Developmental Science (2014), pp 1–9 DOI: 10.1111/desc.12202

SHORT REPORT

Sentence repetition is a measure of children's language skills rather than working memory limitations

Marianne Klem,[1] Monica Melby-Lervåg,[1] Bente Hagtvet,[1] Solveig-Alma Halaas Lyster,[1] Jan-Eric Gustafsson[2] and Charles Hulme[1,3]

1. *Department of Special Needs Education, University of Oslo, Norway*
2. *Department of Education and Special Education, University of Gothenburg, Sweden and Centre for Educational Measurement (CEMO), University of Oslo, Norway*
3. *Division of Psychology and Language Sciences, University College London, UK*

Abstract

Sentence repetition tasks are widely used in the diagnosis and assessment of children with language difficulties. This paper seeks to clarify the nature of sentence repetition tasks and their relationship to other language skills. We present the results from a 2-year longitudinal study of 216 children. Children were assessed on measures of sentence repetition, vocabulary knowledge and grammatical skills three times at approximately yearly intervals starting at age 4. Sentence repetition was not a unique longitudinal predictor of the growth of language skills. A unidimensional language latent factor (defined by sentence repetition, vocabulary knowledge and grammatical skills) provided an excellent fit to the data, and language abilities showed a high degree of longitudinal stability. Sentence repetition is best seen as a reflection of an underlying language ability factor rather than as a measure of a separate construct with a specific role in language processing. Sentence repetition appears to be a valuable tool for language assessment because it draws upon a wide range of language processing skills.

Research highlights

- Sentence repetition tasks are widely used clinically for identifying children with language difficulties, but the underlying abilities measured are poorly understood, and how performance on this test should be interpreted is unclear.
- We investigated theories about the nature of sentence repetition and its longitudinal relationship to other measures of language ability.
- We found no support for a view that sees sentence repetition as a measure of a separate memory component that has a causal influence on the development of language skills.
- We argue that sentence repetition is best seen as a measure of an underlying unitary language construct.

Introduction

Sentence repetition tasks have been used for decades as a tool for investigating language skills (e.g. McDade, Simpson & Lamb, 1982, Rodd & Braine, 1971; Schwartz & Daly, 1978) and are widely recognized as useful measures of individual differences in language ability and as a means for identifying children with language impairments (Vinther, 2002). In fact, it has been suggested that sentence repetition may be the best single test for identifying children with Specific Language Impairment (Conti-Ramsden, Botting & Faragher, 2001). Sentence repetition tasks are widely used clinically and have been considered especially useful because they appear to be sensitive to residual language processing weaknesses which may not be detected on other expressive and receptive language tasks (Conti-Ramsden et al., 2001; Stothard, Snowling, Bishop, Chipchase & Kaplan,

Address for correspondence: Charles Hulme, Division of Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield St, London WCIN 2PF, UK; e-mail: c.hulme@ucl.ac.uk

1998). However, the abilities measured by sentence repetition tasks are as yet poorly understood (Alloway & Gathercole, 2005a; Riches, 2012) and there are different accounts of the relationship between sentence repetition and language skills. The present study aims to clarify the relationship between sentence repetition and other measures of language ability and establish whether sentence repetition is a longitudinal predictor of the growth of language skills in children.

One influential theoretical account suggests that sentence repetition taps a distinct memory system which is 'uniquely linked with language skills' (Alloway & Gathercole, 2005b, p. 279). According to Baddeley's (2000) revised multi-component working memory model, sentence repetition tasks are assumed to measure primarily the capacity of the episodic buffer, and to a lesser extent the phonological loop (Alloway, Gathercole, Willis & Adams, 2004). The episodic buffer is argued to account for the association between the subsystems of working memory and long-term memory (Repovš & Baddeley, 2006), and is seen as a limited capacity temporary storage system responsible for holding integrated chunks or episodic representations in a multidimensional code (Baddeley, 2000, 2012).

According to the working memory framework, the episodic buffer 'may provide an important gateway for learning' (Alloway *et al*., 2004, p. 87) and it has been suggested that the 'episodic buffer is the source of capacity limits in language processing' (Boyle, Lindell & Kidd, 2013, p. 234). Following on from this view, teaching mnemonic strategies has been proposed as a way of treating children with language learning difficulties (Alloway & Gathercole, 2005b). It has also been claimed that working memory training programmes may be useful for treating language difficulties. For example, Holmes, Gathercole and Dunning (2009) reported that CogMed working memory training produced improvements in children's language comprehension abilities. In summary, according to the working memory account, sentence repetition is seen as a measure of a separate component of memory (the episodic buffer) that has a causal influence on the development of language skills in children.

In contrast to this view, others have argued that sentence repetition is best seen as simply one measure of language ability, rather than as tapping a separate component of memory. This view is perhaps most clearly expressed by MacDonald and Christiansen (2002) who argued that language processing tasks and linguistic working memory tasks should be considered as simply different measures of language processing skill. Similarly, Acheson and MacDonald (2009b) argue that the mechanisms used in verbal recall tasks and language production are the same and that language and memory tasks are therefore best seen as reflecting a general language construct. Clearly, in order to repeat a sentence correctly, at a minimum the words in the sentence have to be perceived correctly and a motor programme to articulate the words in the correct order has to be generated. It is likely that intermediate language processing mechanisms will be related to the likelihood of repeating a sentence correctly with repetition being easier for sentences that are understood correctly (Komeili & Marshall, 2013). Understanding the sentence will in turn depend on semantic skills including vocabulary knowledge and grammatical skills. From the perspective of models of language processing, sentence repetition can be expected to be influenced by a wide range of language skills including speech perception, vocabulary knowledge, grammatical processing and speech production. In this light it is hardly surprising that sentence repetition ability is a correlate of many different language processing skills.

Claims of a close relationship between sentence repetition ability and language skills has gained support from studies of immediate verbal memory, as sentence repetition tasks have similar demands to other immediate verbal memory tasks. There is a large literature suggesting that immediate verbal memory performance is intimately related to variations in the functioning of a variety of language processing mechanisms. For example Walker and Hulme (1999) showed that spoken lists of unrelated concrete words were repeated considerably more accurately than lists of abstract words in an immediate serial recall task. This finding provides clear evidence for the importance of semantic representations in immediate memory tasks. Similarly Allen and Hulme (2006) found that word concreteness, word frequency, and word phonological neighborhood size all affected how well spoken word lists could be repeated. In the same study they provided evidence that these effects reflected differences in the efficiency with which speech production codes could be accessed from semantic representations of words.

Further, Acheson and MacDonald (2009a) reported a series of experiments examining the effects of phonological similarity on measures of immediate memory performance and speech production. Previous work had established that phonologically similar (rhyming) words are harder to recall in immediate serial tasks than phonologically distinctive word lists. They replicated this effect and showed that similar patterns of speech errors occurred when rapidly reading lists of confusable items as when recalling them. These results were interpreted as evidence that speech production mechanisms played a critical role in accounting for immediate verbal recall performance.

Finally, Melby-Lervåg and Hulme (2010) showed that teaching children the meanings of words or teaching them to segment spoken words into phonemes both produced improvements in serial recall performance for the trained words. These findings provide evidence for a causal link between the quality of underlying phonological and semantic representations of words and how well children can repeat lists of those words. There seems little doubt that these same word-level mechanisms will be important influences on performance in sentence repetition tasks, but in addition grammatical skills are likely to be an additional influence on performance.

The view that sentence repetition in children is critically dependent on broad oral language skills also gets support from a recent study by Moll, Hulme, Nag and Snowling (2013). These authors found that dyslexic children performed more poorly on a sentence repetition task than typically developing control children. Importantly, however, the difference between the dyslexic and control children appeared to be entirely accounted for by a subset of children with dyslexia who according to parental reports had experienced delays in oral language development and at the time of assessment showed deficits on a test of morphological awareness. In contrast, the differences in sentence repetition ability between groups were not related to differences in measures of word recall or nonword repetition ability. These findings suggest a critical role for non-phonological language abilities in sentence repetition and it was concluded that 'the memory demands of sentence repetition should not be viewed as distinct from those involved in language production' (Moll *et al*., 2013, p. 2). Other recent studies also provide further empirical support for this claim by demonstrating that many aspects of linguistic knowledge (including morphological, grammatical and phonological processes) are involved during sentence repetition (e.g. Nag, Mirkovic & Snowling, 2013).

The view that sentence repetition taps a broad range of language processing skills is also encapsulated in the Regeneration Hypothesis (Lombardi & Potter, 1992; Potter & Lombardi, 1990). According to this hypothesis the process of repeating a sentence starts from a conceptual (meaning-based) representation of the sentence to be recalled and essentially involves all levels of the language production system (Bock & Levelt, 1994). Potter and Lombardi (1990) provided support for the Regeneration Hypothesis through a series of experiments showing that the patterns of intrusions that occur in sentence repetition tasks are consistent with participants using a conceptual representation of the sentence to be repeated. They also showed that if participants included an incorrect word in their repetition of a sentence, the syntax of the sentence was typically modified to accommodate such intrusion errors.

Finally, the literature on the factorial structure of language abilities is relevant to the debate about sentence repetition ability and how it relates to other aspects of language. There is evidence to support a model which sees individual differences in language skill as reflecting a unidimensional construct in young children (Colledge, Bishop, Koeppen-Schomerus, Price, Happé, Eley & Plomin, 2002; Klem, Gustafsson & Hagtvet, in press; Tomblin & Zhang, 2006). Evidence from Tomblin and Zhang is of particular relevance to the current study. They reported analyses of data from a sample of almost 2000 children (average age 6:04) who were given five subtests from the Test of Language Development (TOLD) assessing receptive vocabulary (picture identification), expressive vocabulary (oral vocabulary), receptive grammatical skills (grammatical understanding), and expressive language use (grammatical completion and sentence imitation). The TOLD sentence imitation measure is directly comparable to the test of sentence repetition used in the current study (the child hears a series of sentences and has to repeat each one). Tomblin and Zhang reported a confirmatory factor analysis of their measures which provided support for a unidimensional language factor which accounted for 51% of the variance in language test scores. In this model, it is notable that sentence imitation (i.e. sentence repetition) loaded strongly on the unitary language factor (factor loading of 0.79, which explained 62% of the variance in sentence imitation ability).

In the present paper we report the results from a large-scale longitudinal study of over 200 children assessed three times between the ages of 4 and 6 years. At each time point we assessed sentence repetition ability as well as vocabulary knowledge and grammatical skills. This design allows us to investigate the longitudinal relationship between sentence repetition and these other measures of language ability. If sentence repetition is best thought of as measuring separable working memory systems (the episodic buffer and phonological loop) which are causal influences on the rate of growth in language skills, sentence repetition should be a longitudinal predictor of variations in the growth of other language skills. On the other hand, if sentence repetition is best considered as simply another measure of a broad language ability construct rather than a distinct ability, it should load strongly on a unitary language factor. Thus, the aims of the current study were to assess the extent to which sentence repetition is a longitudinal predictor of variations in other language abilities, and whether it is best conceptualized as a separate ability or simply one measure of a unitary underlying language construct (defined by vocabulary knowledge, grammatical skills and sentence repetition).

Method

Participants

Two hundred and sixteen children (104 girls, 112 boys) were recruited, with parental consent, from Norwegian day-care centers. Based on information from the school psychological services at the beginning of the study, children with Norwegian as a second language, children diagnosed with general learning disabilities or children with sensory impairments were excluded from the study.

Design and procedure

Children were tested individually by trained assistants in a separate room in the children's day-care centers at Time 1 (*M* age = 51.15 months, *SD* = 2.24) and Time 2 (M age = 62.89 months, *SD* = 2.38). At Time 3 children were typically assessed in school (*M* age = 75.04 months, *SD* = 2.28). Tests were administered in a fixed order, over three (at Time 1) or two sessions (at Times 2 and 3) and were part of a comprehensive test battery. Only the measures relevant to the current study are described here.

Tests and materials

Sentence repetition

The children were given an adapted version of the Sentence Memory subtest of the Norwegian language screening test 'Language 6-16' (Språk 6-16; Ottem & Frost, 2005). This includes 21 sentences of increasing length and complexity (see examples in Appendix A). The original 16-item subtest was supplemented with the first five items of the sentence repetition test from the Norwegian version of WPSSI-R (Wechsler, 1990) in order to avoid floor-effects at Time 1. Sentences were spoken by the examiner and the child was required to repeat back exactly what was said to them. Testing was stopped after three consecutive incorrect repetitions.

Vocabulary

A Norwegian version of the British Picture Vocabulary Scale-II (Dunn, Dunn, Whetton & Burley, 1997) was used to measure vocabulary. In this test the child responds by pointing to one of four line drawings that corresponds to the word spoken by the test administrator. The test has 144 items. Testing was stopped when the child erred on eight out of the last 12 items presented.

Grammatical knowledge

The Norwegian version of the 'Grammatic Closure' subtest from the Illinois Test of Psycholinguistic Abilities (ITPA; Kirk, McCarthy & Kirk, 1968) was used to measure expressive grammatical knowledge. In this test the child was shown a series of pictures along with corresponding spoken sentences and then asked to fill in blanks in these unfinished sentences. The test includes 33 items covering a wide range of grammatical constructions (including the inflection of nouns, verbs, and adjectives). Testing was stopped after six consecutive incorrect items at Times 1 and 2, whilst all items were presented at Time 3.

Results

Mean raw scores and standard deviations for all variables at all three time points are shown in Table 1. The correlations between all measures at all three time points are shown in Table 2.

To assess the relationship between Sentence Repetition and other aspects of oral language skills we used a confirmatory approach based on structural equation modeling. To begin with we investigated the concurrent and longitudinal relationships between sentence repetition and language ability (considered as separate constructs). In order to do this we constructed a bivariate simplex model in which a latent language factor was defined by two indicators (vocabulary and grammatical knowledge) and a sentence repetition latent variable was defined by a single indicator. Simplex models estimate the stability of a construct over time by regressing later measures of a construct onto the same construct at previous time points (Bollen & Zimmer, 2010). Stability of a construct refers to the preservation of rank order of individuals relative to the group, across repeated times of measurement. In a bivariate simplex model, the longitudinal cross loadings between the different constructs could provide information about the impact of one construct upon the other 'above and beyond the autoregressive prediction of that construct on itself' (Curran & Bollen, 2001, p. 113), which in turn may be interpreted as evidence of a causal influence. The simplex model used is shown in Figure 1 and fitted the data very well ($\chi^2 = 19.77$, $p = .41$, $df = 19$, RMSEA = 0.01 CI_{90} 0.00–0.06, CFI = 1.00, TLI = 1.00, SRMR = 0.03). The nonsignificant value of χ^2 for this model indicates a very good fit of the data to the model (the

Table 1 *Means, standard deviations, ranges and reliability for all measures at all time points*

	Time 1 (age 4)			Time 2 (age 5)			Time 3 (age 6)		
Measure	Mean (*SD*)	Min–Max	α	Mean (*SD*)	Min–Max	α	Mean (*SD*)	Min–Max	α
BPVS	42.09 (11.11)	15–75	.91	57.05 (11.26)	28–90	.91	72.96 (11.79)	43–104	.91
GramClos	11.05 (3.79)	1–24	.73	14.00 (3.70)	4–24	.69	18.47 (3.75)	8–27	.69
Sentence Rep	6.56 (1.91)	2–12	.63	8.00 (2.11)	1–16	.70	9.17 (2.06)	4–17	.72

Note: All test scores = raw scores (number correct items); BPVS = British Picture Vocabulary Scale-II; GramClos = Grammatic Closure; Sentence Rep = Sentence Repetition; *SD* = Standard deviation; α = Cronbach's alpha; Min–Max = Range of scores in sample.

Table 2 *Correlations between all measures at all time points*

		1	2	3	4	5	6	7	8	9
1	SR T1	–								
2	BPVS T1	.306	–							
3	GC T1	.334	.368	–						
4	SR T2	.479	.207	.266	–					
5	BPVS T2	.338	.518	.288	.274	–				
6	GC T2	.333	.306	.371	.314	.409	–			
7	SR T3	.337	.164	.294	.529	.291	.326	–		
8	BPVS T3	.196	.449	.311	.240	.576	.324	.248	–	
9	GC T3	.320	.267	.363	.173	.337	.560	.298	.404	–

Note: SR = Sentence Repetition; BPVS = British Picture Vocabulary Scale-II; GC = Grammatic Closure; T = time.

observed covariance matrix from the data does not differ significantly from the implied covariance matrix generated from the model).

In this model the error variance for the sentence repetition latent variable was fixed to be equal to the observed error variance on the task at each time point (based on Cronbach's alpha for the sample). Both the language and sentence repetition latent variables show a high degree of longitudinal stability (particularly the language latent variable between Times 2 and 3). We tested for invariance of factor loadings to investigate whether the measures showed equivalent relationships to the latent constructs across time. The model shown in Figure 1 is the constrained model (i.e. where the unstandardized factor loadings between the measures and the latent variable are constrained to be equal over time; the chi-square test of this constrained model against the unconstrained model was not significant $\Delta\chi^2 = 0.29$, $df = 2$). This provides support for factorial invariance of the two latent constructs over time; i.e. the latent variables representing sentence repetition and language ability can be interpreted as representing the same construct at each time point.

If the longitudinal cross loadings between earlier measures of sentence repetition and later measures of language in this model had been significant, they could have been interpreted as support for the hypothesis that sentence repetition may be a causal constraint on the development of language skills. However, neither of the possible cross loadings between earlier measures of sentence repetition and later measures of language were significant. The only statistically significant cross loading here is that between language at Time 2 and sentence repetition at Time 3, indicating that variation in language ability at Time 2 predicts additional variability in sentence repetition at Time 3, over and above the autoregressive effect (i.e. what is already predicted by sentence repetition at the previous time point). This pattern offers some support to the opposite causal hypothesis; that earlier language skills are causally related to later improvements in sentence repetition ability.

In summary, the model shown in Figure 1 postulates that sentence repetition is distinct from a general language factor. However, the model fails to provide support for the theory postulating that sentence repetition is a measure of working memory capacity which has a causal influence on the development of language skills. It is notable that the only significant cross loading we found (language ability at Time 2 predicts additional variance in sentence repetition ability at Time 3) actually provides some evidence for the opposite longitudinal relationship. This finding, in combination with the strong correlation at Time 1 between levels of language

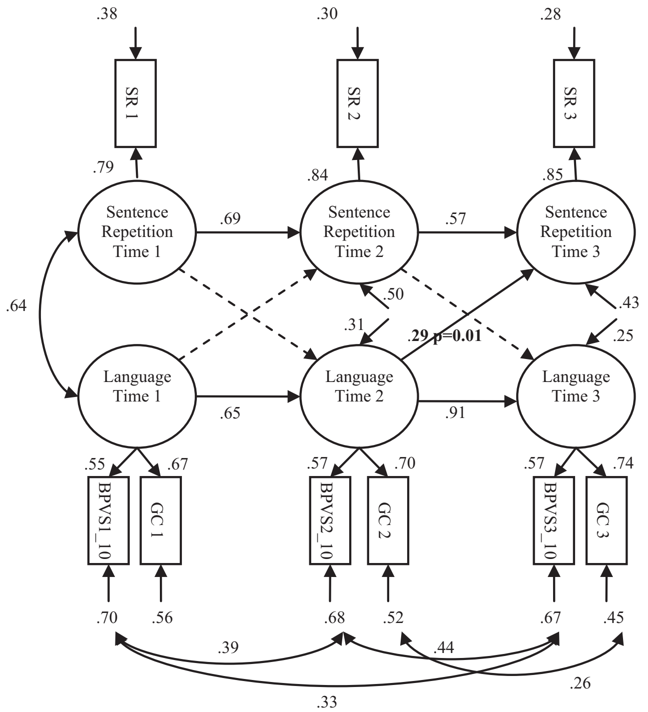

Figure 1 *Bivariate simplex model of the longitudinal relationship between sentence repetition and language ability. Ellipses represent latent variables and rectangles represent observed variables.*
Note: BPVS_10 = British Picture Vocabulary Scale-II/10 (full scale divided by 10), GC = Grammatic Closure, SR = Sentence Repetition. Paths with solid lines = significant at p < .01. Dashed lines = non-significant paths (p > .05).

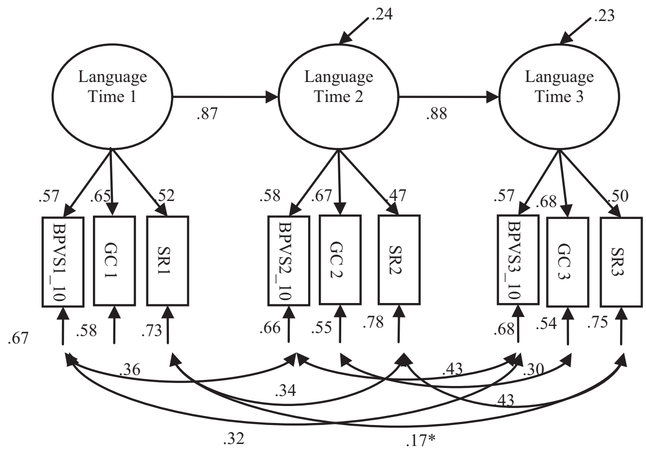

Figure 2 *Unidimensional simplex model of language ability in children. Ellipses represent latent variables and rectangles represent observed variables.*
*Note: BPVS_10 = British Picture Vocabulary Scale-II/10 (full scale divided by 10), GC = Grammatic Closure, SR = Sentence Repetition. Paths with solid lines = significant at p < .01. * = p < .05.*

ability and sentence repetition ability in this model (r = .64), raises the question whether it would not be more parsimonious to simply consider the sentence repetition and language latent variables as measures of the same underlying construct (as proposed by Tomblin & Zhang, 2006).

The model shown in Figure 2 evaluates this alternative theory with a model in which vocabulary, grammatical knowledge and sentence repetition define a single language latent ability factor. The model provides an excellent fit to the data (χ^2 = 21.30, p = .50, df = 22, RMSEA = 0.00 CI_{90} 0.00–0.06, CFI = 1.00, TLI = 1.00, SRMR = 0.03) and the language ability latent factor shows very high longitudinal stability. In this model the unstandardized factor loadings between the latent factor and each measure were constrained to be equal at each time point, and the chi-square difference test revealed no significant difference between this model and the unconstrained model ($\Delta\ \chi^2$ = 1.47, df = 4). Hence we have support for the factorial invariance of this unitary language factor across time.

The models in Figures 1 and 2 are not nested so we cannot test which model provides a statistically better fit to the data (and both models fit the data very well). Arguably, however, the model in Figure 2 provides a more parsimonious account and provides support for earlier theories that postulate that language skills may, to a first approximation, be described as reflecting a unitary factor (Tomblin & Zhang, 2006).

Discussion

We have reported the results of a large-scale longitudinal study of children's language development spanning a period of rapid developmental change (age 4 to 6 years). Our particular focus has been on sentence repetition, a test that has been widely used clinically for identifying children with language impairments. Our specific concern has been with theories about the nature of sentence repetition and how performance on this test should be interpreted.

The main motivation for separating sentence repetition from other measures of language ability (as in the model presented in Figure 1) would be a theory that sees sentence repetition as reflecting a system (or systems) that is believed to be a cognitive 'primitive' that places constraints on the development of other language skills. This is the view advocated by the multicomponent working memory model (Baddeley, 2000, 2012) which sees sentence repetition tasks as depending primarily upon individual differences in the efficiency of an episodic buffer (Alloway *et al*., 2004), which is believed to be a system that places constraints on language acquisition (Boyle *et al*., 2013). This view leads to the prediction that sentence repetition should be a longitudinal predictor of individual differences in language development. In fact, however, there was no evidence from the current study that sentence repetition is a longitudinal predictor of the variations in the growth of language skills.

This absence of longitudinal relationships between earlier sentence repetition and later language skills, and the fact that sentence repetition correlated well with other measures of language ability, leads us to favor the view that sentence repetition is best conceptualized as a measure of language ability (as presented in Figure 2). We would argue that sentence repetition is best seen as a complex linguistic task that reflects the integrity of language processing systems at many different levels (speech perception, lexical (vocabulary) knowledge, grammatical skills and speech production to name but a few). The model evaluated in Figure 2 shows that sentence repetition loads strongly on a unidimensional language factor and in this respect confirms and extends earlier findings from Tomblin and Zhang (2006) using different instruments but similar measures.

Our view of sentence repetition aligns most closely with the Regeneration Hypothesis (Lombardi & Potter, 1992; Potter & Lombardi, 1990). In this view sentence repetition is a multi-faceted task that engages virtually all aspects of language processing. Most critically in relation to the Regeneration Hypothesis, the listener after hearing a sentence creates a conceptual (rather than form-based) representation of the sentence that is to be recalled. Having generated a conceptual representation of the sentence they have heard in order to repeat it, the listener then goes through a series of processes including activating relevant lexical (word) knowledge, grammatical encoding, and the processes involved in phonological realization and speech production.

Our conceptualization of sentence repetition as a reflection of a common underlying language construct is captured in the model depicted in Figure 2. In terms of causality it is important to note that the latent language variable is specified in terms of a reflective measurement model (Edwards & Bagozzi, 2000); that is, the direction of influence is from the latent variable to the indicators (Brown, 2006). It is usual to interpret these relations as causal, in the sense that variations in the observed variables can be seen as being a reflection of variations in a common cause (the latent variable; Bollen, 2002; Borsboom, Mellenbergh & van Heerden, 2003).

In summary, we found no evidence to support the view that sentence repetition taps a causal constraint on the development of language skills in young children. We argue that our findings provide support for the view that sentence repetition is best seen as a reflection of an underlying unitary language construct rather than being conceptualized as a separate construct. Furthermore, we suggest that the Regeneration Hypothesis helps to explain the sensitivity of sentence repetition as an assessment tool for children with language impairments. From this perspective it is clear that sentence repetition, because it involves so many diverse aspects of language comprehension and production, can fail for many different reasons. It is clear that a promising line of research is to examine in greater detail the forms of errors that children make in sentence repetition tasks (see Nag *et al*., 2013; Moll *et al*., 2013). The forms of errors made by children may be diagnostic of the different underlying language processing difficulties seen in different children. We should also stress that, ultimately, longitudinal studies alone cannot provide sufficient evidence to test claims about the possible causal relationship between sentence repetition and other aspects of language skills – such causal theories can only be fully evaluated using training studies.

References

Acheson, D.J., & MacDonald, M.C. (2009a). Twisting tongues and memories: explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language*, **60** (3), 329–350. doi:http://dx.doi.org/10.1016/j.jml.2008.12.002

Acheson, D.J., & MacDonald, M.C. (2009b). Verbal working memory and language production: common approaches to the serial ordering of verbal information. *Psychological Bulletin*, **135** (1), 50–68. doi:10.1037/a0014411

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, **55** (1), 64–88. doi:http://dx.doi.org/10.1016/j.jml.2006.02.002

Alloway, T.P., & Gathercole, S.E. (2005a). Working memory and short-term sentence recall in young children. *European Journal of Cognitive Psychology*, **17** (2), 207–220. doi:10.1080/09541440440000005

Alloway, T.P., & Gathercole, S.E. (2005b). The role of sentence recall in reading and language skills of children with learning difficulties. *Learning and Individual Differences*, **15** (4), 271–282. doi:10.1016/j.lindif.2005.05.001

Alloway, T.P., Gathercole, S.E., Willis, C., & Adams, A.-M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, **87** (2), 85–106. doi:10.1016/j.jecp.2003.10.002

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, **4** (11), 417–423. doi:http://dx.doi.org/10.1016/S1364-6613(00)01538

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, **63**, 1–29. doi:10.1146/annurev-psych-120710-100422

Bock, K., & Levelt, W.J.M. (1994). Language production: grammatical encoding. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Bollen, K.A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, **53** (1), 605–634. doi:10.1146/annurev.psych.53.100901.135239

Bollen, K.A., & Zimmer, C. (2010). An overview of the Autoregressive Latent Trajectory (ALT) model. In K. van Montfort, J. Oud & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 153–175). Berlin: Springer.

Borsboom, D., Mellenbergh, G.J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, **110** (2), 203–219. doi:10.1037/0033-295X.110.2.203

Boyle, W., Lindell, A.K., & Kidd, E. (2013). Investigating the role of verbal working memory in young children's sentence comprehension. *Language Learning*, **63** (2), 211–242. doi:10.1111/lang.12003

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Colledge, E., Bishop, D.V.M., Koeppen-Schomerus, G., Price, T.S., Happé, F.G.E., Eley, T.C., & Plomin, R. (2002). The structure of language abilities at 4 years: a twin study. *Developmental Psychology*, **38** (5), 749–757. doi:10.1037/0012-1649.38.5.749

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology & Psychiatry & Allied Disciplines*, **42** (6), 741–748. doi:10.1111/1469-7610.00770

Curran, P.J., & Bollen, K.A. (2001). The best of both worlds: combining autoregressive and latent curve models. In L. M.C.A.G. Sayer (Ed.), *New methods for the analysis of change* (pp. 107–135). Washington, DC: American Psychological Association.

Dunn, L.M., Dunn, L.M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale (Second ed.)*. London: GL Assessment. The Norwegian version by Lyster, S.-A.H., Horn, E., & Rygvold, A.-L. (2010).

Edwards, J.R., & Bagozzi, R.P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, **5** (2), 155–174. doi:10.1037/1082-989X.5.2.155

Holmes, J., Gathercole, S.E., & Dunning, D.L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, **12**, F9–F15. doi:10.1111/j.1467-7687.2009.00848.x

Kirk, S.A., McCarthy, J.D., & Kirk, W.D. (1968). *Illinois test of psycholinguistic abilities.* Urbana, IL: University of Illinois press. Norwegian version by Gjessing, H.J. & Nygaard, H.D. (1975).

Klem, M., Gustafsson, J.-E., & Hagtvet, B.E. (in press). The dimensionality of language abilities in 4-year-olds: construct validation of a language screening tool. *Scandinavian Journal of Educational Research*. doi: 10.1080/00313831.2014.904416

Komeili, M., & Marshall, C.R. (2013). Sentence repetition as a measure of morphosyntax in monolingual and bilingual children. *Clinical Linguistics & Phonetics*, **27** (2), 152–162. doi:10.3109/02699206.2012.751625

Lombardi, L., & Potter, M.C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language*, **31** (6), 713–733. doi:http://dx.doi.org/10.1016/0749-596X(92)90036-W

McDade, H.L., Simpson, M.A., & Lamb, D.E. (1982). The use of elicited imitation as a measure of expressive grammar: a question of validity. *Journal of Speech and Hearing Disorders*, **47**, 19–24.

MacDonald, M.C., & Christiansen, M.H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, **109** (1), 35–54. doi:10.1037/0033-295X.109.1.35

Melby-Lervåg, M., & Hulme, C. (2010). Serial and free recall in children can be improved by training: evidence for the importance of phonological and semantic representations in immediate memory tasks. *Psychological Science*, **21** (11), 1694–1700. doi:10.1177/0956797610385355

Moll, K., Hulme, C., Nag, S., & Snowling, M.J. (2013). Sentence repetition as a marker of language skills in children with dyslexia. *Applied Psycholinguistics, FirstView*, **1–19**, doi:10.1017/S0142716413000209

Nag, S., Mirkovic, J., & Snowling, M.J. (2013). Refining the understanding of sentence repetition: individual differences and language production subsystems. Manuscript submitted.

Ottem, E., & Frost, J. (2005). *Språk [LANGUAGE] 6–16: screening test: Manual*. Oslo: Bredtvet kompetansesenter.

Potter, M.C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, **29** (6), 633–654. doi:http://dx.doi.org/10.1016/0749-596X(90)90042-X

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience*, **139** (1), 5–22. http://dx.doi.org/10.1016/j.neuroscience.2005.12.061

Riches, N.G. (2012). Sentence repetition in children with specific language impairment: an investigation of underlying mechanisms. *International Journal of Language & Communication Disorders*, **47** (5), 499–510. doi:10.1111/j.1460-6984.2012.00158.x

Rodd, L.J., & Braine, M.D.S. (1971). Children's imitations of syntactic constructions as a measure of linguistic competence. *Journal of Verbal Learning and Verbal Behaviour*, **10**, 430–443.

Schwartz, A.H., & Daly, D.A. (1978). Elicited imitation in language assessment: a tool for formulating and evaluating treatment. *Journal of Communication Disorders*, **11**, 25–35.

Stothard, S.E., Snowling, M.J., Bishop, D., Chipchase, B.B., & Kaplan, C.A. (1998). Language-impaired preschoolers: a follow-up into adolescence. *Journal of Speech, Language and Hearing Research*, **41** (2), 407–418.

Tomblin, J.B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech,* *Language and Hearing Research*, **49** (6), 1193–1208. doi:10.1044/1092-4388(2006/086

Vinther, T. (2002). Elicited imitation: a brief overview. *International Journal of Applied Linguistics*, **12** (1), 54–73. doi:10.1111/1473-4192.00024

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25** (5), 1256–1271. doi:10.1037/0278-7393.25.5.1256

Wechsler, D.M. (1990). *WPPSI-IIl: Wechsler preschool and primary scale of intelligence – revised*. San Antonio, TX: The Psychological Corporation/Harcourt Assessment.

Received: 12 November 2013
Accepted: 16 April 2014

Appendix A

Example items from the sentence repetition task

Item no.	English translation	Norwegian wording
Early items (from WPPSI-R, Norwegian version):		
1.	Fish swim	Fisk svømmer
2.	Per [name] is happy	Per er glad
Later items (from Language 6-16, sentence memory subtest):		
6.	He ran out again	Han sprang ut igjen
9.	The girl kicked the football over the roof	Jenta sparket fotballen over hustaket
11.	Pears in my garden are better than those in the shop	Pærene i hagen min er bedre enn de i butikken
13.	We know that some children stop crying when we give them something to eat	Vi vet at noen barn slutter å gråte når vi gir dem noe å spise