Semantic Representation and Naming in Children With Specific Language Impairment

Karla K. McGregor
Robyn M. Newman
Renée M. Reilly
Nina C. Capone
Northwestern University
Evanston, IL

When 16 children with SLI (mean age = 6;2) and 16 normally developing age-mates named age-appropriate objects, the SLI cohort made more naming errors. For both cohorts, semantic misnaming and indeterminate responses were the predominant error types. The contribution of limited semantic representation to these naming errors was explored. Each participant drew and defined each item from his or her semantic and indeterminate error pools and each item from his or her correctly named pool. When compared, the drawings and definitions of items from the error pools were poorer, suggesting limited semantic knowledge. The profiles of information included in definitions of items from the correct pool and the error pools were highly similar, suggesting that representations associated with misnaming differed quantitatively, but not qualitatively, from those associated with correct naming. Eleven members of the SLI cohort also participated in a forced-choice recognition task. Performance was significantly lower on erroneous targets than on correctly named targets. When performance was compared across all three post-naming tasks (drawing, defining, recognition), the participants evinced sparse semantic knowledge for roughly half of all semantic misnaming and roughly one third of all indeterminate responses. In additional cases, representational gaps were evident. This study demonstrates that the degree of knowledge represented in the child's semantic lexicon makes words more or less vulnerable to retrieval failure and that limited semantic knowledge contributes to the frequent naming errors of children with SLI.

KEY WORDS: lexicon, semantics, naming, specific language impairment

The focus of this paper is the quality of semantic representations in the lexicons of children who have specific language impairment (SLI). It is often reported that children with SLI have smaller lexicons than their age-mates. Late onset of lexical acquisition is frequently the first sign of SLI, and estimates of lexical knowledge obtained from parent-report instruments, standardized vocabulary tests, and spontaneous language samples all serve to differentiate children with SLI from normally developing children (e.g., Bishop, 1997; Watkins, Kelly, Harbers, & Hollis, 1995). In experimental word-learning situations, children with SLI are deficient at establishing initial maps of new words compared to their age-mates. This deficiency is manifested as a small quantity of words learned (e.g., Rice, Buhr, & Nemeth, 1990; Rice, Buhr, & Oetting, 1992; Rice, Oetting, Marquis, Bode, & Pae, 1994).

Much less attention has been paid to the *quality* of stored lexical knowledge of these children. The only relevant studies have focused on the robustness of phonological representations. Dollaghan (1987) demonstrated

1092-4388/02/4505-0998

that children with SLI could map the referent and context of a new word after minimal exposure but, unlike normally developing age-mates, could not map enough of its phonological form to support production. Bishop (1997), reviewing many studies wherein children with SLI demonstrated recognition but not production of new words, suggests that underspecified, global phonological representations characterize the SLI lexicon for a longer than expected period of development. This hypothesis is supported by children's word- recognition performance under conditions of gating (wherein increasingly greater portions of the spoken word are presented). School-age children with SLI must hear significantly more of the spoken word than their normally developing age-mates to recognize newly learned words, but not to recognize familiar words (Dollaghan, 1998). Although semantic representations may also be underdeveloped in children with SLI, empirical data are lacking.

Children with SLI also demonstrate difficulties with retrieval from their long-term lexical memory. When compared to their normally developing age-mates, they make more naming errors during object naming, action naming, and story retelling (McGregor, 1997). Frequently, their errors bear semantic relations to their targets. Semantic coordinate substitutions such as *mouse* for *kangaroo* are especially common. Nearly as often, the relation to the target cannot be determined. These indeterminate responses are typically "don't know" answers (McGregor, 1997; McGregor & Waxman, 1998). Even when children with SLI name correctly, they take a longer time to retrieve names than do their peers (e.g., Katz, Curtiss, & Tallal, 1992; Lahey & Edwards, 1996; Leonard, Nippold, Kail, & Hale, 1983).

We hypothesize that the word-retrieval problems of children with SLI are one manifestation of slow language development in general and underdeveloped semantic representations in long-term lexical memory in particular. This hypothesis parallels posited relations between memory development and increasing automaticity of retrieval in normally developing children. In his theory of children's memory, Bjorklund (1987) argued that growth in the content and organization of semantic memory during childhood influences the ease with which information can be retrieved. He posited that items within a detailed knowledge base are represented robustly in terms of semantic features and semantic relationships and that this robust representation results in low activation thresholds and a high likelihood of spreading activation during retrieval and other aspects of cognitive processing. Among other data, Bjorklund reviewed an interesting case study of a 4-year-old "dinosaur expert" to support his hypothesis. In this case study, conducted by Chi and Koeske (1983), the likelihood of the child's recalling a given dinosaur name during a short-term memory task (list recall) was positively influenced by the number of relationships between that name and others in his dinosaur lexicon. Additional studies reveal that children can retrieve more words from short-term memory than adults can when those words are more familiar to the children (Lindberg, 1980; Roth, 1983). Children demonstrate effects of familiarity on retrieval during tasks that tap long-term lexical memory as well (Leonard et al., 1983).

Further evidence of the interrelatedness of representation and retrieval is evident in a U-shaped developmental phenomenon. After a few months of talking, normal toddlers experience a transient increase in word-retrieval disruptions. These disruptions often manifest as the substitution of a previously retrieved, semantically similar word for a target word (Dapretto & Bjork, 2000; Gershkoff-Stowe, 2001; Gershkoff-Stowe & Smith, 1997). Gerskhoff-Stowe and Smith concluded that this temporary word-retrieval problem results, in part, from the fragility of representation characteristic of all words in the lexicons of children so young, together with the increased opportunities for competing activation that arise when these lexicons are newly expanded.

To date, the relation between robustness of semantic representation and word retrieval in children with SLI has not received direct attention. A monograph on word retrieval in children with SLI (Kail & Leonard, 1986) thoroughly documented short-term and long-term naming deficits but did not include evidence about the quality of stored semantic representations associated with these deficits. Later treatment studies, wherein naming improved after children learned new information about selected words (McGregor & Leonard, 1989; Wright, 1993) did demonstrate a relation between robustness of storage and retrieval; however, these studies did not establish the relation before intervention and, therefore, are not ideal tests.

In an initial attempt to explore the robustness of semantic storage in children and to test for links between representation and retrieval problems, McGregor and Appel (2002) introduced the comparative picture-naming/picture-drawing procedure. A single 5-year-old boy with SLI participated in the prototype procedure. He first named a set of age-appropriate, pictured objects, and then he drew pictures of these same objects from memory. A panel of judges rated the accuracy of the drawings with the result that drawings of objects that were named with semantic errors such as *hat* for *helmet* were significantly poorer than drawings of correctly named objects. Because both naming and drawing tap common semantic representations but involve disparate perceptual and motor processes (Hillis & Caramazza, 1991; Snodgrass, 1984), the association between naming errors and poor drawings suggests that limited semantic representations are related to these problems. Importantly, this case study demonstrated links between sparse representations and poor naming of specific words.

McGregor, Friedman, Reilly, and Newman (2002) extended the comparative picture-naming/picture-drawing procedure to a group of normally developing children. The prediction was that the naming errors of normally developing children, though more rare than in SLI, would also reflect limited semantic knowledge. To test this prediction, 25 5-year-olds named and drew 20 age-appropriate objects. Semantic naming errors were their most frequent error type. As predicted, items from this semantic error pool were drawn significantly less accurately than items from the correctly named pool. The results of a definition task, involving the same 20 items, yielded parallel results. In a second experiment, the drawing results were replicated with a new group of 16 normally developing children. McGregor and colleagues concluded that semantic word retrieval errors of young normally developing children are related to degree of semantic knowledge.

Purpose of the Current Study

In the current study, we extended the comparative picture-naming/picture-drawing procedure to a group of children with SLI, and we used the procedure to examine the source of not only semantic misnaming but of indeterminate errors as well. As in McGregor et al. (2002), we also included a more conventional task—word defining—to provide evidence of concurrent validity for the comparative picture-naming/picture-drawing procedure and to provide additional evidence relevant to our hypothesis. We compared the performance of children with SLI to that of their normally developing (ND) agemates. Our goal was two-fold: to explore the robustness of semantic representations in the lexicons of children with SLI and to test for relationships between semantic representations and naming errors.

Our focus on semantic representation, as opposed to phonological representation, was motivated by several factors. First, the building of semantic representations normally requires a long developmental course (Bloom, 2000), and therefore we would expect underdevelopment to be particularly persistent in children with SLI. Second, semantic naming errors are the most frequent naming error type among children with SLI (Lahey & Edwards, 1999; McGregor, 1997; McGregor & Waxman, 1998). Finally, Leonard (1999) proposed that the lexical storage problems associated with SLI compromise semantic representation. He viewed "…SLI as a type of filter such that some but not all experiences with a word are registered in *semantic* memory" (p. 47, emphasis added). Therefore, an exploration of semantic representation in children with SLI is important for developmental, clinical, and theoretical reasons.

With these motivations and a hypothesized link between semantic representation and retrieval problems in mind, we predicted:

- Children with SLI will make more naming errors than their ND peers.
- For both groups, drawings of misnamed objects will be poorer than drawings of correctly named objects.
- For both groups, definitions of misnamed objects will be poorer than definitions of correctly named objects.

Finally, the information included in the definitions was explored to determine whether there were qualitative differences in the type of stored semantic knowledge associated with correct and erroneous naming.

Method

Participants

Participants were 16 children with SLI and 16 normally developing children. The performance of 10 of the children in the ND group was also reported in Experiment 1 of McGregor et al. (2002). Age was balanced between groups by matching each child with SLI to a child in the ND group by ±3 months. The mean ages for the SLI and ND groups were 6;2 (range = 5;0 to 7;11) and 6;1 (range = 4;9 to 8;1), respectively, with no statistical difference between groups, $t(30) = .18$, $p = .86$. Gender was balanced both within and between groups. Ethnicity was balanced between groups. Within each group, 70% of the sample were Caucasian, and 30% represented minority groups—either African American or Hispanic. This representation closely mimics that in the U.S. population as a whole (U.S. Census Bureau, 2001: Caucasian = 75%; Minority = 25%). Between-group differences in socioeconomic status, as measured by years of maternal education, approached significance [SLI: $M = 14.25$, $SD = 2.38$; ND: $M = 15.75$, $SD = 2.14$; $t(30) = 1.87$, $p = .07$]. For this reason, the between-group statistical comparison of the naming data employed years of maternal education as a covariate.

The children with SLI were selected on the basis of two criteria: (1) current enrollment in language or reading intervention and (2) a score more than 1.3 standard deviations below the expected mean for total phonemes correct on the Nonword Repetition Task (NWRT; Dollaghan & Campbell, 1998). Intervention status was used as a selection criterion because enrollment in intervention reflects decisions made on the basis of more data (e.g., test scores, classroom observation, and parent report), collected in more naturalistic environments, for longer periods of time, and judged in a more integrative manner than the data that could be collected for the purposes of this project alone. Furthermore, such clinical decisions are shown to have moderately high interrater and intrarater reliability (Records & Tomblin, 1994).

The NWRT was used as a selection tool because it is accurate in screening for SLI (Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000) and unbiased against minority participants (Campbell, Dollaghan, Needleman, & Janosky, 1997; Ellis Weismer et al., 2000). Furthermore, deficient nonword repetition ability is a proposed phenotypic marker of SLI (Bishop, North, & Donlan, 1996). Although the norms published by Dollaghan and Campbell (1998) did not include children under 6;0, our pilot testing on 21 normally developing 5-year-olds and 15 normally developing 6-year-olds yielded means of 83% and 84%, respectively, with no significant differences between groups [t (34) = 0.39, p = .70]. The combined mean score for the 5- and 6-year-olds was 84%, with a *SD* of 7.8— figures nearly identical to those reported by Dollaghan and Campbell. Therefore, we used their normative data, and we applied the same –1.3 *SD* cutoff criterion to both 5- and 6-year-old participants. In the current study the mean scores for percentage of total phonemes correct on the NWRT were 68.73% (*SD* = 4.43) and 85.71% (*SD* = 5.68) for the SLI and ND groups, respectively. This was a significant difference [t (34) = 1.7, $p < .001$, d = 2.9].

Children were excluded from participation if they scored at or below 70 on the Columbia Mental Maturity Scale (CMMS; Burgemeister, Blum, & Lorge, 1972); if they failed a pure tone audiometric screening administered per ASHA (1985) guidelines on the first day of selection testing; or if they presented with physical, sensory, or emotional deficits that affect speech and language development via parent report. We set the CMMS cutoff at 70 (rather than the more frequently used cutoff of 85) because children with nonverbal IQ levels of 70 to 84 and children with nonverbal IQs of 85 and above score similarly on standardized language tests (Tomblin & Zhang, 1999), perform similarly on auditory processing measures (Bishop et al., 1999), and demonstrate similar growth in grammatical skills over time (Fey, Long, & Cleave, 1994). The ND children also met all exclusionary criteria, but unlike the SLI children, they had negative histories of language impairment and scores better than 1.3 standard deviations below the mean on the NWRT. Although both the SLI and ND groups scored within normal limits on the CMMS, the SLI group did exhibit significantly lower nonverbal IQs than their peers [SLI: M = 92.88, *SD* = 10.3; ND: M = 106.6, *SD* = 12.72; t(30) = 3.35, p = .002, d = 1.1]. For this reason, nonverbal IQ was also used as a covariate in the between-group statistical comparison of the naming data.

To further describe the language abilities of the two participant groups, the results of a conversational sample, a narrative sample, and three standardized language tests are presented in Table 1. The conversational sample was collected as the examiner and child played with a standard toy set on the first day of testing. Mean length of utterances (MLU) in morphemes and number of different words (NDW) were calculated for 50 consecutive utterances per child. The narrative sample was elicited with the wordless picture book *Frog, where are you?* (Mayer, 1969) on the final day of testing. MLU and length of narrative as measured by the number of different words per story were calculated. The examiner administered the Expressive Vocabulary Test (Williams, 1997), the Peabody Picture Vocabulary Test–III (PPVT-III; Dunn & Dunn, 1997), and one omnibus language test (either the Test of

Table 1. Mean group performance on language measures.

	Conversation[a]		Narrative[b]		Standardized tests[c]		
Group	NDW	MLU	NDW	MLU	EVT	PPVT	Omnibus
SLI							
M	84.25	4.31	57.31*	7.64	84.75*	89.06*	82.00*
SD	12.58	0.81	20.59	1.87	14.56	15.56	11.78
ND							
M	93.19	4.75	74.71	9.31	109.30	111.25	110.19
SD	13.08	0.71	25.23	2.83	15.03	15.33	15.54

Note. NDW = Number of Different Words, MLU = Mean Length of Utterance in Morphemes, EVT = Expressive Vocabulary Test, PPVT = Peabody Picture Vocabulary Test–Revised or Peabody Picture Vocabulary Test–III, Omnibus Test = Woodcock Language Proficiency Battery–Revised, Test of Language Development–Primary (2nd ed.), or Test of Language Development–Primary (3rd ed.).

[a]Calculated per 50 utterances. [b]Calculated per story. [c]Standard scores.

*$p < .04$ (all comparisons are between group).

Language Development–Primary-3 [TOLD-P3; Newcomer & Hammill, 1997] or the Woodcock Language Proficiency Battery–Revised [WLPB-R; Woodcock, 1991]) to all children with the following exception. If a child had taken any of these tests, in their current or previous editions, within a year of participation, we used the score(s) in the child's educational file and did not readminister the test(s). The two participant groups differed on narrative length measured by number of word types, as well as on expressive and receptive vocabulary and overall language ability as measured by standardized tests. Each SLI participant scored poorer than one standard deviation below the mean for his or her chronological age on at least two language sample scores and/or standardized language tests.

Stimuli

Twenty object words from the Snodgrass and Vanderwart (1980) corpus constituted the stimuli (see Appendix A). To balance the need to elicit errors with the likelihood that the children would have some knowledge of the stimuli, we selected words that were relatively low in frequency of occurrence (<50 per million) but age-appropriate (age of acquisition norms: $M = 4.09$ years; range = 3.12 to 5.48 years). Line drawings of the 20 objects provided by Snodgrass and Vanderwart were the stimuli for the picture-naming task.

Data Collection

The participants were tested individually at a school or at the Northwestern University Child Language Laboratory by either a graduate student or a speech-language pathologist with expertise in child language.

There were three tasks: naming, drawing, and defining. Because all tasks involved the same 20 stimuli, the drawing and defining followed the naming task to minimize the child's exposure to the target words before naming. Therefore, we administered the naming task in session one, half of the drawing and defining tasks in session two, and the remaining half of the drawing and defining tasks in session three. The order of the defining and drawing tasks was counterbalanced across participants. Within tasks, items to be drawn or defined during each session were randomly selected, with the caveat that the participants never drew and defined the same item within the same session. There were 2 to 10 days between sessions. The procedures for the three tasks are detailed in McGregor et al. (2002).

Data Analysis

The naming responses were coded by type. Responses coded as semantic errors were those that involved associations (e.g., *milk* for *pitcher*), circumlocutions (e.g., *something that you chop with* for *axe*), novel derivatives (e.g., *chopper* for *axe*), coordinate substitutions (e.g., *mouse* for *kangaroo*), and superordinate substitutions (e.g., *animal* for *kangaroo*). Though not technically errors, superordinate substitutions were counted as such because they are not as specific as the names the stimuli typically elicit. Errors that were not semantic in nature were coded as either indeterminate ("don't know" or nonspecific responses), phonologic (real-word or nonword approximations of the target word form), or other (unintelligible or visual misperceptions). Four randomly selected SLI files and four randomly selected ND files were coded independently for error type by a second coder. Point-to-point agreement between the two coders averaged 98% (range = 86% to 100%). Disagreements were resolved by consensus.

The drawing and defining analyses concerned only the correct, semantic error, and indeterminate error pools. The phonologic and other error pools were not studied further because they were expected to be rare.

Five adults (4 women and 1 man), who were blind to the naming responses of the children, served as drawing raters. They first examined the child's self-portrait to gauge his or her drawing aptitude. This baseline was meant to avoid confounding poor drawing ability with poor semantic knowledge. They then rated the accuracy of each drawing by marking a level of agreement (with 1 being strongest disagreement and 7 being strongest agreement) with the statement, "This young child's drawing reflects accurate and complete knowledge of X." The ratings for each object drawn by 10 children (5 randomly selected SLI, 5 randomly selected ND) were compared across raters. Interrater agreement, measured by Cronbach's alpha, was .92 (range = .77 to .97).

We coded the definitions for the amount and type of accurate information units they contained (see Table 2). In a definition such as, "A kite is a toy that flies," there are two accurate information units: one categorical (toy) and one functional (flies). The definitions of 8 children (4 randomly selected SLI, 4 randomly selected ND) were coded independently by a second coder. Point-to-point agreement between the two coders was 85% (range = 55% to 98%) and 94% (range = 80% to 100%) for number and type, respectively, and any disagreements were resolved by consensus.

Results

Naming

According to a one-way, between-subjects ANCOVA (with maternal education and nonverbal IQ as covariates), children with SLI made significantly fewer

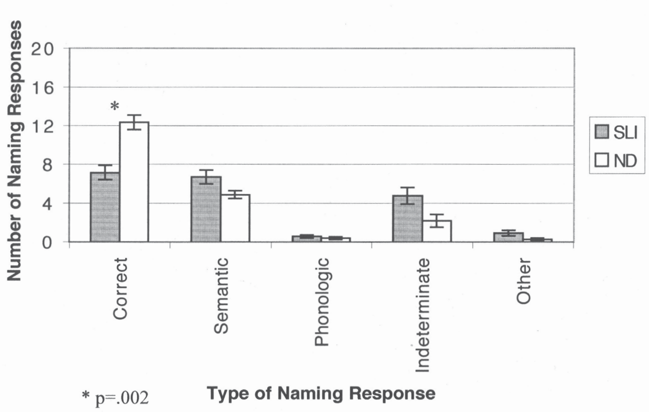
Table 2. Coding scheme for accurate information units in definitions.

Type	Definition	Example
Functional	Purpose of the target, people or instruments that act on the target, the way it is acted upon, or the outcome of the target's actions. Should answer, "Who uses it? How is it used? What does it create?" or "What is it for?"	A ladder is *to reach high.*
Physical	Typical color, size, shape, smell, feel, composition, life cycle, movement, or features of the target. Should answer, "What is it like?"	A mountain is *tall.*
Locative	Characteristic location, circumstances, or time of occurrence. Should answer, "When or where is it found?"	You get strawberries *at the store.*
Evaluative	Gives a common opinion of the target.	Cigarettes are *bad for you.*
Categorical	Category, category coordinate, or exemplar of the target. Should answer, "What kind of thing is it? What is it similar to? or "What is an example?"	A kangaroo is an *animal.*

correct naming responses than their peers [$F(1, 28) = 12.00, p = .002, \eta^2 = .20$]. (See Figure 1.) The effect of the covariates was not significant [$F(2, 28) = 1.24, p = .31$]. For both groups, the majority of errors were semantically related to their targets. Indeterminate responses were the next most frequent error type. In the SLI group, 11 individuals exhibited semantic errors as their most frequent error type, and 5 responded more frequently with indeterminate responses. In the ND group, 14 children exhibited semantic errors as their most frequent error type, and 2 responded more frequently with indeterminate responses. All participants made some semantic errors. Fourteen of the SLI participants and 10 of the ND participants produced indeterminate errors. These were largely "don't know" responses.

If we are correct in thinking that naming performance in children reflects their level of language development

Figure 1. Mean number of responses by group and type on a 20-item naming task.



in general and the integrity of their semantic systems in particular, we should find that language performance measures predict naming performance. We tested this prediction using a multiple regression procedure. Entered into the multiple regression equation were all measures of language performance (standard score on the omnibus language test, PPVT-III and EVT raw scores, conversational and narrative measures, and NWRT percentages) as well as two non-language variables (years of maternal education and nonverbal IQ). The dependent variable was the number of items correct on the naming tasks. Consistent with prediction, two factors (the EVT raw scores and the omnibus standard score) were significant predictors of naming performance (EVT: $R^2 = .63$, $p < .001$; omnibus test: $R^2 = .10$, $p = .002$). Together these scores, one that reflects lexical semantic knowledge and processing and the other that reflects general language development, accounted for 73% of the variance in naming performance.

Drawing

The mean and range of ratings for drawings from the semantic error, indeterminate error, and correct naming response pools are given in Table 3. For all pools, drawing revealed a continuum of semantic knowledge. To test the prediction that drawings of items from the correct naming pool reflect richer semantic knowledge than drawings of items from either the semantic error or indeterminate error pools, a repeated-measures ANOVA with naming pool as the within-subjects factor and mean drawing rating as the dependent variable was conducted. In cases of missing data, group means for the relevant pool were substituted. The result was a

Table 3. The grand mean rating of drawings by group and naming response pool and the proportion of drawings that were rated as low (1–2), moderate (2.01–5.99), or high (6–7) in accuracy.

		Naming response pool					
		Semantic error		Indeterminate error		Correct	
Group		*M*	*SD*	*M*	*SD*	*M*	*SD*
SLI	grand mean	2.03	(0.85)	2.61	(0.91)	3.75	(1.16)
	low	0.68	(0.21)	0.53	(0.27)	0.16	(0.20)
	mod	0.28	(0.18)	0.43	(0.26)	0.72	(0.24)
	high	0.03	(0.09)	0.04	(0.09)	0.11	(0.21)
ND	grand mean	3.08	(1.52)	3.17	(1.77)	4.54	(1.47)
	low	0.47	(0.30)	0.36	(0.37)	0.16	(0.19)
	mod	0.34	(0.25)	0.50	(0.32)	0.56	(0.25)
	high	0.19	(0.26)	0.14	(0.25)	0.29	(0.33)

significant effect for naming pool [$F(2, 62) = 38.74$, $p < .001$, $\eta^2 = .56$]. Post hoc application of Tukey's HSD demonstrated that, compared with the correct naming pool, both the semantic error pool and the indeterminate error pool elicited significantly lower drawing ratings (*p*s < .001). Examples of drawings are given in Appendix B.

For each participant, two drawing difference scores were calculated: one to compare the semantic error pool with the correct naming pool and the other to compare the indeterminate error pool with the correct naming pool. Because the comparisons were within-subject, poor overall drawing ability was minimized as a confound. If we were correct in hypothesizing that errors are associated with more limited semantic representations (and hence poorer drawings) than are correct names, then each participant's difference score should be positive.

For the first difference score, the mean drawing rating (averaged across raters) for items in the semantic error pool was subtracted from the mean drawing rating for items in the correct pool for each participant. The mean difference score for the SLI cohort was 1.72 (*SD* = .84). Scores ranged from .78 to 3.75. For the ND cohort, the mean drawing difference score was 1.46 (*SD* = .93), and the range was .39 to 2.72.

The second drawing difference score was calculated by subtracting the mean drawing rating for items in the indeterminate error pool from the mean drawing rating for items in the correct naming pool. For the SLI group, the mean difference score was 1.13 (*SD* = 1.06), and the range was –.44 to 3.23. All difference scores, except one, were positive. For the ND group, the mean difference score was 1.48 (*SD* = 1.19), with a range of .20 to 3.54. Children who presented with minimal or negative difference scores were typically those who demonstrated floor- or ceiling-level naming performances.

Post Hoc Measures of Visual Complexity

Differential visual complexity between pools could confound our interpretations of the data. To rule out this confound, we conducted post hoc analyses using the visual complexity ratings provided by Snodgrass and Vanderwart (1980). To ensure that naming accuracy was not affected by visual complexity, the 20 stimulus objects were divided by median-split into high and low visual-complexity groups. Analysis across participants revealed that they were no more likely to make errors on high-visual-complexity items than on low-visual-complexity items [Semantic Errors: $t(31) = .34$, $p = .74$, two-tailed; Indeterminate Errors: $t(23) = 0$, $p = 1$, two-tailed]. Their correct-naming responses as well were not influenced by visual complexity [$t(31) = 1.05$, $p = .30$, two-tailed]. Furthermore, there was no correlation between visual-complexity norms for each item and the drawing ratings obtained by the participants on these items ($r = .12$, $p < .613$). Because neither naming nor drawing performance was related to visual complexity of targets, we were more confident that the poor naming and drawings of items from the semantic error and indeterminate error pools reflected limitations at the level of semantic representation. The definition results yield further support for this conclusion.

Defining

The results of the defining task appear in Table 4. As in the drawing data, accurate definitions for all pools represented a range of knowledge. To test the prediction that definitions of items from the correct naming pool reflect richer semantic knowledge than definitions of items from either the semantic error or indeterminate error pools, a repeated-measures ANOVA with naming pool as the within-subjects factor and mean number of information units as the dependent variable was

Table 4. The grand mean number of accurate information units in definitions by group and naming response pool and the proportion of definitions that included low (0), moderate (1–2), or high (3+) numbers of accurate information units.

		Naming response pool					
		Semantic error		Indeterminate error		Correct	
Group		*M*	*SD*	*M*	*SD*	*M*	*SD*
SLI	grand mean	0.86	(0.43)	0.73	(0.51)	1.97	(0.82)
	low	0.49	(0.21)	0.52	(0.25)	0.04	(0.13)
	mod	0.45	(0.20)	0.44	(0.26)	0.74	(0.28)
	high	0.06	(0.12)	0.04	(0.09)	0.22	(0.28)
ND	grand mean	1.29	(0.84)	1.69	(1.23)	2.39	(0.80)
	low	0.40	(0.27)	0.31	(0.34)	0.02	(0.04)
	mod	0.39	(0.21)	0.38	(0.32)	0.58	(0.31)
	high	0.21	(0.27)	0.31	(0.42)	0.40	(0.31)

conducted. In cases of missing data, group means for the relevant pool were substituted. The result was a significant effect for naming pool [$F(2, 62) = 47.61, p < .001, \eta^2 = .61$]. Post hoc application of Tukey's HSD demonstrated that, compared to the correct naming pool, both the semantic error pool and the indeterminate error pool elicited significantly fewer information units (*p*s < .001).

Two within-subject difference scores were calculated to compare definition content for the two error pools relative to the correct naming pool. First, for each participant, the mean number of accurate information units for items in the semantic error pool was subtracted from the mean number of accurate information units for items in the correct pool. For the SLI cohort, the mean definition difference score was 1.11 (*SD* = .75), and the range was 0 to 3.11. For the ND cohort, the mean was 1.10 (*SD* = .44), and the range was .5 to 2.18. The one participant with SLI who demonstrated no difference between definitions of items from the semantic error and correct naming pools also obtained a minimal difference score in the drawing task.

A definition difference score was also calculated for the indeterminate error pool—again with reference to the correct naming pool. For the SLI group, the mean difference score was 1.38 (*SD* = .60), with a range of 0 to 2.69. One child had a difference score of zero. For the ND group, the mean difference score was .64 (*SD* = .23), with a range of –.31 to 1.65. Three children had difference scores that were zero or negative.

The types of accurate information units that occurred in the participants' definitions appear in Figures 2 and 3. Functional and physical properties were the most frequently provided information units, whereas locative, evaluative, and categorical were the least frequent. The profiles of types included in definitions of items from correct, semantic error, and indeterminate

Figure 2. Proportion of accurate definition units by type and naming response pool for the SLI group.

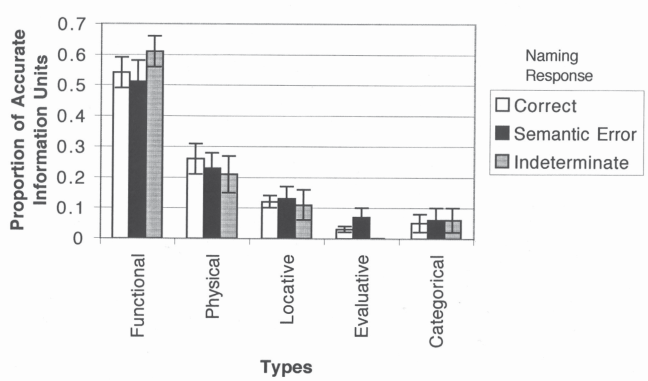
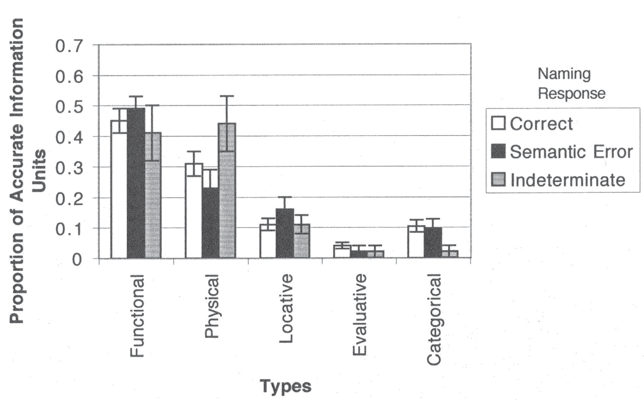


Figure 3. Proportion of accurate definition units by type and naming response pool for the ND group.



error pools were similar, suggesting that representations associated with misnaming, though quantitatively sparser, were not qualitatively different from those associated with correct naming. The two participant cohorts were highly similar in profile of types, suggesting that the children with SLI were not atypical in the types of semantic information they had mapped. The only apparent exception to these conclusions is the high proportion of physical features that the ND group included in definitions of items from the indeterminate error pool.

Post Hoc Analysis of Definition and Naming Mismatch for the Semantic Error Pool

The definitions provided an additional opportunity for examining the semantic representations associated with semantic misnaming. When a child makes a semantic error such as *nail* for *screw*, it could be that he or she has represented the common features of these related objects, but not enough of their unique features to differentiate the two. We explored this possibility by determining the match or mismatch between the children's naming responses and their definitions for each item in the semantic error pool. For example, if a child who produced the error *nail* for *screw* defined a screw as "something you *twist* into wood," her definition was considered a mismatch; it included a feature that characterized the target concept but not the concept that was substituted in the naming probe. If a child with the same naming error defined a screw as "something that you build with," his definition was considered a match; it matched both the target concept and the substituted concept. The latter instances would constitute evidence of sparse semantic representations that do not allow distinctions between close semantic coordinates.

Definition mismatch was also coded independently by a second coder for 4 data sets selected randomly from the SLI group and 4 data sets selected randomly from the ND group. Agreement between coders ranged from 67% to 100%, with a mean of 89%. Disagreements were resolved by consensus.

For the SLI cohort's semantic error pool, 26% (SD = 17) of definitions and substitution errors were matches and 25% (SD = 18) were mismatches. (Recall from Table 4 that 49% of definitions for the semantic error pool contained no accurate information units and therefore could not be classified as matches or mismatches.) For the ND cohort, 18% (SD = 15) of definitions and substitution errors matched, 42% (SD = 25) mismatched, and 40% could not be classified because they lacked accurate information. These results suggest that some sparse representations do not permit distinctions between semantic coordinates. Further, these results, like the previous results, document that gradations of semantic knowledge are associated with semantic errors.

Post Hoc Analysis of Definition and Drawing Concordance

To discern the level of intertask agreement for specific targets, we conducted a Pearson product-moment Q correlation between drawing ratings and definition scores for each participant, treating items as units. Correlations ranged from $r = .05$ to $r = .74$. Twenty-two children had significant ($p < .05$) r values of .44 or above. Those 10 participants who did not demonstrate significant correlations were typically those who presented with near-ceiling or floor-level performance on one of the tasks. We concluded that, for the majority of participants, the definition and drawing tasks led to the same results for specific targets.

Post Hoc Analysis of Comprehension

A primary finding reported above is that semantic and indeterminate naming errors in children with SLI are associated with a (negatively skewed) continuum of semantic knowledge. We had hypothesized that naming errors are frequently related to limited storage of semantic information. Our results support this hypothesis. The low and moderate portions of the continuum suggest the possibility of two types of limitations: missing representations (no knowledge) and sparse representations (minimal knowledge).

To determine to what extent missing versus sparse semantic representations contributed to the semantic naming errors of children with SLI, we asked the last recruited members of the SLI cohort (n = 11) to participate in an additional task. This task, which is traditionally included in word-retrieval assessments, involved recognition of pictured targets within an array of foils. Because forced-choice recognition requires only a minimal semantic representation for successful performance (Carey, 1978; Dollaghan, 1987), we reasoned that correct performance constituted evidence of at least a sparse semantic representation; whereas incorrect performance constituted evidence of a missing representation.

In the task itself, each child was presented with the target and five foils in random order on a single page (see Appendix A). Two of the five foils were chosen from the pool of semantic errors produced by the other participants in this project with the intent of maximizing the possibility of tapping semantic confusions that were relevant to children. As the examiner named each target, the child pointed to indicate his or her answer. This task followed the naming, drawing, and defining tasks.

On the recognition task, the participating subgroup correctly identified 58% (SD = 16) of items from the semantic error pool, 53% (SD = 16) of items from the indeterminate error pool, and 90% (SD = 18) of items from the correct naming pool. The difference between performance on these error pools was significant [$F(2, 20) = 15.83$, $p < .001$, $\eta^2 = .60$]. A post hoc application of Tukey's HSD revealed that there were differences between the correct and semantic error pools ($p < .001$) and between the correct and indeterminate error pools ($p < .001$). Although performances on the semantic and indeterminate error pools were poor, they were better than chance performance of 17% ($p = .056$ [approached significance] and $p = .03$, respectively). Given that 10% of items from the correct pool were misidentified and that children must have had some knowledge of items that they named correctly, we took 10% to be an estimate of error in this task. Therefore, we concluded that approximately 32% of items in the semantic error pool (100% – 58% – 10% = 32%) and 37% of items in the indeterminate error pool (100% – 53% – 10% = 37%) were associated with missing representations.

To further tease apart sparse from missing representations in cases of semantic or indeterminate misnaming, the responses to the recognition task were compared with the responses to the drawing and defining tasks on an item-by-item basis. Whereas our strategy in the planned experimental analysis was to determine the amount of semantic knowledge elicited within each task, in this post hoc analysis we compared between tasks. Our strategy was to determine the number of task modalities that elicited at least minimal knowledge (operationalized as correct picture identification on the recognition task, at least one correct information unit in the definition task, and a drawing rating higher than 2.0) for each of the misnamed targets. The results, along with their interpretations, appear in Table 5. The largest percentage of the semantic error pool (45%) was associated with sparse semantic knowledge of the target—that is, correct performance in only one or two tasks but not all three. For the indeterminate pool, 33% of items were associated with sparse semantic knowledge. Twenty-nine percent of the semantic error pool and 33% of the indeterminate error pool elicited no knowledge in any task modality, suggesting missing representations. These figures are roughly comparable to the 32% and 37% estimates of missing representations based on within-task performance on the recognition probe.

Although the majority of naming errors were best attributed to limited storage of either the sparse or missing variety, 26% of semantic errors and 34% of indeterminate errors were produced despite receptive knowledge demonstrated in all three task modalities. Recall that the within-task comparisons revealed high levels of semantic knowledge for 3% of drawings and 6% of definitions of items that the SLI group named with a semantic error and for 4% of drawings and definitions of items they named with an indeterminate error. Clearly some minority of naming errors represents an exception to our hypothesis; these errors occur despite robust semantic knowledge of the target.

Although not displayed in Table 5, a between-task comparison was also conducted for the correctly named items. Of these, 79% (*SD* = 15%) were associated with correct performance on all three post-naming tasks, whereas 16% (*SD* = 17%) were associated with correct performance on two of the three tasks. Correct naming responses associated with correct post-naming performance on only one task or on no task were negligible. This result is additional evidence of the link between robust semantic storage and successful retrieval.

Table 5. Task results compared.

Descriptive outcome	Numerical outcome: Semantic	Numerical outcome: Indeterminate	Conclusion	Semantic storage plays role in misnaming?
Correct response on all three tasks	*M* = 26%, *SD* = 21	34%, 23	Adequate semantic representation	No
Correct response on one or two tasks	*M* = 45%, *SD* = 24	33%, 19	Sparse semantic representation	Yes. Semantic storage is inadequate to support naming.
Correct response on no task	*M* = 29%, *SD* = 19	33%, 18	Missing representation at one or more levels of the lexicon	Perhaps. Semantic and/or phonological storage may be inadequate to support naming.

Note. This table addresses the following: Given a set of age-appropriate items that elicited semantic naming errors or indeterminate naming errors from children with SLI ($n = 11$), what percentage will elicit a correct response in recognition, drawing, and defining tasks? What conclusions can be drawn about semantic storage and its role in the misnaming of these items? Outcomes were calculated on an item-by-item basis per child and then averaged across children.

Discussion

In this paper we replicated previous research (e.g., Fried-Oken, 1984; Leonard et al., 1983; McGregor, 1997) by demonstrating that children with SLI make more errors when naming than do their age-mates. Also, we documented that for both SLI and ND cohorts, semantic errors and indeterminate errors are the two most common error types. This has been reported previously by Dapretto and Bjork (2000), Lahey and Edwards (1999), McGregor (1997), Wiegel-Crump and Dennis (1986), and Wijnen (1992).

We hypothesized a link between the name-retrieval difficulties of children with SLI and their poor language development. In particular, we predicted that naming errors would relate to impoverished semantic representations. In support of these predictions, we found the children's raw scores on the EVT and their standard scores on either the TOLD-P or the WLPB-R to be significant positive predictors of naming performance. The former test requires the child to provide synonyms in response to pictures and spoken labels provided by the examiner. As such, it measures the integrity of semantic representations and processing. The latter tests are omnibus batteries that measure knowledge and performance in a variety of language domains. The finding that measures of both semantics and overall language development predict naming performance is consistent with the tenets of hypothesized relations between semantic memory and information retrieval as proposed by Kail and Leonard (1986) and Bjorklund (1987).

The Association Between Semantic Representation and Naming Performance

This paper extends our understanding of naming performance in children with SLI. As hypothesized, we found that most naming errors are associated with limited semantic representation. Our initial evidence was collected in a comparative picture-naming/picture-drawing paradigm. This paradigm was developed on the evidence-based assumption that naming and drawing tap the same underlying semantic representations via different input and output modalities (McGregor & Appel, 2002; McGregor et al., 2002). If performance on both is poor, deficiency at the level of semantic representation is the most parsimonious conclusion. If performance is poor on one task but not the other, the deficiency likely lies at some level of input or output representation or processing. When comparing within-subject drawing performance for correctly named and misnamed items, we found the naming errors of SLI and ND groups to be associated with significantly poorer drawings. The majority of the drawings of misnamed objects were low to moderate in accuracy, whereas the majority of drawings of correctly named objects were moderate to high in accuracy. Converging evidence resulted from a definition task in which children in SLI and ND cohorts provided significantly fewer accurate information units for misnamed than correctly named targets. Again, low to moderate and moderate to high amounts of information characterized items from the error pools and the correct naming pool, respectively, for the SLI participants. Because of these significantly different but overlapping distributions, we conclude that limited semantic representations are often, but not always, associated with naming errors.

Finally, for a subset of the children with SLI, recognition of the stimulus items was also assessed. This allowed us to determine the extent to which limited semantic knowledge was a matter of missing or sparse representations. Approximately one third of semantic errors and indeterminate responses were associated with incorrect responses on the recognition probe. When performance per item was compared across tasks (drawing, defining, recognition), again approximately one third of the items from the error pools were associated with incorrect responses in all task modalities. We interpret the one-third figure as a rough estimate of missing representations. In addition, we estimate that roughly half of semantic misnamings and one third of indeterminate responses were associated with sparse semantic representations. This estimate was based on the percentage of items from the error pools that elicited correct performance on only one or two tasks. Therefore, on the basis of within-child comparisons conducted within tasks, between tasks, and between error pools, we conclude that (a) children with SLI have limited semantic knowledge of some age-appropriate words, (b) this limitation affects naming performance, (c) both semantic and indeterminate errors are often associated with limited semantic representation, and (d) this limitation involves both missing and sparse representations. A discussion of the nature of missing and sparse representations follows.

Missing Representations

Some cases of misnaming were associated with missing representations. Given that small vocabularies are characteristic of SLI (e.g., Bishop, 1997), it is not surprising that our participants did not always know age-appropriate object labels. Even errors that bear semantic relations to their targets do not necessarily imply knowledge. The visual stimuli presented during naming may activate semantically and visually similar objects, allowing participants to unconsciously or strategically use the picture cues to name a semantic coordinate of the target. Put simply, pictures afford children a chance to make a good guess.

It should also be noted that the participants' lack of knowledge may not necessarily involve the semantic level. The children may have had semantic representations in some cases but may not have linked those representations to retrievable phonological input and output forms. Because phonological representations were not a focus of this study, we cannot determine the extent to which missing semantic representations versus deficits at other levels of the lexicon played a role in these errors.

Sparse Semantic Representations

In other cases of misnaming, children demonstrated semantic knowledge of a target, but that knowledge was sparse. The nature of these sparse representations is best revealed by comparing the types of information included in definitions of correct and erroneous items. Whether defining items from the semantic error pool, the indeterminate error pool, or the correct pool, children from both SLI and ND cohorts most often mentioned functional and physical object properties. An emphasis on functional and physical properties is a long-noted characteristic of early definitions (Anglin, 1978; Nelson, 1985; Watson & Olson, 1987). Furthermore, infants and toddlers depend upon physical similarity, especially similarity of shape, to extend labels to new objects (Baldwin, 1992; Imai, Gentner, & Uchida, 1994); and they depend upon functional features, especially the role of objects in events, to guide their inferences about conceptual categories (Mandler, 2000). The focus on functional and physical properties in definitions given by both SLI and ND children is continuous with the infant's focus on shape and function during establishment of early semantic representations. Therefore, the sparse semantic representations demonstrated by children with SLI are consistent with their overall developmental delay in language and are best characterized as a problem of limited, but not unusual, information.

Finally, a post hoc comparison of semantic substitution errors and definition content provided additional evidence as to the nature of some sparse semantic representations. For the SLI cohort, roughly one quarter of definitions applied equally well to the target concept or the substituted concept. These may be instances where there is insufficient information represented in semantic memory to allow distinctions between close semantic coordinates.

Robust Semantic Representations

In some cases of naming error, the participants went on to perform well on tasks of drawing, defining, and recognition. Depending on error pool and task, between 3% and 6% of drawings and definitions of misnamed items included high levels of accurate information. Roughly one quarter and one third of items from the semantic and indeterminate error pools, respectively, elicited some correct information on all three post-naming tasks: drawing, defining, and recognition. Therefore, it is unlikely that sparse or missing semantic representations played a role in all errors. The possible sources of these errors are numerous. The most likely would be misperception of the picture stimuli during naming or difficulty finding the correct phonological form because of storage or retrieval breakdown (or their interaction) at the level of the phonological output representation.

Because the lexicon is a memorial system that must be developed entry by entry, it is a certainty that the specific percentages of missing, sparse, and robust representations associated with naming errors will not replicate when a new population sample is tested with new stimuli. The important conclusion is that, given age-appropriate object stimuli, children with SLI will make many naming errors—at least in part—because they do not have well-developed semantic representations for these words. From word to word, this limitation is manifested as either a gap in lexical knowledge or as sparse knowledge. Although the former relationship is obvious (children cannot name objects they do not know), the latter leads to an important conclusion: The *degree* of knowledge represented in the semantic lexicon makes words more or less vulnerable to retrieval failure.

Comparison of Semantic and Indeterminate Naming Errors

Together, semantic and indeterminate errors constituted approximately 90% of all errors produced by the SLI and ND cohorts during the naming task. The post-naming tasks reveal that limited semantic knowledge contributes to both of these error types. However, results on these tasks also suggest some differences between semantic and indeterminate errors. Compared to semantic errors, indeterminate errors were associated with less negatively skewed semantic knowledge continua. This was true of the SLI and ND cohorts for the drawing task and the ND cohort for the definition task. At first this may seem counterintuitive. Because most indeterminate errors were "I don't know" responses, it seems that they should be associated with semantic knowledge continua that are quite negatively skewed. However, our results suggest that children answer "I don't know" in some cases where they do know but are unsure, cautious, forgetful, or tired, among other possibilities. This is not to say that limited semantic knowledge was not a source of indeterminate errors, but rather that the sources of such errors may be more varied than the sources of semantic naming errors.

There was also an apparent between-group difference in indeterminate errors. Compared with the SLI group, the ND group included a larger proportion of physical properties in their definitions of items from the indeterminate error pool. Perhaps the sources of indeterminate errors differ somewhat for the two participant groups. This conclusion is based solely on our post hoc explorations of the data patterns and, as such, requires replication and further study.

Future Directions

This study does not address the causes of the limited lexical semantic knowledge characteristic of SLI; however, we suspect that working memory may be a fruitful place to begin the search. Recall that, as a selection criterion, all of the participants in the SLI cohort had working memory deficits as measured by the Nonword Repetition Task (Dollaghan & Campbell, 1998). Working memory deficits are now a well-documented characteristic of SLI (Dollaghan & Campbell, 1998; Ellis Weismer, Evans, & Hesketh, 1999; Gathercole & Baddeley, 1990; Montgomery, 1995). Although the term *working memory* is typically used to describe the memorial system responsible for short-term processing and storage of information, working memory is thought to play an important role in a range of complex cognitive activities, including successful long-term learning and memory. Gathercole and Baddeley (1989) reported nonword repetition to be correlated with vocabulary size in normally developing 4-, 6-, and 8-year-olds. Furthermore, at 4 years, nonword repetition is predictive of later vocabulary size (Gathercole, Willis, Emslie, & Baddeley, 1992). Baddeley, Gathercole and colleagues account for these relationships by positing that new words must be held in the phonological store of the articulatory loop before representations can be established in the long-term store. They hypothesize that a high capacity or highly efficient phonological store facilitates long-term representation by influencing the quantity and quality of the phonological information retained and the decay rate of that information (Baddeley, Papagno, & Vallar, 1988; Gathercole & Baddeley, 1990).

To date, research on working memory and SLI has focused on the phonological store of the articulatory loop. It is not directly evident how deficits in memory for phonological forms relate to the limitations in storage of semantic knowledge documented in the current study. However, Storkel (2001) presents experimental support for just such a relationship. Storkel manipulated the phonotactic frequency of novel words such that some were composed of common sound sequences and others of rare sound sequences. She then measured the impact of phonotactic frequency differences on normal word learning. In one task, children were given the newly learned words and were asked to pick the correct object referents from arrays. When they made errors in response to words of common sequences, these were typically within-semantic-category, but when they made errors in response to words of rare sequences, these were typically unrelated to the semantic category of the target. Storkel concludes that high phonotactic frequency facilitates the establishment of semantic representations.

Storkel's data are suggestive. When mapping a new phonological form is difficult, either because of its rare phonotactic sequence or, in the case of SLI, because of a phonological memory deficit, perhaps few resources remain for semantic mapping; and thus semantic representations build slowly, if at all. The relationship between phonological working memory and the semantic lexicon in children with SLI merits study. Alternatively, the semantic representation problem associated with SLI may relate to deficits in other working memory components. Given that physical features of objects and their function in the (visual) environment play an important role in the establishment of semantic representations, research on the visual spatial sketch-pad component of working memory in children with SLI may be useful.

Clinical Implications

Several clinical implications follow from this work. First, these data suggest the importance of assessing not only the number of words that children know and use, but also the robustness of their word knowledge (but see Vermeer, 2001). The results of fast-mapping paradigms demonstrate that children require only minimal representations to succeed at forced-choice recognition probes (Carey, 1978; Dollaghan, 1987), the format typical of receptive vocabulary tests. Because lexical semantic storage is a matter of degree, children may know many words well enough to score within normal limits on receptive vocabulary tests but still not know them well enough to retrieve and use them consistently. Indeed, the SLI cohort in this study demonstrated a mean standard score of 89 on the PPVT-III but went on to demonstrate limited semantic knowledge on a substantial proportion of the age-appropriate words we examined experimentally.

Second, it is misleading to describe the word-retrieval deficits of children with SLI as "not being able to find a word even though they know it." In this study, only a minority of semantic naming errors occurred in association with robust semantic knowledge. The phenomenon we describe clinically as a *word retrieval deficit* is more complex and varied than the term implies. It is not as cleanly dissociated from word knowledge as was previously thought.

Finally, the link between semantic knowledge and naming performance suggests a possible intervention strategy. Increasing the robustness of semantic representations should yield improvements in retrieval. This strategy has met some success in previous intervention studies (McGregor & Leonard, 1989; Wright, 1993). With further understanding of the nature of the semantic storage limitation, more refined interventions should be tested.

Acknowledgments

We thank the children who participated, their families, and their speech-language pathologists. We are indebted to the faculty of Walker, Dewey, Greenbriar, and Westgate Schools for their cooperation. Beth Schiff was particularly helpful with subject recruitment, and Rosie Carr and Anne Graham provided invaluable assistance with data analysis. Five faithful Northwestern graduate students served as raters. Credit is due to an anonymous reviewer for suggesting the definition-naming mismatch analysis. Finally, we gratefully acknowledge the National Institute on Deafness and Other Communication Disorders (award #R29 DC 03698) for support of the first author.

References

Anglin, J. (1978). From reference to meaning. *Child Development, 49,* 969–976.

American Speech-Language-Hearing Association. (1985). Guidelines for identification audiometry. *Asha, 27,* 49–52.

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27,* 586–595.

Baldwin, D. A. (1992). Clarifying the role of shape in children's taxonomic assumption.. *Journal of Experimental Child Psychology, 54,* 392–416.

Bishop, D., Bishop, S., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: Evidence from a twin study. *Journal of Speech, Language, and Hearing Research, 42,* 155–168.

Bishop, D., North, T., & Donlan, C. (1996). Nonword repetition as a behavioral marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, 36,* 1–13.

Bishop, D. V. M. (1997). *Uncommon understanding*. East Sussex, UK: Psychology Press.

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review. *Developmental Review, 7,* 93–130.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Burgemeister, B. B., Blum, L. H., & Lorge, I. (1972). *Columbia Mental Maturity Scale*. San Antonio, TX: Harcourt Brace Jovanovich.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research, 40,* 519–525.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–297). Cambridge, MA: MIT Press.

Chi, M. T. H.,, & Koeske, R. D. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology, 19,* 29–39.

Dapretto, M., & Bjork, E. L. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development, 71,* 543–822.

Dollaghan, C. (1987). Fast mapping of normal and language-impaired children. *Journal of Speech and Hearing Disorders, 52,* 218–222.

Dollaghan, C. (1998). Spoken word recognition in children with and without specific language impairment. *Applied Psycholinguistics, 19,* 193–207.

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

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,* 1249–1260.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43,* 865–878.

Fey, M., Long, S., & Cleave, P. (1994). Reconsideration of IQ criteria in the definition of specific language impairment. In R. Watkins & M. Rice (Eds.), *Specific language impairments in children* (Communication and Language Intervention Series, Vol. 4; pp. 161–178). Baltimore: Brookes.

Fried-Oken, M. (1984). *The development of naming skills in normal and language deficient children*. Unpublished doctoral dissertation, Boston University, Boston, MA.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28,* 200–213.

Gathercole, S., & Baddeley, A. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29,* 336–360.

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

Gershkoff-Stowe, L. (2001). The course of children's naming errors in early word learning. *Journal of Cognition and Development, 2,* 131–155.

Gershkoff-Stowe, L., & Smith, L. B. (1997). A curvilinear trend in naming errors as a function of early vocabulary growth. *Cognitive Psychology, 34,* 37–71.

Hillis, A. E., & Caramazza, A. (1991). Mechanisms for accessing lexical representations for output: Evidence from a category-specific semantic deficit. *Brain and Language, 40,* 106–144.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9,* 45–75.

Kail, R., & Leonard, L. B. (1986). *Word-finding abilities in language-impaired children* (ASHA Monographs 25).

Rockville, MD: American Speech-Language-Hearing Association.

Katz, W., Curtiss, S., & Tallal, P. (1992). Rapid automatized naming and gesture by normal and language impaired children. *Brain and Language, 43,* 623–641.

Lahey, M., & Edwards, J. (1996). Why do children with specific language impairment name pictures more slowly than their peers? *Journal of Speech, Language, and Hearing Research, 39,* 1081–1098.

Lahey, M., & Edwards, J. (1999). Naming errors of children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 195–205.

Leonard, L. B. (1999). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L., Nippold, M., Kail, R., & Hale, C. (1983). Picture naming in language impaired children. *Journal of Speech and Hearing Research, 26,* 609–615.

Lindberg, M. A. (1980). Is knowledge base development a necessary and sufficient condition for memory development? *Journal of Experimental Child Psychology, 30,* 401–410.

Mandler, J. M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development, 1,* 3–36.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.

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

McGregor, K. K., & Appel, A. (2002). On the relation between mental representation and naming in a child with specific language impairment. *Clinical Linguistics and Phonetic*s*, 16,* 1–20.

McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45,* 332–346.

McGregor, K. K., & Leonard, L. B. (1989). Facilitating word-finding skills of language-impaired children. *Journal of Speech and Hearing Disorders, 54,* 141–147.

McGregor, K. K., & Waxman, S. R. (1998). Object naming at multiple hierarchical levels: A comparison of preschoolers with and without word-finding deficits. *Journal of Child Language, 25,* 419–430.

Montgomery, J. W. (1995). Examination of phonological working memory in specifically language-impaired children. *Applied Psycholinguistics, 16,* 355–378.

Nelson, K. (1985). *Making sense: The acquisition of shared meaning*. San Diego, CA: Academic Press.

Newcomer, P. L., & Hammill, D. D. (1997). *Test of Language Development–Primary 3*. Austin, TX: Pro-Ed.

Records, N. L., & Tomblin, J. B. (1994). Clinical decision making: Describing the decision rules of practicing speech-language pathologists. *Journal of Speech and Hearing Research, 37,* 144–156.

Rice, M., Buhr, J., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Disorders, 55,* 33–42.

Rice, M., Buhr, J., & Oetting, J. (1992). Specific-language-impaired children's quick incidental learning of words: The effects of a pause. *Journal of Speech and Hearing Research, 35,* 1040–1048.

Rice, M., Oetting, J., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research, 37,* 106–122.

Roth, C. (1983). Factors affecting developmental changes in the speed of processing. *Journal of Experimental Child Psychology, 35,* 509–528.

Snodgrass, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior, 23,* 3–22.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44,* 1321–1337.

Tomblin, J. B., & Zhang, X. (1999). Language patterns and etiology in children with specific language impairment. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders: Developmental cognitive neuroscience* (pp. 361–382). Cambridge, MA: MIT Press.

U.S. Census Bureau. (2001). National Demographic Profile Summary [online]. Available: fttp://www.census.gov/Press-Release/www/2001/demprousmemo.html.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22,* 217–234.

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech, Language, and Hearing Research, 28,* 1349–1355.

Watson, R., & Olson, D. R. (1987). From meaning to definition: A literate bias on the structure of word meaning. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 329–353). New York: Academic Press.

Wiegel-Crump, C. A., & Dennis, M. (1986). Development of word-finding. *Brain and Language, 27,* 1–23.

Wijnen, F. (1992). Incidental word and sound errors in young speakers. *Journal of Memory and Language, 31,* 734–755.

Williams, K. T. (1997). *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Service.

Woodcock, R. W. (1991). *Woodcock Language Proficiency Battery–Revised*. Itasca, IL: Riverside Publishing Co.

Wright, S. H. (1993). Teaching word-finding strategies to severely language-impaired children. *European Journal of Disorders of Communication, 28,* 165–175.

Received October 31, 2001

Accepted June 5, 2002

DOI: 10.1044/1092-4388(2002/081)

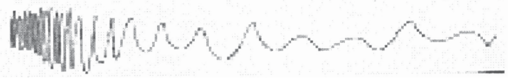
Contact author: Karla K. McGregor, PhD, 2299 N. Campus Drive, Evanston, IL 60208.
E-mail: k-mcgregor@northwestern.edu

Appendix A

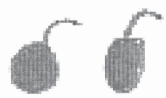
Stimuli for Naming, Drawing, Defining, and Recognition Tasks	Foils for Recognition Task				
	Sem	Sem	Phon	UR	UR
kite	flag	banner	cat	wheel	stove
axe	hammer	jackhammer	X	jar	tie
umbrella	kite	balloon	ambulance	basket	phone
vase	pot	cup	base	arrow	rope
strawberry	cherry	watermelon	straw	car	tent
mountain	hill	volcano	fountain	rake	watch
envelope	letter	postcard	cantaloupe	gloves	radish
blouse	shirt	coat	mouse	spider	lamp
caterpillar	snake	ladybug	water pitcher	chair	wrench
ladder	rail	firetruck	letter	robe	drum
cigarette	smoke	cigar	clarinet	doorknob	key
windmill	golf	pinwheel	treadmill	golfer	peapod
clothespin	paperclip	safety pin	closed pen	stoplight	bear
screw	nail	screwdriver	zoo	fishbowl	hat
iron	clothesdryer	ironing board	island	pencil	star
anchor	hook	ship	hanger	tape	flower
pumpkin	apple	pie	paper	brush	shelves
kangaroo	piglet	reindeer	camera	slippers	lock
pitcher	vase	cup	picture	newspaper	belt
guitar	violin	song	target	measuring cup	fork

Sem = Semantically Related Foil, Phon = Phonologically Related Foil, UR = Unrelated Foil.

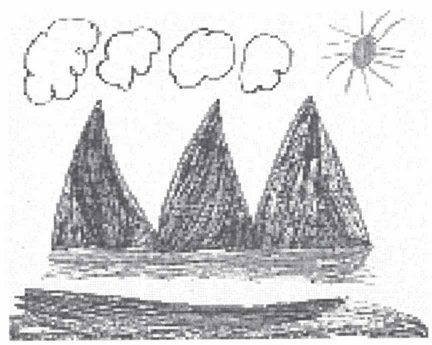
Appendix B. Example drawings from a single child with SLI.



"Pitcher" that was rated as low in accuracy (named with semantic error "milk")



"Strawberry" that was rated as moderate in accuracy (named correctly)



"Mountain" that was rated as high in accuracy (named correctly)