Developmental Science 9:1 (2006), pp 21–32

REPORT

Why girls say 'holded' more than boys

Joshua K. Hartshorne and Michael T. Ullman

Brain and Language Laboratory, Departments of Neuroscience, Psychology, Linguistics and Neurology, Georgetown University, USA

Abstract

Women are better than men at verbal memory tasks, such as remembering word lists. These tasks depend on declarative memory. The declarative/procedural model of language, which posits that the lexicon of stored words is part of declarative memory, while grammatical composition of complex forms depends on procedural memory, predicts a female superiority in aspects of lexical memory. Other neurocognitive models of language have not made this prediction. Here we examine the prediction in past-tense over-regularizations (e.g. holded*) produced by children. We expected that girls would remember irregular past-tense forms (*held*) better than boys, and thus would over-regularize less. To our surprise, girls over-regularized far more than boys. We investigated potential explanations for this sex difference. Analyses showed that in girls but not boys, over-regularization rates correlated with measures of the number of similar-sounding regulars (*folded, molded*). This sex difference in phonological neighborhood effects is taken to suggest that girls tend to produce over-regularizations in associative lexical memory, generalizing over stored neighboring regulars, while boys are more likely to depend upon rule-governed affixation (*hold + -ed*). The finding is consistent with the hypothesis that, likely due to their superior lexical abilities, females tend to retrieve from memory complex forms (*walked*) that men generally compose with the grammatical system (*walk + -ed*). The results suggest that sex may be an important factor in the acquisition and computation of language.*

Introduction

Males and females clearly differ in various perceptual, motor and cognitive domains, both in humans and in other animals (Halpern, 2000; Kimura, 1999). In language, however, sex differences have not been so evident (Halpern, 2000). Such differences might not have been reliably demonstrated for at least two reasons. First, studies have generally not focused on the sub-components of language, and thus could have missed differences limited to one of them. Second, most studies have been exploratory rather than hypothesis-driven, and so might not have probed for specific sex differences predicted by particular theoretical perspectives.

Along with our colleagues, we have been examining sex differences in the neurocognition of language, as predicted by a particular 'dual-system' theoretical perspective. ('Single-system' models of language have not, to our knowledge, made any predictions regarding sex differences; these models are examined in the General Discussion.) According to dual-system models, such as the Words-and-Rules (WR) theory proposed by Pinker and colleagues, language depends on two mental capacities with distinct neurocognitive underpinnings (Clahsen, 1999; Pinker, 1999; Pinker & Ullman, 2002). Idiosyncratic linguistic information (e.g. *hold* takes the irregular past-tense *held*) is memorized in the mental lexicon, whereas the mental grammar underlies the rule-governed combination of lexical forms into complex linguistic representations, such as regular past-tense forms (*walk* + *-ed*), phrases and sentences. Importantly, complex forms that can be put together with the mental grammar can in principle *also* be stored in the lexicon (*walked*). In fact a range of evidence suggests that certain types of regularly inflected forms, such higher frequency regulars, do indeed tend to be memorized (Alegre & Gordon, 1999; Pinker, 1999; Pinker & Ullman, 2002; Ullman, 2001a).

One dual-system theory – the Declarative/Procedural (DP) model – posits that the mental lexicon depends on declarative memory, which subserves the learning and use of knowledge about facts and events, whereas aspects of the mental grammar depend on procedural memory, which subserves the acquisition and expression of motor and cognitive skills (Ullman, 2001b, 2004; Ullman, Corkin, Coppola, Hickok, Growdon, Koroshetz & Pinker, 1997). Thus our knowledge of the two memory

Address for correspondence: Michael Ullman, Brain and Language Laboratory, Department of Neuroscience, Georgetown University, New Research Building, 3970 Reservoir Road, NW, Washington, DC 20057, USA; e-mail: michael@georgetown.edu

systems should lead to specific predictions about lexicon and grammar.

Women and girls are better than men and boys at verbal memory tasks, such as tests of verbal episodic memory (Halpern, 2000; Kimura, 1999; Kramer, Delis, Kaplan, O'Donnell & Prifitera, 1997). These tasks and this sex difference appear to depend on declarative memory (Golomb, Kluger, de Leon, Ferris, Mittelman, Cohen & George, 1996; Pfluger, Weil, Weis, Vollmar, Heiss, Egger, Scheck & Hahn, 1999; Squire & Knowlton, 2000). We therefore posited a female advantage at aspects of lexical memory, possibly including lexical consolidation, access and/or retrieval abilities (Ullman, 2004; Ullman, Estabrooke, Steinhauer, Brovetto, Pancheva, Ozawa, Mordecai & Maki, 2002). Such a sex difference may explain findings that girls learn words faster than boys, and that women are better than men at remembering words and at verbal fluency tasks (Halpern, 2000; Kimura, 1999). We have also hypothesized that, due to their memory advantage, females should tend to remember previously encountered complex forms (*walked*) that men generally compose (*walk* + *-ed*) (Ullman, 2004; Ullman *et al.*, 2002). That is, just as there appear to be differences between different types of *items*, with certain regular forms being more likely to be stored than others, there may also be differences between different types of *subjects*, with certain groups of individuals being more likely than others to remember regulars (and possibly other complex forms as well).

The predicted sex differences can be tested not only in adult processing (Steinhauer & Ullman, 2002; Ullman *et al.*, 2002; Ullman, Hartshorne, Estabrooke, Brovetto & Walenski, under revision; Ullman, Walenski, Prado, Ozawa & Steinhauer, under revision), but also in language acquisition and development. One approach is to examine over-regularization – the application of a regular pattern to the stem of an irregular word (e.g. *hold-holded*). WR, DP and other dual-system perspectives posit that over-regularizations occur when a speaker fails to retrieve the correct irregular form from lexical memory, leading to the rule-governed affixation of the stem. (Single-system models give a different account, which is discussed below.) Such errors are particularly prevalent in children, at least in part because children have not yet been exposed to a large number of irregular forms (Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992). We predicted that girls' advantage in lexical memory would give them an advantage at retrieving irregular forms, thus leading to lower over-regularization rates in girls than boys. We tested this prediction in young girls and boys, who are especially prone to over-regularization.

Methods

In order to examine these novel hypotheses and predictions, we turned to a well-studied data set. We re-analyzed data presented by Marcus, Pinker, Ullman, Hollander, Rosen and Xu (1992). Marcus *et al.* (MEA) reported over-regularization rates for each of 25 children from the CHILDES database of transcripts (MacWhinney, 2000). We analyzed these data separately for the 10 girls and 15 boys. The girls and boys (see Table 1) did not differ statistically in their ages at first or last recording; in the number of sessions recorded for each child; in the total number of utterances (defined in MacWhinney, 2000) produced by each child or by other individuals in each child's recordings; in the number of different irregular verbs (type frequency) that the target child used as a past-marked form (i.e. either as a correct past-tense form or as an over-regularization); or in the token frequency of irregular past-tenses they produced (counted either as correct + over-regularizations, or as correct only). All children spoke standard American English as their native language, and all but one came from middle-class families. Ten of the subjects were black (two girls, eight boys).

Our analyses used the same over-regularization rates and exclusionary criteria as those reported by MEA. These are summarized here; for additional details, see MEA. Over-regularization rates for each child were calculated as follows: (number of over-regularization tokens) / [(number of over-regularization tokens) + (number of correct irregular past tokens)]. For more information on the categorization and tabulation of over-regularizations and correct irregular forms, see MEA. Like MEA, we treated one subject (Abe) as an outlier in analyses comparing over-regularization rates between subjects (his over-regularization rate was 3.8 SDs above the mean, likely due to the testing method; see MEA for details), but not in analyses examining the relative over-regularization rates of verbs with respect to their properties (where he showed the same pattern as other subjects; see below and MEA). All *p*-values in this paper are reported as two-tailed.

Results and discussion

Over-regularization rates in girls and boys

Individual and mean over-regularization rates for the 25 children are displayed in Table 1 (rightmost column). Contrary to our predictions, the girls over-regularized at more than three times the rate of boys (means of 5.7% vs. 1.8%), a difference that was statistically significant

Table 1 *Female and male subjects and their over-regularization rates*

Sex	Subject	Source	Age at 1st recording session[a]	Age at last recording session[a]	Number of recording sessions	Recording frequency	Number of child's recorded utterances	Number of others' recorded utterances	Child irregular past-tense types[b]	Child irregular past-tense tokens[b]	Child correct irregular past-tense tokens	Over-regularization rate
f	Allison	Bloom (1973)	1.4	2.8	6	occasionally	2529	2477	14	33	31	6.1%
f	April	Higginson (1985)	1.8	2.9	6	occasionally	2457	3299	17	54	47	13.0%
f	Eve	Brown (1973)	1.5	2.3	20	2–3/month	11624	14272	33	307	283	7.8%
f	GAT	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	3394	5258	37	169	159	5.9%
f	JUB	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	1823	5661	27	140	132	5.7%
f	MIM	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	2473	6355	14	77	77	0.0%
f	Naomi	Sachs (1983)	1.3	4.8	93	weekly to monthly	17253	11974	39	414	378	8.7%
f	Sarah	Brown (1973)	2.3	5.1	139	weekly	37634	45504	65	1782	1717	3.6%
f	TRH	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	2357	5601	17	50	47	6.0%
f	ZOR	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	1756	2900	22	98	98	0.0%
Female mean (SD)			3.1 (1.5)	4.3 (1.1)	27.4 (48.2)	NA	8330 (11534)	10330 (12942)	28.5 (15.9)	312 (531)	297 (511)	5.7 (3.9)
m	Abe	Kuczaj (1976)	2.5	5.0	210	weekly	22443	22069	70	2350	1786	24.0%[c]
m	Adam	Brown (1973)	2.3	5.2	55	2–3/month	46716	26661	55	2492	2444	1.9%
m	ANC	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	1776	2430	19	81	79	2.5%
m	BOM	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	2832	3158	23	113	112	0.9%
m	BRD	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	2575	5373	17	130	128	1.5%
m	CHJ	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	3105	6066	22	155	151	2.6%
m	DED	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	1667	3750	22	111	106	4.5%
m	JOB	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	1786	4185	26	130	130	0.0%
m	KIF	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	2855	5312	15	100	100	0.0%
m	MAA	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	2450	4462	23	107	105	1.9%
m	Nat	Bohannon & Marquis (1977)	2.7	3.7	21	within 1 month	2494	3697	14	52	52	0.0%
m	Nathaniel	Snow (unpublished)[d]	2.3	3.8	30	weekly	13518	20985	29	257	243	5.4%
m	Peter	Bloom (1973)	1.3	3.1	20	monthly	29497	31167	34	874	853	2.4%
m	TOS	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	2216	3954	23	84	84	0.0%
m	VOH	Hall *et al.* (1984)	4.5	5.0	2	twice in 2 days	1602	2882	15	65	64	1.5%
Male mean (SD)			3.7 (1.1)	4.7 (0.8)	23.7 (53.8)	NA	9169 (13467)	9743 (9944)	27.1 (15.6)	473 (816)	429 (723)	1.8 (1.7)
t-test: males vs. females			$t(23) = 1.24$, $p = .23$	$t(23) = 0.95$, $p = .35$	$t(23) = 0.17$, $p = .86$	NA	$t(23) = 0.16$, $p = .87$	$t(23) = 0.13$, $p = .90$	$t(23) = 0.21$, $p = .83$	$t(23) = 0.55$, $p = .59$	$t(23) = 0.50$, $p = .62$	$t(11.4) = 2.97$[e], $p = .01$

[a] Ages are reported in decimal numbers, not in years and months (e.g. an age of 4 years and 6 months is reported here as 4.5 years).
[b] Correct irregular past-tense forms plus over-regularizations. See Marcus *et al.* (1992).
[c] Excluded from mean and analyses (see text).
[d] See MacWhinney and Snow (1985).
[e] Satterthwaite's method used, due to unequal variances.
Note: The 15 children reported by Hall *et al.* (1984) were each tested over a 2-day period. Their ages at testing ranged from 4.5 to 5.0 years. Because the exact ages of the individual children were not available, here we report the full age range (4.5–5.0) of these children for each child.

(Table 1). Although this result was unexpected, it is robust. Even though there were fewer girls than boys (10 to 14; as elsewhere, with Abe excluded) and fewer past-marked tokens produced by girls (3124 to 4751), the girls produced substantially more over-regularizations (155 to 100). The top seven over-regularizers were girls, whereas four of the six subjects who did not over-regularize at all were boys. The pattern was not limited to particular sources of the data. Indeed, for all three sources with both boys and girls (Bloom, 1973; Brown, 1973; Hall, Nagy & Linn, 1984), girls over-regularized more than boys (Table 1). The pattern was also not due to subjects with few verb tokens, whose low signal could lead to spurious effects caused by noise: the five girls and five boys with the largest total number of verb tokens also showed the over-regularization difference (girls: mean = 6.4%, SD = 2.0%; boys: mean = 2.8%, SD = 1.5%; $t(8) = 3.19$, $p = .01$).

As noted above, the boys and girls did not differ in their ages at first or last recording, in their number of recording sessions, in the number of utterances produced either by themselves or by others in the sessions, or in the number of irregular past-tense types or tokens that they produced; this remained true with Abe excluded (for these variables, in the order presented in Table 1, the p values for the t-test comparisons between boys and girls, with Abe omitted, were .18, .26, .23, .98, .75, .49, .42, .92 and .89). Similarly, there were no differences in social class or linguistic background between the sexes (see Methods). The effect was also not influenced by race: in the 2 (girl/boy) by 2 (black/non-black) ANOVA on over-regularization rates, there was no main effect of race ($F(1, 1) = 1.32$, $p = .26$) and no interaction between sex and race ($F(1, 1) = 1.91$, $p = .18$). Crucially, the main effect of sex remained significant ($F(1, 1) = 4.79$, $p = .04$).

We also examined whether there were any differences between the boys and the girls in the use of irregular past-tense forms by *adults* in the transcripts. If adults used more correct irregular past-tense forms in the boys' recording sessions than in the girls' sessions, then the boys' access to these past-tense forms should be facilitated as compared to the girls' – not only because these irregular forms could be primed more in the case of the boys, but also because such a sex difference in the adult speech in the transcripts might reflect a more general pattern in adult speech to boys and girls, which would correspond to a higher adult irregular past-tense frequency to boys than to girls. Whether from priming and/or frequency differences, the boys' facilitated access to irregular past-tense forms would lead to lower over-regularization rates as compared to girls, as was in fact observed. However, the adults in the boys' transcripts actually used non-significantly *fewer* irregular past-tense tokens than the adults in the girls' transcripts (in each child's recordings, the adult frequencies were calculated over those irregulars that were used in a past-marked form by that target child), both with Abe included (boys: mean = 555, SD = 869; girls: mean = 781, SD = 1427; boys vs. girls: $t(23) = .49$, $p = .63$) and without Abe (boys: mean = 433, SD = 756; girls: same as just above; boys vs. girls: $t(22) = .70$, $p = .49$). This pattern was found in spite of the fact that the number of recording sessions, the number of utterances in these sessions, and the type frequency of past-marked irregular verbs that the children used in the recordings (over which the adult frequencies were calculated) were not greater for the girls than the boys (see Table 1; similarly, without Abe, ps between .23 and .98). Thus, the pattern cannot be explained by larger sample sizes among the girls than the boys, which could inflate the irregular past-tense token frequencies in the girls' transcripts, thereby obscuring an actual higher rate of use in the boys' transcripts. Thus adults did not use more irregular past-tense forms in the boys' recording sessions than in the girls' recording sessions.

Finally, it might be the case that girls only over-regularize more than boys at certain ages, rather than during the entire period of over-regularization. Such a pattern could occur if the over-regularization curves for girls and boys had similar shapes but were offset with respect to each other (e.g. girls over-regularize earlier due to general linguistic precociousness). If the (age-matched) girls and boys were by chance tested at an age when girls over-regularize more than boys, the pattern presented above could be obtained. However, the girls over-regularized more than the boys across a wide age range, covering the entire period during which over-regularizations are generally examined (2 to 5 years old; see MEA). To directly test whether girls over-regularize more than boys at both young and old ages, we split both sets of subjects (as above, excluding Abe) into young (girls: $n = 5$, mean age = 2.62, SD = .74; boys: $n = 4$, mean age = 2.93, SD = .65; $t(7) = .64$, $p = .54$) and old (5 girls, 10 boys, all 4.8 years) age groups, based on the average of the ages at first and last recording for each subject. Only one age-based split was possible, since the 15 children tested by Hall *et al.* (1984) all had a mean of 4.8 years. The 2 (girl/boy) by 2 (young/old) ANOVA on over-regularization rates (young girls: mean = 7.8, SD = 3.5; old girls: mean = 3.5, SD = 3.2; young boys: mean = 2.4, SD = 2.2; old boys: mean = 1.5, SD = 1.4) did not yield an interaction between sex and age ($F(1, 20) = 2.52$, $p = .13$), but did elicit significant main effects of both age ($F(1, 20) = 5.80$, $p < .05$) and sex ($F(1, 20) = 11.74$, $p < .005$). So, over both age groups the girls over-regularized more than the boys, and moreover to a similar degree. Note also that the four youngest girls and three of the five oldest girls all had higher over-regularization rates than *any* of the boys (excluding Abe). These data suggest that it is *not* the case that girls over-regularize more than boys only at younger or older ages, but rather that the sex difference holds over a broad time-span.

In sum, the data show that girls over-regularize reliably more than boys, and that this difference is not explained by a variety of potential confounding factors. So what does explain this sex difference?

Associative over-regularization: an apparent explanation for the sex difference

We hypothesized that the observed sex difference could be explained by a tendency for girls to produce over-regularizations associatively, rather than by rule-governed composition. Although we had not originally considered this hypothesis, such a tendency follows from the WR/DP claim that stored past-tense forms are represented in an associative (superpositional) memory that allows for at least some degree of productivity (Pinker, 1999; Prasada & Pinker, 1993; Ullman, 2001a). Previously, WR/DP has proposed this for irregulars, allowing for memory-based generalizations to new forms like *bring*-*brung* and *spling*-*splung* (cf. *fling*-*flung*). However, *regulars* that are stored could also yield such generalizations. Thus girls, who are predicted to be more likely to remember regulars than boys (see Introduction), should also be more likely than boys to generalize these regulars in associative memory to new forms, on analogy to similar-sounding regulars (e.g. *hold*-*holded*, cf. *fold*-*folded, mold*-*molded*). The increased likelihood of over-regularization due to the influence of these stored similar-sounding regulars could lead to higher over-regularization rates in girls than boys (see General discussion).

Such associative over-regularizations can be empirically distinguished from over-regularizations that are rule-products. In the former case, over-regularizations that are phonologically similar to many regulars should be over-regularized more than those surrounded by fewer regulars. This positive relation should not hold if over-regularizations are rule-products (see MEA for further discussion). MEA found no significant correlations between over-regularization rates and measures of the number of similar-sounding regulars, suggesting the rule-based computation of over-regularizations. We hypothesized that the absence of such a positive relation was attributable to the fact that many of the subjects examined – indeed, the majority – were boys: they are less likely to show the positive relation, and could obscure a positive relation among the girls.

To test this hypothesis, we re-visited the MEA correlations between over-regularization rates and measures of the number of similar-sounding ('neighboring') regular past-tense forms – i.e. measures of 'regular neighborhood strength'. We performed analyses that were almost identical to those reported in the original study; however, we examined regular neighborhood effects separately for boys and girls, and compared the effects between the sexes.

MEA (p. 126) performed correlations, across irregular verbs, between the verbs' over-regularization rates and various measures of their regular neighborhood strength. Separate correlations were carried out for each of the 19 (out of the 25) children who over-regularized at least once. These individual-subject r-values were transformed to Fisher's Z values, which were then used as the dependent measure in t-tests against zero to determine whether the correlation coefficients were reliably positive.

The regular neighborhood strength for a given irregular verb was calculated in MEA as the summed past-tense frequencies (i.e. token frequencies) of all regular verbs whose past-tense forms are phonologically similar to (that is, phonological neighbors of) the over-regularization of that irregular verb. Each irregular verb's regular neighborhood strength was calculated separately for each child, using the regular past-tense frequencies from the adult speech in that child's transcripts.

Neighborhood effects are not well understood, and there is currently no consensus as to which neighborhood measures should yield the strongest effects. Thus MEA performed analyses separately for three different measures of regular neighborhood strength. In the *rhyme* measure, the neighbors of a given irregular were restricted to those regulars whose stems rhyme with the stem of the irregular. For example, regulars such as *linked* and *blinked* are rhyme neighbors of the over-regularization *sinked*. For the *final-coda* measure, neighbors were defined as all regulars whose stems share their final consonant cluster with the irregular stem. Thus not only *linked* and *blinked*, but also *honked*, *yanked* and *flunked* are final-coda neighbors of *sinked*. In the *final-consonant* measure, the neighborhood encompassed all regular stems ending with the same consonant as the irregular stem. In this case, *leaked*, *barked* and *whisked* are *also* neighbors of *sinked*. For the final-coda and final-consonant measures, the word-final coda or consonant was defined as 'null' for verbs with vowel-final stems.

These three regular neighborhood strength measures were calculated twice in MEA: once over only monosyllabic neighboring regular verbs, since all non-prefixed irregulars are monosyllabic, and children may be sensitive to this contingency; and again over polysyllabic as well as monosyllabic verbs (e.g. rhyme neighbors of *sinked* would include not only *blinked* but also *hoodwinked*), to test the generality and robustness of the correlations of interest.

Because MEA showed that the children's over-regularization rates strongly correlated with adult irregular past-tense frequencies (p. 117), this latter variable must be held constant while examining the relation between over-regularization rates and regular neighborhood strength. Therefore we focus here on correlations between over-regularization rates and neighborhood measures while partialling out the natural log-transformed adult irregular past-tense frequencies of each verb – again, using the adult frequencies from each child's transcript, as in MEA (p. 127). In order to avoid the log of zero, we assigned the value 0 to frequency counts of zero, as was done in MEA (p. 118). Additionally, we also partialed out a dummy variable that was coded 1 for adult irregular past-tense counts of zero and 0 otherwise (Muthen & Muthen, 1998–2004); this helps to model the skewness caused by many zero counts.

Thus we performed six partial correlations for each subject: correlations between over-regularization rates and each of the three neighborhood measures (rhyme, final-coda, final-consonant), with these measures calculated once over only monosyllabic neighboring regulars, and once over all neighboring regulars. Of the 19 over-regularizing children, eight were girls and 11 were boys. As in MEA, Abe's *r*-values were not outliers, so he was not excluded from analyses; excluding him yielded an almost identical pattern of results.

All correlations yielded the same basic pattern of results (Table 2). For all six correlations, the mean *r*-values were larger for the girls than the boys. This sex difference reached statistical significance for five of the six correlations (*p*s $\leq .05$), and showed a trend in the other correlation (rhyme measure calculated over monosyllabic verbs; $p = .13$). The girls' mean *r*-values were positive for all six correlations (Table 2). That is, for all six regular neighborhood measures, the girls over-regularized verbs with larger regular neighborhoods more often than those with smaller neighborhoods. The girls' *r*-values were significantly more positive than zero for the four analyses based on final-coda and final-consonant (*p*s $< .05$). For the rhyme measure, this difference showed a trend ($p = .11$, $p = .13$). In contrast, for none of the six correlations were the boys' mean *r*-values even positive. Thus there was a consistent pattern of sex differences, with the girls but not the boys showing positive correlations.

The sex difference in the correlations does not appear to be explained by several potential confounds. First, because there were more boys than girls (11 vs. 8), the lack of positive correlations among the boys cannot be attributed to an insufficient number of male subjects. Second, like the larger sample of 15 boys and 10 girls, these 11 boys and eight girls who over-regularized at least once did not differ statistically on any of the nine variables presented in Table 1 and discussed above (for these nine variables, in the order presented in Table 1, the *p* values were .17, .17, .89, .79, 1.00, .52, .88, .52 and .56).

Third, it might be argued that the boys' low over-regularization rates led to floor effects, which could weaken the correlations between their over-regularization rates and the various neighborhood measures, leading to smaller *r*-values. However, there was no positive relation between the boys' *r*-values and their over-regularization rates (*p*s between .69 and .96, for the six correlations between the boys' over-regularizations rates and each of the six sets of *r*-values). That is, it was not the case that those boys with lower over-regularization rates had smaller *r*-values. In fact, Abe, whose over-regularization rate was by far the highest of *all* the children (boys and girls), had *negative* *r*-values for all six correlations between neighborhood measures and over-regularization rates (Table 2).

Fourth, if the adults in the boys' transcripts produced fewer regular neighbors than the adults in the girls' transcripts, this could lead to floor effects in the boys' neighborhood measures, which might in turn weaken their correlations as compared to those of the girls. To test this hypothesis, for each child we summed the neighborhood values of all their irregular past-marked forms (correct or over-regularizations). That is, for a given child, we computed the total adult token frequency (over the adults in that child's transcripts) of all regular neighbors of all irregular past-marked forms produced by that child. For all six neighborhood measures (rhyme, final-coda and final-consonant, calculated over monosyllabic neighbors only, or over all neighbors), these summed neighborhood values were non-significantly *larger* for the boys than the girls (*p*-values between .74 and .91). Therefore floor effects among the boys' neighborhood values cannot explain the pattern of weaker correlations among the boys than the girls.

Note that the fact that the girls did *not* have larger regular neighborhood values than the boys also argues against another alternative explanation for the finding that the girls over-regularized more than the boys, as reported above. If it *were* the case that the girls had larger neighborhood values than the boys, it would indicate that the girls heard more regular neighbors than the boys in adult speech during the recording sessions. Therefore the girls' regular neighbors might be expected to be more easily accessed than the boys', either because of greater priming of these forms in the recording sessions, or because such a sex difference could be taken

Table 2 r-*values of correlations between over-regularization rates and regular neighborhood strength measures, partialing out adult irregular past-tense frequencies*

		r-values of correlations between over-regularization rates and neighborhood measures based on *token* frequency, calculated over *monosyllabic* neighbors			*r*-values of correlations between over-regularization rates and neighborhood measures based on *token* frequency, calculated over *all* neighbors			*r*-values of correlations between over-regularization rates and neighborhood measures based on *type* frequency, calculated over *monosyllabic* neighbors	
Sex	Subject	Rhyme	Final-coda	Final-consonant	Rhyme	Final-coda	Final-consonant	Rhyme	Final-coda
f	Allison	−0.13	−0.11	−0.11	−0.13	−0.11	−0.11	0.36	0.36
f	April	0.36	0.23	0.23	0.36	0.23	0.23	0.19	−0.15
f	Eve	0.08	0.22	0.18	0.08	0.23	0.20	0.46	−0.02
f	GAT	0.06	0.34	0.26	0.02	0.31	0.21	0.04	0.18
f	JUB	0.08	0.27	0.20	0.09	0.15	0.13	0.19	0.42
f	Naomi	0.04	0.14	0.13	0.06	0.17	0.16	0.04	−0.06
f	Sarah	0.02	0.07	0.04	0.03	0.09	0.07	0.08	0.06
f	TRH	0.26	0.58	0.58	0.20	0.54	0.55	0.29	0.26
Female mean (SD)		.10 (.15)	.22 (.20)	.19 (.20)	.09 (.14)	.20 (.19)	.18 (.18)	.21 (.15)	.13 (.21)
t-test against 0		$t(7) = 1.81$ $p = .11$	$t(7) = 2.94$ $p = .02$	$t(7) = 2.57$ $p = .04$	$t(7) = 1.71$ $p = .13$	$t(7) = 2.94$ $p = .02$	$t(7) = 2.64$ $p = .03$	$t(7) = 3.68$ $p = .01$	$t(7) = 1.80$ $p = .11$
m	Abe	−0.06	−0.10	−0.09	−0.06	−0.10	−0.09	0.06	−0.11
m	Adam	−0.17	0.09	0.16	−0.17	0.09	0.14	−0.07	0.07
m	ANC	0.53	0.28	0.24	0.28	0.25	0.26	0.22	0.15
m	BOM	−0.14	−0.14	−0.14	−0.12	−0.13	−0.27	−0.10	−0.18
m	BRD	0.00	0.11	0.11	0.00	0.23	0.14	0.14	0.14
m	CHJ	−0.25	−0.29	−0.23	−0.25	−0.29	−0.23	−0.28	−0.28
m	DED	−0.18	−0.31	−0.26	−0.24	−0.36	−0.32	0.00	−0.24
m	MAA	−0.19	−0.26	−0.30	−0.29	−0.18	−0.20	−0.14	−0.13
m	Nathaniel	−0.03	0.15	0.17	−0.02	0.16	0.18	0.32	0.00
m	Peter	−0.09	−0.04	−0.08	−0.14	−0.13	−0.17	−0.04	−0.08
m	VOH	0.00	0.31	−0.07	0.00	0.32	−0.05	0.12	−0.08
Male mean (SD)		−.05 (.21)	−.02 (.22)	−.04 (.19)	−.09 (.16)	−.01 (.23)	−.05 (.20)	.02 (.17)	−.07 (.14)
t-test against 0		$t(10) = .70$ $p = .50$	$t(10) = .26$ $p = .80$	$t(10) = .79$ $p = .45$	$t(10) = 1.91$ $p = .08$	$t(10) = .19$ $p = .85$	$t(10) = .88$ $p = .40$	$t(10) = .36$ $p = .72$	$t(10) = 1.58$ $p = .15$
t-test: males vs. females		$t(17) = 1.58$ $p = .13$	$t(17) = 2.37$ $p = .03$	$t(17) = 2.60$ $p = .02$	$t(17) = 2.51$ $p = .02$	$t(17) = 2.15$ $p = .05$	$t(17) = 2.56$ $p = .02$	$t(17) = 2.45$ $p = .03$	$t(17) = 2.50$ $p = .02$

to reflect a more general pattern of greater input of these forms to the girls than the boys. Such differences could lead to stronger neighborhood effects for girls than boys. However, the finding that the girls did *not* have larger neighborhood values strongly argues against this hypothesis. Moreover, as discussed above in the context of adult irregular past-tense frequencies, the lack of sex differences in the number of recording sessions, the number of utterances in the recordings and the type frequency of irregular verbs used, reinforces the conclusion that the girls did not hear more regular neighbors than the boys.

The finding that the girls' correlations were weaker for the rhyme measure than for the two coda-related measures is also of interest. On the one hand, this pattern might be interpreted as revealing aspects of the underlying lexical organization, such that associative generalization occurs across regulars that are similar to the over-regularized verbs in their coda rather than across those that are similar in their entire rhyme. On the other hand, the rhyme-based correlations might be weaker for a much less interesting reason. In particular, since a given over-regularization is likely to have fewer rhyme neighbors than final-coda or final-consonant neighbors, a lack of variability from floor effects is more likely for the rhyme measure than for the other two measures, particularly given that these token-frequency based neighborhood measures were computed over relatively small transcripts. Such floor effects should in turn result in weaker correlations for the rhyme measure than the two coda-related measures.

Indeed, whereas for the monosyllabic rhyme measure 167 of the 249 irregular verbs used by the eight girls in a past-marked form (counting the same verb more than once if used by more than one girl) had a value of zero, for the final-coda measure 74 irregular verbs had zero values, and for the final-consonant measure only 38 verbs had zero values. That is, we counted the number of different irregular verbs that each girl used one or more times as a past-marked form (correct or over-regularized), and then summed these numbers over all eight girls, yielding a total of 249; for 167 of these verbs, no rhyming neighboring regulars were used by adults in the given child's transcript, whereas for only 74 of the verbs were there zero occurrences of neighbors at the level of final-coda. A similar pattern was found for the measures that included polysyllabic neighbors, with 143, 53 and 27 verbs having zero values for the rhyme, final-coda and final-consonant measures, respectively. Thus the weak rhyme-based correlations found for the girls seem likely to be at least partly due to floor effects.

In order to test the reliability and generality of the sex differences observed in the correlations between regular neighborhood strength and over-regularization rates, as well as to further examine the hypothesis that the rhyme measure was a weak predictor due to floor effects, we also examined a measure of neighborhood strength that was not considered in MEA. As discussed above, the regular neighborhood strength measures reported in MEA were computed over the *token frequencies* of neighboring regulars in the adult speech in the transcripts. However, phonological neighborhood measures can also be calculated on the basis of *type frequency* – that is, the number of different words that are neighbors. For example, if *linked* and *blinked* were the only rhyme neighbors of *sinked*, then its type-frequency regular rhyme neighborhood strength would be two. Because token-frequency neighborhood measures take into account not only the type frequencies of neighbors but *also* their individual token frequencies (since these measures are calculated on the basis of the token frequencies of all neighbors), it might be argued that token-frequency neighborhood measures should more accurately reflect associative generalization, and thus should lead to stronger correlations than measures based on type frequency alone. However, the transcripts over which the token frequency measures were calculated are relatively small. Therefore type-frequency neighborhood measures (that are summed over a reasonably large number of words in the language) might actually show stronger correlations than token-frequency neighborhoods, especially for rhyme measures, simply because they avoid floor effects.

We based our analyses on type-frequency regular neighborhoods that we had previously computed for other purposes. For each over-regularization, the type-frequency regular neighborhood value had been calculated as the number of neighbors in a set of 1808 monosyllabic regular verbs drawn from a computerized version of *Webster's Seventh Collegiate Dictionary* (Merriam-Webster, 1967). This procedure had been performed separately for rhyme neighbors and final-coda neighbors. It was not performed for polysyllabic verbs or for final-consonant neighbors. (Note that no type-frequency regular neighborhood measures were calculated on the basis of irregular past-tense types used by the adults in each child's recordings, because floor effects in such measures would be even more problematic than in the token-frequency measures reported above.) Using these type-frequency neighborhood measures, we performed the exact same partial correlations as described above for each of the 19 subjects.

As shown in Table 2 (two rightmost columns), the correlations based on the type-frequency neighborhood measures yielded the same general pattern of results as the correlations based on the token-frequency neighborhood measures. For both the rhyme and the final-coda type-frequency measures, the sex difference in *r*-values was statistically significant (*p*s $< .05$). For both measures the girls elicited positive mean *r*-values. These were significantly greater than zero for the rhyme measure ($p < .01$), and showed a trend in comparison to zero for the final-coda measure ($p = .11$). In contrast, for the boys the mean *r*-value did not differ at all from zero for the rhyme measure ($p = .72$), and was not even positive for the final-coda measure.

Thus, despite the close similarity between these analyses and those based on token-frequency neighborhood measures, the two sets of analyses yielded different results. Unlike the correlations based on token-frequency measures, here the rhyme measure yielded *stronger* correlations than the final-coda measure. Also unlike the token-frequency measures, there was no evidence for floor effects among the type-frequency measures: *none* of the 90 irregular verbs over which the correlations were computed had type-frequency neighborhood counts of zero, for either the rhyme or the final-coda measures. These data strengthen the view that the weaker correlations for the token-frequency rhyme measures were due to floor effects, and suggest that associative generalization to over-regularizations among these girls may in fact have taken place largely at the level of rhyme. Further studies examining these issues are warranted.

In sum, converging evidence from both token- and type-frequency-based regular neighborhood measures, at the level of both rhyme and final coda, consistently suggest a sex difference, with girls but not boys showing positive correlations between regular neighborhood effects and over-regularization rates. These correlations, which do not appear to be explained by a variety of potential confounding factors, suggest that girls have a much stronger tendency than boys to produce over-regularizations by generalizing across stored neighboring regulars.

General discussion

The data presented here suggest a striking and reliable sex difference in over-regularization, with girls over-regularizing at a higher rate than boys. We hypothesized that girls tend to over-regularize as a result of associative generalization, whereas boys' over-regularizations are largely rule-products. In support of this hypothesis, regular neighborhood measures correlated positively with over-regularization rates among girls but not among boys. A variety of potential confounding factors and alternative explanations did not seem to account for the sex differences either in over-regularization rates or in the correlations between over-regularization and regular neighborhood strength.

Girls' higher over-regularization rates may be explained by the apparent sex difference in the manner in which these forms are computed. First, holding other factors constant, associative over-regularization should occur at a higher rate than rule-based over-regularization: whereas in both cases irregular past-tenses with weak memory traces (e.g. those of low frequency) are expected to yield more over-regularizations, in the former case over-regularizations should also be *encouraged* by stored similar-sounding regulars. Even though the hypothesized female advantage at remembering words is expected to provide girls with stronger irregular past-tense memory traces than boys, or better access or retrieval of these forms, girls should *also* have a corresponding advantage for stored similar-sounding regulars – which, depending on the relative type and token frequencies of the regulars and irregulars, may lead to a higher over-regularization rate than boys. Finally, even when over-regularizations cannot be produced associatively (e.g. due to too few regular neighbors), girls can *additionally* rely on the rule. That is, girls seem to have two potential routes to over-regularization, while boys may tend to have only one.

However, our interpretations must be treated with some caution. First, although we have provided evidence for sex differences both in over-regularization rates and in the apparent associative computation of these forms, we have not directly shown that the former depends upon the latter – that is, that girls' higher over-regularization rates are actually due to greater associative generalization. Second, while the sex difference in neighborhood correlations suggests the associative generalization of over-regularizations by girls but not (or much less so) by boys, we have presented no direct evidence that the boys over-regularized via rule-based computation. Similarly, the data do not tell us whether or to what extent the girls may be relying on rule-based computation as well as associative generalization in their production of over-regularizations. Unfortunately, the analytical tools at our disposal do not easily lend themselves to direct testing of rule-based computation in the present data set. Other types of studies, in particular experiments in which inflected verb forms are elicited, may be better suited for investigating this question (Hartshorne, Walenski & Ullman, 2003; Ullman, Walenski, Prado, Ozawa & Steinhauer, 2001). Third, the evidence for associative processing in girls is itself not without problems. The number of subjects was relatively small, and not all of the sex differences in *r*-values for the correlations between over-regularization rates and neighborhood measures reached statistical significance at the .05 level. Moreover, despite our efforts at examining alternative explanations for these correlations, there are probably other accounts that we have not addressed which might also explain the data. (See below for a discussion of single-system perspectives.) Fourth, our starting assumptions about sex differences in declarative and lexical memory are still somewhat tentative. Although a wide range of evidence has suggested a female superiority in verbal declarative memory (see Introduction), we were unable to test this sex difference directly in these subjects (given that the data were acquired many years ago), and therefore we cannot draw the strong conclusion that sex differences in declarative memory were directly related to the findings reported here. Additionally, the independent evidence that females are more likely than males to memorize regulars is still preliminary (Steinhauer & Ullman, 2002; Ullman *et al.*, 2002; Ullman, Hartshorne *et al.*, under revision; Ullman, Walenski *et al.*, under revision).

For all these reasons, it is critical that the findings reported here be replicated, and the broader issues further investigated using a range of experimental approaches. Nevertheless, on the basis of this first examination of sex differences in over-regularization, we believe that the data suggest that girls over-regularize more than boys, and that this pattern can be explained, at least in part, by a greater tendency for girls than boys to compute over-regularizations by associative generalization over similar-sounding stored regular forms.

Dual-system vs. single-system models

This view of lexical memory as a repository of stored regular and irregular forms, over which generalization to new forms can occur, is in some ways similar to the single-system ('single-mechanism') view. On this perspective, all linguistic forms are learned, represented and processed in an associative memory (McClelland & Patterson, 2002; Plunkett & Marchman, 1996; Rumelhart & McClelland, 1986). Over-regularizations occur because irregular past-tenses, being at least somewhat idiosyncratic, cannot always out-compete the regular pattern, particularly when the irregular is of low frequency or is surrounded by many similar-sounding regulars. Young children have not heard each irregular very often, so the regular pattern can overwhelm irregular representations, leading to over-regularizations. In adults this is much less likely to occur because most irregular forms have been encountered sufficiently frequently to yield strong mappings which are resistant to the competition of surrounding regulars.

This logic can be extended to a dual-system perspective in which lexical memory can contain a large number of stored regulars. On this view, over-regularizations in girls, and the lack thereof in women, can be largely explained by such associative memory mechanisms. However, we are *not* claiming that females depend *only* on lexical memory for processing complex forms. Even with their excellent memory abilities, females are expected to compose many types of complex forms, including new and lower frequency regulars, and highly complex linguistic representations, including most phrases and sentences (Ullman, 2005; Ullman, Hartshorne *et al.*, under revision). Thus our dual-system claim that females memorize many regulars differs qualitatively from the single-mechanism claim that regulars can *only* be learned and processed in associative memory. It must also be emphasized that we are not arguing that males never memorize regulars. Indeed, recent preliminary evidence suggests that men memorize higher frequency regulars (Ullman, Hartshorne *et al.*, under revision). Thus the hypothesized sex difference is not posited to be all-or-nothing. Rather, females are predicted to rely more on memorized and associative-memory-generated regulars than males, who thus depend more than females on the rule-governed combination of complex forms.

The results presented above suggest that single-mechanism accounts inspired by Rumelhart and McClelland (1986) may explain a greater portion of the data than earlier dual-system critiques such as MEA have accepted. However, the observed sex differences also strengthen the view that this is not the whole story. As we have argued, they rather support a dual-system model in which many complex forms are stored in lexical memory, at least partially as a function of sex. The data seem to be especially compatible with the DP model's prediction that girls are more likely than boys to remember regular forms. Given that single-mechanism models have not addressed the issue of sex differences, it remains to be seen whether and how the observed patterns can be simulated by such models.

Implications

This study has a number of implications. First, it suggests that sex may be an important factor in the neurocognition of language. Until now, sex has been virtually ignored in studies of the learning, representation, processing and neural bases of language. The data presented here demonstrate the desirability of further investigations of the effects of sex on language, and suggest that this factor should be taken into account in a range of language studies. The consistent introduction of sex as a design factor may help to reduce the heterogeneity of findings across studies, which might have varied in previous investigations partly as a result of inconsistent sex ratios among subjects. Additionally, as in the present study, acknowledging sex differences seems likely to reveal further complexities of the language systems under investigation.

Second, evidence linking estrogen to performance on verbal declarative memory tasks (McEwen, Alves, Bulloch & Weiland, 1998; Woolley & Schwartzkroin, 1998) suggests that the effect of sex hormones in language acquisition should also be investigated. Sex hormone levels vary not only across individuals, but also within subjects, over relatively short periods of time (e.g. across the menstrual cycle in women, and across the seasons in men; see Halpern, 2000; Kimura, 1999), as well as with age (again, in both sexes, increasing in early years, and decreasing in later years; e.g. see Ullman, 2005). Therefore language-related functions may also vary between as well as within individuals, changing over the course of weeks, months and years.

Third, it is important to underscore the point that sex differences in language processing do not preclude individual differences within each sex. Indeed, it would be surprising if such individual variability were not found, particularly given the between- and within-subject heterogeneity in underlying biological factors such as sex hormones. Thus variability in subjects' relative dependence on rule-processing versus associative memory may also be found within each sex. In our view, such variability does not detract from the interest of group differences in language between sexes, even if it turns out (which might or might not be the case) that such within- and between-sex differences in language depend upon the same biological and neurocognitive factors.

Fourth, the observed sex differences in over-regularization rates suggest that superior linguistic abilities can – paradoxically – lead to worse rather than better performance (also see MacDonald, Just & Carpenter, 1992). That is, girls' hypothesized superior lexical memory abilities may lead to more rather than fewer over-regularizations, thanks to their memorization and associative generalization of regular past-tense forms. Such paradoxical effects clearly warrant further investigation, both in inflectional morphology and in other domains of language.

Fifth, this study highlights the importance of further investigations of phonological neighborhood effects. In particular, a better understanding of the predictiveness of token- versus type-based neighborhood measures, as well as the differential effects of measures calculated over different phonological units (e.g. rhyme, final-coda, etc.), would improve the effectiveness of employing phonological neighborhood measures as an analytical tool in the investigation of the computational bases of morphology.

Finally, we would like to emphasize that the work presented here suggests that one should be very cautious in assuming substantial homogeneity across subjects in the neural, cognitive and computational bases of language. Rather, even normal individuals and groups may vary dramatically in the acquisition and use of language. To date, such variability has been largely ignored in the study of language. We believe that theories must address between- as well as within-individual subject differences – not only in performance, but also in the underlying *neurocognitive processes*. Thus in the study reported here, subjects differed not only in their over-regularization rates, but also apparently in how those over-regularizations were computed. We believe that such neurocognitive variability will turn out to be relatively common, and that its investigation may help to resolve differences between competing theoretical accounts. We therefore suggest that subject variability should become a major focus of research in the study of language.

Acknowledgements

We thank Claudia Bonin, Rusan Chen, Matthew Moffa, Elizabeth Prado, Steven Pinker, several anonymous reviewers, and especially Matthew Walenski for helpful suggestions, and Christopher Maloof for many insights and analyses. Support was provided to MTU by NSF SBR-9905273, NIH R01 MH58189 and NIH R01 HD049347.

References

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, **40**, 41–61.

Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax*. The Hague: Mouton.

Bohannon, J.N., & Marquis, A.L. (1977). Children's control of adult speech. *Child Development*, **48**, 1002–1008.

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

Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, **22** (6), 991–1060.

Golomb, J., Kluger, A., de Leon, M.J., Ferris, S.H., Mittelman, M., Cohen, J., & George, A.E. (1996). Hippocampal formation size predicts declining memory performance in normal aging. *Neurology*, **47** (3), 810–813.

Hall, W.S., Nagy, W.E., & Linn, R. (1984). *Spoken words*. Hillsdale, NJ: Erlbaum.

Halpern, D.F. (2000). *Sex differences in cognitive abilities* (3rd edn.). Mahwah, NJ: Lawrence Erlbaum Associates.

Hartshorne, J.K., Walenski, M., & Ullman, M.T. (2003). An investigation of the role of the declarative and procedural memory systems in language using a dual-task paradigm. *Journal of Cognitive Neuroscience, Supplement*, 209.

Higginson, R. (1985). Fixing-assimilation in language acquisition. Unpublished doctoral dissertation, Washington State University.

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: The MIT Press.

Kramer, J.H., Delis, D.C., Kaplan, E., O'Donnell, L., & Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology*, **11** (4), 577–584.

Kuczaj, S.A. (1976). -ing, -s; -ed: a study of the acquisition of certain verb inflections. Unpublished PhD dissertation, University of Minnesota.

McClelland, J.L., & Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences*, **6** (11), 465–472.

MacDonald, M.C., Just, M.A., & Carpenter, M. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, **24**, 56–98.

McEwen, B.S., Alves, S.E., Bulloch, K., & Weiland, N.G. (1998). Clinically relevant basic science studies of gender differences and sex hormone effects. *Psychopharmacology Bulletin*, **34** (3), 251–259.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd edn.). Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B., & Snow, C.E. (1985). The Child Language Data Exchange System. *Journal of Child Language*, **12**, 271–296.

Marcus, G.F., Pinker, S., Ullman, M., Hollander, M., Rosen, T.J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, **57** (4, Serial No. 228), 1–165.

Merriam-Webster (Ed.) (1967). *Webster's Seventh Collegiate Dictionary*. Los Angeles, CA: Library Reproduction Service.

Muthen, L.K., & Muthen, B.G. (1998–2004). *MPlus User's Guide*, 3rd edn. Los Angeles, CA: Muthen & Muthen.

Pfluger, T., Weil, S., Weis, S., Vollmar, C., Heiss, D., Egger, J., Scheck, R., & Hahn, K. (1999). Normative volumetric data of the developing hippocampus in children based on magnetic resonance imaging. *Epilepsia*, **40** (4), 414–423.

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

Pinker, S., & Ullman, M.T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, **6** (11), 456–463.

Plunkett, K., & Marchman, V.A. (1996). Learning from a connectionist model of acquisition of the English past tense. *Cognition*, **61** (3), 299–308.

Prasada, S., & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, **8** (1), 1–56.

Rumelhart, D.E., & McClelland, J.L. (1986). On learning the past tenses of English verbs. In J.L. McClelland, D.E. Rumelhart & PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructures of cognition* (Vol. 2, pp. 216–271). Cambridge, MA: Bradford/MIT press.

Sachs, J. (1983). Talking about there and then: the emergence of displaced reference in parent–child discourse. In K.E. Nelson (Ed.), *Children's language* (Vol. 4, pp. 1–28). Hillsdale, NJ: Erlbaum.

Squire, L.R., & Knowlton, B.J. (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. In M.S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 765–780). Cambridge, MA: MIT Press.

Steinhauer, K., & Ullman, M.T. (2002). Consecutive ERP effects of morpho-phonology and morpho-syntax. *Brain and Language*, **83**, 62–65.

Ullman, M.T. (2001a). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, **30** (1), 37–69.

Ullman, M.T. (2001b). A neurocognitive perspective on language: the declarative/procedural model. *Nature Reviews Neuroscience*, **2**, 717–726.

Ullman, M.T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, **92** (1–2), 231–270.

Ullman, M.T. (2005). A cognitive neuroscience perspective on second language acquisition: The Declarative/Procedural Model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition* (pp. 141–178). Washington, DC: Georgetown University Press.

Ullman, M.T., Corkin, S., Coppola, M., Hickok, G., Growdon, J.H., Koroshetz, W.J., & Pinker, S. (1997). A neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, **9** (2), 266–276.

Ullman, M.T., Estabrooke, I.V., Steinhauer, K., Brovetto, C., Pancheva, R., Ozawa, K., Mordecai, K., & Maki, P. (2002). Sex differences in the neurocognition of language. *Brain and Language*, **83**, 141–143.

Ullman, M.T., Hartshorne, J.K., Estabrooke, I.V., Brovetto, C., & Walenski, M. (under revision). Sex, regularity, frequency and consistency: a study of factors predicting the storage of inflected forms.

Ullman, M.T., Walenski, M., Prado, E., Ozawa, K., & Steinhauer, K. (2001). A new method for investigating the real-time composition of complex linguistic forms. In *Proceedings of the Fourteenth Annual CUNY Conference on Human Sentence Processing* (Vol. 14, p. 64). Philadelphia, PA: CUNY Graduate School and University Center.

Ullman, M.T., Walenski, M., Prado, E., Ozawa, K., & Steinhauer, K. (under revision). The real-time composition of complex forms. *Cognition*.

Woolley, C.S., & Schwartzkroin, P.A. (1998). Hormonal effects on the brain. *Epilepsia*, **39** (8), S2–8.

Received: 19 May 2004
Accepted: 29 April 2005