J. Child Lang., Page 1 of 28.
doi:10.1017/S030500090800888X Printed in the United Kingdom

How the parts relate to the whole: Frequency effects on children's interpretations of novel compounds*

ANDREA KROTT

University of Birmingham, United Kingdom

CHRISTINA L. GAGNÉ AND ELENA NICOLADIS

University of Alberta, Canada

(*Received* 1 *August* 2007. *Revised* 20 *December* 2007)

ABSTRACT

This study explores different frequency effects on children's interpretations of novel noun–noun compounds (e.g. *egg bag* as 'bag FOR eggs'). We investigated whether four- to five-year-olds and adults use their knowledge of related compounds and their modifier–head relations (e.g. *sandwich bag* (*FOR*) or *egg white* (*PART-OF*)) when explaining the meaning of novel compounds and/or whether they are affected by overall frequency of modifier–head relations in their vocabulary. Children's interpretations were affected by their experience with relations in compounds with the same head, but not by overall relation frequency. Adults' interpretations were affected by their experience with relations in compounds with the same modifier, suggesting that children and adults use similar but different knowledge to interpret compounds. Furthermore, only children's interpretations revealed an overuse of visually perceivable relations.

The acquisition of words and linguistic constructions appears to be affected by input frequency. For example, children's first words are often the most frequently used words in their language input (e.g. Harris, Barrett, Jones & Brookes, 1988). Similarly, the first verbs that children correctly mark for the past tense are often the verbs that are high in frequency (e.g. Kuczaj, 1977). There is also evidence that linguistic constructions are learned earlier

[*] We would like to thank Helen Breadmore for her help in conducting the experiment and transcribing participants' responses and Harald Baayen for his help with the statistical analysis. This research was supported by a British Academy Joint Activities Grant, NSERC Discovery Grant 203054 and NSERC Discovery Grant 245058.
Address for correspondence: Andrea Krott, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom. tel: +44 (0)121 414 4903. fax: +44 (0)121 4144897. e-mail a.krott@bham.ac.uk.

when children hear many different exemplars. For example, the passive construction is learned earlier in Inuktitut than in English and one of the likely reasons for this is that passive constructions are more frequent in Inuktitut than in English (Allen, 1996). Similarly, English-speaking children can use the regular past tense morpheme *-ed* productively, even though many regular verbs are rarely used (see Marchman, 1997). In sum, this research suggests that input frequency influences the acquistion of words, morphemes and syntactic phrases. The present study extends previous work by investigating the role of frequency in children's acquisition of a construction that is situated between words and syntactic phrases, i.e. novel English noun–noun compounds such as *egg bag* or *birthday room*[i]. In particular, we focus on children's understanding of the relations between the constituents of these constructions (e.g. 'bag FOR eggs' for *egg bag* or 'soup MADE OF chocolate' for *chocolate soup*) and how this understanding is affected by their experience with similar compounds.

Noun–noun compounds are composed of at least two nouns. The constituent nouns do not play the same semantic role; one plays the role of a modifier, while the other plays the role of the head and specifies the category to which the construct in its entirety belongs. For example *apple juice* is a kind of juice. In English, the noun on the left is usually the modifier and the noun on the right is usually the head. There can be exceptions to this general rule; for example, *a jailbird* is not a type of bird and it is not obvious how *straw* modifies *berry* in *strawberry*. These compounds are (partly) semantically opaque, meaning the right noun is not the superordinate category (*jailbird*) or the meanings of the parts do not fully contribute to the meaning of the construct (see Clark, 1981, for discussion on how transparency is important for children's compound coinages). While children encounter both transparent and opaque compounds, this study focuses on novel transparent compounds.

To understand the relation between the constituents of a transparent compound, children need to know that compounds are complex words. It has been argued that children first do not have such an understanding, but initially learn compounds as holophrases or chunks (Berman, 1987). In other words, while children might understand the meaning of *football, foot* and *ball*, and that one uses it to play, they might not understand that *football* refers to a ball that one kicks with a foot. There is even some evidence that children have some difficulty parsing some compounds into the early school years (e.g. *birthday*; Berko, 1958), suggesting that the chunking stage is not

[i] While some scholars might refer to these constructions as phrases instead of compounds because they do not, for instance, have an established meaning, we opted for the term compound because it is the term previously used in the literature on compound acquisition.

limited to the initial stage in the acquisition of compounds, at least for specific words. Nevertheless, children appear to understand the roles of modifiers and heads within compounds from quite early on, which rather suggests an early understanding of compound complexity. For example, English-speaking children can create novel compounds in spontaneous speech such as *cup egg* for a boiled egg as young as 2;0 (e.g. Clark, 1981). In these early compounds, the rightmost constituent usually indicates the category of the object. This has been taken as evidence that two-year-olds already understand the ordering of modifiers and heads (e.g. E. V. Clark, 1981, 1983; Clark, Gelman & Lane, 1985; Mellenius, 1997).

Children's understanding of compound complexity appears to be affected by frequency, both frequency within a language and an item-based frequency. Across languages, knowledge about compounding appears to be affected by the relative frequency of compounding as a word-formation device. In a language like English, compounds are very common and children reveal knowledge about the role of heads and modifiers in both production and comprehension tasks from the age of 2;0 or 3;0 (e.g. E. V. Clark, 1981, 1983; Clark, Gelman & Lane, 1985; Mellenius, 1997). In French or Hebrew, on the other hand, compounds are much less common and children spontaneously coin their first compounds around 5;0 and perform well in a comprehension experiment from around 3;0 on (Clark, 1998; Clark & Berman, 1987; cf. Nicoladis, 2002).

There is also evidence for an item-based frequency effect on children's understanding of compound complexity. Krott & Nicodadis (2005) and Nicoladis & Krott (2007) showed that four- to five-year-olds' understanding of the structure of familiar compounds appears to be affected by other compounds they know with the same constituents. They asked children to explain the meaning of familiar compounds. Children were likely to mention a compound constituent in their explanations if they knew a number of other compounds with the same constituent, which suggests that this knowledge helped them to recognize that the familiar compound is complex.

To be able to interpret novel compounds, one must not only understand the distinction between heads and modifiers, but also the possible types of subcategorizations, i.e. the possible thematic relations between heads and modifiers. What makes it difficult for children to learn how to infer relations is that there are many possible relations (e.g. a *chocolate muffin* is a muffin that HAS chocolate in it, a *mailbox* is a box FOR mail, a *cardboard box* is a box that is MADE OF cardboard). In addition, a compound can have several likely interpretations (e.g. *chocolate bowl* could be a bowl FOR chocolate, a bowl MADE OF chocolate and a bowl that looks LIKE chocolate because of its colour). Consequently, children must learn how to choose the correct or most likely relation for the compound. For some kinds of compounds, morphological structure can guide the interpretation (Lieber, 1983; Roeper & Siegel, 1978). Our focus is on novel noun–noun compounds such as *chocolate bowl*, which have no morphological markers that indicate how the modifier and head should be linked. In these cases, modifier–head relations need to be inferred on the basis of other knowledge.

Research on adults' conceptual combination has pointed to the importance of the availability of the modifier–head relation in interpretations of novel noun–noun combinations. Gagné & Shoben (1997) demonstrated that knowledge about the types of relations with which the modifier is typically used affects the ease of comprehending a novel combination. For example, the modifier *mountain* is frequently used with the relation LOCATED (e.g. *mountain bird*, as determined by Gagné & Shoben's corpus) but is not often used with the relation ABOUT (e.g. *mountain magazine*). Items that exhibit a high-frequency relation for the modifier (e.g. *mountain bird*) took less time to interpret than items that required a relation that was not frequently used with the modifier (e.g. *mountain magazine*). Furthermore, recent exposure to a phrase with the same modifier and relation as the to-be-interpreted phrase reduces the time required to interpret novel compounds (e.g. Gagné, 2001; Gagné & Shoben, 2002) as well as familiar compounds (Gagné & Spalding, 2004*b*). Although the bulk of the studies have demonstrated only relation priming with the modifier noun, one study has found evidence of relation priming with the head noun (Gagné & Shoben, 2002). Unlike the previous studies, this study used novel combinations that are ambiguous (e.g. a *student vote* can be a vote BY students or a vote FOR students).

Thus far, there has been little research on how children understand the relations between the modifier and head in compounds. It has been suggested that children as young as 2;6 understand that the head and modifier of a compound are related and they can use novel compounds to name several different relations (e.g. Clark *et al.*, 1985). Clark (1981) gives the example of a young child saying *fire-dog* to refer to a dog found at the site of a fire. However, while children might understand relatively early that compounds refer to two concepts that are somehow related (but see Berko, 1958), it appears that their understanding of what a likely relation is still differs from that of adults. For example, for adults a compound is an unlikely label of two objects that do not interact. Nicoladis (2003) found that 30% of the four-year-olds and 35% of the three-year-olds in her study accepted objects that did not interact (e.g. a dragon next to a box). Parault, Schwanenflugel & Haverback (2005) found that even six-year-olds' explanations of novel compounds are not necessarily adult-like. In 10% of their explanations heads and modifiers were not integrated with each other. Instead children used a coordination function, explaining, for instance, *book magazine* as 'a big magazine and a little book'. What remains unclear from previous research is how younger children infer relations for novel noun–noun compounds.

The aim of our research is to examine which sources of information children draw on when they infer a relation for a novel compound without being provided with a context[ii]. We can see several sources of information. Children might use other compounds with the same modifier or head as an analogical basis, being affected by the frequency of relations in these compounds, or they might choose the most frequent relation type in their vocabulary. They might also have a preference for particular relations. They might, for instance, prefer concrete relations such as HAS or LOCATED.

The present study examines the relevance of these different sources of information for children's interpretations of novel compounds, by asking them to explain the meaning of those compounds. First, we investigate the effect of the frequency of relations for specific modifier and head nouns by examining thematic relations in the sets of compounds that share the same constituents with a novel compound. We will refer to these sets as constituent families. For example, the modifier family for *chocolate* consists of compounds of the form *chocolate Y* such as *chocolate cake, chocolate milk* and *chocolate muffin*, while the head family of *bread* consists of compounds of the form *X bread* such as *banana bread, wheat bread* and *sandwich bread* (see also De Jong, Feldman, Schreuder, Pastizzo & Baayen, 2002; Krott, Baayen & Schreuder, 2001). Previous findings suggest that adults' default interpretation of a novel noun–noun combination proceeds by analogy with the modifier family and to some extent with the head family of the novel combination, and that even the processing of familiar compounds is affected by these families (see above). Given the findings of Krott & Nicoladis (2005), we assume that children build up patterns that relate to head families or modifier families (e.g. modifier + *muffin* or *chocolate* + head) and that they can use those to infer relations for new compounds. The present study therefore investigates children's inference processes for novel compound interpretations and compares them with those of adults.

[ii] It has been argued that the meaning of a novel compound is determined by the context it is used in and should therefore not be studied in isolation (e.g. H. H. Clark, 1983). Although it is certainly true that pragmatics and context do play a role in the interpretation of compounds, past research has shown that the ease of processing a compound in a discourse context is still influenced by its out-of-context meaning (Gagné & Spalding, 2004*a*). It has been argued that the function of context is to help readers narrow down a likely set of relationships between the two constituents of a compound to the correct one (e.g. Gagné & Spalding, 2004*a*; Gerrig & Murphy, 1992), but this likely set of relationships needs to be established first. We have opted to present our stimuli in isolation and not in context in order to study whether children can arrive at a meaning without contextual clues.

If children's interpretations are already based on similar inference processes as those of adults, then we expect them to rely more strongly on modifier families than on head families. To directly compare family effects on adults' and children's interpretations, we tested both participant groups on the same compounds and with the same methodology. Note that children's and adults' constituent families might not only be different in size. Earlier studies suggest that children do not necessarily understand familiar compounds the way adults do (Berko, 1958; Krott & Nicoladis, 2005; Nicoladis & Krott, 2007). We cannot be sure that children fully understand that a *lunch box* is a box that one uses to bring lunch in. They might only understand that the box has something to do with lunch but not exactly what that something is. If that was the case, one would not expect constituent families to have any effect on children's interpretations of novel compounds. We assume, however, that the knowledge of different kinds of boxes and the consistency of modifier–head relations in the compounds used to refer to these boxes (*crayon box, cereal box, treasure box* etc.) helps children to understand the relations in these compounds.

The effect of constituent families for children might be twofold. As shown in Krott & Nicoladis' (2005) study, children appear to have a better understanding of the thematic relations in familiar compounds that are part of larger families. A larger family means a better understanding of the relations of the family members as well as a larger analogical basis for the inference of the most likely relation for the novel compound. For example, if a child knows a large modifier family of *chocolate* and if many members of the family (e.g. *chocolate milk, chocolate cake* and *chocolate muffin*) use the relation HAS, he/she might have an increased understanding of the relation HAS in these compounds and he/she might have a larger analogical basis to predict that compounds starting with *chocolate* usually refer to something that HAS chocolate. The findings of Krott & Nicoladis (2005) suggest that four- to five-year-olds have developed such analogical bases. We therefore focus in the present study on four- to five-year-olds.

The second factor affecting relation selection for novel compounds that we consider in the present study is the overall frequency of modifier–head relations in children's vocabulary. If children are more successful with frequent relations than infrequent relations and if they use frequent relations when expected to use other relations, then this would be evidence that they (over)generalize one or more frequent relations. That would be consistent with the hypothesis that children have a general pattern for compounds.

An alternative or additional factor on relation selection is an overall preference for one or more particular relations. In a study that investigated children's interpretations of Swedish compounds, Mellenius (1997) found evidence that children used the LOCATED relation almost as a default relation (e.g. *djuret bor på pälsen* 'the animal lives on the fur' as an explanation for *pälsdjur* 'fur animal'). Because not much is known yet about whether some relations might be more difficult than others for children, we will analyze their success rate and overuse of particular relations.

METHOD

To study what affects compound interpretations, it is first necessary to adopt a categorization scheme for classifying and identifying relations. Linguists and psycholinguists have identified between ten and twenty common relation categories that capture the majority of semantically transparent compounds (Downing, 1977; Gleitman & Gleitman, 1970; Kay & Zimmer, 1976; Lees, 1960; Levi, 1978; Warren, 1978)[iii]. Examples of relations in English compounds include MADE OF (*paper bag* = bag made of paper), FOR (*computer screen* = screen for a computer), and HAS (*chocolate muffin* = muffin that has chocolate in it). We based our relation codings scheme on that of Levi (1978), although we used slightly different categories and category names. A full list of the codes plus examples, based on the codes of the child compound list described below, is shown in Appendix B.

Some scholars have argued that such finite sets of relations are too restrictive. One example that is often cited for this argument is *apple-juice seat*, which Downing (1977) reports as having been used by a friend to refer to a particular place at a table. However, this example and Downing's work in general does not preclude the use of a set of common relations in psychological research. Even though people can (especially with the aid of context) potentially create a wide variety of interpretations, this does not entail that one cannot identify a set of relations that tend to exist for most.[iv] Indeed, even Downing (1977: 836) observed that 'a small set of relationships is generally favored; and the appropriateness of a given relationship is also dependent on its permanence, its predictability in context, and on the semantic class of the head noun' and she lists a set of twelve categories that she believes can account for most compounds (e.g. Composition – *stone furniture*, or Time – *summer dust*).

[iii] Note that the meanings of semantically opaque compounds are not combinations of the meanings of the constituents. Therefore modifier–head relations do not exist for these compounds.

[iv] There is empirical evidence to suggest that, even when interpreted out of context, compounds do not have a wide range of meaning (Štekauer, 2005). Therefore, a theory of meaning predictability is feasible and stands in contrast to previous claims that the number of possible meanings for a novel word out of context is potentially infinite.

Participants

Twenty-seven monolingual British-English-speaking children (age range 4;9 to 5;8, mean 5;4, *SD* 0;3) took part in the experiment. All were deemed typically developing by their parents and teachers. When asked for interpretations of novel compound words, three of the children responded mostly with explanations that did not define the word. For instance, a boy (age 5;4) explained the meaning of *snow seat* as 'uhm, there it's snowing, and it – all boys or girls come out to play outside and make snowballs and throw it at each other'. Because these children were 5;4 or older, one would expect them to be able to identify and relate the constituents of the compounds. Their answers therefore indicate that they did not understand the task very well. Two other children always responded with the same relation (FOR or MADE-OF). We excluded the data from these five children from all analyses.

Apart from the children, thirty-six adults (age 17–26), mostly undergraduate students of the University of Birmingham and all native speakers of British English, took part in the experiment.

Material

We constructed thirty novel noun–noun compounds using constituents that we expected children to know (see Appendix A). Novel compounds can be interpreted in various ways and constituent families support these interpretations to different degrees. In order to investigate whether the family support has an effect, we needed one likely interpretation for each compound against which we could compare participants' (both adults' and children's) responses. We determined likely interpretations for the compounds by presenting our adult participants with randomized lists of the thirty compounds and asking them to write down their interpretations. We selected the relations that dominated their responses as the most likely interpretation (see procedure below and Appendix A). Choosing the adults' preferred relations as the comparison relations also allowed us to directly compare children's and adults' interpretations. The next step was to determine the support that the dominant relations receive from modifier and head families. Because children and adults differ in terms of vocabulary, we needed to determine these families for both groups separately. For the children, we amended a compound database created for Krott & Nicoladis' (2005) study. The original database lists about 2500 North-American compounds that children are likely to know. It is composed of: (a) all compounds occurring in the transcripts of the CHILDES database (MacWhinney, 2000); (b) compounds taken from the CELEX lexical database (Baayen, Piepenbrock & Gullikers, 1995),[v] but only those likely to be known by children; plus (c) further compounds that were added by two native North-American-English speakers (for further details see Krott & Nicoladis, 2005). Two native British speakers deleted from this list all American compounds and added British equivalents and missing compounds. The final list contained 2118 British compounds.

For adults, we created a list of modifier and head families using the CELEX database (Baayen *et al.*, 1995). Two native English speakers deleted from this list all compounds that they did not know. Subsequently they and a fluent English speaker coded the thematic relations of all head and modifier families in both the child list (533 compounds) and the adult list (752 compounds) using our coding scheme in Appendix B. For the child compound list, the three coders agreed on the relation for 82·9% of the compounds (Fleiss' $K=0{\cdot}84$, $p<0{\cdot}001$). For the adult compound list, all coders agreed on the code for 78·7% of all compounds (Fleiss' $K=0{\cdot}80$, $p<0{\cdot}001$). In case of disagreement, the majority choice was taken as the correct relation. In case of no majority, coding differences were resolved by discussion.

The experimental compounds were constructed with varying head family and modifier family support in the child compound list because we were interested in effects of modifier and head families on children's interpretations. A third of the compounds had dominant modifier–head relations that were supported by both a bias of the modifier family in the database (mean number of supporting family members 6·0, *SD* 1·6) and a bias of the head family in the database (mean number of supporting family members 8·5, *SD* 5·4) (condition High–High). Another third (condition Low–High) had dominant relations that were supported by only a bias of the head family (modifier family: 1·0, *SD* 1·7; head family: 10·0, *SD* 4·0) and the remaining third (condition High–Low) had dominant relations that were supported only by a bias of the modifier family (modifier family: 6·6, *SD* 1·4; head family: 0·4, *SD* 0·7). Note that there was no condition Low–Low because it is very difficult to construct easily interpretable compounds that do not have any support of modifier or head families.

We checked the contrasts between the conditions by counting supporting family members that occur in the British part of the CHILDES database (MacWhinney, 2000). This part of CHILDES contains various types of transcripts from 160 children across the United Kingdom, most of which were either of the same age or younger than the children tested in the present study. Twenty-four were older (age 7), but the transcripts of those children amounted to less than 2% of the transcribed speech. Table 1 lists

[v] The CELEX database contains 4843 noun–noun compounds taken from the *Oxford Advanced Learner's Dictionary* and the *Longman Dictionary of Contemporary English*.

TABLE 1. *Mean number of supporting modifier and head family members in CHILDES and according to parental report for novel noun–noun compounds with varying Modifier Support and Head Support (High–High, High–Low, Low–High; standard deviations in parentheses)*

	Mean modifier family size		Mean head family size	
Family bias	CHILDES	Parental report	CHILDES	Parental report
High–High	2·5 (1·8)	2·3 (1·4)	2·5 (2·1)	2·0 (1·7)
Low–High	0·3 (0·4)	0·2 (0·5)	3·1 (3·0)	2·1 (1·9)
High–Low	3·0 (2·4)	2·2 (1·4)	0·6 (1·3)	0·02 (0·05)

the mean number of modifier and head family members for the three subsets of compounds. We checked the differences between the conditions by conducting analyses of variance for the numbers of supportive family members (i.e. supportive modifier or head family) with family bias (High–High, High–Low and Low–High) as the independent variable. In case of the modifier contrast, there was a significant main effect of family bias on supportive modifier family ($F(2, 27)=7{\cdot}0$, $p=0{\cdot}004$, $\eta_p^2=0{\cdot}34$). Planned comparisons of the conditions confirmed that the supportive modifier family of condition Low–High was significantly smaller than the supportive modifier family of conditions High–High and High–Low ($F(1, 27)=13{\cdot}6$, $p=0{\cdot}001$, $\eta_p^2=0{\cdot}34$), and that there was no difference between conditions High–High and High–Low ($F<1$). In case of the head contrasts, there was a significant main effect of family bias on supportive head family ($F(2, 27)=3{\cdot}4$, $p=0{\cdot}049$, $\eta_p^2=0{\cdot}20$). Planned comparisons of the conditions confirmed that the supportive head family of condition High–Low was significantly smaller than the supportive head family of conditions High–High and Low–High ($F(1, 27)=6{\cdot}4$, $p=0{\cdot}017$, $\eta_p^2=0{\cdot}19$), and that there was no difference between conditions High–High and Low–High ($F<1$). Thus, the compounds occurring in CHILDES confirmed the contrasts of our experimental design.

To investigate the effect that overall frequencies of modifier–head relations might have on children's interpretations, we estimated the frequency of thematic relations in children's compound vocabulary. For that, we gathered all noun–noun compounds in both child-speech and child-directed speech in the British part of the CHILDES database (629 compounds). The thematic relations of the compounds were coded by five native English speakers (British), using the same coding scheme as above (Fleiss' $K=0{\cdot}45$, $p<0{\cdot}001$). The majority relation was taken as the correct relation. If there was no majority relation, the compound received the coding 'OTHER'.

TABLE 2. *Family members of* egg bag *on parental checklist*

egg ________	________ bag
egg carton*	sandwich bag*
egg timer*	lunch bag*
eggbeater*	shopping bag*
eggcup*	sports bag*
egg white	shoe bag*
egg yolk	handbag*
eggshell	punching bag*
egg noodles	saddlebag
egg salad	teabag
	canvas bag
	cloth bag
	sandbag
	ice bag

* Indicates the same relation as the dominant relation of the target compound *egg bag* (FOR) according to our compound database.

Parental vocabulary report. We asked the parents of the children to give us information about the actual vocabulary knowledge of the children. For that, we presented to them a vocabulary checklist that contained the constituents of the experimental compounds together with the constituents' modifier and head family members taken from our compound database. We asked the parents to indicate which words they thought their children knew well enough to say spontaneously. This way we were able to determine modifier and head families more accurately than on the basis of our compound database or CHILDES. For example, for each child, we knew his/her modifier family of the experimental compound *egg bag*, i.e. all the compounds he/she knew that contain the modifier *egg*. Table 2 lists the compounds that were included in the questionnaire for the novel combination *egg bag*. The parental report also provided the number of compounds in the modifier family that match the dominant relation of *egg bag* (bag FOR eggs), i.e. the size of the supportive modifier family. Similarly, we determined the head family of *egg bag* for each child by gathering all compounds that the child knew with the head *bag* and determined the number of compounds that contained the dominant relation (FOR), i.e. the size of the supportive head family. The parents of the twenty-two children we included in our analyses had filled in the vocabulary list.

Table 1 lists the mean sizes of the supportive families according to parental report for the three experimental compound sets (High–High, High–Low and Low–High). Similar to the examination of the families occurring in CHILDES, we checked whether the contrasts of our experimental conditions that we had determined on the basis of our compound database were confirmed by the estimates of the parental report. As for the CHILDES compounds, we conducted analyses of variance for the reported numbers of supportive family members (i.e. supportive modifier or head family, averaged over participants) with family bias (High–High, High–Low and Low–High) as the independent variable. Results again confirmed the contrasts initially based on the database. In case of the modifier contrast, there was a significant main effect of family bias on supportive modifier family ($F(2, 27) = 10{\cdot}1$, $p = 0{\cdot}001$, $\eta_p^2 = 0{\cdot}43$). Planned comparisons of the three conditions confirmed that the supportive modifier family of condition Low–High was significantly smaller than the supportive modifier family of conditions High–High and High–Low ($F(1, 27) = 20{\cdot}2$, $p < 0{\cdot}001$, $\eta_p^2 = 0{\cdot}43$), and that there was no difference between conditions High–High and High–Low ($F < 1$). In case of the head contrasts, there was a significant main effect of family bias on supportive head family ($F(2, 27) = 7{\cdot}2$, $p = 0{\cdot}003$, $\eta_p^2 = 0{\cdot}35$). Planned comparisons of the three conditions confirmed that the supportive head family of condition High–Low was significantly smaller than the supportive head family of conditions High–High and Low–High ($F(1, 27) = 13{\cdot}9$, $p = 0{\cdot}001$, $\eta_p^2 = 0{\cdot}34$), and that there was no difference between conditions High–High and Low–High ($F < 1$).

Procedure

Children were tested individually. The experimenter, a native speaker of British English, asked them to explain to her the meaning of the experimental compound words. She introduced them to the task by explaining that we say *coffee cup* because it is a cup for coffee. She then asked the children what *bear trousers* are and corrected the child if he/she did not give a response that explained the relation between head and modifier (e.g. 'trousers for bears'). Following these examples, the experimental compounds were presented to each child in a different random order. For each compound, the experimenter asked 'What does X mean?' If a child did not respond with an explanation of the meaning of the compound, she repeated the same question later in the experiment. As mentioned, adults were tested with a written version of the task, i.e. they were given instructions and the two examples above in writing and asked to write down their interpretations.

Both children's and adults' interpretations were coded by two native English speakers and one fluent English speaker, using the coding scheme in Appendix B. For example, one adult explained the meaning of *fire foil* as 'foil for making fire', which was coded with the relation FOR, while one child said 'it is foil, but it is made out of fire', which was coded as

TABLE 3. *Mean number of dominant relation responses* (*out of* 10) *for combinations of relational biases in modifier family and head families for novel compounds* (*High–High, High–Low, Low–High*)

Family bias	Mean	Standard deviation
High–High	4·5	1·9
Low–High	4·2	1·7
High–Low	3·2	1·7

MADE OF. Another adult explained *birthday room* as 'a room in which birthdays are celebrated', while a child said 'you got a room for your birthday', which were both coded with the relation FOR. The agreement among the three coders for children's responses was 71% (Fleiss' $K = 0{\cdot}68$, $p < 0{\cdot}001$), for adults' responses it was 86% (Fleiss' $K = 87{\cdot}1$, $p < 0{\cdot}001$). The codes for the remaining responses were decided by taking the majority code (15% of all child responses and 13% of all adult responses) or by discussion.

RESULTS

Of all child responses, 39·7% were dominant interpretations, i.e. interpretations that were preferred by adults in our adult pretest of the materials. In 18·5% of the responses children reversed the roles of modifiers and heads in their explanations (e.g. 'an animal what has got a helmet on' for *animal helmet* or 'it's cheese and I think it is made out of fish' for *cheese fish*), suggesting that even five-year-olds still have some problems distinguishing between heads and modifiers when exposed to novel compounds.

Effect of modifier and head families

We first examined whether children base their interpretations on their knowledge of existing noun–noun compounds with either the same modifier or head noun while interpreting the novel compound. We averaged the number of responses with dominant relations for the different conditions of our experiment by participants and conducted an analysis of variance with family support for the dominant relation (High–High, High–Low and Low–High) as the independent variable and number of responses with the dominant relation as the dependent variable. Table 3 lists the mean number of those responses for the conditions. The analysis showed a main effect of family support ($F(2, 42) = 9{\cdot}0$, $p = 0{\cdot}001$, $\eta_p^2 = 0{\cdot}30$). Planned comparisons revealed that the condition High–Low led to fewer responses with the dominant relation than conditions High–High and Low–High ($F(1, 21) = 18{\cdot}7$, $p < 0{\cdot}001$, $\eta_p^2 = 0{\cdot}47$), indicating that a larger number of supportive head family members led to more responses with the supported relation. In contrast, comparing Low–High with High–High and High–Low showed no significant difference ($F(1, 21) = 1{\cdot}5$, $p = 0{\cdot}237$, $\eta_p^2 = 0{\cdot}066$), indicating that the number of supportive modifier family members had no significant effect on the responses.

Parental report and comparison with adults

Our finding for children's interpretations is different from that found for adults in previous studies, which revealed an influence of adults' experience with modifiers in existing combinations (e.g. Gagné, 2001; Gagné & Shoben, 2002). We therefore directly compared effects of modifier and head families on adults' and children's interpretations. We conducted a generalized linear mixed-effect (multilevel) analysis (e.g. Pinheiro & Bates, 2000) because it allows us to predict children's and adults' responses on the basis of different modifier and head families, combining children's and adults' responses in a single analysis. It also has the advantage that we can predict children's responses on the basis of individual modifier and head families, using parental reports. We included the match with the dominant relation (matching or not matching) as the dependent variable into the analysis, participants and items as crossed random effects, and the number of supportive head family members (head support), the number of supportive modifier family members (modifier support), and age group (children vs. adults) as fixed factors. The results revealed significant main effects of modifier support ($F(1, 1734) = 55{\cdot}6$, $p < 0{\cdot}001$)[vi] and age group ($F(1, 1734) = 15{\cdot}3$, $p < 0{\cdot}001$), and the interaction head support × age group ($F(1, 1734) = 6{\cdot}0$, $p = 0{\cdot}01$). Adults responded more often with the dominant relation (70%) than children (39·7%), and a larger modifier support increased the number of matching responses.

Because of the interaction of head support and age group, we conducted separate mixed-effect analyses for children and adults. An analysis of children's responses with response (dominant interpretation versus other interpretation) as the dependent variable, participants and items as crossed random effects, and head support and modifier support as fixed factors showed that head support was a good predictor of children's responses, with larger support leading to more responses that matched the dominant relation ($F(1, 657) = 6{\cdot}5$, $p < 0{\cdot}001$). However, modifier support was not a

[vi] In this and all further analyses, we used the 'Laplace' algorithm to calculate the parameters of the model.

good predictor of responses ($F(1, 657) = 1{\cdot}7$, $p = 0{\cdot}19$). This shows that children's interpretations were affected by their knowledge of supportive head family members, which is in line with the analysis of variance above. The lack of a significant effect of modifier support for children needs to be viewed in the light of the main effect of modifier support and the lack of an interaction of modifier support and age group in the omnibus analysis above. In particular, the random effect of participants and the fixed factor modifier support shared explained variance in the analysis of the child responses. Therefore, we fitted another model, which did not include participants as random factor. For this model, both head support ($F(1, 657) = 9{\cdot}1$, $p < 0{\cdot}001$) and modifier support ($F(1, 657) = 11{\cdot}9$, $p < 0{\cdot}001$) predicted children's responses. The influence of the modifier family is also reflected in the correlations between number of responses with the dominant interpretation and modifier support ($r = 0{\cdot}12$, $p = 0{\cdot}003$). The magnitude of this correlation increases when items with no modifier support (and, hence, with no basis for analogy) are removed ($r = 0{\cdot}19$, $p = 0{\cdot}001$). Likewise, there is a positive correlation between the number of responses with the dominant interpretation and head support ($r = 0{\cdot}19$, $p < 0{\cdot}001$), and without items for which there is no head support ($r = 0{\cdot}20$, $p = 0{\cdot}002$).

The analysis using the adult data alone showed a different pattern; modifier support ($F(1, 1077) = 8{\cdot}0$, $p < 0{\cdot}001$) but not head support was a successful predictor of the responses ($F(1, 1077) = 1{\cdot}9$, $p = 0{\cdot}17$). This finding is consistent with previous adult studies, which report that the speed of interpretation was primarily affected by their experience with modifiers in other noun–noun combinations (Gagné, 2001).

In sum, our analysis reveals that children and adults draw on similar but different knowledge when interpreting novel noun–noun compounds. Both groups use familiar compounds as an analogical base when interpreting novel compounds. However, while adults base their interpretations primarily on knowledge of modifiers in familiar combinations, children base their interpretations primarily on knowledge of heads in familiar combinations.

Overall relation frequency

After having shown the importance of the supportive head family on children's interpretations, our second question was whether children's interpretations were also affected by the overall frequency of relations in the children's compound vocabulary. Did children have a higher success rate with a relation that they were more familiar with? Table 4 shows the distribution of the most frequent thematic relations of compounds that occur in the CHILDES database (summed frequencies in child-speech and

TABLE 4. *Frequency of modifier–head relations in CHILDES compounds and experimental stimuli, as well as success rate with dominant relation and usage if non-dominant relation for children and adults*

Relation	CHILDES (%)	Stimuli		Success rate with dominant relation (%)		Usage if non-dominant relation (%)	
		#	%	Children	Adults	Children	Adults
FOR	40·6	16	53·3	40·9	70·1	4·2	8·3
LOCATED	7·6	1	3·3	59·1	100·0	14·6	0·7
MADE OF	6·6	6	20·0	18·9	65·7	2·7	5·7
PART	5·0	1	3·3	45·5	77·8	0·0	0·1
HAS	4·4	4	13·3	59·1	54·9	17·1	7·7
BE	3·8	1	3·3	31·8	94·3	5·9	0·5
USE	1·9	0	0·0	n/a	n/a	0·9	1·0
LIKE	1·1	0	0·0	n/a	n/a	1·2	4·2
DURING	0·9	1	3·3	63·6	94·4	0·2	0
OTHERS	28·0	0	0·0	n/a	n/a	0·5	8·0

child-directed speech)[vii] and of the experimental stimuli.[viii] Among the compounds in CHILDES, the FOR relation strongly dominates (e.g. *earring* = ring for an ear). If children understand relations better that occur more often in their vocabulary, then they should perform better when explaining a FOR relation than when explaining other relations. For each relation we therefore calculated a success rate, averaging the success rates over children. For example, for sixteen compounds the FOR relation was the dominant and therefore the expected relation. We calculated for each child how many responses out of the sixteen were actually FOR responses and averaged this number over all children. The column labeled 'success rate with dominant relation – children' in Table 4 lists success rates for all relations occurring in our experimental stimuli. As shown in the table, children's success rate was highest for DURING, HAS and LOCATED, followed by PART, FOR and BE, which does not reflect the frequency distribution in the CHILDES compounds. To test this difference, we added the relation frequency in CHILDES as an additional fixed factor to

[vii] Taking only compounds that occur in children's utterances in CHILDES leads to a very similar distribution.

[viii] Note that the group OTHERS slightly differs for children's responses and CHILDES. In the case of children, these are responses that either could not be classified or relations that occurred very seldom. In the case of CHILDES, these are either compounds that are opaque, compounds with relations that do not fall into the main categories, i.e. that rarely occur (in less than 2% of all compounds), or compounds that did not receive a majority code from our coders.

our mixed effect model; we conducted an analysis with response (dominant interpretation versus other interpretation) as a dependent variable, participants and items as crossed random effects, and head support, modifier support, age group (children vs. adults), and CHILDES Frequency as fixed factors. There was no effect of CHILDES Frequency ($F < 1$). Thus, we can conclude that children's success in responding with the dominant relation is not affected by their familiarity with the relation as reflected by the overall frequency of the relation in the CHILDES database.

Relation preference for children and adults

Success rates in Table 4 suggest that children were not equally successful with different relations, and their success rates were different from those of adults. Out of the relations that appear more than once in our experimental material, they seem to be quite successful with the relations HAS (59·1%) and FOR (40·9%), while less successful with MADE OF (18·9%). Adults appear to have success rates that are much more equal (HAS 54·9%, FOR 70·1%, MADE OF 65·7%). We tested whether the relation of a compound has an additional effect on children's and adults' responses (apart from family support effects) by adding relation (FOR, LOCATED, MADE OF, PART, HAS, BE, DURING) as an additional fixed factor into our mixed effect analysis. Results confirmed again main effects of modifier support ($F(1, 1722) = 34{\cdot}6$, $p < 0{\cdot}001$), head support ($F(1, 1722) = 14{\cdot}3$, $p < 0{\cdot}001$) and age group ($F(1, 1722) = 14{\cdot}4$, $p < 0{\cdot}001$), as well as an interaction between age group and head support ($F(1, 1722) = 6{\cdot}6$, $p = 0{\cdot}01$). In addition, there was an interaction between age group and relation ($F(6, 1722) = 8{\cdot}3$, $p < 0{\cdot}001$), indicating that adults and children were not equally successful with different relations.

Due to the interaction, we conducted separate generalized mixed model analyses for children and adults. The analysis of the children's responses showed main effects of head support ($F(1, 651) = 7{\cdot}8$, $p = 0{\cdot}01$) and relation ($F(1, 651) = 4{\cdot}3$, $p < 0{\cdot}001$). Comparing children's responses to compounds with the comparison relation FOR against compounds with other relations, and restricting comparisons to the two relations that occur more often than once in our stimuli (HAS and MADE OF) revealed that children responded more often with the dominant relation HAS than with FOR ($\beta = 0{\cdot}11$, $z = 2{\cdot}5$, $p = 0{\cdot}01$) and more often with the dominant relation FOR than with MADE OF ($\beta = 1{\cdot}1$, $z = 3{\cdot}1$, $p = 0{\cdot}002$). Adults, on the other hand, performed equally successful across the various relation-types; the adult analysis showed only a marginal effect of relation ($F(1, 1071) = 2{\cdot}1$, $p = 0{\cdot}05$), with less successful responses with relation HAS than with FOR ($z = -2{\cdot}2, = 0{\cdot}03$).

Does this finding mean that children have fewer problems explaining compounds with a HAS relation than compounds with a FOR or MADE OF relation because they have a good command of the HAS relation, while they have a weaker command of FOR or MADE OF? Or does their success with HAS result from an overuse of HAS? To answer this question, we examined which relations children (and adults) use when they respond with a non-dominant relation. We calculated the average percentage of responses with a particular relation out of the total number of possible responses with this relation when it was NOT the dominant relation. For example, for fourteen compounds the FOR relation was not the dominant response. We calculated for each child how many responses out of these fourteen were actually FOR responses and therefore non-expected responses. The column 'usage if non-dominant relation – children' in Table 4 lists the mean percentage for each relation when used unexpectedly, i.e. when it is not the dominant relation. Children tended to use the relations HAS and LOCATED more often than adults when producing an unexpected response, while they tended to use FOR and MADE OF less often than adults. To test these differences, we conducted generalized linear mixed effect analyses for the four relations FOR, MADE OF, LOCATED and HAS. For each relation, we based the analysis on the subset of compounds for which the relation would be non-dominant. For example, for the FOR relation we only considered compounds for which FOR would be the non-dominant response. We conducted a mixed effect model with response (e.g. FOR vs. NOT FOR) as dependent variable, participants and items as crossed random effects and age group as fixed factor. There were significant effects of age group for all four relations, confirming that children used FOR ($F(1, 810)=5{\cdot}0$, $p=0{\cdot}03$) and MADE OF ($F(1, 1390)=16{\cdot}0$, $p<0{\cdot}001$) less often unexpectedly (i.e. when the dominant relation was a different relation) than adults and that they used LOCATED ($F(1, 1680)=53{\cdot}1$, $p<0{\cdot}001$) and HAS ($F(1, 1506)=15{\cdot}4$, $p<0{\cdot}001$) more often than adults.

Taking success rates and non-dominant responses together, the following pattern emerges. While adults have a similar success rate for relations HAS, MADE OF and FOR, children have a higher success rate for the HAS relation than for the FOR and MADE OF relations. Children also use the HAS relation more often than adults when responding with a non-dominant relation. Together, these results suggest that the children overused HAS (cf. Clark & Berman, 1987). The same seems to be the case for the LOCATED relation (see also Mellenius, 1997), although children's high success rate for LOCATED (59·1%) is based on a single compound (*side ache*) and therefore needs to be treated with caution. Furthermore, the children used the FOR and MADE OF relations less often than did the adults, which suggests that children underused these relations.

GENERAL DISCUSSION

The purpose of this study was to investigate various ways in which frequency might affect children's selection of thematic relations for novel noun–noun compounds. We addressed the following questions: Do children use their knowledge of modifier and head families and the support for thematic relations in these families when interpreting compounds? If yes, do they do this in the same way adults do? In addition, we investigated whether children's success rates and responses with non-dominant relations are affected by frequency of modifier–head relations in their vocabulary and whether there are relational preferences for children that are different from those for adults.

In our discussion, we highlight four findings, in turn: (1) both children and adults responded with highly frequent relations within compound families when interpreting novel compounds; (2) children and adults relied on families related to different constituents; (3) children's interpretation of novel compounds was not related to the overall frequency of thematic relations in compound words; and (4) in addition to using the most frequent thematic relation of a family, children had a strong tendency to use HAS or LOCATED.

The main finding of our study was that both children's and adults' interpretations of novel compounds was related to high-frequency relations in the compound families. This finding is consistent with other studies showing item-based effects in compounding for children (Krott & Nicoladis, 2005; Neijt, Krebbers & Fikkert, 2002; Nicoladis & Krott, 2007) and adults (Krott *et al.*, 2001). Taken together, these studies support the following account of compound acquisition. We assume that children start to learn compound words initially as individual words, with no knowledge of their internal structure (cf. Berko, 1958; Berman, 1987). As they learn clusters of similar words, children start to understand that compounds have an underlying structure. Initially, this structure is a combination of a specific constituent (e.g. a specific head such as *muffin*) and a specific modifier–head relation (e.g. HAS), which results in a structure such as modifier + *muffin* (HAS). Because specific constituents can be combined with various relations (see *breakfast muffin* = muffin to be eaten for breakfast), closely related structures develop (modifier + *muffin* (HAS), modifier + *muffin* (FOR)), which need to be combined into a single structure (modifier + *muffin* (RELATION)). We have referred to the latter as constituent families. Eventually children learn an abstract pattern (e.g. modifier + head (RELATION)).

Our results concerning adults' compound interpretations suggest that knowledge about the head noun's constituent families gives way to knowledge about the modifier's constituent families and this knowledge is used alongside an abstract compound pattern. This is not an isolated phenomenon. Other researchers have noted that adults can have some more-or-less frozen expressions in their vocabularies, even when they can use the components of those expressions separately in other contexts (Barlow & Kemmer, 2000; Bybee & Hopper, 2001; see also Di Sciullo & Williams, 1987; Langacker, 1987). For example, some expresssions like 'would you' and 'didn't you' might be stored as single units in the lexicon but elements of these constructions (i.e. 'would', 'did', 'not' and 'you') are also stored as units (Bybee & Hopper, 2001).

This account of compound acquisition predicts that children who are still treating compounds as chunks (at least for a majority of compounds) should not be influenced by modifier or head families when inferring thematic relations of novel compounds. Future research will need to confirm this. It is also unclear at what age children have fully developed an abstract compound pattern. Earlier studies on compound production and comprehension suggest that this knowledge is already present for two-year-olds (e.g. Clark *et al.*, 1985). It is important to note, though, that in these studies young children still make a considerable number of errors. For example, Clark *et al.* (1985) asked children to choose the referent for a novel compound (e.g. *mouse-hat*) among the pictures of a mouse, a hat, a hat on a mouse and a hat on a fish. Although significantly above chance level, only 50% of two-year-olds correctly chose the hat on a mouse. Such a performance level can be explained by an approach that assumes that responses do not reflect an abstract knowledge about subcategorization, but that each response depends on the individual knowledge of the child about compounds that have the same modifier (mouse) or head (hat) as the novel target compound. Furthermore, our study revealed that five-year-olds still make reversal mistakes when identifying modifiers and heads in novel compounds (see Nicoladis, 2002, for similar reversals in production in three- and four-year-olds). This finding challenges the assumption that two- and three-year-olds have a complete and robust knowledge of the roles of modifiers and heads, i.e. that they have developed an abstract noun–noun pattern. The finding rather adds to the evidence that this general knowledge develops very slowly and possibly item by item (Krott & Nicoladis, 2005; Nicoladis, 2003; Nicoladis & Krott, 2007). Nevertheless, it is difficult to assess whether children's mistakes in such tasks are due to the lack of a general abstract compound pattern or whether they stem from a different source.

Our account of compound acquisition is very similar to that of Usage-based accounts of syntactic acquisition (e.g. Goldberg, Casenhiser & Sethuraman, 2004; Tomasello, 2000). Usage-based accounts suggest that children initially learn words in specific strings (e.g. 'I love you'), gradually construct underlying schemas with slots that still contain specific words (e.g. 'I ____ you'), and finally generalize to an abstract pattern (e.g. Subject-Verb-Object or Agent-Action-Patient). Although several researchers using a Usage-based framework have focused on children's acquisition of syntactic constructions (e.g. Goldberg *et al.*, 2004; Lieven, Behrens, Speares & Tomasello, 2003), Usage-based theories do not restrict constructions to the level of syntax. A construction can be any high-frequency string (Bybee & Hopper, 2001) or any pairing of form and meaning that is not predictable from its constituents or other constructions (Croft, 2001; Goldberg, 2006). Given these definitions, the term construction also refers to complex words (e.g. Bybee, 1985), including compounds (see also Di Sciullo & Williams, 1987; Plag, 2006).

One should note, however, a difference between typical Usage-based studies and the present study on compound relations. While typical Usage-based explanations are seen as an alternative to a generative approach that assumes abstract syntactic knowledge from early on, this would not make sense for thematic relations in compounds because there is no abstract 'default' relation that children need to learn. In this sense, our results do not present strong evidence against early abstract knowledge of thematic relations. At the same time, our findings are consistent with the idea that item-based knowledge is stored and used by children (and adults) when linking linguistic units to form unified representations.

A second finding from the present study was that children and adults relied on different constituents and their families in interpreting the meaning of novel compounds. Children's interpretations are affected primarily by the relational bias in head families. Although there is some evidence that they are also affected by the bias within modifier families, this evidence is inconclusive. In contrast to the children's data, the data from the adults indicate that adults draw on their experience with modifier families. Why do children and adults rely on different constituent families? The ability to interpret a noun–noun combination involves creating or identifying the concept denoted by the combination. We propose that there are at least two aspects to this process. First, one must understand that the combination is a member of the head noun category, which is the subcategorization function of compounds (cf. Clark, Gelman & Lane, 1985). For instance, a *chocolate bowl* is a kind of bowl. Second, one must understand how the combined concept differs from other members of the category. To illustrate, a *chocolate bowl* is a bowl for chocolate, whereas a *fruit bowl* is a bowl for fruit. In these two examples, the modifier provides information about the purpose or function of the bowl and, in doing so, forms two subcategories of bowls. This second ability involves understanding of how the head noun is modified by the modifier concept. The first aspect of conceptual combination draws more heavily on knowledge about the head noun because it involves understanding the compound's category membership (it is a bowl). The second aspect draws more heavily on the modifier and how it can modify a category.

The finding that children tend to use their knowledge about heads therefore suggests that they use their knowledge about categories and categorization. The effect of the supportive head family on children's compound interpretations suggests that a child is more likely to pick a relation for the novel compound if the child knows a lot of other subcategories of the head with the same relation. For example, he/she knows a lot of other bowls and they all have the function of storing something. The finding that children appear to use knowledge about modifier families to a lesser degree might mean that they do not know very much yet about the possible modifications that a specific modifier (e.g. *chocolate*) can create within a compound. Due to their limited vocabulary, children might not have discovered the predictive power of modifiers and their families, or children's small modifier families might not have a critical mass yet to affect their interpretations.

Another possibility is that preschool children are still developing their understanding of the roles of heads and modifiers and therefore focus their attention on the category to which the compound belongs (i.e. the head noun) because identifying the category is the first step of understanding an unfamiliar compound. In other words, overall our research suggests that preschool children do understand that modifier and head nouns play different roles, which is consistent with past research. However, when they determine how a modifier and head should be linked, they show only weak signs of using knowledge about compounds with the same modifier. It is unclear whether they do not have this knowledge or whether the complex task leads them to focus on head categories.

A third finding to highlight from this study is that the children's interpretations were not sensitive to overall relation frequency in compounds that they are exposed to or that they use themselves. The lack of a general frequency effect of relations in the vocabulary for five-year-olds is unexpected within a Usage-based framework. It might appear surprising that five-year-olds (and even adults) rely on abstract patterns in syntax (see Tomasello, 2000, 2003) and still rely on item-based knowledge when it comes to compounds (the present study). This difference might be due to the fact that the syntactic constructions they learn early are more frequent than compounds. An alternative explanation might be a larger variety of relations for compounds (FOR, HAS, LOCATED, MADE OF etc.), which makes the roles of compound modifiers and heads, especially the abstract roles of being a modifier or the superordinate category, more difficult to understand than those of constituents in syntactic phrases. Note that syntactic phrases most similar to noun–noun compounds, namely adjective–noun phrases, have a much stronger preference for a particular relation, i.e. the IS relation (a *red ball* is a ball that IS red and a *happy girl* is a girl that IS happy), than noun–noun constructs. Even very young children can generalize the meaning of a novel adjective (e.g. Klibanoff & Waxman, 2000).

Another finding of our study was that children, but not adults, tended to overuse HAS (and LOCATED) and to underuse FOR when interpreting compounds. Thus, high frequency alone might not predict children's productive use of some thematic relations (e.g. FOR). We consider briefly three possible explanations for this finding. First, one might argue that children's performance in our task is affected by their metalinguistic awareness, i.e. their competence to express their understanding of modifier–head relations. Children indeed sometimes seemed to struggle to find the right words when explaining a relation and, in general, it might be easier to explain a HAS or LOCATED relation than a FOR relation. However, twenty out of the twenty-two children explained at least once a FOR relation (mean 7·9 times), showing that they had the means of expressing FOR relations. Therefore, metalinguistic considerations alone cannot explain our finding. The second explanation concerns the context of acquisition; children might learn compounds in the context of both objects represented by the compound constituents being perceptually available. This is the case for HAS and LOCATED relations, but not necessarily for FOR relations. For example, while a *stepladder* is a ladder that always HAS steps and a *doormat* is located at a door, a *lipstick* is not always near lips. If the context of compound acquisition leads children to assume that compounds refer to two objects that are both perceptual available, they should be more likely to assume that novel compounds have HAS or LOCATED relations. A third possibility is that children prefer HAS and LOCATED relations because they are conceptually easier. HAS and LOCATED relations are concerned with the physical proximity of objects. FOR relations on the other hand refer to actions in which a head object acts on a modifier object. Because actions are concepts of a higher complexity, it might be more demanding for children to understand a FOR relation than a HAS or LOCATED relation. Children might either overuse HAS and LOCATED relations to avoid complex concepts or because they have not fully understood the FOR relations.

The preference for the perceivable relations HAS and LOCATED over the function FOR is in line with findings for children's development in classifying objects and extending names to objects. There is evidence that in these tasks children are initially strongly driven by perceptual properties (especially the shape) of an object, while the function of the object is used only later in development and depending on the form–function relation (e.g. Gentner, 1982; Landau, Smith & Jones, 1988).[ix] For example, in the classic study by Gentner (1982), children were introduced to two novel objects with two novel names. When they were asked to name a hybrid object that looked like one of the novel objects but had the same function as the other object, they preferred the name of the object that looked like the hybrid object. Our results suggest that children have the same preference for visually perceivable features (HAS and LOCATED versus FOR) when it comes to interpreting novel concept combinations. Therefore, children seem to be not only affected by their knowledge of head (and modifier) families when interpreting novel compounds, but also by relation perceivability.

We chose to present our stimuli without any linguistic or non-linguistic context that might make one or the other interpretation more likely to achieve a clearer picture of the effect of children's experience with compounds on their compound interpretations. Our results show that, at least four- to five-year-olds are not relying on context for their understanding of unknown compounds. They are able to use their knowledge of other compounds to infer an appropriate interpretation, similar to adults. Further studies will need to investigate how children integrate knowledge of constituent families and contextual information.

To conclude, our findings add to the accumulating evidence that the input frequency of linguistic constructions affects children's acquisition of these constructions. Our study investigated the role of thematic relations and distinguished two different types of frequency: local, i.e. item-based frequencies, as reflected by constituent families, and global frequencies across children's compound vocabulary. Our results suggest the effect of only the former, supporting the importance of item-based knowledge in language acquisition and language processing. We have also raised a number of possible variables that might apply to children's learning of any word strings, including syntax. One is the perceptual bias in interpreting relations, which echoes the perceptual bias in early name extension and object categorization.

REFERENCES

Allen, S. E. M. & Crago, M. B. (1996). Early passive acquisition in Inuktitut. *Journal of Child Language* **23**, 129–55.

Baayen, R. H., Piepenbrock, R. & Gullikers, L. (1995). The CELEX Lexical Database (CD-ROM). University of Pennsylvania: Philadelphia, PA: Linguistics Data Consortium.

Barlow, M. & Kemmer, S. (2000). *Usage Based Models of Language*. Stanford, CA: Center for the Study of Language and Information.

[ix] But see Nelson and her colleagues (e.g. Nelson, Russell, Duke & Jones, 2000), who argue that function can be used by younger children when extending names to novel objects. However, the plausibility of the function and the causal link between function and structure of an object influences the effect of function on name extension.

Berko, S. (1958). The child's learning of English morphology. *Word* **14**, 150–77.
Berman, R. A. (1987). A developmental route: learning about the form and use of complex nominals in Hebrew. *Linguistics* **25**, 1057–85.
Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Philadelphia: John Benjamins.
Bybee, J. L. & Hopper, P. (2001). Introduction to frequency and the emergence of linguistic structure. In J. Bybee & P. Hopper (eds), *Frequency and the emergence of linguistic structure*, 1–24. Amsterdam: John Benjamins.
Clark, E. V. (1981). Lexical innovations: How children learn to create new words. In W. Deutsch (eds), *The child's construction of language*, 299–328. London: Academic Press.
Clark, E. V. (1983). Meaning and concepts. In P. H. Mussen, L. H. Flavell & E. M. Markman (eds), *Handbook of child psychology: Vol. 3 Cognitive development*, 787–840. New York: Wiley.
Clark, E. V. (1998). Lexical creativity in French-speaking children. *Cahiers de Psychologie Cognitive* **17**, 513–30.
Clark, E. V. & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language* **14**, 547–67.
Clark, E. V., Gelman, S. A. & Lane, N. M. (1985). Compound nouns and category structure in young children. *Child Development* **56**, 84–94.
Clark, H. H. (1983). Making sense of nonce sense. In G. B. Flores d'Arcais & R. J. Jarvella (eds), *The process of language understanding*, 297–331. New York: Wiley.
Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
De Jong, N. H., Feldman, L. B., Schreuder, R., Pastizzo, M. & Baayen, R. H. (2002). The processing and representation of Dutch and English compounds: Peripheral morphological and central orthographic effects. *Brain and Language* **81**, 555–67.
Di Sciullo, A. M. & Williams, E. (1987). *On the definition of word*. Cambridge, MA: MIT Press.
Downing, P. (1977). On the creation and use of English compound nouns. *Language* **53**(4), 810–42.
Gagné, C. L. (2001). Relation and lexical priming during the interpretation of noun–noun combinations. *Journal of Experimental Psychology – Learning, Memory and Cognition* **27**(1), 236–54.
Gagné, C. L. & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology – Learning, Memory and Cognition* **23**(1), 71–87.
Gagné, C. L. & Shoben, E. J. (2002). Priming relations in ambiguous noun–noun combinations. *Memory and Cognition* **30**(4), 637–46.
Gagné, C. L. & Spalding, T. L. (2004*a*). Effect of discourse context and modifier relation frequency on conceptual combination. *Journal of Memory and Language* **50**(4), 444–55.
Gagné, C. L. & Spalding, T. L. (2004*b*). Effect of relation availability on the interpretation and access of familiar noun–noun compounds. *Brain and Language* **90**, 478–86.
Gentner, D. (1982). A study of early word meaning using artificial objects: What looks like a jiggy but acts like a zimbo? In J. Gardner (ed.), *Readings in developmental psychology*, 137–42. Boston: Little Brown.
Gerrig, R. J. & Murphy, G. L. (1992). Contextual influences on the comprehension of complex concepts. *Language and Cognitive Processes* **7**(3–4), 205–230.
Gleitman, L. & Gleitman, H. (1970). *Phrase and paraphrase: Some innovative uses of language*. New York: Norton.
Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
Goldberg, A. E., Casenhiser, D. M. & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics* **15**(3), 289–316.

Harris, M., Barrett, M., Jones, D. & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language* **15**(1), 77–94.

Kay, P. & Zimmer, K. (1976). On the semantics of compounds and genitives in English. In *Sixth California Linguistics Association Proceedings*, 29–35. San Diego, CA: Campile Press.

Klibanoff, R. S. & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development* **71** (3), 649–59.

Krott, A., Baayen, R. H. & Schreuder, R. (2001). Analogy in morphology: Modeling the choice of linking morphemes in Dutch. *Linguistics* **39**(1), 51–93.

Krott, A. & Nicoladis, E. (2005). Large constituent families help children parse compounds. *Journal of Child Language* **32**(1), 139–58.

Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* **16**, 589–600.

Landau, B., Smith, L. B. & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development* **3**(3), 299–321.

Langacker, R. (1987). *Foundations of cognitive grammar, Vol. 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Lees, R. B. (1960). *The grammar of English nominalizations*. Bloomington, IN: Indiana University.

Levi, J. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.

Lieber, R. (1983). Argument linking and compounds in English. *Linguistic Inquiry* **14**(2), 251–86.

Lieven, E., Behrens, H., Speares, J. & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language* **30**(2), 333–70.

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

Marchman, V. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighbourhood structure. *Cognitive Science* **21**, 283–304.

Mellenius, I. (1997). *The acquisition of nominal compounding in Swedish*. Lund: Lund University Press.

Neijt, A., Krebbers, L. & Fikkert, P. (2002). Rythm and semantics in the selection of linking elements. In H. Broekhuis & P. Fikkert (eds), *Linguistics in the Netherlands 2002*, Vol. 19, 117–27. Amsterdam/Philadelphia: AVT Publications.

Nelson, D. G. K., Russell, R., Duke, N. & Jones, K. (2000). Two-year-olds will name artifacts by their functions. *Child Development* **71**(5), 1271–88.

Nicoladis, E. (2002). What's the difference between 'toilet paper' and 'paper toilet'? French–English bilingual children's crosslinguistic transfer in compound nouns. *Journal of Child Language* **29**, 843–63.

Nicoladis, E. (2003). What compound nouns mean to preschool children. *Brain and Language* **84**, 38–49.

Nicoladis, E. & Krott, A. (2007). Family size and French-speaking children's segmentation of existing compounds. *Language Learning* **57**(2), 201–28.

Parault, S. J., Schwanenflugel, P. J. & Haverback, H. R. (2005). The development of interpretations for novel noun–noun conceptual combinations during the early school years. *Journal of Experimental Child Psychology* **91**(1), 67–87.

Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Plag, I. (2006). The variability of compound stress in English: Structural, semantic, and analogical factors. *English Language and Linguistics* **10**, 143–72.

Roeper, T. & Siegel, M. E. A. (1978). A lexical transformation for verbal compounds. *Linguistic Inquiry* **9**(2), 199–260.

Štekauer, P. (2005). *Meaning predictability in word formation*. Amsterdam/Philadelphia: John Benjamins.

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* **11**(1–2), 61–82.

Tomasello, M. (2003). *Constructing a language: A Usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Warren, B. (1978). *Semantic patterns of noun–noun compounds*. Goteborg: Acta Universitatis Gothoburgensis.

APPENDIX A: STIMULI

Compound	Dominant relation of compound	Relation bias of modifier	Relation bias of head	Family bias	# Responses with dominant relation: Children (max. 22)	Adults (max. 36)
birthday room	for	for	for	High–High	17	33
book basket	for	for	for	High–High	9	28
chocolate bread	has	has	has	High–High	9	18
corn sauce	made of	made of	made of	High–High	1	26
dish table	for	for	for	High–High	14	21
dog shoes	for	for	for	High–High	9	31
paper salad	made of	made of	made of	High–High	7	28
side ache	located	located	located	High–High	13	36
sports rack	for	for	for	High–High	7	35
sun suit	for	for	for	High–High	13	30
apple ring	made of	made of	for	High–Low	3	31
baby soup	for	for	has	High–Low	13	31
car milk	for	for	made of	High–Low	7	27
cheese fish	has	has	like	High–Low	6	12
coffee water	for	for	–	High–Low	7	11
day lesson	during	during	about	High–Low	14	34
fire foil	for	for	made of	High–Low	4	29
hockey powder	for	for	made of	High–Low	5	31
snow seat	made of	made of	located	High–Low	3	19
toy muffin	be	be	has	High–Low	7	34
animal helmet	for	of	for	Low–High	9	26
banana shop	for	has	for	Low–High	6	31
candy cake	made of	for	made of	Low–High	5	23
carrot board	for	has	for	Low–High	6	28
cookie sandwich	has	for	has	Low–High	17	20
egg bag	for	part	for	Low–High	9	22
horse skin	part	for	part	Low–High	10	28
lemon box	for	part	for	Low–High	9	22
pepper bread	has	for	has	Low–High	17	31
rain juice	made of	–	made of	Low–High	6	16

APPENDIX B

RELATIONS AND EXAMPLES FROM CHILDREN'S DATABASE BASED ON MAJORITY CODES PROVIDED BY FIVE BRITISH CODERS

Modifier–Head relation	examples
ABOUT (B is about A)	fairy story, alphabet song, science museum
BE (B is an A)	Barbie doll, baby bear, toy car
CAUSE1 (A causes B)	sunburn, heartbeat, motion sickness
CAUSE2 (B causes A)	light bulb
DURING (B happens during A)	daylight, winter sports, April fool
FOR (B is for A)	postbox, picnic table, baby blanket
FROM (B comes from/is derived from A)	seafood, olive oil, lemon juice
HAS (B has A)	cheese burger, apple tree, fruitcake
LIKE (B is like A)	jellyfish, banana boat, goldfish
LOCATED (B is located at A)	back door, farm animal, seabird
MADE OF (B is made of A)	cardboard box, chocolate bar, snowball
MAKES1 (A makes B)	honeybee, bubble gum
MAKES2 (B makes A)	rabbit-hole, birdnest, chicken egg
OBJECTIVE NOMINALIZATION (A is object of verb B)	haircut, shopkeeper, lawn mover
OPAQUE (A, B and/or whole is opaque)	butterfly, bonfire, ferris wheel
PART (B is part of A)	apple peel, chicken leg, eyelash
SUBJECTIVE NOMINALIZATION (A is subject of verb B)	snakebite, bee sting
USE (B uses A)	pillow-fight, windmill, water pistol
OTHER	weekend, boatman, beauty-spot