



Published in final edited form as:

*Appl Psycholinguist*. 2010 ; 31(03): 463–487. doi:10.1017/S014271641000007X.

## What compound words mean to children with specific language impairment

KARLA K. MCGREGOR, GWYNETH C. ROST, LING YU GUO, and LI SHENG

University of Iowa

### Abstract

Sixteen children (17 age mates, 17 vocabulary mates) with specific language impairment (SLI) participated in two studies. In the first, they named fantasy objects. All groups coined novel noun–noun compounds on a majority of trials but only the SLI group had difficulty ordering the nouns as dictated by semantic context. In the second study, the children described the meaning of conventional noun–noun compounds. The SLI and AM groups did not differ in parsing the nouns, but the SLI group was poorer at explaining the semantic relationships between them. Compared to vocabulary mates, a larger proportion of the SLI group successfully parsed the compounds but a smaller proportion could explain them. These difficulties may reflect problems in the development of links within the semantic lexicon.

---

A categorization of *Webster's Unabridged Dictionary* reveals compounds to be the most common words in the English language. In a sample of 2,076 words, 30% were compounds, 28% were derivations, 24% were single roots, 8% were proper names, and 10% were other word types (Goulden, Nation, & Read, 1990). Of all compounds, those formed from two root nouns (NN) are especially numerous (Carr, 1959). Given their ubiquity, the speaker of English would be unable to communicate precisely and effectively without command of compounds.

In NN compounds, the first root functions as the modifier, the second as the head. The semantic relationship between the modifier and head nouns is one of categorization and specification. The head noun categorizes the referent and the modifier noun specifies the subtype, serving to differentiate it from other members of the category. Thus, for example, a cherry tree is a tree that bears cherries, not apples or pears. The correct syntax of modifier and head in a NN compound is determined semantically. Consider compound pairs like *bumper car* and *car bumper*, *cake pan* and *pan cake*, and *dog show* and *show dog*. Either syntactic order can be correct given an intended meaning.

Children with specific language impairment (SLI) have difficulty ordering NN compounds (Grela, Snyder, & Hiramatsu, 2005). We hypothesize that this difficulty is at least partially semantic in nature; more precisely, these children misorder compounds because they fail to appreciate the semantic relationship between the modifier and the head. In this paper we report comparisons of children with SLI and their unaffected age mates (AMs) and vocabulary mates (VMs) in two studies designed to test this hypothesis. In Study 1, the children coined novel compounds given a semantic context that motivated the order of the modifier and head. In Study 2, these children explained the modifier–head relationships in conventional compounds.

## THE ACQUISITION OF COMPOUNDS

Compounds are represented in the vocabularies of children as young as 1 and 2 years of age. Although the earliest uses are likely rote, children begin to parse compounds at a very young age. Two sources of evidence include children's overt comments on word structure and their combining of roots to coin novel compounds, as seen in the examples from Child D below.

"D(2;4,3 [years; months, days] looking at a toy car): That a motor-car. It got a motor." (Clark, 1993, p. 40)

"D(2;7,1): It's a water-cake ... I made it in the water." (Clark, 1993, p. 40)

One factor influencing the child's analysis of any given compound is the size of the compound family. Modifiers and heads from large compound families promote analysis because the pattern of their combination is highly salient. For example, the modifier and head of the compound *chocolate cake* are from high-frequency families; many compounds include the modifier *chocolate* (e.g., *chocolate + bar, pudding, pie, fudge, candy, milk*) and many include the head *cake* (e.g., *lemon, coconut, pound, layer, cup, griddle, angel food, + cake*). This may well relate to the early coining of "water cake" in the example above. When 3- to 5-year-olds were asked to explain the meaning of NN compounds (i.e., why do we say *chocolate cake*?), they were more likely to mention a root from a large family than from a small family, suggesting that high-frequency patterns promote the child's parsing and analysis via analogy (Krott & Nicoladis, 2005; Nicoladis & Krott, 2007).

As children become increasingly able to analyze and manipulate modifier and head roots, knowledge of the semantic relationship between these roots begins to emerge (Berko, 1958; Nicoladis, 2002). When asked to match compound words to their pictured referents, preschoolers correctly interpret head nouns as indicative of the general category; that is, they select a picture of a knife, not an apple, in response to *apple knife* (Clark, Gelman, & Lane, 1985). When asked to explain the meaning of compounds, preschoolers are more likely to mention the modifier than the head root. This is true both for young speakers of English, where the head appears at the right edge of the compounds (Krott & Nicoladis, 2005) and French, where it appears at the left (Nicoladis & Krott, 2007). Nicoladis and Krott take this as evidence that preschoolers already have some understanding that compounds refer to subcategories and that it is therefore more important to mention the modifier, which specifies the subcategory, than the head, which indicates the general category.

This understanding is more advanced among 4-year-olds than 3-year-olds. In a forced-choice recognition task in which the children heard novel compounds (e.g., *fish shoe*) and then selected the target referent (a shoe decorated with fish) from among three foils (a fish, a shoe, and a shoe surrounded by fish), both 3- and 4-year-olds selected the target on a majority of trials. However, 3-year-olds were more likely than 4-year-olds to select the object named by the head noun (i.e., the shoe), suggesting that they did not yet completely appreciate that NN compounds specify the subtype, not the general type, of the referent (Nicoladis, 2002).

Just as children are better able to parse compounds composed of frequent than infrequent modifiers and heads, they may have a better understanding of modifier-head relationships that are frequent in the language. When faced with determining the semantic relationship within a novel compound (e.g., what does *pepper bread* mean; what does *cheese fish* mean?), 4- and 5-year-olds used their knowledge of semantic relationships within real compounds that shared heads with the novel compounds (Krott, Gagné, & Nicoladis, 2009). The influence of the head was greater for large families (e.g., *bread* is a frequent head) than for small families (e.g., *fish* is an infrequent head). For frequent heads, children tended to interpret the novel compound according to the dominant (most consistent) modifier-head

relationship associated with the head noun. Thus, via analogy to *cinnamon bread*, for example, most of the children interpreted *pepper bread* as bread that has pepper in it (Krott et al., 2009). Frequency continues to affect comprehension of modifier–head relationships in adults but, unlike children, adults are more influenced by modifier than head frequency (Krott et al., 2009).

## ERRORS IN COMPOUND WORD FORMATION AMONG CHILDREN WITH SLI

Despite being ubiquitous in the language and early in development, compounds constitute a challenge for children with SLI. Most extant evidence concerns their ability to obey morphological constraints on the internal structure of compounds. In English, there is a constraint against pluralizing the modifier of a NN compound (e.g., *rat-eater*, but not *rats-eater*, is allowed). School-aged English speakers with SLI obey this constraint less consistently than their unaffected AMs (Oetting & Rice, 1993; Van der Lely & Christian, 2000) and less consistently than their younger language mates in some cases (Van der Lely & Christian, 2000), but not others (Oetting & Rice, 1993). Difficulties with morphophonological aspects of compounding have also been reported for children with SLI who speak Greek (Dalalakis, 1999; Kehayia, 1997) and Japanese (Fukuda & Fukuda, 1999).

Data on syntactic and semantic aspects of compounding are more limited. One exception is the work of Grela and colleagues (2005). They asked English speakers with SLI (ages 4 years, 8 months [4;8] to 7;0) to label fantasy objects such as a car shaped like a shoe. They recognized three potential challenges for the children with SLI. Of most interest here, they examined adherence to a syntactic constraint, namely, that the modifier is ordered before the head (e.g., a hat made of rocks is a *rock hat*, not a *hat rock*). They also examined adherence to a semantic constraint on the nature of the modifier. For example modifiers specifying material are fine (e.g., *rock hat*) but modifiers specifying quantity are not (e.g., if it is a hat for some people one cannot call it a *some hat*). Finally, like previous investigators, they examined the children’s adherence to a morphological constraint barring plurals from the modifier (e.g., a hat made of rocks is a *rock hat* not a *rocks hat*).

The SLI group coined as many NN compounds as a normal comparison group comprising peers who were an average of 8 months younger. As in Oetting and Rice (1993), the children with SLI were as capable as their younger peers in obeying the morphological constraint against plurals within compounds. They were also aware that some meanings cannot be expressed via compounding. They tended not to coin compounds in contexts that expressed quantity, thus obeying that semantic constraint. However, the children with SLI were more prone to misordering of the modifier–head relationship, producing for example, *car shoe* instead of *shoe car*. The children did seem to have some knowledge of the correct order as they performed significantly above chance. Instead, their application of this knowledge was vulnerable and perhaps especially vulnerable given the task demands. Specifically, the nature of the task instructions was such that the examiner always presented the modifier after the head (e.g., “what would we call a hat made of rocks”). This presented a processing load such that the child had to reorder *hat* and *rock* before responding. Grela and colleagues (2005) concluded that, given a high processing load, a weakness in syntactic ability was revealed.

We do not necessarily disagree with this explanation, but we do think it incomplete to refer to the problem as syntactic. What made *shoe car*, rather than *car shoe*, correct in the Grela et al. (2005) study was the semantic context. The picture of the fantasy object *shoe car* was presented along side a picture of a *book car*. In this particular semantic context, *car* must be the head; *shoe* and *book* modify the head thereby specifying its subtype. The syntax of the compound requires the knowledge that object names relate hierarchically such that

subcategories (e.g., *shoe car* and *book car*) exist within more general categories (e.g., *car*). In the current study we hypothesized that the limitation children with SLI display in ordering NN compounds reflects a lack of appreciation of the semantic relationships between the modifier and head.

Our hypothesis is motivated by a growing literature on semantic deficits among children with SLI. As a group, children with SLI begin to convey meaning via spoken words 11 months later than do typical children (Trauner, Wulfeck, Tallal, & Hesselink, 1995). Deficits in receptive vocabulary (Bishop, 1997; Clarke & Leonard, 1996) and expressive vocabulary (Leonard, Miller, & Gerber, 1999; Thal, O'Hanlon, Clemmons, & Fralin, 1999; Watkins, Kelly, Harbers, & Hollis, 1995) characterize many preschoolers with SLI. During the school years, these deficits may become more marked (Haynes, 1992; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). The problem is not only in knowing enough words but also in establishing a rich understanding of the meaning of any given word (McGregor & Appel, 2002; McGregor, Newman, Reilly, & Capone, 2002; Munro, 2007).

Even more pertinent to the semantics of compounding is knowledge of meaningful relationships between words. Children with SLI exhibit difficulty here as well. Compared to their unaffected AMs, school children with SLI are less likely to specify semantic category relationships in noun definitions (Dockrell, Messer, George, & Ralli, 2003); they are slow to recognize category members (Simmonds, Messer, & Dockrell, 2005); and they fail to take advantage of category organization when recalling word lists (Kail & Leonard, 1986). On repeated word association tasks, children with SLI respond with fewer semantically related words and more errors than both same-age peers and younger vocabulary-matched peers, suggesting weaker links between words in the semantic lexicon (Sheng & McGregor, 2008).

In training studies, one sees the emergence of these problems with semantic relationships. For example, over an 8-week period, Munro (2007) taught 5- and 6-year-olds novel words for subcategories of familiar objects (e.g., *tus* referred to a blue wool sock worn in the winter). A posttest included a word association task to tap the semantic networks the children had built during training. In the normal comparison group, most responses were semantic and 17% of those semantic responses reflected knowledge of the noun hierarchy (e.g., *tus* elicited *glove*, a semantic coordinate of *tus*). Despite receiving the same frequency of input during training as the children in the normal comparison group, the children in the SLI group responded most often with a word that was related to the phonological form of the trained word (e.g., *tus* elicited *tug*). Of the semantic responses that did result, only 7% reflected knowledge of the noun hierarchy.

## CURRENT STUDIES

Given their documented semantic deficits, we hypothesized that children with SLI misorder NN compounds in part because they lack an appreciation of the semantic relationship between the modifier and head roots. To test this hypothesis, we compared children with SLI to two groups of unaffected peers, AMs and younger VMs, in two studies measuring semantic knowledge of NN compounds. The first study involved a task similar to that used by Grela et al. (2005) and also Dalalakis (1999) in that children were asked to coin novel NN compounds in response to pictures of fantasy objects. We were interested in whether they included both the modifier and the head in their naming responses and, of more importance, if so, whether they ordered these two roots correctly. It was critical that the correct order was determined by the semantic context provided within the task itself. Given the expected developmental course of NN compound acquisition, we predicted that all groups would be better at including a modifier and head root than at ordering those roots. Given semantic weaknesses on the part of children with SLI and the semantic basis for the

modifier–head order, we predicted that the performance gap between including and ordering would be particularly wide for the SLI group. A secondary goal for this study was to describe the extent to which the deficits exhibited by the children with SLI were because of performance demands. We operationalized performance demand by varying the support of models prior to the child’s ordering of the modifier and head. Performance should be poorest following models of the two root nouns in the reverse order of that required in the compound.

The second study tested the semantic basis of the compounding problem more directly. Following Krott and Nicoladis (2005), the children were asked to parse and explain the meaning of conventional NN compounds. Krott and Nicoladis focused on parsing ability of their young participants, but we expected that our older participants would do well with parsing but would have some difficulty with the developmentally more advanced chore of explaining the modifier–head relationship. We predicted that the performance gap between parsing and explanation would be particularly pronounced for the children with SLI.

## STUDY 1

### Method

**Participants**—Participants were 16 children with SLI (14 boys) and 34 children without. Seventeen of the children without SLI (6 boys) served as AMs; 17 (9 boys) served as expressive VMs. One boy with SLI participated in Study 2 only. All other children participated in both Studies 1 and 2. The studies were administered in counterbalanced order across participants. Forty-one of the 50 participants (all but 2 participants in the SLI group, 1 in the AM group, and 6 in the VM group) also took part in a study of lexical organization that involved word repetition, naming, and word association (Sheng & McGregor, 2008). The ordering of the compounding tasks and the lexical organization tasks was also counterbalanced across participants.

Table 1 compares the demographic and test score characteristics of the three groups. SLI group participants ranged in age from 5;0 (60 months) to 8;6 (102 months), the AM group participants ranged from 5;0 (60 months) to 8;7 (103 months), and the VM group participants ranged from 3;4 (40 months) to 8;6 (102 months). To be included in the SLI group, a child had to be on a current caseload for remediation of oral language deficits and had to score at least 1.3 *SD* below the mean or poorer on at least two of three standardized language measures that we administered prior to the study: the Structured Photographic Expressive Language Test (SPELT; Dawson, Stout, & Eyer, 2003), a measure of morphosyntactic production; the nonword repetition subtest of the NEPSY (Korkman, Kirk, & Kemp, 1998), a measure of phonological short-term memory; and the Test of Narrative Language (Gillam & Pearson, 2004), a measure of the comprehension and production of stories. These tests were selected because deficits in morphosyntax, phonological memory, and narrative discourse are highly characteristic of SLI (Leonard, 1998) and because the tests themselves have good sensitivity and specificity (Gillam & Pearson, 2004; Korkman et al., 1998, p. 225; Perona, Plante, & Vance, 2005). To be selected for the AM or VM groups, a child had to have no history of special services for language and had to score better than 1 *SD* below the mean on all three of the standardized language measures.

To not bias the results of the study, the children with SLI were selected without regard to their lexical–semantic abilities; however, to better describe the participants and interpret their performance, we administered two measures of lexical semantics, the Expressive Vocabulary Test (EVT; Williams, 1997) and the Peabody Picture Vocabulary Test, Third Edition (PPVT-III, Dunn & Dunn, 1997). As a group, the SLI participants scored within 1



*SD* below the mean on these tests, but their standard scores were nevertheless significantly lower than those of the AM and VM groups (see Table 1).

The children in the VM group had age-appropriate standard scores on the EVT and the PPVT-III; however, because they were younger than the SLI group, their raw scores were highly similar to those of the SLI group. The members of the VM group were selected to match the SLI group on raw vocabulary scores so that the performance of the SLI group could be interpreted relative to the performance of normal children with vocabularies of similar size. On the EVT the mean raw score for the SLI group was 57 ( $SE = 3$ ); the mean score for the VM group was 58 ( $SE = 3$ ),  $t(31) = 0.33$ ,  $p = 0.74$ ,  $D = 0.12$ . On the PPVT-III, the mean raw score for the SLI group was 86 ( $SE = 6$ ); the mean for the VM group was 93 ( $SE = 6$ ),  $t(31) = 0.77$ ,  $p = 0.45$ ,  $D = 0.30$ . The children were not selected to match on morphosyntactic ability and it should be noted that the difference in raw scores on the SPELT, which were 35 ( $SE = 3$ ) for the SLI group and 43 ( $SE = 3$ ) for the VM group, approached significance,  $t(31) = 1.97$ ,  $p = 0.06$ ,  $D = 0.75$ .

All participants had to pass a pure-tone hearing screening administered per the American Speech–Language–Hearing Association (ASHA, 1997) guidelines, have normal or corrected to normal vision according to parent report, and have no history of social/emotional deficits characteristic of autism spectrum disorders per parent report. All children had to score at least 80 on a standardized measure of nonverbal cognition, the Kaufman Brief Intelligence Test matrices (K-BIT; Kaufman & Kaufman, 1990). Despite meeting this last criterion, the SLI children as a group scored significantly lower than their AM peers (see Table 1). Finally, we tallied years of maternal education as this socioeconomic variable is known to predict language development (Dollaghan, Campbell, Paradise, et al., 1999). The mothers of the children with SLI had fewer years of education, on average, than the mothers of the children without SLI. Because lower nonverbal IQ and lower socioeconomic status are related to the phenotype of SLI (Bishop, 1992; Miller, Kail, Leonard, & Tomblin, 2001; Tomblin, Smith, & Zhang, 1997), our primary analyses did not address these between-group differences. However, in a secondary analysis, these variables served as covariates so that we could examine their effects on compounding performance. This approach follows that reported in Raitano, Pennington, Tunick, Boada, and Shriberg (2004).

The recruitment, consenting, and testing of all participants were conducted in accord with the policies of the Internal Review Board at the University of Iowa.

**Stimuli**—Forty novel NN compounds, created by combining early acquired root nouns, comprised the stimuli (see Appendix A). Each was an anticipated label for a line drawing created by adding one part of an object conventionally labeled by the modifier noun to one part of an object conventionally labeled by the head noun. For example, the compound *ghost camel* labeled a camel with a ghost-shaped head (see Appendix B). To ensure that the drawings clearly conveyed the intended modifier–head relationship (i.e., that the picture in Appendix B is better labeled *ghost camel* than *camel ghost*), we asked 12 native English-speaking adults to label each stimulus picture. Only drawings that were identified by the compound form we had assigned to them by at least 10 of the 12 (83%) adults were included.

The examiner presented 40 picture plates to elicit the novel compounds. Each plate included a picture of the object(s) named by the modifier, the head, or both; a picture of the object named by the target novel NN compound; and a second picture that could be named by a novel NN compound that shared a head with the target compound (see Appendix B).

The purpose of the pictures of the modifier and/or head objects was twofold: they kept the child from developing a routine wherein all responses were compounds, and they provided the child with varying levels of support for identification and ordering of the potential modifier and head for the target compound responses. Specifically, before 10 targets, the child was first asked to name the modifier noun and then the head noun; before another 10, the child named the modifier only; before another 10, the child named the head only; and finally, before another 10, the child named the head and then the modifier. Notice that the first of these provides a complete model of the roots of the target compound in correct modifier–head order. The last provides a complete model but in the reverse of the expected modifier–head order. During design of the study, the assignment of items to model types was randomized.

The purpose of the second fantasy picture was to provide the child with a context that promoted the correct selection of modifier versus head. Whereas it is possible to name the picture in Appendix B as either a *ghost camel* or a *camel ghost* (although at least 10 of 12 adults preferred *ghost camel*), in the context of *witch camel*, *ghost camel* becomes the only correct answer. The children’s responses to the second picture were not analyzed, the picture was used solely to provide this semantic context. The second picture was as large as the first, it appeared directly beside the first, and many children named it spontaneously. Therefore, we can be reasonably sure that the children processed this context.

The order of the picture plates was randomized prior to the experiment and kept in that particular random order for all children. Within the given plates, order of the pictures meant to elicit modifier, head, and target compound was consistent; this order is displayed in Appendix B.

**Procedure**—The examiner first demonstrated the desired responses by naming three pictures, *tractor*, *toothbrush*, and *toaster toe*. Upon naming the last one, she explained that the end of the toe looked like a toaster so she named it a toaster toe. She then prompted the child to name the conventional objects (e.g., *camel*) and the fantasy objects (e.g., *ghost camel*) on each picture plate. Responses were written by the examiner and also audiorecorded for later verification. The child did not receive any corrective feedback during the task.

**Data coding**—Responses were first classified as compounds, phrases, blends, single morphemes, or “don’t know”/other. Two of these merit explanation. Phrases were multiword descriptions of the picture such as “an umbrella with a zebra stuck on” for *zebra umbrella*. Blends were a merging of the modifier and head roots without regard to their conventional edges, that is, use of a word like *zebrella* in place of the target *zebra umbrella*.

Each attempt at compounding was scored on two levels: the child earned 1 point for including both the modifier and head and 1 point for correctly ordering the two. Scoring was lenient in that any root words related to the target were accepted. For example, a child who said “horse-umbrella” for the target *zebra-umbrella* was given a point for including the modifier and head as well as a point for correctly ordering those roots. The response “umbrella-horse,” while earning a point for inclusion of modifier and head, would not receive a point for order. Finally, “horse” alone or “umbrella” alone would receive no points.

Because phrases and blends also included the modifier and head nouns, inclusion was credited for these response types as well; however, order was not coded. The scoring system is illustrated in Appendix A.

**Reliability**—A second independent coder coded 22% of the response set (11 samples). Mean agreement between the two coders on response type was 96%, with a range of 89% to 100%. Mean agreement on modifiers was 98%, with a range of 94% to 100% across samples. Mean agreement for heads was 97%, with a range of 91% to 100%. Mean agreement for order was 96%, with a range of 89% to 100%.

**Statistical analysis**—The primary analyses tested for differences in the accuracy of two levels of compound response (inclusion and order of roots) and three participant groups (SLI, VM, AM) via a mixed model analysis of variance (ANOVA) by subject ( $F1$ ) and a repeated measures ANOVA by item ( $F2$ ). Because the development of compounding is influenced by frequency of the modifier and head roots in the input language, and because it was, obviously, impossible for us to completely sample all roots, it was important to treat items as random effects. Therefore, the by-item analysis was particularly important.

A secondary analysis by subjects included the K-BIT standard scores and years of maternal education as covariates. By subject, the dependent variables were the proportion of responses that included both the modifier and head roots and the proportion of compound attempts in which the modifier and head roots were correctly ordered. By item, the dependent variable was the proportion of participants who included both modifier and head nouns and the proportion of participants attempting compounds who ordered the modifier and head correctly. Because the number of opportunities for correct ordering varied from child to child and item to item (i.e., varied with the number of compound attempts), all proportions were arc-sine transformed before statistical analysis. The arc-sine transformation was also useful in that it allowed for a more normal distribution.

Untransformed data were plotted in the figures to facilitate interpretation. All significance tests were two tailed. As an indication of effect size, the partial eta squared value ( $\eta_p^2$ ), or the proportion of the effect + error variance that is attributed to the effect, was computed. The Tukey honestly significant difference for unequal total numbers was used for post hoc testing of between-subject differences and the Bonferroni test was used for post hoc testing of within-subject differences.

## Results

On average, the children in all three groups coined novel compounds on the majority of the 40 trials (see Table 2). Phrasal responses (e.g., “an umbrella with a zebra on it”) were rare overall but they did occur in all groups. In particular, phrases were the preferred response strategy of some children in the VM group, with one child providing phrases on 36 of 40 trials. Blends also occurred rarely but in all groups. Blends were a favorite response strategy of two members of the AM group who responded with blends on 25 and 29 of 40 trials. Single roots and “don’t know”/other responses are different from the other three types because they indicate a lack of understanding or ability to name with the specificity required by the semantic context of the task. Numerically, the VM group produced more single roots than the other groups. Again, there was a large amount of within group variation; one VM participant produced single roots on 37 of 40 trials. Don’t know/other responses were rare for all three groups and for all children within those groups. Children in the SLI group were quite able: like their peers, they responded with novel compounds on most trials and no particular child within the SLI group was highly dependent on any other response strategy.

The results of a 3 (Group)  $\times$  2 (Level of Response) mixed model ANOVA revealed no main effect of group,  $F1(2, 46) = 1.27, p = .29, \eta_p^2 = 0.05$ . There was a main effect of level,  $F1(1, 46) = 8.80, p = .005, \eta_p^2 = 0.16$ , with performance on inclusion of the modifier and head ( $M =$



0.85,  $SE = 0.03$ ) superior to ordering of these roots ( $M = 0.78$ ,  $SE = 0.02$ ). As predicted, there was a significant Group  $\times$  Level interaction,  $F(2, 46) = 3.26$ ,  $p = .05$ ,  $\eta_p^2 = 0.12$ , such that the gap between inclusion of the roots and their ordering was significant only for the SLI group ( $p = .02$ ; see Figure 1a). These results were not altered by inclusion of years of maternal education and nonverbal IQ as covariates. Neither covariate was significant, maternal education,  $F(1, 41) = 0.004$ ,  $p = .95$ ,  $\eta_p^2 = 0.0002$ ; nonverbal IQ,  $F(1, 41) = 1.42$ ,  $p = .24$ ,  $\eta_p^2 = 0.03$ , and there were no interactions between the covariates and inclusion or order of the modifier and head ( $ps > .71$ ). In summary, only the children with SLI had more trouble ordering the compounds than using the roots, and this effect was not driven by socioeconomic status or nonverbal IQ.

In the by-item analysis there was a main effect for group,  $F(2, 78) = 18.51$ ,  $p < .0001$ ,  $\eta_p^2 = 0.32$ ; the AM group and SLI group performed similarly overall (AM:  $M = 0.84$ ,  $SE = 0.01$ ; SLI:  $M = 0.81$ ,  $SE = 0.02$ ) and significantly better than the VM group ( $M = 0.77$ ,  $SE = 0.01$ ),  $ps < .0002$ . There was a main effect of level,  $F(2, 39) = 7.37$ ,  $p = .01$ ,  $\eta_p^2 = 0.16$ , with the proportion of children including both modifier and head ( $M = 0.85$ ,  $SE = 0.02$ ) being greater than the proportion correctly ordering those roots ( $M = 0.77$ ,  $SE = 0.02$ ). As predicted, there was a Group  $\times$  Level interaction,  $F(2, 78) = 28.75$ ,  $p < .0001$ ,  $\eta_p^2 = 0.42$  (see Figure 1b). A higher proportion of the SLI and AM group members included the modifier and head than ordered them correctly ( $p < .0001$  and  $.01$ , respectively). In contrast, similar proportions of VM participants included and ordered the modifier and head,  $p = .31$ . The AM group was superior to the SLI group on order ( $p = .03$ ) but superior to the VM group on inclusion of modifier and head ( $p < .0001$ ). No other differences were significant. In summary, the SLI group was better at including the modifier and head than at ordering them and their ordering was worse than that of their AM peers.

To better understand the basis for performance on modifier–head order, we conducted a multiple regression analysis. We entered raw scores from the EVT, a measure of lexical semantic ability, and the SPELT, a measure of morphosyntactic ability, as predictors. There was a significant effect of lexical semantic ability,  $\beta = 0.56$ , partial  $r = .32$ ,  $t(43) = 2.26$ ,  $p = .03$ , but no effect of morphosyntactic ability,  $\beta = -0.14$ , partial  $r = -.09$ ,  $t(43) = -0.57$ ,  $p = .57$ .

To better understand how the task itself affected performance on the modifier–head order, we plotted performance according to the type of model elicited by the examiner (modifier–head, modifier only, head only, or head–modifier; see Figure 2). As is apparent, performance was best following modifier–head models and worst following head–modifier models. It is important that the trends appear highly similar for the three groups. There is no sign that the heavy processing load represented by reverse (head–modifier) models affected the children with SLI to a greater degree than their peers. To test this, we compared the performance of the SLI and AM groups following modifier–head models and head–modifier models; this decision was driven by low available power, that is, to maximize power, we compared only the conditions and groups that should be most sensitive to any processing load differences. A 2 (Model Type)  $\times$  2 (Group) ANOVA with proportion of responses correctly ordered revealed a main effect of group,  $F(1, 21) = 8.14$ ,  $p = .008$ ,  $\eta_p^2 = 0.21$ , and a main effect of model,  $F(1, 21) = 14.54$ ,  $p = .0006$ ,  $\eta_p^2 = 0.32$ . The main effect of group reflects the results of the larger ANOVA above: the children with SLI were less able than their AMs to order compound constituents correctly. The main effect of model type confirms that reverse order models (head–modifier) pose a larger processing burden than correctly ordered models (modifier–head). It was important that there was no Group  $\times$

Model interaction,  $F(1, 21) = 0.09, p = .77, \eta_p^2 = 0.003$ ; therefore, the children with SLI were no more sensitive to the increased processing demands of the head–modifier models than were their AM peers.

## Discussion

Overall, the children with SLI performed remarkably well on this task. Like their AMs and VMs, they coined novel compounds for the majority of the targets. Compounding was the preferred strategy for all of the children with SLI. Most of the VMs also preferred compounds but some liked to use phrases, which are also acceptable. Most of the AMs preferred compounding but some liked to use blends, which, arguably, are also acceptable given models like *brunch*, *spork*, and *skort* in the language.

The primary difference between the groups was that, as predicted, only the children with SLI had particular problems ordering the roots in their NN compounds. This was a robust finding evident in both the by-subject and by-item analyses. By subject, they were the only group to present with a significant gap between the average proportion of responses that included a modifier and head and the average proportion of compound attempts that were correctly ordered. By item, the SLI and AM groups did not differ in the proportion of their members who included modifiers and heads but a lower proportion of the SLI group could order them correctly. This relative weakness occurred despite the use of drawings that consistently elicited modifier–head order from adults and despite the inclusion of an extra drawing on each page to provide a contextual cue for the head noun. This finding accords with Grela et al. (2005), who also reported children with SLI to have difficulty ordering compounds.

Grela and colleagues (2005) attributed this pattern to processing deficits. We did find performance to vary with the processing demands of the task. Performance was best following elicitations of the roots in the modifier–head order and worst following elicitations of roots in the reverse order (recall that all of the root models in Grela et al., 2005, were in the reverse order). However, it is important to note that this variation in processing demand affected the SLI and AM groups equally. Given that the SLI group was not any more sensitive to high processing loads than their AMs, we conclude that their problem with modifier–head order is not processing-load dependent.

In summary, the children with SLI recognized that the task mandated a level of specificity in naming, and they met this mandate by coining novel compounds. Their difficulty was in placing the head at the right edge of the compound. Because the head noun was determined by the stimulus drawing and the semantic context in which it appeared, we argue that their misorderings were semantically based. Scores on a standardized test of lexical semantics, but not scores on a standardized test of morphosyntax, were predictive of modifier–head ordering, which lends support to this argument. Study 2 provided a more direct test of what children with SLI know about the semantics of compounds.

## STUDY 2

### Method

**Participants**—Participants were the same as those reported in Study 1.

**Stimuli**—Twenty-five conventional NN compounds previously used in Krott and Nicoladis (2005) comprised the stimuli (see Appendix C). The compounds were randomized prior to the study and then presented in a consistent order for all participants.

**Procedure**—The procedure was that used by Krott and Nicoladis (2005). Briefly, the examiner introduced the child to a puppet who “doesn’t understand English well and wants to know why we use some words.” The adjective–noun compound *blueberries* and the verb–noun compound *jump rope* were used as examples prior to asking the child for explanations (e.g., “we say blueberries because they are berries that are blue, right?”).

**Data coding**—Each response was scored at two levels: parsing and explanation. To receive a point for parsing, the child had to mention the modifier, the head, or both in the response. Any mention had to be accurate (e.g., “grapes” did not count as a mention of the modifier in “grapefruit juice”) and when both the modifier and head were mentioned, they had to be separated (e.g., “I like peanut better” did not count as a parsing of the modifier *peanut* from the head *butter*). If the child parsed the compound, the second level, explanation, was scored. An explanation earned a point if it correctly (a) indicated the general category and (b) specified the nature of the subcategory. Criteria (a) and (b) must both be met to earn a point. For example, answering “because it is a fort in the snow” to the question, “why do we say snow fort?” would not earn a point. Even though the general category is indicated, the nature of the subcategory is incorrect. A correct answer would be “because it is a fort that is made of snow.” One point would be awarded for this explanation because the phrase “it is a fort” indicates which one of the two roots is the general category and the phrase “that is made of snow” specifies the nature of the subcategory. Referring to the head with a pronoun was acceptable as long as the modifier and the nature of the subcategory were specified; for example, “because it is made of snow” would also receive credit. The scoring system is illustrated in Appendix C.

**Reliability**—For 11 samples, the mean agreement between the two coders on parsing was 99% with a range of 95% to 100% across samples. Mean agreement on explanation of the modifier–head relationship was 92% but the range, 63% to 100%, was wide because of one problematic sample. Coding of this sample was therefore completed by consensus.

**Statistical analysis**—As in Study 1, the primary analysis involved a mixed-model ANOVA for the by-subject analyses (*F1*) and a repeated-measures ANOVA for the by-item analyses (*F2*). K-BIT standard scores and years of maternal education served as covariates in the secondary analysis. The ANOVAs were 3 (Group)  $\times$  2 (Level). By subject the dependent variable was the arc-sine transformed proportion of items that were correctly parsed and the arc-sine transformed proportion of correctly parsed items that were correctly explained. By item, the dependent variable was the arc-sine transformed proportion of participants who parsed and, of those who parsed, the arc-sine transformed proportion who explained the modifier–head relationship correctly.

## Results

There was no main effect of group,  $F(2, 47) = 2.25, p = .12, \eta_p^2 = 0.09$ , but there was a significant main effect of level,  $F(1, 47) = 87.26, p < .0001, \eta_p^2 = 0.65$ , with parsing of the modifier and head ( $M = 0.91, SE = 0.06$ ) being stronger than explanation of the modifier–head relationship ( $M = 0.41, SE = 0.05$ ). The apparent interaction between group and level (see Figure 3a) was not significant,  $F(2, 47) = 1.41, p = .25, \eta_p^2 = 0.06$ ; however, given an a priori hypothesis, we compared the performance of the SLI and AM groups and the SLI and VM groups on parsing and explanation (for this procedure, see Hsu, 1999; for an application in a study of SLI, see Leonard, Davis, & Deevy, 2007). According to the Newman–Keuls test, the SLI and AM groups differed on explanation ( $p = .05$ ) but not parsing ( $p = .26$ ). The SLI and VM groups did not differ on either level of response ( $ps > .20$ ).

The results above remained the same when years of maternal education and nonverbal IQ were added as covariates. In addition, there was a significant effect of nonverbal IQ,  $F(1, 42) = 5.37, p = .03, \eta_p^2 = 0.11$ , but not maternal education,  $F(1, 42) = 0.53, p = .47, \eta_p^2 = 0.01$ . For the groups combined, there was a positive correlation such that the higher the standard nonverbal IQ score, the better the explanation of the modifier–head relationship,  $r = .20, r^2 = .09, t(47) = 2.1, p = .04$ .

The by-item analysis yielded a main effect of group,  $F(2, 48) = 12.79, p < .0001, \eta_p^2 = 0.35$ , with a larger proportion responding correctly in the AM group ( $M = 0.66, SE = 0.03$ ) than the SLI group ( $M = 0.48, SE = 0.03$ ),  $p = .0002$ , and the VM group ( $M = 0.56, SE = 0.03$ ),  $p = .001$ . The SLI and VM groups did not differ ( $p = .55$ ). There was also a main effect of level,  $F(1, 24) = 283.75, p < .0001, \eta_p^2 = 0.92$ , with more children able to parse the compound ( $M = 0.73, SE = 0.02$ ) than explain it ( $M = 0.40, SE = 0.02$ ). The main effects were qualified by the predicted interaction,  $F(2, 48) = 18.43, p < .0001, \eta_p^2 = 0.43$  (see Figure 3b). In parsing of modifier and head, the AM group did not differ from the SLI group ( $p = .44$ ) and both of those groups were superior to the VM group ( $ps < .02$ ). In contrast, for explaining the modifier–head relationship, the AM group did not differ from the VM group ( $p = 1.0$ ), but both of those groups were superior to the SLI group ( $ps < .0001$ ).

To better understand the basis for explanation of the modifier–head relationship, we conducted a multiple regression analysis. We entered raw scores from the EVT, a measure of lexical semantic ability, and the SPELT, a measure of morphosyntactic ability, as predictors. There was a significant effect of lexical semantic ability,  $\beta = 0.44$ , partial  $r = .29, t(44) = 2.00, p = .05$ , but no effect of morphosyntactic ability,  $\beta = 0.14$ , partial  $r = .10, t(44) = 0.66, p = .51$ .

## Discussion

In summary, children were better at parsing the modifier from the head than at explaining the semantics of the modifier–head relationship, and this was especially true of the children with SLI. These children were significantly poorer than their AM peers (by subject and by item) and VM peers (by item only) at explaining relationships, but they did not differ from their AM peers in parsing.

This difficulty is consistent, in a broad sense, with previous reports of semantic deficits among children with SLI. Children with SLI tend to make more naming errors than their unaffected age-mates, and these errors are associated with sparse semantic knowledge. Specifically, children provide less complete drawings (McGregor & Appel, 2002; McGregor et al., 2002) and definitions (McGregor, Berns, Owen, & Michels, 2010; McGregor et al., 2002) of items that they misname than of items that they name correctly. McGregor and colleagues hypothesize that weak or underspecified semantic representations render processing (naming) more vulnerable. Moreover, children with SLI have weaker semantic representations than their peers who are matched in age and thus in word-learning experience (McGregor et al., 2010). As applied to the current study, the children with SLI may have had rich enough representations of the compounds to parse them, but not enough to provide a verbal explanation of their modifier–head relationships.

Explanation of the modifier–head relationship varied not only with language group but also with nonverbal cognitive ability. The small but significant positive correlation between scores on the K-BIT matrices and the explanation task is consistent with previous reports. Purcell and colleagues (2001) found scores on parent reports of verbal and nonverbal abilities among 3,000 2-year-old twins to hold a correlation of +0.44. In that sample, 21.5%

of children who scored at or below the fifth percentile on vocabulary also scored at or below the fifth percentile on nonverbal ability, a much higher concordance than would occur by chance. De-Thorne and Watkins (2006) also reported significant positive correlations between nonverbal IQ scores and scores on standardized tests (but not criterion-referenced measures) of both semantics and morphosyntax. It may be that the weaker non-verbal abilities that characterized the SLI group in the current study contributed to their relative difficulty on modifier–head explanations; however, this difficulty remained when nonverbal IQs were controlled via statistical covariation.

## GENERAL DISCUSSION

In two studies 5- to 8-year-olds with SLI exhibited relative strengths in manipulating root words to build or parse compounds. In the first study, they were as good as their AMs at including two root nouns in their naming responses. Clearly, they were aware that both root words were pertinent for naming at an appropriate level of specificity. They rarely used these roots in phrases or blends. Instead, they combined the roots to coin novel compounds. As a group, they adhered more consistently to this expected response type than either their AMs, some of whom preferred blends instead, or their younger VMs, some of whom preferred phrasal responses. In the second study, children with SLI were as good as their AMs at parsing the modifier and head roots of conventional compounds. Clearly, they recognized these roots and they were aware that the root words were pertinent to the explanation of the compound. In more general terms, their good performance in both studies suggests that they are able to analyze the structure of NN compounds. Their analytical skill actually exceeded their vocabulary development in the sense that, in comparison to their younger VMs, more of the children with SLI could correctly parse compounds. Because they were older than their VMs, they likely had more experience with these compounds.

The children with SLI also exhibited relative weaknesses. In the first study, each fantasy object to be named appeared alongside a second fantasy object that was its semantic neighbor. These neighbors were members of the same semantic category; hence, presentation of the neighbor dictated the choice of head noun. Only the SLI group demonstrated a significant gap between the proportion of responses that included two roots and the proportion that ordered the correct root into head position. One possibility is that these misorderings reflect deficits in knowledge or use of a syntactic paradigm. However, order performance was not predicted by scores on a standardized test of morphosyntax. Given that we used only a single, brief test of morphosyntax (and one heavily weighted toward tapping grammatical inflection rather than word order), we cannot completely rule out effects of syntactic ability, or overall language ability, on performance.

We should point out, however, that the particular syntactic paradigm of interest here, that of right-headed NN compounds, is extremely rich and regular in English. With regard to the application of morphosyntactic paradigms at the phrase and sentence levels, Leonard (1998) has made the case that children with SLI have less difficulty with rich and regular paradigms than with sparse, inconsistent ones. Take, for example, German and English. German has a rich system of subject–verb agreement marking; whether singular or plural, whether first, second, or third person, each stem takes an overt inflection (strong verbs involve changes in the root as well). English has a much sparser morphology. For present tense regular verbs, only the third-person singular is overtly inflected. Consistent with the richness hypothesis, English speakers with SLI make more errors on subject–verb agreement than their German counterparts (Roberts & Leonard, 1997). The two languages also differ in the regularity of their morphosyntactic systems. When it comes to basic sentence construction, English has a very regular subject–verb–object order; in comparison, German is less regular and it is the



German speakers with SLI who present with word order errors (Clahsen, 1991; Hamann, Penner, & Lindner, 1998).

Applying the richness and regularity accounts to the word level, it is unlikely that English speakers with SLI misorder the highly regular modifier–head construction of the ubiquitous NN compound for purely syntactic reasons. Instead, we argue for a second possibility: the syntactic paradigm is misapplied because of faulty semantic processing. In a broader sense, this view is compatible with reports that semantic complexity affects syntactic processing at the sentence level (Windsor, 1999).

In future work, it would be useful to examine modifier–head order in contexts that represent more varied semantic relationships. In the current study, all of the novel compounds conveyed a part–whole relationships (e.g., *ghost camel* = a camel with a ghost head; *pencil ladder* = a ladder with side rails made of pencils). Given the importance of input frequency on the development of compounding, if faulty semantic processing is at play, more frequent semantic relationships may increase correct head placement and less frequent relationships may decrease this placement. Similarly, one could make the semantic cues for designation of the head noun more salient. Rather than a single neighbor (e.g., *witch camel*), multiple neighbors (e.g., *witch camel*, *alien camel*, *monster camel*) might lead to a more reliable assignment of the head noun to the right edge of the compound.

The second study provided more direct evidence of a semantic basis for the compounding problem. On average, compared to both the AM and VM groups, proportionately fewer members of the SLI group could explain the meaningful relationship between the modifier and head roots. This finding is consistent with previous findings from our lab and others. In tasks ranging from word association (Munro, 2007; Sheng & McGregor, 2008), to definition (Dockrell et al., 2003; McGregor et al., 2010), to list recall (Kail & Leonard, 1986), children with SLI do not respond with semantically related words to the same extent as their unaffected peers. We hypothesize that these children have less knowledge of the relationships between words in their semantic lexicons.

One potential explanation for this problem lies in the lower vocabulary scores of the SLI participants. If we take scores on standardized tests of receptive and expressive vocabulary as an estimate of vocabulary size, then it is apparent that the children with SLI knew fewer words than their AMs. Logically, knowing fewer words reduces the possible number of relationships within the semantic lexicon. However, vocabulary size cannot be the only explanation. Compared to younger children who were selected because their vocabularies were similarly small, a lower proportion of children with SLI explained compounds correctly.

The work of Van Helden-Lankhaar (2001) presents an additional explanation. She argues that the cognitive capacity for comparison is behind the establishment of lexical–semantic relationships, and thus, behind the ability to coin and explain compounds. In her developmental account, children learn to compare and contrast coordinate concepts via input rich with contrastive juxtapositions (e.g., “Do you want juice or milk?”; “Give the cup to mama, not daddy.”). By the process of comparison, they build links between coordinated entries in their lexicon.

Van Helden-Lankhaar’s (2001) account yields a parsimonious explanation of the difficulty that children with SLI had constructing novel compounds and explaining old ones. To coin novel compounds, they must make comparisons in the moment. To coin *ghost camel*, the child must recognize the similarities between *ghosts*, *camels*, and the fantastic combination of those two entities in the stimulus picture. To determine that *camel* is the head of the compound, the child must compare *ghost camel* to *witch camel*; their similarity determines

the head, their difference determines the modifier. Likewise, the child's explanation of conventional compounds may reflect knowledge built upon previous comparisons. It should be easier for the child to explain why we say *cheese sandwich* if he or she has knowledge of other coordinates that manifest the same modifier–head relationship, coordinates such as *ham sandwich* and *peanut butter sandwich* (Krott et al., 2009).

Van Helden-Lankhaar (2001) views the capacity for comparison as domain general. This presents an interesting hypothesis for future research on SLI: their deficient knowledge of lexical–semantic relationships is part of a broader deficit that is measurable in nonverbal tasks as well as verbal tasks. The correlation between nonverbal IQ and explanation of modifier–head relationships is tantalizing in this regard but recall that there was no correlation between nonverbal IQ and modifier–head ordering. A larger sample and a more in-depth measure of nonverbal cognition will be required for an adequate test of the hypothesis.

To conclude, 5- to 8-year-olds with SLI were more adept at analyzing roots within NN compounds than ordering them or explaining their relationship, neither of which they could do as well as typically developing peers. We argue that the problems in ordering and explaining reflect limited knowledge of lexical–semantic relationships, and we hypothesize that a smaller vocabulary and poorer comparison abilities may contribute to this limitation.

## Acknowledgments

We thank the participants and their families for their generosity, and Amanda Owen, Amanda Berns, Tracie Machetti, Allison Bean, Tracy Ball, and Rick Arenas for their valuable contributions. The first author gratefully acknowledges the support of NIH-NIDCD Grant 2 R01 DC003698. Some of the data in this paper were presented at the 2007 Symposium for Research in Childhood Language Disorders, Madison, Wisconsin.

## References

- American Speech–Language–Hearing Association. Guidelines for audiologic screening. Rockville, MD: Author; 1997.
- Berko J. The child's learning of English derivational morphology. *Word*. 1958; 14:150–177.
- Bishop DVM. The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry*. 1992; 33:3–66. [PubMed: 1737831]
- Bishop, DVM. *Uncommon understanding: Development and disorders of language comprehension in children*. Hove: Psychology Press; 1997.
- Carr E. Word-compounding in American speech. *Speech Monographs*. 1959; 26:1–20.
- Clahsen, H. *Child language and developmental dysphasia*. Amsterdam: John Benjamins; 1991.
- Clark, EV. *The lexicon in acquisition*. Cambridge: University Press; 1993.
- Clark EV, Gelman SA, Lane NM. Compound nouns and category structure in young children. *Child Development*. 1985; 56:84–94.
- Clarke M, Leonard L. Lexical comprehension and grammatical deficits in children with specific language impairment. *Journal of Communication Disorders*. 1996; 29:95–105. [PubMed: 9157178]
- Dalalakis JE. Morphological representation in specific language impairment: Evidence from Greek word formation. *Folia Phoniatica et Logopaedica*. 1999; 51:20–35. [PubMed: 9973664]
- Dawson, JI.; Stout, CE.; Eyer, JA. *Structured Photographic Expressive Language Test*. 3. DeKalb, IL: Janelle Publications; 2003.
- DeThorne LS, Watkins RV. Language abilities and nonverbal IQ in children with language impairment: Inconsistency across measures. *Clinical Linguistics and Phonetics*. 2006; 20:641–658. [PubMed: 17342874]
- Dockrell JE, Messer D, George R, Ralli A. Beyond naming patterns in children with WFDs—Definitions for nouns and verbs. *Journal of Neurolinguistics*. 2003; 16:191–211.

- Dollaghan CA, Campbell TF, Paradise JL, et al. Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1432–1443.
- Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test. 3. Circle Pines, MN: American Guidance Service; 1997.
- Fukuda SE, Fukuda S. The operation of rendaku in the Japanese specifically language-impaired: a preliminary investigation. *Folia Phoniatica et Logopaedica*. 1999; 51:36–54. [PubMed: 9973665]
- Gillam, RB.; Pearson, NA. Test of Narrative Language. Austin TX: ProEd; 2004.
- Goulden R, Nation P, Read J. How large can a receptive vocabulary be? *Applied Linguistics*. 1990; 11:341–363.
- Grela B, Snyder W, Hiramatsu K. The production of novel root compounds in children with specific language impairment. *Clinical Linguistics and Phonetics*. 2005; 19:701–715. [PubMed: 16147410]
- Hamann C, Penner Z, Lindner K. German impaired grammar: The clause structure revisited. *Language Acquisition*. 1998; 7:193–245.
- Haynes C. Vocabulary deficit: one problem or many? *Child Language Teaching and Therapy*. 1992; 8:1–17.
- Hsu, J. Multiple comparisons: Theory and methods. Boca Raton, FL: Chapman & Hall/CRC Press; 1999.
- Kail R, Leonard LB. Word-finding abilities in language-impaired children. *ASHA Monographs*. 1986; 25
- Kaufman, AS.; Kaufman, NL. Kaufman Brief Intelligence Test. Circle Pines, MN: American Guidance Service; 1990.
- Kehayia E. Lexical access and representation in individuals with developmental language impairment: A cross-linguistic study. *Journal of Neurolinguistics*. 1997; 10:139–149.
- Korkman, M.; Kirk, U.; Kemp, S. A Developmental Neuropsychological Assessment (NEPSY). San Antonio, TX: Psychological Corporation; 1998.
- Krott A, Gagné C, Nicoladis E. How the parts relate to the whole: Frequency effects on children's interpretations of novel compounds. *Journal of Child Language*. 2009; 36:85–112. [PubMed: 18783633]
- Krott A, Nicoladis E. Large constituent families help children parse compounds. *Journal of Child Language*. 2005; 32:139–158. [PubMed: 15779880]
- Leonard L, Miller C, Gerber E. Grammatical morphology and the lexicon in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 1999; 42:678–689.
- Leonard, LB. Children with specific language impairment. Cambridge, MA: MIT Press; 1998.
- Leonard LB, Davis J, Deevy P. Phonotactic probability and past tense use by children with specific language impairment and their typically developing peers. *Clinical Linguistics & Phonetics*. 2007; 21:747–758. [PubMed: 17882693]
- McGregor KK, Appel A. On the relation between mental representation and naming in a child with specific language impairment. *Clinical Linguistics & Phonetics*. 2002; 16:1–20. [PubMed: 11913029]
- McGregor KK, Berns A, Owen AJ, Michels SA. Associations between syntax and the lexicon among children with SLI and ASD. 2010 Unpublished manuscript.
- McGregor KK, Newman RM, Reilly R, Capone NC. Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2002; 45:998–1014.
- Miller C, Kail R, Leonard L, Tomblin JB. Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2001; 44:416–433.
- Munro, N. Unpublished doctoral dissertation. University of Sydney; 2007. The quality of word learning in children with specific language impairment.
- Nicoladis E. The cues that children use in acquiring adjectival phrases and compound nouns: Evidence from bilingual children. *Brain and Language*. 2002; 81:635–648. [PubMed: 12081428]
- Nicoladis E, Krott A. Word family size and French-speaking children segmentation of existing compounds. *Language Learning*. 2007; 57:201–228.

- Oetting JB, Rice ML. Plural acquisition in children with specific language impairment. *Journal of Speech and Hearing Research*. 1993; 36:1236–1248. [PubMed: 8114491]
- Perona K, Plante E, Vance R. Diagnostic accuracy of the Structured Photographic Expressive Language Test (SPELT-3) (3rd ed.). *Language, Speech, and Hearing Services in Schools*. 2005; 36:103–115.
- Purcell S, Eley TC, Dale PS, Oliver B, Petrill SA, Price TS, et al. Comorbidity between verbal and non-verbal cognitive delays in 2-year-olds: A bivariate twin analysis. *Developmental Science*. 2001; 4:195–208.
- Raitano NA, Pennington BF, Tunick RA, Boada R, Shriberg LD. Pre-literacy skills of subgroups of children with speech sound disorders. *Journal of Child Psychology and Psychiatry*. 2004; 45:821. [PubMed: 15056313]
- Roberts SS, Leonard LB. Grammatical deficits in German and English: A crosslinguistic study of children with specific language impairment. *First Language*. 1997; 17:131–150.
- Sheng, L.; McGregor, KK. Lexical–semantic organization in children with SLI. 2008. Unpublished manuscript
- Simmonds, L.; Messer, D.; Dockrell, J. Exploring semantic deficits in children with WFDs. Paper presented at the 10th International Congress for the Study of Child Language; Berlin. 2005 Jul.
- Stothard SE, Snowling MJ, Bishop DVM, Chipchase BB, Kaplan CA. Language impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*. 1998; 41:407–418.
- Thal D, O’Hanlon L, Clemmons M, Fralin L. Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language, and Hearing Research*. 1999; 42:482–496.
- Tomblin JB, Smith E, Zhang X. Epidemiology of specific language impairment: Prenatal and perinatal risk factors. *Journal of Communication Disorders*. 1997; 30:325–344. [PubMed: 9208366]
- Trauner, D.; Wulfeck, B.; Tallal, P.; Hesselink, J. Technical Report No CND-9513. San Diego, CA: University of California San Diego, Center for Research in Language; 1995. Neurologic and MRI profiles of language impaired children.
- Van Der Lely HKJ, Christian V. Lexical word formation in children with grammatical SLI: A grammar-specific versus an input-processing deficit? *Cognition*. 2000; 75:33–63. [PubMed: 10815777]
- Van Helden-Lankhaar M. A connection in lexical development. *Annual Review of Language Acquisition*. 2001; 1:157–190.
- Watkins RV, Kelly DJ, Harbers HM, Hollis W. Measuring children’s lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech, Language, and Hearing Research*. 1995; 38:1349–1355.
- Williams, KT. Expressive Vocabulary Test. Circle Pines, MN: American Guidance Service; 1997.
- Windsor J. Effect of semantic inconsistency on sentence grammaticality judgements for children with and without language-learning disabilities. *Language Testing*. 1999; 16:293–313.

## APPENDIX A

### Study 1 stimuli, example responses, and scores

| Stimulus       | Response             | Response Type        | Inclusion Score | Order Score |
|----------------|----------------------|----------------------|-----------------|-------------|
| ZEBRA UMBRELLA | Umbrella zebra       | Compound             | 1               | 0           |
| SANTA SHEEP    | Santaclaus sheep     | Compound             | 1               | 1           |
| SEAL PHONE     | Phone w/a seal in it | Phrasal <sup>a</sup> | 1               |             |
| BOY BRUSH      | Boy brush            | Compound             | 1               | 1           |
| GIRAFFE CAKE   | Cake                 | Single morpheme      | 0               |             |
| CAR BUTTON     | Car                  | Single morpheme      | 0               |             |

| Stimulus            | Response           | Response Type        | Inclusion Score | Order Score  |
|---------------------|--------------------|----------------------|-----------------|--------------|
| HOUSE KITE          | Home kite          | Compound             | 1               | 1            |
| KEY LAMP            | Don't know         | Don't know           | 0               |              |
| GRAPE HAT           | Wearing            | Other                | 0               |              |
| LION DESK           | Tiger desk         | Compound             | 1               | 1            |
| ARM FENCE           | Fence arm          | Compound             | 1               | 0            |
| BIRD MOBILE         | Bird mobile        | Compound             | 1               | 1            |
| EAR BALL            | Head ball          | Compound             | 1               | 1            |
| SCREW LEGS          | Nail legs          | Compound             | 1               | 1            |
| FLOWER SCISSORS     | Flissors           | Blend <sup>a</sup>   | 1               |              |
| HAND BALLOON        | Balloon hand       | Compound             | 1               | 0            |
| PINEAPPLE PURSE     | Pineapple bag      | Compound             | 1               | 1            |
| NOSE CUP            | Nose cup           | Compound             | 1               | 1            |
| GLOVE PLANT         | Hand plant         | Compound             | 1               | 1            |
| FORK FLOWER         | Fork flower        | Compound             | 1               | 1            |
| LETTER HAIR         | Letter head        | Compound             | 1               | 1            |
| BIKE SKIRT          | Bicycle skirt      | Compound             | 1               | 1            |
| BUNNYBRIDE          | Rabbit bride       | Compound             | 1               | 1            |
| BOTTLE CHIMNEY      | Bottle chimney     | Compound             | 1               | 1            |
| DOCTOR BUILDING     | Man building       | Compound             | 1               | 1            |
| APPLE FISH          | Fishy              | Other                | 0               |              |
| GHOST CAMEL         | Ghost animal       | Compound             | 1               | 1            |
| PIG TRUCK           | Truck pig          | Compound             | 1               | 0            |
| CAMERA SUITCASE     | Camera case        | Compound             | 1               | 1            |
| FLAG HORSE          | Flag horse         | Compound             | 1               | 1            |
| SHOE SOUP           | Shoes are floating | Phrasal <sup>a</sup> | 0               |              |
| BOAT CRUTCH         | Crutch boat        | Compound             | 1               | 0            |
| DOG BREAD           | Dog couch          | Compound             | 1               | 1            |
| PENCIL LADDER       | Pencil ladder      | Compound             | 1               | 1            |
| BROOM SLIDE         | Broom              | Single morpheme      | 0               |              |
| CANDY TREE          | Candy plant        | Compound             | 1               | 1            |
| MONKEY NECKLACE     | Monkey necklace    | Compound             | 1               | 1            |
| MAN BAT             | Boy bat            | Compound             | 1               | 1            |
| DUCK WHEELS         | Wheel duck         | Compound             | 1               | 0            |
| SUN TURTLE          | Sun turtle         | Compound             | 1               | 1            |
| Participant's score |                    |                      | 33/40 = 0.825   | 25/31 = 0.81 |

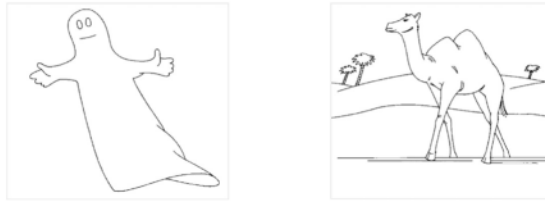
<sup>a</sup>Although both the modifier and head noun are included, these responses are irrelevant to order accuracy because a compound was not attempted.



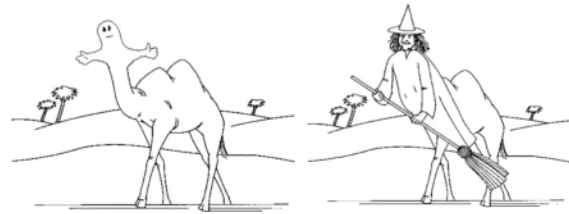
## APPENDIX B

### Example picture stimuli and verbal prompts for Experiment 1

While pointing, examiner prompts, “what’s this?” to elicit *ghost*. “What’s this?” to elicit *camel*.



“What would you call this one?” to elicit *ghost camel*.



Plates varied as to whether both the modifier and head were prompted and in which order. Note that *witch camel* was never elicited, its presence was necessary to cue the compound head, *camel*.

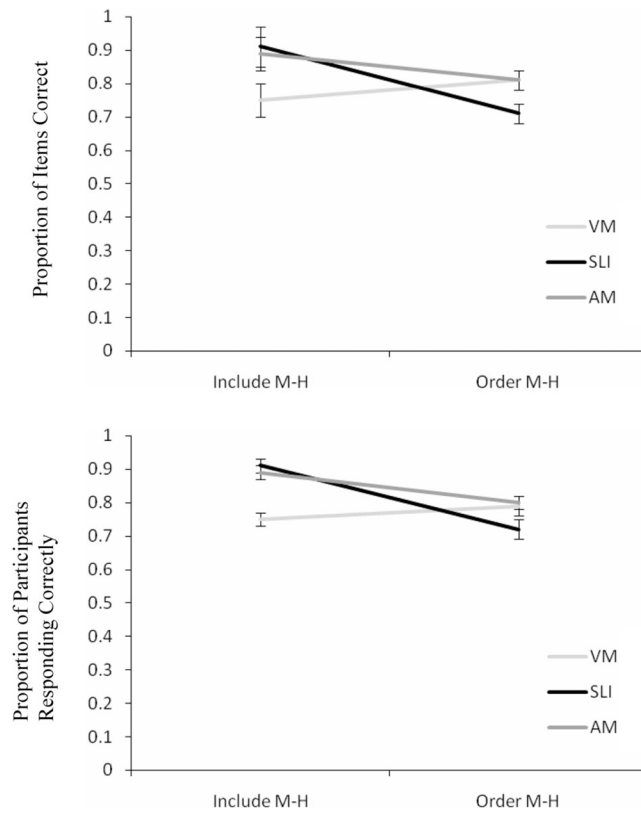
## APPENDIX C

### Study 2 stimuli, example responses, and scores

| Stimulus         | Response                              | Parsing Score | Relationship Score |
|------------------|---------------------------------------|---------------|--------------------|
| CAR DOOR         | B/c it's a car and a door             | 1             | 0                  |
| BOOK SHELF       | B/c books go on shelves               | 1             | 0                  |
| GRAPEFRUIT JUICE | B/c we make juice from grapes         | 0             |                    |
| HOSPITAL BED     | B/c it's a bed in a hospital          | 1             | 1                  |
| FRUIT BASKET     | B/c it holds fruit                    | 1             | 1                  |
| STEP LADDER      | B/c ladders have steps                | 1             | 1                  |
| CARDBOARD BOX    | B/c cards go in it                    | 0             |                    |
| BABY BOOK        | B/c it's a book for/about babies      | 1             | 1                  |
| SNOW FORT        | B/c you make it out of snow           | 1             | 1                  |
| APRON STRINGS    | B/c aprons are made of string         | 1             | 0                  |
| DUCK FEET        | B/c duck have feet                    | 1             | 0                  |
| CHOCOLATE CAKE   | B/c it's a cake made of chocolate     | 1             | 1                  |
| PEANUT BUTTER    | B/c I like peanut butter              | 0             |                    |
| WATER PISTOL     | Don't know                            | 0             |                    |
| CRAYON BOX       | B/c it's a box for crayons            | 1             | 1                  |
| ROOF RACK        | B/c your roof gets racked             | 1             | 0                  |
| CHEESE SANDWICH  | B/c it's a sandwich w/cheese in/on it | 1             | 1                  |

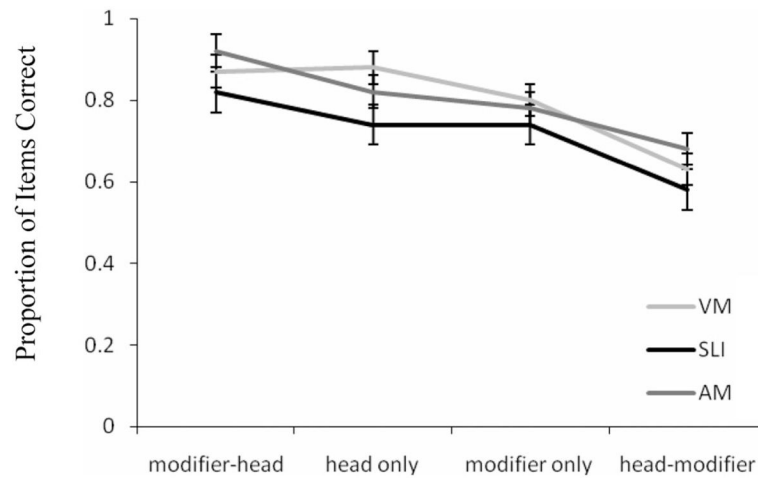
| Stimulus            | Response                               | Parsing Score | Relationship Score |
|---------------------|--|---------------|--------------------|
| PAPER NAPKIN        | B/c you have to wipe your mouth        | 0             |                    |
| HEAT RASH           | B/c it's a rash made of heat           | 1             | 0                  |
| APPLE CORE          | B/c it's the core of/inside an apple   | 1             | 1                  |
| POWER TOOLS         | B/c the tools are made of power        | 1             | 0                  |
| POPSICLE STICK      | B/c when there's a stick in a popsicle | 1             | 1                  |
| BREAKFAST CEREAL    | B/c you eat cereal at breakfast time   | 1             | 1                  |
| CORN BREAD          | B/c it's bread made of corn            | 1             | 1                  |
| WATER SLIDE         | B/c water is for sliding               | 1             | 0                  |
| Participant's score |  | 20/25 = 0.80  | 12/20 = 0.60       |

*Note:* B/c, because.

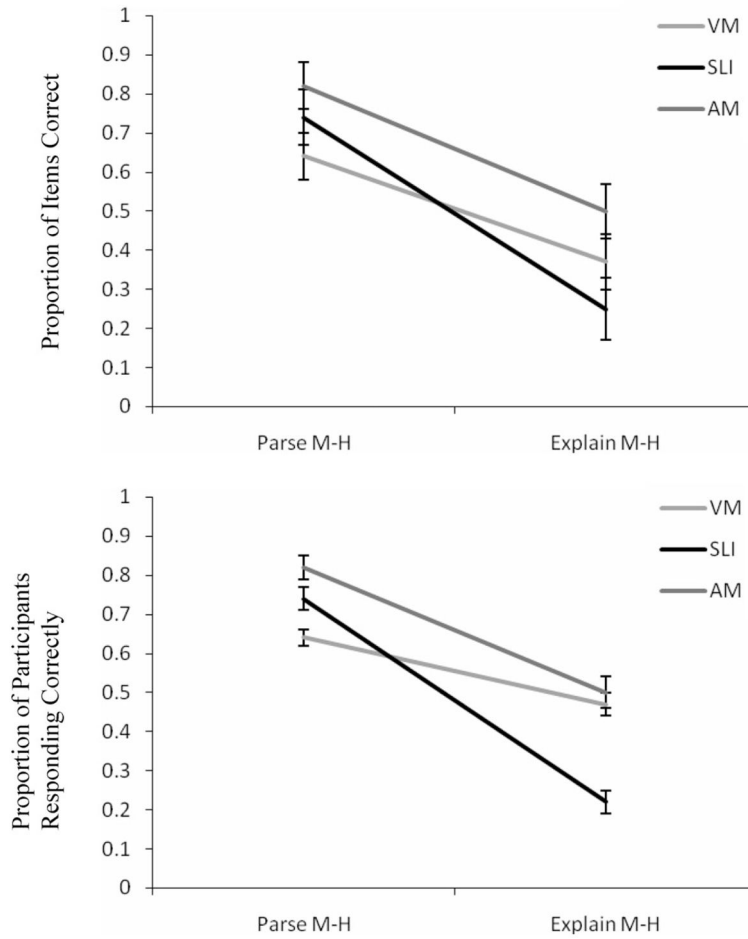


**Figure 1.**

The accuracy of modifier (M)–head (H) inclusion and ordering. The accuracy is plotted for each participant group and is expressed as (a) the proportion of items correct averaged over participants and (b) the proportion of participants who responded correctly averaged over items. Error bars are included.



**Figure 2.** The accuracy of modifier–head order according to the type of model elicited prior to compound naming. The accuracy is plotted for each participant group and is expressed as the proportion of items correct averaged over participants. Error bars are included.



**Figure 3.** The accuracy of parsing the modifier (M) from the head (H) and explaining their relationship. The accuracy is plotted for each participant group and is expressed as (a) the proportion of items correct averaged over participants and (b) the proportion of participants who responded correctly averaged over items. Error bars are included.



Table 1

Demographic information and standardized test scores by participant group

| Measure          | SLI-VM Comparison |               |                           | SLI-AM Comparison      |                       |               |          |
|------------------|-------------------|---------------|---------------------------|------------------------|-----------------------|---------------|----------|
|                  | <i>p</i>          | <i>t</i> (31) | VM Mean ( <i>SD</i> )     | SLI Mean ( <i>SD</i> ) | AM Mean ( <i>SD</i> ) | <i>t</i> (31) | <i>p</i> |
| Age (months)     | .006              | 2.95          | 66 (19)                   | 82 (11.8)              | 82 (12.1)             | <1            | .995     |
| SPELT            | <.0001            | -7.70         | 105 (12.43)               | 73 (10.16)             | 111 (9.27)            | -11.06        | <.0001   |
| TNL-NLAI         | .0007             | -3.93         | 108 <sup>a</sup> (10.19)  | 84 (16.27)             | 108 (11.17)           | -5.04         | <.0001   |
| NEPSY-NWR        | <.0001            | -6.42         | 11.45 <sup>a</sup> (2.25) | 6.81 (1.5)             | 12.06 (1.6)           | -9.66         | <.0001   |
| EVT              | .0005             | -3.92         | 105 (13.37)               | 88 (11.73)             | 110 (8.50)            | -6.27         | <.0001   |
| PPVT             | <.0001            | -5.41         | 117 (11.14)               | 96 (10.58)             | 113 (10.75)           | -4.31         | .0002    |
| Mat. ed. (years) | .0004             | -4.00         | 17 (2.1)                  | 14 (2.0)               | 16 (2.5)              | -2.50         | .02      |
| KBIT-matrices    | .01               | -2.71         | 108 (12.79)               | 96 (12.19)             | 106 (11.94)           | -2.22         | .03      |

*Note:* SLI, specific language impairment; VM, vocabulary mate; AM, age mate; SPELT, Structured Photographic Expressive Language Test; TNL-NLAI, Test of Narrative Language—Narrative Language Ability Index; NEPSY-NWR, NEPSY—Nonword repetition subtest; EVT, Expressive Vocabulary Test; PPVT, Peabody Picture Vocabulary Test (3rd ed.); Mat. ed., maternal education; KBIT-matrices, Kaufman Brief Intelligence Test—Matrices subtest. The comparison of the SLI and VM groups proceeds from the middle to the left, and the comparison of the SLI and AM groups proceeds from the middle to the right. Unless indicated otherwise, all measures are reported as standard scores as determined by the scoring procedures and normative data reported in the test manuals.

<sup>a</sup>The standard scores on the TNL-NLAI and the NEPSY-NWR are based on 9 and 11 children, respectively, because only these children were old enough to take the tests.

**Table 2**

Number of responses (out of 40) by type and group

|                  | VM      | SLI    | AM      |
|------------------|---------|--------|---------|
| Compound         |         |        |         |
| <i>M</i>         | 25.12   | 34.93  | 30.76   |
| <i>SD</i>        | (15.51) | (4.28) | (10.93) |
| Range            | 0–40    | 23–40  | 6–40    |
| Phrases          |         |        |         |
| <i>M</i>         | 4.47    | 0.80   | 1.29    |
| <i>SD</i>        | (9.00)  | (0.86) | (2.49)  |
| Range            | 0–36    | 0–2    | 0–9     |
| Blends           |         |        |         |
| <i>M</i>         | 0.35    | 0.53   | 3.41    |
| <i>SD</i>        | (0.79)  | (1.13) | (8.95)  |
| Range            | 0–3     | 0–4    | 0–29    |
| Single roots     |         |        |         |
| <i>M</i>         | 8.41    | 2.87   | 3.47    |
| <i>SD</i>        | (11.39) | (4.05) | (6.00)  |
| Range            | 0–37    | 0–16   | 0–20    |
| Don't know/other |         |        |         |
| <i>M</i>         | 1.76    | 1.07   | 1.18    |
| <i>SD</i>        | (2.22)  | (1.39) | (1.19)  |
| Range            | 0–6     | 0–5    | 0–4     |

*Note:* VM, vocabulary mate; SLI, specific language impairment; AM, age mate.