Speeded naming, frequency and the development of the lexicon in Williams syndrome

Michael S. C. Thomas,[1,2] Julie E. Dockrell,[3] David Messer,[4] Charlotte Parmigiani,[1] Daniel Ansari,[1,5] and Annette Karmiloff-Smith[1]

Atypical vocabulary has been reported as one of the most notable features of the language of adolescents and adults with Williams syndrome (WS), including use of unusual or low frequency words. Two hypotheses were identified regarding the developmental origins of this phenomenon. The *intra-lexicon* hypothesis views the cause in terms of domain-specific anomalies of structure or activation dynamics in the WS lexicon. The *extra-lexicon* hypothesis views the cause in terms of pragmatic influences, whereby individuals with WS use social engagement devices in their language in service of their ''hyper-social'' profile (Jones et al., 2000), and domain-general deficits to lexical-semantic representations in line with the level of learning disability present in WS. The hypotheses were evaluated using a speeded picture-naming task, in which frequency and semantic category were manipulated as implicit variables. The performance of 16 adolescents and adults with WS was compared with two individually matched control groups, one matched on chronological age (CA) and the other on receptive vocabulary age (RVA). Developmental trajectories were also constructed to assess the relationship between performance and age. Results indicated slower and less accurate naming in the WS group compared with both CA

[1] Neurocognitive Development Unit, Institute of Child Health, London
[2] School of Psychology, Birkbeck College, London
[3] Institute of Education, University of London
[4] Faculty of Education and Language Studies, The Open University, UK
[5] Department of Education, Dartmouth College, USA

Correspondence should be addressed to Dr Michael Thomas, School of Psychology, Birkbeck College, University of London, Malet Street, Bloomsbury, London WC1E 7HX, UK. Email: m.thomas@bbk.ac.uk

This research was supported by MRC Programme Grant No. G9715642 to Annette Karmiloff-Smith and MRC Career Establishment Grant G0300188 to Michael Thomas. We would like to thank John Brock and an anonymous reviewer for their helpful comments on a previous version of this article.

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960500258528

and RVA groups, but frequency and semantic category effects in line with the RVA group. Trajectories were delayed but not atypical. The data support a normal encoding of word frequency (rarity) in the WS lexicon, with rare-word usage explained as a pragmatic aspect of the WS social profile. The finding sets bounds on the domain-specificity of the complex developmental pathway between atypical genotype and atypical phenotype in WS.

INTRODUCTION

Williams syndrome (WS) is a rare neurodevelopmental disorder occurring in approximately 1 in 20,000 live births (Morris et al., 1988), caused by a deletion of a small number of genes from one copy of chromosome 7 at q11.23 (Tassabehji et al., 1999). The disorder is notable for the uneven cognitive profile exhibited in the adult phenotype. While overall IQ scores typically fall between 50 and 70, differential levels of impairment are seen across cognitive abilities. For example, language abilities appear less impaired than visuo-spatial abilities. Performance on standardised tests of face recognition in WS can fall within the normal range, while individuals exhibit difficulties in numerical cognition, problem-solving and planning (Donnai & Karmiloff-Smith, 2000; Mervis, Morris, Bertrand & Robinson, 1999; Paterson, Brown, Gsödl, Johnson & Karmiloff-Smith, 1999). The disorder is also notable for its characteristic social phenotype: individuals with WS are described as over-friendly and engaging, with a ‘‘hypersociable’’ personality profile (Jones, Bellugi, Lai, Chiles, Reilly, Lincoln, & Adolphs, 2000).

The uneven cognitive profile has encouraged some researchers to view WS as offering a *developmental fractionation* of higher cognitive functioning that may have implications for understanding normal development and even, given the genetic basis of WS, for the way that genes influence the development of adult cognitive structures (see Thomas & Karmiloff-Smith, 2005, for discussion). In this article, we concentrate on language ability in WS. This capacity was initially portrayed as exhibiting delay but then developing normally in the disorder. However, subsequent careful investigation has revealed anomalies at many levels of language processing (see Thomas & Karmiloff-Smith, 2003, for review).

One of the most striking features of language in WS is the rare or low frequency words that these individuals are reported to employ in their vocabulary (Bellugi, Wang & Jernigan, 1994; Udwin & Dennis, 1995). Thus WS speech is described as ‘‘peppered with unexpected word choices’’ (Bellugi, Bihrle, Jernigan, Trauner, and Doherty, 1990: 117), whereby individuals are said to show a ‘‘proclivity for unusual words’’ (e.g., Bellugi, Korenberg, & Klima, 2001: 220). These unusual word choices are not always entirely appropriate to the context. Rather, they are ''recherché and slightly off-target'' (Pinker, 1999), ''hyper-specific'' including additional contextually unnecessary detail, or even ''mal-specific'' with extra detail that incorrectly specifies the intended meaning (Rossen, Klima, Bellugi, Bihrle, & Jones, 1996).

In this article, we evaluate two hypotheses to explain this phenomenon. The first we term the *extra-lexicon* hypothesis, which argues that the origins of unusual vocabulary in WS have their developmental roots outside the system that stores and accesses word meanings and their phonological forms. The second we term the *intra-lexicon* hypothesis, which argues that unusual vocabulary is a consequence of domain-specific anomalies present in the WS lexicon during development, anomalies that limit the way that representations can be structured or accessed. The implication of this account is that such lexicon-specific anomalies must be a target for eventual causal links between genotype and phenotype in this disorder.

The *extra-lexicon* account falls under a broader ''conservative'' hypothesis of language development in WS (see Thomas & Karmiloff-Smith, 2003, for a discussion of the competing hypotheses proposed for language development in WS). The conservative hypothesis is essentially a null hypothesis against which to compare claims of highly domain-specific abnormalities in the WS language system. It argues as follows: Deficits in vocabulary, syntax, and pragmatics in WS are what one might expect for the level of learning disability in these individuals. Any anomalies in the WS language system are a consequence of other features of the disorder. Language development from the earliest age reflects the particular interests of the child with WS, such as the strong desire for social interaction (Jones et al., 2000). Language may initially be used more to mediate these interactions than as a referential tool. Subsequent vocabulary development comes to reflect the special interests of the child who has some degree of learning disability. Under this hypothesis, unusual word usage would be viewed as an expression of pragmatic influences, whereby vocabulary is selected to gain attention and mediate social interaction. In line with this view, studies of narrative production have found that children with WS greatly exceed chronological age-matched children in their elaboration and use of evaluative devices (Jones et al., 2000; Losh, Bellugi, Reilly, & Anderson, 2000; Losh, Reilly, Bellugi, Cassady & Klima, 1997; Reilly, Klima, & Bellugi, 1990). These children demonstrated a particular preference for types of evaluation that serve as social engagement devices such as phrases or exclamations that capture addressee attention. Moreover, somewhat atypically, engagement devices were repeated even when re-telling the same story to the same listener. In line with the conservative hypothesis, Losh et al. (2000) viewed the use of engagement devices in language production as a direct reflection of the WS profile of excessive sociability.

The extra-lexicon account does not maintain that the development of the WS lexicon is normal or simply delayed. Indeed, knowledge of lexical-semantic categories may be more superficial, with a greater emphasis on accumulated knowledge instead of abstraction, a pattern characteristic of low IQ (Spitz, 1982). Johnson and Carey (1998) suggested that this pattern characterised conceptual knowledge in WS. Similarly, if the whole cognitive system is inefficient, the lexicon too may function more slowly or inaccurately. Furthermore, some apparently more specific anomalies of WS vocabulary may also be the consequence of extra-lexicon influences. For example, a problem in spatial vocabulary might result from the visuo-spatial deficit in WS (Phillips, Jarrold, Baddeley, Grant & Karmiloff-Smith, 2004). The conservative account thus predicts some degree of atypicality in the language system.

The conservative hypothesis is not without its limitations (Thomas & Karmiloff-Smith, 2003). In particular, it must explain why individuals with WS should show errors in, for example, morphosyntax that are not found in typically developing children (Capirci, Sabbadini, & Volterra, 1996); and why, if the predominantly successful path of language acquisition in WS is tied to their level of learning disability, individuals with other genetic syndromes with comparable learning disabilities do not follow this successful pathway. The conservative hypothesis would presumably place the onus for explaining such dissociations on features of the other genetic disorders (e.g., the less successful language acquisition in Down syndrome (DS) compared with WS might be explained by phonological deficits in DS but not WS; see McDonald, 1997, for discussion). Importantly, the conservative hypothesis makes a clear prediction regarding the role of word frequency in the WS lexicon. Unusual vocabulary is a consequence of the *criterion used to access lexical items during production*, not due to anomalies in the way they are encoded or retrieved. Frequency should play the same role in structuring the WS lexicon as it does in normal development.

The *intra-lexicon hypothesis*, by contrast, sites the cause of unusual vocabulary in WS within the structure and/or activation dynamics of the lexicon. The *intra-lexicon* account falls under the broader "semantics-phonology imbalance" hypothesis of language development in WS, in which the trajectory of development is thought to be deflected by atypical constraints specific to the language system (see Thomas & Karmiloff-Smith, 2003, for discussion). The broader hypothesis can be split into sub-hypotheses regarding the nature of the imbalance, be it a relatively greater influence of phonology in WS language development, or a relatively weaker influence of lexical semantics, or the possible atypicality of both. In the context of unusual vocabulary use in WS, the *intra-lexicon* hypothesis proposes that the cause is a domain-specific abnormality in the cognitive processes underlying the storage or retrieval of words and their meanings. Rare words emerge in spontaneous speech because frequency has not been encoded into the lexicon in the normal way (for example, because the base-rate activation of lexical entries has not been set at a level that appropriately reflects their frequency of usage; see e.g., McClelland & Rumelhart, 1981); or because the lexical retrieval process has developed inadequately to the extent that it mistakenly retrieves not-quite-appropriate, lower-frequency words in a given semantic context. Proposals of this sort include the following claims: (1) that word retrieval is deviant in WS (Bellugi et al., 1990); (2) that while word knowledge may be well organised, inhibitory activation dynamics which integrate current contextual information are abnormal in lexical access (Rossen et al., 1996); (3) that lexical access is fast but sloppy in WS, with inadequately specified semantic representations (Temple, Almazan, & Sherwood, 2002), and more precisely, (4) that children with WS cannot access fine-grained semantic features of lexical items (Clahsen, Ring, & Temple, 2003). Claims of domain-specific atypicalities in the WS lexicon are often associated with a broader perspective of WS as a developmental fractionation of the normal cognitive—and in this case—language system, whereby the WS system is viewed in terms of the architecture of the normal system but with selective components of the system under- or over-developed (Temple & Clahsen, 2002; see Karmiloff-Smith, 1998, and Thomas & Karmiloff-Smith, 2002, for discussion). The proposal under the intra-lexicon hypothesis is that the lexicon develops atypically, against a background of independent, normal development in other components, such as those involved in processing grammar (Clahsen & Almazan, 1998) or phonology (e.g., Barisnikov, Van der Linden, & Poncelet, 1996; see Majerus, 2004, for discussion). In terms of frequency, the intra-lexicon account also makes a clear prediction. Frequency effects in tasks involving lexical access should be atypical, because frequency is either encoded abnormally in the WS lexicon or obscured by a deviant retrieval process.

The difference between the extra-lexicon and intra-lexicon explanations of rare word usage is an important one, because it sets the target level of specificity for the effects of the genetic mutation in WS on the eventual cognitive structures. If we identify (atypical) rare word usage as the outcome of more general structural deficits in the disorder, or deficits in cognitive systems other than the lexicon, then the nature of the ultimate account that links genotype to phenotype will be different from one that involves precise and domain-specific anomalies in the WS lexicon.

Methodologies to explore the WS lexicon

The extra- and intra-lexicon hypotheses have been driven by empirical data emerging from studies using five experimental paradigms: (1) semantic fluency, (2) elicited word associations and definitions, (3) semantic priming, (4) short-term memory and (5) naming. The paradigms have differed in the extent to which they reflect lexical choice or lexical structure. It should be noted that the strength of the claims has sometimes been undermined by use of small participant numbers in some of the studies. The main results from each methodology, along with the participant numbers used in each study, are shown in Table 1.

Viewed as a whole, the evidence for modulated frequency effects in WS is patchy. Initial claims from the semantic fluency task, where individuals with WS were argued to produce more unusual category members, were not replicated in larger studies. Similarly, findings from the verbal short-term memory task initially suggested a reduced influence of frequency in WS word recall. However, despite evidence that lexical-semantics do influence WS word recall in this task, the unusual frequency effects were then not found in a more carefully controlled study.

The different experimental paradigms also have limitations in their ability to reveal lexical structure and the dynamics of access. For example, performance in the semantic fluency task is in part driven by strategies of retrieval, which imply a contribution of executive functioning. Increased levels of repetition in the WS group suggest their fluency performance may be linked to these demands (Jarrold et al., 2000; Stevens, 1996). Indeed, as typically developing children get older, despite the fact that they know more low frequency words, their lists tend to include *fewer* low frequency words (Temple et al., 2002)—consistent with changes in retrieval strategies with increasing age. The word definition task requires additional meta-linguistic ability compared with fluency or naming (Snow, 1990; Snow, Cancino, De Temple, & Schley, 1990), which may lead to poor performance in individuals with learning disabilities. While semantic priming might offer a more sensitive probe of lexical organisation, frequency effects have yet to be explored, and current findings have not pointed to anomalous organisation in WS compared with CA controls. The anomalous findings have been reported in the ERP correlates of semantic priming in the form of exaggerated LPC and N400 components to congruous and incongruous targets respectively, but the functional consequences of these markers are unclear.

In this paper, we focus on frequency effects in naming times and naming accuracy. Interestingly, Temple et al. (2002) have argued that naming error patterns in their children with WS might reveal lexical atypicalities, since there was a tendency for these children to name part of the object pictured

TABLE 1

Methodologies to explore the WS lexicon: Main empirical results compared to chronological age (CA), mental age (MA), or Down syndrome (DS) comparison groups. (BPVS = British Picture Vocabulary Scale; BAS = British Abilities Scale).

Methodology	*Study*	*Participant numbers*	*Main finding*
Semantic fluency	Rossen et al. (1996)	WS = 7	More low frequency words produced compared to CA matches
	Temple et al. (2002)	WS = 4	More low frequency words produced compared to MA matches
	Jarrold et al. (2000)	WS = 13	
	Pezzini, Vicari, Volterra, Milani, & Ozella (1999)	WS = 18	No difference in frequency or typicality of words produced compared to MA matches
	Scott, Mervis, Bertrand, Klein, Armstrong, & Ford (1995) (see Mervis et al. (1999)	WS = 12	
	Stevens (1996)	WS = 15	
Word definitions	Bellugi, Marks, Bihrle & Sabo (1988)	WS = 7	No difference between WS, DS, and MA-matched controls
	Rossen et al. (1996)	WS = 10	WS gave less frequent meaning of homonym (e.g., BANK) more often than DS and MA matches
Semantic priming	Tyler, Karmiloff-Smith, Voice, Stevens, Grant, Udwin, Davies & Howlin (1997)	WS = 12	No difference in semantic priming between WS and CA matches
	Neville, Mills & Bellugi (1994)	WS = 8	Left hemisphere 'LPC' and N400 ERP correlates of auditory semantic priming exaggerated compared to CA matched controls
Verbal short-term memory	Vicari et al. (1996)	WS = 12	Reduced frequency effect in recall of word lists compared to MA controls
	Brock, McCormack & Boucher (in press)	WS = 14	No difference in frequency effects in recall of word lists compared to vocabulary-matched controls
	Laing, Grant, Thomas, Parmigiani, Ewing & Karmiloff-Smith (in press)	WS = 14	Evidence of influence of lexical semantics on recall of word lists
Naming ability	Temple et al. (2002)	WS = 4	Test age on standardised naming task (BAS sub-test) 3 years behind test age on receptive vocabulary test (BPVS)
	Laing et al. (2001)	WS = 15	No difference in picture naming ability compared to BPVS-matched control group
	Bello et al. (2004)	WS = 10	No difference in naming accuracy on Boston Naming Test compared to MA-matched controls
	Thomas & Ansari (unpublished data)[1]	WS = 15	Naming test age (BAS) in line with receptive vocabulary (BPVS)
	Lukács et al. (2001)	WS = 15	Evidence of accuracy frequency effects in picture naming (no exact CA or MA matched controls)

rather than the whole object. Moreover, Temple et al. reported that two of their participants who were able to complete a speeded naming task produced naming times that were faster than MA-matched controls, consistent with unusual lexical access (although these data were also consistent with different speed-accuracy trade-offs in the atypical group). In terms of frequency effects, while these are a salient feature of naming in typically developing children (Newman & German, 2002), they have not been systematically examined in WS. Lukács, Racsmány, and Pléh (2001) offered preliminary data for a picture-naming task with Hungarian children with WS in which children produced fewer errors on high frequency words than low, but there was no specifically matched control group to assess whether the frequency effect was of the expected size given either the participants' chronological or mental ages.

Speeded naming

We employed a speeded picture naming technique to assess frequency effects in lexical access in WS. This was for two reasons. The first is that it is an *implicit* task, able to circumvent possible confounds involving the reduced meta-cognitive abilities of individuals with WS. The second is that it is a technique that has proved sufficient to demonstrate atypical functioning of the lexicon in other developmental disorders, including children with developmental dyslexia, children with poor comprehension, and children with word finding difficulties (WFD). In particular, it appears to be an appropriate tool to address the issue of frequency effects.

Implicit vs. explicit tasks. Karmiloff-Smith, Tyler, Voice, Sims, Udwin, Howlin, & Davies (1998) have argued that implicit and explicit tasks measure different aspects of language processing in WS (see Tyler, 1985, 1992, for general arguments). Implicit tasks are those where the participant's response is closely time-locked to a relevant linguistic variable and where the participant's attention is not explicitly drawn to the linguistic variable. Explicit tasks focus instead on the controlled aspects of processing and are not concerned with the real-time analysis of spoken language. This distinction has proved useful in understanding both normal development and acquired cognitive impairments, with performance on implicit tasks revealing earlier development in children, and language-impaired adult patients usually showing proficient implicit but impaired explicit processing. Karmiloff-Smith et al. (1998) argued that examination of implicit processing may better reveal the nature of language processing in WS, and used such a paradigm to explore syntax comprehension in the disorder. In the current context, we employed a speeded naming paradigm in which naming time was our time-locked variable, and in which the participant was unaware of the manipulation of word frequency. The aim was to avoid possible meta-linguistic influences on the naming of high and low frequency items. Our subsidiary interest was whether manipulations of word category in speeded naming would also have the same effect in WS as in controls, as a window on semantic structure in the WS lexicon.

Speeded naming in developmental disorders. This technique has already proven sufficient to demonstrate atypical functioning of the lexicon in several other developmental disorders. In these cases, empirical differences have been traced back to structural anomalies in a theoretical model of picture naming that distinguishes several stages of processing. In serial order (of description, though processing may itself be interactive) these stages are: Visual Analysis → Object Recognition → Semantic System → Lexical Selection → Name Retrieval → Speech Output (see e.g., Nation, Marshall, & Snowling, 2001; Temple et al., 2002). Frequency effects in this model are thought to act either at the Name Retrieval stage, or perhaps affect the robustness of the representations at the Lexical Selection stage (a set of representations which incorporates the semantic and syntactic specification of each word) (Nation et al., 2001). Research involving children with *developmental dyslexia* has revealed a disadvantage in naming pictures that have long names, generating a disproportionate number of phonological errors (Nation et al., 2001; Swan & Goswami, 1997). This was hypothesised to reflect a deficit at the Name Retrieval stage (Nation et al., 2001). *Poor comprehenders* (defined according to reading comprehension scores) exhibited normal effects of word length but were slower and less accurate at picture naming than chronological (CA) controls, and produced more visually and semantically unrelated errors (Nation et al., 2001). Poor comprehenders also exhibited larger frequency effects than controls in naming accuracy. This was hypothesised to reflect slow and inaccurate semantic processing operating at the Lexical Selection stage. Impoverished semantic representations were taken either to exaggerate frequency effects at Name Retrieval or to interact with the less robust storage of low frequency word meanings at the Lexical Selection stage (Nation et al., 2001). Research on children with word finding difficulties (WFD) demonstrated no difference in basic naming speed (assessed on highly familiar numerals and letters) or comprehension levels compared with CA controls (Dockrell et al., 2001). However, the WFD group was slower and less accurate in naming semantically driven items (Object and Actions) compared with CA controls. The WFD group exhibited smaller frequency effects in the speed of naming Objects compared with the CA group, driven by slower naming of high frequency items. This pattern was hypothesised to reflect an absence of Name Retrieval problems (evidenced by fast numeral/letter naming) but impoverished semantic representations (Dockrell et al., 2001), with frequency effects attenuated because Lexical Selection can occur prematurely reaches a ceiling speed for high frequency items (Dockrell et al., 2001, Table 4). It is quite possible, then, to reveal anomalies of the functional structure of the lexicon through the study of frequency effects in speeded naming.

Outline of the current study

A speeded picture-naming study was used to assess: (1) the basic speed of naming in WS as a measure of lexical access; (2) the encoding of frequency in the lexicon via its effect on naming times; and (3) the encoding of semantic structure via the effect of semantic category on naming times. Fifty pictures, split into objects, actions, letters, and numerals formed the stimulus set. The pictures were familiar to ensure high levels of accuracy in all groups and permit examination of response times. The study evaluated two predictions. The extra-lexicon hypothesis of rare word usage in WS predicted that frequency and perhaps semantic category should play the same role in structuring the WS lexicon as it does in normal development. The intra-lexicon hypothesis of rare word usage predicted that frequency effects in tasks involving lexical access should be atypical, because frequency is either encoded abnormally in the WS lexicon or obscured by a deviant retrieval process.

Our analyses initially compared a WS group to two control groups, one individually matched for chronological age, another individually matched for receptive vocabulary age. However, given that WS is a developmental disorder, we have argued elsewhere that a developmental perspective is more appropriate than static group comparisons (Karmiloff-Smith, 1998; Karmiloff-Smith et al., 2004). The relevant question for developmental disorders is whether the cognitive system is following the typical developmental *trajectory* or not. A trajectory is a function that links performance level with age. Individually matched control groups, though aiming for a similar outcome, nevertheless end up using analytical techniques that involve only performance levels and ignore age. Cross-sectional trajectories take advantage of the wide age and ability ranges often found in available samples of individuals with rare developmental disorders, instead of collapsing across these wide ranges in generating group means. In addition to matched group comparisons, therefore, we also employed our 32 control participants to construct a cross-sectional developmental trajectory on the speeded naming task from 4 to 54 years to assess the typical relationship between naming latency and age. We then explored whether a different relationship between naming and age emerged in the WS group (see Karmiloff-Smith et al., 2004; Thomas et al., 2001, for a similar analytical approach).

METHOD

Participants

Sixteen adolescents and adults with WS were recruited through the Williams Syndrome Foundation UK. All participants had a clinical diagnosis and where application of the FISH test was practical in this older group,[2] those who had been tested were all positive for deletion of the elastin gene (11 individuals). The sample reflected the usual WS cognitive profile that visuospatial ability (assessed on the BAS pattern construction subtest) was significantly worse than vocabulary ability (assessed on the BPVS). The group comprised 8 males and 8 females, with a mean chronological age of 24;7 (range 12;0–53;0), and a mean GCA on the BAS of 46 (range 39–71). The mean verbal mental age, as assessed by the BPVS, was 11;3 (range 4;4–17;6). Individual participant data are included in Appendix A.

Participants with WS were individually matched to two groups of typically developing children and adults. One control group was matched on chronological age (CA), the other on receptive vocabulary age as assessed by the British Picture Vocabulary Scale (henceforth the "RVA" control group). We chose the BPVS for matching on two grounds. The first was cross-study consistency, as this measure (and its American equivalent, the Peabody Picture Vocabulary Test; Dunn & Dunn, 1981) has been widely used to control for verbal mental age in studies on WS, including those with a significant production element (e.g., word repetition: Majerus et al., 2003; Grant et al., 1997; inflectional morphology: Pleh et al., 2003; Thomas et al., 2001; semantic fluency: Jarrold et al., 2000; narrative production: Losh et al., 2000; reading: Laing et al., 2001). The second was that most previous studies have not shown naming ability significantly out

[1] In a pilot study of our own, we contrasted the performance of a group of children between 5 and 12 years of age who had been clinically and genetically diagnosed with WS (N = 15, mean age 7;5, SD 1;8, range 5;5–11;6) on the British Picture Vocabulary Scale (Dunn, Dunn, Whetton, & Pintilie, 1982) and the naming sub-test of the BAS II (Elliott, Smith & McCulloch, 1996). One participant with WS was at floor on the BPVS and so was excluded from the analysis. Of the remainder, test ages on receptive vocabulary and naming ability were significantly correlated: linear regression: $R^2 = .547$, $F(1, 13) = 14.5$, $p = .003$, with no significant disparity in the test ages (regression intercept not significantly different from zero, $p > .5$), and a gradient not significantly different from 1 (naming test age increased by 1.16 months for each month of increase in receptive vocabulary test age, with lower and upper 95% confidence intervals of .49 and 1.82 months respectively).

[2] That is, in those individuals who did not have fear of needles.

of step with BPVS level. We later discuss the implications of this matching choice. The average difference between the chronological age of each participant with WS and their CA-matched control was 4.1 months. The average difference between the receptive vocabulary test age of each participant with WS and their RVA-matched control was 2.4 months. For the CA group, the mean chronological age was 24;9 (range 12;2 to 53;11). The CA group had a mean test age of 16;3 on the BPVS (range 12;5 to 17;6, 11 participants at the ceiling test age of 17;6). For the RVA group, the mean vocabulary age was 11;10 (range 4;4 to 17;6) and chronological age was 11;5 (range 4;0 to 18;1). Comparisons of the mean CA and RVA of the control groups with those of the WS group revealed no significant differences on these indices.

Materials

The materials for the speeded naming task were based on those constructed by Dockrell, Messer, and George (2001) and comprised a set of 40 coloured drawings of objects, 20 coloured drawings of actions, 5 single digit numerals, and 5 letters. Objects were split into four subcategories of Animals, Body parts, Household items, and Clothes. Actions were split into Transitive and Intransitive verbs. In addition, each subcategory within objects and actions was split into high frequency and low frequency items. Frequency was balanced across subcategories but not across objects and actions. Dockrell et al. (2001) constructed this stimulus set to be appropriate for young children. The stimuli were designed not to form a test of vocabulary knowledge (like the BPVS) but to offer a measure of speed of naming for items that the individuals knew. Therefore, stimuli were constructed to fulfil the following criteria: that objects span both natural kinds and artefacts; that items provide sufficient differentiation in frequency ranges; and that items be familiar and interesting to children. Age-of-acquisition counts were available for 32 of the 40 objects, with a mean age of acquisition of 25.4 months for high frequency items and 38.6 months for low frequency items. It was thus anticipated (and turned out to be the case) that this picture set was appropriate to the level of language development of the WS group, and would permit naming at a high level of accuracy.

Following speeded naming, participants were given a comprehension test on the same stimuli. Participants were required to select the correct picture from a set of four. For each target picture, three foil pictures were included which probed for different kinds of potential comprehension errors. One foil was phonologically related to the target name, one foil was semantically related to the target name, and one was unrelated (Dockrell et al., 2001).

Procedure

Participants were seen individually and carried out the speeded naming task prior to the comprehension task. In all cases, both tasks were completed in a single testing session. Stimuli were presented on a laptop computer. For the speeded naming task, items were presented in a random order within their categories, in the order of: objects, numbers, letters, and actions. For the comprehension task, numbers and letters were presented together in random order, followed by objects and actions together in random order. The target and three foils were presented in random quadrants of the screen. Prior to the object naming, participants were instructed to say the name "as quickly as possible without making any mistakes", and were given six practice trials with feedback if necessary. For the numbers and letters, participants were given two practice trials on each category. For the actions, participants were instructed to say "what the person or animal is doing", and were once more given six practice trials with feedback if necessary. No feedback was offered on test items. The testing session lasted approximately 15 minutes, and all participants in both WS and control groups appeared to find the procedure enjoyable.

The measurement of speeded naming times proceeded as follows. Naming performance was recorded using a digital video camera, which captured both the onset of the visual stimulus, the participant's face and lip movements, and the onset of their vocal response. Frame-by-frame analysis using editing software (EditDV) allowed measurement accuracy to within 40 ms on response times generally in the range of 1000–1500 ms. The video track was used to assess stimulus onset and the audio track to assess response onset. This procedure resulted in minimal loss of data due to measurement error compared to the use of a voice-activated microphone.

After collecting the full set of responses, a set of alternative "acceptable" correct responses was determined for this particular picture set. Alternatives arose under three circumstances: near synonyms (e.g., "riding" for "cycling"; "cutting grass" for "mowing"), diminutives (e.g., "doggie" for "dog", "moo-cow" for "cow"), and ambiguous pictures (e.g., "cuddling" for "stroking", "microwave" for "television"). These forms were defined as alternative correct forms and therefore included in the accuracy analysis but excluded from the reaction time analysis. The CA group produced significantly less alternative forms than WS or RVA groups, while the WS and RVA groups did not differ, CA vs. WS: $t(30) = 2.13$, $p = .041$; CA vs. RVA: $t(30) = 2.35$, $p = .025$; WS vs. RVA: $t(30) = .31$, $p = .761$). These responses formed 7.3%, 5.0%, and 7.6% of the total for WS, CA, and RVA groups, respectively. It should be noted that there was no systematic use of less frequent alternatives in the WS group.

Error analysis

The incorrect responses were coded into a number of different error categories. These were as follows: phonological errors, semantic errors, semantic-specific errors, semantic-general errors, thematic errors, don't know/no response errors, and other errors. Phonological errors were those which preserved either the initial or end pattern of the target item, e.g., "crying" for "crawling", "roll" for "bowl", "brazar" for "zebra". Semantic errors were substitutions that preserved the general features of the meaning of the word, were at the same level of specificity, and were of the same grammatical class, e.g., "lion" for "tiger". Semantic-specific errors were semantic errors which included additional specific but incorrect information, e.g., "thigh" for "leg". Semantic-general errors were substitutions where semantic errors were insufficiently specific, e.g., "chair" for "stool". It was possible for participants to provide answers that were considered both phonological and semantic errors, e.g., "mincing" for "mixing", "bath" for "bowl", in which case responses were coded in both error categories. Thematic errors were responses that described some aspect of the picture, usually closely related to the target word but in a thematic rather than semantic way ("bracelet" for "wrist", "tea" for "cup"). Don't know errors were recorded if the participant indicated verbally or otherwise that they did not know the answer, and amounted to 3.0%, 0.4%, and 0.6% of responses for the WS, CA, and RVA groups respectively. Other errors comprised responses not apparently related to the picture and amounted to 0.8%, 0.4%, and 0.7% of responses for the WS, CA, and RVA groups respectively. Two investigators independently coded the full set of errors, and any disagreements were resolved through discussion with a moderator.

RESULTS

The results are presented in five sections. (1) We first analyse speeded Numeral and Letter naming. Given that these stimuli were highly familiar to all participants, these items gave an indication of the basic speed of naming for each group. (2) Next, we examine naming for Objects and Actions in a single analysis, looking for global evidence of frequency effects. (3) We then look at Object naming in more detail, focusing on the influence of individual subcategories. (4) Similarly, we explore Action naming in more detail, concentrating on the effect of subcategories. (5) Finally, we briefly consider the issue of individual variability in the WS group, again with respect to frequency effects.

Numeral and Letter naming

Individually matched analyses

Comprehension accuracy. Comprehension accuracy was 100% for all groups.

Naming accuracy. Naming accuracy on Numbers was 100% for all groups. Naming accuracy on Letters was 100% for the CA group, 99% for the RVA group (representing 1 error by 1 participant), and 96% for the WS group (representing 1 error by 1 participant and 3 errors by another participant). Performance on these stimuli was so good that the two practice trials were included as test data to increase the validity of the reaction time analyses.

Speeded naming. Speeded naming times were analysed only for correct responses. The data for each participant were cropped if a naming time fell outside two standard deviations either side of the mean of that participant's naming times. For Numbers, this cropped 4.5% of the data points for the CA group, 1.1% for the RVA group, and 0% for the WS group. For Letters, this cropped 1.5% of the data for the CA group, 1.1% for the RVA group, and 1.0% for the WS group. Mean naming times for each group are shown in Table 2. An overall mixed-design ANOVA including group (WS vs. RVA vs. CA) and task (Number naming vs. Letter naming) indicated an overall effect of participant group, $F(2, 45) = 7.46$, $p = .002$, but no significant effect of task nor interaction. Separate ANOVAs comparing the groups revealed that the WS group named Numbers and Letters significantly more slowly than both the RVA and CA groups: WS vs. RVA: $F(1, 30) = 4.45$, $p = .043$; WS vs. CA: $F(1, 30) = 11.19$, $p = .002$, while the RVA group named both stimulus types more slowly than the CA group, $F(1, 30) = 5.81$, $p = .022$. None of the planned pair-wise comparisons revealed a significant difference between Number and Letter naming, and there were no other significant interactions. In short, the WS group exhibited slower basic naming speed compared to receptive vocabulary age matched controls and, for the control groups, basic naming speed was greater in the older control group.

Developmental analysis

Regression analyses were used to build developmental trajectories comparing naming time against age for Numbers and Letters. Control participants were merged and a log-log transform was employed to linearise the data, since response times follow a power law with increasing age (Cohen, Dunbar, & McClelland, 1990). The trajectories linking

TABLE 2

Mean scores for comprehension accuracy (%), naming accuracy (%), and naming time (s) for WS, RVA, and CA groups.

Comprehension accuracy (%)

		Objects										Actions							
		Animals		Bodyparts		Household		Clothes		**Mean**		Transitive		Intransitive		**Mean**			
Group		HF	LF	HF	LF	HF	LF	HF	LF	**HF**	**LF**	HF	LF	HF	LF	**HF**	**LF**	**Numbers**	**Letters**
WS	Mean	100	95	100	93	99	95	93	95	**98**	**94**	100	95	98	99	**99**	**97**	**100**	**100**
	SD	–	9	–	12	5	12	10	12	**6**	**11**	–	12	7	5	**5**	**9**	**–**	**–**
RVA	Mean	99	96	100	100	99	100	98	98	**99**	**98**	100	98	100	100	**100**	**99**	**100**	**100**
	SD	5	8	–	–	5	–	7	7	**5**	**5**	–	7	–	–	**–**	**5**	**–**	**–**
CA	Mean	100	100	100	100	100	100	100	100	**100**	**100**	100	100	100	100	**100**	**100**	**100**	**100**
	SD	–	–	–	–	–	–	–	–	**–**	**–**	–	–	–	–	**–**	**–**	**–**	**–**

Naming accuracy (%)

		Objects										Actions							
		Animals		Bodyparts		Household		Clothes		**Mean**		Transitive		Intransitive		**Mean**			
Group		HF	LF	HF	LF	HF	LF	HF	LF	**HF**	**LF**	HF	LF	HF	LF	**HF**	**LF**	**Numbers**	**Letters**
WS	Mean	96	73	90	72	98	66	83	51	**92**	**65**	84	86	86	69	**85**	**78**	**100**	**96**
	SD	11	26	13	30	7	27	16	23	**13**	**28**	17	20	13	25	**15**	**24**	**–**	**11**
RVA	Mean	96	93	98	91	95	85	90	71	**95**	**85**	95	86	98	89	**96**	**88**	**100**	**99**
	SD	8	14	7	15	10	17	13	22	**10**	**19**	12	19	7	15	**9**	**17**	**–**	**4**
CA	Mean	100	93	99	99	100	94	100	93	**100**	**96**	100	95	99	95	**99**	**95**	**100**	**100**
	SD	–	12	5	5	–	12	–	12	**2.5**	**9**	–	9	5	12	**4**	**10**	**–**	**–**

Naming Time (secs.)

		Objects										Actions							
		Animals		Bodyparts		Household		Clothes		**Mean**		Transitive		Intransitive		**Mean**			
Group		HF	LF	HF	LF	HF	LF	HF	LF	**HF**	**LF**	HF	LF	HF	LF	**HF**	**LF**	**Numbers**	**Letters**
WS	Mean	1.21	1.40	1.50	1.81	1.27	1.74	1.60	2.22	**1.39**	**1.79**	1.43	1.65	1.69	1.82	**1.56**	**1.74**	**0.67**	**0.81**
	SD	0.38	0.46	0.61	0.75	0.33	0.70	0.64	1.16	**0.52**	**0.84**	0.40	0.50	1.67	1.35	**1.20**	**1.01**	**0.17**	**0.51**
RVA	Mean	1.06	1.25	1.19	1.50	1.13	1.50	1.16	1.82	**1.14**	**1.51**	1.16	1.19	1.06	1.26	**1.11**	**1.22**	**0.53**	**0.57**
	SD	0.30	0.36	0.35	0.52	0.46	0.58	0.47	0.71	**0.39**	**0.58**	0.30	0.28	0.29	0.31	**0.30**	**0.29**	**0.10**	**0.19**
CA	Mean	0.88	0.90	0.96	1.04	0.86	1.00	0.94	1.14	**0.91**	**1.02**	0.93	0.97	0.90	1.07	**0.91**	**1.02**	**0.45**	**0.47**
	SD	0.19	0.17	0.23	0.22	0.14	0.18	0.21	0.29	**0.19**	**0.23**	0.17	0.14	0.18	0.21	**0.17**	**0.18**	**0.08**	**0.07**

performance and age only accounted for small proportions of the variance, indicating that the reaction time data were noisy. The typical developmental trajectory revealed a significant relationship between age and naming time for both Numbers and Letters: Numbers: $R^2 = .200$, $F(1, 30) = 7.49$, $p = .010$; Letters: $R^2 = .223$, $F(1, 30) = 8.64$, $p = .006$. By contrast, CA did not predict naming times in the WS group at all: Numbers: $R^2 = .000$, $F(1, 14) = .00$, $p = .999$; Letters: $R^2 = .000$, $F(1, 14) = .00$, $p = .980$. The relationship between WS receptive vocabulary age and naming times showed a trend in both analyses: Numbers: $R^2 = .225$, $F(1, 14) = 4.06$, $p = .064$; Letters: $R^2 = .206$, $F(1, 14) = 3.64$, $p = .077$. It seemed likely that response times in the WS group were driven more strongly by language ability than chronological age. This was confirmed by a stepwise regression that initially only included CA as a predictor of naming time in the WS group. For both Numbers and Letters, addition of RVA as a predictor led to a significant increase in the proportion of response-time variance explained: Numbers: change in $R^2 = .352$, $F(1, 15) = 7.68$, $p = .016$; Letters: change in $R^2 = .330$, $F(1, 15) = 6.65$, $p = .023$. By contrast, starting with RVA and adding CA produced no significant increase in explained variance. Figure 1 depicts the respective trajectories relating performance to age in the typically developing groups and in the WS group according to CA and RVA, respectively. When the WS trajectory plotted against RVA was compared with the typically developing trajectory, in contrast to the individually matched analysis, there was no significant group difference either for Numbers or Letters (general linear model predicting log(RT) from log(Age) with group as a between-participants factor; CA was used as a surrogate for RVA for the control group, with ages cropped at the ceiling test age for BPVS; $p > .25$ for main effects of group and interactions of group and age for both Numbers and Letters).

Object and Action naming

Individually matched analyses

Comprehension accuracy. Comprehension accuracy was above 94% for all groups on Objects and Actions of high and low frequency, with the CA group at ceiling. An overall mixed-design ANOVA including group (WS vs. RVA vs. CA), category (Object comprehension vs. Action comprehension), and frequency (high vs. low) indicated an overall effect of participant group, $F(2, 45) = 11.05$, $p < .001$). There was an overall trend towards an effect of frequency, $F(1, 45) = 3.83$, $p = .056$), but analysed individually, no group showed a significant frequency effect in comprehension accuracy. Separate ANOVAs comparing the groups revealed that the WS groups had significantly lower accuracy than the RVA group, 97% vs. 99%, $F(1, 30) = 6.31$, $p = .018$, and the RVA group was in turn less

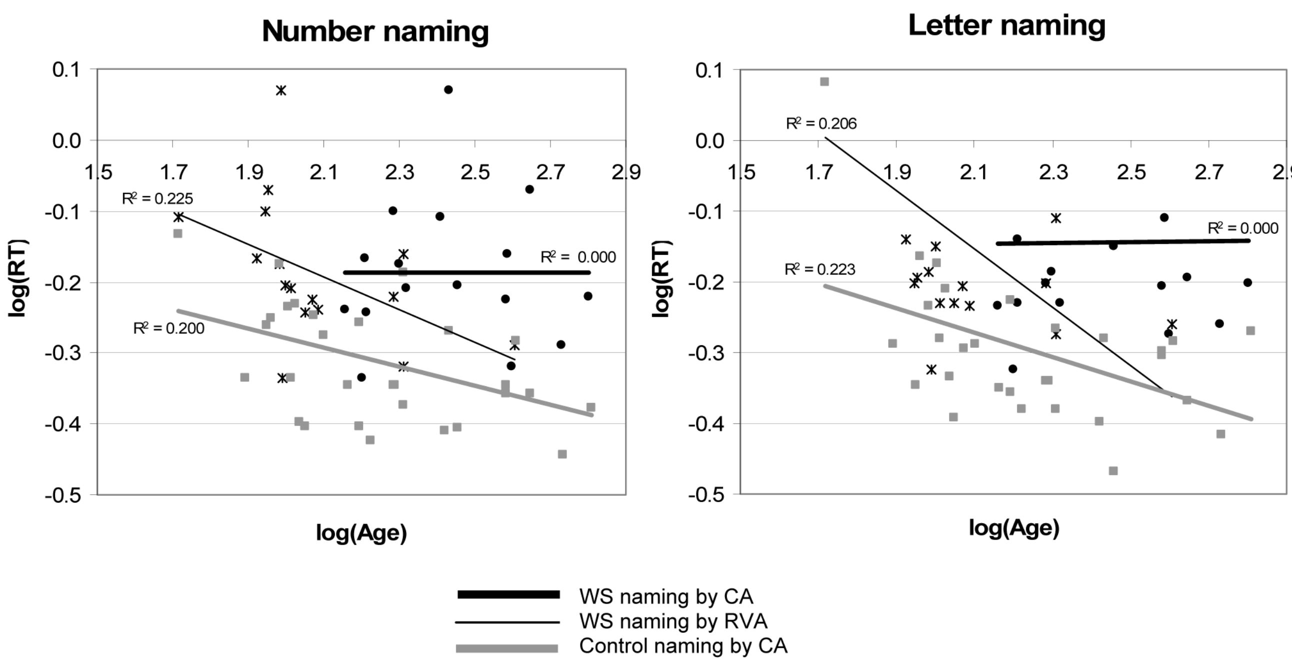

Figure 1. Cross sectional developmental trajectories of response times for WS ($N = 16$) and control groups ($N = 32$) for number naming and letter naming. Trajectories for WS are either plotted against chronological age (CA) or receptive vocabulary age (RVA). R^2 = proportion of variance accounted for by each line.

accurate than the CA group, 99% vs. 100%, $F(1, 30) = 10.97$, $p = .002$. Comprehension errors were too few to justify an analysis of error types, although these errors tended to be semantic rather than phonological or unrelated, in both the WS and RVA groups. Comprehension scores are shown in Table 2, top panel.

Naming accuracy. Following Dockrell et al. (2001), naming accuracy was assessed only for those stimuli that a participant had correctly identified in the comprehension test, so that response accuracy and latency would correspond to "known" items. Mean Object and Action naming accuracy scores are depicted in Table 2, for high frequency (HF) and low frequency (LF) items. A mixed-design ANOVA including group (WS vs. RVA vs. CA), category (Object naming vs. Action naming) and frequency (high vs. low) indicated an overall effect of participant group, $F(2, 45) = 17.72$, $p < .001$. Planned pair-wise comparisons revealed that the WS group named fewer items correctly than the RVA group, $F(1, 30) = 12.29$, p = .001, and the RVA group in turn named fewer items correctly than the CA group, $F(1, 30) = 6.74$, $p = .014$. In the overall analysis, Actions were named more accurately than Objects, $F(1, 45) = 6.62$, $p = .013$, and high frequency items more accurately than low frequency, $F(1, 45) = 49.96$, $p < .001$. In addition, the frequency effect was greater for Objects than Actions, $F(2, 45) = 4.39$, $p = .018$. Note that since frequency was balanced within the Object and Action subcategories but not between Objects and Actions as a whole, the overall category effect must be interpreted with caution. However, the category-by-frequency interaction was predominantly driven by the WS group, who showed a particularly exaggerated frequency effect for the Objects compared to the Actions (3-way interaction of group $\times$ frequency $\times$ stimulus type: $F(2, 45) = 5.89$, $p = .005$. The WS group did, therefore, exhibit an atypical frequency effect in the accuracy of their object naming, but it was an exaggerated rather than attenuated frequency effect.

An ANOVA comparing comprehension and naming accuracy revealed that naming accuracy was lower than comprehension accuracy, $F(1, 45) = 88.39$, $p < .001$. This difference did not interact with stimulus type. However, the difference was more marked in the WS group than both RVA and CA groups: interaction of task $\times$ group: WS vs. RVA: $F(1, 30) = 9.08$, $p = .005$; WS vs. CA: $F(1, 30) = 20.48$, $p < .001$, and more marked in the RVA than CA group, albeit with marginal significance, $F(1, 30) = 4.10$, $p = .053$. The fact that the CA group was close to ceiling on both measures is the likely source of this interaction.

Ceiling effects might also explain the three-way interaction between group, frequency, and stimulus type, given that it was the WS group (the poorest performers) who demonstrated an exaggerated frequency effect for Objects (the more difficult type). However, ceiling effects in accuracy levels for the CA group were in part by design, since we were more interested in the on-line speeded measure and we wanted to ensure that the vocabulary chosen was within the ability range of our participants (in line with Dockrell et al.'s 2001 study of children with word-finding difficulties). We therefore place more weight on the frequency differences in response time data.

Speeded naming. Naming times were analysed only for correct responses and the data were cropped at 2 standard deviations from each participant's mean naming time on the category. For Objects, this eliminated 4.3% of CA data points, 4.2% of RVA data points, and 4.4% of WS data points. For Actions, this eliminated 5.8% of CA data points, 5.2% of RVA data points, and 4.3% of WS data points. The mean naming times for Objects and Actions split by frequency are included in Table 2. An overall mixed-design ANOVA including group (WS vs. RVA vs. CA), category (Object naming vs. Action naming) and frequency (high vs. low) indicated an overall effect of participant group, $F(2, 45) = 11.22$, $p < .001$. As before, the WS group was slower than the RVA group: pair-wise ANOVA: $F(1, 30) = 5.03$, $p = .032$, and the RVA group was slower than the CA group: pair-wise ANOVA: $F(1, 30) = 12.04$, $p = .002$. In the overall analysis, there was no significant difference in the naming times of Objects and Actions and no effect of participant group on this dimension ($p > .5$ and $p > .3$ respectively). There was an overall effect of frequency, with high-frequency items named more quickly than low-frequency items, $F(1, 45) = 58.34$, $p < .001$. This verifies the validity of the frequency manipulation built into the stimulus set. As with the accuracy data, the frequency effect was larger for Objects than it was for Actions: frequency by category interaction: $F(1, 45) = 8.32$, $p = .006$. In addition, the frequency effect differed across groups: frequency by group interaction: $F(2, 45) = 3.91$, $p = .027$. This interaction stemmed from a reduced frequency effect in the CA group, as their naming speeds asymptoted across all stimulus types. From the perspective of our main hypotheses, the most important comparison was whether the WS group demonstrated a different sized frequency effect than would be expected for their level of language ability. A comparison of WS and RVA groups demonstrated no modulation of the frequency effect: $F(1, 30) = .224$, $p = .640$.

Perhaps the only difference between the WS and RVA groups was in basic naming speed? To evaluate this idea, we took each individual's mean naming time on Numerals and Letters—a set of highly familiar stimuli—as an index of their best naming time, and repeated the overall Object-Action analysis dividing each participant's response times by this mean time. Naming times for Objects and Actions were therefore measured in units of the individual's own best naming speed. With differences on basic naming speed factored out, the main effect of frequency and the interaction of frequency and category remained. However, there was now no significant main effect of group or interactions involving group in a comparison of WS and RVA groups. Comparing WS and CA groups, there was a trend of a group × frequency interaction ($p = .067$), stemming from a smaller frequency effect in the CA group.

Given the ceiling effects in the accuracy data, it is possible that if the WS group were faster and/or more able than the language-ability-matched control group, they might have shown more "squeezed out" frequency effects than that group, and therefore apparently smaller modulation by frequency. However, these analyses indicate that the WS group was slower than the RVA group and there was no differential modulation by frequency in the two groups.

Developmental analysis

Developmental trajectories for naming times against age are plotted in Figure 2. This figure plots, respectively, the functions linking performance and age for high frequency Objects, low frequency Objects, high frequency Actions, and low frequency Actions. For the WS group, separate functions were derived to link performance with CA and performance with RVA. As before, data underwent a log-log transform since the relationship between RT and age approximated a power law (Cohen et al., 1990). Trajectories were initially analysed separately by group. A fully factorial, repeated measures ANOVA was employed to compare the four trajectories in each group (see Karmiloff-Smith et al., 2004).[3] Beginning with the normal developmental trajectory (Figure 2a), it is evident that the speed of picture naming increased with age: main effect of age, $F(1, 30) = 25.39$, $p < .001$. Low frequency pictures were named more slowly than high frequency, $F(1, 30) = 13.74$, $p = .001$. This was more marked in Object naming than Action naming for our stimulus set: interaction of category and frequency: $F(1, 30) = 19.18$, $p < .001$. The frequency effect attenuated with increasing age for Objects but stayed roughly constant for Actions (3-way interaction of category × frequency × age: $F(1, 30) = 16.96$, $p < .001$. Turning to the WS group, analysis of the trajectory of naming time against CA produced no significant effects: picture naming times were simply not predictable from CA in this population (Figure 2b). By contrast, Figure 2c demonstrates that receptive vocabulary age successfully predicted naming

[3] This is equivalent to a repeated-measures linear regression design. In contrast to a repeated measures ANOVA, residuals are calculated from each data point to a regression line rather than to a mean.

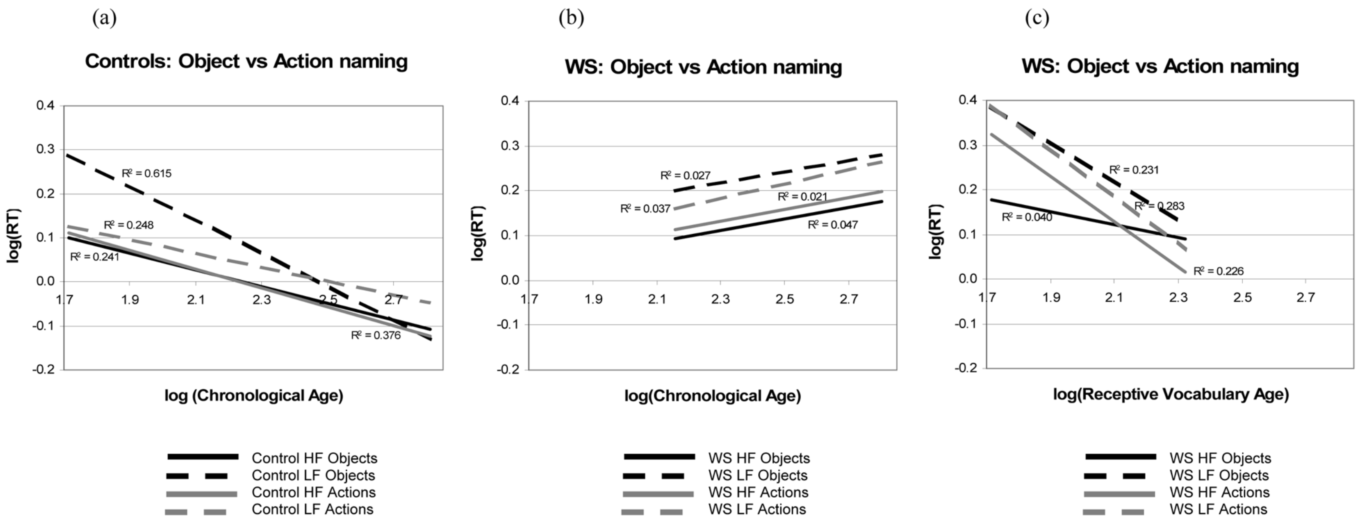

Figure 2. Cross-sectional developmental trajectories for Object and Action picture naming times, split by frequency. (a) Trajectory for typically developing control group ($N = 32$); (b) trajectory for WS group ($N = 16$), plotted against chronological age; (c) trajectory for WS group, plotted against receptive vocabulary age. R^2 = proportion of variance accounted for by each line (for clarity, individual data points are not shown).

time for the WS group, $F(1, 14) = 5.06$, $p = .041$, and produced a pattern not markedly dissimilar from the typical trajectory. There was now a strong trend of a frequency effect, $F(1, 14) = 4.18$, $p = .060$, although the 3-way interaction with category and age did not approach significance, $F(1, 14) = 1.17$, $p = .298$.

When the typical and WS-against-CA trajectories were compared by including group as a between-participants factor in a mixed-design ANOVA, only three effects reached significance. There was a frequency effect that interacted with category, $F(1, 44) = 4.12$, $p = .048$, but was not modulated by group; the WS group was slower, $F(1, 44) = 4.32$, $p = .044$, and did not show the same reduction in naming time with age as the control group, $F(1, 44) = 6.81$, $p = .012$. When the typical and WS-against-RVA trajectories were compared, there was an overall effect of frequency, $F(1, 44) = 14.62$, $p < .001$, which became smaller with increasing age: interaction of frequency and age: $F(1, 44) = 7.64$, $p = .008$. The frequency effect was larger for Objects (interaction of frequency and category: $F(1, 44) = 10.50$, $p = .002$, but this disparity also reduced with age: frequency $\times$ category $\times$ age: $F(1, 44) = 8.24$, $p = .006$. Importantly, group modified neither the main effect of frequency nor its interactions with age and category; the WS trajectory showed the same sensitivity to frequency (all interactions $p > .5$). By contrast, category did show a differential effect across group: the WS group exhibited much slower naming of Actions at younger ages compared with the typical trajectory, but this disparity disappeared with age: group $\times$ category $\times$ age: $F(1, 44) = 4.81$, $p = .034$. This might be taken as evidence for a difficulty in producing verbs over nouns in the WS group at lower levels of language ability.

Lastly, we were concerned that receptive vocabulary age might overestimate the productive language ability of the WS group, if receptive language is a particular strength for them. Although our pilot data on 15 children with WS indicated no significant disparity between British Abilities Scale naming sub-test and British Picture Vocabulary Scale test ages (see Footnote 1) and Bello et al. (2004) found naming ability on the Boston Naming test to be in line with MA, Temple et al. (2002) have nevertheless reported BAS naming ages for 4 children with WS that were 3 years lower than their receptive vocabulary ages on the BPVS. We therefore reran our analyses, docking respectively 1, 2, 3, and 4 years from the verbal mental ages of our WS group, to evaluate their match to a younger section of the normal developmental trajectory. This can only be a rough comparison, because it assumes that any discrepancy between BPVS and naming ability remains constant across development. However, using this manipulation, the main effect of group moved from an (initially non-significant) p-value of .218 using RVA matching, to .279 at RVA minus 1 year, to .375 at RVA-2, .548 at RVA-3, and .962 at RVA-4, creating a greater overlap of WS and control naming times. Reducing the verbal mental age did not produce any significant interactions of group and frequency in any of the four analyses, while the interaction of category, group, and age remained significant in all of them. This is at least suggestive that the normal sensitivity of the WS group to frequency effects was not a consequence of a possible overestimation of their verbal mental age with the BPVS.

Error patterns. Errors were split into the following categories for analysis: (P) phonological, (S) semantic, (S/S) semantic-specific, (S/G) semantic-general, (DK) don't know, (T) thematic, and (O) other. Figure 3 depicts the total number of each type of error in each of the three groups and the proportion of each type of error for those participants who made errors. Statistical comparisons must be interpreted with care, since Figure 3 indicates the different error types have different variances. For this reason, we do not report analyses of variance and instead rely on multiple *t*-tests for each error type. With regard to the overall error counts, 42 post-hoc comparisons were carried out to assess counts of each error type (3 groups $\times$ 7 error types $\times$ 2 stimulus types). Significant differences are shown in Figure 3. A more conservative Bonferroni correction left only three differences that reached statistical significance: the WS group produced more (S) errors for Objects ($t = 3.70$, df = 30, $p = .001$) and more (T) errors for Actions than the CA group ($t = 5.86$, df = 30, $p < .001$); and more (T) errors for Actions than the RVA group ($t = 4.15$, df = 30, $p < .001$).

Turning to the arguably more informative error proportions, post-hoc *t*-tests were again used to compare the proportions of each error type between each group. These indicated no significant differences in phonological or semantic errors between the WS group and either control group. Compared with the RVA group, the WS group exhibited no reliable differences in any error type.[4] Compared with the CA group, there were only two significant differences: the WS group made proportionally more (T) thematic errors for Actions ($t = 6.61$, df = 21, $p < .001$) and the CA group made proportionally more (O) other errors ($t = 2.33$, df = 21, $p = .030$). When a Bonferroni correction for multiple comparisons was applied, only the difference in (T) thematic errors remained significant. In short, once the greater total number of errors in the WS group was factored out, the WS error pattern did not differ markedly from the RVA group, and exhibited only one reliable difference compared with the CA group.

Thematic errors were particularly prevalent in the WS group during Action naming. Temple et al. (2002) argued that individuals with WS have

[4] Note that total error counts included all participants, while error proportions included only those participants who made errors on Object naming or Action naming. Hence the degrees of freedom can differ for the two.

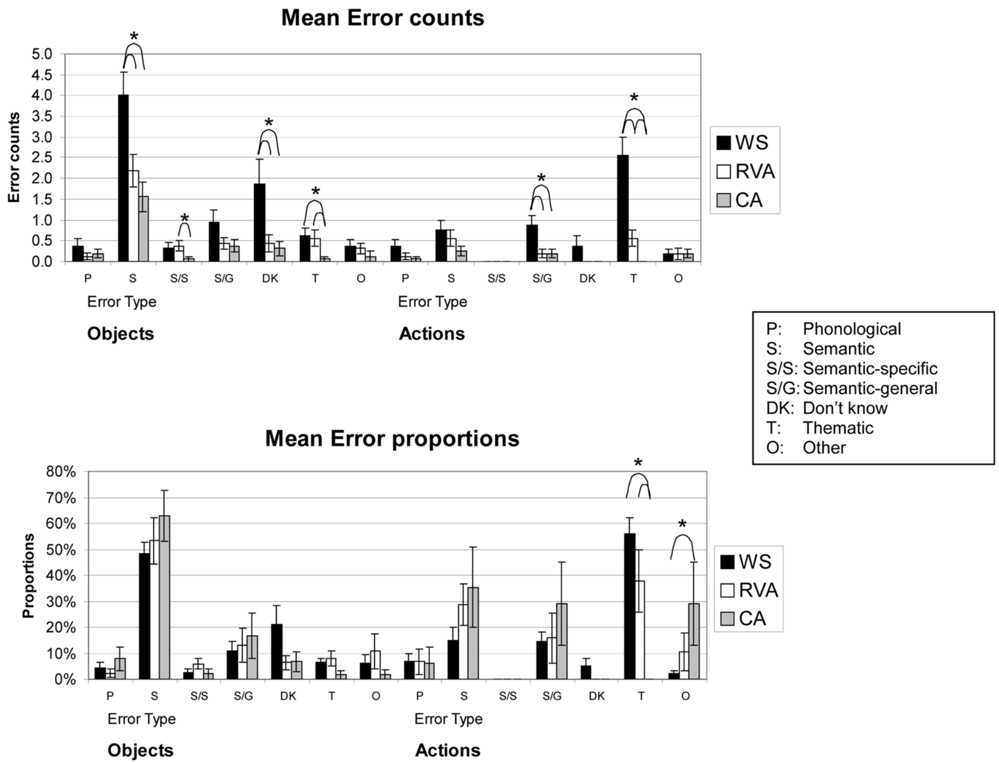

Figure 3. Mean error counts and mean error proportions for the WS group, and the RVA and CA control groups. P = phonological, S = semantic, S/S = semantic-specific, S/G = semantic-general, DK = don't know, T = thematic, and O = other (see text for details).
* = Difference significant at 0.05 level without Bonferroni correction.

a tendency to focus on parts of objects, consistent with their ''local' or ''featural' approach noted in the visuospatial processing. Such part-whole naming errors would have fallen under a thematic classification in our scheme. We therefore looked at the thematic errors more carefully. While some instances of thematic errors in the WS group could be interpreted as part-whole errors, our impression during testing was that individuals with WS were adopting differential strategies for Objects and Actions when they failed to retrieve the target word. When retrieval problems occurred with nouns, participants with WS tended to respond with a ''Don't know''. But where retrieval problems occurred with verbs, thematic errors appeared to reflect substitutions of nouns or adjectives that were appropriate to the picture (e.g., ''cat'' for ''stroking''). This was not due to difficulty in understanding the requirement to produce a verb, since the Action naming accuracy of the WS group was over 70%. Less frequently, individuals with WS produced more general verbs, leading to a greater total count of S/G errors in the WS group on Actions. However, the substitution of a noun or adjective for a target verb occurred in the RVA group as well, albeit to a lesser extent. The full set of thematic errors, along with their distribution across the test groups, is included in Appendix B. Some of these are interpretable either as substitution or part-whole errors. However, the sparseness of the data makes firm conclusions difficult in this respect.

Object naming: Subcategories

In these analyses, for brevity we focus on differences between the WS and RVA groups using an individually matched analysis. Means and standard deviations for Object subcategories are included in Table 2. In terms of comprehension accuracy, mixed-design ANOVAs for each subcategory revealed that the WS group was significantly poorer than the RVA group only in recognising Bodyparts, $F(1, 30) = 5.87$, $p = .022$; that Animals and Bodyparts showed significant frequency effects, $F(1, 30) = 5.00$, $p = .033$; $F(1, 30) = 5.87$, $p = .022$; and that the WS group showed a larger frequency effect in their comprehension accuracy of Bodyparts, $F(1, 30) = 5.87$, $p = .022$. In terms of naming accuracy, planned pair-wise comparison of subcategories revealed that Clothes were named significantly less accurately than Animals, Bodyparts, and Household items, the latter three all named at equivalent levels. This order of difficulty was the same for the WS and RVA groups. The WS group was significantly less accurate in naming Animals, Bodyparts, and Clothes, but not Household items, and produced larger frequency effects for Animals and Household items. When subcategory was included as a 4-level factor, this led to a larger frequency effect for the WS group, $F(1, 30) = 13.60$, $p = .001$. In terms of speeded naming, Animals were named fastest, followed by Household items, then Bodyparts, and finally Clothes (differences between Household Items and Bodyparts, and Bodyparts and Clothes were not significant). However, there were no significant differences between WS and RVA in this order, or in the effects of frequency. In short, where the WS group indicated category-specific modulations of frequency effects compared to the RVA group, these were *larger* effects of frequency (for Bodyparts in comprehension accuracy, for Animals and Household items in naming accuracy). Category-specific differences emerged in a deficit for comprehending Bodyparts in WS, and in all subcategories bar Household items in naming accuracy.

Action naming: subcategories

When comprehension accuracy was evaluated by Action subcategory, there was no effect of the Transitivity dimension on Action comprehension, although Transitive verbs showed a marginally larger frequency effect across the three groups, $F(1, 45) = 8.86$, $p = .005$. This did not interact with participant group. An overall analysis of variance for naming accuracy indicated no significant effect of verb transitivity, interactions of transitivity with group, nor interactions with the robust frequency effect. A similar pattern emerged in the analysis of naming times.

Individual variability

As a syndrome, WS often reveals high levels of individual variability. Given our finding of strong frequency effects in the naming times of our group with WS, we were interested in whether all our participants with WS exhibited such effects. We counted how many participants in each group had exhibited frequency effects in their speeded naming times. This was simply defined as producing a faster mean naming time on high-frequency items than low-frequency items, considered separately for Objects and Actions. For the WS group, 14/16 individuals exhibited frequency effects on Object naming times, compared with 15/16 RVA controls and 14/16 CA controls. On Action naming, 11/16 participants with WS exhibited frequency effects, compared with 11/16 RVA controls and 14/16 CA controls. In short, given the stimuli, frequency effects turned out to be no different across the sample of individuals with WS than they did across the control groups. This argues against the idea that any lexical anomalies are restricted to subgroups of individuals with WS.

DISCUSSION

The appearance of rare words in the productive vocabulary of individuals with WS has been highlighted as a peculiarity of their language. How are we to go about linking this characteristic to the genetic mutation that defines the disorder? We identified two possible accounts to clarify the cognitive status of this phenomenon, the intra-lexicon vs. extra-lexicon hypotheses. Broadly speaking, these claim respectively that frequency is not properly encoded in the WS lexicon, or that rare word usage is primarily an aspect of pragmatics. Previous methodologies for exploring the nature of the WS lexicon (semantic fluency, word definitions, semantic priming, verbal short-term memory, and naming) produced mixed results, often confounded by the meta-cognitive aspects of the task. In the present study, we used an implicit measure of frequency in a speeded picture-naming task, since frequency normally modulates naming times. In addition, we altered the semantic category of pictures (nouns vs. verbs) to offer a window on semantic structure in the lexicon. Our findings are summarised below.

A set of highly familiar items on which there were virtually no errors—letters and numbers—gave us an indication of basic naming speed. The WS group ($N = 16$) proved slower than both their chronological- and receptive-vocabulary-matched control groups. While this superficially disagrees with a previous finding of Temple et al. (2002) based on two individuals with WS (aged 12 and 15) that WS naming can be faster than controls (if less accurate), it should be noted that Temple et al.'s controls were mental-age matched. The receptive vocabulary age of a WS group will be higher than their overall mental age since language is a relative strength in the disorder. As a result, our study compared the WS sample with older controls than in the Temple et al. study. The slower naming speed in WS was not our primary interest, however. Our data cannot indicate whether the longer naming times stem from slower lexical access or from perceptual or motor components of the picture-naming task. For numbers and letters, receptive vocabulary age of individuals with WS predicted their naming speed more strongly than their chronological age. For objects and actions, receptive vocabulary age was a significant predictor of naming speed. However, receptive vocabulary age in our WS group significantly predicted general cognitive ability according to the BAS, $F(1, 14) = 5.06$, $p = .041$, so the closer relationship between naming times and RVA rather than CA is not unexpected.

For the pictures, the WS group was marginally less accurate than the RVA matches on a comprehension test, but these were very small effects and overall levels were very high. The WS group were significantly less accurate than RVA matches on picture naming, a disparity greater than that for comprehension. This indicates a problem with naming, and is consistent with the reports of anomia by Temple et al. (2002). It is inconsistent with our pilot study on a group of 15 children with WS, where naming and receptive vocabulary levels were in line according to two standardised tests (BPVS and BAS naming subtest). This could be due to the age difference of the samples or the lack of sensitivity of the standardised tests used with the younger group. In terms of frequency effects in naming accuracy, the WS group demonstrated *larger* effects compared to RVA and CA groups. Thus the WS group appeared to exhibit an exaggeration of the normal pattern of difficulty on these pictures, rather than attenuation of the frequency effect. The presence of frequency effects in naming accuracy is consistent with a preliminary study of naming in 15 Hungarian children and adolescents with WS (aged 5;9 to 19;6) (Lukács et al., 2001).

Inspection of comprehension and naming accuracy for subcategories of Objects and Actions revealed the same order of difficulty in WS and RVA groups. Some category-specific effects emerged, with the WS group worse in comprehending Bodyparts than RVA (ankles, elbows, and wrists tended to be confused); and in naming, the WS group was significantly worse on Animals, Bodyparts, and Clothes (that is, all but Household items). Frequency effects in the WS group did sometimes differ from the RVA group (comprehension accuracy: Bodyparts; naming accuracy: Animals, Household items), but when they did so, they were always greater in the WS group.

The high levels of accuracy allowed us to focus on our key measure, speed of picture naming, where frequency and semantic category were manipulated as implicit variables. Naming speed indexed dynamics of retrieval and gave us a window onto the role of frequency in the WS lexicon. Naming times were analysed using a static framework (individually matched control groups) and our preferred developmental framework (the construction of cross-sectional developmental trajectories).

The static framework indicated that frequency effects were attenuated in the CA group, consistent with the view that these individuals approach a maximum speed of naming. When the WS group was compared with the RVA controls, frequency produced the same modulation of naming times in each group. This was also the case with the effect of semantic category. The naming of the WS group was simply slower, and if naming speed was factored out (using the mean speed on the highly familiar letters and numbers), then there were no differences at all between WS and RVA groups. Frequency effects in the WS group were subject to a more fine-grained analysis to assess whether a subgroup of individuals with WS might demonstrate attenuation. However, the analysis revealed the same pattern of individual variation in the WS group as in the control groups.

A more theoretically focused question is whether naming develops in the same way in WS as in controls. To evaluate this question, for each group, we constructed cross-sectional trajectories linking performance with age (see Karmiloff-Smith et al., 2004; Thomas et al., 2001). The developmental trajectories for picture naming in the control group revealed that naming became faster with age and trajectories differed depending on frequency. WS naming speed for these pictures showed no relation to CA. When trajectories were plotted against receptive vocabulary age, a more typical pattern emerged. Although naming speed was slower for the WS trajectory, the relationship with frequency was the same as in the typical trajectory. The only difference was an indication of a problem in naming Actions at lower levels of language ability in WS. Anomalous processing of verbs, albeit in receptive syntax, has been reported previously in eight adolescents and adults with WS (Karmiloff-Smith et al., 1998). The construction of trajectories permitted an examination of the sensitivity of group effects to measures of language ability, because it has been argued that receptive vocabulary overestimates the language ability of people with WS. Although some simplifying assumptions were necessary in these analyses, comparisons indicated that the effects of frequency and semantic category on the WS trajectory did not diverge from normal even if receptive vocabulary level overestimated their language ability by 4 years (a year greater than the disparity argued for by Temple et al., 2002).

An analysis of the types of naming errors made by the WS group revealed elevated levels of semantic errors, thematic errors, and don't know errors for Object naming, and more semantic-general and thematic errors for Actions. However, the proportions of the error types did not look markedly atypical (see Bello *et al*., 2004, for similar findings). The excess of thematic errors (a word or phrase thematically related to the picture but not the target word) might have been consistent with a more general ''featural processing'' style in WS, but closer inspection suggested that these were mostly grammatical-class-substitution errors rather than part-instead-of-whole naming preferences.

We can conclude the following: naming in WS is slower and less accurate than would be expected for their receptive vocabulary but, leaving aside a possible (developmentally) early problem with producing verbs, the naming times of the WS group were modulated in the same way by the implicit variables of frequency and semantic category. Frequency effects were sometimes larger in accuracy data, an exaggeration of normal patterns of difficulty. These results argue that rare word usage is not traceable to a lack of encoding frequency in the lexicon: frequency develops in the WS lexicon in the normal way. If one combines this result with evidence of normal patterns of semantic priming in WS (12 adolescents and adults with WS, aged 14;3 to 30;5—Tyler et al., 1997), the indications are that the processing dynamics of the WS lexicon are similar to the normal case. While naming shows deficits in speed and accuracy, the evidence falls in favour of the extra-lexicon explanation of rare word usage in WS, and against domain-specific intra-lexicon effects of the WS genotype on development.

Of course, as with any single measure, there are limits on what the speeded-naming paradigm can reveal. For example, we did not manipulate context, so we cannot rule out a proposal by Rossen et al. (1996) that there is a problem in WS of using context to constrain lexical retrieval. Moreover, we identified the intra-lexicon hypothesis as falling under the broader Semantics-Phonology Imbalance theory of WS language development. That wider theory includes the possibility that the imbalance (unique to WS) involves a particular strength in phonology rather than a particular weakness in lexical semantics. We did not test phonology here, other than demonstrating that phonological naming errors showed the same low levels as in the controls. We cannot, therefore, rule out the possibility that a relatively good phonological short-term memory in WS is instrumental in memorising unusual words.

The intra- vs. extra-lexicon contrast does not map neatly onto an atypical vs. delayed dimension, since neither account proposes representations in the lexicon that are normal. The difference is in whether lexicon itself is the target for syndrome-specific effects. Given that speeded naming data favour the extra-lexicon hypothesis, we conclude by further examining how lexical semantics may be atypical under this account, and then more generally considering the possible relation of rare word usage to pragmatics.

The WS group demonstrated poorer levels of accuracy in naming and, for Objects at least, the errors were predominantly semantic in nature. McGregor (1997) argues that a predominance of errors bearing semantic relations to their targets suggests a robust organization of lexical storage into a network of related information. Further, McGregor et al. (2002) argue that in picture naming tasks, children may make semantic errors for three different reasons. First, they may make semantic errors to fill lexical gaps, that is, representations that are missing from the mental lexicon. Children produce a word that nevertheless captures some features of the picture. Second, children may make such errors not because they lack the relevant representation in the lexicon, but because the representation is fragile: they do not know the target word well enough. Having only partial knowledge means that children cannot choose accurately between the target and other related words and, if a related competitor has a more robust representation, this may be retrieved instead. Third, less often children may make semantic errors as a consequence of temporary access problems, despite having a well-elaborated representation in the lexicon.

With respect to the extra-lexicon account, the representations within the WS lexicon are taken to be poor for domain-general reasons, in line with the lower IQ of the population. Processing across a range of domains may be slow and inefficient. With lower IQ, abstract reasoning is poorer, so one might expect the content of lexical semantics to have a less differentiated similarity structure for abstract properties. Several recent studies have begun to indicate ways in which semantic representations in WS may be poorer. For instance, Temple et al. (2002) found that the format of a receptive vocabulary test such as the BPVS (choose the picture that goes with a target word from a set including three semantically related distracters) could produce problems for children with WS, provided the size of the distracter set was increased and distracters were semantically close. Although children with WS performed better than overall mental-age matched controls on the BPVS, when a large distracter set taken from the Snodgrass and van der Wart (1980) pictures was employed with distracters all taken from the same semantic class as the target, performance fell below that of the controls (Clahsen et al., 2003). This argues for an explanation of naming errors arising from insufficiently delineated semantic representations (see also Johnson & Carey, 1998, for related findings on the absence of abstraction in the development of conceptual knowledge in WS).

What, then, of rare word usage in WS? It is possible that this is a real, if occasional, characteristic of WS language. A typical description of the expressive language of children with WS portrays it as involving ''an over-familiar manner [using] more adult vocabulary and social phrases'' (Udwin and Yule, 1990: 108). Expressive language can include terms that have additional detail that is either unnecessarily specific given the context or specific in a way that is inappropriate given the context (Rossen et al., 1996). It can include frozen and stereotypical phrases, clichés, idioms, and even figurative language that again are not exactly appropriate to the context (Bertrand, Mervis, Armstrong, & Ayers, 1994; Udwin & Yule, 1990). Our focus has been the implication of this behaviour to the functioning of underlying language systems, and the level of specificity of potential anomalies. Given that rarity appears to be encoded in the normal way in WS lexicon, we finish with one caveat and two conclusions.

First the caveat. Our manipulations of frequency have used objective measures. High- versus low-frequency items were selected on the basis of adult-based corpus counts, age of acquisition statistics, and the presence of items in three primary grade source books (Dockrell et al., 2001). The frequency manipulation was given validity by its significant effect on naming time in the typically developing control group. However, in terms of processing, it is subjective frequency that counts: the number of times an individual encounters or uses a term. Given the special interests of children with developmental disabilities, their subjective frequencies may differ quite markedly from norms. One example suffices. When we collected semantic fluency data from one 12-year-old child with WS, her first three responses for the category of Animal were highly unusual: ''reptile'', ''amphibian'', and ''mammal''. This list reminded us of the unusual set of animal names highlighted by Pinker when discussing the syndrome (cf., ''unicorn, pteranodon, yak, ibex, water buffalo, sea lion, sabre-tooth tiger, vulture, koala, dragon, and brontosaurus rex''; Pinker, 1994: 53). However, in our case, it turned out that this child's favourite book (one, indeed, that she carried with her to the testing session) was an encyclopaedia of animals. Thus, to demonstrate that rare word usage is rare from the child's own perspective, some measure of the frequency of vocabulary items in the child's environment is desirable.

Second, for children with WS to employ rare words in their conversation requires at least that they have memorised the relevant phonological forms, even if the semantic representations associated with them are not well elaborated. This may implicate a relative strength for phonological short-term memory in WS (Grant et al., 1997; Majerus, 2004; Majerus et al., 2003; Mervis, Morris, Bertrand, & Robinson, 1999; Vicari, Carlesimo, Brizzolara, & Pezzini, 1996; though see Brock, 2002; Brock et al., in press, for discussion). Although the causes of rare word usage may be extra-lexicon, a full account of their acquisition may involve rejecting the null ''conservative'' hypothesis of WS language development and appealing to a domain-specific imbalance of semantics and phonology in the syndrome (see Thomas & Karmiloff-Smith, 2003, for discussion).

Finally, according to the extra-lexicon account, the deployment of rare words in the expressive vocabulary of individuals with WS is primarily pragmatic, such that the retrieval of the word form is intended and serves a particular role of social engagement. A more detailed account of this type of access is still desirable, including the extent to which it involves a conscious strategy. However, we would anticipate that such an account would have much in common with a processing explanation of the elevated use of evaluative devices such as exclamations, sound effects, character speech, and audience hookers revealed in studies of WS narratives (Jones et al., 2000; Losh et al., 1997, 2000; Reilly et al., 1990). Although rare word usage in WS is atypical, the developmental origins of this behaviour appear to lie more in the atypical social profile of these individuals than in domain-specific anomalies of the language system. Such a finding sets new bounds on the domain specificity of the complex developmental pathway between atypical genotype and atypical phenotype in this disorder.

Manuscript received October 2003
Revised manuscript received July 2005

REFERENCES

Barisnikov, K., Van der Linden, M., & Poncelet, M. (1996). Acquisition of new words and phonological working memory in Williams syndrome: A case study. *Neurocase*, *2*, 395–404.

Bello, A., Capirci, O., & Volterra, V. (2004). Lexical production in children with Williams syndrome: Spontaneous use of gesture in a naming task. *Neuropsychologia*, *42*, 201–213.

Bellugi, U., Bihrle, A., Jernigan, T., Trauner, D., & Doherty, S. (1990). Neuropsychological, neurological, and neuroanatomical profile of Williams syndrome. *American Journal of Medical Genetics*, *6*, 115–125.

Bellugi, U. Bihrle, A., Jernigan, T., Trauner, D., & Doherty, S. (1991). Neuropsychological, neurological, and neuroanatomical profile of Williams Syndrome. *American Journal of Medical Genetics Supplement*, *6*, 115–125.

Bellugi, U., Bihrle, A., Neville, H., Doherty, S., & Jernigan, T. (1992). Language, cognition, and brain organization in a neurodevelopmental disorder. In M. Gunnar & C. Nelson (Eds.), *Developmental behavioral neuroscience: The Minnesota Symposia on Child Psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.

Bellugi, U., Korenberg, J. R., & Klima, E. S. (2001). Williams syndrome: an exploration of neurocognitive and genetic features. *Clinical Neuroscience Research*, *1*, 217–229.

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

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

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

Brock, J. (2002) Language and memory in Williams syndrome. Unpublished Doctoral Thesis, University of Warwick, Coventry, UK.

Brock, J., McCormack, T., & Boucher, J. (in press). Probed serial recall in Williams syndrome: Lexical influences on phonological short-term memory. *Journal of Speech Language and Hearing Research*.

Clahsen, H., & Almazan M. (1998). Syntax and morphology in Williams Syndrome. *Cognition*, *68*, 167–198.

Clahsen, H. & Temple, C. (2003). Words and Rules in Williams Syndrome. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations* (pp. 323–352). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Clahsen, H., Ring, M., & Temple, C. (2003). Lexical and morphological skills in English-speaking children with Williams syndrome. In S. Bartke & J. Siegmüller (Eds.), *Williams syndrome across languages*, (pp. 221–244). Amsterdam: Benjamin Publishing Company.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332–361.

Dockrell, J. E., Messer, D., & George, R. (2001). Patterns of naming objects and actions in children with word finding difficulties. *Language and Cognitive Processes*, *16*, 261–286.

Donnai, D., & Karmiloff-Smith, A. (2000). Williams syndrome: From genotype through to the cognitive phenotype. *American Journal of Medical Genetics: Seminars in Medical Genetics*, *97*, 164–171.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary tests – revised manual*. Circle Pines, MN: American Guidance Service.

Dunn, L. M., Dunn, L. M., Whetton, C. & Pintilie, D. (1982). *British Picture Vocabulary Scale*. Windsor, UK: NFER-Nelson.

Elliott, C. D., Murray, D. J., & Pearson, L. S. (1983). The British Abilities Scales Manual 3. Windsor, UK: NFER-Nelson.

Elliott, C. D., Smith, P., & McCulloch, K. (1996). British Ability Scales II. Windsor, UK: NFER-Nelson.

Francis, W. N. & Kucera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston: Houghton Mifflin.

Gentner, D (1988). Metaphor as structure mapping: The relational shift. Child Development, 59, 47-59.

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.

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.

Johnson, S., & Carey, S. (1998). Knowledge enrichment and conceptual change in folk biology: Evidence from Williams syndrome. *Cognitive Psychology*, *37*, 156–184.

Jones, W., Bellugi, U., Lai, Z., Chiles, M., Reilly, J., Lincoln, A., & Adolphs, R. (2000). Hypersociability in Williams syndrome. *Journal of Cognitive Neuroscience*, *12* (Supplement), 30–46.

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., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Grice, S., Brace, N., Van Duuren, M., Pike, G., & Campbell, R. (2003). Exploring the Williams Syndrome Face Processing Debate: The importance of building developmental trajectories. Manuscript submitted for publication.

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.

Laing, E., Grant, J., Thomas, M. S. C., Parmigiani, C., Ewing, S., & Karmiloff-Smith, A. (in press). Love is . . . an abstract word: The influence of lexical-semantics on verbal short-term memory in Williams syndrome. *Cortex*.

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.

Losh, M., Reilly, J., Bellugi, U., Cassady, C., & Klima, E. (1997). Linguistically encoded affect is abnormally high in Williams syndrome children. *International Behavioural Neuroscience Society Abstract*, *6*, 2–53.

Losh, M., Bellugi, U., Reilly, J., & Anderson, D. (2000). Narrative as a social engagement tool: The excessive use of evaluation in narratives from children with Williams syndrome. *Narrative Inquiry*, *10* (*2*), 265–290.

Lukács, A., Racsmány, M., & Pléh, C. (2001). Vocabulary and morphological patterns in Hungarian children with Williams syndrome: Preliminary Report. *Acta Linguistica Hungarica*, *48*, 243–269.

Majerus, S. (2004). Phonological processing in Williams syndrome. In S. Bartke & J. Siegmüller (Eds.), *Williams syndrome across languages* (pp. 125–142). Amsterdam: Benjamin Publishing Company.

Majerus, S., Barisnikov, K., Vuillemin, I., Poncelet, M., & Van der Linden, M. (2003). An investigation of verbal short-term memory and phonological processing in four children with Williams syndrome. *Neurocase*, *9*, 390–401.

McClelland, J. E., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–107.

McDonald, J. L. (1997). Language acquisition: The acquisition of linguistic structure in normal and special populations. *Annual Review of Psychology*, *48*, 215–241.

McGregor, K. K. (1997). The nature of word-finding errors of preschoolers with and without word-finding difficulties. *Journal of Speech, Language, and Hearing Research*, *40*, 1232–1244.

Mervis, C. B., Morris, C. A., Bertrand, J., & Robinson, B. F. (1999). William Syndrome: Findings from an integrated program of research. In H. Tager-Flusberg (Ed), *Neurodevelopmental disorders* (pp. 65–110). Cambridge, MA: MIT Press.

Mervis, C., & 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.

Morris, C. A., Demsey, S. A., Leonad, C. O., Dilts, C., & Blackburn, B. L. (1988). The natural history of Williams syndrome: Physical characteristics. *Journal of Pediatrics*, *113*, 318–326.

Nation, K., Marshall, C. M., & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skill: Evidence from children with developmental reading disorders. *Language and Cognitive Processes*, *16*, 241–259.

Newman, R. S., & German, D. J. (2002). Effects of lexical factors on lexical access among typical language-learning children and children with word finding difficulties. *Language and Speech*, *45* (*3*), 285–317.

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

Pezzini, G, Vicari, S., Volterra, V., Milani, L., & Ossella, M. T. (1999). Children with Williams syndrome: Is there a single neuropsychological profile? *Developmental Neuropsychology*, *15*, 141–155.

Phillips, C. E., Jarrold, C., Baddeley, A., Grant, J., & Karmiloff-Smith, A. (2004). Comprehension of spatial language terms in Williams syndrome: Evidence for an interaction between domains of strength and weakness. *Cortex*, *40*, 85–101.

Pinker, S. (1994). *The Language Instinct*. London: Penguin Books.

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

Reilly, J., Klima, E. S., & Bellugi, U. (1990). Once more with feeling: Affect and language in atypical populations. *Development and Psychopathology*, *2*, 367–391.

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, NY: Cambridge University Press.

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 [special issue]. *Genetic Counseling*, *6*, 172–173.

Snow, C. (1990). The development of definitional skill. *Journal of Child Language*, *17*, 697–710.

Snow, C., Cancino, H., De Temple, & Schley, S. (1990). Giving definitions: a linguistic or metalinguistic skill? In E. Bialystock (Ed.), *Language processing and language awareness by bilingual children*. New York: Cambridge University Press.

Spitz, H. H. (1982). Intellectual extremes, mental age, and the nature of human intelligence. *Merrill-Palmer Quarterly*, *28*, 167–192.

Stevens, T. (1996). *Lexical constraints on language acquisition in Williams syndrome*. Unpublished PhD, MRC Cognitive Development Unit and University College London.

Swan, D. & Goswami, U. (1997). Picture naming deficits in developmental dyslexia: The phonological representation hypothesis. *Brain and Language*, *56*, 334–353.

Tassabehji, M., Metcalfe, K., Karmiloff-Smith, A., Carette, M. J., Grant, J., Dennis, N., Reardon, W., Splitt, M., Read, A. P., & Donnai, D. (1999). Williams syndrome: Use of chromosomal micro-deletions as a tool to dissect cognitive and physical phenotypes. *American Journal of Human Genetics*, *64*, 118–125.

Temple, C., & Clahsen, H. (2002). How connectionist simulations fail to account for developmental disorders in children. *Behavioral & Brain Sciences*, *25*, 769–770.

Temple, C., Almazan, M., & Sherwood, S. (2002). Lexical skills in Williams syndrome: a cognitive neuropsychological analysis. *Journal of Neurolinguistics*, *15* (*6*), 463–495.

Thomas, M. S. C., & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behavioural and Brain Sciences*, *25* (*6*), 727–788.

Thomas, M. S. C., & Karmiloff-Smith, A. (2003). Modelling language acquisition in atypical phenotypes. *Psychological Review*, *110* (*4*), 647–682.

Thomas, M. S. C., & Karmiloff-Smith, A. (2005). Can developmental disorders reveal the component parts of the human language faculty? *Language Learning and Development*, *1* (*1*), 65–92.

Thomas, M. S. C., Grant, J., Gsödl, M., Laing, E., Barham, Z., 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.

Tyler, L. K. (1985). Real-time comprehension processes in agrammatism: a case study. *Brain and Language*, *26*, 259–275.

Tyler, L. K. (1992). *Spoken language comprehension: An experimental approach to the study of normal and disordered processing*. Cambridge, MA: MIT Press.

Tyler, L., 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.

Udwin, O., & Dennis, J. (1995). Williams syndrome. In G. O'Brian & W. Yule (Eds.), *Clinics in Developmental Medicine No. 128, Behavioural Phenotypes*. London: Mackeith Press.

Udwin, O., & Yule, W. (1990). Expressive language of children with Williams syndrome. *American Journal of Medical Genetics*, *Supplement 6*, 108–114.

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

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

APPENDICES

TABLE A

Individual participant data for the WS group, including chronological age (CA), general cognitive ability on the British Abilities Scale (GCA), test age on the British Picture Vocabulary Scale (BPVS), and on the Pattern Construction subtest of the BAS. Participants are ordered by chronological age.

Participant	*CA*	*BAS GCA*	*BPVS*	*BAS Pattern Construction*
1	12;0	63	10;2	6;1
2	13;3	48	8;2	5;7
3	13;6	44	7;0	4;4
4	13;7	47	9;4	5;4
5	16;0	46	7;4	5;10
6	16;7	41	8;0	5;4
7	17;4	46	8;7	5;4
8	21;4	39	4;4	5;0
9	22;6	39	8;1	4;10
10	23;10	39	8;4	3;4
11	31;10	39	9;9	5;4
12	32;2	39	17;0	5;4
13	33;0	51	17;0	7;1
14	36;10	39	7;6	4;7
15	44;9	73	17;6	8;9
16	53;0	50	16;0	5;0
Mean (SD)	25;1 (12;4)	46.4 (9.6)	10;3 (4;2)	5;5 (1;2)
Range	12;0–53;0	39–3	4;4–17;6	3;4–8;9

TABLE B

Full set of thematic errors for Object and Action pictures, split by group. CA = chronological age match group; RVA = receptive vocabulary age group; WS = Williams syndrome group.

Category	*Thematic error*	*CA*	*RVA*	*WS*	*Context of picture*
Objects	'Chair' for cushion	1	2	4	Only cushion coloured in (see note below)
	'Wristwatch' for wrist		1		
	'Bracelet' for wrist		1		
	'Tea' for cup		1		
	'Teapot' for cup		2		
	'What red-riding-hood wears' for cloak		1	2	Cloak was red
	'Bear' for pyjamas			1	Bear was tucked under arm of pyjamas
	Total	1	8	7	
Actions	'Tug-of-war' for pulling		2		One person pulling on horizontal rope
	'Climbing' for pulling			1	
	'Ballet' for dancing		2	5	Picture of ballerina dancing
	'Feet-in-the-air' for dancing			1	
	'Ballerina' for dancing			1	
	'Headache' for sweating		1	1	Face with droplets of water around it
	'Hot' / 'Tired' / 'Shattered' for sweating		4	4	
	'Washing' for sweating			1	
	'Jelly' for wobbling		1	9	Picture of large jelly with motion lines
	'Bicycle' for cycling		1	4	
	'Sad' / 'tearful' / 'upset' for crying			3	
	'Cat' for stroking			2	Picture of woman stroking a cat
	'Flowers' for picking			2	Picture of girl picking flowers
	'Bird' for flying			1	Picture of bird with motion lines
	'Baby' for crawling			1	Picture of baby crawling
	'Fallen-over' for crawling			2	
	'Presents' for wrapping			1	Picture of woman wrapping present
	'Foot-in-the-air' for hopping			1	
	Total	0	11	40	

Note: Participants were requested to name only the part of the picture that was coloured in. Black and white components of pictures were included to give context.