Language Learning ISSN 0023-8333

Word Family Size and French-Speaking Children's Segmentation of Existing Compounds

Elena Nicoladis
University of Alberta

Andrea Krott
University of Birmingham

The family size of the constituents of compound words, or the number of compounds sharing the constituents, affects English-speaking children's compound segmentation. This finding is consistent with a usage-based theory of language acquisition, whereby children learn abstract underlying linguistic structure through their experience with particular words. The family-size effect is particularly strong for the modifier or the leftmost element. The present study tested whether the effect of family size also holds for left-headed compounds as in French (e.g., *chef de police* "chief of police") and whether the effect is due to headedness or left-to-right processing. Twenty-eight French-speaking children between 3;5 and 5;3 were asked to explain the meaning of existing compounds with constituents of varying family size. The children were more likely to mention a constituent when it came from a large family than a small family, suggesting that children's segmentation of compounds might be facilitated by analogy with existing compounds. Furthermore, as in the previous English study, children mentioned modifiers more often than heads, showing their sensitivity to the semantic roles of the constituents, rather than left-to-right processing.

Keywords word family size; compounding; compound words; compound acquisition; children; French acquisition; word segmentation; French words

When children initially use compound words such as *doghouse*, it is unlikely they know that they are using a word that can be decomposed into the constituents *dog* and *house* (Berko, 1958; Clark & Berman, 1984). As children get

Correspondence concerning this article should be addressed to Elena Nicoladis, Department of Psychology, University of Alberta, P-217 Biological Sciences Building, Edmonton Alberta, T6G 2E9, Canada. Internet: elenan@ualberta.ca

older, they learn the basic principles of noun-noun compounding, realizing that compounds are combinations of two parts, one a head (i.e., the superordinate category [*house*]) and the other one a modifier (*dog*). Evidence for children's understanding comes from studies showing that children can produce and understand novel noun-noun compounds (e.g., Clark, 1993) and that they understand the main function of compounding (i.e., to distinguish between different subcategories [e.g., to differentiate a *coffee cup* from a *tea cup*; see Clark, Gelman, & Lane, 1985]). To do this, children must understand that the two nouns in the compounds play different semantic roles. For English, as for other Germanic languages, they have to discover that compounds consist of a right head and a left modifier.

The present study is concerned with the factors that play a role in children's acquisition of the general principles of compounding. One can learn about these factors by comparing compound acquisition across languages and language families. Crosslinguistic studies have demonstrated that the productivity and frequency of compounds play a crucial role in children's acquisition. In languages in which compounding is highly productive, like English, Chinese, Swedish, and Dutch, children start to coin novel noun-noun compounds, such as *house-smoke* to refer to smoke coming from the chimney of a house, early in development (Becker, 1994; Clark, 1993, 2003). Novel noun-noun compounds occur in English-speaking children's spontaneous speech before 2 years of age (Clark, 1981, 1983, 2003; see also Mellenius, 1997, for Swedish). In elicitation experiments, English children produce many novel compounds by the age of 3;0, Swedish children by the age of 4;0 (Clark et al., 1985; Mellenius). Nicoladis and Yin (2002) showed that Chinese-English bilingual children produced the kinds of compounds (e.g., noun-noun compounds, adjective-noun compounds, verb-noun compounds) that appeared most frequently in their input. In contrast to children who acquire "compounding" languages, children acquiring languages in which compounding is not very productive, like Hebrew, have been shown to form novel noun-noun compounds later in development (Clark & Berman, 1987). Hebrew-speaking children spontaneously produce novel compounds around 5;0, and children perform well in an elicitation experiment from the age of 6;0 (Clark & Berman). In comprehension experiments, children acquiring compounding languages are very accurate at identifying the referent of novel noun-noun compounds from the age of 3;0 (Clark et al., 1985, Mellenius; Nicoladis, 2003), while children acquiring Hebrew reach similar rates of accuracy around the age of 4;0 (Berman & Clark, 1989).

Taken together, studies on children's ability to form and understand novel compounds suggest that they understand the basic principle of noun-noun compounding at quite an early age, if compounding is productive in their input (Clark, 1993). However, does this mean that, once the basic principle is understood, all compounds are recognized as compounds (i.e., as having internal morphological structure) and that they are analyzed correctly? There is evidence that this is not the case. For example, Berko (1958) reported that English-speaking children between 4 and 7 years of age, who are assumed to master the principle of compounding, had difficulties parsing and explaining existing compounds like *birthday*. Instead of mentioning the constituents in their explanations, they often mentioned a major function or a salient feature of the compound. It seems that even children as old as 7 years might have some difficulty parsing existing compounds, in spite of the fact that they understand and use novel compounds much earlier. These results suggest that there must be other factors influencing the recognition of the internal structure of a compound. Krott and Nicoladis (2005) examined the effect of children's knowledge of similar compounds. They focused on the effect of the number of compounds that overlap with a target compound at a constituent. Such a set of compounds has been called a constituent family (e.g., De Jong, Feldman, Schreuder, Pastizzo, & Baayen, 2002). For example, the modifier constituent family of *paperclip* includes *paper bag*, *paper napkin*, *paperweight*, *paper chase*, and so on, whereas the head constituent family includes *hair clip, nose clip, tie clip*, and so on.

Family Size and Compounding

The idea that the size of constituent families might affect children's compound processing originates from research on adults' compound processing. Constituent family size has, for instance, been shown to affect the speed with which visually presented compounds are recognized. In visual lexical decision experiments, adults recognized English and Dutch compounds faster when the modifier or head had a large family size than when it had a small family size (De Jong et al., 2002; Krott, 2007). Modifier families also influence sense-nonsense decisions of novel compounds and the creation of novel compounds in Dutch and German (Krott, Baayen, & Schreuder, 2001; Krott, Schreuder, & Baayen, 2002; Krott, Schreuder, Baayen, & Dressler, 2004). The effect of the family size has been explained by coactivation of the constituent family members during the processing of the target compound. For the task of lexical decision of visually presented words, there is evidence that the effect of family size arises not because of formal overlap of the target word with its morphological family but because of conceptual overlap with its family (see also Bertram, Schreuder, & Baayen, 2000; De Jong, Schreuder, & Baayen, 2000; Schreuder

& Baayen, 1997). This is also in line with a finding by Krott (2007) that the effect is independent of the occurrence of spaces between the constituents in English compounds (but see De Jong et al., 2002).

The nature of the two effects is not clear. One possibility is that the two measures are effective on different levels of visual processing. Family frequency might be relevant for an early processing stage since it stands for the probability that a constituent appears in the left or right position of a compound, while family size probably relates to a more central level of processing.

There is evidence that constituent families are not only important for adults but also for preschool children. Post hoc analyses of a compound production experiment revealed that Dutch-speaking 4-year-olds are partly guided by constituent families when they create novel compounds (Neijt, Krebbers, & Fikkert, 2002). Dutch compounds can contain interfixes (e.g., -s- and –en- in schaap + s + hond "sheepdog" and schaap + en + kaas "sheep cheese"). Analyses of children's responses showed that the choice of interfixes for novel compounds (e.g., -en- for banana + ? + soep "banana soup") can partly be predicted by the distribution of interfixes in familiar compounds (banana + en + schil "banana peel," banana + en + ijs "banana ice cream," etc.).

Most important for the present study is the finding that family size affects children's segmentation of existing noun-noun compounds (Krott & Nicoladis, 2005). English-speaking children (3–5 years of age) were asked to explain to an alien puppet why we say certain compound words (e.g., *chocolate cake*). The compounds varied in family size of both modifiers and heads. The children were more likely to mention a constituent if its family size was large. These results suggest that family size (i.e., the knowledge of similar compounds) is a cue for segmentation of existing compounds for English-speaking preschool children. As mentioned, for adults' compound processing, it is assumed that constituent families are coactivated during the processing of a target compound. The effect of family size in children's compound processing can be understood along similar lines. When explaining the meaning of a compound, similar compounds (i.e., the modifier and head families) might be active and help to parse the compound.

To the extent that children use knowledge about the specific words within a construction, the results from the Krott and Nicoladis (2005) study are consistent with a usage-based theory of language acquisition (e.g., Tomasello, 2000). According to this theory, children initially use words that they have heard without knowledge of the underlying linguistic structure. With the acquisition of more and more words, children make conservative generalizations on the basis of how they have heard or used particular words. The higher the number of constructions in which those particular words are used, the more likely are children to generalize the use of those words in similar constructions (see, e.g., Bybee, 1995). As evidence, Tomasello (1992) found that his daughter's early use of particular verbs was similar to how she had previously used those verbs in other sentences (see also Akhtar, 1999; Matthews, Lieven, Theakston, & Tomasello, 2005). For example, at the age of 22 months and 8 days, she said "Laura gave that for me." The previous day she had said "Granmommy gave that for Mommy." The latter example could have served as a model for the former, with the child changing only the people who were the agent and receiver of the giving (see Tomasello, 1992, for other examples). Children might learn to generalize the linguistic rules first on the basis of particular words before they generalize on the basis of a more abstract pattern (see also Braine et al., 1990).

If children's interpretations of compounds are shown to be related to family size, then it is possible to conclude that children are acquiring compounds by analogy with other compounds in their vocabularies; that is, children acquire general rules about compounding by first learning how existing, familiar compounds derive their meaning.

An unresolved issue of Krott and Nicoladis's (2005) study is the result that the children were more likely to mention modifiers than heads in their explanations. Children often mentioned the head nouns only by referring to them with the pronoun *it* or not at all. This suggests that the children focused more on the modifiers than the heads. However, why were the modifiers more important than the heads? Was it because they appear first in the input stream and are therefore processed first? There is evidence for a left-to-right processing of morphologically complex words when presented visually or auditorily (Cutler, Hawkins, & Gilligan, 1985; Green, Hammond, & Supramaniam, 1983; Hudson & Buijs, 1995; Järvikivi & Niemi, 1999; Segui & Zubizarreta, 1985). However, directionality is clearly not the only factor that influences adults' and children's language processing (see, e.g., Clancy, Lee, & Zoh, 1986). The difference between modifiers and heads might, for instance, also be due to their functional difference. Modifiers might be more salient because they express the features that distinguish compounds from each other. Interestingly, modifier families have been shown to be more important in adults' compound processing as well (De Jong et al., 2002; Krott et al., 2001, 2002, 2004). Note, though, that these studies have all examined right-headed compounds and therefore cannot rule out that the importance of the modifier is due to a left-to-right parsing (see, though, importance of modifiers in Storms & Wisniewski, 2005). In order to rule out that the greater importance of modifier families in Krott and Nicoladis's experiment can simply be explained by a left-to-right processing, it is necessary to look at a language with left-headed compounds. For the present study, we therefore asked children to explain French noun-noun compounds, which are left-headed. In this way we were able to check whether the effect of the constituent family size on children's parsing of compounds can be generalized to another language and whether the greater importance of the modifier can be replicated for a left-headed language.

French Noun-Noun Compounds

There has been some debate as to whether there are nominal compounds in French at all.[1] There are two possible structures that are candidates for compounds[2] : noun-noun structures (like *timbre-poste* "stamp-post + office" meaning a postage stamp or *homme-grenouille* "man-frog" meaning a diver) and noun-preposition-noun structures (like *sac à main* "bag for hand" meaning a handbag or *chef de police* "chief of police" meaning police chief). The latter type is more frequent than the former, which is restricted in semantics and not very productive (see Nicoladis, 2001). Both structures are left-headed.

Di Sciullo and Williams (1987) proposed a universal Right-headed Rule, ruling out the existence of left-headed compounds and therefore ruling out the existence of nominal compounds in French. However, the Right-headed Rule has been ignored or denounced by the majority of linguists (e.g., Beard, 1995; Bréal, 1964; Fabb, 1998; Robinson, 1979; Selkirk, 1982; Spence, 1980; ten Hacken, 1994), leaving French minimally with noun-noun compounds. This seems to be also Clark's (1998) position when she described French as having few compounds.

There is no general agreement as to whether noun-preposition-noun structures are compounds (see discussion in Robinson, 1979), at least partially because there is no general agreement on what constitutes a compound word (see Fabb, 1998). To further complicate the problem, some researchers have argued that some noun-preposition-noun structures are compounds and some not (Robinson). For the purpose of this article, we will assume that noun-preposition-noun structures are compounds (following Fabb; Robinson). Like noun-noun compounds in both French and English, noun-preposition-noun structures serve the same function (i.e., they provide names for objects) and typically allow no intervening linguistic material. For example, it is not possible to insert an adjective in the middle of many noun-preposition-noun structures; for the lexical *sac à main* "bag for hand" to refer to a handbag, one cannot say *sac à grande main* "a bag for big hand." There is some psycholinguistic evidence to back up our assumption that noun-preposition-noun structures are compounds. French-speaking preschool children order the two nouns in novel noun-noun and novel noun-preposition-noun constructions in a similar manner, suggesting that the children do not differentiate between the two structures (Nicoladis, 2001).

The most common prepositions to appear in noun-preposition-noun compounds are *de* (e.g., *bottes de neige* "boots of snow" meaning snow boots) and *à* (e.g., *cuiller à thé* "spoon for tea" meaning a teaspoon; see Nicoladis, 2001). Robinson (1979) pointed out that the choice of preposition in a compound can help to interpret the compound as a whole. For example, *tasse à café* "cup for coffee" means a coffee cup, whereas *tasse de café* "cup of coffee" means a cup with coffee in it. However, the choice of preposition does not guarantee that the interpretation will be straightforward (Robinson). For example, *sac de papier* "bag of paper" could be a bag made out of paper or a bag filled with paper. Preschool children appear not to benefit from the meaning of the preposition in producing novel noun-preposition-noun compounds. In Nicoladis, even when French-speaking children included a preposition in their novel nominal compounds, they did not necessarily use the preposition in novel compound nouns as adults do, preferring to use the more frequent *de* where adults would say *à*. For that reason, when calculating family size in this study, we have ignored whether there was a preposition in the compound and have not distinguished between prepositions. We have even included some items with a preposition + determiner (e.g., *des* "of + the" and *au* "'with + the") because these words often appear between two nouns related to food and might not reveal any difference to the children. For example, consider *gâteau au chocolat* "cake with + the chocolate" referring to chocolate cake and *barre de chocolat* "bar of chocolate" referring to a chocolate bar. Although the word between the two nouns is different, both the bar and the cake have chocolate in them.

In sum, this study is predicated on several assumptions about French compounds. First, we assume that French has left-headed compounds, both noun-noun and noun-preposition-noun. Second, we assume that whether a compound has a preposition makes no difference to preschool children when counting family size. Third, we assume that the choice of preposition (including some examples of preposition + determiner) makes no difference to children when counting family size. We have justified these assumptions where possible with reference to previous studies.

Bilingualism and Language Development

The primary focus of our study was children's acquisition of French. Most of the children who participated in this study were bilingual (French and English); that is, all but five children understood English well enough to perform above chance on a test of English comprehension vocabulary. Although there were variations in the degree of the children's English and French knowledge, according to parental reports, the children who participated in this study were, on average, more proficient in French than in English. Before turning to our design of this study, we briefly review how the children's bilingualism might affect our results. Bilingualism has documented effects on both language and cognitive development (see Genesee & Nicoladis, 2007). Bilingualism can affect the rate of development of particular linguistic structures, either through delay (e.g., Nicoladis, Palmer, & Marentette, 2007) or acceleration (e.g., Nicoladis, 2003). One way to control for possible differences between monolinguals and bilinguals is to take into account their degree of linguistic development, as in their score on a vocabulary test (e.g., Nicoladis, 2003). In this study, we analyzed the relationship between the children's French vocabulary scores and their parsing of compounds. Another way in which bilinguals' language development differs from that of monolinguals is by systematic influence of one language on the other, or crosslinguistic transfer. Studies on crosslinguistic transfer have generally shown that transfer is more likely when the two languages share an underlying structure (Müller & Hulk, 2001). All documented cases of crosslinguistic transfer have been in terms of linguistic structure, such as choice of phoneme or ordering of linguistic elements, and in production rather than comprehension (see review in Nicoladis, 2006). Given our current understanding of crosslinguistic transfer, we cannot see how the children's knowledge of English could influence their segmentation of French compounds. Nevertheless, we checked for a possible effect of English knowledge by checking for the effect of English family size on the children's performance. One final way in which bilingual children are known to differ from monolingual children is in some aspects of cognitive development, particularly the ability to control their attention to the relevant aspects of a problem (Bialystok, 2001). Even if the children in this study were advanced in cognitive development relative to monolingual children, there would be no reason to suspect that the *pattern* of results we observed would not also be true of monolingual children.

In sum, although there are differences between bilingual and monolingual children in development, we have made every attempt to account for the known differences in our analyses. Whereas bilingual children might produce novel compounds that differ from monolinguals in terms of headedness (Nicoladis, 2002), the pattern of results in terms of the influence of frequency of head and modifier is unlikely to be due to crosslinguistic influence. For this reason, we argue that we can make conclusions about how French-speaking children in general acquire compounds on the basis of our data with bilingual children. Naturally, these conclusions would be strengthened if our results could be replicated with monolingual children.

This Study

The primary purpose of the present study is to see whether constituent families affect preschool children's segmentation of familiar French compounds. We expected to replicate the findings of Krott and Nicoladis (2005), showing that preschool children are more likely to mention a compound constituent when they know other compounds with the same constituents.

We asked children to explain the meaning of familiar compound words composed minimally of two nouns (following Berko, 1958; Krott & Nicoladis, 2005; Mellenius, 1997). Children as young as 2 years of age have been noted to spontaneously explain the meaning of compounds (Clark, 2003). An explanation of a compound naturally contains the two constituents because of its subcategorizing function (Berman & Clark, 1989; Clark & Berman, 1987). The meaning of a compound is often adequately captured by the meaning of the higher category (the head), the distinguishing feature of the subcategory (the modifier), and the relation between both constituents (*chapeau-soleil* "hat-sun" meaning a sun hat is a hat FOR the sun). Given this structure, the most straightforward way of explaining a compound is to name both head and modifier and to state the relation between them. The dependent measure is therefore whether the child explicitly (or implicitly) mentions the head or the modifier.

We refer to this task as a segmentation task because if children mention the constituents of a compound in their explanation, they need to have segmented the compound. However, the task used in this study underestimates whether the children have indeed segmented the compounds. The failure to mention the parts and the semantic relation can have different reasons. First, a child might not have decomposed a compound into its constituents. In this case, one expects the child to mention a major function or a salient feature of the compound as in Berko's (1958) experiment. Second, a child might have recognized only one of the two constituents. For example, if the child recognizes *animaux* in *animaux de cirque* "animals of circus" meaning circus animals but does not know what *cirque* means, the child might reject the possibility that *animaux de* *cirque* is a compound and therefore will mention a major function or a salient feature instead. Alternatively, the child might only mention *animaux*. Third, a child might have recognized both parts of the compound without being able to relate the two constituents. In this case, the child might decide not to mention the constituents. Therefore, not mentioning a constituent does not necessarily mean that the compound has not been correctly segmented. Nevertheless, if children's responses differ by compounds with different family sizes, then we have evidence that family size indeed affects their compound processing.

Krott and Nicoladis (2005) found that English-speaking children were more likely to mention modifiers than heads in a similar segmentation task. Because English compounds are right-headed, it was not clear if the children mentioned modifiers more because they were the first constituent or because of the semantic role. By looking at left-headed French compounds, we can clarify this result. Thus, a secondary goal of this study is to see if French-speaking children are more likely to mention the first constituent (the head) or the modifier (the rightmost constituent).

As mentioned, in contrast to English compounds, French compounds frequently contain a preposition such as *de* or *à*. The occurrence of these prepositions might help children to recognize that they are dealing with a complex structure and to find the boundary between the constituents (cf. Nicoladis, 2001). French-speaking children might therefore perform better in the task than the children in the previous English study (Krott & Nicoladis, 2005).

Apart from family size, children's responses might also be influenced by other factors—for instance, by their knowledge of the compounds. In order to estimate individual children's familiarity with compounds, we measured their vocabulary in French. The rationale was that the greater the number of words known by a child, the greater the number of compound words he/she would know. We decided against using measures of familiarity by previous studies. The measure of compound familiarity used in Krott and Nicoladis (2005) was a parental checklist of (among others) the target compounds. That measure had no relationship at all to children's performance (see Appendix B). In contrast, some aspects of the acquisition of compounding are related to vocabulary test scores in French (Nicoladis, 2003).

Methods

Participants

Twenty-eight French-speaking children took part in this study. The children were aged 3;5 to 5;3 (mean = 4;5, *SD* = 0;6). Ten of them were boys and 18 were girls.

Two other children were tested for English proficiency but left the preschool before we could test their French. All of the children lived in Edmonton, Alberta, Canada, a part of Canada where English is the majority language. There is a small but active French-speaking community in Edmonton. French-speaking children in Edmonton can be monolingual in the preschool years but are usually bilingual in French and English by the school years. The children for this study were recruited through French daycare centers and preschools.

According to parental reports, five of the children spoke no English at all. Of the remaining 23 bilingual children, their parents reported that most spoke French better than English. We tested the language proficiency of the 23 bilingual children in both English and French using vocabulary tests to see if their language proficiency influenced the children's performance. In English we used the Peabody Picture Vocabulary (PPVT) Test-III version A (Dunn & Dunn, 1997). In French we used a translated version of the same test (version B). The translated version has not been standardized, but the raw scores have been found to correlate with age for monolingual French-speaking children (Nicoladis, 2007).

Overall, the children achieved higher scores in the French vocabulary test (mean = 59.3, *SD* = 11.9) than in the English vocabulary test (mean = 45.5, *SD* = 29.1). Note, though, that the raw scores might not be compatible in French and English because the French test is a translation of the English one. Nevertheless, because the vocabulary scores follow the same trend as the parental reports, the fact that the mean score for the French vocabulary test is higher suggests that most of the children are French-dominant.

Materials

Following Krott and Nicoladis (2005), we first gathered an extensive list of French noun-noun compounds (both with and without prepositions) that children might know. For this, we started by translating a database of English compounds that children might know, constructed by Krott and Nicoladis. We then asked a fluent French speaker and a native French Canadian speaker, both with psycholinguistic training and experience with children, to add compounds that children might know and remove compounds that were not used in Canadian French or that children were unlikely to know.

For each combination in this list, we calculated the head (modifier) family size by determining the number of other combinations in the list that contain the same heads (modifiers). For example, the head family of *cuiller à thé* "teaspoon" was the set consisting of *cuiller d'argent* "silver spoon," *cuiller à dessert* "dessert spoon," *cuiller de service* "serving spoon," *cuiller à soupe* "soup spoon," and *cuiller à table* "tablespoon," whereas the modifier family of *chef de police* "police chief" was the set consisting of *agent de police* "police agent," *caserne de police* "police station," *commissariat de police* "police commission," *poste de police* "police station," *uniforme de police* "police uniform," and *voiture de police* "police car." We then selected from our list four sets with each eight noun-noun combinations, listed in Appendix A. The combinations of set 1 had a high head family size (mean = 7.9 family members, $SD = 1.5$) and a high modifier family size (mean = 10.1, $SD = 3.7$). The combinations of set 2 had a high head family size (mean = 12.5, $SD = 7.5$) and a low modifier family size (mean = 1, $SD = 0$). Set 3 consisted of combinations with a low head family size (mean = 1.1, $SD = 0.4$) and a high modifier family size (mean = 12.0, $SD = 5.2$). Finally, set 4 consisted of eight combinations with a low head family size (mean = 1, $SD = 0$) and a low modifier family size (mean = 1, $SD = 0$). In total, there were 32 items that children were asked to explain (see Appendix A).

Procedure

The procedure was based on that used in Krott and Nicoladis (2005). All children were tested in a quiet room or corner in their daycare center or preschool. A native speaker of Canadian French asked the children to explain the meaning of the noun-noun compounds. The compounds were presented in a different random order for every participant. The experimenter first introduced them to the task by giving them the example *tasse à thé* "cup for tea," explaining in French that it was called a teacup because it was a cup for tea. To elicit the answers, she first used the question *Qu'est-ce que* ______ *veut dire*? "What does ____ mean?" If the child did not answer, she followed up by asking *Pourquoi est-ce qu'on dit* ______? "Why do we say ____?" If the child still did not answer or claimed not to know, she went on to the next item and did not return to unanswered items. The children's responses were tape-recorded and all speech relevant to the task was transcribed.

Scoring Children's Responses

The responses were coded by the first author. For the coding, we again followed the procedure of Krott and Nicoladis (2005). We coded heads and modifiers separately. If a child explicitly mentioned a constituent (head or modifier) in the explanation, the constituent received two points. This was, for example, the case for the modifier *neige* "snow" in the explanation *c'est pour la neige* "it's for the snow" for *pantalon de neige* "pants of snow" referring to snow pants. If a constituent was mentioned indirectly, as *lunch* "lunch") in the explanation *c'est pour manger* "it's for eating" for *boîte à lunch* "lunch box," it received one point. We did not credit children with partial credit if they used a subject pronoun (as in the previous example) because we assumed that children might use a subject pronoun for grammatical reasons alone. We did, however, award partial credit if a child used an object pronoun (e.g., one point would be counted for *boîte* "box" in *tu mets le lunch dans ça* "you put lunch in that"). If the child had not mentioned the constituent at all, as *boîte* "box" in the example "*c'est pour manger*," we gave zero points. In this way, each constituent (head and modifier) in each answer received two, one, or zero points.

To check on the interrater reliability of our codes, an independent coder coded three children's responses. Both coders agreed on an average of 82.8% of the children's responses. Most disagreements were due to whether to give children partial credit (i.e., one point) for indirect mention, with the first coder tending to give more partial credit than the second coder. We decided to retain the first coder's choice in order to remain consistent with the other children's codes.

Results

Modifier Dominance

In Krott and Nicoladis's (2005) English study, children had mentioned compound modifiers more often than compound heads. When comparing modifier scores and head scores in the present study, we found that the modifier scores (mean = 0.58, $SD = 0.28$) were again, on average, higher than the head scores (mean = 0.41, $SD = 0.24$). This difference was significant, $t(31) = -2.54$, $p = .016$, two-tailed. As in the English study (Krott & Nicoladis), this was not due to a higher average modifier family size (6.0) than an average head family size (5.6), $t(31) = -0.25$, $p = .81$. The scores in the present study were lower than those in the English study (mean modifier score: 1.2; mean head score: 0.8; see Krott & Nicoladis). This might be due to the fact that the children in the present study were on average 5 months younger and bilingual. Given these differences it is difficult to draw any strong conclusions. Nevertheless, there is no evidence that the prepositions contained in French compounds helped the children to decompose the compounds.

Effect of Family Size

To check the influence of the head and modifier family size on head and modifier scores, we calculated a mean score for each stimulus and a mean score for each subject for each combination of head and modifier family size (see Table 1).

Table 1 Mean head and modifier scores (*SD*) for all four experimental conditions

Head Fam	Mod Fam	Mean head scores (*SD*)	Mean modifier scores (*SD*)
High	High	0.476 (0.3)	0.674 (0.2)
	Low	0.478 (0.2)	0.443 (0.3)
Low	High	0.408 (0.3)	0.750 (0.2)
	Low	0.281 (0.2)	0.453 (0.2)

For both item means (F_2) and subject means (F_1), we conducted two head family size (2) × modifier family size (2) ANOVAs: one with head scores as the dependent variable and one with modifier scores. For the head scores, there was only a significant effect of head family size in the subject analysis, $F_1(1, 27) = 9.4, p = .005$, not in the item analysis, $F_2(1, 31) = 2.4, p = .131$, with higher mean scores for high family size compounds (0.48) than for low head family size compounds (0.34). This means that a higher head family size led to higher head scores (i.e., to a higher likelihood that the children mentioned the head in their explanations). There was no effect of modifier family size, either in the subject analysis, $F_1(1, 27) = 2.1, p = .164$, or in the item analysis, $F_2(1, 31) < 1$, and no interaction of head family size and modifier family size, $F_1(1, 27) = 2.4, p = .134$; $F_2(1, 31) < 1$.

Analyses of variance for modifier scores showed similar results. There was a significant effect of modifier family size on modifier scores, in both subject, $F_1(1, 27) = 30.2, p < .001$ and item analysis, $F_2(1, 31) = 8.6, p = .007$, and there was no effect of head family size, neither an interaction of the two factors (all F's < 1). When collapsing high and low head groups, scores for compounds with high modifier family size (mean 0.71) were significantly higher than scores for compounds with low modifier family size. Thus, the results show that the children mentioned those modifiers more often that had a higher family size.

Age

In the English study (Krott & Nicoladis, 2005) the effect of head family size had been present only for 5-year-olds, not for 4-year-olds, whereas the effect of modifier family size had been present for all children. To look for differences in the present study, we split the group of children into 3-, 4-, and 5-year-olds. Both the group of 3-year-olds (five children) and of 5-year-olds (five children) did not show any significant effect of head family size on head scores (5-year-olds: $t(4) < 1$, one-tailed). There was only a trend for 3-year-olds, $t(4) = 1.9, p = .06$, one-tailed. The group of the 18 four-year-olds revealed a significant effect of head family size on head scores, $t(17) = 2.9$, $p = .006$, one-tailed. As the groups of 3- and 5-year-olds are very small, we cannot rule out that their responses were also affected by head family size. In contrast to the head scores, modifier scores showed a significant effect of modifier family size for all three age groups: 3-year-olds, $t(4) = 3.3$, $p = .016$; 4-year-olds, $t(17) = 3.4$, $p = .002$; 5-year-olds: $t(4) = 4.3$, $p = .006$; all one-tailed. As the three groups did not differ very much from each other (at least not with respect to the effect of modifier families), one wonders whether the children's performance increased with age at all. This seemed to have been the case, though, as there was a significant correlation between the children's age (in months) and their scores for heads (Spearman r's $= .55$, $p = .001$, one-tailed)[3] as well as between age and scores for modifiers (Spearman r's $= .50$, $p = .003$, one-tailed). These relations are also shown in Figure 1.

French Vocabulary Size

Apart from the relationship between age and children's performance, we wanted to know whether their vocabulary size played a role. For that, we looked at the correlations between the French PPVT scores and the head (modifier) scores (see Figure 2). A significant correlation between the PPVT scores and the head scores (Spearman r's $= .38$, $p = .024$, both one-tailed) as well as between PPVT scores and modifier scores (Spearman r's $= .44$, $p = .010$, both one-tailed) showed that the children do seem to have performed better when they had a larger French vocabulary.

English Vocabulary

As the noun-noun combinations tested in this study have English counterparts that bilingual children might know as well, we checked on how far the children's

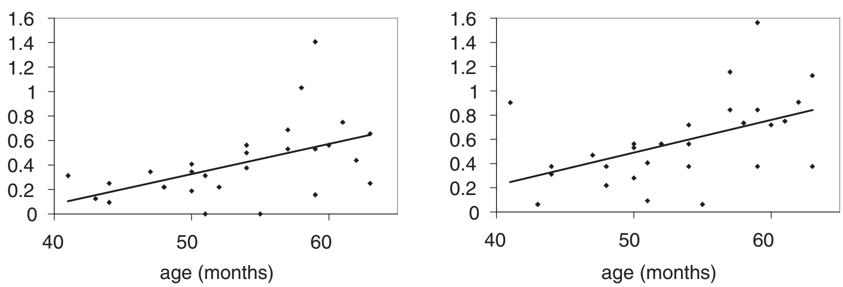

Figure 1 Relation between children's age and head scores and modifier scores plus linear trend line.

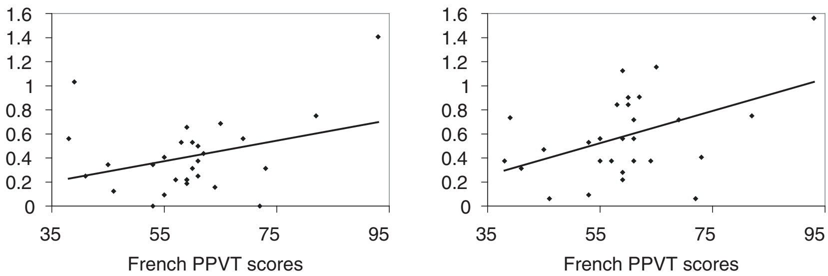

Figure 2 Relation between children's French proficiency (PPVT) and head scores and modifier scores plus linear trend line.

knowledge of English affected their performance. For that, we examined the relation between the English PPVT scores and children's scores and the possible additional effect of English family size on children's scores. A correlation analysis revealed that there was no significant correlation between the English PPVT scores and head scores (Spearman r's $= .14$, $p = .245$, one-tailed) or modifier scores (Spearman r's $= -.067$, $p = .467$, one-tailed). Thus, English vocabulary was not related to task performance.

We also examined whether English family size was related to the children's performance. For that, we determined the corresponding English modifier families of each French compound as the modifier family of the English translation equivalent, and the corresponding English head families of a French compound as the head family of the English translation equivalent. Correlation analyses showed that in the case of children's scores for heads, only French family size (Spearman r's $= .35$, $p = .048$) correlated significantly with the scores, whereas there was a trend for English family size (Spearman r's $= .33$, $p = .069$). In the case of children's scores for modifiers, English family size (Spearman r's $=$ $.62$, $p < .001$) correlated slightly stronger with the scores than French family sizes (Spearman r's $= .56$, $p = .001$). However, French and English family sizes were moderately to highly correlated with each other (heads: Spearman r's $= .56$, $p = .001$; modifiers: Spearman r's $= .80$, $p < .001$). These correlations indicate a close resemblance of the two languages. It most likely shows that both French noun-preposition-noun combinations and English noun-noun compounds denote objects that children experience in their environment and it therefore supports our above claim that these two linguistic structures serve the same function. In fact, the two vocabularies resemble each other so closely that, when partialing out English or French family size, neither factor correlated with children's scores (head scores and French family size: $r = .06$, $p = .752$;

head scores and English family size: $r = .13$, $p = .475$; modifier scores and French family size: $r = .18$, $p = .328$; modifier scores and English family size: $r = .22$, $p = .236$). Therefore, these analyses cannot differentiate whether it was the English or French family size that had an effect on children's responses. However, because the children were mostly French dominant and because their performance correlated with the French, not the English, vocabulary scores, it is more likely that it was the French family size and that the English family size does not add to their performance.

Age Versus French Vocabulary

As we have seen, children's age and their French PPVT scores are both correlated with head and modifier scores. In addition, these two variables are correlated with each other (Spearman r's $= .55$, $p = .001$). Thus, the question arises whether both factors influence the performance or whether it is only one of them. To answer this question, we partialed out the contribution of each factor. When controlling for the PPVT, age does correlate with both head scores ($r = .50$, $p = .004$) and modifier scores ($r = .39$, $p = .022$), whereas when controlling for age, PPVT does *not* correlate with head scores ($r = -.11$, $p = .289$) or modifier scores ($r = .16$, $p = .208$). These results suggest that vocabulary size might not be relevant for our task.

Frequency Versus Family Size

The family size of a constituent is typically correlated with the frequency of the constituent as an independent word. In the English study we found that children tended to name the head noun in their explanations more often if the head was more frequent (i.e., more familiar to them), whereas there was no such indication for the modifier (Krott & Nicoladis, 2005). Therefore, we were interested in whether the children in the present study mentioned a constituent in their responses because the constituent was familiar rather than because it occurred in many other compounds. We estimated the familiarity of the constituents of our compounds by determining their frequency in the French part of the CHILDES database (MacWhinney, 2000; see Appendix B), counting their occurrences in both adults' and children's utterances. Adding these frequencies (head frequency or modifier frequency) as covariates to our by-item ANOVAs led to very similar results as the ones in the English study: There was a main effect of head frequency for the head scores, $F_2(1, 29) = 7.4$, $p = .011$, no effect of head family size, $F_2(1, 29) < 1$, and no interaction ($F_2 < 1$). This suggests that the marginal significant effect of head family size in the item analysis was probably due to differences in head frequencies. In the case of modifier scores, however, there was a main effect of both modifier frequency, $F_2(1, 28) = 4.9$, $p = .035$, and modifier family size, $F_2(1, 28) = 6.1, p = .020$, and no interaction ($F_2 < 1$), suggesting that both higher modifier family size and higher modifier frequency led to increased modifier scores. Thus, children named the modifier nouns more often in their explanations when they were more familiar with the nouns and when these nouns occurred more often in other familiar compounds, whereas they mentioned the head nouns more often when they were more familiar with the nouns as independent words. The latter replicates previous findings for English (Krott & Nicoladis).

Discussion

We have shown that French-speaking children are more likely to mention a constituent (particularly a modifier) in their explanation of a compound if the constituent has a high family size. This indicates that children's compound segmentations are affected by (modifier) family size (i.e., by their knowledge of other compounds that are similar in terms of form and meaning).

The present results are very similar to those for English-speaking children (Krott & Nicoladis, 2005) despite differences between English and French compounds. In contrast to English compounds, French compounds are left-headed and contain explicit segmentation cues (i.e., prepositions between the constituents). Nevertheless, overall performance in the French experiment was not higher than that in the English study. In both studies, children's responses were strongly affected especially by modifier family size. The effect of family size therefore appears to generalize crosslinguistically.

However, what exactly is the nature of the effect? We have seen that for French-English bilingual children, French and English family sizes are correlated and account for the same variance in the children's responses, meaning that English family size had no additional effect on children's performance. Because English families and French families of particular modifiers are not identical (they contain different compounds [i.e., concepts]), this result suggests that the family size effect does not arise from semantic overlap with the target compounds alone but from both semantic and form overlap. In other words, it is not the number of concepts with a specific modifier (e.g., TEA) but the number of compounds of the target language with a specific modifier (the modifier *thé* "tea") that affects children's compound segmentation. Because of the importance of both form and meaning overlap, the family-size effect in our task is a morphological effect.

Our findings suggest that the knowledge of compounds belonging to a constituent family can reveal the internal structure of compounds to children. Some of the children's responses in this study, along with others (e.g., Berko, 1958; Clark & Berman, 1984; Krott & Nicoladis, 2005), suggest that children initially acquire compound words as whole units without internal structure. For example, one of the children in our study (age 3;5) explained the compound *sapin de Noël* "fir + tree of Christmas" meaning a Christmas tree as *on décore*! "we decorate!" It is clear that the child knew what a Christmas tree is and what its function is, but it was not clear that she knew that the word was decomposable. We argue that once this child has realized that there are other words modified by *Noël* and that all these words have something to do with Christmas, she might segment *sapin de Noël* into its constituents. These results are consistent with a usage-based theory of language acquisition (Tomasello, 2000); that is, children are more likely to segment existing compounds if they have experience with the compounds' constituents in other compounds. A usage-based theory would further predict that children's segmentation of existing compounds would precede and predict their ability to use novel compounds. A future study could test that prediction.

Krott and Nicoladis (2005) considered the possibility that left-to-right processing could account for English-speaking children's tendency to mention modifiers (the leftmost element) more than heads (i.e., to focus more on the modifiers) when explaining compounds. In the present study, French-speaking children were also more likely to mention modifiers (here, the rightmost element) than heads. Taken together, these suggest that the role of a constituent might be more important in this task than its position in the compound. Our results support the earlier finding that preschool children already understand the function of compounds to refer to subcategories (Clark et al., 1985). For that reason, they regarded it to be more important to mention the modifier (i.e., the constituent that distinguishes a compound from other compounds with the same head) than the superordinate category (i.e., the head).

How do the results for children's compound processing relate to adults' compound processing? Adult studies have shown that constituents play an important role during compound processing. There is accumulating evidence that both visually (Andrews, Miller, & Rayner, 2004; Hyönä & Pollatsek, 1998; Inhoff, Radach, & Heller, 2000; Juhasz, Inhoff, & Rayner, 2005) and auditorily presented compounds (Koester, Gunter, Wagner, & Friederici, 2004) are decomposed into constituents and that constituents play a role, especially in an early phase of recognition. For instance, eye-movement studies have shown that participants focused longer on low-frequency constituents than on high-frequency constituents (Andrews et al.; Hyönä & Pollatsek; Pollatsek & Hyönä, 2005). On the other hand, there is evidence that compounds are processed as wholes because the frequency of the whole affects processing latencies for visual stimuli (e.g., De Jong et al., 2002; Pollatsek, Hyönä, & Bertram, 2000; Van Jaarsveld & Rattink, 1988). Because of the importance of constituents and full forms in compound processing for adults, it is often assumed that compounds are processed via two routes (i.e., they are accessed as full forms and via their constituents), and it is assumed that these routes are followed in parallel (e.g., De Jong et al.; Koester et al.; Krott, 2007). The importance of compound constituents is also evident in the effect of the constituent family size. For example, adults recognize visually presented compounds faster when the modifier and/or head have large families (De Jong et al.; Krott, 2007). This effect appears to be independent of orthographic conventions because the effect has been found for English compounds with and without spaces (Krott, 2007).

The factors that have been found to affect adults' compound processing appear to play a role in children's compound processing as well, but not necessarily in the same way. Both the present French study and the previous English study (Krott & Nicoladis, 2005) suggest that the decomposition and processing of compounds depend on the frequency of the heads (and modifiers) and the family size of constituents, especially the size of modifier families. In accordance with the conclusion for adults that family-size effects are independent of orthographic conventions (see Krott, 2007), results are very similar for English compounds without spaces and French compounds that are often written with spaces.[4] In contrast to adult studies, though, the findings suggest that family size affects children's decomposition of compounds. Adult studies have concluded that effects of constituent families arise due to lexical/conceptual relatedness in the mental lexicon (Bertram et al., 2000; De Jong et al., 2000; Schreuder & Baayen, 1997). Krott (2007) even argues that family-size effects of compounds only arise when compounds are processed on a semantic level. We argued earlier that children's segmentation might be a morphological effect (i.e., due to the choice of constituent) rather than a conceptual effect. If this interpretation is correct, then there might be a developmental sequence from word-based analysis to more abstract category-based analysis of linguistic structures (Tomasello, 2000).

To conclude, the present results have demonstrated the crosslinguistic generalizability of the importance of family size in children's acquisition of existing compounds. The importance of family size in children's segmentation reinforces an argument from a usage-based theory of language acquisition (e.g., Tomasello, 2000). Our results suggest that the type frequency of particular words within particular structures is related to productivity in acquisition; that is, it appears to be rather the local constituent family that affects children's segmentation of existing compounds and not so much their general knowledge of compounds and compounding. Our results can also explain previous findings that children are able to understand and produce novel compounds from an early age while still having problems with decomposing existing compounds years later (e.g., Berko, 1958; Krott & Nicoladis, 2005; Mellenius, 1997); that is, it might be compounds with small or no family sizes that resist decomposition for school-aged children. Furthermore, as the modifier played a more important role in both English right-headed and French left-headed compounds, our study confirms previous findings that children are sensitive to the different roles played by the modifier and head noun in compounds (Clark et al., 1985).

Revised version accepted 3 March 2006

Notes

1 There is little debate that deverbal compounds like *casse-noisettes* "break-hazelnuts" meaning nutcracker exist in French (although those are rare and not very productive).
2 This discussion is limited to endocentric compounds and nonborrowed compounds. As Spence (1980) pointed out, exocentric and borrowed compounds might function differently. For example, as a general rule, the gender of the leftmost noun of a compound determines the gender of the compound as a whole. Exceptions to this general rule are almost all exocentric (e.g., *rouge-gorge* "red-throat" means a robin and is masculine even though *gorge* is feminine) or borrowed from either Latin or English.
3 The data of our study contained outliers and/or were skewed. We therefore report Spearman correlation coefficients.
4 Note, however, that the subjects of our study were preliterate children, who could not know whether the compounds we presented are written with spaces.

References

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driving learning of syntactic structure. *Journal of Child Language*, *26*, 339–356.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*, 285–311.

Beard, R. (1995). *Lexeme-morpheme base morphology*. Albany: State University of New York Press.

Becker, J. A. (1994). "Sneak-shoes", "sworders" and "nose-beards": A case study of lexical innovation. *First Language*, *14*, 195–211.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150–177.

Berman, R. A., & Clark, E. V. (1989). Learning to use compounds for contrast: Data from Hebrew. *First Language*, *9*, 247–270.

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 489–511.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.

Braine, M. D. S., Brody, R. E., Brooks, P. J., Sudhalter, V., Ross, J. A., Catalano, L., et al. (1990). Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of Memory and Language*, *29*, 591–610.

Bréal, M. (1964). *Semantics: Studies in the science of meaning*. New York: Dover.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425–455.

Clancy, P. M., Lee, H., & Zoh, M.-H. (1986). Processing strategies in the acquisition of relative clauses: Universal principles and language-specific realizations. *Cognition*, *24*, 225–262.

Clark, E. V. (1981). Lexical innovations: How children learn to create new words. In W. Deutsch (Ed.), *The child's construction of language* (pp. 299–328). London: Academic Press.

Clark, E. V. (1983). Meanings and concepts. In P. H. Mussen (Series Ed.), L. H. Flavell & E. M. Markman (Vol. Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 787–840). New York: Wiley.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Clark, E. V. (1998). Lexical creativity in French-speaking children. *Cahiers de Psychologie Cognitive*, *17*, 513–530.

Clark, E. V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Clark, E. V., & Berman, R. A. (1984). Structure and use in the acquisition of word formation. *Language*, *60*, 542–590.

Clark, E. V., & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language*, *14*, 547–567.

Clark, E. V., Gelman, S. A., & Lane, N. M. (1985). Compound nouns and category structure in young children. *Child Development*, *56*, 84–94.

Cutler, A., Hawkins, J. A., & Gilligan, G. (1985). The suffixing preference: A processing explanation. *Linguistics*, *23*, 723–758.

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*, 555–567.

De Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, *15*, 329–365.

Di Sciullo, A. M., & Williams, E. (1987). *On the definition of word*. Cambridge, MA: MIT Press.

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

Fabb, N. (1998). Compounding. In A. Spencer & A. M. Zwicky (Eds.), *The handbook of morphology* (pp. 66–83). Oxford: Blackwell Publishers.

Genesee, F., & Nicoladis, E. (2007). Bilingual first language acquisition. In E. Hoff and M. Shatz (Eds.), *Handbook of language development* (pp. 324–342). Oxford: Blackwell.

Green, D. W., Hammond, E. J., & Supramaniam, S. (1983). Letters and shapes: Developmental changes in search strategies. *British Journal of Psychology*, *74*, 11–16.

Hudson, P. T. W., & Buijs, D. (1995). Left-to-right processing of derivational morphology. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 383–396). Hillsdale, NJ: Erlbaum.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1612–1627.

Inhoff, A. W., Radach, R., & Heller, D. (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory and Language*, *42*, 23–50.

Järvikivi, J., & Niemi, J. (1999). Linearity and morphological structure in derived words: Evidence from category decision. *Brain and Language*, *68*, 340–346.

Juhasz, B. J., Inhoff, A. W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes*, *20*, 291–316.

Koester, D., Gunter, T. C., Wagner, S., & Friederici, A. D. (2004). Morphosyntax, prosody, and linking elements: The auditory processing of German nominal compounds. *Journal of Cognitive Neuroscience*, *16*, 1647–1668.

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

Krott, A. (2007). *Access and representation of open and concatenated English compounds: The role of the constituent family revisited*. Manuscript submitted for publication.

Krott, A., & Nicoladis, E. (2005). Large constituent families help children parse compounds. *Journal of Child Language*, *32*, 139–158.

Krott, A., Schreuder, R., & Baayen, R. H. (2002). Linking elements in Dutch noun-noun compounds: Constituent families as analogical predictors for response latencies. *Brain and Language*, *81*, 723–735.

Krott, A., Schreuder, R., Baayen, R. H., & Dressler, W. U. (2007). Analogical effects on linking elements in German compounds. *Language and Cognitive Processes*, *22*, 25–57.

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

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, *20*, 121–136.

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

Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, *4*, 1–21.

Neijt, A., Krebbers, L., & Fikkert, P. (2002). Rhythm and semantics in the selection of linking elements. In H. Broekhuis & P. Fikkert (Eds.), *Linguistics in the Netherlands 2002* (pp. 117–127). Amsterdam: Benjamins.

Nicoladis, E. (2001). When is a preposition a linking element? Bilingual children's acquisition of French compound nouns. *Folia Linguistica*, *36*, 45–63.

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*, 843–863.

Nicoladis, E. (2003). Cross-linguistic transfer in deverbal compounds of preschool bilingual children. *Bilingualism: Language and Cognition*, *6*, 17–31.

Nicoladis, E. (2006). Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Bilingualism: Language and Cognition*, *9*, 15–32.

Nicoladis, E. (2007). *Acquisition of deverbal compounds by French-speaking preschoolers*. *Mental Lexicon*, *2*, 79–102.

Nicoladis, E., Palmer, A., & Marentette, P. (2007). The role of type and token frequency in using past tense morphemes correctly. *Developmental Science*, *10*, 237–254.

Nicoladis, E., & Yin, H. (2002). The role of frequency in acquisition of English and Chinese compounds by bilingual children. *Proceedings of the 26th annual Boston University Conference on Language Development* (pp. 441–452). Somerville, MA: Cascadilla Press.

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*, *20*, 261–290.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 820–833.

Robinson, A. H. (1979). Observations on some deficiencies in the transformational model as applied to particular compound types in French. *Cahiers de Lexicologie*, *35*, 107–115.
Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory And Language*, *37*, 118–139.
Segui, J., & Zubizarreta, M.-L. (1985). Mental representation of morphologically complex words and lexical access. *Linguistics*, *23*, 759–774.
Selkirk, E. O. (1982). *The syntax of words*. Cambridge, MA: MIT Press.
Spence, N. C. W. (1980). The gender of French compounds. *Zeitschrift für Romanische Philologie*, *96*, 68–91.
Storms, G., & Wisniewski, E. J. (2005). Does the order of head noun and modifier explain response times in conceptual combination? *Memory and Cognition*, *33*, 852–861.
ten Hacken, P. (1994). *Defining morphology: A principled approach to determining the boundaries of compounding, derivation and inflection*. Hildesheim: Georg Olms Verlag.
Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, *11*, 61–82.
Van Jaarsveld, H. J., & Rattink, G. E. (1988). Frequency effects in the processing of lexicalized and novel compounds. *Journal of Psycholinguistic Research*, *17*, 447–473.

Appendix A

Table A1 Experimental Stimuli

Stimulus	Meaning	Head family	Modifier family	Head family size	Modifier family size
bouteille d'eau	Water bottle	8	18	High	High
chapeau-soleil	Sun hat	7	8	High	High
chef de police	Police chief	8	7	High	High
couteau à pain	Bread knife	7	8	High	High
cuiller à thé	Teaspoon	6	11	High	High
feu de bois	Wood fire	10	12	High	High
gâteau de mariage	Wedding cake	10	7	High	High
poudre pour bébé	Baby powder	7	10	High	High
boîte à lunch	Lunch box	30	1	High	Low

(Continued)

Appendix A *Continued*

Stimulus	Meaning	Head family	Modifier family	Head family size	Modifier family size
fête des mères	Mothers' Day	8	1	High	Low
jeu de blocs	Game of blocks	15	1	High	Low
jus de canneberge	Cranberry juice	10	1	High	Low
livre d'image	Picture book	12	1	High	Low
peau de banane	Banana peel	8	1	High	Low
sandwich au jambon	Ham sandwich	8	1	High	Low
soupe aux fèves	Bean soup	9	1	High	Low
arrêt d'autobus	Bus stop	1	12	Low	High
gants de hockey	Hockey gloves	1	8	Low	High
hamburger au fromage	Cheeseburger	1	9	Low	High
jouets de bain	Bath toys	1	13	Low	High
niche de chien	Doghouse	1	11	Low	High
pantalon de neige	Snow pants	2	11	Low	High
punch aux fruits	Fruit punch	1	8	Low	High
sapin de Noël	Christmas tree	1	24	Low	High
animaux de cirque	Circus animals	1	1	Low	Low
coussins de divan	Sofa cushions	1	1	Low	Low
échelle de corde	Rope ladder	1	1	Low	Low
enseignante de garderie	Daycare teacher	1	1	Low	Low
outils de jardinage	Garden tools	1	1	Low	Low
pastille de menthe	Mint	1	1	Low	Low
pattes de chaise	Chair legs	1	1	Low	Low
sonnerie d'alarme	Alarm bell	1	1	Low	Low

Appendix B

Brief Report: Do Children Report That They Are Familiar With Existing Compounds?

In order to see if preschool children report that they are familiar with existing French compounds, we asked 18 French-speaking 4-year-olds if they were familiar with the existing compounds in our study. To verify that their reports were accurate, we also asked them if they were familiar with some invented compounds (most often invented out of the parts of the existing compounds in our study; see below for the words we used). We administered children one of two versions of this test, each with half of the existing compounds in our study ($N = 16$) and half novel compounds ($N = 16$); see the Table B1 for a list of

Table B1 Items included in familiarity test

Version 1		Version 2	
Item	Category	Item	Category
bouteille d'eau	Existing	bouteille de corde	Novel
chapeau-soleil	Existing	chapeau-fromage	Novel
chef de police	Existing	chef de lunch	Novel
couteau à pain	Existing	couteau à canneberge	Novel
cuiller à lait	Novel	cuiller à thé	Existing
feu de pneus	Novel	feu de bois	Existing
gâteau de papier	Novel	gâteau de mariage	Existing
poudre pour ordinateur	Novel	poudre pour bébé	Existing
boîte à lunch	Existing	boîte à soleil	Novel
fête des mères	Existing	fête des bains	Novel
jeu de blocs	Existing	jeu d'autobus	Novel
jus de canneberge	Existing	jus d'eau	Novel
livre d'auto	Novel	livre d'image	Existing
peau de mariage	Novel	peau de banane	Existing
sandwich au bois	Novel	sandwich au jambon	Existing
soupe au thé	Novel	soupe aux fèves	Existing
arrêt d'autobus	Existing	arrêt de pain	Novel
gants de hockey	Existing	gants de divan	Novel
hamburger au fromage	Existing	hamburger aux mères	Novel
jouets de bain	Existing	jouets de pneus	Novel
niche de fèves	Novel	niche de chien	Existing
pantalon d'image	Novel	pantalon de neige	Existing
punch au jambon	Novel	punch aux fruits	Existing
sapin de banane	Novel	sapin de Noël	Existing
animaux de cirque	Existing	animaux de hockey	Novel
coussins de divan	Existing	coussins de police	Novel
échelle de corde	Existing	échelle de garderie	Novel
enseignante de garderie	Existing	enseignante de blocs	Novel
outils de fruits	Novel	outils de jardinage	Existing
pastille de neige	Novel	pastille de menthe	Existing
pattes de Noël	Novel	pattes de chaise	Existing
sonnerie de chien	Novel	sonnerie d'alarme	Existing

the compounds in each of the versions. The order of the items was randomized once for each version and the same order was administered to all children. Seven children were administered Version 1 and 11 children were administered

Version 2. The experimenter was a fluent speaker of French—for each item, she asked them if they had already heard the word. Before administering the test items, the experimenter gave two examples of compound words, one existing in French (i.e., *un camion de pompier* "a truck of fire-fighter") and one novel (i.e., *une tasse à pluie* "a cup for rain"). All of the children's responses were one of these three: yes, no, I don't know. As our dependent measure, we calculated the number of yes responses out of the total responses given by each child. All children answered all of the questions.

The children reported that they were familiar with our existing compounds for an average of 99.0% of the items. However, they also reported that they were familiar with an average of 98.6% of our novel compounds. Overall, a total of 98.8% of the children's 576 responses were yes. We would like to stress that the novel compounds were usually nonsensical like *pantalon d'image* literally "pants of picture" or *soupe au thé* literally, "soup with tea" (more naturally in English: tea soup).

Our conclusion is that this test did not yield valid results about whether the children were actually familiar with the existing compounds in our study.