J. Child Lang. **35** (2008), 247–268.
doi:10.1017/S0305000907008367 Printed in the United Kingdom

Grammaticality judgments in children: The role of age, working memory and phonological ability*

JANET L. McDONALD
Louisiana State University

(*Received* 11 *December* 2006. *Revised* 8 *May* 2007)

ABSTRACT

This paper examines the role of age, working memory span and phonological ability in the mastery of ten different grammatical constructions. Six- through eleven-year-old children ($n=68$) and adults ($n=19$) performed a grammaticality judgment task as well as tests of working memory capacity and receptive phonological ability. Children showed early mastery of some grammatical structures (e.g. word order, article omissions) while even the oldest children differed from adults on others (e.g. past tense, third person singular agreement). Working memory capacity and phonological ability accounted for variance in grammaticality judgments above and beyond age effects. In particular, working memory capacity correlated with structures involving verb morphology and word order; phonological ability was important for structures with low phonetic substance. Children's relative difficulty with the different constructions showed parallels to adult performance under memory load stress, indicating working memory capacity may be a limiting factor in their performance. Implications for performance by memory and phonologically impaired populations are discussed.

INTRODUCTION

Researchers in child language have long been interested in the order in which various grammatical devices are mastered (e.g. Brown, 1973). This paper, while also interested in order of mastery, goes beyond this to

[*] This research was supported by a Faculty Research Grant from Louisiana State University to the author. A preliminary version of this data was presented at the 45th annual meeting of the Psychonomic Society, Minneapolis, MN. I thank Chase Faucheaux, Patrick Johnson, Stephanie Peterson and Marie Nebel-Schwalm for help in gathering the data. Thanks also to the children and parents who participated, and the University Lab School and Trinity Episcopal School for their help. Address for correspondence: Janet L. McDonald, Department of Psychology, Louisiana State University, Baton Rouge, LA 70803, USA. Tel: (225) 578-4116. Fax: (225) 578-4125. e-mail: psmcdo@lsu.edu

investigate the influence of the factors of working memory capacity and phonological ability on mastery of different grammatical constructions. Exploration of these factors is relevant both to the relationship between performance by children and adults, and that between typically developing children and impaired populations. Below we review how constructions may differ in difficulty and how working memory span and phonological ability may selectively impact certain structures.

Construction difficulty in grammaticality judgment tasks

In order to investigate the order in which children master various grammatical constructions, we tested ten different constructions in a grammaticality judgment task. In grammaticality judgment, sentences, both grammatical and ungrammatical, are presented to a child, and he or she judges them for well-formedness. While initial work with grammaticality judgments (or acceptability tasks) found that children often judged sentences for semantic content rather than grammaticality, these tests were often done on preschool aged children (e.g. deVilliers & deVilliers, 1972). Tests on older children showed that they were able to do the task but that grammaticality judgment performance differed by construction type (Hakes, 1980; Sutter & Johnson, 1990; Wulfeck, 1993; Kail, 2004; Wulfeck, Bates, Krupa-Kwiatkowski & Saltzman, 2004). Although most studies have only tested a small subset of constructions, looking across studies it is possible to piece together a rough order of construction difficulty for children performing grammaticality judgments.

Several studies have had six- to twelve-year-olds judge intraphrasal word order violations – these involved the inversion of a modifier and noun (e.g. *man the*) or auxiliary and verb (e.g. *running is*) (Wulfeck, 1993; Kail, 2004; Wulfeck *et al.*, 2004). While not necessarily at adult levels, grammaticality judgment of these types of word order errors was better than other constructions that were tested in these studies such as omission of determiners or auxiliaries, or agreement errors. While intraphrasal word order violations were easy to detect, a study conducted in French, a freer word order language than English, found that children had much more difficulty detecting interphrasal word order errors (Kail, 2004). We test interphrasal word order violations in the current study, but given the stricter word order of English, may still find children are fairly good at detecting these types of errors.

Another structure that seems to be fairly easy to master in grammaticality judgment tasks is the present progressive morpheme *-ing*. Typically developing four- to six-year-old children found it easier to detect a missing *-ing* (e.g. *he is run*) than missing auxiliaries (e.g. *he running*) (Rice, Wexler & Redmond, 1999).

Missing auxiliaries (e.g. *the man running*) and missing determiners (e.g. *man is running*) both seem to be moderate in difficulty for elementary school aged children. Grammaticality judgment scores for these omissions were intermediate in difficulty between intraphrasal word order changes and errors in agreement morphology (Wulfeck *et al.*, 2004).

Across studies, a construction that has proven to be difficult for children to correctly judge is third person singular subject–verb agreement (e.g. *the man are* or *the man run*) (Wulfeck, 1993; Kail, 2004; Wulfeck *et al.*, 2004). Wulfeck (1993) found third person agreement more difficult than plural agreement between determiners and nouns (e.g. *a boys*) and intraphrasal word order violations for a group of six- to seven-year-olds, while Wulfeck *et al.* (2004) found both third person agreement and plural agreement harder than intraphrasal word order violations across an age range of seven to twelve. The difficulty of third person agreement is not specific to grammaticality judgment tasks. Indeed, the optional production of this morpheme in obligatory contexts by children before the ages of three or four has caused theorists to propose a structural account for this difficulty. Known as the Optional Infinitive account, it predicts variability in the production of both agreement and tense markers on verbs for a period of time during development (Wexler, 1994). The account predicts greater difficulty for the *-s* morpheme marking verb agreement in English, than for the *-s* morpheme marking plural, and thus predicts the pattern of results of Wulfeck (1993) rather than Wulfeck *et al.* (2004).

Detection of errors involving irregular forms may also be late to develop. In a study by Redmond & Rice (2001), typically developing five- to eight-year-old children often accepted over-regularized irregular past tense verbs (i.e. *falled* instead of *fell*) in grammaticality judgment, resulting in poorer grammaticality scores than on third person singular subject–verb agreement items. Interestingly, the ability to correct over-regularized forms may develop earlier for irregular nouns than irregular verbs (Cox, 1989).

Thus, gleaning across studies, a rough ordering of mastery of various grammatical structures for children performing a grammaticality judgment task may be (1) simple word order changes, (2) the present progressive morpheme, (3) omitted determiners and auxiliaries, (4) agreement errors, especially those involving the third person singular subject–verb agreement, which may or may not be easier than plural agreement, and (5) irregular forms. For the last, irregular noun plurals may be mastered sooner than irregular verb past tense.

Working memory

Beyond age, another factor that may play into the mastery of grammatical constructions is working memory span. Span increases throughout the elementary school years (Gaulin & Campbell, 1994), and thus structures with high working memory demands may not be mastered until late due to working memory limitations. Working memory has been implicated in the processing of complex syntactic sentences in adults, with individuals with higher memory capacity better able to deal with syntactic ambiguity, and to give more correct sentence interpretations than individuals with low spans (King & Just, 1991; Vos, Gunter, Schriefers & Friederici, 2001). In children, working memory span has been found to be correlated with some aspects of language ability in both typical and atypical populations. For example, in typically developing populations, working memory span significantly correlated to both grammaticality judgment and the ability to correct ungrammatical sentences in third graders (Gottardo, Stanovich & Siegel, 1996), to receptive syntax ability in six- to nine-year-olds (Ellis Weismer, Evans & Hesketh, 1999), and to sentence comprehension in seven- to eight-year-olds (Montgomery, 2000*b*). Children with specific language impairment (SLI) have been shown to have smaller working memory capacity than their peers (Ellis Weismer *et al.*, 1999; Montgomery, 2000*a*, 2000*b*; Marton & Schwartz, 2003), and do more poorly than these peers on tests of grammatical mastery. Thus, working memory may contribute to children's ability to do a grammaticality judgment task, particularly on structures that have higher memory demands – these may be structures that involve comparing across sentential components.

Evidence for what structures may suffer when working memory is inadequate can be gained by looking at the performance of stressed speakers on grammaticality judgment tasks. For example, if adult speakers are placed under working memory stress (i.e. asked to retain digit strings) while performing a grammaticality judgment task, certain structures are more impacted than others. In one study, a small working memory load was found to impact the detection of agreement errors, while determiner omission, auxiliary omission and word order errors were still well detected (Blackwell & Bates, 1995). Another study on adults found that working memory load impacted past tense and irregular plurals while structures such as regular plurals and word order were not significantly affected (McDonald, 2006). In a study that tested five- to seven-year-old children under memory stress, in this case realized by longer sentences rather than digit load, memory stress lowered performance on third person agreement and past tense, but was not so influential on plurals (Hayiou-Thomas, Bishop & Plunkett, 2004). Further evidence for working memory impacting agreement constructions comes from production data with adults – when adults with low working memory spans were placed under memory load they made significantly more errors in subject–verb agreement than high span adults under the same load (Hartsuiker & Barkhuysen, 2006). Thus, structures involving third person agreement and past tense may be the most vulnerable for speakers with small working memory capacity, while plurals, determiner or auxiliary omissions, and word order errors may not be affected.

Phonological ability

Another factor that may influence mastery of particular grammatical constructions is perceptual or receptive phonological ability. In particular, constructions that have hard to hear, or low phonetic substance markings, may be mastered late due to perceptual difficulties. Phonological ability, although measured in a variety of ways in the literature, does relate to grammatical mastery. For example, in typically developing children, phonological sensitivity significantly correlated to grammatical judgment and error correction in third graders (Gottardo *et al.*, 1996), and the ability to identify phonemes correlated significantly to the ability to correct grammatical errors in five-year-olds (Bowey, 2005). Children with SLI have been shown to have impairment in both their ability to discriminate phonemes, and their grammaticality judgment scores for structures that have low phonetic substance markers (i.e. the *-s* of third person singular agreement and the *-ed* of regular past tense), while performance on structures with higher phonetic substance morphemes (i.e. the *-ing* of present progressive) was strong (Montgomery & Leonard, 1998).

We can also look at which constructions may suffer due to poor phonological ability by looking at phonologically stressed speakers performing comprehension or grammaticality judgment tasks. For example, adults have been asked to perform such tasks with an overlay of white noise, which makes phonological processing more difficult. When given a comprehension task under noise conditions, adult speakers of strong morphological languages showed decreased use of subject–verb agreement morphology, while SVO word order was minimally impacted (Kilborn, 1991; Dick, Bates & Ferstl, 2003; note however, interpretations of other word orders may be impacted by noise: e.g. Dick, Bates, Wulfeck, Utman, Dronkers & Gernsbacher, 2001). In performing grammaticality judgments through noise, English-speaking adults were significantly worse than controls at detecting errors in structures whose ungrammatical form was missing an unaccented syllable or a bound morphological marker (e.g. omitted articles or auxiliaries, third person agreement, regular plurals, regular past tense, present progressive) while structures involving more salient errors (word order violations, some irregular forms) were unaffected (McDonald, 2006). In general, constructions whose ungrammatical form differs from its grammatical form by a low phonetic substance unit (i.e. an unaccented function word or short bound morpheme) may cause the most difficulty for speakers with poor phonological ability.

The current study

In the following study, children and adults perform a grammaticality judgment task over ten different grammatical constructions. These include items that involve word order changes, function word omissions, bound regular morphology omissions including plurals, present progressive, past tense, third person agreement, and over-regularized irregular forms for plurals and past tense. Order of mastery is predicted to roughly follow that gleaned from the literature, with word order items showing adult levels of mastery early on, present progressives and function word omissions mastered next, and third person agreement and irregular past tense mastered late.

Working memory capacity and phonological ability are also measured. Working memory measures should correlate to grammatical judgment performance on structures involving high memory load such as third person agreement and past tense. Phonological ability should correlate to performance on items where a hard to hear difference distinguishes grammatical from ungrammatical sentences – these would include structures involving omitted function words, and low phonetic substance bound regular morphemes.

While children's scores on the grammaticality judgment task are likely to be lower than those of adults, it is also interesting to ask if the relative ordering of construction difficulty differs between the two groups. However, since children have lower working memory capacity (Gaulin & Campbell, 1994) and may have poorer phonological ability than adults, their order of construction difficulty may not match that of normal adults. Rather, they may resemble adults who perform grammaticality judgments with impaired working memory capacity or impaired phonological ability. McDonald (2006) had unstressed adults and adults under a variety of different types of cognitive stress perform the same grammaticality judgment task as used here. Therefore, we are able to compare the performance of children not only to the adults in the current study, but also to the unstressed and variously stressed adults in that study. Finding a high correlation in construction difficulty between either unstressed or a particular stressed adult group will give further insight into factors limiting construction mastery in children.

METHOD

Participants

Sixty-eight grade school native English-speaking children participated. The children attended one of two private schools in the community, and thus were largely upper or middle class. Both parental consent and child assent were obtained. Parents were asked if their child had any diagnosed language problem, and if they were seeing a speech/language pathologist. Data from children who had referrals for any language problem other than minor pronunciation difficulties were not used. The children were divided up into three approximately equal groups based on age. The first group (6–7) had 22 six- to seven-year-old children ranging in age from 6;3 to 7;11 (mean 7;1). The second group (8–9½) had 23 eight- to nine-and-a-half-year-old children ranging in age from 8;1 to 9;5 (mean 8;9). The third group (9½–11) had 23 nine-and-a-half- to eleven-year-old children ranging in age from 9;7 to 11;1 (mean 10;3). Nineteen adults ranging in age from 18;2 to 22;6 (mean 20;3) also participated in the experiment. All participants had two native English-speaking parents.

Stimuli

Working memory span. Working memory capacity was measured by a SIZE JUDGMENT TASK, versions of which have successfully been used with children (Montgomery, 2000*a*, 2000*b*). In this task, learners were given a list of concrete nouns which they were instructed to reorder in terms of size of the referent from smallest to largest, and report this ordering. For example, if they heard *pig*, *door*, *butterfly*, they would have to repeat back *butterfly*, *pig*, *door*. Participants got 3 sets each of size 2, 3, 4 and 5 items for a total of 12 sets. Working memory span was computed by taking the highest set size where all three sets were recalled, with an additional half point credit if at least one set at the next highest level was recalled.

Phonological ability. Perceptual or receptive phonological ability was measured in a GATING TASK, wherein participants attempted to identify a word from ever-increasing snippets from its beginning. Each of the five words to be recognized (*television, refrigerator, vegetable, helicopter, thermometer*) was recorded by a native speaker, and then digitized on a computer for presentation. Each word was then broken up into eight 100 ms segments. The first gate consisted of the first 100 ms segment, and each successive gate added an additional 100 ms segment. As each gate was played, participants were asked to guess the word. The words were always administered in the same order, as listed above. The dependent measure of interest for the current study was the overlap between the participant's guess at the initial gate, and the first two phonemes of the target word. This measures perceptual phonological ability in that it captures how well participants could identify phonemes from limited input, and doesn't rely on participants' familiarity with the particular vocabulary items.

Grammatical judgment task. The grammaticality judgment task was the same as the one used in McDonald (2006). This consisted of 100 total sentences – a grammatical and ungrammatical version of fifty different sentences. Ten different structures were tested, including subject-verb-object

TABLE 1. *Example stimuli in the grammaticality judgment task*

Construction	Number of sentence pairs	Ungrammatical example
Word order	4	The teacher the tests graded.
Yes/no questions	4	Drives the teacher a really fancy red car?
Articles	4	The lady drove same car for the past twenty years.
wh-questions	4	What you think about the new coach?
Regular plural	8	There are twenty flute in our marching band.
Present progressive	4	The little girl is play with her dolls.
Regular past tense	8	Last night my friend walk home after dark.
Third person agreement	8	The boy jump whenever he is startled.
Irregular plural	2	Several of the mans decided not to go to the football game.
Irregular past tense	4	Last week the pilot flied to Paris.

word order (4 pairs), yes/no questions (4 pairs), articles (4 pairs), *wh*-questions (4 pairs), regular plurals (8 pairs), present progressive (4 pairs), regular past tense (8 pairs), regular third person singular (8 pairs), irregular plurals (2 pairs) and irregular past tense (4 pairs). Within each sentence type, the sentences were made ungrammatical in the same way. Word order was tested in the word order items by changing the subject-verb-object order to either verb-object-subject or subject-object-verb. Word order and auxiliary omission was tested in yes/no questions by inverting the order of the subject and verb and omitted the auxiliary *do*. Article and auxiliary omissions were tested in article and *wh*-question items by omitting articles and the auxiliary *do*, respectively. Mastery of various regular bound morphology was tested for plurals, present progressive, past tense and third person by omitting the regular morpheme. Items testing irregular morphology regularized the irregular form (e.g. *mans* for *men* or *flied* for *flew*). Example ungrammatical sentences are shown in Table 1. The entire test is available in McDonald (2006).

A further manipulation in the stimuli occurred for regular plurals, regular past tense and third person agreement constructions. For these constructions, half of the grammatical items realized the appropriate bound morpheme with a single phoneme (i.e. /s/ or /z/ for plurals and third person agreement, and /d/ or /t/ for past tense), while the other half used a syllabic realization (i.e. /ɯz/ or /ɯd/ respectively). Thus, we are able to compare performance on items where the construction is marked by a harder to hear realization (e.g. for plurals, *flutes*) to ones where it is marked by an easier to hear realization (e.g. *glasses*). This difference is relevant to the relationship between phonological ability and phonetic substance.

Sentences were recorded by a native speaker and stored in digital form in the computer. Sentences were played in a random order to each participant, with the restriction that the grammatical and ungrammatical version of the same sentence occurred in different halves of the presentation.

Procedure

All tasks were auditorially presented on a Macintosh PowerBook laptop computer equipped with headphones, using Superlab software from Cedrus. Participants were tested at a room at their school (children), or in a testing room at the university (adults).

All participants performed the size judgment task, followed by the gating task and finally the grammaticality judgment task. All instructions were given orally by the experimenter. For the size judgment task they were told after they heard a list of words through the headphones, stars would appear on the computer screen and then they were to say the words in the list back, ordering them in size of the object from smallest to largest. An example with two words in it was given by the experimenter, and the participant asked to do the task using these words. If the participant could not do this correctly, the task was re-explained, until it was understood. Then the participant proceeded on to the computer administered task. After each list, the participant's answer was recorded by the experimenter on an answer sheet, and the experimenter then hit a button to allow the computer to play the next list. The same order of list administration was used for all subjects, with list sizes proceeding from two items up through five items; all participants did all lists, regardless of success on previous lists.

The gating task was presented as a fun guessing game. Participants were told to listen carefully as they would be hearing just the very beginning of a word, and would be asked to guess what word they thought it could be. Then they would hear a longer piece of the same word and be asked to guess again. They were to give an answer each time, even if they weren't sure of what the word could be. The experimenter wrote down each guess after each gate, and then had the computer present the next gate. All participants heard all eight gates for each word. The start of a new word was marked by the appearance of stars on the computer screen.

For the grammaticality judgment task, participants were instructed to listen to each sentence and push the Z key (marked by a smiley face sticker) if it was a correct, grammatical English sentence, and the / key (marked by a sad face sticker) if it was an incorrect, ungrammatical sentence. The computer first presented four practice sentences testing constructions other than those tested in the main part of the task, and the experimenter gave feedback during the practice phase if it was necessary. The presentation then moved seamlessly into the main test. Response to each item was

TABLE 2. *Mean performance (and standard deviations) across age groups on basic measures*

	6–7	8–9½	9½–11	adults
Working memory span	3·80[a]	3·98[a]	4·07[a]	4·50[b]
	(·55)	(·51)	(·43)	(·37)
Phonological ability	·56[a]	·77[b]	·91[b]	·91[b]
	(·29)	(·37)	(·26)	(·23)
A′	·89[a]	·93[b]	·95[b]	·98[c]
	(·08)	(·04)	(·02)	(·03)

NOTE: Different superscripts indicate significant differences between means by Newman–Keuls post-hoc tests.

recorded by the computer. Halfway through the task a break was given to allow the participant time to relax, and to help relieve boredom.

Participants' ages and confirmation of their native speaker status was determined either at the beginning or end of the experiment, or during the break in the grammaticality judgment task. Participants were thanked for their participation. Child participants received either a small monetary reward ($5) or a small toy for their participation; adult participants received extra credit towards their psychology course.

RESULTS

First we compare the age groups on working memory span, phonological ability and overall accuracy on the grammaticality judgment task. All these analyses are done with a one-way ANOVA with age group (6–7, 8–9½, 9½–11, adults) as a between-subject factor (F_1 statistic); for the grammaticality judgment task corresponding items analyses are also done (F_2 statistic), and these two F statistics were combined into the min F' statistic (Clark, 1973). Significant results are followed up with a Newman–Keuls post-hoc test. We then break the grammaticality judgment task down into the ten constructions, and similarly analyze the data for each construction. Then regressions using age, working memory and phonological ability measures to predict performance on the grammaticality judgment task, both as a whole, and more interestingly, on the individual structures, are presented. Finally, we look at correlations in construction difficulty between the children and adults, and the stressed adults of McDonald (2006).

Working memory span. As shown in Table 2, span performance differed across the groups ($F(3, 83) = 7{\cdot}93$, $p < 0{\cdot}001$, partial $\eta^2 = 0{\cdot}22$). Post-hoc tests showed that all of the child groups, which did not differ significantly from each other, had significantly lower spans than the adults.

Phonological ability. Also shown in Table 2, the ability to accurately identify the first two phonemes of the target word on the initial gate differed between groups ($F(3, 83) = 6.65$, $p < 0.001$, partial $\eta^2 = 0.19$). Post-hoc tests showed that the six- to seven-year-olds guessed fewer correct phonemes than all other groups.

Grammaticality judgment accuracy. To correct for bias from guessing, hits to grammatical sentences and false alarms to ungrammatical sentences were combined into the A' measure (Pollack & Norman, 1964; Donaldson, 1992). Perfect discrimination between grammatical and ungrammatical sentences results in a score of 1, whereas as chance performance results in a score of 0·5. A' scores were generally good across all groups, and in order to differentiate performance near the ceiling of 1, an additional arcsine transformation was applied to the A' scores; for ease of interpretation the untransformed scores are given in Table 2. There was a significant increase in A' performance with increasing age group ($F_1(3, 83) = 16.92$, $p < 0.001$, partial $\eta^2 = 0.38$; $F_2(3, 147) = 70.02$, $p < 0.001$; min $F'(3, 124) = 13.63$, $p < 0.001$). Post hoc tests showed that the six- to seven-year-olds had lower A' scores than all the other groups, and the eight- to nine-and-a half-year-olds and the nine-and-a half- to eleven-year-olds had lower performance than the adults. Thus, even by the end of the grade school years, children had not quite achieved adult level performance on the grammaticality judgment task.

Individual structures

In this section we break down the grammaticality judgment task into the ten different constructions tested. A two-way ANOVA on group (between-subjects) and construction (within-subjects) on the arcsine of the A' scores, and a corresponding items analysis found main effects of both group and structure, which were qualified by an interaction between the two, significant in both the subjects and items analysis ($F_1(27, 747) = 3.73$, $p < 0.001$, partial $\eta^2 = 0.12$; $F_2(27, 120) = 1.71$, $p = 0.026$; min $F'(27, 247) = 1.17$, ns). As can be seen in Table 3, the different structures showed different trends across the age groups. Mastered early at the adult level were two structures involving word order changes (word order, yes/no questions), as well as one involving omissions (articles). Even the six- to seven-year-olds did not differ from adults on these three constructions. Next to be mastered were regular and irregular plurals – the six- to seven-year-olds were significantly worse than the other age groups at these structures, but adult levels were reached by the eight- to nine-and-a half-year-olds. Then progressives and *wh*-questions were mastered – both the six- to seven-year-olds and eight- to nine-and-a half-year-olds performed worse than the adults on these structures, with the nine-and-a

TABLE 3. *Mean A′ scores (and standard deviations) by age group on the ten constructions in the grammaticality judgment task*

	6–7	8–9½	9½–11	adult	F_1	F_2	min F'	Partial η^2
Word order	·96[a] (·06)	·97[a] (·06)	·99[a] (·03)	·99[a] (·02)	$F_1(3, 83) = 2{\cdot}17$, ns	$F_2(3, 9) = 2{\cdot}97$, ns	min $F'(3, 42) = 1{\cdot}25$, ns	·07
Yes/no question	·95[a] (·11)	·96[a] (·07)	·99[a] (·03)	·98[a] (·03)	$F_1(3, 83) = 1{\cdot}48$, ns	$F_2(3, 9) = 3{\cdot}93$, $p = .048$	min $F'(3, 68) = 1{\cdot}08$, ns	·05
Articles	·93[a] (·11)	·95[a] (·11)	·97[a] (·04)	·97[a] (·04)	$F_1(3, 83) = .66$, ns	$F_2(3, 9) = 1{\cdot}39$, ns	min $F'(3, 59) = .45$, ns	·02
wh-question	·94[a] (·05)	·96[b] (·05)	·98[bc] (·04)	·99[c] (·02)	$F_1(3, 83) = 7{\cdot}01$, $p < .001$	$F_2(3, 9) = 11{\cdot}66$, $p = .002$	min $F'(3, 49) = 4{\cdot}38$, $p = .008$	·20
Regular plural	·89[a] (·10)	·95[b] (·05)	·95[b] (·05)	·97[b] (·05)	$F_1(3, 83) = 7{\cdot}75$, $p < .001$	$F_2(3, 21) = 16{\cdot}96$, $p < .001$	min $F'(3, 97) = 5{\cdot}32$, $p = .002$	·22
Progressive	·92[a] (·08)	·89[a] (·11)	·95[ab] (·04)	·97[b] (·04)	$F_1(3, 83) = 4{\cdot}96$, $p = .003$	$F_2(3, 9) = 3{\cdot}00$, ns	min $F'(3, 22) = 1{\cdot}87$, ns	·15
Regular past tense	·86[a] (·12)	·91[a] (·08)	·92[a] (·05)	·97[b] (·04)	$F_1(3, 83) = 8{\cdot}10$, $p < .001$	$F_2(3, 21) = 17{\cdot}08$, $p < .001$	min $F'(3, 95) = 5{\cdot}49$, $p = .002$	·23
Third person agreement	·85[a] (·16)	·91[ab] (·07)	·94[b] (·05)	·99[c] (·03)	$F_1(3, 83) = 13{\cdot}86$, $p < .001$	$F_2(3, 21) = 18{\cdot}52$, $p < .001$	min $F'(3, 79) = 7{\cdot}93$, $p < .001$	·33
Irregular plural	·82[a] (·21)	·93[b] (·11)	·99[b] (·03)	·93[b] (·16)	$F_1(3, 83) = 7{\cdot}85$, $p < .001$	$F_2(3, 3) = 7{\cdot}67$, ns	min $F'(3, 11) = 3{\cdot}88$, $p = .041$	·22
Irregular past tense	·75[a] (·19)	·86[b] (·18)	·94[b] (·06)	·98[c] (·07)	$F_1(3, 83) = 14{\cdot}80$, $p < .001$	$F_2(3, 9) = 15{\cdot}47$, $p < .001$	min $F'(3, 34) = 7{\cdot}56$, $p < .001$	·35

NOTES: Different superscripts indicate significant differences between means by Newman–Keuls post-hoc tests.
F_1 statistics treat participants as random effects, so a significant result indicates the effect on the tested items would generalize to other participants. F_2 statistics treat the linguistic items, in this case, sentences, as random effects, to see if the effect for the tested participants would generalize to other, non-tested items. The min F' statistic combines these two and is a conservative test of generalization to both other participants and other items.

TABLE 4. *Results of regression analyses*

	Age			Working memory			Phonological ability		
	β	t	p	β	t	p	β	t	p
Overall A'	·26	2·59	·011	·23	2·21	·030	·32	3·46	·001
Word order	−·05	−.39	ns	·35	2·88	·005	·14	1·32	ns
yes/no questions	·02	·21	ns	·03	·28	ns	·35	3·12	·003
Articles	·05	·43	ns	−.02	−·18	ns	·30	2·62	·011
wh-questions	·26	3·39	·024	·13	1·13	ns	·14	1·28	ns
Regular plurals	·17	1·46	ns	·09	·76	ns	·26	2·41	·018
Progressive	·17	1·46	ns	·32	2·64	·010	−.09	−·83	ns
Regular past	·23	2·13	·036	·23	2·02	·047	·19	1·81	ns
Third person agreement	·22	2·17	·033	·23	2·12	·037	·30	3·09	·003
Irregular plurals	·04	·31	ns	·18	1·45	ns	·13	1·13	ns
Irregular past	·31	2·94	·004	·10	·95	ns	·29	2·86	·005

half- to eleven-year-olds' performance not differing from either younger groups or the adults. Finally, three structures, regular and irregular past tense and third person singular subject–verb agreement, did not show adult mastery even by the oldest child group.

Regression analyses

In this section we examine the contributions of age, working memory capacity and phonological ability to performance on the grammaticality judgment task. In particular, we were interested in whether working memory and/or phonological ability accounted for variance in performance beyond any effects of age. We therefore ran a regression analysis on the overall A' scores on the grammaticality judgment task, with age, working memory and phonological ability as predictors. The results, shown in Table 4, indicate that all three predictors made significant contributions to the regression equation. Thus, working memory capacity and phonological ability did play a role in grammaticality judgment performance above and beyond effects of age.

We did similar regressions to predict the performance on each of the individual construction types in the grammaticality judgment task. These results, also shown in Table 4, show several interesting facts. First, working memory was a significant predictor of grammaticality judgment performance on four different structures (word order, progressives, regular past and third person agreement). Working memory may play a role in the detection of word order errors because one has to have the elements of the sentence and their order in mind to make the judgment. All constructions involving regular verb morphology errors (progressives, regular past and third person agreement) also showed an impact of working memory capacity. These verb morphology structures (especially, as noted in the Introduction, past tense and third person agreement) seem to be high in working memory processing demands.

Second, phonological ability was a significant predictor of grammaticality judgment performance on five different structures (yes/no questions, articles, regular plurals, third person agreement and irregular past tense). In addition, the effect of phonological ability on the regular past tense was in the same direction as for other low salience morphemes (see positive beta in Table 4; $p=0.073$), although it was not statistically significant. Thus, as predicted, most constructions involving the omission of function words and omission of lower phonetic substance morphemes were affected by phonological ability. The relation of phonological ability to irregular past tense, however, was not predicted, as the irregular and over-regularized forms of these verbs used on the test were quite phonologically distinct.

It is possible to do an even more refined test of the relationship of phonological ability to grammaticality judgment for the items involving regular morphology. Recall that the amount of phonetic substance was systematically manipulated in regular plurals, regular past tense and third person singular items. Within these constructions, half of the stimuli contained morphological realizations of a single phoneme, while the other half contained realizations of a syllable (e.g. compare *flutes* vs. *glasses*, *walked* vs. *ended*, *jumps* vs. *watches*). If phonological ability plays a role in performance, we would expect a stronger relationship between phonological ability and performance on the single phoneme versions than on the syllabic ones. This is indeed what we found for two of the constructions – in regressions using age, working memory and phonological ability as predictors, phonological ability was a significant predictor for the phoneme realization for both regular plurals ($\beta=0.27$, $t(83)=2.50$, $p=0.014$) and third person agreement ($\beta=0.38$, $t(83)=3.90$, $p<0.001$); phonological ability did not significantly predict performance on the syllable realization of regular plurals ($\beta=0.17$, $t(83)=1.46$, ns) or third person agreement ($\beta=0.16$, $t(83)=1.48$, ns). Phonological ability did not reach significance for either the phoneme or syllable realization of past tense.

Comparison to adults under stress

We now examine the relative difficulty of the constructions for the children in this study, and the adults in this study and those in McDonald (2006). Recall that the same grammaticality judgment task used here was also administered to groups of adults under a variety of stress conditions in McDonald (2006). In addition to a control group, there were two groups that performed under working memory stress (low memory load

TABLE 5. *Correlations between grammaticality judgment performance in children and adults under stress from McDonald* (2006)

	Current	McDonald (2006)					
	Unstressed adults	Unstressed adults	Low memory	High memory	Noise	Deadline	Compressed speech
6–7	·41	·09	·47	·52	−·30	−·03	·53
8–9½	23	·51	·67*	·64*	−·18	·00	·68*
9½–11	−·04	·55	·87***	·75*	·19	·32	·92***
all children	·30	·34	·69*	·68*	−·18	·06	·74*

* $p < 0{\cdot}05$.
*** $p < 0{\cdot}001$.

(remembering a different four digit number for each sentence judged), and high memory load (remembering a different seven digit number for each sentence)), one that performed under noise stress (sentences were played through white noise), one that performed under deadline stress (answers to be given within a restricting time limit) and one that performed under compressed speech stress (sentences played at twice the speed, with pitch normalization). These various types of stressors impacted the ten constructions in the grammaticality judgment test differently – for example, noise impacted eight of the ten constructions, with omitted articles showing the greatest decrease, while compressed speech affected six of the structures, with past tense most impacted.

To see if the relative difficulty that the children had on the ten constructions was mirrored in the adult data, we computed the correlations between the A' performance of the children on the ten constructions and that of the unstressed adults in the current experiment as well as that of the unstressed adults and five stressed adult groups from McDonald (2006). As shown in Table 5, the ordering of construction difficulty for the children did not significantly correlate with either of the adult unstressed groups, or the noise stress or deadline stress group. Starting with the eight- to nine-and-a-half-year-olds, however, it did significantly correlate with the two memory stress groups. The children's ordering also correlated to the compressed speech group. However, the order of this group was strongly correlated to both the low memory ($r = 0{\cdot}93$) and high memory ($r = 0{\cdot}79$) stress groups, and thus patterns together with them. This suggests that processing under memory load or speeded speech may have a common underlying construct – perhaps a deficit of processing resources to compute accurate judgments in real time. Hayiou-Thomas *et al.* (2004) also found that memory load and compressed speech affect constructions similarly. In any case, children had the same relative difficulty with constructions as adults with processing limitations in terms of memory capacity. This adds credence to the argument that mastery of certain constructions is not wholly a function of age or language exposure. Rather, mastery may also be determined by other basic cognitive factors such as memory or processing capacity.

DISCUSSION

The current study tested typically developing children on their mastery of ten different grammatical constructions in a grammaticality judgment task. Even the youngest children tested here, six- to seven-year-olds, performed at adult levels on constructions involving word order changes (word order, yes/no questions items) and omission of articles. Next mastered were regular and irregular plurals, followed by constructions omitting auxiliaries (*wh*-questions) and the present progressive suffix *-ing*. Constructions not mastered at the adult level even by the oldest children tested, nine-and-a-half- to eleven-year-olds, involved regular and irregular past tense and third person singular agreement.

This ordering of mastery was consistent in general outline with the predictions made based on past literature, although there were also a few differences. As predicted, adult levels of grammaticality judgment was achieved early for word order violations, and late for third person singular agreement. For irregular forms, plurals were mastered before past tense. However, although some past literature suggested that present progressive *-ing* omissions were easier than auxiliary omissions (Rice *et al.*, 1999), we found them to be mastered at about the same age. Prior literature was conflicted about the difficulty of regular plurals in comparison to third person agreement. Wulfeck (1993) found plurals easier than third person agreement, while Wulfeck *et al.* (2004) found them equivalent in difficulty. Our results are consistent with Wulfeck (1993) – both regular and irregular plurals were mastered at a younger age than third person agreement.

In addition to age, the mastery of certain constructions – word order, present progressive, regular past tense and third person agreement – was significantly predicted by working memory capacity. That is, better performance was evident on these constructions by people with larger working memory capacities, above and beyond the effects of age and phonological ability. Recall that adults and children under working memory load also had trouble with regular past tense and third person agreement, and few problems with function word omission, regular plurals or word order errors (Blackwell & Bates, 1995; Hayiou-Thomas *et al.*, 2004; McDonald, 2006). Our results, consistent with this, indicate that working memory span plays a role in grammaticality judgment mastery, but also only for particular constructions. These constructions may be ones that require participants to keep in mind multiple sentence parts, and check if information across these parts is consistent. It may be that integrating information about the verb morphology with other information in the sentence is particularly high in demand on working memory, as all of our structures involving verb morphology correlated with working memory span. It is worth noting that although word order items showed correlations to working memory in the current study, performance was uniformly high – which is consistent with the high performance also shown on these items by memory stressed speakers. Thus, while working memory capacity may be relevant for the judgment of word order items, the relatively strong mastery of this construction may buffer it from showing a large impact of diminished capacity.

Mastery of several (but not all – see *wh*-questions) constructions involving the omission of function words (articles, auxiliaries in yes/no questions) and shorter bound morphemes (regular plurals, third person agreement) was significantly related to phonological ability above and beyond age and working memory contributions. As predicted, phonological ability did not significantly predict performance on high phonetic substance bound morphemes (e.g. present progressive) or for pure word order items. But perhaps the best evidence for phonological ability affecting grammaticality judgment performance came when construction type was held constant, and the effect of single phoneme vs. syllabic realization of a morpheme was tested. For both regular plurals and third person agreement, phonological ability predicted judgments on the single phoneme realizations, and not on the syllabic realization. That is, people with poor phonological ability may have had trouble hearing whether the low phonetic substance realization was present or not. When the easier to hear syllabic realizations were used, the poor phonological ability participants were no longer at a disadvantage.

When relative construction difficulty for the children was compared to that of unstressed and stressed adults, it was found that children resembled adults under memory load. Recall that all child groups in the current study had memory span scores significantly less than the adults. It appears, then, that this smaller working memory capacity did impact performance in the same way as limiting working memory capacity does in adults. Interestingly, children's order of construction difficulty did not correspond to that of adults under noise stress. This is interesting for two reasons. First, it indicates that the phonological difficulties that children had – recall that the six- to seven-year-olds did have poorer phonological ability scores than adults – were different either in nature or magnitude from a simple overlay of noise. Second, McDonald (2006) found late (at or after the age of 12) second-language learners of English showed an order of relative construction difficulty quite similar to that of adult native speakers under noise. Thus, any phonological difficulties experienced by native speaking children appear to be quite different in nature from those affecting late second-language learners.

The relationships of working memory and phonological ability to performance on certain constructions and the correlation in construction difficulty between children and adults under memory load shows that factors other than grammatical knowledge are important in grammaticality judgment, and probably other language tasks as well. Thus, poor performance on a particular construction is not necessarily due to poor knowledge about that structure's grammar. Rather, low working memory span, poor phonological ability or other factors could be responsible for the low performance. If so, children may be able to evidence mastery of a structure in tasks that decrease memory load, or that increase the clarity of phonological input.

Indeed, one could view the grammaticality judgment task as a high demand task in and of itself. Participants not only need to comprehend the orally presented sentence, they must also make a metalinguistic judgment. This may be why we do not see mastery of certain structures at the adult level in this task even by the 9½- to 11-year-old children. One could vary the demands within the grammaticality judgment task itself by varying whether or not repetitions are allowed, or whether time pressure to answer is present to see how that impacts performance, or one would compare performance in the grammaticality judgment task to a less demanding comprehension task to see at which ages adult mastery of the various structures is reached in this type of task.

There were two structures that we tested that seemed to cause particular trouble for the children: third person agreement and regular past tense were acquired at adult levels late – not even the oldest group had reached adult levels. Interestingly, these two structures also showed significant effects in the regression analysis of multiple factors on performance. It is interesting to note that in spontaneous speech samples of preschool children, these structures were also late in mastery level (Brown, 1973). Also, when atypically developing populations are compared to typically developing populations, these two structures are even more delayed. For example, children with SLI show both delayed productive mastery of these items, and delayed mastery in a grammaticality judgment task (Montgomery & Leonard, 1998; Rice *et al.*, 1999; Eadie, Fey, Douglas & Parsons 2002; Laws & Bishop, 2003; Wulfeck *et al.*, 2004). Children with Down Syndrome also showed delayed production of these morphemes relative to language matched children (Eadie *et al.*, 2002; Laws & Bishop, 2003), as do children learning English as their second language (Paradis, 2005). The fact that such diverse populations show similarities in their language delay is intriguing, and raises the possibility that some common factor could underlie these performance patterns. Indeed, literature suggests that these populations may all have impaired working memory (or in the case of second-language learning children, lower working memory in their second than first language), and various types of perceptual phonological or hearing problems relative to typically developing or native speaking peers (Kay-Raining Bird & Chapman, 1994; Montgomery & Leonard, 1998; Ellis Weismer *et al.*, 1999; Montgomery, 2000*a*, 2000*b*; Laws & Bishop, 2003; Marton & Schwartz, 2003; Gutiérrez-Clellen, Calderón & Ellis Weismer, 2004; Nelson, Kohnert, Sabur & Shaw, 2005).

There are more structural based accounts for the problems at least one of these populations has with third person agreement and past tense. The Extended Optional Infinitive account (Rice, Wexler & Cleave, 1995; Rice & Wexler, 1996), an extension of the Optional Infinitive account of typically developing children (Wexler, 1994), claims that children with SLI go through a prolonged period of only sometimes producing agreement and tense morphology in obligatory contexts. This account accurately predicts that children with SLI will master the *-s* of regular plural morphology before the *-s* of third person agreement. However, the current results with typically developing children also easily account for this difference. While phonological ability predicted performance on both plurals and agreement sentences (and recall, more specifically for those sentences with the /s/ or /z/ realizations, rather than the /ɷz/ realization), working memory capacity did not predict performance on plurals, but did on agreement sentences. Thus, for any population with restricted working memory capacity (e.g. typically developing children as compared to adults, children with SLI as compared to typically developing children, adults under memory load as compared to unstressed adults), agreement sentences should be more impaired than plurals, simply due to the former's greater demand on working memory.

Child language researchers interested in the order of acquisition of various grammatical devices have posited several factors that may underlie this ordering. For example, Brown (1973) posited that the order of English morpheme acquisition in production is determined by the semantic and syntactic complexity of the morphemes. More recently, Goldschneider & DeKeyser (2005) conducted a meta-analysis of studies on the order of morpheme acquisition in second-language learners of English. They found they could account for 71 percent of the variance in accuracy on morpheme production by a combination of five factors, including frequency, phonological salience, semantic complexity, morphophonological regularity and syntactic category. Both these studies sought to explain order of acquisitions in terms of the characteristics of the morphemes. In contrast, the current study looks to characteristics of the learners in explaining the mastery of various grammatical devices in a grammaticality judgment task. These methods complement each other, and factors show obvious correspondences. For example, complex morphemes should tax working memory more, and low salience morphemes will be problematic for those with poor phonological ability.

Finally, the current research adds to a recent trend in the language literature in using stressed groups of normal speakers as models for other populations. For example, Dick *et al.* (2001) succeeded in modeling the behavior of adult aphasics on a language comprehension task by having normal adults listen to the same stimuli with low pass filtering and speech compression. McDonald (2006) was able to model the grammaticality judgment pattern of late second-language learners with native speakers judging the same stimuli under noise stress. Hayiou-Thomas *et al.* (2004) modeled several characteristics of children with SLI by having typically developing children perform grammaticality judgments under the stress of increased sentence length; similar results were found when typically developing children had the stress of compressed speech stimuli. The current study shows that grade school aged children can be successfully modeled by adults under memory stress. Thus, developmental changes as well as impairments can be modeled by groups under stress. This approach, that is modeling impaired or developing populations by putting end state learners under stress, comes from a tradition of examining how deficits in various domain general systems (such as memory) can have a specific impact on language and point to these domain general processes as possible underlying causes for the linguistic behavior seen in both impaired and developing populations.

REFERENCES

Blackwell, A. & Bates, E. (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience* **7**, 228–57.

Bowey, J. A. (2005). Grammatical sensitivity: Its origins and potential contribution to early word reading skill. *Journal of Experimental Child Psychology* **90**, 318–43.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior* **12**, 335–59.

Cox, M. V. (1989). Children's over-regularization of nouns and verbs. *Journal of Child Language* **16**, 203–6.

deVilliers, P. A. & deVilliers, J. (1972). Early judgments of semantic and syntactic acceptability by children. *Journal of Psycholinguistic Research* **1**, 299–310.

Dick, F., Bates, E. & Ferstl, E. C. (2003). Spectral and temporal degradation of speech as a simulation of morphosyntactic deficits in English and German. *Brain & Language* **85**, 535–42.

Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N. & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review* **108**, 759–88.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General* **121**, 275–7.

Eadie, P. A., Fey, M. E., Douglas, J. M. & Parsons, C. L. (2002). Profiles of grammatical morphology and sentence imitation in children with specific language impairment and Down syndrome. *Journal of Speech, Language, and Hearing Research* **45**, 720–32.

Ellis Weismer, S., Evans, J. & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research* **42**, 1249–60.

Gaulin, C. A. & Campbell, T. F. (1994). Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Perceptual and Motor Skills* **79**, 55–64.

Goldschneider, J. M. & DeKeyser, R. M. (2005). Explaining the 'natural order of L2 morpheme acquisition' in English: A meta-analysis of multiple determinants. *Language Learning* **55**, 27–77.

Gottardo, A., Stanovich, K. E. & Siegel, L. S. (1996). The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology* **63**, 563–82.

Gutiérrez-Clellen, V. F., Calderón, J. & Ellis Weismer, S. (2004). Verbal working memory in bilingual children. *Journal of Speech, Language and Hearing Research* **47**, 863–76.

Hakes, D. T. (1980). *The development of metalinguistic abilities in children.* Berlin: Springer Verlag.

Hartsuiker, R. J. & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes* **21**, 181–204.

Hayiou-Thomas, M. E., Bishop, D. V. M. & Plunkett, K. (2004). Simulating SLI: General cognitive processing stressors can produce a specific linguistic profile. *Journal of Speech, Language, and Hearing Research* **47**, 1347–62.

Kail, M. (2004). On-line grammaticality judgments in French children and adults: A crosslinguistic perspective. *Journal of Child Language* **31**, 713–37.

Kay-Raining Bird, E. & Chapman, R. S. (1994). Sequential recall in individuals with Down Syndrome. *Journal of Speech and Hearing Research* **37**, 1369–80.

Kilborn, K. (1991). Selective impairment of grammatical morphology due to induced stress in normal listeners: Implications for aphasia. *Brain & Language* **41**, 275–88.

King, J. & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory & Language* **30**, 580–602.

Laws, G. & Bishop, D. V. M. (2003). The comparison of language abilities in adolescents with Down syndrome and children with specific language impairment. *Journal of Speech, Language, and Hearing Research* **46**, 1324–39.

Marton, K. & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research* **46**, 1138–53.

McDonald, J. L. (2006). Alternatives to the Critical Period Hypothesis: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory & Language* **55**, 381–401.

Montgomery, J. W. (2000*a*). Relation of working memory to off-line and real time sentence processing in children with specific language impairment. *Applied Psycholinguistics* **21**, 117–48.

Montgomery, J. W. (2000*b*). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research* **43**, 293–308.

Montgomery, J. W. & Leonard, L. B. (1998). Real time inflectional processing by children with specific language impairment: Effects of phonetic substance. *Journal of Speech, Language, and Hearing Research* **41**, 1432–43.

Nelson, P., Kohnert, K., Sabur, S. & Shaw, D. (2005). Classroom noise and children learning through a second language: Double jeopardy? *Language, Speech, and Hearing Services in Schools* **36**, 219–29.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech and Hearing Services in Schools* **36**, 172–87.

Pollack, I. & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science* **1**, 125–6.

Redmond, S. M. & Rice, M. L. (2001). Detection of irregular verb violations by children with and without SLI. *Journal of Speech, Language, and Hearing Research* **44**, 655–69.

Rice, M. L. & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English speaking children. *Journal of Speech and Hearing Research* **39**, 1239–57.

Rice, M. L., Wexler, K. & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research* **38**, 850–63.

Rice, M., Wexler, K. & Redmond, S. M. (1999). Grammaticality judgments of an extended optional infinitive grammar: evidence from English speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research* **42**, 943–61.

Sutter, J. C. & Johnson, C. J. (1990). School-age children's metalinguistic awareness of grammaticality in verb form. *Journal of Speech and Hearing Research* **33**, 84–95.

Vos, S. H., Gunter, T. C., Schriefers, H. & Friederici, A. D. (2001). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes* **16**, 65–103.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivation in child grammar. In D. Lightfoot & N. Hornstein (eds), *Verb movement*, 305–50. Cambridge: Cambridge University Press.

Wulfeck, B. (1993). A reaction time study of grammaticality judgments in children. *Journal of Speech and Hearing Research* **36**, 1208–15.

Wulfeck, B., Bates, E., Krupa-Kwiatkowski, M. & Saltzman, D. (2004). Grammaticality sensitivity in children with early focal brain injury and children with specific language impairment. *Brain & Language* **88**, 215–28.