spatial prepositions such as 'the pen is above the book' or 'the tree is in front of the house' (Clahsen & Almazan, 1998; Lichtenberger & Bellugi, 1998; Phillips, Jarrold, Baddeley, Grant, & Karmiloff-Smith, 2002; Whittle, Chang, Thomas, & Mervis, 2001). Phillips et al. argue that these difficulties are caused by general deficits in visuo-spatial cognition in WS, although the study by Whittle et al. suggests that there may be a more general problem with relational terms.

Consistent evidence has also been found for specific difficulties with complex grammatical constructs including embedded clauses such as 'the boy the dog chases is big' (Grant, Valian, & Karmiloff-Smith, 2002; Karmiloff-Smith et al., 1997; Stojanovik et al., 2001). Zukowski (2001) investigated the production of relative clauses in 10 children with WS and 10 TD controls matched on the vocabulary and matrices subtests of the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990). Both groups found it easier to produce subject-gap relative clauses (e.g. 'the boy who is pointing to the cow') than object-gap clauses (e.g. 'the cow that the boy is pointing to'), but the size of this effect was greater for the WS group. Nevertheless, all but one of the WS group produced at least one object-gap relative clause, suggesting that knowledge of object-gap relative clauses was intact. Zukowski proposed that retrieval of lexical material is unusually slow in WS so that, while TD controls are able to prepare material for both the subject and the object before producing a response, children with WS can only have material for the object position ready, and so have increased difficulties with object-gap clauses.

Similar conclusions were reached by Karmiloff-Smith et al. (1998) who investigated the sensitivity of people with WS to different syntactical errors using an on-line word-monitoring task. Typically, when participants are required to monitor spoken sentences for a target word, response latencies increase if the target is preceded by a syntactic violation (Marslen-Wilson, Vanhalen, & Moss, 1988). Individuals with WS were sensitive to violations in the formation of coherent noun phrases (e.g. 'new the test' rather than 'the new test') or auxiliary and main verbs (e.g. 'might expecting' rather than 'might be expecting') but, unlike normal adults, they were insensitive to violations of subcategory structures (e.g. 'struggle the dog' rather than 'struggle with the dog'). Karmiloff-Smith et al. suggested that people with WS are slow to integrate lexically specified syntactic constraints, and that the specific deficit in processing subcategory structures arises because these structures are relatively difficult to integrate on-line.

Thus, contrary to earlier claims, and consistent with the null hypothesis, there is little evidence that overall syntactical abilities in WS are any better than might be expected given the level of mental retardation. However, there is some evidence that certain syntactical structures prove unusually difficult for individuals with WS. Difficulties with spatial relations probably reflect the general deficit in spatial cognition in WS rather than a specific linguistic deficit. In contrast, problems with embedded clauses and subcategory structures suggest that syntactical abilities are not only delayed but may also be deviant. Nevertheless it is difficult to rule out the possibility that these apparently specific deficits simply reflect a general syntactical delay combined with ceiling effects for easier items. Moreover, without data from learning disabled control groups it is unclear how specific these findings are to WS.

2.3.3: Morphology

Contrary to the idea that language is preserved in WS, numerous studies have noted difficulties with morphology. For example, Rubba and Klima (1991) reported difficulties with preposition use in English, while Caceres, Heinze, and Mendez (1999) found problems in noun / determiner agreement and gender and number anaphor agreement in Spanish-speaking individuals. Karmiloff-Smith et al. (1997) reported that French-speaking individuals with WS had difficulties with the complex cue-based system of gender assignment which were even greater than those experienced by TD controls who had lower receptive vocabulary and grammar ages than the WS group. Consequently, these problems could not be explained in terms of general language delay. Similarly, Volterra et al. (1996; see also Capirci, Sabbadini, & Volterra, 1996) reported that Italian children with WS showed errors of gender agreement and verb conjugation which were qualitatively different from those found in normally developing children of the same syntactical level.

A number of studies have suggested that while regular morphology is relatively intact, there is a specific deficit in irregular morphology in WS (Bromberg et al., 1994; Clahsen & Almazan, 1998, 2001; Clahsen & Temple, 2002; Zukowski, 2001). According to modular linguistic accounts (e.g. Pinker, 1991, 1994), regular morphology reflects the operation of a rule (e.g. add -ed to walk to create the past tense), while irregular morphology requires the use of lexical-semantic information to block the rule and provide the appropriate form (e.g. go becomes went). Deficits in irregular morphology in WS have therefore been explained in terms of an impaired lexical-semantic system (Clahsen & Almazan, 1998; Clahsen & Temple, 2002).

However, the evidence to support this distinction between regulars and irregulars is not entirely convincing. Bromberg et al. (1994) reported that individuals with WS were better at inflecting regular plurals and past tense than irregular forms, although they did not report whether this difference was statistically significant. Moreover, they only tested six people with WS and gave no details of the control group. Clahsen and colleagues (Clahsen & Almazan, 1998, 2001; Clahsen & Temple, 2002) have found consistent evidence for selective deficits in the formation of irregular past tense, irregular plurals and irregular comparatives (e.g. bad becomes worse). However, their studies only involved four children with WS, and ceiling effects for regulars combined with a general deficit in both regular and irregular morphology in WS could account for the apparently selective deficits for irregulars (cf. Section 2.2.5).

A larger study was conducted by Zukowski (2001) who tested 12 children with WS and 12 TD controls matched on both the vocabulary and matrices subtests of the K-BIT. Both groups were close to ceiling for regular noun plurals, and while the WS group were worse than controls on irregulars, the difference was non-significant. Again, ceiling effects could account for the trend towards an apparently selective deficit for irregulars. Finally, Thomas et al. (2001) compared past tense formation in 18 individuals with WS and groups of TD children and adults of various ages. When vocabulary age was controlled for by regression, individuals with WS were worse than controls overall, but there was no evidence for a selective deficit in the irregular past tense.

Several of these studies have also looked at generalization to novel verbs and nouns. Clahsen and Almazan (1998) reported that children with WS were more likely than controls to regularise novel verbs that rhymed with existing irregulars (e.g. crive becoming crived rather than crove by analogy with drive and drove), but Thomas et al. (2001) failed to find such a group difference. Instead the WS group showed a reduced ability to generalise the regular add –ed rule to novel verbs. Neural network simulation suggested that this could be caused by increasing the discriminability of phonological representations in the network and may be related to abnormal auditory perception in WS (Thomas & Karmiloff-Smith, 2002). However, in a similar analysis Zukowski (2001) found that children with WS and controls were equally likely to provide the regular plural for novel nouns.

Thomas et al. (2001) also reported that their WS group showed an advantage for high- compared with low-imageability irregular verbs which was not present in controls. They argued that this may indicate weak semantics in WS which, combined with the weaker semantic input provided by low imageability verbs, takes the semantic information below some critical level and impairs performance on these verbs, a view supported by neural network simulation (Thomas & Karmiloff-Smith, 2002). However, given the inconsistency of other findings concerning morphology in WS, this finding really needs to be replicated.

To summarise, the general weakness in morphology provides further evidence against the idea that language as an entity is selectively preserved in WS. The Thomas et al. (2001) findings suggest that morphological skills are below the level predicted by receptive vocabulary, although it is not clear how they relate to overall mental age. The evidence for a specific deficit in irregular morphology (and therefore the evidence for a deficit in lexical-semantic processing) is not convincing as many of the studies are plagued by ceiling effects and have small participant groups. The evidence for impairments in generalisation is mixed and the reduced imageability effect reported by Thomas et al. needs to be replicated. Consequently, although there are a number of interesting findings, further research is required before any strong claims can be made about atypical development of morphology in WS.

2.4: Vocabulary

While there is little evidence that syntactical and morphological abilities are superior to those predicted by overall mental age, there is consistent evidence that performance on receptive vocabulary tests like the British Picture Vocabulary Scale (BPVS; Dunn, Dunn, Whetton, & Burley, 1997; Dunn, Dunn, Whetton, & Pintilie, 1982) or Peabody Picture Vocabulary Test (PPVT-R; Dunn & Dunn, 1981) is a relative strength. In the largest study to date, Mervis et al. (1999) tested 127 children and adults with WS and found that, although most individuals had full-scale IQs in the mentally-retarded range according to the DAS, 42% scored in the normal range (standard score over 70) on the PPVT-R. In general, studies with adults and older children with WS have reported that vocabulary scores are significantly higher than predicted by overall mental age measured using the WISC or the WAIS (Bellugi et al., 1988, 1990, 1992; 1994; Bellugi & Wang, 1998; Rossen et al., 1996; Temple, Almazan, & Sherwood, 2002; Tyler et al., 1997) and are also superior to grammatical abilities as measured by the TROG or the IPSyn (Clahsen & Almazan, 1998; Grant et al., 1997; Karmiloff-Smith et al., 1997; Klein et al., 1996, cited in Mervis et al., 1999; Volterra et al., 1996; Zukowski, 2001). However, studies with younger children with WS have generally found that vocabulary age is at the same level as overall mental age (Gosch, Stading, & Pankau, 1994; Kataria, Goldstein, & Kushnick, 1984; Klein & Mervis, 1999; Udwin & Yule, 1990; Volterra et al., 1996).

A similar pattern emerges from studies comparing WS and Down syndrome. Older children and adults with WS have consistently been found to have superior vocabulary knowledge to age- and IQ-matched controls with Down syndrome (Bellugi et al., 1988, 1990, 1992, 1994; Bellugi & Wang, 1998; Rossen et al., 1996). However, Paterson et al. (1999) tested 24- to 36-month-old infants with WS and Down syndrome and younger TD controls matched on mental age according to the Bayley Scales of Infant Development (Bayley, 1993). When children were shown pairs of photographs of everyday objects and told to 'Look, look at the X' all three groups showed the same preference for looking at the corresponding picture, indicating equivalent vocabulary knowledge. Similarly, Singer–Harris et al. (1997) found that 12- to 76- month-old children with WS and Down syndrome matched on chronological age showed no difference in overall language age according to the MacArthur Communicative Development Inventory (CDI; Fenson et al., 1993). Other studies have shown that there is no difference in the receptive vocabulary of mental-age-matched children with WS and Down syndrome at nine- and ten-year-olds of age (Klein & Mervis, 1999) and that children with WS only begin to outperform those with Down syndrome during early adolescence (Jones, Singer, Rossen, & Bellugi, 1993).

Jarrold, Baddeley, Hewes, and Phillips (2002; see also Jarrold et al., 1998) reported a longitudinal study of WS in which age equivalent scores for receptive vocabulary increased faster than those for Pattern Construction subtest of the DAS. Similarly, Jones, Hickock, Rossen, and Bellugi (1998) reported a cross-sectional study in which vocabulary scores began low but showed a sharp increase with age, while performance on a drawing task remained consistently low. Such findings have been taken to show that, although initially poor, verbal abilities develop at a faster rate than nonverbal abilities (Jarrold et al., 1998, 2001). However, given the lack of evidence that syntax or morphology are relative strengths at any age, it seems premature to generalise from receptive vocabulary to language in general.

Despite their relatively good performance on receptive vocabulary tests, individuals with WS have much more difficulties on other tests of vocabulary knowledge. In particular performance on tests in which they are required to produce a definition for a word or describe the similarities between word pairs (e.g. family and tribe) is worse than that for receptive vocabulary (Bellugi et al., 1990; Klein & Mervis, 1999; Mervis et al., 2000; although Temple et al., 2002 reported difficulties with definitions but not similarities). Temple et al. also reported that, when presented with an array of 24 semantically related pictures and required to point to a named item, children with WS performed worse than mental age-matched controls, despite having superior receptive vocabulary. They therefore argued that receptive vocabulary tests only require superficial semantic knowledge and individuals with WS perform poorly when more precise semantic information is required. However, it may simply be the case that individuals with WS have difficulties with the meta-cognitive requirements of forming definitions or identifying the crucial similarities (cf. Bellugi et al., 1988) and have difficulty choosing from a large number of responses.

Nevertheless, the idea that people with WS have a weak grasp of the meaning of words fits well with observations that, in spontaneous speech, they have a tendency to use low-frequency words, social phrases, clichés and idioms in contexts that are not entirely appropriate (Bellugi et al., 1990, 2000; Bradley & Udwin, 1989; Howlin et al., 1998; Rossen et al., 1996; Udwin et al., 1987; Udwin & Yule, 1990). It also fits well with studies of early language acquisition in WS. Normally young children comprehend more words than they actually produce, but Paterson (2000) reported that children with WS did not show this asymmetry. Similarly, Singer-Harris et al. (1997) noted that, although the relationship between production and comprehension in WS measured by parental reports using the MacArthur CDI was not significantly different to that in typical development, several parents reported that their child produced words that they did not understand[2].

To summarise, there is consistent and convincing evidence that individuals with WS accrue vocabulary knowledge at a rate faster than expected given their overall level of mental retardation. However there is a suggestion that although they may be proficient at learning the phonological forms of words and sufficient semantic information to be able to perform well on receptive vocabulary tests, they have difficulty when more precise semantic information is required. Thomas and Karmiloff-Smith (2002) have referred to this as an 'imbalance' between phonology and semantics in WS, although the precise nature and cause of this imbalance is unclear. The following sections review studies of lexical-semantic processing, constraints on vocabulary acquisition, early language development, and verbal memory that potentially address these issues.

2.5: Lexical-semantic processing

Temple et al. (2002) investigated naming abilities in four children with WS and found that age-equivalent scores on the Naming subtest of the British Abilities Scale (BAS; Elliot, Smith, & McCulloch, 1996) were significantly below those for receptive vocabulary. Naming reaction times were also measured for two of these children who were found to be faster but less accurate than TD controls matched on overall mental age. Temple et al. therefore proposed that the parameters governing lexical-semantic access have looser criteria for target identification but more rapid arrival at selected targets (see Rossen et al., 1996 for a similar hypothesis). However, other larger studies have failed to find similar effects. Rossen et al. (1996) reported that the naming accuracy of adolescents with WS, measured using the Expressive One-Word Vocabulary Test (Gardner, 1983), was in line with their receptive vocabulary. Similarly, Laing, Hulme, Grant, and Karmiloff-Smith (2001) found that naming abilities were equivalent in WS and TD controls matched on the BPVS. Thomas (2001) did report that a group of 16 adolescents and adults with WS were less accurate than vocabulary-matched controls on naming of objects and actions but, contrary to the findings reported by Temple et al., the WS group were also slower than controls.

A number of studies have looked at performance on word-fluency tasks, in which participants are required to think of as many different category exemplars as possible (e.g. 'animals'). Individuals with WS have consistently been found to produce an unusual number of low frequency words compared with controls matched on overall mental age (Bellugi et al., 1994; Rossen et al., 1996; Temple et al., 2002; Wang & Bellugi, 1993) which may suggest unusual semantic organization or access. However, the WS groups in these studies probably had better vocabularies than controls (as vocabulary in WS is typically in advance of overall mental age) and would therefore have known more low frequency words. Indeed, studies in which controls were matched on vocabulary rather than overall mental age have failed to find any group differences in the number of low frequency words produced (Jarrold, Hartley, Phillips, & Baddeley, 2000; Scott et al., 1995; Stephens, 1996; see Jarrold et al. for a discussion of the methodological issues).

Rossen et al. (1996) investigated semantic processing in six children with WS using a variety of tasks involving homonyms, words (e.g. 'bank') that have two meanings. When required to say the first word they thought of upon hearing the homonym, children with WS showed the same bias for the more common 'primary' meaning as TD controls and children with Down syndrome. When presented with the homonym and words related to its alternative meanings (e.g. 'bank', 'river' and 'money') and asked to decide which word pair went together best, both WS and Down syndrome children showed a reduced bias towards primary meanings compared with TD children. Finally, when required to produce definitions for the homonyms, children with WS produced the same number of primary definitions as both control groups but were more likely to produce the secondary definition as well. Rossen et al. suggested that children with WS show a reduced bias towards primary meanings when task demands are high. However, it is not entirely clear why this should be the case and the small sample size means that this study really needs to be replicated.

Tyler et al. (1997) investigated semantic priming in adults with WS and age-matched controls. In priming tasks, the processing of a word is usually facilitated by the earlier presentation of a semantically related prime word. Adults with WS showed normal priming effects for both functionally related items (e.g. broom – floor), and taxonomically related items. (e.g. coat – hat), suggesting normal semantic organization. In a related study, Neville, Mills, and Bellugi (1994) recorded event related potentials (ERPs) during processing of visual and auditory sentences. There was little difference in the ERP waveform for visual sentences but for auditory presentation, people with WS showed a slightly increased positive ERP response to words which were primed by the preceding sentential context (e.g. 'After the show they clapped their hands.'). Tyler et al. suggested that these findings could be reconciled by assuming that semantic organization is normal but people with WS have difficulty in integrating word meaning as they process sentences online. However, Neville et al.'s results indicate that, if anything, priming by sentential context is enhanced in WS, so there is, in fact, no indication of a deficit in semantic integration.

To summarise, several studies have suggested that semantic representations or access to semantic information is somehow impaired or atypical in WS. However, subsequent studies addressing methodological issues such as the group size and the control measures have failed to replicate these findings. There is therefore little evidence that semantic organization or access in WS is in any way unusual.

2.6: Constraints on vocabulary acquisition

Two studies have looked at the constraints used by individuals with WS when determining the meaning of new words. Mervis and Bertrand (1997) investigated the link between fast mapping (assuming that a novel word refers to an object which has no known name) and spontaneous exhaustive sorting (putting all the objects from a basic-level category together). In TD children and children with Down syndrome, the onset of both behaviours tends to coincide with the onset of a rapid growth in vocabulary. However, in WS, although the onsets of fast mapping and spontaneous sorting coincided, all the children had already experienced some form of vocabulary spurt. This could mean that children with WS acquire vocabulary in an unusual way. Alternatively, it could mean that the link between the vocabulary spurt and fast-mapping and spontaneous sorting in TD children is coincidental.

Stevens and Karmiloff-Smith (1997) conducted four studies in which they tested between 10 and 14 participants with WS with an age range of 7 to 32 years. Individuals with WS obeyed the fast-mapping constraint. However, because they were much older than the children tested by Mervis and Bertrand (1997), it was not possible to confirm any delay in its onset. They also obeyed the mutual exclusivity constraint, which stipulates that an object cannot have more than one name. However, they did not obey the whole object constraint, according to which a novel word heard in the presence of a novel object is taken to refer to the whole object rather than one of its features. They also failed to obey the taxonomic constraint. In other words, when given a novel word, X, for a known object and asked for another X, they failed to choose an object from the same category.

Together these findings suggest that individuals with WS may acquire vocabulary in a different way to TD children. Moreover, findings such as the delayed onset of categorical sorting relative to the vocabulary spurt and the failure to adhere to the taxonomic constraint in vocabulary acquisition would appear to fit well with suggestions that individuals with WS are better at learning the phonological forms of words than their meanings. Nevertheless, the precise nature of any causal relationship between these findings has not been determined, and, as Stevens and Karmiloff-Smith (1997) point out, it is not entirely clear how specific these findings are to WS.

2.7: Early nonverbal communication

Several studies have investigated early nonverbal communication in WS for evidence that language development follows an atypical trajectory. Singer-Harris et al. (1997) found that children with WS produced the same range of gestures as predicted by norms for TD children at the same language age. However, Masataka (2000) reported that, whereas TD two-year-olds showed a link between gestures and choice of a novel object, WS children matched on the size of their lexicon showed no evidence of such a link.

Mervis and Bertrand (1997) conducted a longitudinal study of eight toddlers with WS combining observations and parental reports. In all cases, the use of referential language appeared before the onset of production or comprehension of referential pointing, the opposite order to that observed in TD toddlers. Instead of pointing, parents often used alternative means for establishing joint attention such as placing an object in the location where the child was already looking or tapping the object. The authors argued that there were important practical implications of these findings. In particular, it is commonly assumed that children who do not use or comprehend referential pointing will not benefit from interventions to facilitate referential language, but this study shows that this is not the case in WS.

Consistent results were recently reported by Laing, Butterworth et al. (2002) who found that young children with WS initiated and responded to joint attention bids less than TD children matched on mental age using the Bayley Scales of Infant Development. The WS children also produced fewer instrumental and declarative gestures, but made considerable use of referential language. Moreover, social referencing behaviour in these children was not correlated with language ability as it was in the TD group. These studies of the precursors to language in WS provide further evidence that language development is not only delayed but is also atypical from the start.

2.8: Summary and theoretical interpretations

2.8.1: Summary

Two important claims have been made about language in WS. The first is that language is relatively spared in comparison with general cognitive abilities. The second claim is that there are specific strengths and weaknesses within the linguistic profile, indicative of atypical language development. In this chapter these two claims were evaluated relative to a null hypothesis that stated that language abilities in WS are what would be expected given overall abilities. While there is consistent evidence that receptive vocabulary is a relative strength in older children and adults, suggesting that vocabulary develops at a faster rate than overall abilities, there is little evidence that syntactical or morphological skills are any better than predicted by general cognitive abilities at any stage of development. It therefore appears incorrect to speak of language in WS as being relatively preserved. Instead, the relative advantage for receptive vocabulary over most other measures of linguistic competence supports the claim that the language profile in WS is uneven and therefore does not follow a typical developmental trajectory.

Several other claims have been made regarding specific deficits in irregular morphology and lexical-semantic processing. However, the evidence in support of these claims is less than convincing and often, when methodological difficulties are ironed out, people with WS appear no different to TD controls at the same language level. One claim that does seem to have some empirical support is that individuals with WS are relatively good at learning the sound patterns of words but have a somewhat limited grasp of the meanings of the words.

2.8.2: Modular accounts of language in WS

As mentioned in Section 1.6, there are two competing accounts of the cognitive profile in WS. According to the modular approach, the cognitive profile reflects the juxtaposition of modules that are differentially spared and impaired. The uneven linguistic profile in WS argues against the idea that individuals with WS have a relatively preserved 'language module'. It is, however, consistent with the idea that there are multiple different modules within the language faculty and these may be independently preserved and impaired (cf. Levy, 1996).

In linguistic theory it has been argued that the language faculty consists of two components, a lexicon which stores information about category membership and idiosyncratic information regarding form and meaning, and a computational system of combinatorial operations to form larger expressions (e.g. Pinker, 1994). Several authors (e.g. Clahsen & Almazan, 1998; Pinker, 1991; Temple et al., 2002) have suggested that in WS, the lexical system is impaired but the computational system is contact, and that this mirrors the pattern observed in SLI (e.g. van der Lely & Ullman, 2001). However, the idea that different linguistic modules can develop independently is questionable (cf. Bishop, 1997; Karmiloff-Smith, 1997, 1998). Moreover, there is little empirical evidence for relative strengths in syntactical abilities in WS (see sections 2.3.1 and 2.3.2) and, while earlier studies provided evidence for selective deficits in irregular morphology (which, according to modular accounts, is assumed to rely on the lexical system) and for impaired lexical-semantic processing in WS (see sections 2.3.3 and 2.5), later studies addressing several important methodological issues have failed to replicate their findings.

2.8.3: Neuroconstructivist accounts of language in WS

In contrast to the modular approach, the neuroconstructivist approach (e.g. Karmiloff-Smith, 1998; Thomas & Karmiloff-Smith, 2002) considers that cognitive development in WS proceeds under an altered set of constraints. Although they are non-specific, these constraints may nevertheless influence certain abilities more than others leading to dissociations in overall performance across domains (e.g. a general advantage for verbal over nonverbal abilities). However, because the underlying constraints are different, this account predicts that closer inspection should reveal within-domain dissociations in performance. The neuroconstructivist account is therefore consistent with the evidence reviewed in this chapter, but the nature of the constraints relevant to WS has not been specified. Thomas and Karmiloff-Smith (2002) have recently argued that there is some form of imbalance between relatively preserved phonological processing and impaired semantic processing but, as these authors admit, the precise nature of this imbalance is unclear.

2.8.4: Language and verbal memory in WS

One possible way of investigating the constraints on language acquisition in WS is to consider the relationship between language and verbal memory abilities. Numerous studies have identified verbal STM as a relative strength in WS (e.g. Babb, Gage, Perales, Huntley-Fenner, & Hickok, 1998; Finnegan, Smith, Meschino, Vallance, & Sitarenios, 1995; Klein & Mervis, 1999; Udwin & Yule, 1991; Wang & Bellugi, 1994). For example, Mervis et al. (1999) reported that 73% of a sample of 104 children and adults with WS scored in the normal range for the digit span subtest of the DAS, despite the fact that most of these individuals had full-scale IQs in the mentally retarded range. Moreover, standard scores for digit span were greater than those for receptive vocabulary or grammar. Jarrold et al. (1999) found that once vocabulary was controlled for by covariation, digit span was the same in WS, MLD and TD children, suggesting that vocabulary and STM are at equivalent developmental levels in WS. Similarly, Laing et al. (2001) reported that individuals with WS and TD controls matched on vocabulary had equivalent nonword repetition performance, although Grant et al. (1997) found that age-equivalent scores for nonword repetition were at a lower level than vocabulary.

Gathercole and Baddeley (e.g. Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1993) have argued that verbal STM plays an important role in language learning, particularly in the acquisition of new vocabulary. Consequently, the relatively strong performance on STM tasks in WS has led several authors to argue that linguistic strengths in WS are a consequence of preserved STM (Bishop, 1999; Grant et al., 1997; Klein & Mervis, 1999; Mervis et al., 1999; Vicari, Carlesimo, Brizzolara, & Pezzini, 1996).

Further studies have suggested that the processes involved in memory performance in WS are atypical. Individuals with WS have been found to show normal effects of phonological factors such as word length and phonological similarity, but reduced effects of lexical-semantic factors such as lexicality and word-frequency (Majerus, Palmisano, van der Linden, Barisnikov, & Poncelet, 2001; Vicari, Carlesimo et al., 1996). Other studies have noted a reduced tendency for individuals with WS to misrepeat nonwords as real words (Karmiloff-Smith et al., 1997), and a reduced influence of wordlikeness on recall of nonwords (Majerus et al., 2001) suggesting, perhaps, that in WS phonological memory for nonwords is not influenced by knowledge of real words as it is in TD individuals. Finally, in a free recall study, individuals with WS showed a significant recency effect but no evidence of primacy, which is consistent with intact phonological STM and impaired semantic long-term memory (LTM; Vicari, Brizzolara et al., 1996).

Together these findings suggest a possible dissociation (or imbalance) between relatively good memory for phonological material and impaired memory for semantic material. However, the relationship between language and STM is complex and interactive so the nature of any causal relationship is unclear. Moreover, there are several methodological concerns with previous studies of STM in WS. These theoretical and methodological issues are explored in Parts III and IV of this thesis.

2.8.5: Language and temporal processing in WS

Before memory abilities in WS are investigated, Part II of this thesis investigates the hypothesis that the linguistic strengths in WS are a consequence of preserved temporal processing mechanisms. Temporal processing mechanisms are a good example of what Karmiloff-Smith (1998) terms 'domain-relevant' as opposed to 'domain-specific' mechanisms. That is, they are important for language but are also likely to be involved to a lesser extent in many other cognitive functions. As such, temporal processing mechanisms have not evolved specifically to enable language processing (indeed, there is compelling evidence that many non-human animals have similar temporal processing mechanisms; e.g. Gallistel, 1990). Nevertheless, the relative impairment or preservation of these mechanisms in a developmental disorder is likely to have a greater influence on linguistic abilities than on other cognitive faculties. Moreover, temporal processing is likely to have a greater influence on phonological processing, which is intrinsically temporal, than it will on semantic processing (cf. Boucher, 1999). Consequently, the relative preservation of temporal processing mechanisms could not only explain the overall advantage for verbal over nonverbal abilities in WS, but could also account for the apparent imbalance between phonology and semantics.

---

[2] The CDI only asks if words are 'understood' or 'understood and produced'.

Part II: Temporal discrimination

Chapter 3: Temporal processing and language

3.1: Overview of Part II

The evidence reviewed in Chapter 2 suggests that, while overall language abilities in WS may be superior to nonverbal abilities, there are certain strengths and weaknesses within the linguistic profile. The studies reported in Part II of this thesis investigated the hypothesis that linguistic strengths in WS are a consequence of the relative preservation of temporal processing mechanisms.

Language processing is intrinsically temporal in nature (cf. Elman, 1990; Federmeier & Kutas, 2000; Gupta & Dell, 1999). Analysing speech input involves processing a temporally extended acoustic signal, and converting it into a sequence of phonemes, words and sentences, while producing speech output requires an orderly sequence of motor instructions. Similarly, verbal STM tasks by definition require the representation, maintenance and output of arbitrary sequences of items. As such, language and verbal STM are fundamentally dependent on temporal processing abilities. Consistent with this view, Boucher (1999) has argued that deficits in temporal processing are the cause of SLI. Williams syndrome could therefore represent the opposite of SLI in that relatively preserved temporal processing abilities are the cause of linguistic strengths[3,4].

This chapter begins by looking at evidence in support of this hypothesis. Next, it reviews oscillator-based models of temporal processing upon which Boucher (1999) has based her account of SLI, and considers their implications for language and memory abilities in WS. It concludes with an overview of the timing experiments that are presented in the following two chapters.

3.2: Evidence for preserved temporal processing mechanisms in WS

3.2.1: Preservation of frontal and cerebellar regions in WS

Evidence from brain imaging studies, patients with acquired neurological impairments, and animal lesion studies has consistently implicated the cerebellum and the frontal cortex in temporal processing abilities, including the production of rhythmic motor patterns and the discrimination of temporal durations (Gibbon, Malpani, Dale, & Gallistel, 1997; Hazeltine, Helmuth, & Ivry, 1997; Ivry, 1993; Mangels, Ivry, & Shimizu, 1998; Nichelli, Alway, & Grafman, 1996; Penhune, Zatorre, & Evans, 1998). Studies of brain morphology in WS have shown that, although overall brain size is reduced relative to age-matched controls, frontal and cerebellar regions are relatively preserved (e.g. Jernigan et al., 1993; Laakmann, 2001; Reiss et al., 2000) [5] It has been suggested that the relative sparing of these structures might contribute to relative linguistic competence in WS (Bellugi, Lichtenberger et al., 1999) and, in support of this view, Laakmann (2001) reported that the surface area of the neocerebellum in WS was significantly correlated with performance on a syntax task. One possibility, is that this relationship is mediated by relatively preserved temporal processing.

3.2.2: Musical and rhythmical abilities in WS

Like language, music is intrinsically temporal. The fact that musical and rhythmical abilities seem to be a particular strength in WS therefore provides further indirect evidence for relatively good temporal processing abilities. Many individuals with WS have a fascination with music, whether it be listening, singing or playing an instrument (e.g. Lenhoff, 1998; Lenhoff, Wang, Greenberg, & Bellugi, 1997; Levitin & Bellugi, 1999; Marriage, 1996; Sacks, 1995). Hopyan, Dennis, Weksberg, and Cytrynbaum (2001) reported that children with WS were less able to discriminate pitch and rhythm than were controls matched on chronological age, suggesting that musical abilities are not preserved in the strict sense. However, Don, Schellenberg, and Rourke (1999) reported that individuals with WS performed comparably to TD controls matched on receptive vocabulary on tests requiring the discrimination of pairs of melodic or rhythmic fragments. Since vocabulary is a relative strength in WS, this suggests that musical and rhythmic abilities are also a relative strength. Moreover, Don et al. found significant correlations between measures of language and musical ability in both groups suggesting that similar mechanisms are involved.

3.3: Oscillator-based models of temporal processing

3.3.1: Boucher's time-parsing hypothesis

Gallistel (1990) noted that many species demonstrate oscillatory brain activity with periods ranging from milliseconds to years. He argued that temporal position can be encoded by recording the corresponding states of a number of oscillators, and the duration of an event can be computed by comparing the states of the oscillators at its beginning and end. Based on Gallistel's work, Boucher (1999) argued that SLI could be caused by deficits in oscillatory timing mechanisms. Specifically, she suggested that oscillators of varying frequencies provide a hierarchy of temporal windows which allow the parsing of incoming speech into its component words, rimes, and phonemes, and that the absence or late development of high-frequency oscillators could lead to phonological SLI. If this argument is extended to WS then the relatively good phonological skills of people with WS could be a consequence of the relative preservation of these high-frequency oscillators.

However, while formal oscillator-based models have been applied to temporal discrimination (Church & Broadbent, 1990), serial order memory (Brown, Preece, & Hulme, 2000) and speech production (Vousden, Brown & Harley, 2000), they have not yet been applied to speech perception[6] , so it is unclear exactly how oscillators may be involved in parsing of speech input. Nevertheless, if high-frequency oscillators are deficient in SLI and relatively spared in WS, these more formal models do allow certain predictions to be made regarding performance on temporal discrimination and serial order memory tasks. The following sections therefore provide a brief overview of oscillator-based models of timing and serial order memory and their implications for WS and SLI.

3.3.2: OSCAR

Brown et al. (2000) have recently proposed OSCAR (standing for OSCillator-based Associative Recall), a computational model of serial order memory. In OSCAR, learning a list involves the formation of associations between vectors representing list items and a so-called 'temporal context vector' which represents the states of an array of endogenous oscillators. Each element of the temporal context vector represents the product of a subset of oscillators at a particular time. Recall is achieved by reinstating the temporal context for the start of the list and allowing the temporal context vector to proceed of its own accord. Throughout recall, items compete to be output but the competition is influenced by the temporal context such that the item associated with the current context is most strongly activated and therefore the most likely to be produced.

OSCAR is similar to several other recent models of serial order memory (e.g. Burgess & Hitch, 1999; Henson, 1998) although it differs in the precise nature of the temporal context signal. However, it contrasts with chaining models (e.g. Lewandowsky & Murdock, 1989) which assume that each item in a sequence provides the cue for the next element. It also contrasts with graded activation models (e.g. Page & Norris, 2000) in which earlier responses have higher activation levels so are performed first before being inhibited to allow later items to be output.

OSCAR has been applied to a wide range of data including developmental changes in serial recall performance (Brown, Vousden, McCormack, & Hulme, 1999). McCormack, Brown, Vousden, and Henson (2000) found that developmental improvement in serial recall performance was mainly the result of a reduction in the number of movement errors (producing list items in the incorrect position). Moreover, movement gradients showing the proportion of movement errors at a particular distance from the correct position increased in steepness with developmental level. In other words, whereas most of the adults' movement errors involved short-distance displacements, young children made a much larger number of movement errors with large displacements. Brown et al. (1999) found that the best fit to these data was obtained by varying the distinctiveness of the temporal context signal. In OSCAR this corresponds to the speed with which the learning-context vector changes over time, which is determined by the frequency of highest-frequency (fastest changing) oscillator.

If, as Boucher (1999) suggests, high-frequency oscillators are relatively slow to develop in SLI, then individuals with SLI would be expected to show increased movement errors and shallow movement gradients. Conversely, individuals with WS might be expected to show the opposite pattern. This prediction is tested in chapter 7 which reports a study of probed serial recall in WS using a procedure designed to allow investigation of order errors.

3.3.3: Church and Broadbent's model of time perception

Church and Broadbent (1990) proposed a computational model of timing behaviour according to which duration is represented by the correlation between the states of each of a number of oscillators with different periods. The main principle is that when two durations are similar, the correlations between the states of the oscillators should also be similar and this allows comparisons to be made between the duration of the current stimulus and previously experienced durations.

The major alternatives to this model are those which assume that the duration of an event is represented by the number of pulses produced by a pacemaker in that interval (e.g. Creelman, 1962; Gibbon, 1977; Gibbon, Church, & Meck, 1984; Treisman, 1963). Both oscillator- and pacemaker-based models have been successfully applied to a wide range of timing data in humans and other animals (see Wearden, 1999). However, pacemaker models require some form of accumulator to record the number of pulses produced during an event and it is not clear how an accumulator process might be realised in terms of known neurological mechanisms (cf. Matell & Meck, 2000).

One consequence of Church and Broadbent's model is that accurate representation of an interval requires oscillators with similar periods to its duration. This is because the state of an oscillator with a much longer period would be similar at the start and end of the interval so would provide poor discrimination (it would be like trying to time 3 seconds using the hour hand of a watch), while an oscillator with a much shorter period would have changed state so many times that any noise in the system would essentially leave it in a random state. Consequently, deficits in high-frequency oscillators in SLI would lead to poor discrimination of temporal intervals in the corresponding range. In contrast, individuals with WS should perform relatively well on the same tasks.

3.4: Summary and overview of timing experiments

The hypothesis proposed in this chapter is that individuals with WS have relatively preserved timing abilities. This hypothesis is motivated by the intrinsically temporal nature of language and the idea that selectively preserved temporal processing abilities could have a causal role in the relatively good language abilities found in WS. Further support for the hypothesis comes from the relative strengths in music and rhythm which also rely on temporal processing, and the relative preservation of frontal and cerebellar regions which are involved in timing behaviour. Finally, relatively preserved timing abilities in WS would also contrast with hypothesised temporal processing deficits in SLI. In particular, Boucher (1999) has proposed that SLI is caused by deficits in oscillatory timing mechanisms, and it is possible that these same mechanisms are relatively preserved in WS.

Chapter 4 reports a study that tested the prediction that individuals with WS will perform relatively well on tasks requiring the discrimination of temporal durations using two paradigms, temporal bisection and temporal generalization, which have been developed for use with children. Chapter 4 also presents a follow-up study using a modified temporal generalization task, while chapter 5 presents a study with three novel timing tasks which address problems with the bisection and generalization tasks.

---

[3] Boucher's (1999) temporal processing account of SLI is different to that proposed by Tallal and colleagues (e.g. Tallal & Piercy, 1973) based on the finding that children with SLI perform poorly on tasks requiring judgements of the temporal order of rapidly presented stimuli. In fact, it has been argued that such deficits reflect a more general problem with rapid information processing in SLI (e.g. Studdert-Kennedy & Mody, 1995). Moreover, performance of controls on these 'temporal discrimination' tasks is typically at ceiling, so it would be difficult to show that children with WS perform better than controls.

[4] Several authors (e.g. Clahsen & Almazan, 1998; Pinker, 1999) have previously argued that WS and SLI provide a developmental double dissociation between lexical (impaired in WS but intact in SLI) and computational (intact in WS but impaired in SLI) modules. However, such a dissociation is at the level of highly specific linguistic modules, and this contrasts with the more general temporal processing mechanisms considered here. It is nevertheless conceivable that differences in temporal processing abilities could be the underlying cause of an apparent modular linguistic dissociation, but this is beyond the scope of the current chapter.

[5] Rae et al. (1998) reported significant biochemical abnormalities in the cerebellum of individuals with WS but it is unclear whether these results reflect abnormalities which are specific to the cerebellum or reflect global brain biochemistry in WS.

[6] Current models of speech perception have difficulties representing time (see e.g. Protopapas, 1999) so this could represent a promising application of oscillator-based models.

Chapter 4: Temporal bisection and generalisation

4.1: Introduction

Chapter 4 introduced the hypothesis that individuals with WS have relatively preserved timing abilities. This hypothesis could potentially explain the relatively good language abilities in WS, and fits well with the good rhythmical and musical abilities found in WS, as well as being consistent with the relative preservation of the cerebellum which is known to be involved in timing behaviour. In this chapter and the following one, a series of studies are reported which investigated this hypothesis.

When talking about relatively preserved abilities, the crucial question is 'relative to what?' Although temporal processing abilities may have a crucial role in determining linguistic competence, there are clearly many other important factors. Consequently, if the argument is that certain language abilities in WS are relatively good because temporal processing abilities are relatively spared, then one might expect performance on timing tasks to be above the level predicted by language ability because weaknesses in other contributory factors will impair language abilities. In the studies in this chapter and in chapter 5, the performance of children with WS on various timing tasks was therefore compared with that of controls matched on BPVS scores.

A further consideration was the range of durations to be discriminated. If, as Boucher (1999) has suggested, phonological abilities rely on high-frequency oscillators and these are relatively impaired in SLI but relatively preserved in WS, then oscillator-based models of timing (Church & Broadbent, 1990) predict that individuals with WS should perform relatively well on discrimination of very short durations. Another reason for choosing short durations is that with durations much above one second, participants may start to use a counting strategy to aid discrimination. Potential group differences may therefore reflect differences in strategy use rather than pure temporal discrimination abilities.

The choice of the timing tasks was also important. Several paradigms have been developed to assess judgements of temporal duration but most of these are unsuitable for use with children. For example, asking participants to estimate the duration of a presented stimulus relies on their ability to use a scale, while requiring participants to reproduce a response of a certain duration or responses separated by the target duration is confounded with children's ability to initiate or inhibit motor responses. Instead, recent studies have investigated developmental changes in temporal discrimination abilities using two paradigms, temporal bisection and temporal generalisation, derived from studies of animal learning.

The following sections present two preliminary studies of timing abilities in WS using the temporal bisection and temporal generalisation tasks developed for use with children in a previous study by McCormack, Brown, Maylor, Darby, and Green (1999).

4.2: Experiment 1: Temporal Bisection and Generalisation

4.2.1: Introduction

In the temporal bisection paradigm (e.g. Allan & Gibson, 1991; Church & Deluty, 1977; Wearden, 1991), participants are given two standard durations, one long and one short. They are then presented with a range of test durations and have to decide whether each test duration is more similar to the short or the long standard duration. When the proportion of 'long' responses is plotted against the actual duration of the stimulus, an S-shaped curve is usually observed, with short durations receiving relatively few 'long' responses and long durations receiving mostly 'long' responses. A measure of temporal discrimination is provided by the steepness of the curve, a steep curve being indicative of good discrimination between durations. McCormack et al. (1999) recently adapted the paradigm for children using tones of different durations and found that five-year-olds showed flatter curves than eight-year-olds and older children and adults, indicating developmental improvement in temporal discrimination. Similarly, Droit-Volet and Wearden (2001) reported that eight-year-olds showed steeper curves than three- or five-year-olds in a task using visual stimuli.

In the temporal generalisation paradigm (Church & Gibbon, 1982; Wearden, 1992), participants learn one standard duration and are then presented with different test durations which they must classify as either the same as or different to the standard duration. When the proportion of positive responses are plotted against the duration of the test tone, a bell-shaped curve is typically observed which is usually centred at or near to the actual standard duration, and the steepness of the curve gives a measure of the participant's ability to discriminate between different durations. McCormack et al. (1999) adapted this paradigm for children using auditory stimuli of differing durations, and found that ten-year-olds showed steeper curves than five-year-olds. Similarly, Droit-Volet, Clement and Wearden (2001) showed developmental increases in steepness of the generalisation curve using visual stimuli.

In Experiment 1, children with WS and MLD controls matched on BPVS scores and chronological age were tested on the bisection and generalisation tasks developed by McCormack et al. (1999) for use with children. The prediction was that individuals with WS should produce steeper bisection and generalisation curves than these controls. It was also possible to compare the performance of these two groups with that of TD five- and eight-year-olds tested by McCormack et al..

Table 4.1 Participant details for Experiment 1

	WS N=14		MLD N=14	
	M	SD	M	SD
Chronological age (years; months)	12;5	2;7	12;11	1;7
BPVS age (years; months)	7;5	1;4	7;9	1;7

4.2.2: Method

4.2.2.1: Participants

Two groups of children took part in this study. 14 children with WS (9 boys) were recruited via the Williams Syndrome Foundation. The MLD group were 14 children (6 boys) attending a special school for children with learning disabilities. None of the children in either group had any history of hearing impairment. The groups were matched on receptive vocabulary using the BPVS and on chronological age. Group characteristics are shown in Table 4.1.

4.2.2.2: Apparatus and Stimuli

The experiment was run on a Macintosh PowerBook G3 laptop computer and stimulus presentation was controlled using the PsyScope software package (Cohen, MacWhinney, Flatt, & Provost, 1993). The auditory stimuli were the same as those used by McCormack et al. (1999). All stimuli were 500 Hz pure tones and were generated using the software package SoundEdit and presented via the computer's internal speaker as some of the WS children were unwilling to wear headphones. For the bisection task, the short standard duration was 200ms, the long standard was 800 ms and the nonstandard stimuli durations were 300, 400, 500, 600 and 700 ms. For the generalisation task, the standard duration was 500ms with nonstandard durations of 125, 250, 375, 625, 750 and 875 ms.

4.2.2.3: Procedure

Children with WS were tested in a quiet room at home. Children in the MLD group were tested in a quiet room in school. Half the children in each group performed the bisection task first, while the other half performed the generalisation task first.

Temporal bisection

The procedure was the same as that followed by McCormack et al. (1999). In the initial exposure phase, participants were shown a display in which two birds, one small and one large, appeared side by side on the computer screen (Figure 4.1 upper panel). The children were informed that the birds made sounds of different lengths, with the small bird making a short sound and the large bird making a long sound. They were then given five alternating presentations of the short and long standards, with the appropriate bird appearing on the screen during the presentation of the sound.

Participants were then told that they would hear some more sounds and would have to judge whether the sounds were made by the big bird or the little bird. If they thought the sound was made by the big bird they should point at the big bird or say 'big bird'. If they thought the sound was made by the little bird they should point at the little bird or say 'little bird'. There then followed a practice session in which participants heard each of the seven durations in a random order while both birds were displayed on the screen. The experimenter recorded the responses by pressing '1' if the participant indicated that the big bird had made the sound and pressing '2' for the little bird. There was no feedback. The test stage consisted of five further blocks identical to the practice stage.

Temporal generalisation

The procedure was identical that followed by McCormack et al. (1999) when testing children, except in terms of the number of blocks in the test phase. In the first phase, a picture of an owl appeared in the centre of the screen. Participants were told that this was Barney and that he always made a sound of the same length. They were instructed to listen to Barney's sound and to try and remember how long his sound was. The standard tone was then presented five times.

In the second phase, three more owls appeared on the screen. The child was told that these were Barney's friends and that they all made different sounds to Barney, some longer and some shorter. The experimenter then demonstrated the task. Two pictures appeared side-by-side on the screen. On the left was a picture of Barney; on the right, a picture of Barney crossed out with a large semi-transparent red cross (Figure 4.1 lower panel). The 750ms tone was played and the experimenter pointed to Barney crossed out, explaining that this was because the sound was too long and therefore was not Barney's sound but one of his friends'. This was followed by presentation of the standard tone whereupon the experimenter pointed to Barney, and further demonstrations with a 250ms tone and another standard tone.

In phase 3, participants heard tones of duration 125, 375, 625, 875 and 500ms in that order and were instructed to point to the corresponding picture or say 'Barney' or 'not Barney'. Once a response had been made, participants were informed whether it was correct, too long or too short.

In phase 4, Barney's sound was played five more times with his picture again displayed in the centre of the screen.

Phase 5, the test stage, consisted of five blocks (McCormack et al. tested participants on eight). Each block contained two standard tones and one each of the nonstandard tones in a random order. After each trial the message 'That was Barney's sound' or 'That was not Barney's sound' appeared at the bottom of the screen and was read out by the experimenter.

Figure 4.1 Displays for the temporal bisection (upper panel) and temporal generalisation (lower panel) tasks

4.2.3: Results

Temporal bisection

The upper panel of Figure 4.2 shows the proportion of 'long' responses to each tone in the bisection task. The WS group had a steeper bisection curve than MLD controls. McCormack et al. (1999) omitted data from participants if the difference in proportion of 'big' responses to any two stimuli was never greater than 0.6, indicating a random response pattern. All of the children with WS passed this criterion. However, one child in the MLD group failed and a further two members of the MLD group only passed because they gave very few 'big' responses to one of the two longest stimuli, indicating systematically incorrect responding. Because of the small size of the groups, results are reported with all participants included, except where excluding participants led to qualitative differences.

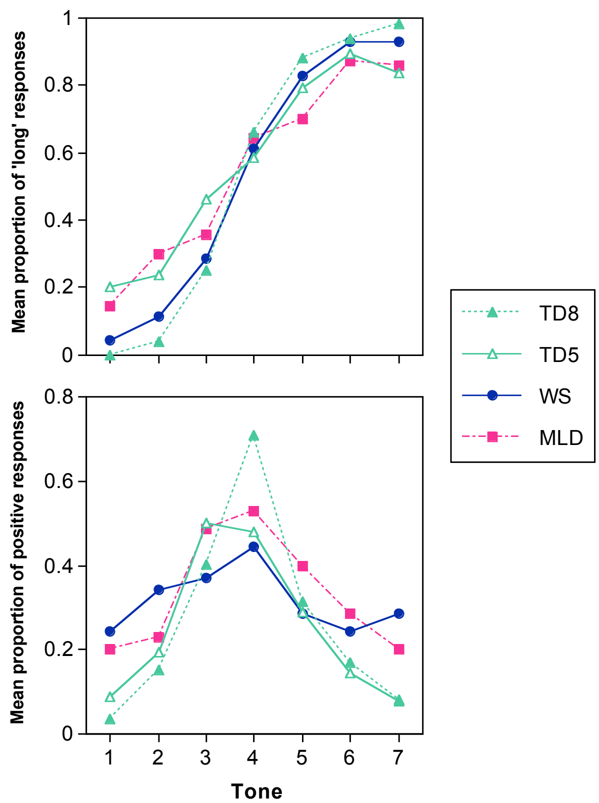

Figure 4.2 Bisection (upper panel) and generalisation (lower panel) curves for TD eight-year-olds (TD8), TD five-year-olds (TD5), children with WS and MLD controls

Following McCormack et al. (1999), the results were analysed by looking at the interaction between group and duration. However, because of ceiling and floor effects, the results for the very long and very short durations were significantly skewed, and since they were skewed in opposite directions, it was not possible to find an appropriate transform. However, this problem was resolved by collapsing results for the two shortest durations and for the two longest durations. There was a significant effect of duration (F(4,104)=58.86, p<.001), but the interaction between duration and group failed to reach significance (F(4,104)=1.78, p=.14). When the three MLD children who failed to pass the criteria were excluded, the interaction between group and duration was further reduced (F(4,92)=0.55).

Temporal generalisation

The lower panel of Figure 4.2 shows the proportion of positive responses to each tone in the generalisation task.[7] Contrary to predictions, the WS group performed worse than the MLD group showing a flatter generalization curve. McCormack et al. (1999) omitted participants' data if the largest difference in proportion of positive responses was never greater than 0.5. Using this criterion, three children with WS and two MLD children could be rejected. However, data from a further three WS and two MLD children could be rejected if it is further required that the stimulus with the most positive responses must be one of the middle three, thus eliminating systematically incorrect as well as random response patterns.

Following McCormack et al. (1999), the results were subjected to ANOVA with group and duration as factors. There was a significant effect of duration (F(6,156)=4.02, p<.001) but no effect of group (F(1,26)=0.18) and no group by duration interaction (F(6,156)=0.94). To determine whether the individual groups were performing above chance levels, a one-way ANOVA was performed for each group with duration as a repeated measure. This proved significant for the MLD group (F(6,78)=5.36, p<.001) reflecting a greater proportion of positive responses to the standard tone than to the two shortest and two longest tones. In contrast, the effect of duration was not significant for the WS group (F(6,78)=0.86).

Comparison with data collected by McCormack et al. (1999)

Figure 4.2 also shows data from 26 TD five-year-olds and 32 TD eight-year-olds tested by McCormack et al. (1999). The WS group showed a steeper bisection curve than the five-year-olds, but a slightly shallower bisection curve than eight-year-olds. In contrast, they showed much flatter generalisation curves than either of the TD groups.

The performance of the WS group was compared with that of the TD five-year-olds.[8] On the bisection task there was a significant effect of duration (F(2.77,152)=69.70, p<.001; Greenhouse-Geisser correction applied), but no significant effect of group (F(1,38)=0.67). The interaction between group and duration approached significance (F(2.77, 152)=2.28, p=.089) but fell well short of significance when four TD children who failed to pass the criterion were excluded (F(2.57, 136)=1.13).

On the generalisation task, the TD five-year-olds made very few positive responses to the extreme short and long stimuli. It was therefore necessary to collapse results for the two shortest and the two longest stimuli in order to allow comparison with the WS group. ANOVA revealed a significant effect of duration (F(2.99,152)=12.84, p<.01; Greenhouse-Geisser correction applied) but no main effect of group (F(1,38)=0.65). The interaction between duration and group was significant (F(2.99, 152)=3.13, p<.05), and analysis of simple effects of group showed that the WS group made more positive responses to the two shortest stimuli (t(38)=2.47, p<.05) although, when the number of comparisons was taken into account, this difference was non-significant. When the six children with WS who failed to pass the criteria were excluded, the interaction between group and duration failed to reach significance (F(2.176, 76)=2.11, p=.13; Greenhouse-Geisser correction applied).

4.2.4: Discussion

Experiment 1 used temporal bisection and generalisation tasks to investigate the hypothesis that individuals with WS have relatively spared temporal discrimination abilities. The hypothesis predicted that children with WS would show steeper bisection and generalisation curves than MLD controls.

In the bisection task, the trend was in this direction but the interaction between group and duration failed to reach significance. It is possible that a more sensitive task would result in a significant group difference. Alternatively, it may be that more of the participants in the control groups failed to understand the task properly. Indeed, when the control children who failed to meet the criteria for inclusion were eliminated, there was very little evidence for any group differences. However, in such circumstances it is unclear whether children have failed to meet the criteria because they did not understand the task or because they had poor temporal discrimination.

In the generalisation task, the WS group actually showed shallower generalization curves than MLD controls and their response distributions were not significantly above chance, although the interaction between group and duration was again nonsignificant. The poor performance of the WS group on the generalisation task could reflect impaired temporal discrimination in WS (the opposite of the proposed hypothesis) but this is difficult to reconcile with the relatively good performance on the bisection task. Indeed, the WS group also showed significantly shallower generalisation curves than TD five-year-olds tested by McCormack et al. (1999), despite showing steeper bisection curves.

Two main issues remain to be addressed. The first concerns the original hypothesis that individuals with WS have relatively good temporal processing abilities. The results of these studies neither supported this hypothesis, nor did they convincingly rule it out. The bisection and generalisation tasks were chosen because they have previously been used successfully in studies with young children, but with hindsight there are several problems with the paradigms which perhaps make them unsuitable for testing the current hypothesis. One problem affecting both tasks is that group differences in performance are determined by looking for an interaction between group and duration rather than a main effect of group, thus making the tasks relatively insensitive. Scoring systems can be devised which provide a way of looking at main effects of group, but are not particularly satisfactory. For example, scoring the bisection task is difficult because there are no objectively correct responses for all but the extreme stimuli.

A further difficulty particular to the bisection task is that averaging across participants to produce bisection curves can underestimate discrimination (see Figure 4.3). Various techniques have been derived for calculating the steepness of bisection curves on an individual basis. Unfortunately, these require a much greater number of trials than were used in Experiment 1.

A problem with the generalisation task is that discrimination performance is confounded with response bias because there are far more non-standard trials than there are standard trials. Thus a participant with a positive response bias is likely to perform worse than one with a negative bias even if they have the same discrimination abilities. There was little evidence for group differences in response biases in Experiment 1, as indicated by the lack of any group main effects. Nevertheless, individual variation in response biases makes it less likely that any significant interactions will be found, and controlling for this by normalizing the data would have led to problems of non-independence.

In sum, tasks aimed at testing the temporal discrimination hypothesis should have objectively correct responses for each trial thus allowing participants to be given an individual score and also ensuring that the steepness of averaged response curves reflect discrimination abilities and not variations in subjective cut-off points. Moreover, perfect performance on these tasks should also require an equal number of responses of each type thus avoiding the confounding influence of response bias on discrimination.

The second issue concerns the reason for the relatively poor performance of the participants with WS on the generalisation task. One possible explanation is that the children with WS did not understand negatives (i.e. 'not Barney'), but several studies have shown that people with WS have no problem comprehending negatives (Bellugi et al., 1990, 1992; Clahsen & Almazan, 1998). Another possibility is that individuals with WS had an increased tendency to produce random responses. Droit-Volet et al. (2001) found that developmental changes in generalisation performance could be accounted for by a combination of increased discrimination abilities and reduced random responding. However, it is unclear why children with WS should produce more random responses than controls in the generalisation task but clearly produced few such responses in the bisection task.

One way of understanding this pattern of results is to look at the differences between the bisection and generalisation tasks. For example, generalisation can be seen as requiring a judgement of absolute duration (standard or non-standard) rather than a categorical decision (long or short). Alternatively, generalisation can be seen as involving a three-way (too short, standard, too long) rather than two-way (long or short) categorical decision. Finally, the bisection tasks requires a simple mapping from decision to response (short = 'little bird', long = 'big bird') whereas generalisation requires a more complex mapping (standard = 'Barney', too short or too long = 'not Barney'). It is possible that these differences can explain the relative difficulty faced by the children with WS.

Chapter 5 reports a study of timing abilities in WS that addresses these two issues using three new temporal discrimination tasks.

---

[7] McCormack et al. (1999) reported normalized data whereby the number of positive responses to each tone was divided by the total number of positive responses across all tones. This approach avoids the possible confounding influences of response bias, but leads to a problem with the non-independence of the data. Given that there was no evidence of group differences in response bias (see below), the un-normalized data are reported here.

[8] It was not possible to find an appropriate transformation to allow comparison of the WS group and the TD eight-year-olds on the bisection task. For the generalisation task, the significant interaction between Group and Duration in the comparison of the WS and TD five-year-olds renders the comparison with the eight-year-olds unneccessary.

Chapter 5: Temporal discrimination

5.1: Issues arising from Chapter 4

Chapter 4 reported studies of temporal bisection and generalization in WS that tested the hypothesis that individuals with WS have relatively preserved temporal processing mechanisms. However, the results were somewhat inconclusive in that children with WS performed better than controls on one task, temporal bisection, but worse on the other, temporal generalisation. The two main issues arising from these studies were (a) the inadequacies of the two tasks for identifying group differences in temporal discrimination abilities, and (b) the discrepancies between performance on the bisection and generalisation tasks. These issues were addressed in Experiment 2.

5.2: Experiment 2: Temporal discrimination

5.2.1: Introduction

Given the above discussion (see also section 4.2.4), three novel temporal discrimination tasks were developed, each of which involved responding to six tones which varied in duration. Discrimination of duration follows Weber's law such that performance depends on the ratio of the durations rather than the absolute difference in their magnitude (e.g. Gibbon, 1977). The tones in Experiment 2 were therefore arranged on a geometric scale (i.e. each tone was a fixed ratio longer than the next shortest tone) so that the tones were evenly distributed in psychological space (c.f. Brown, Neath, & Chater, 2002).

In the first timing task the three shortest sounds were made by a small hippo while the three longest tones were produced by a large hippo. In the second, the two shortest sounds were made by a small bear, the two longest by a large bear and the remaining middle two tones were produced by a medium-sized bear. In the third task, there were six frogs of increasing size. Each frog had one sound and the bigger the frog the longer its sound. Thus for all three tasks, each trial had a correct response, and each tone was presented equally often so each possible response was correct an equal number of times.

Comparison of performance on the three tasks also allowed investigation of several of the possible explanations discussed in section 4.2.4 for the relatively poor performance of children with WS on the temporal generalisation task. Thus, if children with WS had a problem with three-way as opposed to two-way categorisation they should perform poorly on the bears task compared with the hippos. If, however, their difficulty was with absolute rather than categorical judgements then the frogs task should prove particularly difficult. Finally, if the difficulty with the generalisation task resulted from problems understanding the mapping from decision to response then there should be no group by task interaction as all three tasks involve one-to-one mapping.

Whether or not any of these explanations proved correct, the hypothesis that individuals with WS have relatively spared temporal processing abilities predicted that the WS group would perform better than controls on the hippos task. To ensure that any group differences on the hippos task were specific to discrimination of duration, participants were also tested on a control pitch discrimination task. This was similar to the hippos task but featured two chickens, the larger of which made the three lowest sounds and the smaller of which made the three highest sounds.

Participants in the WS group were the same children who had taken part in the first experiment. A new group of MLD control children were recruited who were matched to the WS group on receptive vocabulary and were of similar chronological age. In addition, a group of TD children were also tested. These children were matched to the two other groups in terms of vocabulary knowledge but were significantly younger.

Table 5.1 Participant details for Experiment 2

	WS N=14		TD N=20		MLD N=13	
	M	SD	M	SD	M	SD
Chronological age (y;m)	13;5	2;5	7;5	0;9	12;6	1;9
BPVS–II age (y;m)	8;2	1;2	8;2	1;4	8;1	1;3
BPVS-II scaled score	N/A	N/A	106.2	8.7	N/A	N/A
Ravens matrices age (y;m) †	7;6	1;8	8;9	1;2	8;1	2;0

† Based on N=13 for the WS group and N=12 for the MLD group

5.2.2: Method

5.2.2.1: Participants

The WS group were the same 14 children (9 boys) who had taken part in Experiment 1. Sixteen children with other learning disabilities (MLD) were recruited from the same school as the MLD children from Experiment 1. However, three were unable to complete the bears and frogs tasks so were not included. Of the remaining 13 MLD children (8 of whom were boys), 4 had taken part in Experiment 1. Finally 20 TD controls (10 boys) were recruited from two mainstream primary schools. These children were selected by their teachers to represent the middle ability range of children in their class. Participant details are shown in Table 5.1. Groups were closely matched on vocabulary age but the TD group was significantly younger than the other two groups. The WS group were slightly older than the MLD group but not significantly so ($t(25)=1.45$). A one-way ANOVA revealed significant group differences in performance on the Ravens Coloured Progressive Matrices (Raven, 1993; $F(2,44)=3.278$, $p<.05$), with the WS group performing significantly worse than TD controls (Tukey HSD).

5.2.2.2: Stimuli and Apparatus

The experiment was run on a Macintosh PowerBook G3 laptop computer and stimulus presentation was controlled using the PsyScope software package. Auditory stimuli for all four tasks were generated using SoundEdit and presented via the computer's internal speaker. For the temporal discrimination tasks, the stimuli were six 400 Hz pure tones with durations of 300, 375, 469, 586, 732, and 916 ms such that each tone was 1.25 times the duration of the next shortest duration. For the pitch discrimination task, the stimuli were six 500 ms pure tones with frequencies of 400.0, 410, 420.3, 430.8, 441.5, and 452.6 Hz such that each tone was 1.025 times the frequency of the next lowest tone.

5.2.2.3: Procedure

Testing took place in a quiet room either at home in the case of the WS children or in school in the case of controls. The three temporal discrimination tasks were presented in counter-balanced order such that an approximately equal number of participants within each group completed the tasks in each of the six possible orders. Testing was conducted at the same time as the probed recall study reported in chapter 7 (Experiment 3), with the participants completing a block of the probed recall study between each temporal discrimination task. Where possible, all three tasks were completed within the same session although practical constraints entailed that this was not possible for some of the control children. The pitch discrimination (chickens) task was performed after the three timing tasks. In most cases this involved a separate session, although some of the older TD control children performed all four tasks within one session. For all but one of the WS children, the chickens task was undertaken in a separate session to the temporal discrimination tasks.

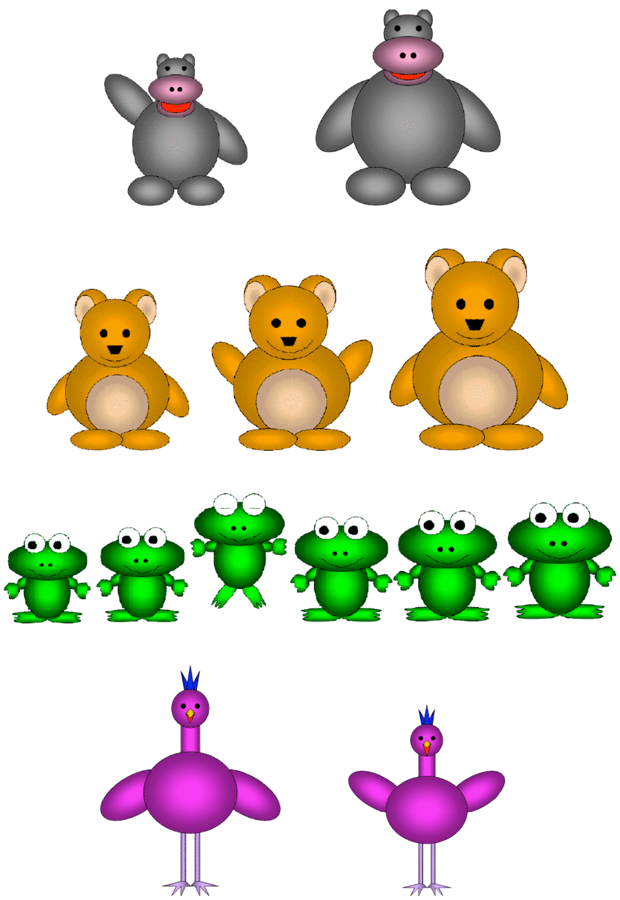

Figure 5.1 Displays for the hippos, bears, frogs and chickens tasks

The hippos task began with both hippos side-by-side on the screen. The experimenter explained that the big hippo made long sounds and the little hippo made short sounds. The six tones were then played in order of increasing duration. After each tone there was a delay of 500 ms and then the corresponding hippo waved (see Figure 5.1). The participant was then told that they would hear lots of different sounds and would have to decide which hippo had made the sound. The hippo that really had made the sound would then wave and they would be able to see whether they had answered correctly. In the test phase there were seven blocks although there were no obvious breaks between blocks. In each block all six tones were presented once in a random order. The first block was treated as a practice block and the responses on these trials were not analysed.

The procedures for the bears, frogs and chickens tasks were almost identical to that for the hippos task. The differences were in the display (see Figure 5.1), the explanations of which animal made which sound (see above), and the fact that frogs jumped and chickens flapped rather than waved.

5.2.3: Results

Overall performance

The left panel of Figure 5.2 shows mean performance for each timing task in terms of the proportion of correct responses. The pattern of performance across the three timing tasks was similar for each group, with the proportion of correct responses being greatest in the hippos task and least in the frogs task. A mixed-design ANOVA was conducted with group (WS vs TD vs MLD) as a between-subjects factor and task (hippos vs bears vs frogs) as a repeated measure. There was a significant main effect of task ($F(2,88)=299.75$, $p<.001$), and post-hoc tests revealed that performance on the hippos task was better than that on the bears task which in turn was better than performance on the frogs task. However, there was no main effect of group ($F(2,44)=0.07$) and no interaction between group and task ($F(4,88)=0.37$). These results suggest little difference in the temporal discrimination abilities of the three groups.

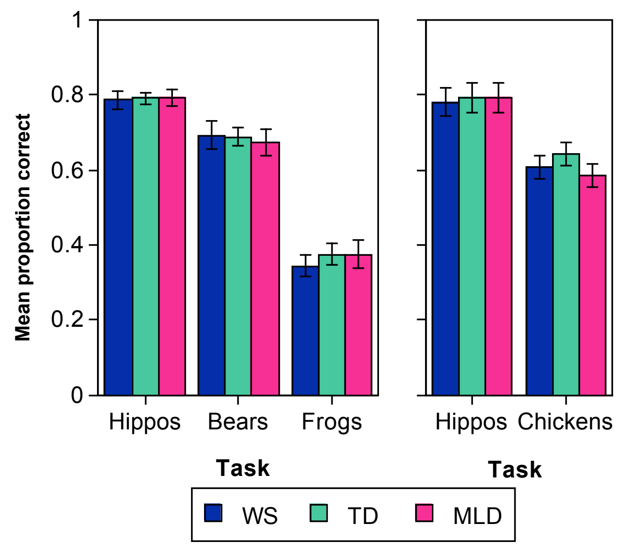

Figure 5.2 Overall performance on hippos, bears, frogs, and chickens tasks. Errors bars show standard errors

Due to time constraints, two children in the WS were unable to complete the chickens task. The right panel of Figure 5.2 shows the performance of the remaining participants on the chickens task compared with their performance on the hippos task. The chickens task was clearly the more difficult of the two, with many participants scoring close to or even below chance levels. ANOVA with group as between-subjects factor and task as a repeated measure confirmed a significant effect of task ($F(1,42)=68.30$, $p<.001$), but there was no significant effect of group ($F(2,42)=0.68$) and no interaction ($F(2,42)=0.66$).

The overall performance of the three groups across the four tasks gave no indication of any group differences. The following analyses looked at performance on the individual tasks to see whether there was any evidence for group differences in the way that overall scores were achieved.

Hippos task

Figure 5.3 shows the distribution of 'long' responses to the six tones in the hippos task. There is no suggestion of any group differences in performance, and this was confirmed by ANOVA. Performance on the middle four tones only was analysed thus avoiding problems with kurtosis and skewness due to ceiling and floor effects for extreme items, and ensuring independence of the data. There was a significant effect of tone ($F(3,132)=123.68$, $p<.001$) but no effect of group indicating equivalent performance on the extreme tones ($F(2,44)=0.82$). The interaction between group and tone was non-significant ($F(6,132)=0.45$).

Bears task

Figure 5.4 shows the distribution of 'middle' responses to the six tones in the bears task. As with the hippos task there were problems of kurtosis for the highest and lowest tones so ANOVA was performed on the middle four tones. The effect of tone was significant ($F(2.22,132)=11.65$, $p<.001$; Greenhouse-Geisser correction applied) but there was no effect of group ($F(2,44)=1.03$) and no tone by group interaction ($F(6,132)=0.46$).

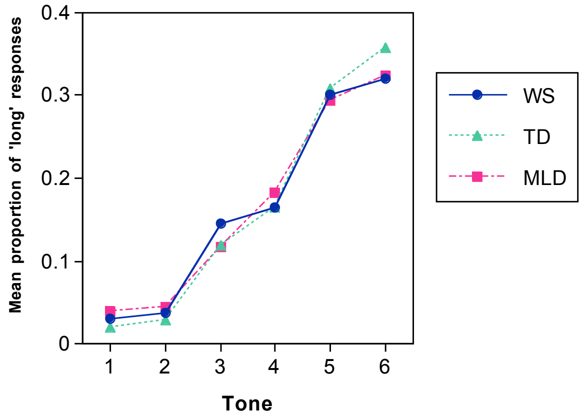

Figure 5.3 Distribution of 'long' responses in the hippos task

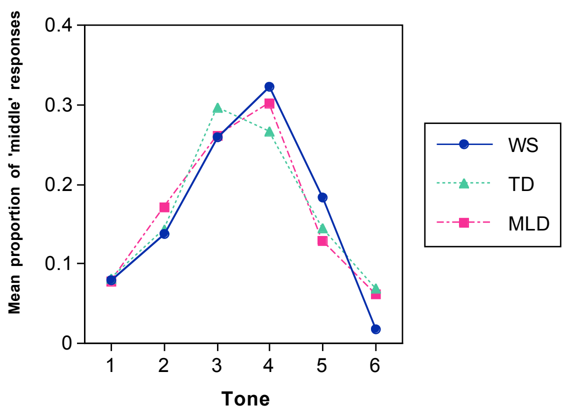

Figure 5.4 Distribution of 'middle' responses in the bears task

Frogs task

Performance on the frogs task was analysed in two ways. All three groups showed strong response biases towards the first and last frogs, making analysis of performance as a function of the presented tone problematic. Instead, Figure 5.5 shows the number of times each response was given correctly as a proportion of the total number of times that response was given. Even after this normalization procedure, all three groups showed a U-shaped curve with better performance on the extreme tones. ANOVA with group as a between-subjects factor and response as a repeated measure revealed a significant effect of response (F(3.87,220)=34.71, p<.001; Greenhouse-Geisser correction applied) which post-hoc tests showed was due to a higher proportion of correct responses for the first and last frogs than for the middle four frogs. The effect of group was not significant (F(2,44)=0.68), nor was there a significant interaction between response and group (F(10,220)=1.08). Performance was also analysed by looking at the error on each trial. Thus, for example, if the participant responded by pointing to the second frog following presentation of tone five then an error of minus three would be recorded. Figure 5.6 shows the error profiles for each group, as well as the error pattern that would be expected given a random response pattern. Clearly all groups responded well above chance levels but there is little difference between groups, although the TD group appear to make fewer extreme errors than either learning-disabled group. ANOVA was performed on error patterns with group as between-subjects factor and error as a repeated measure. Because there were very few extreme errors, only errors from minus three to plus three were included. This maintained independence of the remaining errors while a main effect of group would indicate differences in the proportion of extreme errors. In fact, the effect of group approached significance (F(2,44)=2.51, p=.093) but there was no significant interaction between group and error (F(12,264)=0.84).

Chickens task

Figure 5.7 shows the proportion of 'high' responses for each tone in the chickens task (pitch increases with tone number). ANOVA with group as a

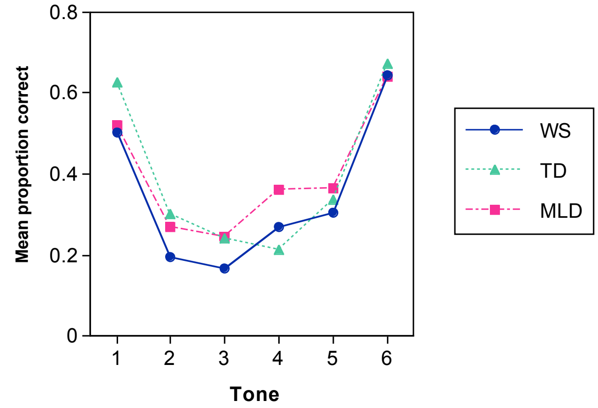

Figure 5.5 Correct responses in the frogs task

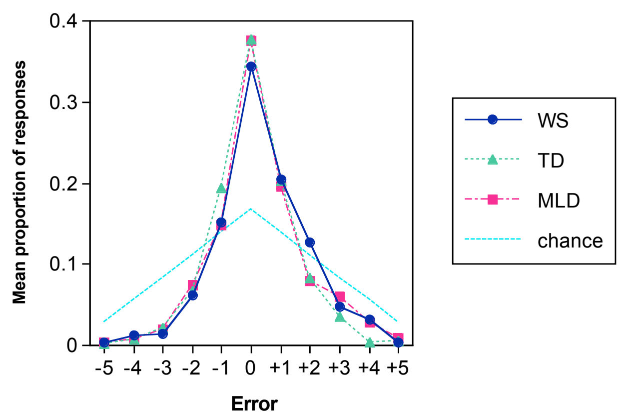

Figure 5.6 Error distributions in the frogs task

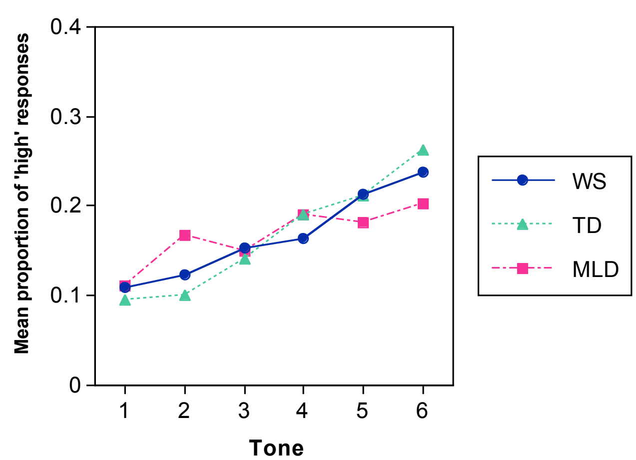

Figure 5.7 Distribution of 'high' responses in the chickens task

between-subjects factor and tone as a repeated measure showed significant main effects of tone (F(5,210)=15.23, p<.001) and group (F(2,42)=4.11, p<.05) with more 'high' responses in the MLD group than in the TD group (Tukey HSD). However, there was no significant interaction between tone and group (F(10,210)=0.79).

Correlations with temporal discrimination

If relatively good language abilities in WS are related to good temporal discrimination abilities, then there should be significant positive correlations between measures of the two abilities. A composite temporal discrimination score was calculated for each participant by taking the mean performance across the three timing tasks. Table 5.2 shows the correlations between this temporal discrimination score and vocabulary age. This approached significance when all the participants were considered as a single group (p=.128) and when the TD group were analysed separately (p=.148)[9]. Table 5.2 also shows that there was no significant correlation between temporal discrimination and the vocabulary standard score for the TD group. Because standard scores were not available for many of the children in the WS and MLD groups, vocabulary IQ was calculated for each child by dividing vocabulary age by chronological age (cf. Jarrold, Baddeley, Hewes, Leeke, & Phillips, 2002). However, there were again no significant correlations with temporal discrimination scores. Finally, Table 5.2 also shows that there were no significant correlations between temporal discrimination and Ravens matrices raw scores.

Table 5.2 Correlations with temporal discrimination score

	All	WS	TD	MLD
BPVS age	.225	.204	.336	.084
BPVS standard score	N/A	N/A	.220	N/A
Vocabulary IQ [a]	.116	.042	.199	.162
Ravens matrices †	.116	.013	.090	.208

* Correlation significant p<.05 (two-tailed)
** Correlation significant p<.01 (two-tailed)
[a] BPVS age divided by chronological age
† Based on N=13 for the WS group and N=12 for the MLD group

5.2.4: Discussion

Experiment 2 investigated timing abilities in WS using three novel tasks. These tasks had several advantages over the bisection and generalisation tasks used in Experiment 1. In particular there were objectively correct responses and there were equal numbers of correct responses of each type. Thus participants could be given individual scores which reflected discrimination abilities rather than response biases or subjective cut-off points. In all three tasks, children with WS performed at an equivalent level to MLD and TD controls matched on receptive vocabulary. More detailed analysis of within-task response patterns also found little evidence of group differences. The results of Experiment 2 therefore failed to support the prediction that individuals with WS have relatively preserved temporal discrimination abilities.

Experiment 2 also included a control pitch discrimination task. If there had been any group differences in performance on the temporal discrimination tasks, it would have been important to show that these differences were specific to the temporal domain and did not reflect other task demands. In fact the three groups also performed similarly on this control task, although most participants found the task much more difficult than the equivalent temporal discrimination task, with a large number performing at or below chance levels.

It may be that the discrimination of the pitches was more difficult than discrimination of the different durations. However, one would still expect virtually perfect performance on the extreme stimuli as they were highly distinctive from one another, but performance on these tones was at a much lower level than in the equivalent temporal task. This suggests that the pitch discrimination task was more difficult to understand. Perhaps for the participants in this study, mapping short sounds to small animals and long sounds to large animals is more intuitive than mapping high sounds onto small animals and low sounds onto large animals. If participants were frequently forgetting the response mappings then this could explain their poor performance on the easiest stimuli. Whichever explanation is correct, the near-floor levels of performance on the chickens task preclude any claims being made about pitch discrimination abilities in WS relative to language age or to temporal discrimination abilities.

5.3: Evaluation of the temporal processing hypothesis

In chapter 3 the hypothesis was proposed that individuals with WS have relatively preserved temporal processing abilities, and that this could explain their relatively good language abilities. The results of Experiment 1 were inconclusive, but Experiment 2 demonstrated that temporal discrimination abilities are at the same level as those in controls matched on vocabulary. Given that receptive vocabulary is regarded as a peak ability in WS, this implies that temporal processing abilities are also better than predicted by overall cognitive abilities. Performance on the Ravens matrices in the WS group was significantly below the level of that in TD controls matched on vocabulary, so comparison with controls matched on Ravens matrices would probably have produced a significant advantage for the WS group. However, in the absence of such data, it is not possible to state definitively whether or not temporal discrimination abilities are a relative strength.

The fact that temporal discrimination abilities appear to be a relative strength in WS would seem to be consistent with the idea that relatively preserved temporal processing abilities contribute to linguistic strengths in WS. However in section 4.1 it was argued that one might expect temporal discrimination abilities to be above the level predicted by vocabulary. The logic of this argument was that temporal processing is only likely to be one of a number of factors which determine language abilities, and if temporal processing abilities are preserved relative to these other factors then these other factors will pull language abilities down. In particular, vocabulary knowledge would be expected to be at a lower level than temporal discrimination abilities because it is dragged down by poor knowledge of the meaning of the words. If this argument holds then the fact that vocabulary and temporal processing abilities are at the same level in WS does not really provide strong evidence to support the idea that temporal processing abilities contribute to linguistic strengths in WS.

Nevertheless, receptive vocabulary in WS appears to be superior to almost every other measure of language attainment (see Section 2.4), so this would appear to be a particularly conservative test. If, for example, controls had been matched on performance on the TROG, there might well have been significant group differences. Moreover, it has been argued that the BPVS and other receptive vocabulary tests provide a weak test of semantic knowledge (Temple et al., 2002), so even if semantic processing in WS is inferior to temporal processing, it may have relatively little effect on BPVS scores.

A second way to support the claim that temporal processing is related to linguistic strengths in WS would be to show that there is a positive correlation between performance on the timing tasks and vocabulary scores. Preliminary analysis in Experiment 2 suggested that, at least in TD children, there is such a relationship, although the correlations narrowly failed to achieve significance. Given the small sample size in this analysis, this finding is perhaps not surprising.

However, even assuming that this correlation is reliable, the precise nature of the relationship is unclear. It would seem more likely that low-level perceptual abilities would contribute to language abilities than vice-versa, but this assumes that performance on the timing tasks is a pure measure of temporal discrimination abilities when in fact it is likely to depend on other factors (such as attention and the ability to remember the instructions) that may mediate the relationship between task performance and vocabulary scores. Ideally, future correlational studies should include a closely matched control task, performance on which could be partialled out. If the relationship between vocabulary and temporal discrimination remained then this would suggest that temporal discrimination abilities do contribute to language skills. Unfortunately, many of the participants in Experiment 2 performed at close-to-chance levels on the pitch discrimination task, indicating that they had failed to understand or remember the instructions. It was not therefore possible to use performance on this task as a partial correlate in the analysis for this experiment.

Evidence against the idea that temporal discrimination mechanisms contribute to vocabulary acquisition comes from considering group differences in the rate of vocabulary acquisition. Jarrold et al. (2002) have recently argued that the proposed causal influence of STM abilities on vocabulary acquisition implies that performance on STM tasks should be associated with the rate of vocabulary acquisition (i.e. vocabulary IQ) rather than its absolute level. Although these authors only considered the relationship between STM and language, the same argument can also be applied to the hypothesis that temporal processing abilities determine vocabulary acquisition. In Experiment 2, the three groups were matched in terms of vocabulary level, but the TD group were much younger than the other two groups and therefore had a much faster rate of vocabulary acquisition. Consequently, if temporal processing has a causal influence on the rate of vocabulary acquisition, the TD group should have superior temporal processing abilities to the MLD group. The absence of any group differences in temporal discrimination therefore suggests that there is no causal relationship between temporal discrimination and vocabulary acquisition.[10]

The final problem for the temporal processing hypothesis is that there is no formal theoretical link between temporal discrimination and vocabulary acquisition. According to Boucher's (1999) temporal processing account of SLI, oscillatory timing mechanisms are involved in the parsing of speech. However, formal oscillator-based models of speech perception have not been developed, so the relationship between oscillatory timing mechanisms and speech perception is only intuitive at present.

An alternative approach is to look at the relationship between timing mechanisms and STM. In chapter 6, it is argued that serial order memory plays an important role in the acquisition of vocabulary and, according to OSCAR (Brown et al., 2000; see section 3.3.2), serial order memory depends on oscillatory timing mechanisms. This suggests a possible link between vocabulary acquisition and temporal discrimination abilities. Part III of this thesis investigates the link between STM and language abilities in WS. Of particular relevance to this discussion is Experiment 3 (conducted at the same time as Experiment 2 with the same participants), which investigated serial order memory in WS. It was therefore possible to compare serial order memory, temporal discrimination abilities and vocabulary knowledge and explore the associations between these abilities.

---

[9] P-values for two-tailed tests are reported for consistency. Given that a positive correlation was predicted, a one-tailed test may have been more appropriate. Nevertheless, the correlations were still non-significant (p-values are halved).

[10] A further implication of this argument is that the WS group would be unlikely to outperform the TD group. Nevertheless, if linguistic strengths in WS are a consequence of preserved temporal processing mechanisms, children with WS should still outperform MLD controls matched on vocabulary level and chronological age (and therefore on vocabulary IQ).

Part III: Short-term memory

Chapter 6: Language and short-term memory

6.1: Overview of Part III

The studies reported in Part III of this thesis investigated the relationship between language and verbal STM abilities in WS. As noted in section 2.8.4, verbal STM is considered to be a relative strength in WS and, given the evidence from other populations that STM has a causal influence on the acquisition of language (cf. Baddeley et al., 1998), it has been suggested that linguistic strengths in WS are a consequence of relatively good STM abilities (e.g. Bishop, 1999; Grant et al., 1997; Mervis et al., 1999). A second issue concerns evidence that the influence of lexical-semantic knowledge on STM is reduced in WS (Karmiloff-Smith et al., 1997; Majerus et al., 2001; Vicari, Carlesimo et al., 1996) and the potential link between these findings and the deficits in semantic processing in WS. These two issues are addressed in Experiments 3, 4, 5, 6, and 7. However, before these experiments are reported, this chapter looks at the different theoretical accounts of verbal STM, its relationship with language abilities, and the possible implications of previous studies of STM in WS.

Section 6.2 begins by asking the question 'What is STM?', and contrasts two different approaches, the first which sees STM as separate from (but connected to) LTM representations, and the second which sees STM as simply the activation of representations in LTM.

Section 6.3 looks at the relationship between verbal STM and vocabulary acquisition, and attempts to reconcile the argument that vocabulary acquisition depends on phonological STM (i.e. the phonological loop), with the idea that the relationship between vocabulary and STM is mediated by the quality of underlying phonological representations. The suggestion is that STM and vocabulary acquisition both depend upon serial order mechanisms which allow the representation of arbitrary sequences, whether these be sequences of digits in a digit span task or sequences of phonemes in a novel word. Applying this argument to WS, this leads to the hypothesis that people with WS have relatively preserved serial order mechanisms, leading to their relative strengths in digit span and vocabulary knowledge.

Section 6.4 looks at the effects of lexical-semantic knowledge on STM performance. Evidence suggests that long-term knowledge of the phonological forms of words can be used to 'fill-in' partial STM traces in a process termed 'redintegration'. There is also evidence that phonological STM is influenced by semantic word-knowledge. Studies suggesting that there is a reduced influence of lexical-semantic factors on STM in WS are then reviewed. As noted above, their findings have been related to impaired semantic processing in WS. However, they all face methodological concerns and it is not clear whether they reflect impaired semantic memory or impaired redintegration.

6.2: What is verbal STM?

The two major issues explored in this chapter are (a) the influence of STM on long-term learning of vocabulary, and (b) the influence of long-term vocabulary knowledge on performance on STM tasks. Essentially, there are two different accounts of verbal STM and its relationship with long-term language knowledge.

The first account sees STM as being separate from, but connected to, language processing mechanisms. The most widely applied model of verbal STM is the phonological loop, introduced by Baddeley and Hitch (1974; Baddeley, 1986). The phonological loop is assumed to be a short-term buffer, separate from language processing mechanisms, that is used for storing phonological information. Long-term learning of novel vocabulary is thought to be dependent on the phonological loop because phonological traces must be held in STM before they can enter LTM. Conversely, STM performance can be influenced by LTM representations of the phonological forms of words (Baddeley, Gathercole, & Papagno, 1998).

Along similar lines, Martin, Shelton, and Yaffee (1994) have postulated the existence of a semantic buffer, operating in parallel to the phonological loop, which is involved in comprehension and gist memory[11]. Both models therefore assume a physical distinction between STM and LTM. Moreover, Freedman and Martin (2001) have argued that, whereas the phonological loop is involved in learning the phonological forms of words, the semantic buffer is involved in learning their meaning.

Both of these models assume a physical distinction between STM and LTM. However, this view is largely contradicted by functional imaging studies in humans, by single cell studies in non-human primates, and by artificial neural network models, which all suggest that the neural substrates of memory and processing are one and the same (Nadeau, 2001). The alternative account, therefore, is that verbal STM simply reflects the activation of long-term phonological and semantic representations within the language system (Gathercole & Martin, 1996; Gupta & MacWhinney, 1997; Martin & Saffran, 1997; Nadeau, 2001). According to this view, the relationship between STM and LTM occurs because performance on STM tasks is influenced by the nature of the activated long-term representations, while long-term learning depends on the activation of these representations.

6.3: Verbal STM and vocabulary acquisition

6.3.1: Evidence for a causal relationship

Gathercole and Baddeley (1993; Baddeley et al., 1998) have argued that learning the phonological form of a word is dependent on the ability to retain the novel phonological trace in STM. Consistent with this view, numerous studies have shown that there is a strong association between children's receptive vocabulary, as measured by performance on tests such as the BPVS and their performance on verbal STM tasks such as digit span, nonword repetition and serial order recognition (Gathercole & Adams 1993; Gathercole & Baddeley, 1989, 1990; Gathercole, Hitch, Service, & Martin, 1997; Gathercole, Pickering, Hall, & Peaker, 2001; Gathercole, Service, Hitch, Adams, & Martin, 1999; Gathercole, Willis, & Baddeley, 1991; Gathercole, Willis, Emslie, & Baddeley, 1992; Michas & Henry, 1994). This association remains strong even when controlling for nonverbal abilities suggesting that it cannot be explained simply in terms of the influence of overall IQ on both vocabulary acquisition and STM (cf. Baddeley et al., 1998).

However, correlation does not prove causation. As the studies to be reviewed in section 6.3 demonstrate, performance on STM tasks is influenced by vocabulary knowledge, thus providing an alternative causal relationship. Nevertheless, support for Gathercole and Baddeley's (1993; Baddeley et al., 1998) argument comes from studies showing associations between STM abilities and performance on experimental word-learning tasks. For example, learning of word-nonword pairs (but not word-word pairs) is impaired in patients with STM deficits (Baddeley, Papagno, & Vallar, 1988; Freedman & Martin, 2001; Martin, 1993; Trojano & Grossi, 1995). Similarly, word-nonword learning in neurologically intact adults is disrupted by requiring participants to articulate irrelevant material, increasing word-length or using phonologically similar items, manipulations which also impair serial recall performance (Papagno, Valentine, & Baddeley, 1991; Papagno & Vallar, 1992). Nonword learning is also correlated with STM abilities in TD five year-olds (Gathercole et al., 1997), while learning of unfamiliar names for toys (e.g. Meeton) but not familiar names (e.g. Peter) is associated with STM abilities in both TD five year-olds (Gathercole & Baddeley, 1990) and children with learning disabilities (Jarrold et al., 2002).

6.3.2: The phonological loop model

Baddeley et al. (1998) have attempted to explain the association between verbal STM abilities and vocabulary acquisition by arguing that the phonological loop component of working memory (Baddeley, 1986; Baddeley & Hitch, 1974) has evolved as a device for the acquisition of language. The phonological loop model comprises of a phonological store which represents material in a phonological code which decays over time, and a rehearsal process which refreshes decaying representations in the phonological store and allows recoding of nonphonological inputs such as printed words or pictures. Subvocal rehearsal and recoding of nonspeech input does not appear to emerge until around the age of seven years (Cowan & Kail, 1996; Gathercole & Hitch, 1993), but STM performance is closely correlated with vocabulary knowledge from a much earlier age (e.g. Gathercole & Adams, 1993), implying that the phonological store rather than the rehearsal component is fundamental to vocabulary acquisition (Baddeley et al., 1998). Baddeley et al. proposed that words enter LTM via the phonological store which provides a precise registration of phonological sequences while they are recoded into a more durable form (cf. Brown & Hulme, 1996).

6.3.3: Phonological awareness, STM and vocabulary acquisition

The view that vocabulary acquisition depends on STM has, however, been challenged by researchers arguing that the relationship between vocabulary and STM is mediated by the quality of underlying phonological representations (e.g. Fowler, 1991). Consistent with this view, performance on phonological awareness tasks is highly correlated with STM abilities and with vocabulary knowledge (e.g. Hansen & Bowey, 1994; Wagner, Torgesen, & Rashotte, 1994). Moreover, Bowey (1996) reported a significant relationship between children's vocabulary and their phonological awareness even when chronological age, nonverbal ability, nonword repetition and digit span had been partialled out (see also Metsala & Stanovich, 1995).

There are two explanations for the association between vocabulary knowledge and phonological awareness. Metsala (1997) reported evidence from a speech-perception study showing that, as vocabulary knowledge increases, lexical representations that are initially holistic become increasingly segmentalized in order to allow discrimination between similar-sounding words. Thus it can be argued that increasing vocabulary knowledge is associated with an increased tendency to represent words as sequences of phonemes, leading to improved phonological awareness. However, there is also evidence that phonological awareness contributes to vocabulary acquisition. De Jong, Seveke and van Veen (2000) reported that even when variance due to age, nonverbal ability, vocabulary and nonword repetition had been accounted for, phonological awareness was significantly correlated with learning of unfamiliar names for toys but not with learning of familiar names. Moreover, when children were given phonological awareness training, their learning of unfamiliar names improved but there was no effect on learning of familiar names. As such, the relationship between vocabulary knowledge and phonological awareness is likely to be reciprocal (cf. Brown & Hulme, 1996).

The second part of the argument concerns the relationship between STM and phonological awareness. Several authors have suggested that nonword repetition performance depends on the ability to identify the constituent phonemes in a novel word (cf. Brady, 1997; Snowling, Goulandris, Bowlby, & Howell, 1986). However, it is unclear how phonological awareness is related to performance on tasks such as digit span or serial order recognition. Consequently it is difficult to see how phonological awareness could mediate the relationship between vocabulary knowledge and performance on these tasks. Indeed, Bowey (1996) reported that digit span in children was strongly related to vocabulary even when age, nonverbal ability, nonword repetition and phonological awareness had been partialled out.

A further difficulty for this account is that certain patients have been identified who have impaired STM and difficulty acquiring new vocabulary but nevertheless perform well on phonological awareness tests (Baddeley et al., 1988). This suggests that these patients have some other impairment that leads to poor performance on verbal STM tasks and also directly or indirectly disrupts their learning of vocabulary.

6.3.4: Serial order mechanisms and vocabulary acquisition

Most errors on tasks like digit span involve misordering of the items (e.g. Aaronson, 1968; Bjork & Healy, 1974), so individual variation in digit span can be taken as an index of memory for serial order or, more generally, of the ability to represent arbitrary sequences. Similarly, serial order recognition only requires the participant to compare the serial order of two lists of the same items so is also a relatively pure measure of serial order memory[12]. The association between vocabulary and digit span and serial recognition therefore suggests that the ability to represent serial order contributes to vocabulary acquisition. Consistent with this argument, McCormack (unpublished analysis of data from a study by McCormack et al., 2000) analysed errors in serial recall in TD children and found that receptive vocabulary was more strongly related to the number of order errors than to the number of item errors.

Intuitively it makes sense that serial order mechanisms should contribute to vocabulary acquisition. Representing a novel word requires not only the representation of the constituent phonemes but also the serial order of those phonemes. Indeed Hartley and Houghton (1996) have proposed a connectionist model of nonword repetition that includes an explicit mechanism for representing serial order. From the phonological loop perspective, the ability to represent such sequences in STM would be necessary for long-term learning. Alternatively, if STM involves activation of LTM then STM for new words will primarily involve activation of phonological representations, but representing the order of the phonemes requires a mechanism to control the temporal patterns of activation. Long-term learning could occur as associations are made between the activated phonological representations, but the precise pattern of these associations will depend on the serial activation of these representations.

Figure 6.1 shows how the pattern of associations between performance on phonological awareness, STM and vocabulary tests could possibly be explained. It is assumed that there is an interactive relationship between vocabulary acquisition and the development of segmentalized lexical representations, but that vocabulary acquisition also depends on memory for serial order. Thus receptive vocabulary is independently predicted by phonological awareness and by digit span because vocabulary acquisition depends on both the quality of phonological (i.e. segmentalized) representations and on serial order mechanisms. Nonword repetition performance is determined by all three underlying factors as it requires the identification and representation of the constituent phonemes of the nonword and their sequential order and is also influenced by long-term knowledge of word-structure (as evidenced by the wordlikeness effect, see section 6.4.1). As such, it is a very good predictor of vocabulary but does not predict unique variance in vocabulary.

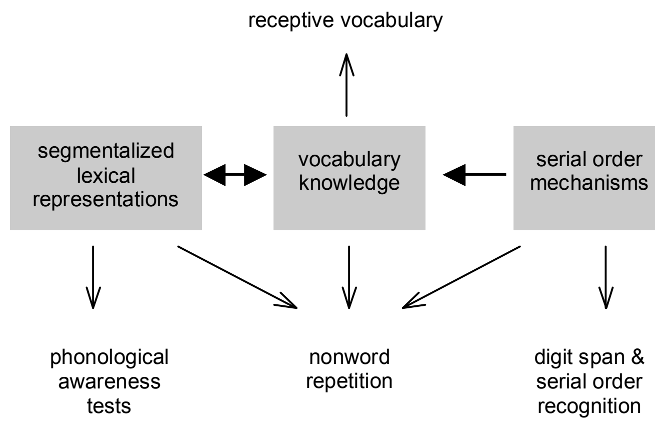

Figure 6.1 The hypothesized relationship between memory for serial order, vocabulary acquisition and the development of segmentalized lexical representations. Grey boxes represent underlying cognitive processes and representations. Thick arrows represent causal relationships, whereas thin arrows show that an underlying cognitive variable contributes to individual variation in performance on a particular test.

Clearly there will be other factors that influence vocabulary acquisition. However, the aim of Figure 6.1 is not to provide a complete account of vocabulary acquisition but simply to elucidate the associations found between performance on these tasks. Nevertheless, Figure 6.1 is able to explain correlational patterns that neither the phonological awareness nor the STM accounts can explain. Moreover, it suggests the possibility that patients who are impaired on STM tasks and are unable to learn new words but still perform well on phonological awareness tests (cf. Baddeley et al., 1988) have specifically impaired serial order mechanisms.

6.3.5: Implications for Williams syndrome

Given the above discussion, there would appear to be two possible hypotheses concerning the relative strengths in STM and vocabulary in WS. The first is that phonological representations are relatively well-developed. The second is that serial order mechanisms are relatively preserved in WS.

Evidence against the first hypothesis comes from a recent study by Laing et al. (2001) who found that people with WS performed worse than TD controls on a range of phonological awareness tests, despite the fact that the groups were matched on reading abilities and the WS group had slightly superior vocabulary scores. As the authors admit, these findings must be treated with caution because the poor performance of the WS group may have been due to difficulties with understanding the phonological awareness tasks. Nevertheless, there is a further problem with this hypothesis in that it cannot readily explain why digit span appears to be a strength in WS. Indeed Majerus et al. (2001) reported that children with WS performed within two standard deviations of the mean of that achieved by age-matched controls on serial recall tasks but were well below this level of performance on a range of phonological awareness tests

The alternative hypothesis that serial order mechanisms are relatively preserved in WS would appear more promising. It could account for the relatively good receptive vocabulary, digit span and nonword repetition in WS.

It would also fit well with the temporal processing hypothesis proposed in chapter 3.

Experimental support for this hypothesis comes from a study of 97 children and adults with WS by Mervis et al. (1999) in which standard scores on the DAS Digit Recall subtest were significantly greater than those for the PPVT-R vocabulary test. In SLI, standard scores for digit span are typically below those for receptive vocabulary, suggesting that poor STM in SLI is not just a function of poor language, but is in fact the cause of language impairments (cf. Baddeley et al., 1998). By the same logic, the fact that Mervis et al. found the opposite pattern in WS can be seen as strengthening the argument that relatively preserved STM and serial order memory in particular has a causal role in the relatively good language abilities found in WS.

Nevertheless, as mentioned in section 2.2.1, studies that simply compare standard or age-equivalent scores across tests must be treated with caution. Because Mervis et al. (1999) did not compare standard scores to those of a control group it is not clear whether their results reflect relatively good STM in WS or just the fact that the two tests were normalized on different populations. In a much smaller study, Jarrold et al. (1999) compared digit span in children with WS and TD and MLD controls matched on pattern construction abilities. Once vocabulary had been controlled for by covariation, digit span was the same in WS, MLD and TD children, indicating that digit span was not at a higher developmental level than vocabulary. Moreover, Laing et al. (2001) reported that the performance of individuals with WS on the Digit Span subtest of the DAS was significantly below that of TD controls matched on the BPVS. [13]

The evidence for or against the serial order hypothesis is therefore inconclusive. Experiment 3, reported in the next chapter, compared serial order memory in children with WS with that of TD and MLD children matched on receptive vocabulary. Children were tested on digit span and a probed serial recall task that is assumed to be a sensitive measure of serial order memory. Strong evidence to support the serial order hypothesis would be provided if individuals with WS outperformed controls on these tasks.

The studies reported in the following chapters also allowed further investigation of claims that STM performance in WS is less influenced by lexical-semantic knowledge than in normal development. The theoretical implications of these claims are discussed below.

6.4: The effect of vocabulary knowledge on verbal STM

6.4.1: Lexicality, word-frequency, and wordlikeness effects on STM

The lexicality effect refers to the finding that serial recall of words is superior to that for nonwords (e.g. Gathercole et al., 2001; Hulme, Maughan, & Brown, 1991; Turner, Henry, & Smith, 2000). This effect demonstrates that STM performance can be influenced by long-term vocabulary knowledge, as confirmed by studies by Hulme and colleagues (Hulme et al., 1991; Hulme, Roodenrys, Brown, & Mercer, 1995) who showed that teaching participants the meaning of unfamiliar foreign words or even simply giving them exposure to those words resulted in significant increases in serial recall performance for those words. Before testing, these words were effectively nonwords and the only difference between initial and final testing was the participants' experience of those words. Turner et al. (2000) have recently investigated developmental changes in the lexicality effect and found that the effect size increases with chronological age. Again this presumably reflects increasing vocabulary knowledge and familiarity with the stimuli.

A related finding is the word-frequency effect whereby serial recall of high-frequency words is better than that for low-frequency words (e.g. Hulme et al., 1997; Poirier & St Aubin, 1995; Watkins & Watkins, 1977). The assumption is that there is greater support from long-term knowledge for high- compared with low-frequency words. Essentially, the lexicality effect can be seen as an extreme version of the word-frequency effect with nonwords representing extremely low-frequency words.

Long-term vocabulary knowledge also influences STM for nonwords. The wordlikeness effect refers to the finding that repetition of nonwords rated as highly similar to real words is superior to repetition of nonwords rated low in wordlikeness (e.g. Gathercole, 1995; Gathercole, Willis, Emslie, & Baddeley, 1991; Munson, 2001; van Bon & van der Pijl, 1997). Gathercole (1995) reported that the size of the wordlikeness effect was greater at age five than at age four, suggesting that the influence of vocabulary knowledge on nonword repetition increases as the extent of that knowledge increases.

Several studies have investigated these effects in WS and these studies are described in detail below. However, in order to understand their relevance to language development in WS, it is first necessary to look at the theoretical accounts of these effects.

6.4.2: Redintegration

According to the phonological loop model (Baddeley et al., 1998), lexicality, word-frequency, and wordlikeness effects occur because representations in STM that are incomplete due to decay or interference can be filled-in using phonological LTM in a process which has been termed 'redintegration' (cf. Brown & Hulme, 1996; Hulme et al., 1991). Thus for example, the lexicality effect arises because the partial trace of a word (e.g. hippo_otamus) can be completed using long-term vocabulary knowledge, but the partial trace of a nonword (e.g. tarro_adarus) cannot.

The word-frequency effect can be accounted for in similar terms. Brown, Metsala and Hulme (2001) found that the probability that a word was correctly recalled in a serial recall task depended on whether or not a word had a higher-frequency neighbour in phonological space. This suggests that there is competition between lexical items, with the winner being the highest frequency word compatible with the phonological trace. High-frequency words are less likely than low-frequency words to have even higher-frequency neighbours so are more likely to be correctly redintegrated.

The redintegration account can also potentially explain the wordlikeness effect. Bailey and Hahn (2001) reported that wordlikeness ratings were predicted independently by both phonotactic frequency (i.e. the frequency of the nonwords' phoneme combinations in the native language) and by the density of the phonological neighbourhood which the nonwords occupied (i.e. the summed similarity of the nonword to existing words). Gathercole, Frankish, Pickering, and Peaker (1999) found that serial recall of nonwords was sensitive to phonotactics and therefore argued that knowledge of phonotactic combinations could be used to reconstruct degraded phonological traces of nonwords. However, in a more recent study, Roodenrys and Hinton (2002) reported that this phonotactic frequency effect disappeared when the density of the phonological neighbourhood was controlled for. In contrast, nonwords from a high-density neighbourhood were better recalled than those from a low-density neighbourhood, even when phonotactic frequency was controlled for. They therefore argued that redintegration of nonwords is supported by the activation of similar sounding words.

Hulme et al. (1991, 1997) have argued that redintegration reflects the operation of processes involved in speech perception and production. Consistent with this view, response speeds in repetition tasks are faster for words than for nonwords and for nonwords of high- compared with low-phonotactic frequency (Vitevitch & Luce, 1999). Moreover, Brown et al. (2001) found that the ease with which a word could be identified in a speech perception task predicted its likelihood of correct recall in a serial recall task. According to the phonological loop account (Baddeley et al., 1998), STM is separate from speech processing mechanisms but is influenced by long-term knowledge represented in these mechanisms. However, by the alternative view that STM involves activation of normal speech processing mechanisms, effects of lexicality, word-frequency and wordlikeness can be considered as consequences of top-down influences in speech processing.

Gathercole and Martin (1996) suggested that redintegration processes can be understood in terms of connectionist models of speech perception. For example, according to McClelland and Elman's (1986) TRACE model, the phonological structure of incoming speech activates consistent representations in the mental lexicon which can in turn activate consistent phonological representations. Thus representations at the phonological level receive greater top-down activation if they are consistent with representations at higher levels, so words receive more activation than nonwords, while high-wordlike nonwords receive more activation than low-wordlike nonwords. In fact, connectionist models of speech production provide similar accounts. In Dell's (1986) spreading activation model, for example, nodes corresponding to the intended morpheme activate connected nodes at the syllable, rime and phoneme levels. Activity from these nodes then spreads up to related morphemes that in turn provide top-down activation of the syllables, rhymes and phonemes that are its component parts. Redintegration effects could therefore reflect speech output processes as much as input processes.

6.4.3: Semantic effects on STM

The redintegration account explains the lexicality, word-frequency and wordlikeness effects in terms of long-term phonological memory. However, semantic factors also play a role in performance on STM tasks. For example, concrete or imageable words are better recalled than abstract words (e.g. Bourassa, & Besner, 1994; Nation, Adams, Bowyer-Crane, & Snowling, 1999; Walker & Hulme, 1999), and content words are better recalled than function words (e.g. Caza & Belleville, 1999), even when they are matched for frequency.

Semantic memory could also contribute to the lexicality effect as recall of words benefits from semantic memory traces whereas recall of nonwords depends exclusively on retrieval of the phonological trace. Similarly, the word-frequency effect could plausibly result from increased activation or accessibility of semantic traces for high- compared with low-frequency words. Although simply exposing participants to the phonological forms of unfamiliar words without teaching their meaning has been shown to improve subsequent serial recall (Hulme et al., 1995), this does not rule out the possibility that semantic memory plays a supplementary role in lexicality and word-frequency effects.

Further support for this view comes from studies of patients with acquired neurological disorders. For example, Martin et al. (1994) tested two patients, A.B., who had difficulty with sentence comprehension, and E.A., who had difficulty with rote repetition of sentences but not with producing their gist. When required to listen to a word list and then judge whether a probe word was in the same semantic category as any list member, A.B. performed worse than E.A., suggesting that he had a deficit in semantic STM. In contrast, when asked to judge whether the probe word rhymed with any list members, A.B. performed better than E.A., suggesting that she had a deficit in phonological STM. Consistent with the idea that the lexicality effect involves semantic memory, A.B. showed little evidence of a lexicality effect despite a significant phonological similarity effect, whereas E.A. showed the opposite pattern.

6.4.4: Lexical-semantic effects on STM in WS

6.4.4.1: The word-frequency effect

Vicari, Carlesimo et al. (1996) investigated the effect of word-frequency on serial recall span in WS, testing 12 children with WS (mean age 9;11) and 12 TD controls whose mean chronological age (5;2 years) was similar to the mean nonverbal mental age of the WS group (5;6 years, measured by the Leiter Intelligence Performance Scale; Leiter, 1979). There were no significant differences between groups in terms of overall performance and both groups showed equivalent effects of word-length, having longer spans for two- compared with four-syllable words. Both groups also showed significant word-frequency effects, having longer spans for high- compared with low-frequency words, but there was a significant group by condition interaction whereby the WS group showed a reduced word-frequency effect compared with controls.

Vicari, Carlesimo et al. (1996) argued that this finding reflected a reduced influence of lexical-semantic knowledge on STM in WS. However, it is not clear whether this reflects impaired redintegration or impaired semantic memory. Moreover, the WS group were much older than controls and, since their mean nonverbal mental age was higher than the mean chronological age of the controls, they almost certainly had superior vocabulary knowledge. On the face of it, this is not a problem because the lexicality effect increases with age and (one assumes) with vocabulary knowledge (Turner et al., 2000). If anything, one might therefore have expected the WS group to have shown an increased rather than a reduced effect of word-frequency. Nevertheless, the WS group may have been more familiar with the low-frequency words than controls (because they were older and had better vocabulary knowledge), while both groups may have been equally familiar with the high-frequency words. The WS group would then have performed relatively well on recall of low-frequency words leading to a reduced word-frequency effect.

6.4.4.2: The lexicality effect

Such a problem in interpretation is not found with the lexicality effect because the nonwords are necessarily unfamiliar for all participants, so the size of the effect is entirely determined by the support offered by knowledge of the real words. In fact, Majerus et al. (2001) did find evidence for a reduced lexicality effect in WS, suggesting that either redintegration or semantic memory is impaired in WS. Unfortunately there are again concerns with their choice of control group. They tested four 10- to 12-year-old children with WS and 22 TD control children from the same age range, but the WS group had much lower vocabulary scores. As such, their reduced lexicality effect could simply reflect reduced familiarity with the words. Moreover, the overall performance of the WS group was much worse than that of controls, reducing the possible effect sizes that may have been found in the WS group.

6.4.4.3: Lexicalization errors in nonword repetition in WS

Karmiloff-Smith et al. (1997) investigated French grammatical gender acquisition in 14 individuals with WS (age range 9;0 to 22;6 years, mean 15;9) and 18 TD control children (age range 4;6 to 5;11, mean 5;1). Part of the task required participants to repeat aloud a nonword (before marking the article and adjective to complete a sentence), and it was noted that participants with WS were far superior to controls at this repetition element of the task. Moreover, whereas TD children made numerous lexicalisation errors (misrepeating nonwords as similar sounding real words), participants with WS almost always repeated the words correctly.

Karmiloff-Smith et al. (1997) suggested that when controls heard a nonword it activated representations of similar-sounding words and meanings which were sometimes erroneously produced, but this did not happen in WS. Although the authors linked the reduction in lexicalization errors to impaired semantics in WS, nonwords have no meaning so cannot activate semantic representations directly. As such, this finding favours the impaired redintegration account over the impaired semantic memory account.

Again, however, the choice of control group could provide an alternative explanation. Because the study aimed to show that difficulties in grammar acquisition in WS could not be explained in terms of overall delay in language acquisition, the control group were deliberately chosen to have lower vocabulary scores than the WS group. The authors essentially argued that reduced influence of vocabulary knowledge led to fewer lexicalization errors and therefore better nonword repetition. However, it is possible that the reverse argument is in fact true. In other words, if the WS group had superior STM (and this seems likely given that they had much higher vocabulary scores), they would have made fewer errors per se, and consequently they would have made fewer lexicalization errors.

6.4.4.4: Wordlikeness effect

Grant, Karmiloff-Smith, Berthoud, and Christophe (1996) reported that people with WS recalled nonwords that resembled words from their native language better than they recalled nonwords resembling foreign vocabulary. Thus individuals with WS showed a wordlikeness effect but it was not clear from this study whether the effect-size was normal.

In a subsequent study using the Children's Test of Nonword Repetition (CNRep; Gathercole & Baddeley, 1996; Gathercole, Willis, Baddeley, & Emslie, 1994), Grant et al. (1997) individually matched 15 participants with WS to TD five-year-olds on the basis of Ravens matrices scores and found no significant difference in the size of the wordlikeness effect (J. Grant, personal communication). This finding suggests a normal influence of redintegration processes in WS. However, the WS group in the Grant et al. study were not only older than the controls but almost certainly had higher BPVS ages than the controls. Given that the wordlikeness effect increases with age (Gathercole, 1995), it remains possible that the WS group may have shown a reduced wordlikeness effect compared to controls matched more closely on age and / or vocabulary knowledge.

More recently, Majerus et al. (2001) tested four children with WS and reported that they showed a reduced phonotactic frequency effect in serial recall of nonwords compared with a group of 20 children from the same age range. Given that phonotactic frequency may be confounded with the density of the phonological neighbourhood (cf. Roodenrys & Hinton, 2002), this could reflect a reduced influence of either lexical or sublexical redintegration processes in WS. However, the small size of the WS group, together with the fact that the WS group had much poorer vocabulary knowledge and overall performance (hence reducing the possible effect size that they could show) means that this finding should also be treated with caution.

Table 6.1 Support for the impaired redintegration and impaired semantic memory hypotheses

	Impaired redintegration	Impaired semantic memory
Reduced word-frequency effect [a]	✓	✓
Reduced lexicality effect [b]	✓	✓
Reduced lexicalization errors [c]	✓	
Normal wordlikeness effect [d]	✗	
Reduced phonotactic frequency effect [b]	✓	

a Vicari, Carlesimo et al. (1996)
b Majerus et al. (2001)
c Karmiloff-Smith et al. (1997)
d Grant et al. (1997)
✓ Finding supports the hypothesis
✗ Finding is inconsistent with the hypothesis

6.4.4.5: Theoretical interpretations

The findings reviewed above would appear to provide relatively consistent evidence that there is a reduced influence of lexical-semantic factors on STM in WS. However, the interpretation of these findings is unclear. Essentially, there are two hypotheses; the first is that redintegration processes are impaired in WS, while the second is that there is some form of semantic memory impairment. Table 6.1 summarises the findings, indicating whether or not they support each hypothesis.

Establishing the validity of these hypotheses is potentially important as it could provide insight into the development of language in WS. For example, an interesting comparison can be drawn between individuals with WS and patients with acquired impairments of semantic STM. These patients show a reduced influence of lexicality in serial recall (Martin et al., 1994), and also have difficulty learning the semantic form of new words, suggesting that semantic STM is important for learning the meaning of words (Freedman & Martin, 2001). Impaired semantic STM could help explain why, for example, children with WS are better at learning the phonological forms of words than they are at learning their meanings (cf. Paterson, 2000; Singer-Harris et al., 1997).

If, however, these findings indicate a reduced influence of redintegration processes on STM in WS, this would have quite different implications. As noted above, redintegration is assumed to reflect the operation of normal speech-processing mechanisms so impaired redintegration would indicate that individuals with WS process speech differently to TD individuals. Nevertheless, there are methodological concerns with many of the studies, so it remains possible that neither hypothesis is correct. The following set of experiments aimed first of all to replicate previous studies investigating lexical-semantic effects on STM in WS while addressing these methodological concerns. The second aim was to try and tease apart the two hypotheses. Experiment 3 investigated the lexicality effect in probed serial recall and lexicalization errors in the probed serial recall of nonwords, while Experiments 4 and 5 looked at the word-frequency effect in probed and forward serial recall respectively. Experiment 6 investigated the wordlikeness effect in nonword repetition, while Experiment 7 looked at the concreteness effect in forward serial recall. The predictions of the two hypotheses are shown in Table 6.2. Experiment 3 also allowed investigation of the serial order hypothesis proposed in section 6.3.5.

---

[11] A similar concept is the Episodic Buffer proposed by Baddeley (2000).

[12] Neath (1997) has argued that serial order recognition is not a pure test of order memory because the participant must remember the items to determine whether they are in the correct order. However, for phonologically distinct items, the participant does not need to remember the whole item, and the majority of individual variation in performance is likely to be due to order memory.

[13] Several other studies (e.g. Grant et al., 1997; Howlin et al., 1998) have looked at performance on the Digit Span subtest of the WISC or WAIS, but only reported composite scores from forward and backward span.

Chapter 7: Probed recall

7.1: Overview

Two major claims have been made about verbal STM in WS. First, it has been argued that the relative strengths in language in WS are a consequence of relatively preserved verbal STM. In chapter 6 it was suggested that memory for serial order plays an important role in the relationship between STM and vocabulary acquisition and therefore that memory for serial order may be a particular strength in WS. The second claim is that there is a reduced influence of lexical-semantic knowledge on STM performance. In particular, it has been reported that individuals with WS show reduced effects of word-frequency (Vicari, Carlesimo et al., 1996) and lexicality (Majerus et al., 2001), and a reduced tendency to make lexicalization errors in nonword repetition (Karmiloff-Smith et al., 1997).

Two experiments are reported in this chapter. Experiment 3 utilised a series of probed recall tasks to investigate item and order memory and the effect of lexicality in children with WS and controls. It was also possible to investigate the tendency of individuals with WS to produce lexicalization errors when attempting to recall nonwords. Experiment 4 used a probed recall task to look at the effect of word-frequency.

7.2: The probed recall paradigm

There are two variations of the probed recall paradigm. In the probed item recall task, participants are presented with a list of items and are required to recall the item that occurred at a given position in a list. Thus participants must recall both the item and its order (position). In the probed position recall task participants are presented with a list and are then given an item and asked for its position in the list. Although participants need to recall some item information, they only need to be able to recall enough information to distinguish the item from the others in the list. Consequently, variation in performance is primarily determined by order memory.

There are several advantages of probed recall tasks over traditional serial recall tasks. The first is that analysis of the factors contributing to performance is relatively straightforward, and the cause of any group differences is correspondingly more transparent. In serial recall, for example, performance on later words may be influenced by output interference whereby recall of earlier items impairs recall of later items (cf. Cowan et al., 1992; Hulme et al., 1997). A further complication in serial recall is the temporally extended recall process. In particular, the delay between presentation of an item and its recall depends on the rate of output, so performance is dependent on speech rate (cf. Brown et al., 2002). By comparison, in probed recall, output interference is eliminated as only one item is output per trial, and the delay between presentation and output is relatively constant, so speech rate should not influence performance.

A second advantage of probed recall is that it allows relatively simple analysis of error patterns because only one response is made. In contrast, in serial recall, an ambiguous combination of errors may occur. For example, recalling ABCDE as ADCE could represent several different combinations of transpositions and omissions.

The probed recall tasks used in Experiment 3 were adapted from those used by Turner et al. (2000) in a study of developmental changes in the size of the lexicality effect in TD children. Turner et al. tested five-, seven- and ten-year-olds on recall of words and nonwords using both item and position recall tasks. They reported that lexicality had no significant effect on position recall in any age group[14]. The second main finding was that in the item recall task, word recall was significantly better than nonword recall, and the size of this lexicality effect increased with age.

7.3: Experiment 3: Probed recall for words and nonwords

7.3.1: Introduction

In Experiment 3, the probed serial recall performance of children with WS was compared with that of TD children and a group of MLD children. All groups were matched on receptive vocabulary using the BPVS. This allowed investigation of the lexicality effect while controlling for the extent of lexical-semantic knowledge, and it also allowed direct comparison between STM and vocabulary knowledge in WS.

Memory for serial order was tested by looking at overall performance on the two probed position recall tasks. It was also possible to investigate item and order memory by looking at the pattern of errors in the probed item recall tasks. Mervis et al. (1999) reported that in WS, standard scores for digit span were superior to those for receptive vocabulary. In chapter 6 it was argued that this finding is consistent with the hypothesis that relatively preserved STM, and serial order memory in particular, contributes to linguistic strengths in WS, because good performance on the digit span task could not be explained in terms of generally good language abilities (cf. Baddeley et al., 1998 for a similar argument regarding STM deficits in SLI).

Experiment 3 offered the opportunity to test this hypothesis, the prediction being that individuals with WS would perform better than vocabulary-matched controls on the two probed position recall tasks, and would make fewer order errors in the two probed item recall tasks. Participants were also tested on forward digit span from the Digit Span subtest of the WISC. This allowed direct comparison with the study by Mervis et al. (1999). Moreover, by comparing performance on the digit span and probed recall tasks, it was possible to investigate the claim that digit span is a test of serial order memory. It was also possible to look at the relationship between the serial order memory of participants and their temporal discrimination abilities that were assessed in Experiment 2.

The influence of lexical knowledge on STM was examined by comparing performance on the probed item recall of words and nonwords. Previous studies have reported reduced effects of word-frequency (Vicari, Carlesimo et al., 1996) and lexicality (Majerus et al., 2001) in WS, but both studies were compromised by the fact that control groups were not matched on vocabulary knowledge, so group differences could simply reflect the extent of lexical knowledge. If there really is a reduced influence of lexical knowledge then the children with WS should show a reduced lexicality effect in the probed item recall task compared with controls matched on vocabulary. It was also possible to investigate the claim that individuals with WS make fewer lexicalization errors than controls (Karmiloff-Smith et al., 1997) by looking at error patterns in probed item recall for nonwords.

7.3.2: Method

7.3.2.1: Participants

The participants were the same 14 children with WS (9 boys), 20 TD children (10 boys), and 13 MLD children (8 boys) who took part in Experiment 2. Participant details are shown in Table 7.1. As noted in section 5.2.2, the TD children were significantly younger than the other groups, and the WS group performed significantly worse than the TD group on the Ravens matrices but the groups were closely matched on receptive vocabulary.

Table 7.1 Participant details for Experiment 3

	WS N=14		TD N=20		MLD N=13	
	M	SD	M	SD	M	SD
Chronological age (y; m)	13;5	2;5	7;5	0;9	12;6	1;9
BPVS–II age (y; m)	8;2	1;2	8;2	1;4	8;1	1;3
Ravens matrices age (y; m) †	7;6	1;8	8;9	1;2	8;1	2;0

† Based on N=13 for the WS group and N=12 for the MLD group

7.3.2.2: Materials

For the practice session of the probed recall task, nine picture cards were used. On one side of each card was an identical star shape. On the other side, each card had a different picture taken from the Snodgrass Picture Inventory (hat, pig, bus, duck, tree, lamp, shoe, kite and watch; Snodgrass & Vanderwart, 1980).

For the word-recall tasks, two sets (A and B) of 20 one-syllable CVC words were constructed (see Appendix 1). These words all had high familiarity ratings (greater than 500) and low concreteness and imageability ratings (less than 500; Quinlan, 1992) and sets A and B were equated for mean familiarity and concreteness ratings. For the nonword-recall tasks, two sets of nonwords (A and B) were constructed by changing the initial consonant sounds in the corresponding list of words. These words and nonwords were recorded by a native English-speaking female and stored individually as SDII 16 bit sound files.

Test lists were devised for each of the four sets of words or nonwords in the probed recall tasks. Each list consisted of two practice trials with two items (each position probed once), followed by nine test trials with three items (each position probed three times), and then twelve test trials with four items (each position probed three times). The choice of items within each list was pseudo-random, subject to the constraint that no item appeared more than once in any list or in consecutive lists. Furthermore, within each trial, the same phoneme did not appear more than once in the same position within an item. Each item appeared an approximately equal number of times, although some items necessarily appeared one more time than others did.

Probed recall tasks were presented on a MacIntosh PowerBook G3 laptop computer using the PsyScope application (Cohen et al., 1993). Items were presented via the computer's internal speakers as several of the WS children did not like wearing headphones.

7.3.2.3: Procedure

Children with WS were tested in a quiet room in their own home. The four conditions of the probed recall task, the background measures and the Forward Digit Span subtest of the WISC were all administered within the same session, although short breaks were allowed when the child requested. Both control groups were tested in a quiet room at their own school. Where possible, all testing occurred within the same session. However, this was not possible in all cases, especially when younger and less able controls were concerned.

Testing of probed recall consisted of two blocks, corresponding to probed item recall or probed position recall, with an unrelated filler task in between[15]. Each test block was always completed within one session. The order of testing of item recall and position recall was counterbalanced across groups. However, the first block always consisted of training followed by word recall and then nonword recall, while the second block always consisted of training followed by nonword recall and then word recall. Each set of words or nonwords was used once and, within each block, the nonwords used were those derived from the words used in the other block. There were therefore four possible test sequences (for example item recall training, item recall words A, item recall nonwords B, filler task, position recall training, position recall nonwords A, position recall words B) which were counterbalanced across groups.

Each block began with a training phase. The experimenter first introduced the training cards, showing the child that they had different pictures on one side but all had stars on the other. The child was told that they would play a memory game and had to collect five cards. Two cards were then chosen at random, and placed face down side-by-side on the table in front of the child. The experimenter touched the back of each card in turn from left to right, naming the card as it was touched. In the item recall condition, the experimenter pointed to one of the cards and asked the participant to name it, whereas in the position recall condition, the experimenter named one of the cards and asked the child to point to it. The cards were then turned over and if the participant was correct they were allowed to keep the card. The procedure was repeated until the participant had correctly identified or pointed to one card in each of the two positions. The game was then repeated with three cards and continued until one card had been correctly identified in each of the three positions. Upon completing the training phase, the participant was told that they could now play the same game on the computer. In the words version they played it with an English lady who said English words, whereas in the nonwords version, they played it with a Martian lady who said 'funny Martian words'. On each trial, two, three or four cards appeared in a row on the screen, a hand moved across the bottom of the screen, pointing to each card in turn (see Figure 7.1), and the participant heard each card named as it was pointed to. Items were heard at a rate of one every 1.5 seconds (inter-onset times).

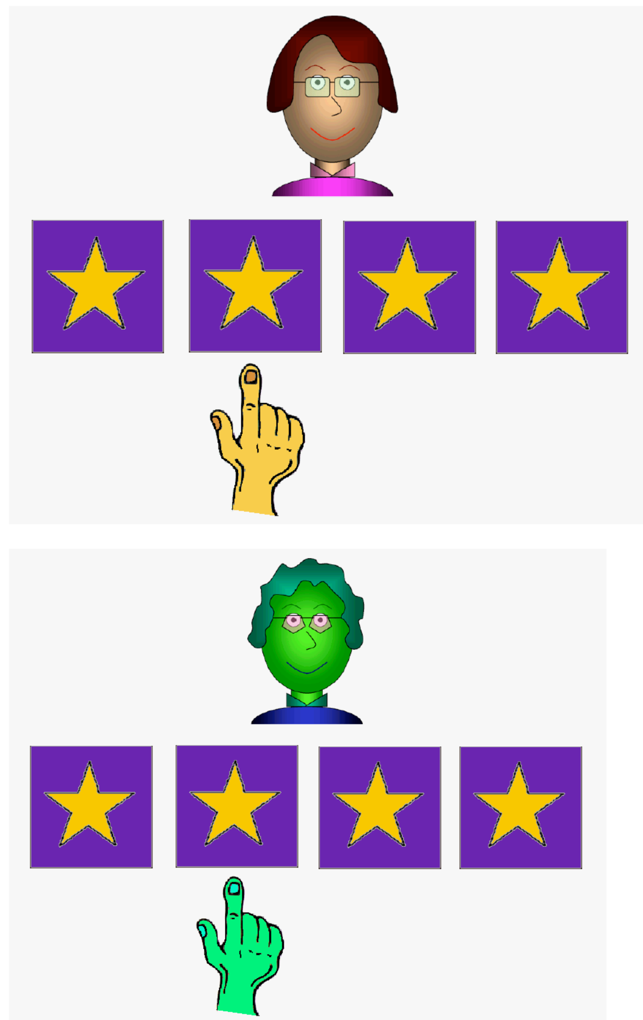

Figure 7.1 Displays for the probed recall of words (upper panel) and nonwords (lower panel). The lady and the hand in the lower panel are green.

The probe was presented 2.5 seconds after the onset of the last word. In the position recall task, the child heard one of the words again and was asked to point to the corresponding card. The experimenter entered a 1, 2, 3 or 4 on the keyboard corresponding to the position indicated by the child and the hand re-appeared pointing to the correct card. In the item recall task, the hand pointed at one of the cards and the child was asked to name the card. The experimenter then pressed the space bar and the child heard the correct answer. The experimenter pressed 1 on the keyboard to indicate a correct response or 2 to indicate an error and also noted the participant's response.

7.3.3: Results

Digit span

Figure 7.2 shows performance on the digit span task. The upper panel shows the mean span (i.e. the longest list correctly repeated) for the three groups, whereas the lower panel shows mean the score (i.e. the number of lists correctly repeated). One-way ANOVAs revealed a significant effect of group on both span ($F(2,44)=4.32$, $p<.05$) and score ($F(2,44)=4.21$, $p<.05$), with the TD

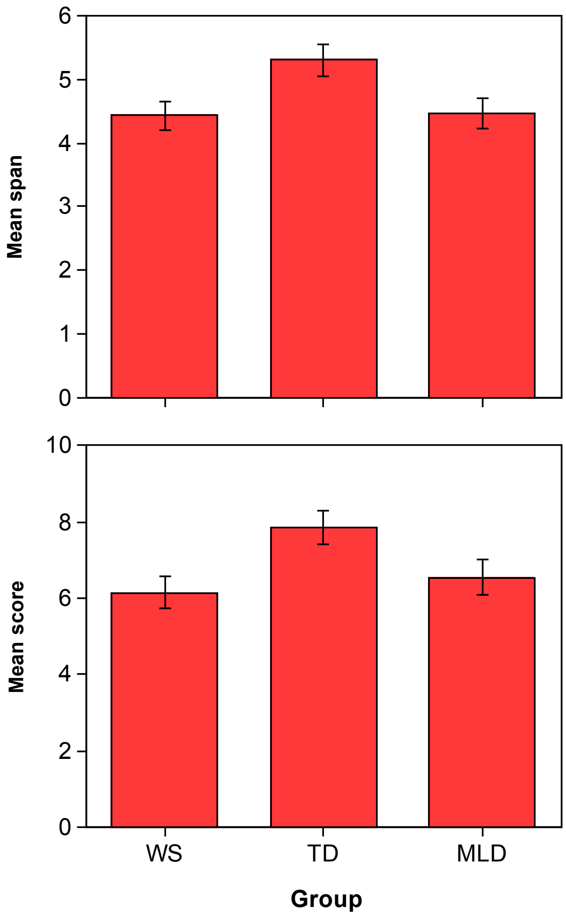

Figure 7.2 Digit span. Upper panel shows performance in terms of span. Lower panel shows performance in terms of score (i.e. the number of lists correctly recalled).

group significantly outperforming the WS group on both measures (Tukey HSD).

Table 7.2 shows correlations between digit span score and various other measures. There was a significant correlation between digit span and BPVS age, although when individual groups were considered separately, this was only significant in the TD group. Digit span was also significantly correlated with standard scores for the BPVS in the TD group. However, standard scores were not available for many of the children in the WS and MLD groups. Following Jarrold et al. (2002), vocabulary IQ was therefore calculated for each child by dividing vocabulary age by chronological age. Digit span was significantly correlated with vocabulary IQ in the overall group and, although there was a significant correlation in the TD group, there was little evidence of correlations in the two learning-disabled groups.

Table 7.2 Correlations with digit span

	All	WS	TD	MLD
BPVS age	.365*	.105	.565**	.328
BPVS standard score	N/A	N/A	.570**	N/A
Vocabulary IQ [a]	.503**	.068	.570**	.136
Temporal discrimination score [b]	.075	-.345	.083	.395
Ravens matrices raw score †	.296*	.359	.112	.158

* Correlation significant p<.05 (two-tailed)
** Correlation significant p<.01 (two-tailed)
[a] BPVS age divided by chronological age
[b] Mean score on the three temporal discrimination tasks from Experiment 2
† Based on N=13 for the WS group and N=12 for the MLD group

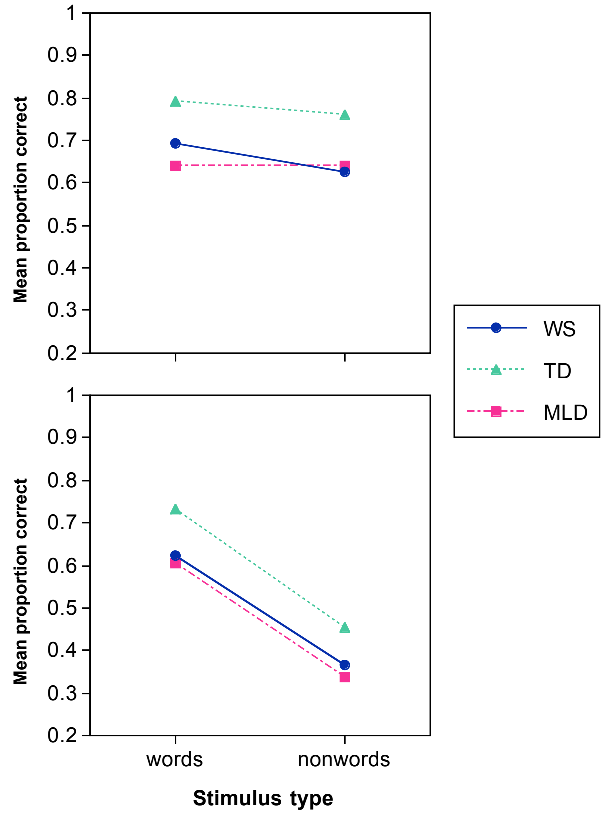

Figure 7.3 The effect of lexicality on performance on the probed recall tasks. The upper panel shows probed position recall. The lower panel shows probed item recall

There was no significant correlation between digit span and temporal discrimination performance from Experiment 2. Digit span was significantly correlated with Ravens matrices raw score. Analyses were repeated with Ravens score as a partial correlate, but the only qualitative effect on the pattern of correlations was that the correlation between BPVS age and digit span for the overall group failed to reach significance.

Overall performance on the probed recall tasks

Overall performance in terms of the proportion of trials correct is shown in Figure 7.3. As in the Turner et al. (2000) study, the effect of lexicality was only present in the probed item recall task. In both tasks, the TD group outperformed the other two groups, and the size of the lexicality effect in the probed item recall task was similar for all groups.

These observations were confirmed by a three-way mixed design ANOVA with group (WS vs TD vs MLD) as a between-subjects factor and lexicality (words vs nonwords) and recall type (item vs position) as repeated measures. Recall of words was significantly superior to recall of nonwords ($F(1,44)=150.87$, $p<.001$) and position recall was significantly better than item recall ($F(1,44)=122.25$, $p<.001$). There was also a significant effect of group ($F(2,44)=5.96$, $p<.01$), which reflected an advantage for the TD group over the WS and MLD groups (Tukey's HSD). There were no significant interactions between group and lexicality ($F(2,44)=0.17$) or between group and recall type ($F(2,44)=0.37$). There was, however, a significant interaction between lexicality and recall type ($F(1,44)=67.72$, $p<.001$), which occurred because the magnitude of the lexicality effect (proportion of words correct minus proportion of nonwords correct) was greater for item recall than for position recall ($t(46)=8.48$, $p<.001$). The three-way interaction between group, lexicality and recall type was non-significant ($F(2,45)=0.59$).

There was therefore little evidence for a reduced effect of lexicality as predicted by previous studies of WS. However, it was expected that an effect of lexicality would only be observed in the item recall task, and it is possible that group effects in the item recall task may have been masked by conducting a three-way ANOVA which included data from the position recall task. To rule out this possibility, overall performance was analysed separately for the two types of recall.

For position recall, there was a significant effect of group (F(2,44)=6.37, p<.01), with the TD group outperforming the WS and MLD groups (Tukey's HSD). The effect of lexicality was nonsignificant (F(1,44)=2.60), as was the interaction between group and lexicality (F(2,44)=0.83).

For item recall, there was again a significant effect of group (F(2,44)=4.11, p<.05) with the TD group outperforming the MLD group, but no significant difference between the WS group and either control group (Tukey's HSD). Words were recalled significantly better than were nonwords (F(1,44)=200.62, p<.001), but there was no significant interaction between group and lexicality (F(2,44)=0.08).

Given that performance on the item recall task is assumed to depend on both item and order memory, but lexicality is only thought to affect item memory, the analysis of performance on the item recall task was repeated with overall performance on the Position Recall task as a covariate to control for order memory. Position recall accounted for significant variation in item recall (F(1,43)=46.26, p<.001). Once this variation had been removed, there was still a significant effect of lexicality (F(1,43)=12.89, p<.001) but no effect of group (F(2,43)=0.08). There was no interaction between lexicality and position recall (F(1,43)=1.43) and, as before, there was no interaction between group and lexicality (F(2,43)=0.40). Thus group differences in the item recall task appear to be attributable to the superior order memory of the TD group.

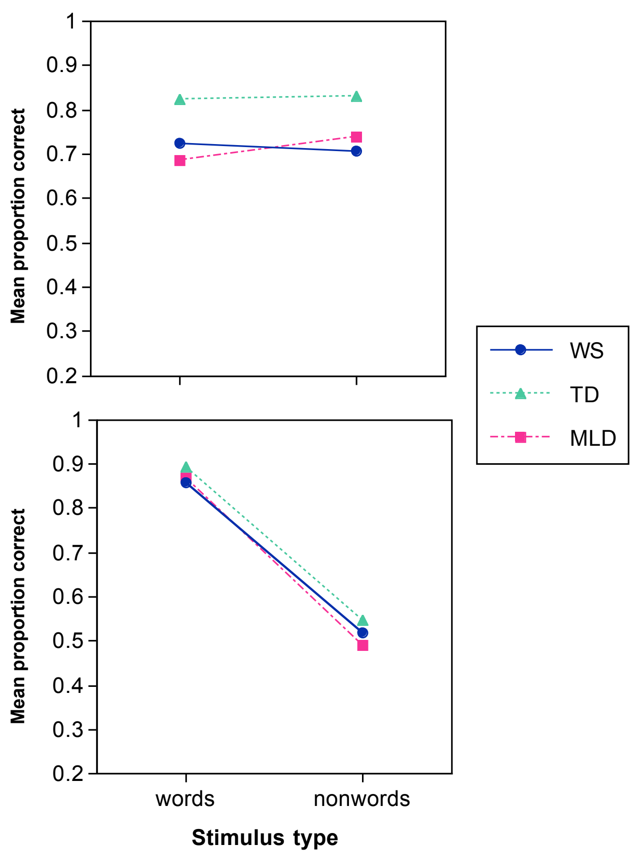

Figure 7.4 Estimates of order memory (upper panel) and item memory (lower panel) obtained from error analysis of the probed item recall task

Error analysis - estimated item and order memory

Error analysis of responses in the item recall task revealed a similar pattern. Responses were classified as correct, movement error (another item from the same list) or other error (including extra-list intrusions and omissions). Given that a correct response requires that both the item and the order are correctly encoded, while a movement error requires that the item is correctly encoded but the order is wrongly encoded, and assuming that item and order memory are independent (e.g. Healy, 1974) then

P (correct) = P (item correct) x P (order correct)

P (movement error) = P (item correct) x (1 – P (order correct))

Thus for each participant it is possible to estimate the probability of correctly recalling the item and order.

P (item correct) = P (correct) + P (movement error)

P (order correct) = P (correct) / [P (correct) + P (movement error)]

Mean estimated probabilities of correct item and order recall are shown in Figure 7.4. These results were subjected to two-way mixed design ANOVA with Group as between-subjects factor and Lexicality as a within-subjects factor. The probability of correct item recall was significantly higher for words than for nonwords (F(1,44)=165.41, p<.001) but there was no significant effect of group (F(2,44)=0.62) and no interaction between lexicality and group (F(2,44)=0.18). There was no significant effect of lexicality on the probability of correct order recall (F(1,44)=0.54) but there was, however, a significant effect of group (F(2,44)=3.80, p<.05) although no pairwise differences were significant (Tukey HSD). The interaction between lexicality and group was non-significant (F(2,44)=0.98).

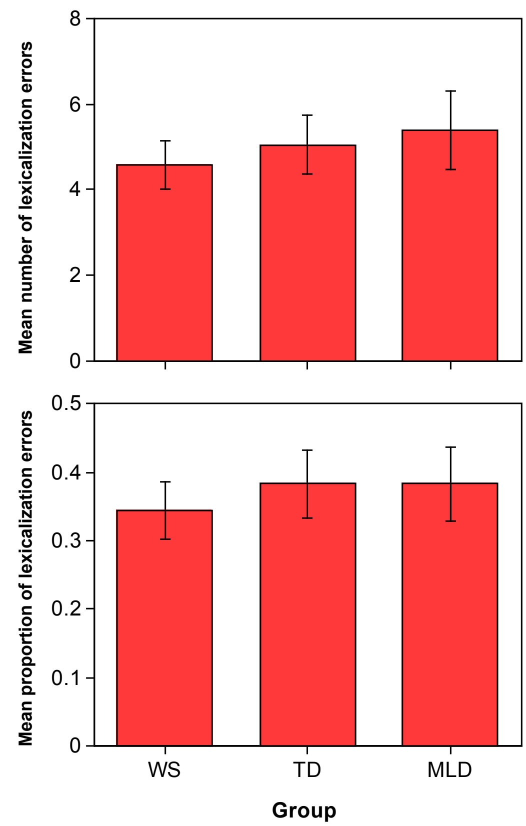

Figure 7.5 Lexicalization errors in the probed item recall of nonwords. Upper panel shows the number of lexicalization errors. Lower panel shows lexicalization errors as a proportion of total errors. Error bars show standard errors.

Error analysis - lexicalization errors

Error analysis of nonword recall in the probed item task also allowed investigation of the claim that individuals with WS make relatively few lexicalization errors. Erroneous responses were classified as being either lexical or non-lexical. Figure 7.5 (upper panel) shows that individuals with WS made slightly fewer lexicalization errors than TD or MLD controls. However, one-way ANOVA failed to find any significant group difference (F(2,44)=0.27). A similar pattern was found when looking at lexicalization errors as a proportion of total errors (lower panel of Figure 7.5), but the difference was again non-significant (F(2,44)=0.85). The above analyses were repeated using the less conservative method of comparing the WS group with the other two groups separately, but in all cases there were no significant group effects.

Serial position effects for probed position recall

The upper panel of Figure 7.6 shows the serial position curves for the four-item position recall task collapsed across levels of lexicality.[16] For all three groups performance was much better for the last two items than for the first two, but there was little difference between the third and fourth positions.

However, closer analysis of the distribution of responses showed that children in all three groups had a tendency to point to the third position more frequently than expected, and the first position relatively infrequently. Because of this response bias, performance on the third item may appear better than it really is, and performance on the first item may appear particularly poor.

The serial position curves were therefore normalized for each participant by dividing the number of correct responses at each position by the number of times they had given that response. Participants were only included in this analysis if they had made at least three responses of each type. This left seven children with WS, fourteen TD children and seven MLD children. The lower panel of Figure 7.6 shows the normalized serial position curves for these remaining participants. All three groups showed strong recency effects and moderate primacy effects. ANOVA was performed with group as a between-subjects factor and position as a repeated measure. There was a significant effect of position (F(3,75)=16.15, p<.001) reflecting superior performance on the final position compared with the other three positions (Bonferroni correction applied). However, there was no significant effect of group (F(2,25)=1.97) and no significant interaction between position and group (F(6,75)=0.87).

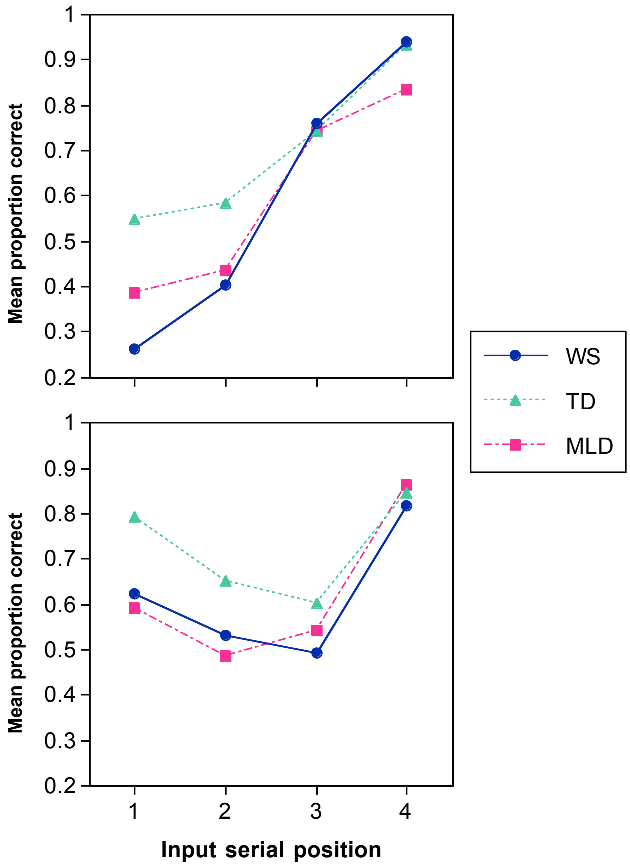

Figure 7.6 Serial position curves for the probed position recall task. Upper panel shows un-normalized data. Lower panel shows normalized data

Correlations with position recall

Table 7.3 shows the correlations of various measures with overall performance on the position recall task. First, looking at estimates of item and order memory obtained from the error analysis, position recall was highly correlated with order memory for both words and nonwords, and with item memory for nonwords but not with item memory for words. Position recall was also significantly correlated with digit span but not with BPVS age, with Ravens matrices score, or with temporal discrimination abilities.

The pattern of correlations was also similar when the individual groups were considered, although because of the smaller group sizes, many of the correlations failed to achieve significance. The one surprising finding was a significant negative correlation between position recall and estimated item memory for words in the TD group. Correlational analyses were repeated with Ravens matrices raw score as a partial correlate, but this did not have any qualitative effect on the pattern of correlations.

There was a significant correlation between position recall and vocabulary IQ (cf. Jarrold et al., 2002; see above) but when the individual groups were considered there were no significant correlations. Moreover, the correlation between position recall and BPVS standard score in the TD group was negative and non-significant.

Table 7.3 Correlations with position recall

	All	WS	TD	MLD
Order memory (words)	.798**	.673**	.781**	.800**
Order memory (nonwords)	.607**	.421	.647**	.591*
Item memory (words)	-.041	.086	-.481*	.129
Item memory (nonwords)	.366*	.166	.425	.398
Digit span score	.515**	.430	.322	.584*
BPVS age equivalent	.133	-.187	.107	.464
BPVS standard score	N/A	N/A	-.089	N/A
Vocabulary IQ [a]	.441**	.183	-.047	.237
Temporal discrimination score [b]	.269	.065	.364	.316
Ravens matrices raw score †	.261	-.305	.326	.230

* Correlation significant p<.05 (two-tailed)
** Correlation significant p<.01 (two-tailed)
a BPVS age divided by chronological age
b Mean score on the three temporal discrimination tasks from Exp 2
† Based on N=13 for the WS group and N=12 for the MLD group

7.3.4: Discussion

Experiment 3 investigated item and order memory and the effect of lexicality using a probed recall task.

Overall performance

The overall performance of the WS group on the probed recall tasks was significantly lower than that of the TD group, but at a comparable level to that of the MLD group. Similarly, the TD group significantly outperformed the WS group on the forward digit span subtest of the WISC, but there was no significant difference between the WS and MLD groups. This result contradicts the findings of Mervis et al. (1999) who reported that standard scores for digit span in WS were superior to those for vocabulary. This may reflect the fact that Mervis et al. did not include a control group and relied on the assumption that standard scores are equivalent across different tests (cf. Bishop, 1997; section 2.2.1). Alternatively, it may reflect the much broader age range of the participants tested by Mervis et al..

The results of Experiment 3 therefore failed to find evidence for relatively preserved verbal STM in WS. However, as noted in section 5.3, Jarrold et al. (2002) have recently argued that, if verbal STM is involved in vocabulary acquisition, then performance on STM tasks should be related to vocabulary IQ (i.e. the rate of vocabulary acquisition) rather than the absolute level of vocabulary. As such, it is perhaps not surprising that children with WS had inferior STM to TD children because, despite being matched on vocabulary level, the WS group were much older and therefore had much slower rates of vocabulary acquisition. In fact, a more valid comparison may be with the MLD controls who were of similar age and therefore had comparable rates of vocabulary acquisition. The failure to find any differences between the WS and MLD groups on any STM measure suggests that STM abilities are in line with vocabulary IQ but not above.

The findings of Experiment 3 therefore provide some support for Jarrold et al.'s (2002) argument. These authors tested their argument by comparing two groups of MLD children who were matched on vocabulary age but differed in terms of chronological age. The younger (high vocabulary IQ) group outperformed the older (low IQ) group on digit span, nonword repetition and serial order recognition. The results of Experiment 3 followed a similar pattern insofar as the groups were matched on vocabulary but the younger group (i.e. the TD children) outperformed the two older groups (i.e. the WS and MLD children), on probed recall and digit span.

Item versus order memory

Although there were group differences in both tasks, these appear to reflect differences in order memory, with groups being equivalent in terms of item memory. More specifically, the TD group appeared to have superior order memory to the WS and MLD groups with little evidence for differences between the two learning-disabled groups. The position recall task is assumed to assess order memory and the TD group significantly outperformed the other groups on this task. In contrast, the item recall task is assumed to rely on both item and order memory, and once order memory had been controlled for by covarying out performance on the position recall task, there were no group differences in performance on this task. These findings were complimented by error analysis of the item recall task which showed a significant effect of group on estimated order memory but not on estimated item memory.

This pattern of results suggests that order memory is related to rate of vocabulary acquisition rather than level of vocabulary knowledge. Following Jarrold et al.'s (2002) argument, this is consistent with the idea that order memory has a causal influence on vocabulary acquisition. However, there are other possible explanations for this pattern of results. For example, Pennington (2001) has recently reported that the performance of young and elderly adults on a similar probed position recall task was strongly predicted by fluid ability (AH4 score; Heim 1968). Although the groups in Experiment 3 were matched on vocabulary, the TD group may have had superior fluid intelligence[17], and their superior order memory could simply be a reflection of this.

Lexicality effect

The results of Experiment 3 were broadly compatible with the findings of Turner et al. (2000) using the same paradigm. First, in the probed position recall task, there was no effect of lexicality, but in the probed item task, words were better recalled than nonwords. Of primary interest, though, was the finding that in the probed item recall task, children with WS showed the same sized lexicality effect as TD and MLD controls. This result stands in contrast to the reduced word-frequency effect identified by Vicari, Carlesimo et al. (1996) and the reduced lexicality effect reported by Majerus et al. (2001), and provides evidence against both the impaired redintegration and impaired semantic memory hypotheses.

One possible concern may be that Experiment 3 lacked statistical power and that a more powerful design would have produced the predicted interaction between group and lexicality. However, there were more participants in the WS group (N=14) than in either the Vicari et al. study (N=12) or the Majerus et al. study (N=4). Furthermore, there was a much greater distribution of scores in the current study than was possible in the Vicari et al. study. Alternative explanations for the findings of Vicari et al. are explored in the following two experiments.

Lexicalization errors

Error analysis of the nonwords condition in the probed item recall task permitted investigation of the finding by Karmiloff-Smith et al. (1997) that in a nonword repetition task, individuals with WS made fewer lexicalisation errors than controls. The main difficulty with the Karmiloff-Smith et al. study was that controls had a significantly lower verbal mental age, and therefore, it may be assumed, poorer STM. Models of redintegration (e.g. Schweickert, Chen, & Poirier, 1999) assume that long-term vocabulary knowledge is only employed if phonological traces become degraded, so a reduction in lexicalization errors may reflect superior STM rather than impaired redintegration. In the current study, the three groups were matched on vocabulary knowledge and, it transpired, on item memory. The lack of any group differences in the proportion of lexicalization errors therefore suggests that the findings from the Karmiloff-Smith study reflect the matching procedure rather than any fundamental differences in memory processes in WS.

Serial position effects

Serial position effects were analysed for the position recall tasks. Once response biases had been accounted for by normalizing the data, all three groups showed similar bow-shaped serial position curves with large recency effects and small primacy effects. The performance advantage for the TD group was most pronounced for earlier positions but there was no significant group by position interaction. Unfortunately a large number of participants had to be excluded from this analysis because of their extreme response biases, and it is unclear if group effects would have arisen if more participants could have been included.

Correlations with position recall and digit span

Correlational analyses revealed an interesting pattern of associations, although all results should be treated with caution given the heterogeneous nature of the overall participant group and the small sizes of the individual groups.

Overall performance on the position recall task was highly correlated with estimates of order memory obtained from error analysis of the item recall task, thus validating the assumption that the position recall task is a measure of order memory. Position recall was also significantly correlated with item memory for nonwords but not with item memory for nonwords. This is consistent with the argument put forward in section 6.3.4 that representing nonwords in STM relies on some form of serial order mechanism.

It was also suggested in section 6.3.4 that digit span performance is primarily determined by memory for serial order. Consistent with this suggestion, position recall and digit span were significantly correlated. However, while vocabulary age was significantly correlated with digit span, there was no significant correlation between vocabulary age and position recall. One important difference between the digit span and probed position recall tasks is in terms of their speech output demands. It could therefore be argued that the relationship between digit span and vocabulary is mediated by speech output processes. However, Gathercole, Service et al. (1999) reported a significant association between vocabulary and performance on a serial recognition test that, like the probed position recall task, has limited output demands and can be considered as a relatively pure test of order memory. Moreover, both position recall and digit span were highly correlated with vocabulary IQ. This is consistent with the argument that serial order memory determines the rate of vocabulary acquisition rather than the absolute level (cf. Jarrold et al., 2002).

The relationship between order memory and temporal discrimination

Experiment 3 also allowed comparison between performance on the probed position recall task and performance on the temporal discrimination tasks in Experiment 2. It has been proposed that temporal discrimination abilities and order memory both depend on oscillatory timing mechanisms (Brown et al., 2000; Church & Broadbent, 1990), and in section 5.3 it was suggested that serial order memory could provide a link between temporal discrimination and language abilities. However, there were no significant correlations between position recall and temporal discrimination. Moreover, the pattern of performance on the two tasks was different; while order memory was better in the TD group than in the two learning-disabled groups, all three groups had equivalent temporal discrimination abilities.[18]

There is therefore little support for the proposed association between temporal discrimination and serial order memory. It may be that the processes involved are completely different, or that the processes are similar but the actual mechanisms are independent. Alternatively, there could be some overlap in the mechanisms involved, but most variation in performance arises from non-shared processes. For example, Brown et al. (1999) demonstrated that developmental improvements in serial order memory could be modelled by OSCAR, not only by increasing the speed of the fastest oscillators, but also by reducing noise in the retrieved temporal context signal. The same oscillatory mechanisms could therefore be involved in temporal discrimination and *encoding* of serial order, but group differences in Experiment 3 could reflect problems with *retrieving* the temporal context signal in the WS and MLD groups relative to the TD group. This latter explanation leaves open the possibility that other groups (e.g. individuals with SLI or Down syndrome) may have deficits in STM and language that are caused by impaired oscillatory timing mechanisms (cf. Boucher, 1999).

Summary

The results of Experiment 3 failed to support the hypothesis that serial order memory (or STM abilities as a whole) are a relative strength in WS. Order memory in WS children was inferior to that of TD children matched on vocabulary, and was at the same level as that of MLD children matched on vocabulary and chronological age (and therefore on rate of vocabulary acquisition).

There was also no evidence for a reduced influence of lexical knowledge on STM in WS. Children with WS and controls showed similar tendencies to make lexicalization errors in the recall of nonwords, and demonstrated similar lexicality effects in probed item recall. This latter finding contrasts strongly with the reduced word-frequency effect in WS reported by Vicari, Carlesimo et al. (1996). Possible reasons for this contradiction are explored in Experiments 4 and 5.

7.4: Experiment 4: Probed recall for low-frequency words

7.4.1: Introduction

One important difference between Experiment 3 and the study by Vicari, Carlesimo et al. (1996) is that Vicari et al. compared high-frequency words with low-frequency words, whereas in Experiment 3, recall of words was compared with recall of nonwords. Given the general assumption that the lexicality effect is simply an extreme version of the word-frequency effect (cf. Hulme et al., 1991), it seems unlikely that lexicality and word-frequency effects will be affected differentially in WS. However, it has been suggested that individuals with WS have a reduced sensitivity to word-frequency (e.g. Rossen et al., 1996). It remains possible, therefore, that individuals with WS show a reduced word-frequency effect but a normal lexicality effect. This possibility was explored in Experiment 4.

Lexicality effects in Experiment 3 were only present in the item recall task and the words used in Experiment 3 were all high-frequency. Therefore, in Experiment 4, the same participants were tested on item recall for low-frequency words. Results from Experiment 4 could then be compared with item recall of high-frequency words and nonwords in Experiment 3. Furthermore, by testing participants on further trials of the item recall task, it became possible to investigate serial position effects in item recall by collapsing across different types of stimulus.

7.4.2: Method

7.4.2.1: Participants

Participants were the same as those tested in Experiment 3 (see Table 7.1). Control children in both groups were tested shortly after testing for Experiment 3, but practical constraints entailed that most children with WS were tested up to nine months after testing for Experiment 3. Consequently, some of the WS children had to be re-tested on the BPVS. The mean BPVS age of the WS group at the time of testing for Experiment 4 was 9;1 years (SD=1;7), which was higher than that for both control groups, although a one-way ANOVA failed to show any significant effect of Group on BPVS age ($F(2,44)=2.41$, $p=.102$).

7.4.2.2: Materials

Two sets (A and B) of 20 one-syllable CVC words were constructed (see Appendix 1). These words all had low familiarity ratings (less than 500) and low imageability ratings (less than 500) according to the MRC Linguistic Database (Quinlan, 1992). These words were matched to the high-frequency words in Experiment 4 on imageability ratings but had significantly lower familiarity ratings. The words in Experiment 4 also had lower written frequencies according to both the Kucera and Francis (1967) and Thorndike and Lorge (1944) norms. Words were recorded by the same native English-speaking female as in Experiment 3 and stored individually as SDII computer sound files. Materials were presented using the same computer as in Experiment 3.

7.4.2.3: Procedure

When testing for Experiments 3 and 4 occurred in separate sessions, the practice task was repeated with real cards before conducting the computerised recall of low-frequency words. Participants were randomly assigned one of the two word lists such that an approximately equal number of children in each group received each list. Apart from the items used, the procedure for Experiment 4 was exactly the same as for Item Recall for words in Experiment 3.

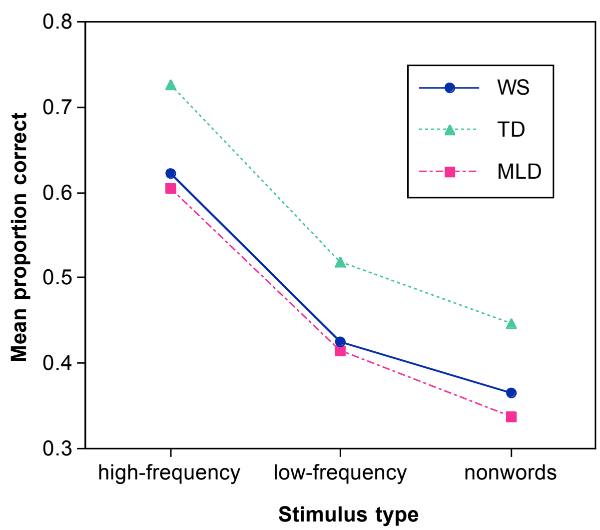

Figure 7.7 The effect of word-frequency on performance on the probed item recall task

7.4.3: Results

Overall performance

The overall performance for item recall for high-frequency words (Experiment 3), low-frequency words (Experiment 4) and nonwords (Experiment 3) is shown in Figure 7.7. As might be expected, performance on the low-frequency words was better than for nonwords but worse than for high-frequency words. Figure 7.7 also provides little evidence for an interaction between Group and Word-frequency.

These observations were confirmed by a two-way mixed design ANOVA with group (WS vs TD vs MLD) as a between-subjects factor and word-frequency (High-frequency words vs Low-frequency words vs Nonwords) as a repeated measure. There was a significant main effect of group ($F(2,44)=4.42$, $p<.05$) which was due to significant differences between the TD and MLD groups (Tukey HSD). There was also a significant effect of word-frequency ($F(2,88)=109.08$, $p<.001$) with recall of high-frequency words being superior to recall of low-frequency words which was in turn superior to recall of nonwords (Bonferroni correction applied). However, there was no significant interaction between group and word-frequency ($F(4,94)=0.07$).

It is possible that the use of three-level factors in ANOVA may mask differences which may be found if certain two-way comparisons were made. The above analyses were therefore repeated in a series of nine ANOVAs for all possible two-way comparisons between groups, combined factorially with all possible two-way comparisons between high- and low-frequency words and nonwords. Despite the nonconservative nature of this approach there was no evidence for an interaction between group and word-frequency in any of these analyses.

As mentioned above, most of the children in the WS group were tested on the low-frequency words several months after testing on the high-frequency words and nonwords, whereas the controls were all tested on all the stimuli within a few weeks. If the size of the word-frequency effect is dependent on vocabulary knowledge, this could have distorted the findings. BPVS ages at the time of testing for Experiment 4 were therefore taken as covariates in an ANOVA. The dependent variable was the size of the word-frequency effect (performance on high- minus performance on low-frequency words), and group was a between-subjects factor. However, there was no effect of BPVS age ($F(1,43)=0.00$) and no effect of group ($F(2,43)=0.02$).

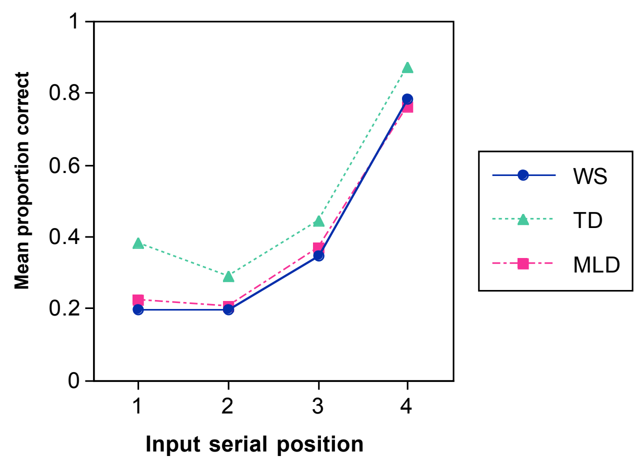

Figure 7.8 Serial position curves for the probed item recall task

Serial position effects

In order to analyse serial position effects, results were collapsed across the three levels of word-frequency. Figure 7.8 shows the number of items correctly recalled as a function of serial position in the four-item trials (ceiling effects precluded analysis of serial position effects in the three-item trials). All three groups showed strong recency effects with little evidence of any primacy. A mixed design ANOVA was performed with group as a between-subjects factor and serial position as a within-subjects factor. There was a significant effect of group ($F(2,44)=4.83$, $p<.05$) reflecting an advantage for the TD group over the other two groups (Tukey HSD). There was also a significant effect of position ($F(3,132)=129.67$, $p<.001$). With the Bonferroni correction, all pairwise comparisons were significant, apart from that between the first and second serial positions. However, there was no significant interaction between group and serial position ($F(6,132)=0.48$).

7.4.4: Discussion

In Experiment 3, there was no evidence for a reduced lexicality effect in WS. This finding was extended in Experiment 4 which found no evidence for a reduced word-frequency effect in WS. One potential concern with Experiment 4 is that the delay between testing on high-frequency words and nonwords (Experiment 3) and testing on low-frequency words (Experiment 4) was on average much larger in the WS group. Thus, if the children with WS had performed relatively well on the low-frequency words (as Vicari, Carlesimo et al.'s (1996) results would predict), this could have been caused by the increase in their vocabulary knowledge between the first and second test stage. However, the absence of a group by word-frequency interaction suggests that this is not a major concern and certainly does not change the interpretation of the results. Moreover, when current BPVS score was taken as a covariate, there remained no effect of Group on the size of the word-frequency effect.

Experiment 4 therefore provided evidence that the discrepancies between the reduced word-frequency effect in WS reported by Vicari, Carlesimo et al. (1996) and the normal lexicality effect reported in Experiment 3 were not caused by the fact that the former investigated the word-frequency effect, whereas the latter investigated the lexicality effect. Moreover, the congruence of findings in Experiments 3 and 4 is consistent with the assumption that word-frequency and lexicality effects in traditional serial recall tasks reflect the same underlying processes and that the lexicality effect is simply an extreme version of the word-frequency effect with nonwords representing words with a frequency of zero (cf. Hulme et al., 1991).

Experiment 4 also allowed the analysis of serial position effects in four-item trials by collapsing performance across the three levels of word-frequency. All three groups showed significant recency effects, performing significantly better when the target word was in the final position than when it was in the other three positions, but showed little evidence of a primacy effect (relatively good performance on the first item). Importantly, the lack of a significant Group by Serial Position suggests that, although the TD children performed better than the other groups, all three groups performed the task in a similar fashion.

It still remains, therefore, to explain why the results of Experiments 3 and 4 should differ from those reported by Vicari, Carlesimo et al. (1996). One difference between the studies that has not yet been considered is the recall paradigm used. Vicari et al. employed a traditional serial recall paradigm, whereas a probed recall task was used in Experiments 3 and 4. The advantages of the probed recall task were outlined earlier but it is possible that the effect reported by Vicari et al. is specific to the serial recall paradigm. This possibility was explored in Experiment 5 which attempted a more direct replication of the Vicari et al. study. This experiment is reported in the next chapter.

---

[14] This is consistent with the finding that lexicality has little effect on performance on a serial order recognition task which, like the probed position task, can be considered as a test of order rather than item memory (Gathercole et al., 2001).

[15] This was one of the three temporal discrimination tasks from Experiment 2.

[16] Ceiling effects precluded similar analysis of the three-item task. Experiment 4 looked at word-frequency effects in probed item recall, so serial position curves for the item recall task are reported in Experiment 4 with data from Experiments 3 and 4 being combined.

[17] Vocabulary is a measure of crystallized intelligence (i.e. accumulated knowledge), so the three groups could still be matched on vocabulary because the TD group were much younger than the other two groups.

[18] A similar pattern of results was reported by Pennington (2001). Young adults outperformed elderly adults on a probed position recall task, but there was no group difference in performance on a temporal discrimination task.

Chapter 8: The word-frequency effect in forward serial recall

8.1: Overview

Experiments 3 and 4 investigated probed serial recall performance in WS. However, in contrast to previous studies using traditional forward serial recall tasks (Majerus et al., 2001; Vicari, Carlesimo et al., 1996), there was no evidence for a reduced influence of lexical-semantic knowledge on performance in the WS group. Experiment 5 investigated the possibility that this discrepancy reflected the particular recall paradigm that was used.

8.2: Experiment 5: Word-frequency effects in serial recall

8.2.1: Introduction

In Experiment 5, the participants who had taken part in Experiments 3 and 4 performed a traditional serial recall span task, similar to that used by Vicari, Carlesimo et al. (1996). Vicari et al. compared recall of high- and low-frequency words but because they tested Italian children, it was not possible to use exactly the same stimuli in Experiment 5. As a result, the familiarity of the stimuli in the two studies may have been different. In Experiment 5 there were therefore three levels of word-frequency (high, medium, and low). Given Vicari et al.'s findings, one would predict a reduced effect of word-frequency in WS in at least one of the three possible pairwise comparisons.

Vicari, Carlesimo et al. (1996) measured word span in each condition by giving participants lists of increasing length starting with two words. If a list was correctly repeated then the child was tested on a longer list. If the child failed twice at a particular length then testing halted and span was calculated as the longest list-length correctly repeated. However, from the reported means and standard deviations it is clear that there was a relatively narrow range of spans obtained, with, in some conditions, participants scoring either 0 (i.e. they were unable to repeat any lists) or 2. Because it was impossible to score 1, this means that group differences in these conditions may be amplified relative to conditions in which intermediate scores were possible.

In Experiment 5, the task was designed such that three different measures of recall performance could be taken. However, this required that participants were tested on an equal number of lists for each condition. They were therefore presented with six lists (i.e. two each of high-, medium-, and low-frequency words) up to the point at which they failed to correctly recall any list in any condition. This contrasted with the Vicari, Carlesimo et al. (1996) study in which each condition was tested in a separate block, so that participants were often tested on a different number of lists for each condition.

The first measure, 'span', was defined as the list length preceding that at which the child first made two errors in that condition (even if subsequently they correctly recalled a longer list in that condition). This allowed direct comparison with the Vicari, Carlesimo et al. (1996) study, because these authors stopped testing once two errors had been made in a particular condition. The second measure, 'lists', was simply the number of lists correctly recalled in each condition. Finally, 'items' was the proportion of items correctly recalled. Omission errors often make it difficult to decide whether an item is recalled in its correct position (for example, if ABCDE is recalled as ABCE then the item E is not strictly recalled in the correct position). However, word-frequency effects are generally found to affect memory for items but not memory for order (Saint-Aubin & Poirier, 1999), so items were considered to be correctly recalled even if they were recalled in the incorrect position. The proportion of items recalled was analysed rather than the number of items because the absolute size of the word-frequency effect is likely to be larger for participants who perform better and are therefore tested on more items.

8.2.2: Method

8.2.2.1: Participants

The participants were the same 14 children with WS (9 boys), 20 TD children (10 boys), and 13 MLD children (8 boys) who were tested in Experiments 3 and 4 (see Table 7.1). Testing took place at the same time as for Experiment 4 so the WS group in fact had slightly higher BPVS ages than controls (cf. Section 7.3.2.1).

8.2.2.2: Materials

Stimuli were 60 two-syllable words, comprising 20 high-, 20 medium- and 20 low-frequency words (see Appendix 2). Word-frequency was determined according to the frequency of occurrence in children's reading schemes. This was calculated by summing occurrences in four reading schemes; Children's Story Books, Rhymeworld, Oxford Reading Tree and Keystage 1 Recommended Texts (see Solity & Vousden, 2002). Low-frequency words had an incidence of 0 and therefore did not appear in any of the reading schemes. All medium-frequency words had an incidence of 1 or 2. High- frequency words had occurrences of greater than 10. Word groups also differed significantly according to measures of written frequency, age of acquisition and familiarity as reported in the MRC Psycholinguistic Database (Quinlan, 1992) but were matched on number of phonemes, concreteness and imageability.

Word lists were arranged in blocks of increasing length, starting from two-word-lists. There were six lists in each block, and each block consisted of two sub-blocks containing one list of each word-type. The order of the three conditions within each sub-block cycled randomly without replacement through the six possible orders. Words never appeared more than once in the same block nor did they appear twice in consecutive blocks. To ensure this constraint was met, it was practical to test all participants with the same test sheet. Consequently, the same lists were presented in the same pseudo-random order for each participant.

8.2.2.3: Procedure

Children with WS were tested individually in a quiet room in their own home. Both control groups were tested individually in a quiet room at their own school. Participants were told that they would be playing a memory game, that the experimenter would say a list of words, and that they should try and repeat those words in the same order. Words were presented at a rate of approximately one item every 1.5 seconds. If the participant repeated at least one list in a block correctly then they proceeded to the next block with longer lists. Testing was discontinued when all six lists in a block were repeated incorrectly.

8.2.3: Results

The upper panel of Figure 8.1 shows performance according to span. All three groups showed similar effects of word-frequency, although there is some

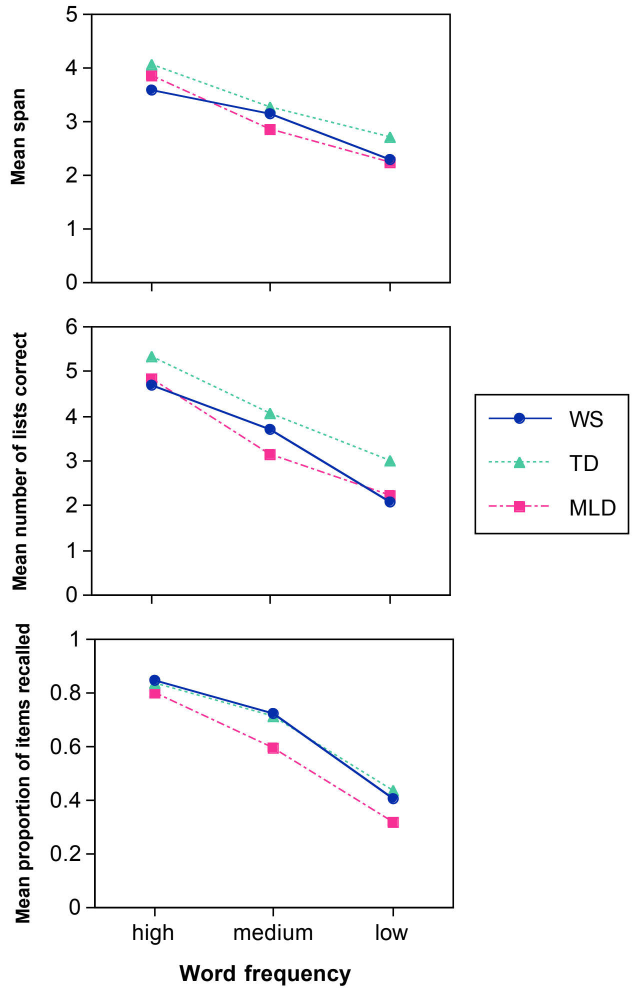

Figure 8.1 The effects of word-frequency on forward serial recall

indication of relatively good performance on the medium-frequency words by the WS group. Performance across all conditions was slightly better in the TD group than in the other groups. Scores were analysed with a mixed design ANOVA with group (WS vs TD vs MLD) as a between-subjects factor and word-frequency (high vs medium vs low) as a repeated measure. There was a significant effect of word-frequency (F(2,88)=47.44, p<.001) with all pairwise comparisons being significant (Bonferroni correction applied). However, the main effect of group was not significant (F(2,44)=1.61), and there was no significant interaction between group and word-frequency (F(4,88)=0.60).

The middle panel of Figure 8.1 shows performance according to lists, and a similar pattern to that observed for span. ANOVA again revealed a significant effect of word-frequency (F(2,88)=84.22, p<.001) with all pairwise comparisons being significant (Bonferroni correction), a non-significant main effect of group (F(2,44)=2.04), and no interaction between group and word-frequency (F(4,88)=0.80). Finally, the lower panel of Figure 8.1 shows performance according to items, and again all three groups demonstrated similar effects of word-frequency. Compared with the other two groups, the MLD children recalled a lower proportion of items in all conditions, although because results were divided by the total number of items presented, this does not give an indication of overall performance. There was a significant effect of word-frequency (F(1.72,88)=217.16, p<.001; Greenhouse-Geisser correction applied), and all pairwise differences were significant. The main effect of group narrowly failed to reach significance (F(2,44)=2.712, p=.078), but there was no significant interaction between group and word-frequency (F(4,88)=1.094; sphericity assumed).

Analyses were repeated for each measure and for each of the possible two-by-two comparisons of group and word-frequency. Despite the non-conservative nature of this procedure, there were still no significant interactions. Because the WS group had slightly higher vocabulary scores than the control groups, the analyses were also repeated with BPVS age as a covariate. However, the findings were almost exactly the same, with F-statistics for the group by word-frequency interactions decreasing in each case.

8.2.4: Discussion

The results of Experiment 5 complemented and extended those of the two previous experiments. Children with WS and MLD controls again performed worse than TD controls although, unlike in Experiments 3 and 4, the group difference was not significant. Moreover, in Experiment 5, children with WS showed the same-sized word-frequency effect in serial recall as did TD and MLD controls matched on vocabulary. This finding is consistent with Experiments 3 and 4 in which equivalent lexicality and word-frequency effects were found in WS and control groups using a probed recall paradigm, and suggests that the discrepancies between Experiments 3 and 4 and the earlier study of serial recall by Vicari, Carlesimo et al. (1996) cannot be explained in terms of the specific recall paradigm used.

Three different measures of performance were used and the results were consistent across measures. The first, span, was equivalent to that used by Vicari, Carlesimo et al. (1996), although they averaged span over two different word-lengths whereas in Experiment 5, only two-syllable words were tested. In this respect, the Vicari et al. study had more statistical power, and this could potentially explain why they reported a significant interaction whereas the interaction in Experiment 5 did not reach significance. However, the second measure, lists has twice the number of possible scores as Span (as there were two trials per condition at each list length) and therefore had the same statistical power as the measure of span employed by Vicari et al.. Moreover, it avoided the difficulty with span that it is impossible to score 1 although possible to score 0. Again there was no interaction between word-frequency and group. Finally, items had the greatest range of possible scores and there was no evidence at all for an interaction between word-frequency and group.

It still remains to be explained why Experiment 5 failed to replicate Vicari, Carlesimo et al.'s (1996) finding of a significant reduction in the word-frequency effect in WS. In Section 6.4.4.1, it was argued that Vicari et al.'s results were questionable because their WS group probably had much better vocabulary knowledge than controls and may therefore have been more familiar with the low-frequency words than were controls. The WS group in Experiment 5 also had slightly higher vocabulary ages than controls but this did not appear to affect the results. One possibility is that the discrepancy in vocabulary scores was much bigger in the Vicari et al. study than in Experiment 5, and this could explain why they found group differences.

Another possibility is that the choice of stimuli may be particularly important. Unfortunately, because Vicari, Carlesimo et al. (1996) tested Italian children, it was impossible to replicate their study exactly with English-speaking children. Nevertheless, this explanation is supported by the observation that there was a non-significantly reduced word-frequency effect in WS when span and lists were the dependent variables and when high- and medium-frequency words were compared, but this was not apparent in any other two-way contrasts between high-, medium- or low-frequency words. Thus with small group sizes, and with dependent variables which offer a small range of scores, it may be that differences in knowledge of particular words can sometimes lead to statistically significant effects.

8.3: Summary

There were two main findings in Experiment 5, which complement those from Experiments 3 and 4. First, there was no evidence for superior STM in WS. Second, individuals with WS showed a normal word-frequency effect. The results of Experiments 3, 4, and 5 therefore fail to support either the impaired redintegration hypothesis or the impaired semantic memory hypothesis.

Chapter 9: The concreteness effect in forward serial recall

9.1: Overview

Vicari, Carlesimo et al. (1996) reported that children with WS showed a reduced word-frequency effect in serial recall in WS, and argued that there is a reduced contribution of semantic information to STM performance in WS. However, while the word-frequency and lexicality effects may in part reflect the availability of semantic information, the main cause of these effects appears to be that participants are able to use knowledge of the sound patterns of the words to complete degraded traces (e.g. Hulme et al., 1997). As such, it is far from clear that reduced word-frequency effects in WS are related to semantic processing. By the same token, the fact that Experiments 3, 4, and 5 failed to find any evidence for reduced word-frequency or lexicality effects in WS does not necessarily rule out the possibility that there is some form of semantic memory impairment. Experiment 6 therefore investigated semantic influences on STM in WS by looking at the concreteness effect on serial recall.

9.2: Semantic effects on STM

Two main findings have been used to demonstrate the influence of semantic factors on STM. First, serial recall of lists of semantically related words is superior to that for lists of unrelated words (Huttenlocher & Newcombe, 1976; Poirier & Saint-Aubin, 1995; Wetherick, 1975; but see Baddeley, 1966). Poirier and St Aubin (1995) suggested that this effect occurs because retrieval of a word is aided by the ability to narrow the potential candidates down to a particular semantic category. If individuals with WS do rely more on the phonological than the semantic trace then one would expect them to show a reduced effect of semantic similarity. However, because there appears to be a strategic element to the semantic similarity effect, one could not be entirely sure that a reduced effect represented a deficit in semantic processing or simply a failure to utilise this strategy.

A second finding is that concrete or imageable words are better recalled than abstract words (e.g. Bourassa, & Besner, 1994; Nation et al., 1999; Walker & Hulme, 1999). Walker and Hulme (1999) suggested that this concreteness effect could be understood in terms of the Feature Model (e.g. Neath & Nairne, 1995), according to which, primary memory (STM) traces consist of features representing different aspects of the stimulus. Compared with traces of abstract words, traces of concrete words are assumed to have more meaning-based features associated with them (cf. Plaut & Shallice, 1993), and will therefore be more distinctive and resistant to interference. Walker and Hulme also suggested that the redintegration account could be extended to include semantic aspects of stimuli. Thus, temporary phonological and semantic traces can be redintegrated by comparison with permanent phonological and semantic traces respectively. Under the assumption that representations of concrete words contain more unique information than those of abstract words (cf. Plaut & Shallice, 1993), the success of redintegration should increase as a function of concreteness.

If children with WS have impaired semantic memory then one might expect them to be less susceptible to the influence of concreteness, and consequently they should show a reduced concreteness effect. This prediction contrasts with that made by the impaired redintegration hypothesis. So long as abstract and concrete words are equated for frequency, this predicts that children with WS should show a normal concreteness effect.

9.3: Experiment 6: The concreteness effect in serial recall

9.3.1: Introduction

Experiment 6 was a preliminary study of the concreteness effect in which serial recall for concrete and abstract words was compared in children with WS and TD control children.[19] The procedure was based on that used by Nation et al. (1999) because it had been shown to work with children, and to produce reliable group differences in the size of the concreteness effect. The task was a traditional forward recall task, with each participant receiving a number of five-item lists and performance being measured in terms of the number of items correctly recalled. The words were all monosyllabic and of relatively high written frequency, and each list contained either all concrete words or all abstract words.

In Experiment 5, most of the WS group had spans of three or four for two-syllable high-frequency words. Given that shorter words are recalled better than longer words (e.g. Baddeley, Thomson, & Buchanan, 1975), it was assumed that these children would have slightly longer spans for monosyllabic words. All participants were therefore tested on five-item lists. The list length was set slightly higher than expected span because ceiling effects could distort the results. Participants would be unlikely to recall no items at all so there should have been no problem with floor effects whatever list length was tested. Moreover, testing all participants on exactly the same list length removed the need to adjust scores to control for the number of items tested, and allowed analysis of serial position effects. The effect of word-frequency is greatest at later serial positions (Hulme et al., 1997), whereas concreteness does not interact with serial position (Walker & Hulme, 1999), which suggests that the two effects have different origins. Analysing serial position effects could therefore provide information about the source of the effect in the two groups.

The other important issue here is the procedure used for matching controls. It was decided to match groups on recall of concrete words because potential group differences in effect size could not then be explained in terms of differences in overall performance, and analysis of serial position effects would be relatively straightforward. In the absence of any evidence for relationships between the size of the concreteness effect and chronological or language age, receptive vocabulary was taken as a background measure which could be controlled for by covariation if necessary.

9.3.2: Method

9.3.2.1: Participants

The participants were 12 of the children with WS (8 boys) who had taken part in Experiments 1 to 5. Twelve TD control children (4 boys), two of whom had previously taken part in Experiments 2 to 5, were matched to the WS group on overall performance in the concrete words condition. Participant details are shown in Table 9.1.

Table 9.1 Participant details for Experiment 6

	WS N=12		TD N=12	
	M	SD	M	SD
Chronological age (years; months)	13;11	2;4	6;6	0;4
BPVS age (years; months)	9;2	1;7	6;10	1;1
Ravens matrices age (years; months) †	7;7	1;9	7;3	0;11

† Based on N=11 for the WS group

9.3.2.2: Materials

The words used for testing were the same 16 one-syllable concrete words and 16 one-syllable abstract words used by Nation et al. (1999) and by Walker and Hulme (1999)[20]. The concrete words received significantly higher concreteness ratings than abstract words (Walker & Hulme), but the two groups of words were matched in terms of written frequency (Kucera & Francis, 1967; see Walker & Hulme) and frequency in children's reading materials (Carroll, Davies, & Richman, 1971; see Nation et al.).

Following Nation et al. (1999), 32 five-word lists were constructed, half of which contained only concrete words, the other half containing only abstract words. Words were never repeated within a list and each word appeared once in each of the five possible serial positions. Nation et al. tested participants separately on concrete and abstract words but in the current experiment, the word lists were divided into two sets, A and B, each containing 8 concrete and 8 abstract lists. These were presented in a pseudo-random order subject to the constraint that the same word never appeared in consecutive lists and the distribution of concrete and abstract words in one list mirrored that in the other.

9.3.2.3: Procedure

Participants were tested individually in a quiet room. As with all other experiments, WS participants were tested at home whereas controls were tested in school. Testing was conducted in two blocks, with half of the children in each group being tested on set A first and half being tested first on set B. Between the two blocks, participants performed a filler task. In the WS group this task was a computerised game undertaken as a pilot study of auditory streaming. Because the pilot study was not successful, the filler task for the TD children was either the BPVS or the Ravens matrices, depending upon time constraints. In all cases, the two blocks of the experiment were completed within the same session.

Participants were told that they would hear lists of five words and should try and recall them all in the correct order. Following Nation et al. (1999), they were told that if they could not remember a word in a particular position they should say 'blank' and then move on to the next word. The experimenter then read each list aloud at a rate of approximately one word every 1.5 seconds and recorded the output position of each response. This aspect of the procedure contrasted with the Nation et al. study in which presentation was computerised.

9.3.3: Results

Few of the participants in either group were able to recall any lists completely correctly. Unfortunately, most of the participants were also unable to comply with the instructions to say 'blank' when they could not recall a word in a particular position; rather they tended to simply miss items out. Furthermore, a number of participants consistently began recall with the last item, indicating that they were not even trying to recall items in the correct order. This made it difficult to analyse performance in terms of the number of items recalled in their correct position and prevented analysis of serial position effects. Performance was therefore analysed in terms of the proportion of items correctly recalled, regardless of whether they were output in the correct position.

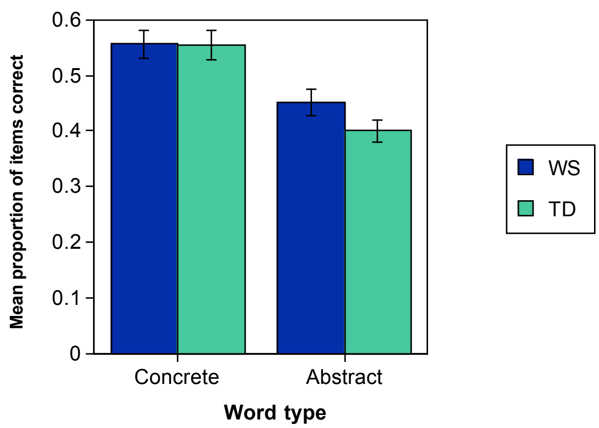

Figure 9.1 The effect of concreteness on forward serial recall. Error bars show standard errors.

Overall performance

Figure 9.1 shows the mean proportion of items correctly recalled by children in each group as a function of concreteness. Both groups performed better on the concrete words than on the abstract words, although the size of this concreteness effect was smaller in the WS group. Results were subjected to a two-way ANOVA with group as a between-subjects factor and concreteness (concrete vs abstract) as a repeated measure. There was no main effect of group ($F(1,22)=0.70$) but a significant effect of concreteness ($F(1,22)=57.07$, $p<.001$), with both the WS group ($t(11)=3.80$, $p<.01$) and the TD group ($t(11)=7.55$, $p<.001$) recalling significantly more concrete words than abstract words. However, the interaction between group and concreteness was non-significant ($F(1,22)=2.10$).

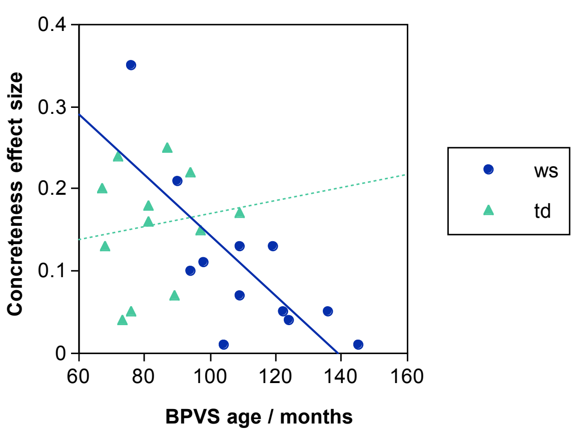

Figure 9.2 The relationship between vocabulary age and the size of the concreteness effect (i.e. the proportion of concrete words recalled minus the proportion of abstract words recalled). Solid and dashed lines show the regression slopes for the WS and TD groups respectively.

The effect of vocabulary age

Because groups were not matched on vocabulary, it was important to determine whether the size of the concreteness effect depended on vocabulary knowledge. Figure 9.2 shows the relationship between BPVS age and the size of the concreteness effect (i.e. the difference between recall of concrete and abstract words). Regression analysis was performed with BPVS age as the independent variable and the size of the concreteness effect as the dependent variable. The results are shown in Table 9.2 which reveals a significant reduction in the size of the concreteness effect as BPVS age increases. However, when the two groups were considered separately, this relationship was only present in the WS group[21].

ANOVA was therefore performed with group as a between-subjects factor, concreteness as a repeated measure and BPVS age as a covariate. As with the previous analysis, concrete words were recalled better than abstract words ($F(1,21)=13.62$, $p<.001$) and the main effect of group was nonsignificant ($F(1,21)=0.07$). The effect of BPVS age ($F(1,21)=0.31$) was non-significant, but the interaction between concreteness and BPVS age was significant ($F(1,21)=5.61$, $p<.05$) reflecting a reduced effect of concreteness in individuals with higher vocabulary ages (see regression above). With BPVS age covaried out, there was no evidence for an interaction between concreteness and group ($F(1,21)=0.12$), suggesting that the trend towards a reduced concreteness effect in WS could be an artefact of the higher vocabulary age of this group.

Table 9.2 Regression coefficients for vocabulary as a predictor of the size of the concreteness effect

	All	WS	TD
B	-0.00210	-0.00371	0.00100
SE B	0.001	0.001	0.002
β	-.525**	-.767**	.179

The effect of recall strategy

The final analyses concerned the effect of differences in strategy on the size of the concreteness effect. In other words, did the size of the concreteness effect depend on whether the participant attempted to recall items in the correct order or simply treated the task as a free recall task, and how did this affect the performance of the two groups? This was investigated by looking at the proportion of times each participant began recall with the first presented item. This value was smaller for the WS group ($M=.297$, $SD=.221$) than for controls ($M=.430$, $SD=.207$), but the difference was not significant ($t(22)=1.52$, $p=.14$). Moreover, there was no significant correlation between this value and the size of the concreteness effect, either when all participants were considered together or when groups were analysed separately.

9.3.4: Discussion

Overall results

The aim of Experiment 6 was to investigate the hypothesis that semantic memory is impaired in WS. The prediction was that if children with WS rely relatively heavily on phonological as opposed to semantic memory, they would show a reduced concreteness effect. Both children with WS and TD controls recalled more items from lists composed of concrete words than from those composed of abstract words, thus replicating the findings of Nation et al. (1999) who used the same procedure with TD children and poor comprehenders. However, while there was a trend towards a reduced concreteness effect in WS, the interaction between group and word-type was not significant.

These results are consistent with a recent study by Laing, Grant, Thomas, and Karmiloff-Smith (2002) who investigated the concreteness effect in serial recall in 14 children and adults with WS (age range 10;11 to 52;1 years; mean 21;7) and 14 TD controls matched on digit span. Participants were tested on lists of increasing length until they were unable to correctly recall any lists, and were then scored on the number of lists correctly recalled. Laing, Grant et al. found that, although there was a trend towards a reduced concreteness effect in WS, the interaction between concreteness and group was not significant ($F(1,26)=0.36$).

However, there were several problems with the procedure employed in Experiment 6, most notably the fact that many participants clearly found the task too difficult and resorted to recalling as many words as possible in any order, rather than trying to recall words in the correct serial order. There are also several other methodological issues that need to be addressed.

Speech rate

The first issue concerns speech rate. Individuals with faster speech rates tend to have superior serial recall to slower individuals (e.g. Baddeley et al., 1975; Roodenrys, Hulme, & Brown, 1993). A possible concern, therefore, is that there may be group differences in speech rate and this may interact with the concreteness effect. However, previous studies have failed to find any evidence for an interaction between speech rate and the concreteness effect (Nation et al., 1999; Walker & Hulme, 1999). Moreover, Laing, Grant et al. (2002) found that individuals with WS and controls had similar speech rates for both concrete and abstract words, suggesting that speech rate is not an important factor.

Item and order memory

The second issue concerns the distinction between item and order memory. Nation et al. (1999) measured performance in terms of the number of items recalled in the correct position, but in the current study, items were considered to be correct irrespective of their output position. This was because participants in Experiment 6 were unable to follow the instruction to say 'blank' when they could not remember an item so, once an omission error had been committed, it was often impossible to tell whether or not subsequent items had been recalled in the correct position. Moreover, a significant number of participants did not appear to even to attempt to recall in serial order, instead treating the task as a free recall task.

Nevertheless, it seems unlikely that difficulties recalling items in serial order would have influenced the overall findings of Experiment 6. Walker and Hulme (1999) found that concreteness primarily affected item rather than order errors in serial recall, and the concreteness effect was eliminated in a serial order recognition task, which can be considered as a test of order memory. Thus concreteness seems to influence item memory rather than order memory, and scoring on correct items, irrespective of order recall, should make little difference to the conclusions.

A slightly different concern is that concreteness may have a different influence on free as opposed to serial recall, so the effect size may depend on whether or not participants are actually trying to recall items in the correct order. For example, participants who were attempting to recall in serial order may have been more likely to give up when they knew they had made an error early in recall. However, although the TD group seemed to be trying harder to recall items in the correct order, correlational analysis failed to show any link between a participant's tendency to begin recall with the first presented item and the size of the concreteness effect.

Vocabulary knowledge

The third issue concerns the effect of vocabulary knowledge on the size of the concreteness effect. Because of the lack of evidence to show that there might be a link between these factors, and the desirability of equating overall performance on the tasks, groups were matched on recall of concrete words. This entailed that the TD group had much lower receptive vocabulary ages than the WS group. Subsequent regression analysis showed a significant decrease in the concreteness effect-size with increasing vocabulary knowledge, and when BPVS age was treated as a covariate, there was no evidence for an interaction between group and concreteness, suggesting that the trend towards a reduced concreteness effect in WS in the initial analysis was an artefact of the group differences in vocabulary knowledge.

Nevertheless, the results of the ANCOVA should be treated with caution. It could be that the WS group had higher vocabulary ages and smaller concreteness effects and these facts are unrelated, but when looking at both groups together it appears that there is a relationship whereby individuals with high vocabulary scores have small concreteness effects and vice versa. Because vocabulary age was strongly associated with group membership, covarying out vocabulary age would be likely to remove variation in the concreteness effect-size associated with group membership. However, the fact that there was a significant association between vocabulary and effect-size within the WS group suggests that the relationship is genuine.

The question then is why the size of the concreteness effect should decrease with increasing receptive vocabulary. One possibility is that the extent of semantic support takes longer to reach an asymptotic level for abstract words than for concrete words. Thus semantic support for concrete words may be at ceiling, but for abstract words the extent of support will increase with language level. A possible complication here is the fact that the relationship between BPVS age and the size of the concreteness effect was only found in WS. However, it is difficult to see why the concreteness effect might interact with BPVS age in WS but not in TD children. Instead, this finding could simply reflect the greater range of vocabulary scores in the WS group. If TD children with higher vocabulary scores had been tested, they may also have shown relatively small concreteness effects.

These observations nevertheless demonstrate the need to further investigate developmental changes in the concreteness effect and also indicate that any future studies of concreteness effects in WS should match groups on vocabulary knowledge. Given the relatively poor STM of children with WS relative to vocabulary-matched TD controls, a different paradigm such as the probed item recall task used in Experiments 3 and 4 may be more appropriate.

Comparison with poor comprehenders

A final point of interest is the comparison between children with WS and poor comprehenders, that is, children who have poor reading comprehension despite normal decoding skills (Nation & Snowling, 1997; Stothard & Hulme, 1995). These children appear similar to individuals with WS in the sense that they seem to have impaired semantic processing (reading comprehension) relative to their phonological skills (decoding). However, using the same task as that employed in Experiment 6, Nation et al. (1999) found that poor comprehenders showed an increased concreteness effect relative to controls. These authors interpreted this finding in terms of a semantic processing deficit that was more severe for abstract than for concrete words. In both Experiment 6 and the study by Laing, Grant et al. (2002), individuals with WS showed a trend towards a reduced concreteness effect, suggesting that the semantic processing deficit in WS is very different to that found that in poor comprehenders. Further confirmation of this view comes from semantic priming studies. Nation and Snowling (1999) reported that poor comprehenders showed reduced priming for taxonomically related words, but in a similar study, Tyler et al. (1997) found normal priming of taxonomically related items in WS.

9.4: Summary

Experiment 6 investigated the hypothesis that individuals with WS show impaired semantic memory by looking at the concreteness effect in serial recall. Consistent with a recent study by Laing, Grant et al. (2002), the size of the concreteness effect was not significantly different to that found in controls. However, there were a number of methodological concerns which future studies might address. In particular, the task proved too difficult so many participants did not attempt to recall items in the correct order, and there was an indication that the size of the concreteness effect might be confounded with vocabulary knowledge.

While the results of Experiment 6 are not entirely convincing, the failure to find evidence for a reduced influence of semantic knowledge on STM performance is consistent with the results of Experiments 3, 4, and 5 which showed no difference in the size of the word-frequency or lexicality effects in WS. These effects are assumed to reflect the influence of both phonological knowledge (i.e. redintegration) and semantic knowledge. The fact that word-frequency and lexicality effects were normal in WS implies that there is a normal contribution of both phonological and semantic knowledge to STM in WS.

---

[19] Due to time constraints it was not possible to test an MLD control group.

[20] The two lists of words were reported by Walker and Hulme (1999).

[21] One participant in the WS group had a relatively low vocabulary age and showed a particularly strong concreteness effect. Given the small numbers of participants, it is possible that this one child may have been responsible for the significant negative relationship between effect size and vocabulary. However, the relationship remained significant in the WS group even when this data point was removed.

Chapter 10: Nonword repetition

10.1: Overview

Experiment 7 investigated nonword repetition abilities in WS. Children with WS and TD and MLD controls matched on vocabulary were tested using the CNRep and three main issues were addressed. First, analysis of overall performance allowed further investigation of the claim that STM is a relative strength in WS. Second, comparison of the wordlikeness effect in the three groups provided a test of the hypothesis that redintegration processes are impaired in WS. Third, correlational analyses permitted investigation of the causal relationship between nonword repetition and vocabulary.

10.2: Overall performance

The first issue concerns the level of overall performance of individuals with WS on the CNRep. Two studies have previously reported relevant data. In the first of these, Grant et al. (1997) tested 17 individuals with WS who were between 8 and 35 years of age (mean age 18;7). Age-equivalent scores for the CNRep were below those for receptive vocabulary (BPVS), but at the same level as those for receptive grammar (TROG) and nonverbal reasoning (Ravens matrices).

In the second study, Laing et al. (2001) tested 15 individuals with WS who were between 9 and 28 years of age (mean age 15;1) and 15 TD five- to nine-year-olds (mean age 6;9). There was no significant difference between the two groups in terms of performance on the CNRep, the BPVS, or the matrices subtest of the BAS. This study therefore contradicts the findings of Grant et al. (1997) insofar as vocabulary and nonword repetition in WS were at equivalent levels.

Experiment 7 was similar to the study by Laing et al. (2001) in that controls were matched on vocabulary. The advantage of Experiment 7 was that it also included an MLD control group who were matched to the WS group on both vocabulary and chronological age. This is potentially important given recent arguments concerning the relationship between STM and the rate of vocabulary acquisition (cf. Jarrold et al., 2002).

10.3: The wordlikeness effect

The second issue concerns the wordlikeness effect. This effect provides evidence that memory for nonwords is influenced by knowledge of the phonological forms of words (see section 6.4.2), so impaired redintegration would predict a reduction in the size of the effect in WS.

The wordlikeness effect in WS was investigated by Grant et al. (1997) who individually matched 15 individuals with WS to TD five-year-olds on the basis of Ravens matrices scores[22] and found no significant difference in the size of the wordlikeness effect (J. Grant, personal communication). This finding fits well with those of Experiments 3, 4, and 5 which showed no evidence for a reduction in the influence of lexical knowledge on STM in WS. However, the Grant et al. (1997) study did not have its own control group. Instead, individuals with WS were compared with TD children tested by Gathercole (1995), and there were several differences between the procedures used in the two studies. For example, Grant et al. presented items on a tape, whereas Gathercole gave live presentations.

A further concern is that the WS group in the Grant et al. (1997) study were not only older than the controls but almost certainly had higher BPVS ages than the controls. Given that the wordlikeness effect increases with age (Gathercole, 1995), it remains possible that the WS group may have shown a reduced wordlikeness effect compared to controls matched more closely on age and / or vocabulary knowledge. Experiment 7 allowed comparison of the size of the wordlikeness effect in WS with that in controls matched on vocabulary and, in the case of the MLD group, on chronological age.

10.4: Correlations between vocabulary and nonword repetition

The third issue is the causal relationship between nonword repetition and vocabulary. Gathercole (1995) reported a longitudinal study of repetition of high- and low-wordlike nonword repetition and its relationship with vocabulary knowledge in four- and five-year-olds. Utilising a cross-lagged correlational design, she showed that low-wordlike repetition had a causal effect on vocabulary knowledge, and that vocabulary knowledge had a causal influence on repetition of high-wordlike nonwords.

Gathercole (1995) also reported that in four-year-olds, there was a significant correlation between repetition of low-wordlike nonwords and vocabulary even when repetition of high-wordlike nonwords and performance on the Ravens matrices were partialled out. In contrast, the correlation between high-wordlike repetition and vocabulary was non-significant when low-wordlike repetition and Ravens matrices performance were partialled out. By the age of five, neither partial correlation was significant. Gathercole interpreted this pattern of findings as evidence that phonological memory and vocabulary knowledge are closely associated at age four but the relationship disappears by the age of five. However, one could argue that, because repetition of the other kind of nonword was partialled out, these findings do not necessarily indicate a reduced association between vocabulary and low-wordlike repetition - they could equally reflect an increasing association between vocabulary knowledge and high-wordlike repetition.

Grant et al. (1997) conducted a similar analysis with the 17 individuals with WS that they tested, partialling out repetition of the other nonword type and performance on the TROG. The pattern of correlations was similar to that shown by TD four-year-olds in that the correlation with low-wordlike repetition was stronger than that with high-wordlike repetition. Grant et al. therefore concluded that, like TD four-year-olds, 'individuals with WS continue to rely heavily on the ability to retain phonological material in short-term memory to support word learning' (p. 95). However, this relies on the validity of Gathercole's (1995) interpretation of her results from TD four- and five-year-olds.

As well as concerns over the theoretical interpretation of the results in WS, there are also a number of methodological concerns. First, as noted above, there were several differences between the procedures used by Grant et al. (1997) and by Gathercole (1995) for administering the CNRep, so comparisons between the studies should be treated with caution. Second, Grant et al. used the Long Form of the BPVS whereas Gathercole et al. used the Short Form. Third, correlations between vocabulary knowledge and repetition of high- and low- wordlike nonwords were calculated based on subsets of high- and low-wordlike nonwords that were matched on phonological and syllabic length and phonological complexity, but different subsets were used for the WS and TD groups. Finally, in analysing partial correlations, Gathercole controlled for Ravens matrices scores but Grant et al. controlled for TROG scores instead. These difficulties were avoided in Experiment 7 by the inclusion of the two control groups. All three groups were tested following the same procedure, and results were analysed in the same way.

10.5: Experiment 7: Nonword repetition

10.5.1: Introduction

In Experiment 7, children with WS and TD and MLD controls were tested using the CNRep. The groups were matched in terms of BPVS age, so effect sizes should not be influenced by differences in vocabulary knowledge. Furthermore, the WS and MLD groups were also matched on chronological age, and so had similar language experience and similar rates of vocabulary acquisition.

10.5.2: Method

Participants were the same children who took part in Experiments 2 to 5 (see Table 10.1 for details). Testing took place within the same sessions as Experiments 2 and 3, although the order that tasks were completed varied depending on practical constraints. Children were played the tape provided in the CNRep. On this tape, a female native-English speaker told the participant that they would hear some 'funny made-up words' and would have to repeat them back. The child was then tested on two practice nonwords and 40 test nonwords presented in the same voice with a three second gap. The nonwords comprised 10 each of two-, three-, four- and five-syllable nonwords in a random order. If no response was given then the tape was paused. Nonwords were only replayed if the participant's response from the previous item interfered with the current item. Responses were scored strictly as either completely correct or wrong.

Table 10.1 Participant details for Experiment 7

	WS N=14		TD N=20		MLD N=13	
	M	SD	M	SD	M	SD
Chronological age (y;m)	13;5	2;5	7;5	0;9	12;6	1;9
BPVS–II age (y;m)	8;2	1;2	8;2	1;4	8;1	1;3
Ravens matrices age (y;m) †	7;6	1;8	8;9	1;2	8;1	2;0

† Based on N=13 for the WS group and N=12 for the MLD group

10.5.3: Results

Overall performance and the effect of nonword length

Figure 10.1 shows the mean proportion of items correctly repeated by the three groups as a function of the syllabic length of the nonwords. The WS group performed at a similar level to TD controls and better than MLD controls. Moreover, all three groups showed similar effects of nonword-length, with performance declining as length increased from two to four syllables, but improving for five-syllable nonwords. Results were subjected to two-way ANOVA with group as between-subjects factor and nonword length as a repeated measure. The effect of group was not in fact significant ($F(2,44)=1.00$), but the effect of nonword length was significant ($F(2.61,132)=44.29$, $p<0.001$, Greenhouse-Geisser correction applied). Post-hoc analysis revealed that all pairwise differences were significant, other than that between three- and five-syllable nonwords (Bonferroni correction applied). There was no group by nonword length interaction ($F(6,132)=0.427$).

Effect of wordlikeness

The effect of wordlikeness was investigated by looking at repetition performance for two subsets of 12 high- and 12 low-wordlike nonwords that were matched on phonological complexity, and phonological and syllabic length. These were the same subsets analysed by Grant et al. (1997; J. Grant, personal communication). Figure 10.2 shows the proportion of items correctly repeated as a function of wordlikeness. All three groups performed slightly better on the high-wordlike nonwords, but this effect was greater in the WS group than in the other two groups. ANOVA was performed with group as a between-subjects factor and wordlikeness as a repeated measure. The main effect of wordlikeness narrowly failed to reach significance ($F(1,44)=3.60$, $p=.064$), and there was no effect of group ($F(2,44)=1.44$) and no interaction between wordlikeness and group ($F(2,44)=0.824$).

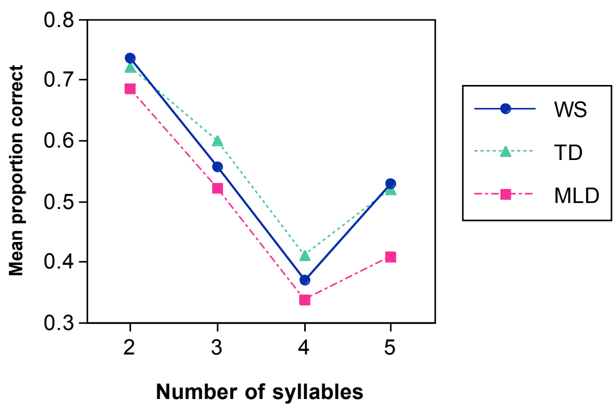

Figure 10.1 The effect of syllabic length on nonword repetition

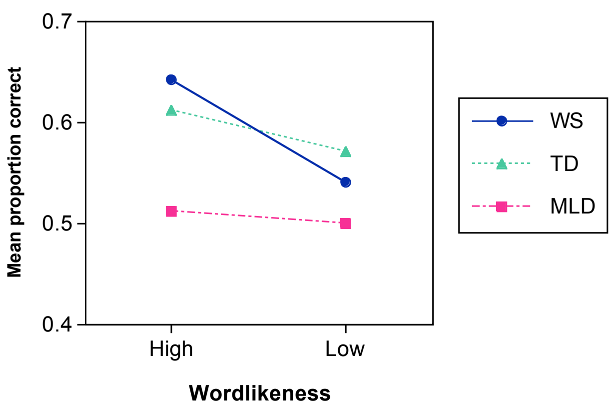

Figure 10.2 The effect of wordlikeness on nonword repetition

Correlations between nonword repetition and vocabulary

Table 10.2 shows the pattern of correlations between BPVS age, overall nonword repetition performance, and repetition of the 12 high- and 12 low-wordlike nonwords. Generally, correlational patterns were similar for the individual groups and for the combined participant group, although sometimes the correlations for the individual groups failed to reach significance. Correlational analyses were also performed with Ravens matrices performance as a partial correlate, but this had no effect on the pattern of correlations.

When all participants were considered as a single group, vocabulary age was significantly correlated with overall performance on the CNRep. The correlation was also significant for the TD group but failed to reach significance for either the WS group or the MLD group.

Vocabulary age was more strongly correlated with high- than with low-wordlike repetition. Using the procedure for comparing non-independent correlation coefficients (Williams, 1959; see Howell, 1992), this difference was significant in the WS group ($t(11)=3.32$, $p<.01$) but not in the TD ($t(17)=0.58$) or MLD ($t(10)=0.21$) groups. When all participants were grouped together, the difference approached significance ($t(44)=1.85$).

Partial correlations confirmed that vocabulary was more strongly correlated with high- than with low-wordlike repetition. When high-wordlike repetition was partialled out, repetition of low-wordlike nonwords was not significantly correlated with vocabulary age in any group. In contrast, the correlation between vocabulary and high-wordlike repetition for the overall group and for the WS group remained even once low-wordlike repetition had been partialled out.

Table 10.2 Correlations between vocabulary age and nonword repetition

Correlate	Partial correlates	All	WS	TD	MLD
Overall		.483**	.308	.662**	.397
Overall †	Ravens	.460**	.313	.642**	.372
High		.549**	.676**	.625**	.389
High	Low	.479**	.746**	.451	.278
High †	Low, Ravens	.496**	.666*	.454	.496
Low		.312*	-.138	.529*	.328
Low	High	.058	-.446	.244	.173
Low †	High, Ravens	.005	-.377	.208	-.117

*	Correlation significant p<.05 (two-tailed)
**	Correlation significant p<.01 (two-tailed)
High	Repetition of high-wordlike nonwords (proportion correct)
Low	Repetition of low-wordlike nonwords (proportion correct)
Ravens	Ravens matrices (raw score)
†	Based on N=13 for the WS group and N=12 for the MLD group

10.5.4: Discussion

In Experiment 7, children with WS were tested using the CNRep and were compared with MLD and TD control children matched on vocabulary. This allowed comparison between overall nonword repetition performance and vocabulary knowledge, investigation of the wordlikeness effect, and correlational analysis of the relationship between nonword repetition and vocabulary.

Overall performance

There were no significant differences between the three groups in terms of overall performance. This implies that nonword repetition in WS is at the level predicted by vocabulary knowledge. These results are consistent with those of Laing et al. (2001) who reported that individuals with WS and TD controls performed at comparable levels on the BPVS and on the CNRep. However, they contrast with the findings of Grant et al. (1997) who reported that age-equivalent scores for the CNRep were below those for the BPVS. It is unlikely that this discrepancy reflects differences in the ages of the participants, as the age-ranges in the Laing et al. study and the Grant et al. study were comparable. One possible explanation is that Grant et al.'s findings simply reflect differences in the standardization of the two tasks.

However, results of Experiment 7 contrast with those of Experiments 3, 4, and 5 (involving the same participants), in which the TD group outperformed the other two groups. In Experiment 3, it was shown that the group differences reflected superior order memory in the TD group. If, as was argued in section 6.3.4, serial order memory is involved in the representation of nonwords, one might have expected that the TD group would have performed relatively well on the CNRep. However, CNRep performance is also directly influenced by vocabulary knowledge and by phonological awareness, which may be in turn, be influenced by vocabulary knowledge. This could explain the difference between the pattern of results for the CNRep and purer tests of order memory such as the probed recall, digit span, and forward serial recall tasks employed in Experiments 3, 4, and 5.

Wordlikeness effect size

There was no evidence for a reduced wordlikeness effect in WS. If anything, the WS group showed a stronger wordlikeness effect than either control group. This finding is consistent with the normal wordlikeness effect reported by Grant et al. (1997), and disconfirms the suggestion made in Section 10.3 that a reduced effect in WS may have been masked by the superior vocabulary knowledge of the WS group.

However, both sets of results do conflict with the findings of Majerus et al. (2001) who reported a reduced effect of phonotactic frequency in the serial recall of nonwords in WS. As noted in section 6.4.4.4, Majerus et al.'s findings may simply be a consequence of the small size of their WS group or the fact that the groups were mismatched on vocabulary knowledge and overall performance.

Perhaps the most surprising result from Experiment 7 is that, while repetition was better for high- compared with low-wordlike nonwords, the main effect of wordlikeness was not actually significant. Given that the size of the wordlikeness effect increases between the ages of four and five (Gathercole, 1995), it is especially surprising that the TD group who had a mean age of 7;5 years showed little evidence of a wordlikeness effect. This perhaps reflects the complicated nature of the relationship between vocabulary and nonword repetition. Although increasing vocabulary knowledge should lead to an increased influence of vocabulary knowledge on repetition (benefiting the recall of high-wordlike nonwords), it may also lead to increasingly segmentalized phonological representations (cf. Metsala, 1997) which may have a particularly beneficial effect on low-wordlike repetition. At this particular developmental stage, it may be that the two effects cancel each other out.

Correlations with vocabulary knowledge

The third main finding concerned the correlations between vocabulary knowledge and repetition of high- and low-wordlike nonwords. Grant et al. (1997) reported that vocabulary knowledge in WS was more highly correlated with low- than with high-wordlike repetition. In contrast, in Experiment 7, vocabulary knowledge was more strongly correlated with high- than with low-wordlike repetition in all three groups.

It is difficult to explain why the results of Experiment 7 and the study by Grant et al. (1997) should be so different. Given that the mean chronological and vocabulary ages of the WS participants were lower in Experiment 8 than in the Grant et al. study, one might have expected that the association between vocabulary and high-wordlike repetition would be weaker in Experiment 7 than in the Grant et al. study. One explanation is that sizes of the WS groups in both the Grant et al. study (N=17) and Experiment 7 (N=14) were too small to allow correlational analysis, comparing unfavourably with the 70 TD children tested by Gathercole (1995). Small groups are likely to lead to noisy results, especially when extra degrees of freedom are lost due to partial correlates and when the objective is to compare the relative sizes of correlations.

A further difficulty with the correlational analysis is that the WS participants covered a wide range of ages (8 to 39 years in the Grant et al. study; 9 to 16 years in Experiment 7), whereas the TD children in this study and the Gathercole study were from a much narrower age range. Given the complex nature of the relationship between STM, vocabulary, and rate of vocabulary acquisition, it is extremely difficult to interpret correlational data from a group with such a large age range.

Summary

In Experiment 7, overall performance on the CNRep in WS was comparable with that of controls matched on vocabulary. As in the study by Grant et al. (1997), there was no significant difference between the WS group and controls in the size of the wordlikeness effect. However, Experiment 7 failed to replicate Grant et al.'s finding that vocabulary was more strongly correlated with low- rather than with high-wordlike repetition. Experiment 7 therefore provided further that verbal STM abilities in WS are not atypical.

10.6: Summary of STM studies

Experiments 3, 4, 5, 6, and 7 investigated STM abilities in WS. The first main finding was that the performance of children with WS on tasks with a strong serial order component (i.e. probed recall, digit span and forward serial recall of words) was inferior to that of TD children matched on vocabulary and at the same level as in MLD controls matched on age as well as vocabulary. Analysis of results from the probed recall tasks in Experiment 3 would appear to confirm that it is the serial order component of these tasks that leads to the relatively poor performance of both learning-disabled groups. This in turn suggests that serial order memory is not a relative strength in WS but may constrain the rate of vocabulary acquisition.

The second main finding was that, contrary to previous reports, there was no evidence for a reduced influence of lexical-semantic knowledge on STM performance in WS. In each case, previous findings in support of these hypotheses can be explained in terms of methodological concerns that were eliminated in the current set of experiments. As such, there was no evidence to support the hypothesis that redintegration processes are impaired in WS. Nor was there any evidence for impaired semantic memory. However, a study of free recall in WS conducted by Vicari, Brizzolara et al. (1996) does provide some evidence for semantic memory impairments in WS. This study is followed up in Part IV of the thesis.

---

[22] These participants were a subset of the 17 participants tested (see section 10.2 for details).

Part IV: Free recall

Chapter 11: Free recall

11.1: Overview

Experiments 8 and 9 reported in this chapter followed up a study of free recall in WS conducted by Vicari, Brizzolara et al. (1996). The findings of this study have been used as evidence for impaired semantic memory in WS.

In the free recall paradigm, participants are presented with a list of words that is longer than their span for serial recall and are required to recall as many words as possible in any order. Typically, a bow-shaped serial position curve is observed, with words from the early 'primacy' and late 'recency' portions of the list having a higher probability of recall than words from the midlist portion (e.g. Kirkpatrick, 1894; Murdock, 1962).

According to dual-store accounts of memory (e.g. Atkinson & Shiffrin, 1968; Waugh & Norman, 1965), the primacy effect reflects the output of a long-term store (LTS) which encodes semantic information, whereas the recency effect reflects the output of a short-term store (STS) which encodes phonological information [23]. Vicari et al. reported that, compared with TD controls matched on vocabulary, children with WS showed a normal recency effect but a reduced primacy effect. These authors therefore concluded that individuals with WS have intact phonological STM but impaired semantic LTM, and linked the reduced primacy effect in WS with other findings suggesting that semantic processing is impaired in WS (see chapter 2).

However, this conclusion relies on the validity of dual-store models, and these have been widely criticised in recent years. Even if dual-store models are correct, there are several alternative explanations for the reduced primacy effect in the Vicari, Brizzolara et al. (1996) study. This chapter therefore begins with a review of the theoretical interpretations of the serial position curve in free recall, and of studies of free recall in TD children. This is followed by a critique of the methodology of the Vicari et al. study, and two experiments that investigate alternative explanations for the reduced primacy effect.

11.2: Dual-store models of memory

As mentioned above, dual-store models of episodic memory and free recall (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965) postulate the existence of separate short- and long-term stores, and assume that the recency effect in free recall reflects the output of the STS whereas primacy reflects LTS output. This view is supported by the finding that the recency effect is abolished by a filled delay between presentation and recall, whereas this manipulation has little effect on recall of primacy items (e.g. Glanzer & Cunitz, 1966; Postman & Phillips, 1965). The idea is that when recall is delayed, these items are 'pushed out' of the STS by the intervening task. In contrast, primacy items are retrieved from the LTS, which is assumed to be relatively stable and resistant to interference. Consequently, the introduction of a delay has little effect on primacy.

The STS is assumed to encode information primarily in a phonological form, whereas the LTS encodes mainly semantic information (Atkinson & Shiffrin, 1968). Thus intrusions in recall are more likely to be phonologically similar to recency items than to primacy or midlist items (Craik, 1968; Shallice, 1975), while semantic factors such as word-frequency and imageability influence recall of primacy and midlist items but not recency items (Glanzer, 1972). Craik and Lockhart (1972) suggested that these effects arise because participants encode early items semantically as this results in a durable memory trace, but employ a more superficial phonological encoding strategy for later items. An alternative explanation is that phonological and semantic traces may be encoded for all items but the phonological trace deteriorates more rapidly than the semantic trace (Baddeley, 1986; cf. Conrad, 1967).

Dual-store models assume that access to the LTS is via the STS and that the longer an item spends in the STS, the more likely it is to be transferred to the LTS (Atkinson & Shiffrin, 1968). Primacy effects arise because earlier items can be rehearsed more often, enabling them to remain longer in the STS and increasing their chances of entering the LTS. This view is supported by studies using the overt rehearsal paradigm in which participants are required to rehearse aloud (e.g. Rundus, 1971). Participants tend to engage in so-called 'cumulative rehearsal' whereby as many of the previously presented items are rehearsed as possible. This means that early items are rehearsed more often than midlist items (e.g. compare A and D in Figure 11.1a). Indeed, when participants are required to rehearse only the most recent item and therefore rehearse each item an equal amount (compare A and D in Figure 11.1b), primacy effects are reduced or even disappear (Fischler, Rundus, & Atkinson, 1970; Glanzer & Meinzer, 1967). The importance of rehearsal is also demonstrated by studies showing that primacy effects are attenuated by various manipulations which prevent rehearsal, such as distracter tasks or rapid presentation rates (Bjork & Whitten, 1974; Cuvo, 1975; Glanzer & Cunitz, 1966; Richardson & Baddeley, 1975; Tzeng, 1973).

Despite the apparent success of dual-store models in accounting for free recall data, a number of findings have proved difficult to accommodate. In particular, there is evidence from various sources that recency effects do not reflect STS output. First, requiring participants to rehearse six-digit lists during free recall should interfere with the STS and remove recency effects but this is not the case (Baddeley & Hitch, 1977). Second, while recency is abolished by a period of distracter activity between presentation and recall (see above), it can be reinstated if a distracter task follows the presentation of each item (e.g. Bjork & Whitten, 1974; Tzeng, 1973). Finally, recency effects are observed in the recall of events separated by days, weeks or months (Baddeley & Hitch, 1977; Pinto & Baddeley, 1991) which clearly do not involve STM.

(a) **A** A A A A **B** B A B A **C** C A B A **D** D A B C **E** E A B C **F** F A B C **G** G A B C **H**

(b) **A** A A A A **B** B B B B **C** C C C C **D** D D D D **E** E E E E **F** F F F F **G** G G G G **H**

Figure 11.1 Examples of (a) cumulative rehearsal and (b) rehearsal of the most recently presented item. Large letters in bold represent items presented by the experimenter. Small letters in italics indicate rehearsals. See text for details.

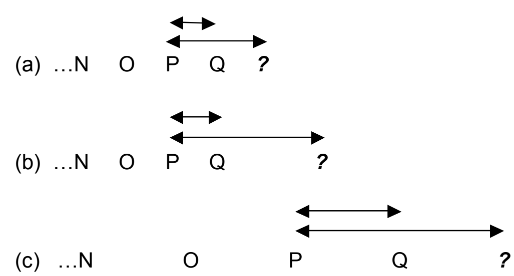

Figure 11.2 Illustration of the ratio rule. Letters represent the last few items of a list, question marks represent the start of recall, and arrows represent the duration of the retention intervals. When the items are close together but there is little delay before recall (a), the retention interval for Q is much shorter than for P, so Q is much more likely to be recalled than P. Increasing the retention interval (b) means that the retention interval for P and Q are relatively similar so Q has less advantage and recency is reduced. However, increasing the spacing of list items (c) restores the advantage for recall of Q and consequently the recency effect returns.

11.3: Single-store models of episodic memory

Bjork and Whitten (1974) argued that recency effects can be explained by assuming that the probability of recall of an item depends on the ratio of the inter-item interval and the time since presentation of that item. This 'ratio rule' is able to explain why recency can occur over any time-scale. It can also explain why recency is reduced by increasing the gap between the last item and the beginning of recall, but is enhanced by increasing the spacing of items (see Figure 11.2). Single store models of episodic memory (e.g. Brown et al., 2002; Crowder, 1993; Glenberg et al., 1980; Neath, 1993) have sought to capture this ratio rule by abandoning the distinction between long- and short-term stores, instead postulating that retrieval of an item depends (at least in part) on the similarity between the contexts at encoding and at retrieval. The suggestion is that the greater the delay between the encoding of an item and its retrieval, the more different the contexts will be and the less likely the item is to be retrieved.

Primacy effects can be accounted for by assuming that retrieval depends not only on the recency of the presentation of each item but also on the recency of each rehearsal of that item (see Tan & Ward, 2000). When participants employ a cumulative rehearsal strategy, items presented earlier in a list are not only rehearsed more frequently than midlist items but are also rehearsed later (compare A and D in Figure 11.1a). Thus, when recall is plotted as a function of last rehearsal position rather than order of presentation, large and extended recency effects are found, primacy effects are virtually abolished (e.g. Brodie, 1975; Modigliani & Hedges, 1987; Rundus, 1971; Tan & Ward, 2000), and the interaction between word-frequency and serial position is eliminated (Tan & Ward, 2000). Moreover, if rehearsal is prevented or participants rehearse only the most recent item (Figure 11.1b) then there is no re-ordering and consequently no primacy.

11.4: Free recall in typical development

A number of studies have investigated changes in free recall performance with age (e.g. Cole, Frankel, & Sharp, 1971; Cuvo, 1975; Hagen & Kingsley, 1968; Ornstein, Naus, & Prince-Stone, 1977). Unsurprisingly, overall performance improves with age. However, while performance on primacy increases steadily, there appears to be little improvement in the recall of recency items. For example, Ornstein et al. found that six-year-olds showed recency but not primacy, but by the age of nine, children showed both primacy and recency. Ornstein et al. provided evidence that these developmental changes in primacy recall reflect changes in rehearsal strategy. Under overt rehearsal conditions, six-year-old children spontaneously rehearsed only the most recently presented item and, as would be expected, showed no evidence of primacy. In contrast, nine-year-olds tended to engage in cumulative rehearsal and showed strong primacy effects. Moreover, when six-year-olds were given instructions that encouraged them to use cumulative rehearsal, they showed strong primacy effects.

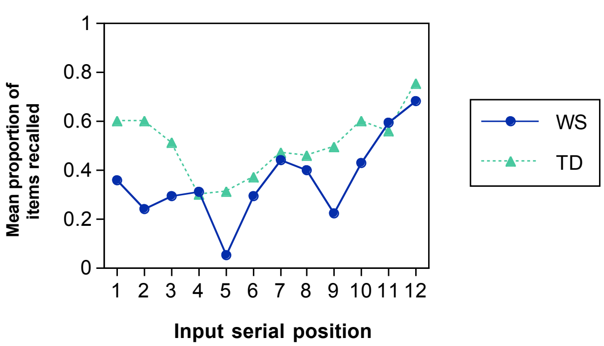

Figure 11.3 Serial position curves in a free recall task. Data replotted from Vicari, Brizzolara, et al. (1996)

11.5: Free recall in WS

With these theoretical considerations in mind, it is worth taking a closer look at the study of free recall in WS conducted by Vicari, Brizzolara et al. (1996). These authors tested 16 children with WS (mean age 10;2 years) and 16 TD controls (mean age 5;5 years) matched on receptive vocabulary using the PPVT-R. Participants heard lists of 12 high-frequency words and were then required to recall as many words as possible without time limit. Both groups showed recency effects but only the TD group showed a primacy effect (see Figure 11.3). Results were analysed by comparing performance on the first three items (primacy), the middle six items (midlist) and the final three items (recency). There was no significant group difference in terms of the number of items recalled from the midlist and recency positions, but the TD group recalled significantly more items from the primacy portion than the WS group.

Vicari, Brizzolara et al. (1996) interpreted these findings as evidence that individuals with WS have an intact phonological STS but an impaired semantic LTS. However, this conclusion relies on the validity of dual-store accounts of free recall which, as reviewed above, face a number of difficulties that are overcome to a large extent by models which assume a single episodic memory system. Moreover, there are several alternative explanations for the reduced primacy effect in WS.

First, the WS group recalled significantly fewer words than the controls overall, and were in fact worse at 10 out of the 12 serial positions. Given that the developmental studies reviewed above show an association between increasing overall performance and the extent of the primacy effect, a reduction in primacy could simply be a function of reduced overall performance.

Second, in most studies of free recall, including those with children, participants are tested on different words on each trial, or on the same words but in a different order. However, Vicari, Brizzolara et al. (1996) tested each participant on the same 12 words in the same order on each of five trials. In other words, they had one word list that was used for every trial for every participant (S. Vicari, personal communication) [24]. Serial order effects were therefore confounded with the saliency of the particular words. This might explain why the serial position curves in Figure 11.3 are so noisy (note for example the almost total failure of WS participants to recall a word from position 5, despite reasonable performance on positions 4 and 6).

Third, since other studies have shown that TD six- and seven-year-olds do not show primacy effects (Cuvo, 1975; Ornstein et al., 1977), the lack of primacy in the WS group studied by Vicari, Brizzolara et al. (1996) is entirely consistent with their verbal and nonverbal mental ages of around five years. The surprise is that the five-year-old TD controls did show a primacy effect, and this may reflect the unusual procedure adopted by Vicari et al.. Cole et al. (1971) tested free recall in TD children using the same words on each trial but in a different order and noted that participants showed a small primacy effect on the first trial but this reduced from trial to trial (see Helstrup, 1984; Huang 1986, for similar results in adults). However, if words are tested in the same order on each trial, primacy effects from the first trial may be maintained across subsequent trials because items recalled on the first trial probably have a greater chance of being recalled in subsequent trials. The absence of a primacy effect in WS under such conditions may therefore reflect a failure to benefit from previous trials with the same items.

Finally, both single- and dual-store models regard primacy as resulting (at least in part) from the use of cumulative rehearsal processes. If children with WS do not rehearse at all or only rehearse the most recently presented item, they would not be expected to show a primacy effect.

Experiments 8 and 9 addressed these issues. The two studies were conducted at the same time but are reported separately for ease of exposition. Experiment 8 looked at whether Vicari, Brizzolara et al.'s (1996) unusual procedure was responsible for the reduced primacy effect in WS. Experiment 9 investigated the effect of encouraging participants to engage in overt cumulative rehearsal.

11.6: Experiment 8: Free recall

11.6.1: Introduction

In Experiment 8, an attempt was made to replicate the findings of Vicari, Brizzolara et al. (1996) while eliminating some of the methodological difficulties. Free recall performance of children with WS was compared with that of TD children under three conditions. The control condition employed the standard procedure used in adult studies of free recall by presenting different words on each trial. In the same items condition, the same words were used on each trial but were presented in a different order, thus replicating the methodology of other studies with young children (e.g. Cole et al., 1971). In the same order condition, the same words were presented in the same order on each trial (i.e. the same procedure used by Vicari et al.). In all conditions, participants were required to repeat the items as soon as they were presented. This ensured that the participants remained 'on-task' throughout presentation.

Groups were matched in terms of their overall performance (i.e. total number of words correctly recalled) in the control condition so that potential differences in primacy effects could not be interpreted in terms of differences in overall performance. Comparisons between the serial position curves for the three conditions should show whether Vicari's results were a consequence of their unusual procedure. Moreover, the analysis of changes in performance over trials within a condition should indicate whether there is any impairment in long-term learning in WS as speculated above. Specifically, if children with WS fail to benefit from previous trials with the same items then they should improve at a slower rate than controls in both the same items and the same order conditions.

11.6.2: Method

11.6.2.1: Participants

Twelve of the 14 children with WS (7 boys) from the previous studies were tested, together with 24 TD controls (9 boys). The 12 oldest controls also took part in Experiment 6. Participant details are shown in Table 11.1. The WS group were older and had higher vocabulary ages than controls although there was no significant difference in performance on the Ravens matrices.

Table 11.1 Participant details for Experiment 8

	WS N=12		TD N=24	
	M	SD	M	SD
Chronological age (years; months)	13;10	2;6	6;1	0;6
BPVS age (years; months)	9;3	1;8	6;6	1;0
Ravens matrices score †	20.0	5.4	17.3	3.7

† Based on N=11 for the WS group. Raw scores reported because the performance of many of the TD children was below that for which age-equivalent scores were available

11.6.2.2: Stimuli

One hundred and forty-four one-syllable nouns were taken from the MRC Psycholinguistic database (Quinlan, 1992). All words had early age of acquisition ratings (four years or less), high concreteness ratings (300 or greater), high familiarity ratings (greater than 400) and high written frequencies (at least 30 words per million; Kucera & Francis, 1967). For each participant the words were randomly split into 12 lists of 12 words. Five lists were used for the control condition, one for the same items condition and one for the same order condition. The remaining five lists were used in the overt rehearsal condition in Experiment 9.

11.6.2.3: Procedure

Children with WS were tested at home in a quiet room with a parent present if required. Typically developing controls were tested in the library of their school. Each participant performed free recall of twelve-item lists in three blocks of five trials corresponding to the three conditions. The order of testing was counter-balanced such that an equal number of participants within each group were allocated to each of the six possible orders of testing. The three conditions were always completed within the same session. A filler task was completed between each recall block. This was either a computer game from another study, the BPVS or the Ravens matrices.

Participants were told that they would hear a list of words and should repeat each word aloud upon hearing it. At the end of the list, the experimenter would say 'OK, Go' and they should try and remember as many of the words as possible. Following Vicari, Brizzolara et al. (1996), the experimenter read each word aloud at a rate of approximately one word every two seconds. In the control condition, a different list was used on every trial. In the same items condition, the same list was used on each trial but words were presented in a different order. In the same order condition, the same list was presented in the same order on each trial.

11.6.3: Results

Effects of serial position

Figure 11.4 shows the serial position curves for the control, same items and same order conditions. For each condition, results were subjected to ANOVA with group and serial position as between- and within-subjects factors respectively. Following Vicari, Brizzolara et al. (1996), results were collapsed into three serial positions - primacy (mean of the first three items), midlist (mean of the middle six items), and recency (mean of the last three items) to avoid problems of ceiling and floor effects as much as possible. For each condition, the simple effect of group on recall of primacy items is also reported. Although it is common practice to only report simple effects when there is a significant interaction (Howell, 1992), in order to be as certain as possible that there is no effect of group on primacy recall, it is necessary to employ the least conservative statistical tests.

The upper panel of Figure 11.4 shows similar serial position curves for the control condition. Both groups demonstrated similar serial position effects, with large recency effects and little evidence of primacy. The effect of serial position was significant ($F(1.47,68)=184.27$, $p<.001$; Greenhouse Geisser correction applied) resulting from a significant advantage for recency items over primacy and midlist items (Bonferroni correction applied), but there was no significant effect of group ($F(1,34)=0.06$), and the interaction between group and serial position was non-significant ($F(2,68)=1.14$, $p=.325$). However, the data for the recall of primacy items in the WS group suffered from both positive kurtosis and skewness. No transformation could be found which did not result in kurtosis and skewness in at least one cell, but an arcsine transformation enabled inspection of the effect of group on primacy recall, which proved to be non-significant ($t(34)=-0.94$).

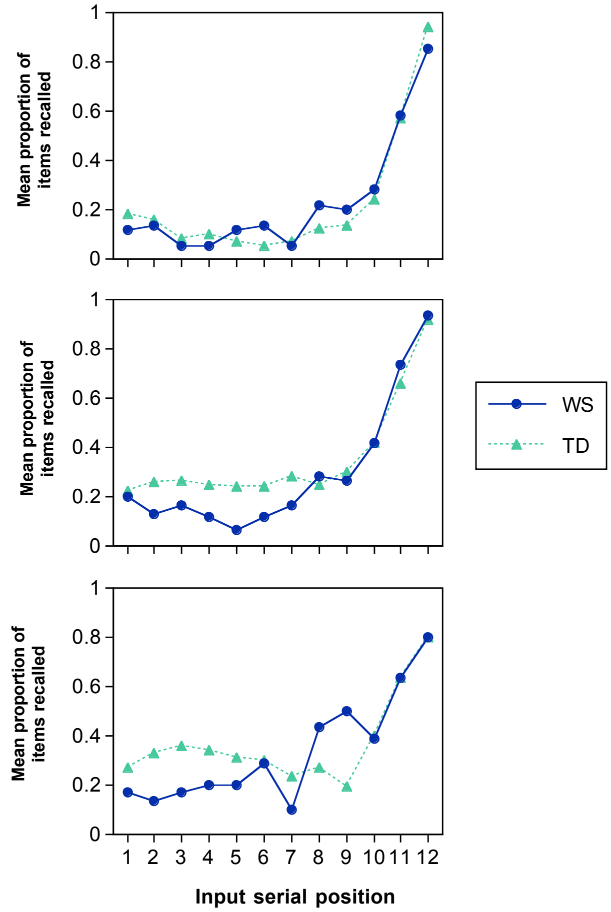

Figure 11.4 Serial position curves for the control (upper panel), same items (middle panel), and same order (lower panel) conditions of the free recall task

For the same items condition, the middle panel of Figure 11.4 shows similar recency effects in both groups but the TD group performed slightly better on the primacy and midlist items. As in the control condition, the effect of serial position was significant ($F(1.58,68)=140.95$, $p<.001$) resulting from a significant advantage for recency items over primacy and midlist items (Bonferroni correction), and there was no significant effect of group ($F(1,34)=1.76$). The interaction between group and serial position was non-significant ($F(2,68)=2.31$, $p=.107$; sphericity assumed), as was the effect of group on primacy recall ($t(34)=-1.40$).

For the same order condition, the lower panel of Figure 11.4 shows much noisier serial position curves, especially for the WS group. The TD group performed relatively well on the first seven items, whereas the WS group performed relatively well on the eighth and ninth items, and the two groups performed comparably on the three recency items. As in the previous conditions, the effect of serial position was significant ($F(2,68)=33.98$, $p<.001$) resulting from a significant advantage for recency items over primacy and midlist items (Bonferroni correction), and the effect of group ($F(1,34)=1.71$) and the interaction between group and serial position ($F(2,68)=1.81$, $p=.171$) were both non-significant. As in the control condition, floor effects meant that the data for recall of primacy items by the WS group suffered from positive skewness but this problem was removed using an arcsine transformation. The ANOVA on the transformed data, however, produced similar results, with a significant effect of position ($F(2,68)=31.59$, $p<.001$) and no effect of group ($F(1,34)=2.59$). The interaction between group and serial position approached significance ($F(2,68)=2.62$, $p=.080$), and the WS group performed significantly worse than controls on primacy items ($t(34)=-2.52$, $p<.05$).

Learning effects

Figure 11.5 shows overall performance in the three conditions. This confirms that the groups were closely matched on overall performance in the control condition. However, the TD group performed better in the same items and same order conditions. Results were analysed by way of an ANOVA with group as a between-subjects factor and condition as a repeated measure. There was no main effect of group (F(1,34)=1.43), but there was a significant effect of condition (F(2,68)=26.81, p<.001) resulting from inferior performance in the control condition compared with the other two conditions (Bonferroni correction applied). However, the interaction between group and condition was non-significant (F(2,68)=1.69, p=.192).

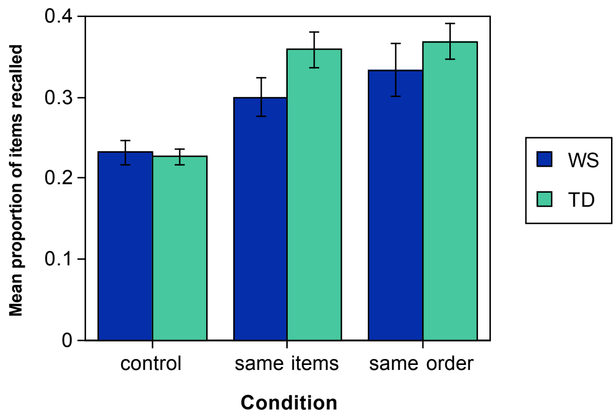

Figure 11.5 Overall performance for the control, same items, and same order conditions of the free recall task. Error bars show standard errors.

Figure 11.6 shows overall performance in the control, same items and same order conditions respectively for each of the five trials within a block. Because of the low numbers involved, results were analysed by comparing performance on the first and second trials with performance on the fourth and fifth trials. Results were then subjected to two-way ANOVA with group as a between-subjects factor and trial (1 & 2 vs 4 & 5) as a repeated measure.

The upper panel of Figure 11.6 shows that, in the control condition, the WS group recalled a fairly constant number of items from trial to trial, whereas the TD group performed relatively well on the first trial and relatively poorly on the last trial. Unsurprisingly, given that groups were deliberately matched on overall performance in this condition, there was no main effect of group (F(1,34)=0.25), but there was a significant effect of trial (F(1,34)=4.66, p<.05) and the interaction between group and trial approached significance (F(1,34)=3.70, p=.063).

The middle panel of Figure 11.6 shows the effect of trial for the same items condition. Both groups improved steadily from trial to trial, and although the TD group performed better overall, the rate of improvement was comparable in the two groups. These observations were confirmed by ANOVA. There was a significant effect of trial (F(1,34)=18.44, p<.001), an effect of group which approached significance (F(1,34)=3.89, p=.057), but no interaction between trial and group (F(1,34)=0.61).

The lower panel of Figure 11.6 shows the effect of trial in the same order condition. Results were similar to the same items condition although learning effects were slightly greater. There was a significant effect of trial (F(1,34)=57.41, p<.001), but the effect of group (F(1,34)=0.70) and the interaction between group and trial (F(1,34)=0.09) were non-significant.

The increased learning effect in the same order condition relative to the same items condition was confirmed by a three-way ANOVA with group as a between-subjects factor and trial and condition (same items vs same order) as repeated measures. There were no significant effects other than an interaction between trial and task (F(1,34)=8.08, p<.01) which reflected a significant effect of task for the fourth and fifth trials (t(35)=2.35, p<.05) but not for the first and second (t(35)=-0.67).

Performance on the first trial of each block

Performance on the first trial of each block should have been the same across conditions as there would have been no learning or interference effects from previous trials. This was confirmed by a two-way ANOVA with group as a between-subjects factor and condition (control vs same items vs same order) as a repeated measure. This showed no effect of condition (F(2,68)=0.72) and no task by group interaction (F(2,68)=0.02). There was, however, an effect of group that narrowly failed to reach significance (F(1,34)=4.14, p=.050).

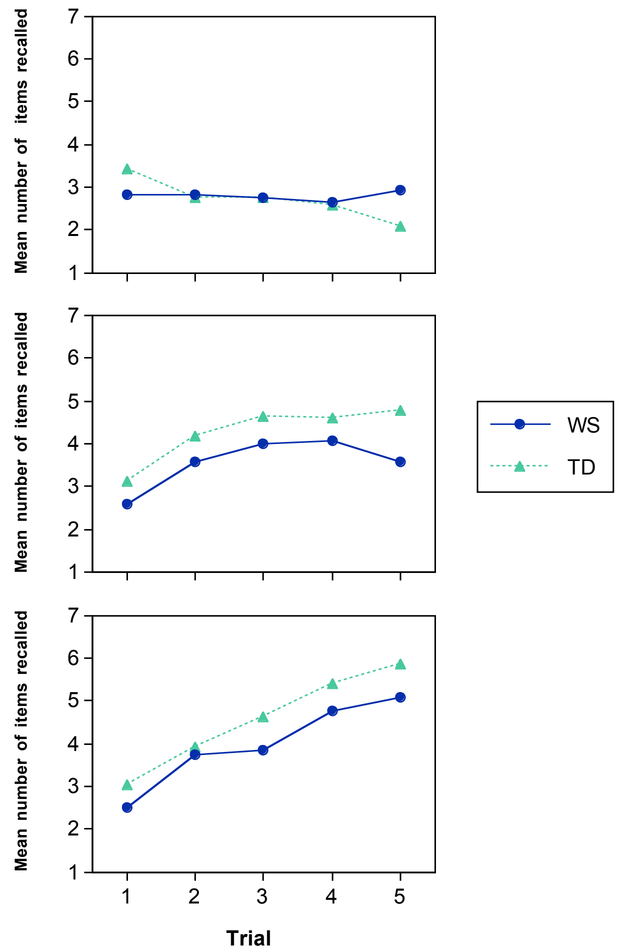

Figure 11.6 Learning effects for the control (upper panel), same items (middle panel), and same order (lower panel) conditions of the free recall task

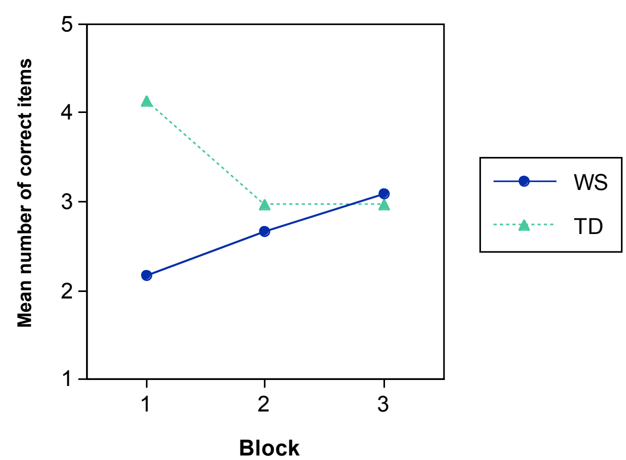

Figure 11.7 Overall performance on the first trial of each block of the free recall task

Because the first trial of each condition was equivalent, it was possible to look at performance on the first trial of each block as a function of the order of testing (i.e. disregarding the condition). Figure 11.7 shows that although the groups performed comparably on the first trial of the second and third blocks, the TD group were much better on the first trial of the first block. This was confirmed by a two-way ANOVA with group as a between-subjects factor and block as a repeated measure. There was no main effect of block (F(2,68)=0.77) but the TD group performed better overall (F(1,34)=7.15, p<.05) and there was a significant interaction between group and block (F(2,68)=8.21, p<.001). Two-tailed t-tests showed a significant effect of group for the first block (t(34)=5.36, p<.001) but not for the second (t(34)=0.83) or third (t(34)=0.25) blocks. One-way ANOVAs were performed to look at the effect of block for the two groups separately. The TD group showed a significant effect of block (F(1.59,46)=8.60, p<.01; Greenhouse-Geisser correction applied), performing better on the first block than on the second and third blocks (Bonferroni correction applied). However, the improvement of the WS group across blocks was non-significant (F(2,22)=2.50, p=.105).

11.6.4: Discussion

In Experiment 8, children with WS and TD controls performed free recall under three conditions. In the same order condition, the same words were presented in the same order on each trial. This replicated the procedure used by Vicari, Brizzolara et al. (1996). Their findings were corroborated to a certain extent in that the interaction between group and serial position was almost significant, with the TD group recalling more primacy items than the WS group. However, Vicari, Brizzolara et al. reported that TD controls showed an advantage for primacy over midlist items while WS children showed no such difference. In contrast, in the same order condition of Experiment 8, the TD controls showed no difference between recall of primacy and midlist items whereas the WS group showed a disadvantage for primacy items.

The reason for this discrepancy is unclear. One difference between the two studies is that, whereas Vicari, Brizzolara et al. (1996) tested all participants on the same word list, thus confounding serial position with word identity, in the same order condition each participant was given a different list. Vicari et al. may have chosen particularly salient words for the primacy items leading to relatively good performance of both groups on these items. A further difference between the studies is that in Experiment 8, participants were required to repeat each item aloud after it had been presented, whereas Vicari et al. had a silent interval between items. Typically developing children in the Vicari et al. study may have used this silent interval to rehearse, leading to their primacy effect.

The relative deficit in recall of primacy items in WS reported here and in the Vicari, Brizzolara et al. (1996) study does, however, appear to be peculiar to the particular recall paradigm whereby the same word list is used on each trial. In the control condition, in which different words were presented on each trial, both groups showed substantial recency effects but neither group showed a primacy effect. In the same items condition, the same words were presented on each trial but in a different order each time. Again there were strong effects of recency but little evidence of a primacy effect in either group. Although the TD group recalled more words from both primacy and midlist positions than did WS children, this difference was not significant. Moreover, it is possible that this trend reflects the fact that, despite being matched on overall performance in the control condition, overall performance in the same items condition was slightly higher in the TD group than in the WS group.

Experiment 8 also allowed the investigation of long-term learning in WS. It was suggested that impaired learning of words from previous trials could have contributed to the reduced primacy in WS observed by Vicari, Brizzolara et al. (1996). Since groups were matched on performance on the control condition, this would predict relatively poor overall performance in the WS group on the same order and same items conditions. While the group differences were not significant, the trend was towards impaired performance in the WS group. However, in both the same order and same items conditions, both groups showed comparable increases in overall performance with trials, suggesting that the children with WS did benefit as much as controls from repeated testing of the same material.

The trend towards worse overall performance in WS in these two conditions appears to reflect an advantage for controls on the first trial that then continued throughout subsequent trials. This arose because, although the groups were matched on overall performance on the control condition (averaged across five trials), the TD group performed better than the WS group on the first trial of the control condition. Since performance on the first trial should theoretically be the same for each condition, this gave the TD group a slight advantage on the first trial of all conditions.

The decline in performance with trials of the TD group in the control condition could reflect proactive interference, whereby items from previous lists interfere with recall of the current list. In contrast, the WS group showed no such effect. Although the interaction between group and trial narrowly failed to reach significance, this finding suggests that there is a reduced effect of proactive interference in WS. Further support for this view comes from analysis of performance on the first trial of each block as a function of order of testing (rather than condition). The TD group performed significantly better on the first trial of the first block than on the first trial of the second and third blocks, suggesting that they suffered from proactive interference from previous blocks[25]. In contrast, the WS group actually improved slightly from block to block. Unfortunately these effects are confounded with effects of practice and fatigue, and the tasks completed between blocks were not always the same. Consequently the evidence for group differences in proactive interference is suggestive at best.

The main conclusion from Experiment 8, however, was that when standard free recall procedures were employed (i.e. when the same word list is not repeatedly tested), neither children with WS nor controls showed evidence of primacy. These findings are consistent with previous studies using similar procedures which show that young TD children aged six and seven show recency but not primacy (Cole et al., 1971; Ornstein et al., 1977). Since developmental changes in primacy are related to changes in rehearsal strategy, it seems likely that both the WS and TD groups in Experiment 1 failed to show primacy effects because neither group employed a cumulative rehearsal strategy. However, this does not eliminate the proposal by Vicari, Brizzolara et al. (1996) that children with WS have an impaired LTS. Assuming for argument's sake that dual-store models are correct, then primacy effects depend on both the use of cumulative rehearsal and the integrity of LTS processes. The absence of primacy could therefore reflect immature rehearsal strategy or deficits in LTS or both. To eliminate the possibility that the LTS is impaired in WS it is necessary to show that children with WS do show primacy effects when they use cumulative rehearsal. If the LTS is impaired in WS then children with WS should show no primacy effect even when they employ such a strategy.

11.7: Experiment 9: Free recall with overt rehearsal

11.7.1: Introduction

In Experiment 9, children with WS and TD controls were tested using the overt rehearsal paradigm introduced by Rundus (1971) and subsequently used with children by Ornstein et al. (1977). As reviewed earlier, Ornstein et al. reported that six-year-olds spontaneously rehearsed only the most recently presented item and consequently showed no primacy effects in recall. However, when instructed to rehearse as many different items as possible, they did show significant primacy effects. Therefore, in Experiment 9, participants were explicitly instructed to rehearse as many different items as possible. Performance on this task was then compared with performance on the control task from Experiment 8.

11.7.2: Method

11.7.2.1: Participants

The participants were 10 of the children with WS (5 boys) and 11 of the TD controls (4 boys) who were tested in Experiment 8. The remaining participants from Experiment 8 were unable to comply with the task demands. Participant details are shown in Table 11.2. The children in the TD group had performed slightly better than those in the WS group on the control condition of Experiment 8, but this difference was not significant (t(19)=0.46). The WS group were older than the TD group and had higher vocabulary ages but there was no significant group difference in Ravens matrices scores.

Table 11.2 Participant details for Experiment 9

	WS N=10		TD N=11	
	M	SD	M	SD
Chronological age (years; months)	13;3	2;2	6;6	0;3
BPVS age (years; months)	9;3	1;10	6;10	1;1
Ravens matrices age (years; months) †	7;5	2;0	7;3	0;11

† Based on N=9 for the WS group

11.7.2.2: Materials

Participants were tested using the remaining 60 words (five lists of 12) from the pool of 144 words that they had not been tested on in Experiment 8.

11.7.2.3: Procedure

The experimental setting was the same as in Experiment 8. In most cases, Experiment 9 took place later within the same session as Experiment 8 with a filler task interpolated between the two experiments. However, for some of the TD children the constraints of the school timetable meant that this was not possible and testing on Experiment 9 occurred during a later session. Before testing, participants were told that it would help them to remember more words if they could practice remembering as many words as they could after the experimenter read each word aloud. An example was given.

'If I said 'apple' then you would say 'apple', then if I said 'banana' you would say 'banana, apple', then if I said 'orange', you might say 'orange, apple, banana', then if I said 'pear', you would say 'pear, apple, orange, banana''.

It had been intended to follow the methodology of Ornstein et al. (1977) and present words at a rate of one every five seconds. However, pilot studies showed that the participants were unable to rehearse overtly at this rate. The experimenter therefore gave up to 10 seconds for rehearsal but continued if the child indicated that they could not recall any more items. Five lists were presented, each containing different words, as in the control condition of Experiment 8. The experiment was recorded on audio tape to allow later analysis of rehearsal strategies but unfortunately the recording did not work for several of the participants.

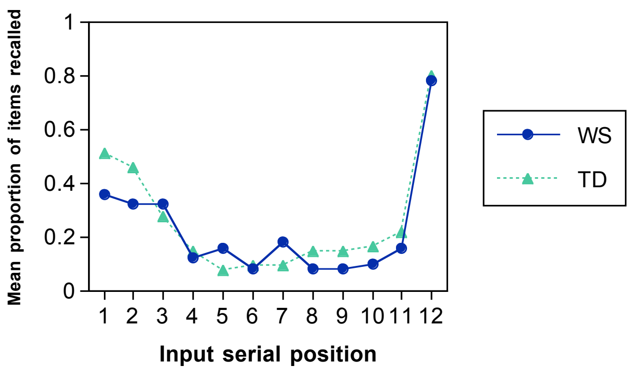

Figure 11.8 Serial position curves for the overt rehearsal condition of the free recall task

11.7.3: Results

Figure 11.8 shows the serial position curves for the overt rehearsal condition for the WS and TD groups. Both groups showed primacy and recency effects, although the primacy effect was slightly larger in the control group than in the WS group. As in Experiment 8, these results were subjected to a mixed-design ANOVA with group as between-subjects factor and serial position (primacy vs midlist vs recency) as repeated measures. There was a significant effect of serial position ($F(1.37,38)=27.23$, $p<.001$; Greenhouse-Geisser correction applied) resulting from superior recall of items from primacy and recency positions compared with midlist items. The effect of Group was non-significant ($F(1,19)=2.01$), as was the interaction between group and serial position ($F(2,38)=0.43$; sphericity assumed) and the effect of group on recall of primacy items ($t(19)=-1.08$).

The performance of the WS group in the control and overt rehearsal conditions was compared by way of a two-way ANOVA with condition and serial position as repeated measures. There was a significant interaction between serial position and condition ($F(1.63,18)=37.49$, $p<.001$). This occurred because in the overt rehearsal condition, recall of primacy items was better ($t(9)=5.65$, $p<.001$) and recall of recency items was worse ($t(9)=-7.00$, $p<.001$) than in the control condition. There was no effect of condition on recall of midlist items ($t(9)=-0.78$). Confirmation of primacy effects in the overt rehearsal condition in WS was found in the significant advantage for primacy items over midlist items ($t(9)=4.32$, $p<.01$), an effect absent in the control condition ($t(11)=-0.68$).

Finally, the effect of rehearsal strategy on overall performance (total number of items recalled) for the two groups was investigated by means of a two-way ANOVA with group as a between-subjects factor and condition (control vs overt rehearsal) as a repeated measure. There was no effect of condition ($F(1,19)=0.03$) or group ($F(1,19)=0.88$) and no group by condition interaction ($F(1,19)=0.55$).

11.7.4: Discussion

In Experiment 9, children with WS and controls performed free recall using the overt rehearsal paradigm. Participants were encouraged to use a sophisticated cumulative rehearsal strategy whereby they rehearsed a number of different items between presentations. This is the strategy typically used by adults and older children and is associated with significant primacy effects in these groups (Ornstein et al., 1977; Rundus, 1971). It was therefore expected that TD controls would also show a primacy effect under these conditions. However, if, as Vicari, Brizzolara et al. (1996) argued, the LTS is impaired in WS, the WS group should continue to show no evidence of primacy. In fact, both groups showed significant primacy effects under overt rehearsal conditions.

Although there was no significant interaction between group and serial position in the overt rehearsal condition, Figure 11.8 shows a slight advantage for the TD group in recall of primacy items. It is possible that testing of more participants would reveal a significant deficit in primacy recall in WS that could then be taken as evidence for a relatively impaired LTS. However, the control group also performed slightly better on recency items so this trend may reflect a difference in overall performance. Moreover, although all the participants included in Experiment 9 engaged in cumulative rehearsal, there may have been group differences in the number of rehearsals completed, and this could have affected the extent of the primacy effect. Unfortunately, due to technical problems, tape recordings of rehearsals were not available for many of the participants so it is not posible to objectively verify this suggestion.

A final interesting point concerns the fact that engaging in overt rehearsal did not actually improve overall performance. Instead there was a trade-off with more items being recalled from primacy positions and less from recency. Similar findings were reported by Ornstein et al. (1977) who found that instructing six-year-olds to rehearse multiple items had little effect on overall performance, and the increase in primacy was mirrored by a reduction in recency. Such findings fit well with single-store models, according to which, rehearsal effectively moves primacy items into recency positions (Tan & Ward, 2000), thereby making recall of those recency items less probable. In contrast, dual-store models postulate that rehearsal allows primacy items to enter the LTS (e.g. Atkinson & Shiffrin, 1968), and since items can be recalled from either the LTS or the STS, one might expect rehearsal to result in improved overall performance.

11.8: Summary

Vicari, Brizzolara et al. (1996) reported that, compared with TD controls, children with WS showed a reduced primacy effect in free recall. These authors interpreted this finding in terms of impaired semantic memory in WS. However, there are several methodological concerns with this study, and alternative interpretations of its results. The methodological concerns were addressed in Experiment 8. The main finding was that, when different items were used on each trial, and when the same words were used but in a different order each time, children with WS and TD controls matched on overall performance showed strong effects of recency but no primacy effect. In Experiment 9, children with WS and TD controls were instructed to engage in overt cumulative rehearsal, and showed strong and significant primacy effects. If the LTS was impaired in WS, then rehearsal strategy should not have made any difference.

The final important issue concerns the question of whether or not semantic memory is impaired in WS. Vicari, Brizzolara et al. (1996) did not explicitly manipulate semantic factors. Rather, they noted a reduced primacy effect in WS and concluded that this reflected impared semantic memory. Experiments 8 and 9 appear to show that this explanation of Vicari et al.'s findings is incorrect. However, semantic factors were not directly manipulated in these experiments either, so it is premature to conclude that semantic memory is unimpaired in WS. Nevertheless, given that none of the studies in Part II or Part III of this thesis have found any evidence for impaired sematnic memory in WS, there is little reason to suspect that manipulating semantic factors in a free recall task would lead to significant group differences.

---

[23] Atkinson and Shiffrin used the terms STS and LTS to distinguish the underlying memory stores from the nature of the task (e.g. a STM task involves memorizing material over a short time-scale).

[24] A literature search failed to find any other studies that have used this procedure.

[25] This finding suggests that the deterioration in performance of the TD group across trials in the control condition may only be found in participants who performed the control condition first. Consequently, a significant difference in proactive interference might have been found if all participants had been tested on the control condition first.

Part V: Discussion

Chapter 12: Discussion

12.1: Overview of the thesis

The aim of this thesis was to investigate some of the factors that may contribute to the linguistic profile in WS. In chapter 2 it was noted that, although language in WS is often described as being 'relatively preserved', there is in fact little evidence to suggest that any linguistic abilities other than receptive vocabulary are above the level predicted by overall mental age. Moreover, there is a consistent suggestion that individuals with WS are better at learning the phonological forms of words than they are at learning their meanings. Such findings have led to the suggestion that there is some form of imbalance in WS between relatively preserved phonological processing and relatively impaired semantic processing (cf. Thomas & Karmiloff-Smith, 2002).

The studies in this thesis addressed two issues. The first concerned the underlying cause of the relative strength in receptive vocabulary. Part II of the thesis explored the hypothesis that temporal processing mechanisms are relatively preserved in WS, while Part III investigated the hypothesis that serial order mechanisms are relatively preserved. The second issue concerned evidence from previous studies suggesting that semantic memory is impaired in WS. As well as investigating serial order memory, the experiments in Part III also investigated claims from previous studies that the influence of lexical-semantic knowledge on STM performance is reduced in WS. Part IV, meanwhile, investigated the claim that individuals with WS showed a reduced primacy effect in free recall and therefore have impaired semantic LTM. The main results from these experiments are summarised in Table 12.1 and are discussed below.

12.2: Why is vocabulary a relative strength in WS?

12.2.1: Determining relative strengths

The first issues concerns the cause of the relative strengths in vocabulary in WS, particularly in older children and adults. Determining causal relationships is problematic, particularly in a syndrome like WS in which it is difficult to separate specific causal relationships from the effects of general mental retardation. Although it was possible to conduct some preliminary correlational analyses, given the homogenous nature of the overall group of participants, and the small size of the individuals groups, the conclusions from these analyses were necessarily tentative. Consequently, the main aim of the studies in Parts II and III was simply to determine whether temporal discrimination and serial order memory are relative strengths in WS.

For these analyses, the choice of control group was critical. In all of the relevant experiments, children with WS were compared with TD controls matched on vocabulary age, and with MLD controls matched on vocabulary and chronological age. The logic, put forward in Part II, was that language abilities are likely to be determined by a number of different factors. If, for example, language is a relative strength in WS because of relatively preserved temporal processing abilities, then other causal factors, which (according to this hypothesis) are at a lower developmental level than temporal processing, will pull language abilities down below the level predicted by temporal processing abilities. The hypothesis therefore predicted that children with WS should actually perform better than language-matched controls on tests of temporal processing. A similar argument holds for the hypotheses that linguistic strengths are a consequence of relatively preserved serial order memory.

12.2.2: Is temporal discrimination a strength in WS?

The results of Experiment 1 were inconclusive. Children with WS performed better than MLD controls on a temporal bisection task, but worse on a temporal generalisation task, although neither difference was significant. In Experiment 2, children with WS, and TD and MLD controls performed comparably on three different temporal discrimination tasks. It was argued that, because vocabulary is a relative strength in WS, these findings imply that temporal discrimination abilities are also a relative strength in WS (see section 5.3). However, the children with WS were not better than either control group so it is difficult to argue that language abilities are a relative strength in WS *because* of good temporal processing abilities. Moreover, there was little evidence to suggest that a causal relationship exists between temporal discrimination mechanisms and vocabulary acquisition. First of all, there were no significant correlations between temporal discrimination abilities and either vocabulary age or vocabulary IQ. Second, it was argued that if temporal processing has a causal influence on vocabulary acquisition, temporal discrimination performance should predict rate of vocabulary acquisition. However, despite their faster rate of vocabulary acquisition, the TD group performed no better than the MLD group. In summary, while temporal discrimination may be a relative strength in WS, there was no evidence that it is linked to vocabulary acquisition.

Table 12.1 Summary of results

Exp	Task	Test variable	Cntrls	Matching	Res
1	Temporal	Steepness of	MLD	BPVS	=
	bisection	bisection curve	TD5[a]	N/A	=
	Temporal	Steepness of	MLD	BPVS	=
	generalisation	generalisation	TD5[a]	N/A	↓
		curve			
2	Temporal	Proportion correct	TD	BPVS	=
	discrimination		MLD	BPVS&age	=
3	Digit span	Span and score	TD	BPVS	↓
			MLD	BPVS&age	=
	Probed position	Proportion correct	TD	BPVS	↓
	recall		MLD	BPVS&age	=
	Probed item	Proportion correct	TD	BPVS	↓
	recall		MLD	BPVS&age	=
		Estimated order	TD	BPVS	↓
		memory	MLD	BPVS&age	=
		Estimated item	TD	BPVS	=
		memory	MLD	BPVS&age	=
		Lexicality effect	TD	BPVS	=
		size	MLD	BPVS&age	=
		Lexicalization	TD	BPVS	=
		errors	MLD	BPVS&age	=
4	Probed item	Word-frequency	TD	BPVS	=
	recall	effect size	MLD	BPVS&age	=
5	Forward serial	Span & number of	TD	BPVS	↓
	recall	lists correct	MLD	BPVS&age	=
		Word-frequency	TD	BPVS	=
		effect size	MLD	BPVS&age	=
6	Forward serial	Concreteness	TD	BPVS	=
	recall	effect size			
7	Nonword	Overall	TD	BPVS	=
	repetition	performance	MLD	BPVS&age	=
		Wordlikeness	TD	BPVS	=
		effect size	MLD	BPVS&age	=
8	Free recall	Primacy effect size	TD	Overall	=
				perform [c]	
		Extent of learning	TD	Overall	=
		effect		perform [c]	
9	Free recall with	Primacy effect size	TD	Overall	=
	overt rehearsal			perform [c]	

Res Result of comparison with WS group
= no group difference
↓ WS group significantly worse than controls
[a] TD five-year-olds tested by McCormack et al. (1999)
[b] BPVS taken as a covariate
[c] Participants matched on overall performance in the control condition

12.2.3: Is serial order memory a strength?

Overall performance on the STM tasks showed a relatively consistent pattern. On digit span (Experiment 3), probed recall (Experiments 3 and 4), and forward serial recall of words (Experiment 5), the WS group performed worse than the TD group but comparably with the MLD group[26]. The one exception to this pattern of performance was the nonword repetition test (CNRep; Experiment 9), on which the three groups performed comparably[27].

In Experiment 3 it was possible to separate the item and order memory components of the probed recall task. Item memory was at the same level in all three groups. In contrast, order memory in the WS group was significantly worse than that of the TD group and at the same level as that in the MLD group. This suggests that the relatively good performance of the TD group on the other STM tasks (except nonword repetition) was a consequence of their superior order memory. The superior order memory of the TD group is consistent with the idea that order memory has a causal influence on vocabulary acquisition and therefore determines the rate of vocabulary acquisition. However, there are other possible explanations for this pattern of results (see section 7.3.4).

If serial order memory is associated with the rate of vocabulary acquisition, it is perhaps not surprising that the WS group did not outperform the TD group on measures of order memory. Nevertheless, if their relatively good language is a consequence of relatively preserved serial order memory, they should still have outperformed the MLD controls who had comparable vocabulary IQs. This prediction was not supported. Given that the same pattern of performance was also observed for overall performance on the STM tasks, there is also no evidence from the studies in Part III that relatively preserved STM abilities in general are responsible for linguistic strengths in WS.

The implication here is that serial order mechanisms in WS are at the same level as the other factors that contribute to vocabulary acquisition. However, this does rely on the assumption that there are other factors that have a direct influence on the rate of vocabulary acquisition. It has been argued that receptive vocabulary tests provide a weak test of semantic knowledge (Temple et al., 2002), which would imply that the ability to learn the phonological form of words is the main constraint on receptive vocabulary. If learning the phonological form is primarily determined by serial order mechanisms, one would expect serial order memory to predict the rate of vocabulary acquisition, and therefore that WS and MLD controls matched on age and vocabulary should have comparable serial order memory.

12.3: Is semantic memory impaired in WS?

12.3.1: Is there a reduced contribution of lexical-semantic knowledge to STM?

The second main issue in this thesis concerned the suggestion that there is some form of semantic memory impairment in WS. The experiments in Part III allowed investigation of the contribution of lexical-semantic knowledge to performance on STM tasks. Previous studies reporting reduced effects of lexical-semantic factors on STM performance in WS had been interpreted in terms of impaired semantic memory and linked to evidence for impaired semantic processing in WS. However, the precise nature of the link was not specified, and it was not clear from the pattern of results that they did indeed reflect impaired semantic memory. Moreover, there were methodological concerns with many of these previous studies.

Two hypotheses were therefore tested. The first was that semantic memory is impaired in WS. Patients with acquired semantic memory impairments have difficulty learning the meaning of words (Freedman & Martin, 2001), so this hypothesis could possibly help explain why children with WS seem to be better at learning the phonological forms of words than their meanings (cf. Paterson, 2000; Singer-Harris et al., 1997). The second hypothesis was that redintegration processes are impaired in WS: i.e., people with WS do not make use of knowledge of the phonological structure of words to fill in degraded phonological STM traces. Redintegration is assumed to involve normal speech processing mechanisms (e.g. Hulme et al., 1997) so, if this hypothesis were true, it would not help explain semantic processing problems in WS, but would suggest differences in the way that individuals with WS process speech.

The major concern with most of the previous studies of lexical-semantic effects on STM in WS related to the choice of control group. Thus, Vicari, Carlesimo et al. (1996) reported a reduced effect of word-frequency in WS, but the WS group were much older than controls and probably had superior vocabulary knowledge. As such they were likely to be more familiar with the low-frequency words, and this could explain why they performed relatively well on recall of these items, leading to the reduced word-frequency effect. Majerus et al. (2001) reported reduced effects of lexicality and phonotactic frequency in serial recall, but they compared children with WS to age-matched controls, so their findings could reflect the reduced familiarity of the WS group with the stimuli. Karmiloff-Smith et al. (1997) reported a reduced tendency to make lexicalization errors in repetition of nonwords, but the WS group had much higher language ages than controls and were therefore likely to have superior STM. The reduction in lexicalization errors could, therefore, simply have reflected an overall reduction in errors. Finally, Grant et al. (1997) reported a normal effect of wordlikeness on nonword repetition, but the WS group were much older and had superior vocabulary knowledge, so it was unclear whether they would show a reduced wordlikeness effect compared with controls with similar language experience.

The studies in Part III therefore controlled for vocabulary knowledge. In Experiments 3, 4, 5, and 7, this was achieved by matching children with WS to TD and MLD controls on their BPVS ages. In Experiment 6, children with WS were compared with TD controls who had lower vocabulary ages, but vocabulary age was controlled for by covariation. The WS group showed normal effects of lexicality (Experiment 3), word-frequency (Experiments 4 and 5), concreteness (Experiment 6), and wordlikeness (Experiment 7). They also showed no significant reduction in the number or proportion of lexicalization errors in recall of nonwords (Experiment 3). As such, there was no evidence to support either the impaired semantic memory hypothesis or the impaired redintegration hypothesis.

12.3.2: Is semantic LTM impaired in WS?

Part IV of the thesis investigated free recall in WS. Vicari, Brizzolara et al. (1996) reported that children with WS showed a reduced primacy effect in free recall and this was taken as evidence of impaired semantic LTM, on the basis that the extent of the primacy effect in TD adults is influenced by semantic factors such as word-frequency and concreteness. These authors therefore argued that this finding was related to deficits in semantic processing, although the precise nature of the relationship was not specified.

There were, however, a number of alternative explanations for the finding. First, the control group performed significantly better than the WS group overall. The extent of the primacy effect increases with developmental improvement in overall performance, so the reduced primacy effect could simply reflect reduced overall performance. Second, Vicari, Brizzolara et al. (1996) adopted an unusual procedure whereby participants were repeatedly tested on the same items in the same order, and the reduced primacy effect could have arisen because the WS group did not benefit from previous trials with the same stimuli. Third, primacy effects depend on the adoption of a cumulative rehearsal strategy, so the reduction in the effect in WS could have reflected group differences in rehearsal strategy.

In Experiment 8, controls were matched on overall performance and neither children with WS nor TD controls showed a significant primacy effect. Both groups showed comparable improvements in performance when tested repeatedly on the same items. In Experiment 9, participants were encouraged to engage in overt cumulative rehearsal, and this led to significant primacy effects in both groups. There was therefore no evidence for impaired semantic LTM in WS.

12.4: Implications

12.4.1: STM and language in developmental disabilities

The studies in Part III of this thesis could have important implications for research investigating the relationship between language and STM deficits in children with other developmental disabilities such as SLI and Down syndrome.

Hulme and Roodenrys (1995) have argued that STM deficits in developmental disorders are a consequence of language impairment rather than its cause. However, while poor nonword repetition may be caused by poor vocabulary knowledge or phonological awareness (cf. Figure 6.1), and poor digit span could plausibly be a consequence of deficits in speech production, it is difficult to see how this account could explain the finding that children with WS and MLD controls performed worse than TD controls on the probed position recall task which has no verbal output and is not influenced by lexical knowledge.

The alternative account is that deficits in STM are the primary cause of language impairment in these disorders (cf. Baddeley et al., 1998). Numerous studies, based on the phonological loop / working memory framework, have investigated whether the phonological store or the rehearsal process is impaired in SLI and Down syndrome (cf. Jarrold, 2001). However, the available evidence suggests that rehearsal does not play a role in language acquisition (Baddeley et al., 1998; see section 6.3.2), so it seems unlikely that language impairments can be explained in terms of rehearsal deficits. By default, this implies that impairments of the phonological store are the cause of language delay.

Unfortunately, the phonological loop model does not specify the mechanisms involved in maintaining representations in the phonological store. Consequently, further insights into STM deficits in developmental disorders may have to rely on more detailed models of STM such as OSCAR (Brown et al., 2000). Part III of this thesis highlighted the possible importance of serial order mechanisms in determining the rate of vocabulary acquisition. It therefore seems plausible that children with SLI or Down syndrome may have an underlying deficit in serial order mechanisms. According to models like OSCAR, this would imply a deficit in the encoding or retrieval of a temporal context signal (cf. Boucher, 1999; see section 3.3.2).

12.4.2: Implications for theories of language in WS

None of the studies in this thesis found any significant differences between children with WS and MLD controls, and the only difference between children with WS and TD controls was in terms of the superior serial order memory of the TD group. For the temporal processing studies, the failure to find any group differences is interesting insofar as it suggests that temporal discrimination abilities are a relative strength in WS. However, the lack of any evidence for a causal relationship between temporal discrimination and vocabulary means that the theoretical relevance of this finding unclear.

It was suggested in section 2.8.5 that temporal processing mechanisms are a good example of what Karmiloff-Smith (1998) terms 'domain relevant' mechanisms, in that they are likely to be more important for language processing than for other cognitive functions. Had there been evidence that temporal processing mechanisms contribute to linguistic strengths in WS, this would have provided support for the neuroconstructivist account of WS. However, the failure to find support for the temporal processing hypothesis does not constitute evidence against the neuroconstructivist account.

In contrast to the results from the temporal processing studies, the negative findings reported in the memory studies have direct implications for the understanding of language abilities in WS. First, they consistently contradicted the suggestion made in previous studies that STM is a relative strength in WS. Second, the absence of group differences in lexical-semantic effects on STM or in primacy effects in free recall directly contradict findings from previous studies. As such, the results of this thesis provide evidence against the suggestion that language development in WS is unusually dependent on STM and that semantic processing deficits are related to impairments in semantic memory (cf. Grant et al., 1997; Vicari, Carlesimo et al., 1996). Consequently, they weaken the argument that language development in WS proceeds under an atypical set of constraints (cf. Karmiloff-Smith et al., 2002; Thomas & Karmiloff-Smith, 2002).

However, the findings do not, of course, rule out the possibility that other constraints may be atypical in WS. For example, studies with very young children with WS (e.g. Laing et al., 2002; Mervis & Bertrand, 1997; Paterson, 2000; Paterson et al. 1999; Singer-Harris et al., 1997) suggest that influences on early language development in WS are different to those in typical development. A further area of research that has received relatively little attention is the impact of auditory hyperacusis on language development[28].

12.4.3: Methodological implications

The studies reported in this thesis are similar to a number of other recent studies investigating language abilities in WS (e.g. Jarrold et al., 2000; Thomas et al., 2001). These studies addressed methodological concerns with earlier studies whose results had been used to make important theoretical claims about language in WS, and failed to replicate their findings. Together, they highlight the necessity of considering alternative explanations for interesting findings.

The most important methodological issue would appear to be the choice of control measure. In section 2.2.3, it was argued that this should depend on the specific research question that is being addressed. However, early studies of language and memory in WS were necessarily exploratory, so it was not clear at the outset what the best control measure would be. Moroever, a number of findings that have been taken as evidence for atypical language and memory processes in WS were not actually the main focus of the study in which they were reported, so groups were not matched in the most appropriate way (see for example, the reduction in lexicalization errors in nonword repetition reported by Karmiloff-Smith et al., 1997).

A further important issue, highlighted by the comparison between vocabulary and STM in Part III of this thesis, is the influence of a general delay in language acquisition on the linguistic profile. In chapter 2, a null hypothesis was proposed that stated that language abilities in WS are exactly what would be expected given overall abilities. It was assumed that if this hypothesis were correct, performance across a range of linguistic tasks would be at roughly the same age- equivalent level. However, if performance on one measure (e.g. verbal STM) is associated with the rate of change in another (e.g. vocabulary), it is likely that a general delay in langauge development will entail that age-equivalent performance on the first task will be at a different level to that for the second. Consequently, it is possible for the null hypothesis to be rejected, even though the dissociation simply reflects a general delay.[29] This highlights the importance of including a control group matched on chronological age as well as on relevant measures of ability (cf. Section 2.2.2).

12.5: Final thoughts

The studies in this thesis have looked at possible causes of the uneven cognitive and linguistic profile in WS. However, there was no evidence for atypical memory or temporal processing abilities that could not be explained in terms of a general delay in language acquisition. The assumption in this thesis and in many other accounts of WS has been that, while overall abilities are retarded, some factor (whether it be a specific linguistic module or a non-specific factor such as serial order mechanisms) is relatively preserved, and that this leads to relatively good language abilities in WS. Perhaps a more informative approach would be to look at the causes of the specific weaknesses in the cognitive profile (especially visuo-spatial construction) and attempt to determine why they have a relatively minor impact on language.

[26] The difference was not significant in the forward serial recall task

[27] As noted in section 10.5.4, this probably reflects the fact that nonword repetition depends on the use of knowledge of real words and the ability to identify the constituent phonemes in the nonword, factors that are likely to be related to the level of vocabulary knowledge rather than rate of acquisition.

[28] Thomas and Karmiloff-Smith (2002) have speculated that auditory perceptual anomolies in WS could result in atypical phonological representations, and consequently to impaired morphology.

[29] The possibility that a general delay in overall abilities combined with interactions between different linguistic components could lead to an apparent dissociation in performance on two tasks is entirely consistent with a neuroconstructivist account of WS.

References

Aaronson, D. (1968). Temporal course of perception in an immediate recall task. *Journal of Experimental Psychology, 76*, 129-140.

Allan, L. G., & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning and Motivation, 22*, 39-58.

Anderson, M. (1998). Mental retardation, general intelligence and modularity. *Learning and Individual Differences, 10*, 159-178.

Arber, S., Barbayannis, F. A., Hanser, H., Schneider, C., Stanyon, C. A., Bernard, O., & Caroni, P. (1998). Regulation of actin dynamics through phosphorylation of cofilin by LIM-kinase. *Nature, 393*, 805-809.

Arnold, R., Yule, W., & Martin, N. (1985). The psychological characteristics of infantile hypercalcaemia: A preliminary investigation. *Developmental Medicine and Child Neurology, 27*.

Atkinson, J., King, J., Braddick, O., Nokes, L., Anker, S., & Braddick, F. (1997). A specific deficit of dorsal stream function in Williams' syndrome. *NeuroReport, 8*, 1919-1922.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89-195). New York: Academic Press.

Babb, K., Gage, N., Perales, O., Huntley-Fenner, G., & Hickok, G. (1998). Tonal, verbal, and spatial memory span in Williams syndrome. *Journal of Cognitive Neuroscience*, 79.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362-365.

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Baddeley, A. D., & Hitch, G. J. (1977). Recency re-examined. In S. Dornic (Ed.), *Attention and performance VI* (pp. 647-667). Hillsdale, NJ: Erlbaum.

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long term learning depends on short-term storage. *Journal of Memory and Language, 27*, 586-595.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word-length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589.

Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language, 44*, 568-591.

Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29-79). London: LEA.

Bayley, N. (1993). *Bayley scales of infant development* (2nd ed.). San Antonio, TX: Psychological Corporation.

Bellugi, U., Adolphs, R., Cassady, C., & Chiles, M. (1999). Towards the neural basis for hypersociability in a genetic syndrome. *Neuroreport, 10*, 1653-1657.

Bellugi, U., Bihrle, A., Jernigan, T., Trauner, D., & Doherty, S. (1990). Neuropsychological, neurological, and neuroanatomical profile of Williams syndrome. *American Journal of Medical Genetics*, 115-125.

Bellugi, U., Bihrle, A., Neville, H., Doherty, S., & Jernigan, T. (1992). Language, cognition, and brain organization in a neurodevelopmental disorder. *Minnesota Symposia On Child Psychology, 24*, 201-232.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St George, M. (2000). The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience, 12*, 7-29.

Bellugi, U., Lichtenberger, L., Mills, D., Galaburda, A., & Korenberg, J. R. (1999). Bridging cognition, the brain and molecular genetics: evidence from Williams syndrome. *Trends in Neurosciences, 22*, 197-207.

Bellugi, U., Marks, S., Bihrle, A., & Sabo, H. (1988). Dissociation between language and cognitive functions in Williams syndrome. In D. Bishop & K. Mogford (Eds.), *Language development in exceptional circumstances* (pp. 177-189). London: Churchill Livingstone.

Bellugi, U., Mills, D., Jernigan, T., Hickock, G., & Galaburda, A. (1999). Linking cognition, brain structure and brain function in Williams syndrome. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders: Contributions to a new framework from the cognitive neurosciences* (pp. 111-136). Cambridge, MA: MIT Press.

Bellugi, U., Wang, P., & Jernigan, T. L. (1994). Williams syndrome: an unusual neuropsychological profile. In S. Broman & J. Grafman (Eds.), *Atypical cognitive deficits in developmental disorders: Implications for brain function* (pp. 23-56). London: Erlbaum.

Bellugi, U., & Wang, P. P. (1998). Williams syndrome: From cognition to brain to gene. In G. Edelman & B. H. Smith (Eds.), *Encyclopaedia of Neuroscience* (CD-ROM version). Amsterdam, the Netherlands: Elsevier.

Benton, A. L., Hamsher, K. d. S., Varney, N. R., & Spreen, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.

Bertrand, J., & Mervis, C. B. (1996). Longitudinal analysis of drawings by children with Williams syndrome: preliminary results. *Visual Arts Research, 22*, 19-34.

Bertrand, J., Mervis, C. B., & Eisenberg, J. D. (1997). Drawing by children with Williams syndrome: A developmental perspective. *Developmental Neuropsychology, 13*, 41-67.

Beuren, A. J., Schulze, C., Eberle, P., Harmjanz, D., & Apitz, J. (1964). The syndrome of supravalvular aortic stenosis, peripheral pulmonary stenosis, mental retardation and similar facial appearance. *American Journal of Cardiology, 13*, 471-482.

Bihrle, A., Bellugi, U., Delis, D., & Marks, S. (1989). Seeing either the forest or the trees: Dissociation in visuo-spatial processing. *Brain and Cognition, 11*, 37-49.

Bishop, D. (1989). *Test for the Reception of Grammar* (2nd ed.). Manchester: Chapel Press.

Bishop, D. V. M. (1997). Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *Quarterly Journal of Experimental Psychology, 50A*, 899-923.

Bishop, D. V. M. (1999). Cognition - An innate basis for language? *Science, 286*, 2283-2284.

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior, 13*, 80-97.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6*, 173-189.

Bongiovanni, A. M., Eberlein, W. R., & Jones, I. T. (1957). Idiopathic hypercalcaemia of infancy, with failure to thrive. *New England Journal of Medicine, 257*, 951-958.

Boucher, J. (1999). Time parsing, normal language acquisition and language-related developmental disorders. In M. Perkins & S. Howard (Eds.), *New directions in language development and disorders* (pp. 13-23). London: Plenum Press.

Bourassa, D., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin and Review, 1*, 122-125.

Bowey, J. A. (1996). On the association between phonological memory and receptive vocabulary in five-year-olds. *Journal of Experimental Child Psychology, 63*, 44-78.

Braddick, O., Atkinson, J., Wattam-Bell, J., & Anker, S. (1997). Global motion and form processing in Williams syndrome children. *Perception, 26*, 762.

Bradley, E. A., & Udwin, O. (1989). Williams Syndrome in Adulthood - a Case-Study Focusing On Psychological and Psychiatric Aspects. *Journal of Mental Deficiency Research, 33*, 175-184.

Brady, S. A. (1997). Ability to encode phonological representations: An underlying difficulty of poor readers. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 21-47). Mahwah, NJ: Erlbaum.

Brodie, D. A. (1975). Free recall measures of short-term store: Are rehearsal and order of recall data necessary? *Memory and Cognition, 3*, 653-662.

Bromberg, H., Ullman, M., Coppola, M., Marcus, G., Kelley, K., & Levine, K. (1994). *A dissociation of lexical memory and grammar in Williams syndrome: Evidence from inflectional morphology.* Paper presented at the Sixth International Professional Conference of the Williams Syndrome Association, San Diego, CA.

Brown, G. D. A., & Hulme, C. (1996). Nonword repetition, STM, and word age-of-acquisition: A computational model. In S. E. Gathercole (Ed.),

Models of short-term memory (pp. 129-148). Hove, England: Psychology Press.
Brown, G. D. A., Metsala, J. L., & Hulme, C. (2001). *Redintegration in short-term memory: Effects of lexical neighbours*. Unpublished manuscript.
Brown, G. D. A., Neath, I., & Chater, N. (2002). *A ratio model of scale-invariant memory and identification*. Manuscript submitted for publication.
Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107*, 127-181.
Brown, G. D. A., Vousden, J. I., McCormack, T., & Hulme, C. (1999). The development of memory for serial order: A temporal-contextual distinctiveness model. *International Journal of Psychology, 34*, 389-402.
Burack, J. A., Iarocci, G., Bowler, D., & Mottron, L. (2002). Benefits and pitfalls in the merging of disciplines: The example of developmental psychopathology and the study of persons with autism. *Development and Psychopathology, 14*, 225-237.
Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.
Caceres, J. A., Heinze, E. G., & Mendez, M. S. (1999). Preliminary evaluation of some cognitive, social, linguistic, and personality characteristics in a Spanish sample with Williams syndrome. Unpublished manuscript.
Capirci, O., Sabbadini, L., & Volterra, V. (1996). Language development in Williams syndrome: A case study. *Cognitive Neuropsychology, 13*, 1017-1039.
Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book*. New York: American Heritage.
Caza, N., & Belleville, S. (1999). Semantic contribution to immediate serial recall using an unlimited set of items: Evidence for a multi-level capacity view of short-term memory. *International Journal of Psychology, 34*, 334-338.
Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition, 37*, 55-81.
Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 3*, 216-228.
Church, R. M., & Gibbon, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes, 8*, 165-186.
Clahsen, H., & Almazan, M. (1998). Syntax and morphology in Williams syndrome. *Cognition, 68*, 167-198.
Clahsen, H., & Almazan, M. (2001). Compounding and inflection in language impairment: Evidence from Williams Syndrome (and SLI). *Lingua, 111*, 729-757.
Clahsen, H., & Temple, C. M. (2002). Words and rules in children with Williams syndrome. In Y. Levy & J. Schaeffer (Eds.), *Towards a definition of specific language impairment* (in press). Dordrecht: Kluwer Academic Publishers.
Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257-271.
Cole, M., Frankel, F., & Sharp, D. (1971). Development of free recall learning in children. *Developmental Psychology, 4*, 109-123.
Conrad, R. (1967). Interference or decay over short retention intervals. *Journal of Verbal Learning and Verbal Behavior, 6*, 49-54.
Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word-length on immediate memory. *Journal of Memory and Language, 31*, 1-17.
Cowan, N., & Kail, R. (1996). Covert processes and their development in short-term memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 29-50). Hove, England: Psychology Press.
Craik, F. I. M. (1968). Types of error in free recall. *Psychonomic Science, 10*, 353-354.
Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.
Creelman, C. D. (1962). Human discrimination of auditory duration. *Acoustical Society of America, 34*, 582-593.
Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory and Cognition, 21*, 142-145.
Cuvo, A. J. (1975). Developmental differences in rehearsal and free recall. *Journal of Experimental Child Psychology, 19*, 265-278.
Dall'oglio, A. M., & Milani, L. (1995). Analysis of the cognitive development in Italian children with Williams syndrome. *Genetic Counselling, 6*, 175-176.
De Jong, P. F., Seveke, M. J., & van Veen, M. (2000). Phonological sensitivity and the acquisition of new words in children. *Journal of Experimental Child Psychology, 76*, 275-301.
Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283-321.
Deruelle, C., Mancini, J., Livet, M. O., Casse-Perrot, C., & de Schonen, S. (1999). Configural and local processing of faces in children with Williams syndrome. *Brain and Cognition, 41*, 276-298.
Don, A. J., Schellenberg, E. G., & Rourke, B. P. (1999). Music and language skills of children with Williams syndrome. *Child Neuropsychology, 5*, 154-170.
Donnai, D., & Karmiloff-Smith, A. (2000). Williams syndrome: From genotype through to the cognitive phenotype. *American Journal of Medical Genetics, 97*, 164-171.
Droit-Volet, S., Clement, A., & Wearden, J. H. (2001). Temporal generalization in 3- to 8-year-old children. *Journal of Experimental Child Psychology, 80*, 271-288.
Droit-Volet, S., & Wearden, J. H. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology, 80*, 142-159.
Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test - revised* (2nd ed.). Circle Pines, MN: American Guidance Service.
Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale* (2nd ed.). Windsor, U.K.: NFER-Nelson.
Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). *British picture vocabulary scale*. Windsor, U.K.: NFER-Nelson.
Einfeld, S. L., Tonge, B. J., & Florio, T. (1997). Behavioral and emotional disturbance in individuals with Williams syndrome. *American Journal On Mental Retardation, 102*, 45-53.
Elliot, C. D. (1990). *Differential abilities scales*. New York: Psychological Corporation.
Elliot, C. D., Smith, P., & McCulloch, K. (1996). *British ability scales*. Windsor, U.K.: NFER-Nelson.
Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness*. Cambridge, MA: MIT Press.
Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.
Ewart, A. K., Morris, C. A., Atkinson, D., Jin, W. S., Sternes, K., Spallone, P., Stock, A. D., Leppert, M., & Keating, M. T. (1993). Hemizygosity At the Elastin Locus in a Developmental Disorder, Williams-Syndrome. *Nature Genetics, 5*, 11-16.
Ewart, A. K., Morris, C. A., Ensing, G. J., Loker, J., Moore, C., Leppert, M., & Keating, M. (1993). A human vascular disorder, supravalvular aortic-stenosis, maps to chromosome-7. *Proceedings of the National Academy of Sciences of the United States of America, 90*, 3226-3230.
Fabbro, F., Alberti, B., Gagliardi, C., & Borgatti, R. (2002). Differences in native and foreign language repetition tasks between subjects with William's and Down's syndromes. *Journal of Neurolinguistics, 15*, 1-10.
Farran, E. K., & Jarrold, C. (2002). Visuo-spatial cognition in Williams syndrome: Reviewing and accounting for the strengths and weaknesses in performance. *Developmental Neuropsychology*, in press.
Farran, E. K., Jarrold, C., & Gathercole, S. E. (2001). Block design performance in the Williams syndrome phenotype: A problem with mental imagery? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*, 719-728.
Farran, E. K., Jarrold, C., & Gathercole, S. E. (2002). Divided attention, selective attention and drawing: Processing preferences in Williams syndrome are dependent on the task administered. Manuscript submitted for publication.
Federmeier, K. D., & Kutas, M. (2000). It's about time. *Brain and Language, 71*, 62-64.
Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *MacArthur communicative development inventories: Technical manual*. San Diego, CA: Singular Publishing Group.
Finegan, J.-A., Smith, M. L., Meschino, W. S., Vallance, P. L., & Sitarenios, G. (1995). *Verbal memory in children with Williams syndrome.* Paper presented at the Society for Research in Child Development, Indianapolis, IN.
Fischler, I., Rundus, D., & Atkinson, R. C. (1970). Effects of overt rehearsal procedures on free recall. *Psychonomic Science, 19*, 249-250.
Fowler, A. E. (1990). Language abilities in children with Down syndrome. In D. Cicchetti & M. Beeghly (Eds.), *Children with Down syndrome: the developmental perspective* (pp. 302-328). New York: Cambridge University Press.

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy* (pp. 97-117). Hillsdale, NJ: Erlbaum.

Frangiskakis, J. M., Ewart, A. K., Morris, C. A., Mervis, C. B., Bertrand, J., Robinson, B. F., Klein, B. P., Ensing, G. J., Everett, L. A., Green, E. D., Proschel, C., Gutowski, N. J., Noble, M., Atkinson, D. L., Odelberg, S. J., & Keating, M. T. (1996). LIM-kinase 1 hemizygosity implicated in impaired visuospatial constructive cognition. *Cell, 86*.

Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology, 18*, 193-226.

Galaburda, A. M., & Bellugi, U. (2000). Multi-level analysis of cortical neuroanatomy in Williams syndrome. *Journal of Cognitive Neuroscience, 12*, 74-88.

Galaburda, A. M., Wang, P. P., Bellugi, U., & Rossen, M. (1994). Cytoarchitectonic anomalies in a genetically based disorder - Williams syndrome. *Neuroreport, 5*, 753-757.

Gallistel, C. R. (1990). *The organisation of learning*. Cambridge, MA: MIT Press.

Garcia, R. E., Friedman, W. F., Kaback, M. M., & Rowe, R. D. (1964). Idiopathic hypercalcaemia and supravalvular aortic stenosis. *New England Journal of Medicine, 271*, 117-120.

Gardner, M. F. (1983). *Expressive one-word picture vocabulary test*. Novato, CA: Academic Therapy Publications.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge - it all depends on the nonwords. *Memory & Cognition, 23*, 83-94.

Gathercole, S. E., & Adams, A. M. (1993). Phonological working-memory in very young children. *Developmental Psychology, 29*, 770-778.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Langauge, 28*, 200-213.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition - a study of young-children learning new names. *British Journal of Psychology, 81*, 439-454.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove: LEA.

Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of nonword repetition*. London: Psychological Corporation.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology:Learning Memory and Cognition, 25*, 84-95.

Gathercole, S. E., & Hitch, G. J. (1993). Developmental changes in short-term memory: A revised working memory perspective. In A. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 189-210). Hillsdale, NJ: Erlbaum.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33*, 966-979.

Gathercole, S. E., & Martin, A. J. (1996). Interactive processes in phonological memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 73-100). Hove, England: Psychology Press.

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology, 54A*, 1-30.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*, 65-77.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme - reading and vocabulary development in children. *British Journal of Psychology, 82*, 387-406.

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

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on childrens repetition of nonwords. *Applied Psycholinguistics, 12*, 349-367.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years - a longitudinal study. *Developmental Psychology, 28*, 887-898.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279-325.

Gibbon, J., Church, R. M., & Meck, W. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Annals of the New York Academy of Sciences, 423, Timing and time perception* (pp. 52-77). New York: New York Academy of Sciences.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7*, 170-184.

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory, Vol. V* (pp. 129-193). New York: Academic Press.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351-360.

Glanzer, M., & Meinzer, A. (1967). The effects of intralist activity on free recall. *Journal of Verbal Learning and Verbal Behavior, 6*, 928-935.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., Gretz, A. L., Fish, J. H., & Turpin, B. M. (1980). A two-process account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 355-369.

Gosch, A., & Pankau, R. (1997). Personality characteristics and behaviour problems in individuals of different ages with Williams syndrome. *Developmental Medicine and Child Neurology, 39*, 527-533.

Gosch, A., Stading, G., & Pankau, R. (1994). Linguistic abilities in children with Williams-Beuren syndrome. *American Journal of Medical Genetics, 52*, 291-296.

Grant, J., Karmiloff-Smith, A., Berthoud, I., & Christophe, A. (1996). Is the language of people with Williams syndrome mere mimicry? Phonological short-term memory in a foreign language. *Cahiers De Psychologie Cognitive-Current Psychology of Cognition, 15*, 615-628.

Grant, J., Karmiloff-Smith, A., Gathercole, S. A., Paterson, S., Howlin, P., Davies, M., & Udwin, O. (1997). Phonological short-term memory and its relationship to language in Williams syndrome. *Cognitive Neuropsychiatry, 2*, 81-99.

Grant, J., Valian, V., & Karmiloff-Smith, A. (2002). *A study of relative clauses in Williams syndrome*. Manuscript submitted for publication.

Gupta, P., & Dell, G. S. (1999). The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.), *The emergence of language* (pp. 447-481). Mahwah, NJ: Erlbaum.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Hagen, J. W., & Kingsley, P. R. (1968). Labeling effects in short-term memory. *Child Development, 39*, 113-121.

Hansen, J., & Bowey, J. A. (1994). Phonological analysis skills, verbal working-memory, and reading-ability in 2nd-grade children. *Child Development, 65*, 938-950.

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language, 35*, 1-31.

Hazeltine, E., Helmuth, L. L., & Ivry, R. B. (1997). Neural mechanisms of timing. *Trends in Cognitive Sciences, 1*, 163-169.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 644-655.

Heim, A. W. (1968). *AH4 test*. Windsor, U.K.: NFER-Nelson.

Helstrup, T. (1984). Serial position phenomena - training and retrieval effects. *Scandinavian Journal of Psychology, 25*, 227-250.

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36*, 73-137.

Hickock, G., Neville, H., Mills, D., Jones, W., Rossen, M., & Bellugi, U. (1995). Electrophysiological and quantitative MR analysis of the cortical auditory system in Williams syndrome. *Cognitive Neuroscience Society Abstracts, 2*, 66.

Hoffman, J. E., & Landau, B. (2000). Spared object recognition despite profound spatial deficits: Evidence from children with Williams syndrome. *Journal of Cognitive Neuroscience*, 22C.

Hopyan, T., Dennis, M., Weksberg, R., & Cytrynbaum, C. (2001). Music skills and the expressive interpretation of music in children with Williams-Beuren syndrome: pitch, rhythm, melodic imagery, phrasing, and musical affect. *Neuropsychology, development, and cognition. Section C, Child neuropsychology: a journal on normal and abnormal development in childhood and adolescence, 7*, 42-53.

Howell, D. C. (1992). *Statistical methods for psychology* (3rd ed.). Belmont, CA: Duxbury Press.

Howlin, P., Davies, M., & Udwin, O. (1998). Cognitive functioning in adults with Williams syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 39*, 183-189.

Huang, I. N. (1986). Transitory changes of primacy and recency in successive single- trial free-recall. *Journal of General Psychology, 113*, 5-21.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words - evidence for a long- term-memory contribution to short-term-memory span. *Journal of Memory and Language, 30*, 685-701.

Hulme, C., & Roodenrys, S. (1995). Practitioner review - Verbal working-memory development and its disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*, 373-398.

Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term-memory mechanisms in memory span. *British Journal of Psychology, 86*, 527-536.

Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term-memory mechanisms in memory span. *British Journal of Psychology, 86*, 527-536.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning Memory and Cognition, 23*, 1217-1232.

Huttenlocher, J., & Newcombe, N. (1976). Semantic effects on ordered recall. *Journal of Verbal Learning and Verbal Behavior, 15*, 387-399.

Ivry, R. (1993). Cerebellar involvement in the explicit representation of temporal information. *Annals of the New York Academy of Science, 682*, 214-230.

Jarrold, C. (2001). Applying the working memory model to the study of atypical development. In J. Andrade (Ed.), *Working memory in perspective* (pp. 126-150). Hove, U.K.: Psychology Press.

Jarrold, C., Baddeley, A. D., & Hewes, A. K. (1998). Verbal and nonverbal abilities in the Williams syndrome phenotype: Evidence for diverging developmental trajectories. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 39*, 511-523.

Jarrold, C., Baddeley, A. D., & Hewes, A. K. (1999). Genetically dissociated components of working memory: evidence from Down's and Williams syndrome. *Neuropsychologia, 37*, 637-651.

Jarrold, C., Baddeley, A. D., Hewes, A. K., Leeke, T. C., & Phillips, C. E. (2002). *Does verbal short-term memory drive vocabulary acquisition? Evidence from individuals with learning disability*. Manuscript submitted for publication.

Jarrold, C., Baddeley, A. D., Hewes, A. K., & Phillips, C. (2001). A longitudinal assessment of diverging verbal and non-verbal abilities in the Williams syndrome phenotype. *Cortex, 37*, 423-431.

Jarrold, C., Hartley, S. J., Phillips, C., & Baddeley, A. D. (2000). Word fluency in Williams syndrome: Evidence for unusual semantic organisation? *Cognitive Neuropsychiatry, 5*, 293-319.

Jernigan, T. L., Bellugi, U., Sowell, E., Doherty, S., & Hesselink, J. R. (1993). Cerebral morphologic distinctions between Williams and Down syndromes. *Archives of Neurology, 50*, 186-191.

Johnson, S. C. (2000). Williams syndrome and conceptual change in the domain of number. *Journal of Cognitive Neuroscience*, 84D.

Jones, W., Hickock, G., Rossen, M. L., & Bellugi, U. (1998). *Dissociations in cognitive development: Differential effects from two genetically based syndromes* . San Diego, CA: University of California, Center for Research in Language.

Jones, W., Singer, N., Rossen, M. L., & Bellugi, U. (1993). *Fractionations of higher cognitive functions in Williams syndrome: Developmental trajectories.* Paper presented at the ASHA National Convention, Anaheim, CA.

Joseph, M. C., & Parrott, D. (1958). Severe infantile hypercalcaemia with specific reference to the facies. *Archives of Diseases of Childhood, 33*, 385-395.

Karmiloff-Smith, A. (1992). *Beyond modularity*. Cambridge, MA: MIT Press.

Karmiloff-Smith, A. (1997). Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology, 13*, 513-524.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends In Cognitive Sciences, 2*, 389-398.

Karmiloff-Smith, A., Grant, J., Berthoud, I., Davies, M., Howlin, P., & Udwin, O. (1997). Language and Williams syndrome: How intact is "intact"? *Child Development, 68*, 246-262.

Karmiloff-Smith, A., Klima, E., Bellugi, U., Grant, J., & Baron-Cohen, S. (1995). Is there a social module - language, face processing, and theory of mind in individuals with Williams syndrome. *Journal of Cognitive Neuroscience, 7*, 196-208.

Karmiloff-Smith, A., Scerif, G., & Thomas, M. (2002). Different approaches to relating genotype to phenotype in developmental disorders. *Developmental Psychobiology, 40*, 311-322.

Karmiloff-Smith, A., Tyler, L. K., Voice, K., Sims, K., Udwin, O., Howlin, P., & Davies, M. (1998). Linguistic dissociations in Williams syndrome: evaluating receptive syntax in on-line and off-line tasks. *Neuropsychologia, 36*, 343-351.

Kataria, S., Goldstein, D. J., & Kushnick, T. (1984). Developmental delays in Williams (elfin facies) syndrome. *Applied Research in Mental Retardation, 5*, 419-423.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman brief intelligence test*. Circle Pines, MN: American Guidance Service.

King, J., Atkinson, J., Braddick, O., Hartley, T., Nokes, L., & Braddick, F. (1997). Dissociation of visually guided action from visual judgment in Williams syndrome children. *Perception, 26*, 762-763.

Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review, 1*.

Klein, B. P., & Mervis, C. B. (1999). Contrasting patterns of cognitive abilities of 9-and 10-year-olds with Williams syndrome or Down syndrome. *Developmental Neuropsychology, 16*, 177-196.

Kucera, H., & Francis, V. W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Laakmann, W. J. (2001). Cerebellar contributions to cognition: Evidence from Williams syndrome. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 62*, 2490.

Laing, E., Butterworth, G., Ansari, D., Gsodl, M., Longhi, E., Panagiotaki, G., Paterson, S., & Karmiloff-Smith, A. (2002). Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science, 5*, 233-246.

Laing, E., Grant, J., Thomas, M., & Karmiloff-Smith, A. (2002). *The influence of phonological and semantic factors on verbal short-term memory in Williams syndrome*. Manuscript submitted for publication.

Laing, E., Hulme, C., Grant, J., & Karmiloff-Smith, A. (2001). Learning to read in Williams syndrome: Looking beneath the surface of atypical reading development. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*, 729-739.

Leiter, R. G. (1979). *Leiter international performance scale, revised*. Chicago: Ed Stoelting Co.

Lenhoff, H. M. (1998). Insights into the musical potential of cognitively impaired people diagnosed with Williams syndrome. *Music Therapy Perspectives, 16*, 33-36.

Lenhoff, H. M., Wang, P. P., Greenberg, F., & Bellugi, U. (1997). Williams syndrome and the brain. *Scientific American, 277*, 68-73.

Levitin, D. J., & Bellugi, U. (1998). Musical abilities in individuals with Williams syndrome. *Music Perception, 15*, 357-389.

Levy, Y. (1996). Modularity of language reconsidered. *Brain and Language, 55*, 240-263.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review, 96*, 25-57.

Lichtenberger, L., & Bellugi, U. (1998). The intersection of spatial cognition and language in Williams syndrome. *Journal of Cognitive Neuroscience*, SS80.

Lowery, M. C., Morris, C. A., Ewart, A., Brothman, L. J., Zhu, X. L., Leonard, C. O., Carey, J. C., Keating, M., & Brothman, A. R. (1995). Strong correlation of elastin deletions, detected by FISH, with Williams syndrome. *American Journal of Human Genetics, 57*, 49-53.

Majerus, S., Palmisano, I., van der Linden, M., Barisnikov, K., & Poncelet, M. (2001). An investigation of phonological processing in Williams syndrome. *Journal of the International Society, 7*, 153.

Mangels, J. A., Ivry, R. B., & Shimizu, N. (1998). Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. *Cognitive Brain Research, 7*, 15-39.

Mari, A., Amati, F., Mingarelli, R., Giannotti, A., Sebastio, G., Colloridi, V., Novelli, G., & Dallapiccola, B. (1995). Analysis of the Elastin Gene in 60 Patients With Clinical- Diagnosis of Williams-Syndrome. *Human Genetics, 96*, 444-448.

Marriage, J. E. (1996). *Hyperacusis in Williams syndrome.* Unpublished Doctoral Thesis, University of Manchester, Manchester.

Marslen-Wilson, W. D., Vanhalen, S., & Moss, H. (1988). Distance and competition in lexical access. *Bulletin of the Psychonomic Society, 26*, 490-491.

Martin, N., & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology, 14*, 641-682.

Martin, N. D. T., Snodgrass, G. J. A. I., & Cohen, R. D. (1984). Idiopathic infantile hypercalcaemia - a continuing enigma. *Archives of Diseases in Childhood, 59*, 605-613.

Martin, R. C. (1993). Short-term-memory and sentence processing - evidence from neuropsychology. *Memory & Cognition, 21*, 176-183.

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working-memory - neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language, 33*, 83-111.

Masataka, N. (2000). Information from speech and gesture is integrated when meaning of new words are categorized in normal young children, but not in children with Williams syndrome. *Cognitive Studies, 7*, 37-51.

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays, 22*, 94-103.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech-perception. *Cognitive Psychology, 18*, 1-86.

McCormack, T., Brown, G. D. A., Maylor, E. A., Darby, R. J., & Green, D. (1999). Developmental changes in time estimation: Comparing childhood and old age. *Developmental Psychology, 35*, 1143-1155.

McCormack, T., Brown, G. D. A., Vousden, J. I., & Henson, R. N. A. (2000). Children's serial recall errors: Implications for theories of short-term memory development. *Journal of Experimental Child Psychology, 76*, 222-252.

Mervis, C. B., & Bertrand, J. (1997). Developmental relations between cognition and language: Evidence from Williams syndrome. In L. B. Adamson & M. A. Romski (Eds.), *Research on communication and language disorders: Contributions to theories of language development* (pp. 75-106). New York: Brookes.

Mervis, C. B., Morris, C. A., Bertrand, J., & Robinson, B. F. (1999). Williams syndrome: Findings from an integrated program of research. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders: Contributions to a new framework from the cognitive neurosciences* (pp. 65-110). Cambridge, MA: MIT Press.

Mervis, C. B., Robinson, B. F., Bertrand, J., Morris, C. A., Klein-Tasman, B. P., & Armstrong, S. C. (2000). The Williams syndrome cognitive profile. *Brain and Cognition, 44*, 604-628.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition, 25*, 47-56.

Metsala, J. L., & Stanovich, K. E. (1995). *An examination of young children's phonological processing as a function of lexical development.* Paper presented at the Annual American Educational Research Association, San Francisco, CA.

Michas, I. C., & Henry, L. A. (1994). The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology, 12*, 147-163.

Mills, D. L. (1998). Electrophysiological markers of Williams syndrome. *Journal of Cognitive Neuroscience*, SS10.

Mills, D. L., Alvarez, T. D., St George, M., Appelbaum, L. G., Bellugi, U., & Neville, H. (2000). Electrophysiological studies of face processing in Williams syndrome. *Journal of Cognitive Neuroscience, 12*, 47-64.

Modigliani, V., & Hedges, D. G. (1987). Distributed rehearsals and the primacy effect in single-trial free-recall. *Journal of Experimental Psychology: Learning Memory and Cognition, 13*, 426-436.

Morris, C. A., Demsey, S. A., Leonard, C. O., Dilts, C., & Blackburn, B. L. (1988). Natural history of Williams syndrome: Physical characterisitics. *Journal of Pediatrics, 113*, 318-326.

Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech Language and Hearing Research, 44*, 778-792.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*, 482-488.

Nadeau, S. E. (2001). Phonology: A review and proposals from a connectionist perspective. *Brain and Language, 79*, 511-579.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139-158.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: the validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67*, 359-370.

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition, 21*, 689-698.

Neath, I. (1997). Modality, concreteness, and set-size effects in a free reconstruction of order task. *Memory & Cognition, 25*, 256-263.

Neath, I., & Nairne, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review, 2*, 429-441.

Neville, H. J., Mills, D. L., & Bellugi, U. (1994). Effects of altered auditory sensitivity and age of language acquisition on the development of language-relevant neural systems: Preliminary studies of Williams Syndrome. In S. Broman (Ed.), *Cognitive Deficits in Developmental Disorders: Implications for Brain Function* (pp. 67-83). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nichelli, P., Alway, D., & Grafman, J. (1996). Perceptual timing in cerebellar degeneration. *Neuropsychologia, 34*, 863-871.

Ornstein, P. A., Naus, M. J., & Prince-Stone, B. (1977). Rehearsal training and developmental differences in memory. *Developmental Psychology, 13*, 15-24.

Osborne, L. R. (1999). Williams-Beuren syndrome: Unraveling the mysteries of a microdeletion disorder. *Molecular Genetics and Metabolism, 67*, 1-10.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105*, 761-781.

Pani, J. R., Mervis, C. B., & Robinson, B. F. (1999). Global spatial organization by individuals with Williams syndrome. *Psychological Science, 10*, 453-458.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term-memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*, 331-347.

Papagno, C., & Vallar, G. (1992). Phonological short-term-memory and the learning of novel words - the effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology, 44A*, 47-67.

Paterson, S. J. (2000). The development of language and number understanding in Williams syndrome and Down syndrome: Evidence from the infant and mature phenotype. Unpublished Doctoral Thesis, University College London.

Paterson, S. J., Brown, J. H., Gsodl, M. K., Johnson, M. H., & Karmiloff-Smith, A. (1999). Cognitive modularity and genetic disorders. *Science, 286*, 2355-2358.

Penhune, V. B., Zatorre, R. J., & Evans, A. C. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience, 10*, 752-765.

Pennington, E. A. (2001). *Temporal memory in ageing.* Unpublished Doctoral Thesis, University of Warwick, Coventry.

Phillips, C. E., Jarrold, C., Baddeley, A. D., Grant, J., & Karmiloff-Smith, A. (2002). Comprehension of spatial language terms in Williams syndrome: Evidence for an interaction between domains of strength and weakness. Manuscript submitted for publication.

Pinker, S. (1991). Rules of language. *Science, 253*, 530-535.

Pinker, S. (1994). *The language instinct*. New York: Harper Perennial.

Pinker, S. (1999). *Words and rules*. London: Weidenfeld & Nicolson.

Pinto, A. D. C., & Baddeley, A. D. (1991). Where did you park your car? Analysis of a naturalistic long-term recency effect. *European Journal of Cognitive Psychology, 3*, 297-313.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia - a case-study of connectionist neuropsychology. *Cognitive Neuropsychology, 10*, 377-500.

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words - further evidence on the influence of semantic factors in immediate serial-recall. *Quarterly Journal of Experimental Psychology, 48A*, 384-404.

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology, 17*, 132-138.

Proschel, C., Blouin, M.-J., Gutowski, N. J., Ludwig, R., & Noble, M. (1995). LIMK1 is predominantly expressed in neural tissues and phosphylorates serine, threonine and tyrosine residues in vitro. *Oncogene, 11*, 1271-1281.

Protopapas, A. (1999). Connectionist modeling of speech perception. *Psychological Bulletin, 125*, 410-436.

Quinlan, P. T. (1992). *The Oxford psycholinguistic database*. Oxford, England: Oxford Electronic Publishing, Oxford University Press.

Rae, C., Karmiloff-Smith, A., Lee, M. A., Dixon, R. M., Grant, J., Blamire, A. M., Thompson, C. H., Styles, P., & Radda, G. K. (1998). Brain biochemistry in Williams syndrome - Evidence for a role of the cerebellum in cognition? *Neurology, 51*, 33-40.

Raven, J. C. (1993). *Coloured progressive matrices*. Oxford, U.K.: Information Press Ltd.

Reilly, J., Klima, E. S., & Bellugi, U. A. (1990). Once more with feeling: Affect and language in atypical populations. *Development and Psychopathology, 2*, 367-391.
Reiss, A. L., Eliez, S., Schmitt, J. E., Straus, E., Lai, Z., Jones, W., & Bellugi, U. (2000). Neuroanatomy of Williams syndrome: A high-resolution MRI study. *Journal of Cognitive Neuroscience, 12*, 65-73.
Rey, A. (1968). *Reattivo della Figura Complessa. Manuale*. Firenze: Organizzazioni Speciali.
Richardson, J. T. E., & Baddeley, A. D. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior, 14*, 623-629.
Robinson, W. P., Waslynka, J., Bernasconi, F., Wang, M., Clark, S., Kotzot, D., & Schinzel, A. (1996). Delineation of 7q11.2 deletions associated with Williams-Beuren syndrome and mapping of a repetitive sequence to within and to either side of the common deletion. *Genomics, 34*, 17-23.
Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition., 28*, 29-33.
Roodenrys, S., Hulme, C., & Brown, G. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology, 56*, 431-442.
Rossen, M., Klima, E. S., Bellugi, U., Bihrle, A., & Jones, W. (1996). Interaction between language and cognition: Evidence from Williams syndrome. In J. H. Beitchman, N. Cohen, M. Konstantareas, & R. Tannock (Eds.), *Language learning and behaviour* (pp. 367-392). New York: Cambridge University Press.
Rubba, J., & Klima, E. S. (1991). Preposition use in a speaker with Williams syndrome: some cognitive grammar proposals. *Center of Research in Language Newsletter, University of California, La Jolla, CA, 3*, 3-12.
Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89*, 63-77.
Sacks, O. (1995). Musical ability. *Science, 268*, 621-622.
Saint-Aubin, J., & Poirier, M. (1999). The influence of long-term memory factors on immediate serial recall: An item and order analysis. *International Journal of Psychology, 34*, 347-352.
Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.
Sarimski, K. (1997). Behavioral phenotypes and family stress in three mental retardation syndromes. *European Journal of Child and Adolescent Psychiatry, 6*, 26-31.
Scarborough, H. S. (1990). Index of productive syntax. *Applied Psycholinguistics, 11*, 1-22.
Schweickert, R., Chen, S. B., & Poirier, M. (1999). Redintegration and the useful lifetime of the verbal memory representation. *International Journal of Psychology, 34*, 447-453.
Scott, P., Mervis, C. B., Bertrand, J., Klein, B. P., Armstrong, S. C., & Ford, A. L. (1995). Semantic organization and word fluency in 9- and 10-year-old children with Williams syndrome. *Genetic Counselling, 6*, 172-173.
Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 34*, 1351-1364.
Shallice, T. (1975). On the contents of primary memory. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 269-280). London: Academic Press.
Singer-Harris, N. G., Bellugi, U., Bates, E., Jones, W., & Rossen, M. (1997). Contrasting profiles of language development in children with Williams and Down syndromes. *Developmental Neuropsychology, 13*, 345-370.
Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174-215.
Snowling, M., Goulandris, N., Bowlby, M., & Howell, P. (1986). Segmentation and speech-perception in relation to reading skill - a developmental analysis. *Journal of Experimental Child Psychology, 41*, 489-507.
Solity, J. E., & Vousden, J. I. (2002). *Learning to read using a rational analytic approach*. Manuscript in preparation.
Stevens, T. (1996). *Lexical constraints on language acquisition in Williams syndrome.* Unpublished Doctoral Thesis, MRC Cognitive Development Unit and University College London.
Stevens, T., & Karmiloff-Smith, A. (1997). Word learning in a special population: do individuals with Williams syndrome obey lexical constraints? *Journal of Child Language, 24*, 737-765.
Stiers, P., Willekens, D., Borghgraef, M., Fryns, J. P., & Vandenbussche, E. (2000). *Visuo-perceptual and visuo-constructive ability in persons with Williams syndrome.* Paper presented at the 8th International Professional Conference on Williams Syndrome, Dearborn, MI.
Stojanovik, V., Perkins, M., & Howard, S. (2001). Language and conversational abilities in Williams syndrome: How good is good? *International Journal of Language & Communication Disorders, 36*, 234-239.
Stothard, S. E., & Hulme, C. (1995). A comparison of phonological skills in children with reading- comprehension difficulties and children with decoding difficulties. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*, 399-408.
Studdert-Kennedy, M., & Mody, M. (1995). Auditory temporal perception deficits in the reading-impaired: A critical review of the evidence. *Psychonomic Bulletin and Review, 2*, 508-514.
Tager-Flusberg, H., Boshart, J., & Baron-Cohen, S. (1998). Reading the windows to the soul: Evidence of domain-specific sparing in Williams syndrome. *Journal of Cognitive Neuroscience, 10*, 631-639.
Tager-Flusberg, H., & Sullivan, K. (1999). *Are children with Williams syndrome spared in theory of mind?* Paper presented at the Society for Research in Child Development, Albuquerque, NM.
Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: evidence from Williams syndrome. *Cognition, 76*, 59-89.
Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature, 241*, 468-469.
Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning Memory and Cognition, 26*, 1589-1625.
Tassabehji, M., Metcalfe, K., KarmiloffSmith, A., Carette, M. J., Grant, J., Dennis, N., Reardon, W., Splitt, M., Read, A. P., & Donnai, D. (1999). Williams syndrome: Use of chromosomal microdeletions as a tool to dissect cognitive and physical phenotypes. *American Journal of Human Genetics, 64*, 118-125.
Temple, C. M., Almazan, M., & Sherwood, S. (2002). Lexical skills in Williams Syndrome: a cognitive neuropsychological analysis. *Journal of Neurolinguistics, 15*, 463-495.
Thomas, M. S. C. (2001). *Semantic representations in individuals with Williams syndrome*, Seminar presentation at the Centre for Brain and Cognitive Development, Birkbeck College, London.
Thomas, M. S. C., Grant, J., Barham, Z., Gsodl, M., Laing, E., Lakusta, L., Tyler, L. K., Grice, S., Paterson, S., & Karmiloff-Smith, A. (2001). Past tense formation in Williams syndrome. *Language and Cognitive Processes, 16*, 143-176.
Thomas, M. S. C., & Karmiloff-Smith, A. (2002). *Modeling language acquisition in atypical phenotypes.* Manuscript submitted for publication.
Thorndike, E. L., & Lorge, I. (1944). *The teacher's book of 30,000 words*. New York: Teachers College Press.
Tomc, S. A., Williamson, N. K., & Pauli, R. M. (1990). Temperament in Williams syndrome. *American Journal of Medical Genetics, 36*, 345-352.
Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs, 77*, 1-31.
Trojano, L., & Grossi, D. (1995). Phonological and lexical coding in verbal short-term-memory and learning. *Brain and Language, 51*, 336-354.
Turner, J. E., Henry, L. A., & Smith, P. T. (2000). The development of the use of long-term knowledge to assist short-term recall. *Quarterly Journal of Experimental Psychology, 53A*, 457-478.
Tyler, L. K., Karmiloff-Smith, A., Voice, J. K., Stevens, T., Grant, J., Udwin, O., Davies, M., & Howlin, P. (1997). Do individuals with Williams syndrome have bizarre semantics? Evidence for lexical organization using an on-line task. *Cortex, 33*, 515-527.
Tzeng, O. J. L. (1973). Positive recency effect in a delayed free recall. *Journal of Verbal Learning and Verbal Behavior, 12*, 436-439.
Udwin, O. (1990). A survey of adults with Williams syndrome and idiopathic infantile hypercalcemia. *Developmental Medicine and Child Neurology, 32*, 129-141.
Udwin, O., Davies, M., & Howlin, P. (1996). A longitudinal study of cognitive abilities and educational attainment in Williams syndrome. *Developmental Medicine and Child Neurology, 38*, 1020-1029.
Udwin, O., & Yule, W. (1990). Expressive language of children with Williams syndrome. *American Journal of Medical Genetics*, 108-114.

Udwin, O., & Yule, W. (1991). A cognitive and behavioral-phenotype in Williams syndrome. *Journal of Clinical and Experimental Neuropsychology, 13*, 232-244.

Udwin, O., & Yule, W. (1998). *Williams Syndrome: Guidelines for parents.* Tonbridge, U.K.: Williams Syndrome Foundation.

Udwin, O., Yule, W., & Martin, N. D. T. (1987). Cognitive abilities and behavioural characteristics of children with idiopathic infantile hypercalcaemia. *Journal of Child Psychology and Psychiatry, 28*, 297-309.

Van Bon, W. H. J., & Van der Pijl, J. M. L. (1997). Effects of word length and wordlikeness on pseudoword repetition by poor and normal readers. *Applied Psycholinguistics, 18*, 101-114.

Van Borsel, J., Curfs, L. M. G., & Fryns, J. P. (1997). Hyperacusis in Williams syndrome: A sample survey study. *Genetic Counselling, 8*, 121-126.

Van der Lely, H. K. J., & Ullman, M. T. (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes, 16*, 177-217.

Vicari, S., Brizzolara, D., Carlesimo, G. A., Pezzini, G., & Volterra, V. (1996). Memory abilities in children with Williams syndrome. *Cortex, 32*, 503-514.

Vicari, S., Carlesimo, G., Brizzolara, D., & Pezzini, G. (1996). Short term memory in children with Williams syndrome: A reduced contribution of lexical-semantic knowledge to word span. *Neuropsychologia, 34*, 919-925.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*, 374-408.

Volterra, V., Capirci, O., Pezzini, G., Sabbadini, L., & Vicari, S. (1996). Linguistic abilities in Italian children with Williams syndrome. *Cortex, 32*, 663-677.

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology, 41*, 101-175.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Review, 30*, 73-87.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning Memory and Cognition, 25*, 1256-1271.

Wang, P. P., & Bellugi, U. (1993). Williams syndrome, Down syndrome, and cognitive neuroscience. *American Journal of Diseases of Children, 147*, 1246-1251.

Wang, P. P., & Bellugi, U. (1994). Evidence from 2 genetic syndromes for a dissociation between verbal and visual-spatial short-term-memory. *Journal of Clinical and Experimental Neuropsychology, 16*, 317-322.

Wang, P. P., Doherty, S., Rourke, S. B., & Bellugi, U. (1995). Unique profile of visuo-perceptual skills in a genetic syndrome. *Brain and Cognition, 29*, 54-65.

Wang, P. P., Hesselink, J. R., Jernigan, T. L., Doherty, S., & Bellugi, U. (1992). Specific neurobehavioral profile of Williams syndrome is associated with neocerebellar hemispheric preservation. *Neurology, 42*, 1999-2002.

Watkins, O. C., & Watkins, M. J. (1977). Serial recall and the modality effect: Effects of word frequency. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 712-718.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72*, 89-104.

Wearden, J. H. (1991). Human-performance on an analog of an interval bisection task. *Quarterly Journal of Experimental Psychology, 43B*, 59-81.

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 18*, 134-144.

Wearden, J. H. (1999). "Beyond the fields we know...": Exploring and developing scalar timing theory. *Behavioural Processes, 45*, 3-21.

Wechsler, D. (1974). *Wechsler intelligence scale for children - revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1986). Wechsler adult intelligence scale - revised, U.K. edition. Windsor, U.K.: NFER-Nelson.

Wetherick, N. E. (1975). The role of semantic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 471-480.

Whittle, M. B. Z., Chang, F., Thomas, M. L., & Mervis, C. B. (2001). *Acquisition of concrete and relational vocabulary by children with Williams syndrome.* Paper presented at the Society for Research in Child Development, Minneapolis, MN.

Williams, E. J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society (Series B), 21*, 396-399.

Williams, J. C. P., Barratt-Boyes, B. G., & Lowe, J. B. (1961). Supravalvular aortic stenosis. *Circulation, 24*, 1311-1318.

Zukowski, A. (2001). *Uncovering grammatical competence in children with Williams syndrome.* Unpublished Doctoral Thesis, Boston University.

Acknowledgements

I would first of all like to thank the Williams Syndrome Foundation (WSF) for funding my Ph.D., and my supervisors, Jill Boucher and Teresa McCormack, for their support, encouragement and all the other things that supervisors do. Thanks also go to Gordon Brown for being my unofficial third supervisor, to Glynn Collis for supervising me in my first year, and to George Dunbar who was nominally my mentor (whatever that means) but was really my stats guru.

A number of people have also helped with my work, providing information, discussing methodologies or trying to work out why we've got completely different results. These people include Chris Jarrold, Julia Grant, Stefano Vicari, Steve Majerus, Judy Turner, Fred Thomas, and Emma Laing.

Special thanks go to John and Pam Nelson at the WSF for helping me recruit the children with WS, and to the WS families whose weekends and holidays I've interrupted and who have provided me with vast quantities of tea and biscuits. Thanks also to the teachers and parents of Round Oak, Woodloes, Westgate, Sparrowdale, and Inscape House schools, particularly Ms Pocock at Round Oak and Mrs Clough at Woodloes Infants.

I would also like to acknowledge the various people who have indirectly contributed to this thesis. Thanks to my office-mates Janet, Rachel, and Mark who had the pleasure of my company and the displeasure of my lack of tidiness for the past three and a bit years. To all the other postgrads and post-docs, including Sarah, Neil, Lewis, Gaynor, Caroline, Liz, and Chris for their varying degrees of encouragement and abuse (special thanks to Sarah for proof-reading my thesis - I'll be repaying the favour soon!). To the mutifarious residents of 21 Kenilworth Road, including Nick, Cam, Delph, Duncans 1 & 2, Marks I & II, Jon, Hywel, Nicky, Pascal, Fishy and about 30 others. To my Mum and Dad for their encouragement and support and all the other things that parents do. To my Gran for entrusting me with her little car which I would have been stuck without.

Finally, I would like to thank all the children who took part in my research and put up with the beeping owls, flapping chickens, and scary talking sheep, even though they weren't quite up to PlayStation standards.

Appendix 1: Stimuli for Experiments 3 and 4

Set	HFW	FAM	CNC	IMG	KFF	TLF	NW	LFW	FAM	CNC	IMG	KFF	TLF
A	cheat	549	329	457	3	1020	reet	coil	498	490	492	6	540
	fine	560	328	392	161	1078	hine	dip	466	479	456	6	1280
	get	604	290	264	750	69750	ret	gig	425	525	437	1	30
	gone	546	321	313	195	18590	chon	keel	337	515	451	6	150
	have	597	251	271	3941	244561	nav	lathe	295	481	356	1	20
	keep	584	339	284	264	20810	meep	loon	303	481	348	2	150
	large	634	-	449	361	16970	targe	nave	203	402	282	0	0
	live	608	436	446	177	43070	tiv	perch	396	576	462	1	680
	loud	577	413	448	20	2140	goud	psalm	418	409	427	4	80
	make	618	299	322	794	71430	pake	puck	281	472	370	0	60
	nice	583	279	375	75	8300	wice	rack	486	535	439	9	1370
	part	579	339	340	500	15850	gart	sap	377	540	451	1	330
	pull	565	360	446	51	9360	chull	sash	385	540	488	3	520
	rude	550	-	390	6	910	kood	shone	482	363	384	5	910
	same	569	248	326	686	10000	hame	siege	345	417	445	6	170
	shock	560	395	471	31	3440	gock	tun	141	211	235	0	0
	shoot	536	445	494	27	3150	doot	vile	404	379	444	5	450
	touch	590	417	456	87	10160	ruch	wad	354	479	370	0	200
	when	609	219	231	2331	147750	shen	whirl	423	402	499	3	1560
	wish	556	270	399	110	12200	gish	whoop	320	383	426	1	270
	MEAN	579	332	379	529	35527		MEAN	367	454	413	3	439
B	cool	567	364	429	62	5390	sool	debt	494	416	384	13	2200
	cut	581	430	460	192	9430	lut	dell	244	513	397	5	1180
	down	546	339	459	895	55340	chown	dirge	282	403	262	2	40
	feel	588	324	363	216	15610	cheal	gush	383	396	483	1	270
	got	565	252	286	482	41100	fot	jerk	452	441	479	2	1470
	hard	595	425	460	202	19090	sard	keen	449	336	335	11	1680
	less	541	266	286	438	13570	wess	lewd	290	-	324	3	30
	look	607	376	395	399	92150	wook	lodge	429	538	464	19	930
	miss	586	372	447	258	24080	niss	mass	473	397	484	110	3230
	move	572	390	413	171	18680	koove	mousse	363	492	439	0	0
	need	589	314	327	360	25390	keed	numb	487	379	477	4	550
	race	543	463	457	103	5520	gace	reap	441	373	408	3	120
	sad	589	360	419	35	2020	wad	rhyme	480	434	475	3	300
	sharp	572	-	495	72	3240	marp	sage	424	462	434	2	480
	sit	581	437	487	67	10000	yit	sop	259	373	294	1	100
	tell	596	306	350	268	38000	kell	tarn	251	399	299	0	10
	time	604	343	413	1599	85590	hime	till	442	365	364	50	4530
	top	571	435	486	204	8960	rop	toil	468	386	393	1	750
	wait	577	317	357	94	22160	nate	vale	334	467	421	4	80
	wide	569	348	455	125	5930	mide	whiff	426	413	461	1	140
	MEAN	577	361	412	312	25063		MEAN	394	420	404	12	905

FAM – Familiarity ratings (Quinlan, 1992)
CNC – Concreteness ratings (Quinlan, 1992)
IMG – Imageability ratings (Quinlan, 1992)
NW - nonword
KFF – Written frequencies (Kucera & Francis, 1967)
TLF – Written-frequencies (Thorndike & Lorge, 1944)
HFW - high-frequncy word
LFW - low-frequency word

Appendix 2 – Stimuli for Experiment 5

HFW	SVF	KFF	AOA	FAM	CNC
people	81	847	281	628	540
lady	38	80	231	573	564
uncle	27	57	192	557	580
supper	13	37	250	593	563
paper	29	157	229	635	599
letter	10	145	256	610	577
money	20	265	247	631	574
honey	37	25	286	533	611
mirror	15	27	258	593	605
water	75	442	153	641	616
mother	54	216	144	632	579
trouble	13	134	322	590	310
wonder	24	67	339	530	305
magic	174	37	281	481	257
pocket	11	46	228	590	578
giant	41	23	256	469	515
village	33	72	317	524	576
brother	15	73	219	598	585
window	43	119	231	621	609
island	26	167	289	507	596
MEAN	38.9	152	250	577	537

MFW	SVF	KFF	AOA	FAM	CNC
copper	1	13	428	491	547
willow	1	9	386	425	589
cellar	1	26	361	467	572
singer	1	10	314	548	553
organ	1	12	356	510	596
kennel	2	3	322	449	611
kettle	2	3	274	551	602
rubber	1	15	289	547	596
elbow	1	10	237	564	607
ankle	2	8	264	543	608
ocean	1	34	317	526	593
damage	1	33	358	558	406
delight	1	29	414	510	282
temper	2	12	333	575	353
polish	2	19	336	485	535
coral	1	5	434	425	572
ribbon	1	12	286	480	600
mermaid	1	1	322	391	494
soldier	1	39	275	517	578
salad	1	9	342	554	595
MEAN	1.3	15.1	332	506	544

LFW	KFF	AOA	FAM	CNC
labour	4	506	559	406
algae	7	631	317	545
abyss	4	597	293	450
abode	4	553	334	498
nozzle	4	411	412	555
cedar	1	425	381	608
ulcer	5	517	423	558
knuckle	3	356	491	586
basin	7	250	504	602
riot	7	456	490	414
maiden	2	429	374	545
halter	1	511	374	550
havoc	3	469	462	338
thicket	1	469	361	571
herring	2	397	425	617
torture	3	408	471	437
linen	6	386	515	581
bristle	3	383	461	558
monarch	3	456	428	525
bosom	8	489	425	552
MEAN	3.9	455	425	525

SVF – Children's written frequencies (Solity & Vousden, 2002)
KFF – Written frequencies (Kucera & Francis, 1967)
AOA – Age of acquisition ratings (Quinlan, 1992)
MFW - medium-frequency word
FAM – Familiarity ratings (Quinlan, 1992)
CNC – Concreteness ratings (Quinlan, 1992)
HFW - high-frequency word
LFW - low-frequency word