Williams, D., Payne, H., Marshall., C. (2013). Non-word repetition impairment in autism and specific language impairment: Evidence for distinct underlying neuro-cognitive causes. *Journal of Autism and Developmental Disorders, 43,* 404-417.

David Williams, Department of Psychology, Durham University, Durham, UK.

Heather Payne, Department of Language and Communication Science, City University London, UK.

Chloe Marshall, Department of Psychology and Human Development, Institute of Education, London, UK.

Abstract

Language-impaired individuals with autism perform poorly on tests such as non-word repetition that are sensitive clinical markers of specific language impairment (SLI). This has fuelled the theory that language impairment in autism represents co-morbid SLI. However, the underlying cause of these deficits may be different in each disorder. In a novel task, we manipulated non-word stimuli in three ways known to influence the repetition accuracy of children with SLI. Participants with SLI were affected differently by these manipulations to children with autism. Children with autism performed similarly to language-matched typical children in terms of levels and patterns of performance, and types of error made, suggesting that the underlying cognitive cause of non-word repetition deficits is different in each disorder.

Keywords: Autism; specific language impairment; non-word repetition; clinical markers

Abbreviations: language-impaired individuals with autism spectrum disorder (ASD-LI); verbal mental age-matched typically developing individuals (VMA-TD); chronological age-matched typically developing individuals (CA-TD); specific language impairment (SLI); Clinical Evaluation of Language Fundamentals (CELF); Core language score (CLS); contactin associated protein-like 2 (CNTNAP2); ATPase, Ca++ transporting, type 2C, member 2 (ATP2C2); c-Maf inducing protein (CMIP)

Author Note

This study was conducted during a City University Research Fellowship awarded to the first author. The third author was funded by a Leverhulme Early Career Fellowship (ECF/2008/0469). An oral presentation of the results of this paper was delivered by the first

author at the biennial meeting of the Society for Research in Child Development, Montreal, Canada, April 2011.

Correspondence concerning this article should be addressed to Chloe Marshall, Department of Psychology and Human Development, Institute of Education, University of London, 25 Woburn Square, WC1H 0AA. Email: c.marshall@ioe.ac.uk. Telephone: +44 (0)20 7612 6509. Alternatively, correspondence can be addressed to David Williams, School of Psychology, Keynes College, University of Kent, Canterbury, Kent, CT2 7NP. Tel. +44 (0)1227 827652. Email: d.m.williams@kent.ac.uk;

Non-word repetition impairment in autism and specific language impairment: Evidence for distinct underlying cognitive causes

Although a deficit in structural language (phonology/grammar/semantics) is not core to autism spectrum disorder (ASD), over 50% of individuals with an ASD diagnosis manifest clinically significant impairment in this domain (e.g., Baird et al., 2006), and this can occur independently of diminished nonverbal IQ (NVIQ). In this regard, ASD has been compared to another developmental disorder, specific language impairment (SLI). SLI is diagnosed in children who display markedly impaired spoken language functioning with test scores at least 1.25 *SD*s below the mean, despite normal NVIQ, and no frank sensory or neurological dysfunction (American Psychiatric Association, 2000; Tomblin et al., 1997). Currently, ASD and SLI are mutually exclusive diagnoses, although there is a long-standing suggestion that they overlap at some level of description (Bartak et al., 1975, 1977). Indeed, some have suggested that when language impairment occurs in ASD it represents co-morbid SLI (see Lindgren, Folstein, Tomblin, & Tager-Flusberg, 2009). However, similarities in the surface presentation of two disorders do not necessarily mean that the disorders are qualitatively the same (Williams, Botting, & Boucher, 2008); observed behaviour may have different underlying *causes* at the cognitive, neurobiological, and/or etiologic levels of explanation. In this case, Bishop and Snowling (following Morton & Frith, 1995) argue that "researchers and practitioners need to move away from global measures of reading or language and use instead measures of underlying component processes" (2004; p.859). In this regard, a great deal of research into SLI has focussed on the performance of affected individuals on clinical marker tasks that appear to index the underlying cognitive liability/endophenotype of the disorder (e.g., Bishop, North, & Donlan, 1996; Rice, Wexler, & Cleave, 1995). In particular, "non-word repetition" skills have been investigated thoroughly among individuals with SLI (Coady & Evans, 2008).

In a non-word repetition test, the participant listens to non-words spoken by the tester and repeats each immediately after hearing it. Diminished non-word repetition distinguishes children with SLI from typically developing (TD) children in over 80% of cases (Conti-Ramsden et al., 2001), characterises "resolved cases" of SLI (e.g., Bishop et al., 1996), is highly heritable in this disorder (e.g., Barry, Yasin, & Bishop, 2007), and is associated with specific polymorphisms of identified genes (see below). For these reasons, poor non-word repetition is widely regarded as a cognitive marker of heritable SLI (e.g., Bishop et al., 1996).

Strikingly, several studies of non-word repetition in ASD have observed diminished test performance only among those individuals who also have impaired structural language (ASD-LI; see Whitehouse, Barry, & Bishop, 2008). Clearly, this could represent a significant link between language impairment in ASD and language impairment in SLI (Tager-Flusberg, 2004). However, there could be many different underlying causes of poor non-word repetition performance. Only if the underlying cause of poor non-word repetition in ASD-LI is the same as the underlying cause of poor non-word repetition in SLI can the overall diminution seen among both groups of individuals be accurately said to represent more than a superficial overlap (see Williams et al., 2008). As Bishop and Snowling (2004) suggest, one way to investigate this is to focus on qualitative *patterns* of non-word repetition performance and *types of error* made during unsuccessful repetition, in addition to overall levels of repetition accuracy. The key point is this: If the underlying cognitive cause of poor non-word repetition is the same in ASD-LI as it is in SLI, then manipulations to the structure of a non-word repetition task should influence the overall levels of accuracy and specific error patterns in both groups in a similar manner. Put another way, if the underlying causes of poor non-word repetition are the same in ASD-LI and SLI, then qualitative and quantitative similarity

in performance between individuals with each disorder should be observed. Alternatively, if SLI represents a qualitatively distinct disorder in which non-word repetition deficits have mainly unique underlying causes (that are different from the underlying causes of normal variation in non-word repetition ability), then the performance of individuals with this disorder should not pattern like that of individuals with ASD-LI.

Underlying cause of poor non-word repetition in SLI

Ever since Gathercole and Baddeley's (1990) seminal finding that children with SLI have a selective difficulty in repeating long non-words (3 or more syllables), but not short ones (1 or 2 syllables), many have argued that reduced short-term memory capacity underlies poor non-word repetition in SLI (see Coady & Evans, 2008). However, a number of other factors influence the non-word repetition performance of children with SLI, leading some researchers to question whether diminished short-term memory can fully explain non-word repetition impairment in this disorder (Archibald & Gathercole, 2007; Snowling et al, 1991; van der Lely & Howard, 1993). For example, short, but phonologically complex, non-words (i.e., those containing consonant clusters or unusual stress patterns) can cause difficulties for children with SLI (e.g., Gallon, Harris, & van der Lely, 2007).

Crucially, then, it is not just the amount of phonological material that influences how accurately a non-word is repeated, but also the way in which that material is structured. For example, irrespective of non-word length, the position of a cluster within an item affects how accurately children with SLI and children with dyslexia are able to repeat it (Marshall & van der Lely, 2009). Children with SLI and children with dyslexia are less likely to repeat a cluster accurately when it is word-medial (as in fe*bl*itorgist) as opposed to word-initial (as in *fl*ebitorgist), whereas TD children are comparatively unaffected by cluster position. Following Marshall and van der Lely, we speculate that this reflects a core deficit in both SLI and dyslexia with the construction/short-term retention/retrieval of phonological representations (as argued by Snowling et al., 1991; van der Lely & Howard, 1993). Further evidence in favour of this interpretation is provided by Marshall and van der Lely's findings that children with SLI (and dyslexia) frequently *created* clusters in non-target positions, which is consistent with difficulty in forming representations, but not with a difficulty in perception (i.e., children must have perceived clusters, but misremembered their position in the nonword) or articulation (i.e., they *are* able to articulate clusters, just not always in the correct position).

A third factor that affects non-word repetition accuracy in SLI is word-likeness. Although non-words are not stored in the lexicon, the extent to which they are similar to existing lexical representations does impact on repetition accuracy (Gathercole et al, 1991). Given the lexical and morphological deficits that characterise SLI (see Leonard, 2000), this group might be expected to benefit less from the word-likeness of non-word stimuli, and there is some evidence to support this. Archibald and Gathercole (2006) found that SLI children were impaired relative to verbal ability-matched controls on non-words that were word-like as indexed by the fact that many contain real words within them and morphological endings, but were not impaired relative to that group on non-words that were less word-like.

Finally, it is important to discuss briefly the findings of recent molecular genetic studies of non-word repetition in SLI. Variants of the CNTNAP2 gene on chromosome 7 (e.g., Vernes et al., 2008), and the CMIP and ATP2C2 genes on chromosome 16q (e.g., Newbury et al., 2009) have been implicated in non-word repetition impairment (taken as an index of heritable language impairment) in SLI. Variants of CNTNAP2 are associated with normal variation in language ability in the population as a whole (Whitehouse et al., 2011), and are implicated in multiple disorders of learning, including attention deficit hyperactivity

disorder (Elia et al., 2009) and mental retardation (Ballarati et al., 2009). Thus, CNTNAP2 may represent the kind of "generalist gene" that does not contribute to specific disorder, as such, but to normal *variation* in ability (e.g., Butcher, Kennedy, & Plomin, 2006). Thus, it is unlikely to explain the specific *profile* of non-word repetition difficulty observed among individuals with SLI. On the other hand, variants of CMIP and ATP2C2 appear to be associated with non-word repetition performance/language ability among individuals with SLI only (Newbury et al., 2009; Newbury et al., 2011). As Newbury et al. (2009, p.270) argue, unlike with the case of generalist genes, "genetic variants might have selective effects in specific populations…SLI represents a distinct disorder caused by genetic variants that are distinct from those that influence language ability in the general population". If this is the case, then it could explain why, in addition to showing a general diminution of non-word repetition ability, children with SLI show a unique profile of performance on this task. The question to be addressed in this paper, of course, is whether children with ASD-LI show a similar profile of non-word repetition performance to that shown by individuals with SLI.

Underlying cause of poor non-word repetition in ASD-LI

In one of the only studies to explore qualitative patterns of non-word repetition performance in ASD, Whitehouse et al. (2008) found individuals with SLI (n = 18) performed less well than participants with ASD-LI (n = 9) only when stimuli were ≥ 4 syllables in length. In other words, individuals with ASD-LI were relatively less affected by stimulus length than individuals with SLI. Such a different pattern of non-word repetition performance in SLI than in ASD-LI performance led Whitehouse et al. to conclude that the mechanism(s) underlying poor non-word repetition performance in ASD-LI may be different to those underlying poor non-word repetition performance in SLI. As Whitehouse et al. argued, this weakens the broader claim that language impairment in ASD represents a co-morbid SLI.

Recently, Riches, Loucas, Baird, Charman, and Simonoff (2011) explored non-word repetition skills among adolescents with ASD-LI and adolescents with SLI. Like Whitehouse et al., (2008), Riches et al. report that participants with SLI perform less well (in terms of the number of phonemic errors made) than participants with ASD-LI only when stimuli were ≥ 4 syllables in length. Such qualitative differences in non-word repetition performance between the groups led Riches et al. (p.10) to the same conclusion as Whitehouse et al., that "the claim for a phenotypic overlap between SLI and ALI [ASD-LI] may have been overstated". However, in terms of a) overall error rates, and b) percentage of phonemic errors made by participants that specifically affected the syllable structure of the stimuli, Riches et al. did not observe any significant differences between the two groups. As such, this may yet suggest an overlapping cognitive cause of non-word repetition deficits in ASD-LI and SLI.

The study by Riches et al. (2011) provides a welcome focus on the cognitive and linguistic underpinnings of non-word repetition deficits in ASD-LI and SLI, and it is methodologically rigorous in many respects. However, the results of the study may have been influenced by the fact that participants with ASD-LI were not matched with participants with SLI for chronological age or (as a result) verbal mental age (VMA). Thus, the comparable levels of non-word repetition performance in ASD-LI and SLI may simply have been due to participants with ASD-LI having lower verbal mental ages than participants with SLI. Moreover, and most importantly, to our knowledge, no study of non-word repetition in ASD-LI has included a group of typically developing children who are matched with clinical participants for verbal mental age, even though this is commonplace in studies of non-word repetition among children with SLI. Thus, any observed difficulty with non-word repetition in ASD-LI may reflect only a developmental delay that is in keeping with overall developmental (language) level, rather than any deviance that might reflect a specific deficit

in the cognitive mechanisms that underpin non-word repetition in SLI. Clearly, it is essential for our understanding of non-word repetition ability in ASD that comparison participants are matched closely for verbal mental age.

Several things lead us to question a) whether non-word repetition will be impaired in ASD-LI relative to such a verbal mental age-matched typical group and b) whether children with ASD-LI will show a similar profile of non-word repetition performance to their peers with SLI. First, immediate/rote/short-term memory, which is clearly one important component of non-word repetition, is long-considered to be an area of strength among individuals with ASD, including individuals with ASD-LI (e.g., Boucher & Warrington, 1976; Williams, Happé, & Jarrold, 2008; Williams, Bowler, & Jarrold, 2012). Second, non-functional echolalia is a common feature of ASD, including among individuals with ASD-LI. An individual who echoes words that they do not understand is essentially engaging in real-world non-word repetition. The fact that even young, low-functioning children with ASD can echo words accurately leads us to believe that non-word repetition is, relatively speaking, not particularly impaired in ASD-LI (i.e., not impaired more than would be expected on the basis of structural language level). Third, difficulties with phonology are not a prominent feature of language impairment in ASD; even when phonological impairments are observed in very young children with ASD, they resolve by the time these children enter school, unlike in the majority of cases of SLI where deficits in phonology are persistent and pervasive (see Williams et al., 2008). Finally, non-word repetition performance shows no sign of being heritable in ASD-LI, as it is in SLI. Family studies have consistently demonstrated a lack of familial aggregation of non-word repetition deficits in ASD-LI (for a review, see Williams & Lind, forthcoming), and molecular genetic studies have not (as far as we know) identified loci on chromosome 16q as harbouring susceptibility genes for (language impairment in) ASD. Common variants of CNTNAP2 have been implicated in ASD (Arking et al., 2008), with the association strongest when analyses are restricted to language-delayed samples (but note: in these studies, "language delayed" refers to individuals who have not uttered first words by 12 months of age; not necessarily individuals who would be classified as language -impaired later in life; e.g., Alarcón et al., 2008). This finding has been taken to support the notion that ASD-LI are genetically comorbid and that language impairment in ASD-LI has (at least partially) the same (genetic) basis as it does in SLI (see Bishop, 2010 for a discussion). However, these variants of CNTNAP2 have general effects on language and cognition among the typical population also (Whitehouse et al., 2011). Thus, even if CNTNAP2 does contribute to language impairment in ASD, it may not result in a pattern of non-word repetition performance that is fundamentally "atypical", such as that seen in SLI. As Bishop (2010, p.626) argues, with regard to the question of whether language impairment in ASD is comorbid SLI, or whether it is merely a phenomimic of SLI: "we need more studies of qualitative aspects of language phenotypes in ASD and SLI to test this hypothesis convincingly." We suggest that the current study represents the kind of qualitative investigation of language phenotypes that Bishop (2010) argues could shed light on this debate.

In the present study, we employed a novel non-word repetition task, in which stimuli were manipulated systematically for three factors that have previously been shown to reveal differences in repetition accuracy between children with SLI and typically developing children: length (3 syllables versus 4 syllables), consonant cluster position (initial versus medial), and word-likeness (presence versus absence of a morphological suffix). We assessed performance on this task among closely matched groups of participants with ASD-LI and SLI, as well as among a TD comparison group matched with the clinical groups for chronological age, and a second typically developing comparison group matched with the clinical groups for verbal mental age. Our predictions were as follows:

1) Children with ASD-LI will show significantly diminished non-word repetition performance relative to age-matched typically developing participants, replicating other studies (e.g., Riches et al., 2011). However;

2) Relative to verbal mental age-matched typically developing children, children with ASD-LI will show *no* significant differences in either levels or patterns of non-word repetition performance. Both the verbal mental age-matched typically developing group and the ASD-LI group will show more accurate levels of non-word repetition compared to the SLI group, and will be less affected by the manipulation of stimulus length, consonant cluster position and word-likeness.

3) Patterns of error shown by children with ASD-LI will be qualitatively similar to the pattern shown by the verbal mental age-matched typically developing group, reflecting the similar processes underlying test performance in each group. Both the verbal mental age-matched typically developing group and the ASD-LI group are predicted to make fewer errors on consonant clusters and to create clusters in non-target positions less often than the SLI group.

Method

Participants

This research was approved by the appropriate University Research Ethics Committee and all participants took part after informed consent had been gained from parents/guardians. Seventeen children with ASD-LI (16 male), 15 children with SLI (13 male), 19 young typically developing children (14 male), and 19 older typically developing (all male) took part. All participants were native speakers of English. Participant characteristics are shown in Table 1.

Participants in the ASD-LI group had received formal diagnoses by a psychiatrist or paediatrician of autistic disorder according to established criteria (American Psychiatric Association, 2000). Participants in this group attended specialist ASD schools in the UK that required children to have a formal diagnosis for entry into the school. The Social Communication Questionnaire (SCQ; Rutter et al., 2003) was employed as a measure of ASD feature severity[1]. Participants' verbal abilities were assessed using the Clinical Evaluation of Language Fundamentals – Forth Edition UK (CELF; Semel, Wiig, & Secord, 2006). All participants in this group had a "Core Language Score" (CLS) on the CELF < 78 (i.e., the seventh centile, or below), indicating significant structural language impairment. Non-verbal IQ was determined using the Perceptual Reasoning Index of the Wechsler Intelligence Scales for Children – Fourth Edition UK (WISC-IV; Wechsler, 2004). These measures of autism feature severity, language ability, and NVIQ are all widely employed in studies of ASD and SLI. All participants had a NVIQ > 80 (n = 13) or a minimum 15 point discrepancy between their CLS and their NVIQ (n = 4).

Participants in the SLI group had confirmed clinical diagnoses of language disorder, and were recruited from specialist schools for children with speech and language disorder. At these schools, pupils' articulation/oral-motor functioning and hearing, as well as linguistic/cognitive functioning, are routinely tested by speech and language therapists. In the current study, children were pre-selected for inclusion by speech and language therapists from each school on the basis that none had difficulties with articulation, or any current or previous history of hearing loss. None had any documented ASD-like features, and all scored below the ASD cut-off on the SCQ. All participants achieved a CLS on the CELF < 78, and a NVIQ > 80 (n = 11) or a minimum 15 point discrepancy between their CLS and their NVIQ (n = 4)[2].

Table 2 shows the standard scores on each subscale of the CELF among these participants and highlights how similar the profile of language impairment was in each clinical group (ASD-LI and SLI). The effect size for each contrast was small, except for that relating to the Recalling Sentences subtest, which was moderate (and for which the associated *p* value approached significance).

Participants in the verbal mental age-matched typically developing (VMA-TD) group were recruited from mainstream primary schools. No child had any reported developmental difficulties. Verbal abilities were assessed using the Concepts and Following Directions, and the Recalling Sentences subtests from the CELF. A prorated NVIQ was derived from performance on the Block Design and Matrix Reasoning subtests of the WISC. Finally, VMA-TD participants, as well as participants with ASD-LI and participants with SLI, completed the Forwards Digit Span (DS) subtest of the WISC as a measure of basic short-term memory capacity.

Participants in the chronological age-matched typically developing (CA-TD) group were recruited from a mainstream secondary school and none had any reported developmental difficulties. Verbal and non-verbal abilities were not assessed in these participants. However, IQ testing was carried out as a matter of routine at the school from which these participants were recruited. Inspection of each participant's IQ score (which we do not have permission to include details of) revealed that all participants in this group had IQs in the normal range.

One-way ANOVA indicated significant differences between the four groups in chronological age, $F(3, 70) = 95.26$, $p < .001$. Independent *t*-tests revealed that the VMA-TD group was younger than in each of the other groups (all *p*s <.001, all Cohen's *d*s > 3.98). There were no significant differences in chronological age between the other groups (although participants in the CA-TD group were somewhat younger than those in the ASD-LI and SLI groups), all *p*s >.07, all Cohen's *d*s < 0.70.

There were no significant differences in verbal mental age between the ASD-LI, SLI, and VMA-TD groups, $F(2, 50) = 0.04$, $p = .96$. Moreover, differences between these three groups in digit span only approached significance, $F(2, 47) = 2.60$, $p = .09$, although there was a clear tendency for participants with SLI to manifest lower spans than VMA-TD participants (but not than ASD-LI participants; see Table 1). The groups did differ significantly in CLS, $F(2, 50) = 101.30$, $p < .001$, and NVIQ, $F(2, 50) = 9.89$, $p < .001$. The clinical groups were closely matched on both variables (all *p*s >.55, all *d*s < 0.22), whereas the VMA-TD group had a significantly higher CLS (all *p*s <.001, all *d*s > 3.90) and a significantly higher NVIQ (all *p*s ≤.001, all *d*s > 1.18).

Tables 1 and 2 here

Stimuli and procedures

The non-word repetition stimuli were based on those created by Marshall and van der Lely (2009), but in addition to manipulating cluster position, as Marshall and van der Lely did, we also manipulated length and the presence versus absence of a suffix. They consisted of eight basic non-word stems containing obstruent + liquid clusters (e.g., kr, kl, dr, fl) that were constructed in a 2 (Length: 3 versus 4 syllables) x 2 (Cluster position: Initial versus medial) x 2 (Morphological suffix: Presence versus absence) manner. The suffixes were attested in English (e.g., -ist, -ing). Unlike Marshall and van der Lely, we did not manipulate stress, as their data showed that this did not affect children with SLI. Instead, our stimuli had constant stress: for three syllable stimuli, this was on the first syllable (as in the real word 'chrysalis'),

and for four syllable items, primary stress was on the first syllable and secondary stress on the third (as in the real word 'caterpillar'). In total, 64 test items were created, along with 29 one and two syllable filler items. The test items are presented in Appendix 1.

For the non-word repetition task, participants were instructed that they would hear some made-up words and they should repeat the words to the experimenter. The stimuli were split into two 3 minute segments, separated by a brief break. Stimuli were presented in .mp3 format through Sennheiser CX 300-II Precision noise-isolating headphones and answers were recorded on a Tascam DR-1 digital recorder in .wav format. These were transcribed and scored from the recording on a separate occasion. The entire dataset was transcribed by a trained transcriber (the second author), and a random 25% of the data (comprising data from six participants with ASD-LI, four participants with SLI, four VMA-TD participants, and four CA-TD participants) was additionally transcribed by another trained transcriber (the third author), who was blind to group membership.

Scoring and Analyses

Two main measures of performance were investigated. First, overall repetition accuracy was analysed. For this analysis, each non-word correctly repeated was given a score of 1. Any reproduction errors on a given item resulted in a score of 0 for that item, with two exceptions: a) participants were not penalised for voicing errors (e.g., /g/ realised as [k]), because transcription of these with certainty was not always possible; b) minor vowel alterations, /r/ realised as [w], or the ending /-ing/ realised as [-in], were not penalised, as these reflect common dialectical variation in the part of the UK from which participants originated and are usually credited in non-word repetition studies. Alteration of stress was not accepted. Second, in order to compare our results to those of Marshall and van der Lely, consonant cluster repetition was analysed independent of overall repetition. Hence, accurate cluster repetition within an item was awarded a score of 1 even if the rest of the non-word was not repeated correctly. Inter-rater reliability on 25% of the data was near perfect according to Landis and Koch's (1977) criteria, for both whole non-words (correct/incorrect), $\kappa = .85$ and consonant clusters, $\kappa = .82$.

In addition to these analyses of performance, we also explored the *patterns of error* made by participants on items that were not repeated correctly. As noted above, distinct patterns of repetition error have been observed in children with SLI (e.g., Marshall & van der Lely, 2009), and previous studies of error patterns on a variety of clinical marker tasks among individuals with SLI have yielded a rich source of information regarding the underlying cognitive deficits that characterise the disorder (see Williams et al., 2008). Therefore, following Marshall and van der Lely, we analyzed two types of error with respect to consonant cluster repetition. First, we explored errors on the target cluster itself. In this respect, five types of error were possible, where C1 refers to the first consonant in the cluster and C2 to the second: C1 deleted: e.g. *flebitorgist – lebitorgist;* C2 deleted: e.g. *flebitorgist – febitorgist;* C1 substituted: e.g. *flebitorgist – lebitorgist;* C2 substituted: e.g. *flebitorgist – frebitorgist;* Other errors (i.e. errors that did not fit into any of the aforementioned categories, and which resisted an easy explanation): e.g. *flebitorgist – tebitorgist*.

Second, we explored errors characterised by the *creation* of a novel consonant cluster in the non-word. This kind of error could occur even when the target cluster in an item was repeated accurately. For example, a participant who incorrectly repeated the item *feblitorgist* as *fleblitorgist* would be credited for correctly reproducing the medial target cluster [bl], but would have added a cluster [fl] at the beginning of the item.

Results

Overall non-word repetition accuracy

A 2 (Length) x 2 (Suffix) x 4 (Group) mixed ANOVA was conducted using the percentage of non-words repeated correctly as the dependent variable. There were significant mains effects of Group, $F(3, 67) = 29.97$, $p < .001$, Length, $F(1, 67) = 189.70$, $p < .001$, and Suffix, $F(1, 67) = 16.58$, $p < .001$. However, these main effects were qualified by significant interactions between Group and Length, $F(3, 67) = 4.50$, $p = .006$, and between Length and Suffix, $F(1, 67) = 5.42$, $p = .02$.

To break down the interaction between Group and Length (see Figure 1), a series of within-participant and between-participant *t*-tests were conducted. Among each group of participants, a significant effect of non-word length was present, with four syllable non-words being less likely to be repeated correctly than three syllable non-words (all ps $\leq .002$). However, the effect size for this contrast was notably smaller among participants with SLI ($d = 0.54$) than among participants with ASD-LI ($d = 1.12$), VMA-TD participants ($d = 1.36$), or CA-TD participants ($d = 1.29$). Thus, stimulus length appeared to affect the repetition performance of children with SLI somewhat *less* than it affected the performance of the other groups of participants. This result reflects the poor non-word repetition performance of participants with SLI even when stimuli were only three syllables in length. Thus, participants with SLI performed less well than each of the other groups not only when stimuli were 4 syllables in length (all ps $\leq .03$, all ds ≥ 0.80), but also when stimuli were 3 syllables in length (all ps $\leq .02$, all ds ≥ 0.94). There were no significant differences in performance between participants with ASD and VMA-TD participants either when stimuli were 3 syllables in length, $t(34) = 1.28$, $p = .21$, $d = 0.42$, or 4 syllables in length, $t(34) = 1.31$, $p = .20$, $d = 0.44$. However, the performance of CA-TD participants was superior to that of participants from all other groups for both 3 syllable stimuli (all ps $<.001$, all ds ≥ 1.96) and 4 syllable stimuli (all ps $<.001$, all ds ≥ 2.25).

To break down the interaction between Length and Suffix, within-participant *t*-tests among all groups collapsed were conducted. Items with a suffix were repeated significantly more reliably than items without a suffix when items were four syllables in length, $t(70) = 4.61$, $p < .001$, $d = 0.35$, but only marginally significantly when items were three syllables in length, $t(70) = 2.03$, $p = .05$, $d = 0.16$.

The interactions between Group and Suffix, and between Group, Suffix, and Length were non-significant, all Fs < 0.85, all ps $>.47$.

Figure 1 here

Association between non-word repetition and short-term memory

To explore the contribution of short-term memory to non-word repetition performance, correlation analyses were conducted among each group of participants exploring the association between overall non-word repetition (with repetition of three and four syllable non-words collapsed) and digit span. ASD-LI and VMA-TD participants were similar in showing a moderate-to-large association between these variables, $r_s = .52$, $p = .03$, and $r_s = .42$, $p = .05$ (one-tailed), respectively. In contrast, the association between these variables was only small and non-significant among participants with SLI, $r_s = .23$, $p = .40$.

Cluster repetition accuracy

A 2 (Cluster position) x 4 (Group) mixed ANOVA was conducted using the percentage of consonant clusters correctly repeated as the dependent variable. There were significant mains effects of Group, $F(3, 67) = 20.95$, $p < .001$, and Cluster position, $F(1, 67) = 45.59$, $p < .001$. However, these were qualified by a significant interaction between Group and Cluster position, $F(3, 67) = 10.44$, $p < .001$.

To break down the interaction between Group and Cluster position (see Figure 2), a series of within-participant and between-participant *t*-tests was conducted. Medial clusters were significantly less likely to be repeated accurately than initial clusters among both participants with SLI, $t(14) = 4.89$, $p < .001$ .21, $d = 1.59$, and (relatively less so) VMA-TD participants, $t(18) = 3.56$, $p = .002$, $d = 1.22$. Among participants with ASD-LI, medial clusters were repeated less frequently than initial clusters, although this difference only approached significance, with only a moderate effect size $t(16) = 2.05$, $p = .06$, $d = 0.60$. Among CA-TD participants, medial clusters were repeated as well as initial clusters, $t(19) = 0.68$, $p = .51$, $d = 0.16$. Critically, between-participant analyses revealed that children with SLI reproduced medial clusters significantly less well than both participants with ASD-LI, $t(30) = 2.97$, $p = .006$, $d = 1.59$, and VMA-TD participants, $t(19.26) = 3.42$, $p = .003$, $d = 1.97$. However, among participants with SLI, repetition of initial clusters was relatively unimpaired, compared to participants with ASD-LI, $t(30) = 0.15$, $p = .88$, $d = 0.05$, and VMA-TD participants, $t(21.65) = 1.83$, $p = .08$, $d = 0.65$. There were no significant differences between the ASD-LI and VMA-TD groups in the reproduction of either initial clusters, $t(34) = 1.61$, $p = .12$, $d = 0.53$, or medial clusters, $t(34) = 0.37$, $p = .77$, $d = 0.12$. Among CA-TD participants, repetition of both initial and medial clusters was significantly superior to that observed in participants from each of the other groups (all ps $\leq .004$, all ds ≥ 1.03).

Figure 2 here

Error patterns

Given that so few errors were made by participants in the CA-TD group, we analysed error patterns among the ASD-LI, SLI, and VMA-TD groups only.

Errors in cluster production

Table 3 shows the proportion of each kind of error in cluster production made by participants in each diagnostic group, collapsed across conditions (Length, Cluster position, Suffix). Using Pillai's trace, A MANOVA indicated a significant effect of Group on the type of error made, $V = 0.36$, $F(8, 90) = 2.20$, $p = .03$. Separate one-way ANOVAs revealed significant differences between the groups in the proportion of Other errors, $F(2, 47) = 3.39$, $p = .04$, and in the proportion of C2 substitution errors, $F(2, 47) = 3.80$, $p = .04$ (all other ps > .07). Follow-up *t*-tests on these variables indicated the following. As a proportion of their total number of errors, participants with SLI made significantly more errors of the Other error type than participants with ASD-LI, $t(19.52) = 2.10$, $p < .05$, $d = 0.78$. Participants with SLI also made a marginally higher proportion the Other error kind than VMA-TD participants, $t(32) = 1.97$, $p < .06$, $d = 0.68$. ASD-LI and VMA-TD participants did not differ from each other in this respect, $t(33) = 0.29$, $p = .78$, $d < 0.01$. With respect to C2 substitution errors, the VMA-TD group made a significantly higher proportion of errors of this type than participants from either clinical group (all ps < .05, all ds > 0.71), whereas the groups of clinical participants did not differ from each other, $t(29) = 0.14$, $p = .89$, $d = 0.07$.

Table 3 here

Errors in cluster creation

On average, participants with SLI created new consonant clusters on 11.73 (SD = 7.20) items. This compared to 4.57 (SD = 3.70) items among VMA-TD participants and 6.41 (SD = 4.24) items among ASD-LI participants. A one-way ANOVA revealed a significant difference between the groups, in this respect, $F(2, 50) = 8.56$, $p = .001$. Follow-up *t*-tests indicated that participants with SLI were significantly more likely to create a new consonant cluster than VMA-TD participants, $t(19.79) = 3.50$, $p = .002$, $d = 1.25$, or participants with ASD, $t(22.09) = 2.51$, $p = .02$, $d = 0.90$. Participants with ASD and VMA-TD participants did not differ significantly in this respect, $t(34) = 1.38$, $p = .18$, $d = 0.46$.

Discussion

This study presents the first comparison of non-word repetition in children with specific language impairment (SLI), language-impaired children with autism spectrum disorder (ASD-LI), and their Verbal Mental Age Typically Developing controls (VMA-TD), using a new set of stimuli that manipulate a range of phonological characteristics, namely stimulus length, cluster position and word-likeness. Two findings of the current study stand out as particularly important. On the one hand, there was a striking *difference* in non-word repetition performance between children with ASD-LI and children with SLI. Compared to participants with ASD-LI, those with SLI a) performed significantly less well overall, b) were more affected by the position of the consonant cluster in the non-word, c) were more likely to create novel clusters in incorrect positions of the non-word, and e) differed from participants with ASD-LI in *not* showing a reliable association between non-word repetition performance and short-term memory capacity.

On the other hand, there was a striking *similarity* in non-word repetition performance between children with ASD-LI and verbal mental age-matched typically developing children. The two groups performed similarly regardless of stimuli-length, or whether clusters were located word-initially or word-medially. They also showed a similar pattern of errors in the repetition of clusters, differing significantly in only one way (VMA-TD participants were more proportionately likely to substitute the second consonant in a cluster). Moreover, both groups showed a similar pattern and magnitude of association between non-word repetition performance and short-term memory capacity.

Before considering the implications of these findings, it is important to consider whether the results reflect methodological flaws in the design of the study, rather than substantive discoveries. With regard to the *similarities* in non-word repetition performance between ASD-LI and VMA-TD participants, we did not employ what are seen as the "gold standard" instruments for diagnosing ASD, namely the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) or the ADI (Le Couteur et al., 2003). Therefore, it could be questioned whether the similarity in non-word repetition performance shown by ASD-LI and VMA-TD participants could be accounted for by unreliable ASD diagnoses among participants in the ASD-LI group. We feel strongly that this is unlikely to be an explanation for our findings. First, if participants in the ASD-LI group did not really have ASD, then the diagnosis that would be most obvious from their cognitive profile would be SLI. If this was the case, then they could have been reasonably expected to perform similarly on the experimental task to the (closely-matched) participants in the SLI group. Instead, they performed quite differently. In fact, however, participants in the ASD-LI group did have formal ASD diagnoses from a psychiatrist/paediatrician. Indeed, several participants in this

group had been diagnosed at a single, UK-leading ASD clinic. Therefore, we are confident in the accuracy of participants' ASD diagnoses even though we did not employ the ADOS or ADI. Indeed, in our view, the use of ADOS/ADI in research studies is not to "independently confirm the participants' ASD diagnoses", as some suggest, but rather to provide a useful indicator of the severity of ASD features. That is why we employed the SCQ with our clinical participants (as well as to allow exploration of the association between severity of ASD features and performance on the experimental task). Even if we had employed the ADOS or ADI, none of us is a trained psychiatrist or paediatrician and, thus, whatever scores were obtained by a participant on these measures would not give us the right to confirm or disconfirm the diagnosis[3].

With regard to the *differences* in non-word repetition performance between ASD-LI and SLI participants, we did not employ a standardised measure of articulation skills, such as the Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 2000). It is well established that individuals with SLI more frequently have significant articulation difficulties than do children with ASD-LI (see Williams et al., 2008). Thus, it could be questioned whether the only real reason for differences in non-word repetition performance between these two groups was because of significant articulation problems in the latter group, but not the former group. Again, we feel that this is an unlikely explanation for our results. We recruited participants with SLI from nationally-recognised schools for children with developmental language disorders. The speech, language, and intelligence of these children are tested extensively by professional speech and language therapists upon entry to the school, and then subsequently on a semi-regular basis. We specifically asked the head speech and language therapist at these schools to pre-select children who were free from articulation/oro-motor difficulties. We did not formally assess articulation because we felt that the judgements of speech and language therapists who work with the children each day and who had full access to each child's medical history were superior to any judgment that we could make on the basis of a five minute assessment using a standardised task. Therefore, we are confident that these participants did not have concurrent articulation problems (indeed, had there been articulation problems, we would have picked these up when conducting language assessments), although they may well have had articulation problems early in life. In fact, our finding that children with SLI *produced* significantly more novel consonant clusters than any other group is entirely inconsistent with articulation problems in this group. Articulation difficulties could result in an across-the-board difficulty producing clusters, but children with SLI in the current sample had difficulties only in producing *target* clusters (arguably suggesting a representational deficit; see below). Nor do published studies of non-word repetition in SLI of this age routinely include articulation measures (e.g., Catts, Adlof, Hogan & Weismer, 2005; Marshall & van der Lely, 2009; Riches et al., 2011).

In sum, this is the first study of non-word repetition in ASD-LI to include a group of typically developing comparison children *matched for verbal mental age*. Although we replicated previous findings of non-word repetition impairment in ASD-LI relative to age-matched typical participants (e.g., Riches et al., 2011), the current results suggest that language-impaired children with ASD show no more of a deficit in non-word repetition than would be expected on the basis of their developmental (language) level – *they are developmentally delayed, but not developmentally deviant*, in this respect. This is in stark contrast to children with SLI, who have been shown in multiple studies of non-word repetition (including the current study) to perform not only less well than typical children with an equivalent verbal mental age, but also to show notable atypicality in patterns of performance, suggesting a more fundamental deviance in the mechanisms underpinning non-word repetition ability (Conti-Ramsden & Hesketh, 2003; Gallon, Harris & van der Lely, 2007; Gathercole & Baddeley, 1990).

Although the central focus of the paper was on non-word repetition in ASD, it is important to provide an interpretation of the non-word repetition performance of participants with SLI. That our results replicate the findings of Marshall and van der Lely is important, given heterogeneity within this population. Following Marshall and van der Lely (2009), we suggest four sources of evidence point to a deficit in phonological representations as an important underlying cognitive cause of non-word repetition deficits in SLI. First, children with SLI were particularly poor at repeating medial clusters compared to initial clusters, which indicates that it is not the overall presence versus absence of a cluster *per se* that affects repetition accuracy, but where in the phonological representation that cluster occurs. Second, with respect to errors, our group of children with SLI make proportionally more cluster errors that are not straightforward deletions or substitutions of consonants, but rather are difficult to describe and were therefore allocated to a catch-all "other error" category. Other errors all have in common the characteristic that both the structural integrity of the cluster and its segmental content are changed (the number of consonants is reduced from two to one, and neither of the original segments is preserved). Therefore, our interpretation is that these are particularly severe errors at the level of the phonological representation. Third, the error whereby children *create* clusters elsewhere in the non-word is again more common in children with SLI. We interpret these errors along the lines of the "misattachment" errors described by Marshall, Harris, and van der Lely (2003), whereby the child knows that there is a cluster somewhere in the non-word but, because they are relatively less able to assign a full prosodic template to the representation of the non-word they are more likely to rely on guessing (not always successfully) where the cluster was located. Finally, non-word repetition performance among participants was not significantly associated with a widely-used measure of short-term memory capacity, namely digit span. While this does not show that a primary deficit in phonological representations is the main cause of non-word repetition impairment in SLI, it does add to the growing body of evidence suggesting that impaired non-word repetition in SLI cannot be accounted for solely by diminished short-term memory (e.g., Archibald & Gathercole, 2007), which paves the way for alternative theories.

Therefore, our findings are consistent with van der Lely's long-argued position that the locus of the non-word repetition deficit in SLI is in the creation of phonological representations, which itself then makes the short-term retention and retrieval of phonological sequences problematic. However, our findings suggest no significant representational deficit in children with ASD-LI, which is in keeping with the vast majority of the literature on school-aged children with this disorder (see Williams et al., 2008). Non-word repetition performance in children with ASD-LI resembles that seen in language-matched typically developing children and therefore any non-word repetition deficit is very likely to reflect merely a general, non-specific delay in non-word repetition ability that is *in keeping with* general language level; among individuals with ASD-LI, non-word repetition ability is not deviant, as it is in SLI. There is no quantitative or qualitative difference in non-word repetition ability among individuals with ASD-LI and their mental age typically developing peers. The combination of factors that influence non-word repetition ability among typically developing individuals (e.g., short-term memory, quality of phonological representations, lexical access etc.) appear to influence non-word repetition ability among individuals with ASD *in equal measure*; there is a quantitative diminution of non-word repetition ability *relative to age-matched peers*, but this is not in the least surprising. Given that non-word repetition ability is so inherently tied up with language, it is unlikely that any language-impaired group, regardless of their primary diagnosis (i.e., ASD, SLI, Down syndrome etc.), could perform at a level entirely comparable to their age peers on a non-word repetition task.

The current results have wider implications for the debate regarding the relation between ASD-LI and SLI. Williams et al. (2008) argued against the theory that language

impairment in ASD-LI represents a co-morbid SLI, suggesting that behavioural similarities between the two disorders are only superficial (a form of "phenomimicry"; Bishop, 2010) and do not share the same *core* underlying causes (cf. Riches et al., 2011; Whitehouse et al., 2008). The current results support this latter view insofar as they suggest qualitatively different underlying cognitive causes of non-word repetition difficulty in ASD-LI and SLI. Let us be clear about logic here: non-word repetition impairments are widely accepted as a clinical marker of heritable SLI. As noted in the Introduction, in SLI, these deficits are persistent (even when superficial language impairment has resolved), highly heritable, and associated with specific polymorphisms of identified genes. These facts have driven the search for the cognitive and (more recently) etiological causes of SLI for some time (e.g., Newbury et al., 2009; Scerri et al., 2011). Now, if a picture of the core underlying cause of SLI is captured by performance on non-word repetition tasks, and if language impairment in ASD is straightforwardly comorbid SLI, then the logical conclusion is that individuals with ASD-LI should show similar levels *and* patterns of non-word repetition performance to their peers with SLI. In the current study, participants with ASD-LI were very similar to participants with SLI in terms of age, NVIQ, overall language ability, and even language profile (in that the groups were equated on all four language subtests from the CELF, as well as on the overall Core Language Score). Yet, despite the close similarity of these individuals in terms of their language impairment, the two groups performed strikingly differently on the non-word repetition task. At the very least, this finding is inconsistent with the hypothesis that language impairment in ASD is comorbid SLI. Of course, these findings will require replication, ideally among a larger sample of participants (although we should note that this is the largest study of non-word repetition in ASD and SLI that we know of). However, if replicable, these results suggest that non-word repetition performance probably does not provide the key insight into language impairment in ASD that it provides in SLI, and should not be a central focus of genetic studies attempting to uncover the etiological causes of language impairment in ASD.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th edition, text revised) (DSM-IV-TR).* Washington DC: American Psychiatric Association

Archibald, L.M., & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic Bulletin and Review, 14,* 919-924.

Archibald, L.M., & Gathercole, S. E. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research, 49,* 970-983.

Arking, D.E., Cutler, D.J., Brune, C. W., Teslovich, T. M., West, K., Ikeda, M., et al. (2008). A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *American Journal of Human Genetics, 82,* 160-164.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., et al. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet, 368*(9531), 210-215.

Ballarati, L., Recalcati, M. P., Bedeschi, M. F., Lalatta, F., Valtorta, C., Bellini, M., Finelli, P., Larizza, L., Giardino, D. (2009). Cytogenetic, FISH and array-CGH characterization of a complex chromosomal rearrangement carried by a mentally and language impaired patient. *European Journal of Medical Genetics, 52,* 218 -223.

Barry, J. G., Yasin, I., & Bishop, D. V. M. (2007). Heritable risk factors associated with language impairments. *Genes Brain and Behavior, 6,* 66-76.

Bartak, L., Rutter, M., & Cox, A. (1975). Comparative study of infantile autism and specific developmental receptive language disorder 1: Children. *British Journal of Psychiatry, 126,* 127-145.

Bartak, L., Rutter, M., & Cox, A. (1977). Comparative study of infantile autism and specific developmental receptive language disorder 3: Discriminant function analysis. *Journal of Autism and Childhood Schizophrenia, 7*(4), 383-396.

Bishop, D. (2011, May 30). Are our 'gold standard' autism diagnostic instruments fit for purpose? [Web log message] Retrieved from http://deevybee.blogspot.com/2011/05/are-our-gold-standard-autism-diagnostic.html

Bishop, D.V.M. (2010). Overlaps between autism and language impairment: Phenomimicry or shared etiology? *Behavior Genetics, 40,* 618-629.

Bishop, D. V. M., & Snowling, M. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130,* 858–886.

Bishop, D.V.M., Maybery, M., Wong, D., Maley, A., Hill, W., & Hallmayer, J. (2004). Are phonological processing deficits part of the broad autism phenotype? *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 128B,* 54-60.

Bishop, D.V.M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37*, 391-403.

Boucher, J., & Warrington, E. K. (1976). Memory deficits in early infantile autism: Some similarities to the amnesic syndrome. *British Journal of Psychology, 67*, 73-87.

Butcher, L. M., Kennedy, J. K. J., Plomin, R. Generalist genes and cognitive neuroscience. *Current Opinion in Neurobiology, 16*, 14 5 – 151

Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language and Hearing Research, 48*, 1378-1396.

Coady, J. ., & Evans, J.L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*, 1-40.

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

Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: Outcomes at 11 years of age. *International Journal of Language & Communication Disorders, 36*, 207-219.

Constantino, J. N. (2002). *The Social Responsiveness Scale*. Los Angeles, CA: Western Psychological Services.

Gallon, N,, Harris, J. & van der Lely, H.K.J. (2007). Nonword repetition: An investigation of phonological complexity in children with Grammatical-SLI. *Clinical Linguistics and Phonetics, 21*, 435-455.

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

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

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The Autism Diagnostic Interview – Revised (ADI-R).* Los Angeles, CA: Western Psychological Services.

Lindgren, K.A., Folstein, S.E., Tomblin, J.B., & Tager-Flusberg, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research, 2,* 22-38

Lord, C., Risi, S., Lambrecht, L., Cook, E. Leventhal, B., DiLavore, B., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule – Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205-224.

Marshall, C. R., Harris, J. & van der Lely, H.K.J. (2003). The nature of phonological representations in children with Grammatical-Specific Language Impairment (G-SLI). In D. Hall, T. Markopoulos, A. Salamoura & S. Skoufaki (eds.) *Proceedings of the University of Cambridge First Postgraduate Conference in Language Research,* 1, 511-517.

Marshall, C.R., & van der Lely, H.K.J. (2009). Effects of word position and stress on onset cluster production: Evidence from typical development, specific language impairment, and dyslexia. *Language, 85,* 39-57.

Mawhood, L., Howlin, P., & Rutter, M. (2000). Autism and developmental receptive language disorder - a comparative follow-up in early adult life. I: Cognitive and language outcomes. *Journal of Child Psychology and Psychiatry, 41,* 547-559.

Morton, J. & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, volume 1, theory and methods* (pp. 357-390). New York: John Wiley & Sons Inc.

Newbury, D. F., Winchester, L., Addis, L., Paracchini, S., Buckingham, L., Clark, A., Cohen, W., Cowie, H., Dworzynski, K., Everitt, A., Goodyer, I. M., Hennessy, E., Kindley, A. D., Miller, L. L., Nasir, J., O'Hare, A., Shaw, D., Simkin, Z., Siminoff, E., Slonims, V., Watson, J., Ragoussis, J., Fisher, S. E., Seckl, J. R., Helms, P. J., Bolton, P. F., Pickles, A., Conti-Ramsden, G., Baird, G., Bishop, D. V. M., Monaco, A. P. (2009). CMIP and ATP2C2 modulate phonological short-term memory in language impairment. *The American Journal of Human Genetics, 85,* 264 – 272.

Newbury, D. F., Paracchini, S., Scerri, T. S., Winchester, L., Addis, L., Richardson, A. J., Walter, J., Stein, J. F., Talcott, J. B., Monaco, A. P. (2011). Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects. *Behavior Genetics, 41,* 90 – 104.

Rice, M.L., Wexler, K., & Cleave, P.L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38,* 850-863.

Riches, N.G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2011). Non-word repetition in adolescents with specific language impairment and autism plus language impairments: A qualitative analysis. *Journal of Communication Disorders, 44,* 23-36.

Rutter, M., Bailey, A., Berument, S.K., Le Couteur, A., Lord, C. & Pickles, A. (2003) *Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.

Semel, E., Wiig, E., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals-Fourth Edition (UK).* San Antonio, TX: Psychological Corporation.

Skuse, D., Warrington, R., Bishop, D., Chowdhury, U., Lau, J., Mandy, W., et al. (2004). The developmental, dimensional and diagnostic interview (3di): A novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 548-558.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics, 12*, 369-373.

Tager-Flusberg, H. (2004). Strategies for conducting research on language in autism. *Journal of Autism and Developmental Disorders, 34*, 75-80.

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

van der Lely, H.K.J., & Howard, D. (1993). Specifically language impaired children: Linguistic impairment or short-term memory deficit? *Journal of Speech and Hearing Research*, 37, 1193-1207.

Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., et al. (2008). A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine, 359*, 2337-2345.

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children - Fourth Edition* (UK). London: Pearson Psychological Corporation.

Whitehouse, A.J.O., Barry, J.G., & Bishop, D.V.M. (2008). Further defining the language impairment of autism: Is there a specific language impairment subtype? *Journal of Communication Disorders, 41,* 319-336.

Whitehouse, A. J. O., Bishop, D. V. M., Ang, Q. W., Pennel, C. E., & Fisher, S. E. (2011) CNTNAP2 variants affect early language development in the general population. *Genes, Brain and Behavior, 10,* 451 – 456

Williams, D.M., Botting, N., & Boucher, J. (2008). Language in autism and specific language impairment: Where are the links? *Psychological Bulletin, 134*, 944-963.

Williams, D.M., & Lind, S.E. (forthcoming). Comorbidity and diagnosis of developmental disorders: What we know and what we need to know. In C. R. Marshall (Ed.), *Current issues in developmental psychology*. London: Psychology Press.

Williams, D., Happe, F., & Jarrold, C. (2008). Intact inner speech use in autism spectrum disorder: Evidence from a short-term memory task. *Journal of Child Psychology and Psychiatry, 49*(1), 51-58.

Table 1: Participant characteristics

	Group				
	CA-TD (n = 20)	VMA-TD (n = 19)	ASD-LI (n = 17)	SLI (n = 15)	Cohen's d[a]
Age	11.69 (0.36)	6.69 (0.33)	12.38 (1.46)	12.73 (2.12)	0.19
VMA	-	7.61 (0.95)	7.62 (1.13)	7.52 (1.34)	0.08
CLS	-	114.21 (16.78)	57.82 (11.50)	59.40 (10.59)	0.14
NVIQ	-	106.68 (12.01)	92.47 (12.12)	89.87 (12.18)	0.21
DS	-	5.31 (1.08)[b]	4.71 (0.99)	4.60 (0.74)	0.13
SCQ	-		18.65 (7.66)	7.00 (4.02)	1.91

CLS = Core language score; DS = Digit span

[a]Effect size for the contrast between the clinical (ASD-LI & SLI) groups

[b]Based on 16/19 VMA-TD participants

Table 2: Mean standard scores on each subtest of the CELF among participants with ASD-LI and participants with SLI

	Group			
	ASD-LI	SLI	*p*	Cohen's d
Subtest				
Concepts & Following Directions	2.36 (1.79)	3.44 (3.09)	.37	0.43
Recalling Sentences	3.71 (2.57)	2.27 (1.54)	.06	0.68
Formulated Sentences	2.94 (2.22)	2.67 (2.19)	.73	0.12
Word Classes Receptive	4.18 (2.01)	4.79 (1.89)	.40	0.31
Word Classes Expressive	3.88 (2.40)	4.57 (2.41)	.43	0.29

Table 3: Proportion (SD) of each kind of cluster production error made within each group

Error type	Group		
	ASD-LI	SLI	VMA-TD
C1 deletion	.02 (.06)	< .01 (.01)	.01 (.03)
C2 deletion	.66 (.18)	.49 (.23)	.48 (.29)
C1 substitution	.11 (.08)	.15 (.12)	.14 (.13)
C2 substitution	.10 (.13)	.11 (.15)	.24 (.21)
Other error	.11 (.11)	.25 (.23)	.12 (.14)

Figure 1: Percentage of 3 syllable and 4 syllable non-words repeated correctly by participants from each diagnostic group

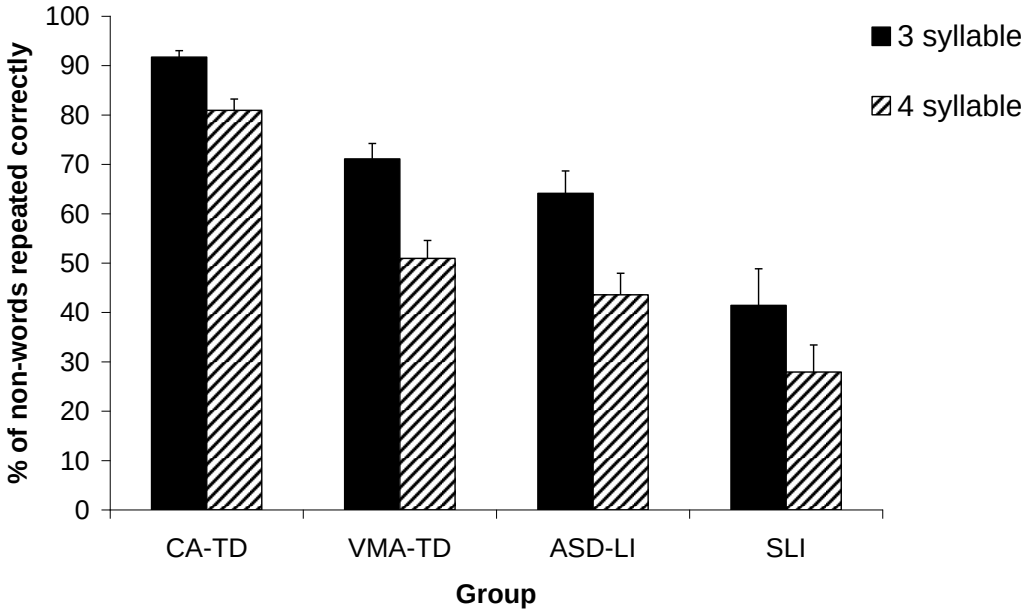
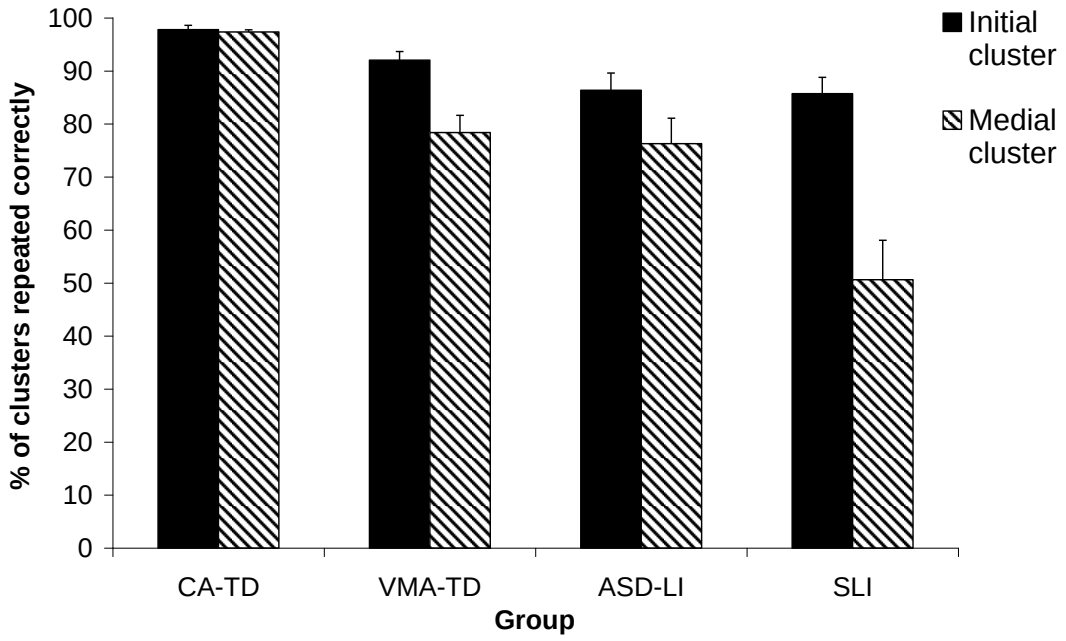


Figure 2: Percentage of initially-positioned and medially-positioned consonant clusters repeated correctly by participants from each diagnostic group



Footnotes

1. Three participants with ASD scored under the ASD cut-off of 15 on the SCQ, with scores of 12, 13, and 4, respectively. Two of these participants (who scored 12 and 13) were taking part in another of our studies and, as a result, had data from another measure of ASD severity; the Developmental, Dimensional and Diagnostic Interview (3di; Skuse et al., 2004), a more detailed parent interview schedule that is based on the Autism Diagnostic Interview (Le Couteur et al., 2003). Each of these participants scored well above the ASD cut-off on the 3di. The final participant (who scored 4 on the SCQ) did not have any other parent report data available. However, this participant had been diagnosed by a UK-leading clinician in London. This, combined with our clinical impression of the child, make us entirely confident of his diagnosis, regardless of his score on the SCQ.

2. It may strike the reader as out of keeping with a diagnosis of *specific* language impairment that four participants in the SLI group had a NVIQ score outside of the normal range. However, a notable proportion of individuals who manifest SLI early in life (during the period in which most receive a diagnosis) show below average NVIQ later in life (e.g., Conti-Ramsden, Botting et al. 2001; Mawhood, Howlin, & Rutter, 2000), with NVIQ dropping as much as 20 points across time (Botting, 2005). Therefore, to obtain a representative sample of children who receive a diagnosis of SLI, we did not set as an inclusion criterion that NVIQ be in the normal range. We did, however, decide that in cases where NVIQ was below 80, non-verbal abilities must be substantially superior to verbal abilities, reflecting the fact that language problems are identified as the most clinically significant feature among all participants with SLI.

3. In fact, it is important to note that ADOS-G and ADI-R appear to have a surprisingly low specificity, which raises questions about their suitability for use in research studies. In the largest study of its kind, Risi, Lord, Gotham, Corsello, Chrysler et al. (2006) found that, if used in isolation, the specificity of each measure was less than 50%, with identifying around 29% of non-spectrum children as having autism. If used together, specificity is improved, but in over 15% of cases the instruments disagree on spectrum vs. non-spectrum diagnoses. These measures (particularly the ADI) have come under recent scrutiny, with some offering what we view as persuasive arguments that measures such as the SCQ and 3di are preferable to ADOS and ADI for research purposes (Bishop, 2011, May 30).

Appendix 1: Nonword stimuli

3 syllable, initial cluster, suffix	3 syllable, initial cluster, no suffix	3 syllable, medial cluster, suffix	3 syllable, medial cluster, no suffix
krifyist	krifyimp	kifryist	kifryimp
drupōling	drupōlif	duprōling	duprōlif
blofitid	blofitim	boflitid	boflitim
klopishiz	klopishiv	koplishiz	koplishiv
frakōping	frakōpif	fakrōping	fakrōpif
pridisiz	pridisif	pidrisiz	pidrisif
flebitist	flebitimp	feblitist	feblitimp
plakytid	plakytif	paklytid	paklytif

4 syllable, initial cluster, suffix	4 syllable, initial cluster, no suffix	4 syllable, medial cluster, suffix	4 syllable, medial cluster, no suffix
krifytadist	krifytadimp	kifrytadist	kifrytadimp
drupōlating	drupōlatif	duprōlating	duprōlatif
blofisetid	blofisetim	boflisetid	boflisetim
klopimishiz	klopimishiv	koplimishiz	koplimishiv
frakōleping	frakōlepif	fakrōleping	fakrōlepif
pridisiliz	pridisilif	pidrisiliz	pidrisilif
flebitorgist	flebitorgimp	feblitorgist	feblitorgimp
plakynūtid	plakynūtif	paklynūtid	paklynūtif