University of Pennsylvania Working Papers in Linguistics

Volume 16
Issue 1 *Proceedings of the 33rd Annual Penn Linguistics Colloquium*

Article 7

1-1-2010

Examining the `Noun Bias': A Structural Approach

Rajdip Dhillon
Yale University, rajdip.dhillon@yale.edu

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/pwpl/vol16/iss1/7
For more information, please contact repository@pobox.upenn.edu.

Examining the `Noun Bias': A Structural Approach

Abstract
The current study examines whether a difference exists in the emergence of nouns and verbs in children's early vocabularies in languages possessing different parameter settings with respect to null arguments. Using data from CHILDES, proportions of noun and verb types are examined in the speech of native English-, Spanish-, and Mandarin-speaking children with ages ranging from 1;7-2;0, 2;1-2;5, and 2;6-2;11 and their respective caregivers. English- and Spanish-speaking children exhibit a noun bias across all age groups, yet consistent results are not obtained for Mandarin-speaking children in the first two age groups, although a verb bias is found in the third. While results for caregivers indicate that the input does differ across language types, their within-language results do not match those for children. The findings indicate that the structure of a language is more dominant than the input in determining the existence of a noun bias in children's early vocabularies.

This working paper is available in University of Pennsylvania Working Papers in Linguistics: http://repository.upenn.edu/pwpl/vol16/iss1/7

Examining the 'Noun Bias': A Structural Approach

Rajdip Dhillon

1 Introduction

The existence of a noun bias in children's early vocabularies is a point of much contention as far as whether a universal noun bias actually exists and how the noun bias is defined. Kauschke and Hofmeister (2002) offer four diverging views on what the noun bias indicates: (1) nouns are acquired earlier than verbs and other word classes; (2) nouns form the majority of children's early vocabularies; (3) nouns in children's early vocabulary are predominantly object labels; (4) a preference for nouns promotes further language development.

For Gentner (1982), nouns are acquired earlier as they offer less complexity than verbs both formally and conceptually as they tend to be more concrete and thus easier to comprehend. On the basis of Italian and English, Caselli et al. (1995) take a slightly different approach than Gentner in that they claim that a universal sequence exists such that children first acquire 'routines' (e.g. 'Hi' and 'Night-night') then nouns then verbs. However, Tardif (1996) argues that whether nouns precede verbs in acquisition or vice versa depends on the nature of the input and on linguistic factors such as the position of the noun or verb in a sentence. Furthermore, Bassano (2000) reports that verbs and nouns appear simultaneously in French, and Choi and Gopnik (1995) have found that a verb spurt occurs before a noun spurt in the early vocabulary of children acquiring Korean, as evidenced by Korean-speaking children's consistent use of verbs from an early age.

Regarding (2), the issue of how to quantify 'majority' arises. For some, in order for a noun bias to exist, the number of noun types must occupy over 50% of a child's entire early vocabulary (Bloom, Tinker and Margulis, 1993). For others, a noun bias exists so long as the number of noun types within a child's vocabulary outweighs all other categories (Pine, Lieven and Rowland, 1996).

The position taken in (3) is also supported by Gentner (1982), yet has been shown not to be representative of children's early vocabularies. Nelson, Hampson and Kessler Shaw (1993) found that nouns in children's early vocabularies are not predominantly object labels or basic-level terms. In fact, Nelson et al. determined that about 40% of the nouns children produced did not refer to basic-level objects but rather were non-object nouns, referring to generic, superordinate terms, events, locations, person roles, and natural phenomena. Bloom et al. (1993) discovered that words for basic-level objects represented about one third of the different words children learned.

Regarding (4), Snyder, Bates and Bretherton (1981) observed positive correlations between a child's vocabulary size and a large proportion of nouns at age 1;1. However, Bates et al. (1994) found no relation between an increased use of nouns in the early stages of lexical development and later linguistic abilities.

In addition to the difficulty in settling upon one definition of the noun bias, other issues and questions continue to linger. One large issue is whether the noun bias is indeed universal. Another issue involves why a difference exists in the acquisition of nouns and verbs.

The central purpose of this study is to investigate whether a difference exists in the emergence of nouns and verbs in children's early vocabularies in languages possessing different parameter settings with respect to the appearance of nouns. In terms of what constitutes a noun bias, the view expressed in (2) – that nouns form the majority of children's early vocabularies – is adopted in keeping with Pine et al.'s (1996) interpretation that the number of noun types within a child's vocabulary outweighs all other categories. Using data from CHILDES (MacWhinney, 2000), proportions of nouns and verbs used in productive speech are calculated for both children and caregivers whose native language is English, Spanish, or Mandarin. These particular languages were selected as they all differ on the dimension of the presence of null arguments and the degree to which null arguments permeate through the language. Null arguments are prohibited in English, whereas Spanish is a pro-drop language which allows for the omission of subject pronouns. Mandarin, on the other hand, is a topic drop language in which both subjects and objects can be omitted. The findings indicate that the structure of a language does determine whether or not a noun bias exists and to what extent it exists for that particular language, thus lending to the view that the noun bias is not actually universal, as has previously been claimed by Gentner (1982) and Caselli et al.

(1995). What is particularly unique about the current study is that it questions the very existence of a universal noun bias and investigates the degree to which a noun bias exists in the context of three languages which differ in their parameter settings with respect to null arguments. Previous studies have not examined the noun bias within this exact context, making the current study among the first to make a contribution to how language typology and the noun bias interact.

2 Is there a Cross-linguistic Noun Bias?

There are currently a number of varying accounts either confirming or denying the cross-linguistic nature of the noun bias. A noun bias has been detected in English (Bornstein et al., 2004; Caselli et al., 1995; Gentner, 1982), French (Bassano, 2000; Bornstein et al., 2004; Parisse and Le Normand, 2000; Poulin-Dubois, Graham and Sippola, 1995), Dutch (Bornstein et al., 2004; De Houwer and Gillis, 1998; Verlinden and Gillis, 1988), German (Gentner, 1982), Italian (Bornstein et al., 2004; Caselli et al., 1995; Tardif, Shatz and Naigles, 1997), Spanish (Bornstein et al., 2004; Jackson-Maldonado, Thal, Marchman, Bates and Gutierrez-Clellen, 1993), Hebrew (Bornstein et al., 2004; Maital, Dromi, Sagi and Bornstein, 2000), Kaluli (Gentner, 1982), and Japanese (Gentner, 1982; Sakurai, 1998; Yamashita, 1999).

Despite the overwhelming number of studies confirming a noun bias, conflicting data arises in examining Korean, Mandarin, and Turkish (Au, Dapretto and Song, 1994; Bornstein et al., 2004; Choi and Gopnik, 1995; Gentner, 1982; Gopnik and Choi, 1995; Pae, 1993; Tardif, 1996; Tardif, Gelman and Xu, 1999; Türkay, 2005). Such data leads to two possible conclusions: (1) the noun bias observed in early speech may actually be language dependent and not universal as has been previously claimed; (2) the methodology behind these studies may not be consistent or appropriate for measuring a noun bias.

Regarding (1), the claim taken up in this paper is that the noun bias is actually language dependent and is determined by a language's parameter setting governing the use of null arguments. Specifically, a language such as English, in which null arguments are prohibited, should exhibit a noun bias, and a language such as Spanish, in which subject pronouns can be omitted, should also exhibit a noun bias, but one which is not as strong as that of English. In Mandarin, a topic-drop language, a verb bias is expected. Because of other structural factors in Mandarin – such as morphological transparency – it is not predicted that a clear-cut noun bias or verb bias will initially emerge but rather that Mandarin-speaking children will eventually move toward a verb bias.

Returning to the contrasting results mentioned earlier, how is it that children acquiring a language such as Mandarin can be found to exhibit a noun bias in some studies yet a verb bias in others? In Gentner's (1982) cross-linguistic study, a noun bias was found for Mandarin. Gentner's results prove to be unreliable on account of small sample sizes and inconsistent methods of data collection. Furthermore, Choi and Gopnik (1995) and Gopnik and Choi (1995) suggest that the proposed universality of the noun bias may be due to the methods employed in measuring the noun bias.

3 Why is there a Difference in the Acquisition of Nouns and Verbs?

In proposing that the noun bias is not universal, the consequence that either nouns or verbs emerge earlier in some languages – or that both emerge simultaneously – arises. In addition to nativist approaches which claim that syntactic categories are innately given, the difference in the acquisition of nouns and verbs is often attributed to semantic and structural saliency, phonetic saliency, pragmatic factors, and input factors.

Gentner (1982) suggests that nouns are universally more accessible to children acquiring language than verbs. For Gentner, nouns are mapped onto cohesive perceptual entities and allow for the partitioning of objects into natural categories more so than verbs, while verbs are more complex in that they label more language-specific categories, and thus require language input and more time to acquire them. However, in verb-friendly languages such as Mandarin and Korean, given the studies indicating that a verb bias exists in these languages as well as Bassano's (2000) claim that verbs and nouns are acquired simultaneously in French, Gentner's notions of verb acquisition is not cross-linguistically accurate.

Maratsos (1991) proposes a distinction between semantics and morphology to account for the existence of a noun bias. For Maratsos, the core of the noun category involves semantic properties and the core of the verb category involves shared morphological properties in terms of inflection. The semantic core of verbs is not particularly substantial in the acquisition of verbs. While children acquire nouns on the basis of their semantic properties and their abilities to use nouns as a means for concrete object reference, children acquire verbs on the basis of similar morphological features. Since the morphological transparency of verbs in Mandarin operates as a counterpoint to Maratsos' account, it becomes apparent that Maratsos' explanation may only apply to those languages which exhibit a definitive noun bias and are morphologically complex.

Purely structural differences which are likely to affect the acquisition of nouns and verbs revolve around the morphological and syntactic properties of a language. Yamashita (1999) argues that children acquiring Japanese acquire nouns earlier since nouns have the least amount of variation in the linguistic input. Verbs in Japanese, however, are acquired later as they possess the most morphological variation. In English, nouns have lesser morphological variation in comparison to verbs, and so it is no surprise that a noun bias is found in English. Mandarin is morphologically transparent and so there are no morphological markings on either nouns or verbs which serve to assist the language-learning child in distinguishing between the two categories. Based solely on Yamashita's argument for Japanese, neither nouns nor verbs in Mandarin should have an advantage initially in terms of which comprise the bulk of a child's early vocabulary.

Another structural factor involves the structural saliency of nouns. In a language such as English which deems subject nouns and object nouns obligatory, nouns are highly salient. However, in a pro-drop language such as Spanish and a topic-drop language such as Mandarin, nouns lose some amount of structural saliency. The level of structural saliency of nouns will be present in the input a child receives given the particular parameters governing the appearance of nouns in the language of input. As a result, it would be expected that a child acquiring English would exhibit a larger noun bias compared to a child acquiring Mandarin on account of the structural saliency of nouns in the input, as well as other contributing factors. So to say that structural saliency is a factor in predicting whether or not a noun bias exists is to say that it is the input a child receives which determines whether or not the noun bias will appear in a child's early vocabulary.

Regarding phonetic saliency, Slobin (1973) postulates that words in the utterance-final position highly salient as they are bound by silence and that the utterance-final position exaggerates pitch, which together aid the extraction of words from the speech stream. Mandarin is an SOV language and the verb is often in the utterance-final and phonetically salient position. These characteristics of Mandarin contribute to the higher degree of saliency of verbs in the language and perhaps even to a verb bias. In an examination of word order in caregiver language and its relation to the proportion of nouns and verbs in children's early vocabulary, Tardif et al. (1997) observed that Mandarin-speaking children produced more verbs than nouns and English- and Italian-speaking children did the opposite. Tardif et al. concluded that the tendencies of the children to produce more nouns or vice versa matched the likelihood of a noun or verb to occur in the utterance-final position in the caregivers' speech. Despite this, Slobin's claim bears a number of ramifications. The structural properties of English are such that English is an SVO language, and so object nouns and intransitive verbs occupy the phonetically salient position. Is this to say that children acquiring language have more access to object nouns than subject nouns, indicating that the content of object nouns is easier for children to access and acquire? This seems rather unlikely. Furthermore, how does Slobin's notion of phonetic saliency translate to languages with varying word order if morphological complexity is not an issue?

Pragmatic and input factors are also applicable and operate in tandem. An example of pragmatic differences that have the potential to affect the acquisition of nouns and verbs is that English-speaking children may learn nouns early because nouns are salient in the input pragmatically as English-speaking mothers may engage in object-naming activities with their children more (Choi and Gopnik, 1995) than Mandarin-speaking mothers do.

A number of studies have been conducted to examine caregivers' input in a variety of languages. Kim, McGregor and Thompson (2000) examined caregivers' input in English and Korean, finding that Korean-speaking caregivers used more verb types than noun types and English-speaking caregivers did the opposite. Kim et al. also detected that Korean-speaking caregivers used more verb tokens on average than English-speaking caregivers. Choi and Gopnik (1995) ob-

served that Korean-speaking mothers provide more action verbs but fewer object nouns than English-speaking mothers, and Korean-speaking mothers engage in activity-oriented conversation more than English-speaking mothers. Choi and Gopnik's study suggested that early lexical development is influenced by a caregivers' language-specific input and, specifically, that verbs, like nouns, are available to children early on and may be acquired early by children who receive the appropriate encouragement to do so by their language-specific grammar and the input they receive.

4 Current Study

The current study examines the noun bias in the context of English, Spanish, and Mandarin – three typologically different languages. The primary difference among these languages with which this study is concerned is their tolerance of null arguments. English prohibits null arguments except in imperatives and restricted contexts such as diary and instructional registers. In these cases, it is the subject which is the null argument and not the object. Outside of these contexts, null arguments generally do not appear in English. Spanish is a pro-drop language which allows for null subject pronouns. Subject pronouns are generally dropped but are used to emphasize the subject of a verb or to indicate contrast. Spanish is morphologically complex and so its verbs are marked with number, person, and tense. Since the verb carries sufficient inflection, it is possible to identify the person and number of the subject when the subject has been dropped. Mandarin is a topic-drop language and so it allows for both subject and object drop. Mandarin is morphologically transparent and lacks agreement morphology. Mandarin verbs are marked only for aspect with no subject-verb agreement and Mandarin nouns are marked with a numeral-classifier compound that precedes the noun (Türkay, 2005). Neither aspect nor the numeral-classifier compound alters the verb or noun stem. While Spanish uses its agreement morphology to identify the number and person of a dropped subject, dropped arguments in Mandarin must be made obvious from context. In examining the speech of Mandarin-speaking adults, Wang, Lillo-Martin, Best and Levitt (1992) found that subjects were dropped in roughly 35% of all sentences while objects were dropped in about 10%. For Mandarin-speaking children at 2;0, subjects were dropped in 47% of all sentences and objects were dropped in 23%. Wang et al.'s results indicate that, while children encounter both null subjects and null objects in the input, they come across more null subjects and consequently produce more null subjects than null objects in their productive speech.

Wang et al. (1992) also examine the use of null arguments in child language and find that English-speaking children at 2;0 drop subjects in 33% of their sentences while they drop objects in only 4%. In comparison to Mandarin-speaking children at 2;0, who drop subjects in 47% of their sentences and drop objects in 23%, Wang et al.'s findings suggest that the rates of null arguments in child language are related to a language's tolerance of null arguments.

Method

Data

Files containing observational data for English-speaking, Spanish-speaking, and Mandarin-speaking children in naturalistic settings or toy-play contexts were selected from the CHILDES database. Three age levels were established for the children: 1;7 to 2;0, 2;1 to 2;5, and 2;6 to 2;11. The first age level is referred to as Group 1, the second as Group 2, and the third as Group 3. Data for ten children at each of the three age levels for English and Mandarin was used. For Spanish, data was used for three children in Group 1, three children in Group 2, and four children in Group 3. Data on caregivers' speech was elicited using the same file set as for the children. The files containing caregiver data were separated based on the age levels established for the children. Data for ten caregivers at each of the children's three age levels for English and Mandarin was used. For Spanish, data was used for three caregivers with children in Group 1, two with children in Group 2, and four with children in Group 3.

Procedure

Analysis of the files was required to determine the inclusion and exclusion of particular word

types. Three categories of word types were established: 'nouns', 'verbs', and 'other'. The 'other' category included adjectives, adverbs, conjunctions, prepositions, determiners, and the like.

Previous noun bias studies have counted noun and verb categories in the manner that they are counted in adult language (Caselli et al., 1995), and so the same method of counting nouns and verbs was adopted in the current study. In examining nouns, types and tokens were counted. Within the counts for noun types and tokens, proper nouns, pronouns, concrete nouns, and abstract nouns were included. Within the counts for verb types and tokens, auxiliary verbs and gerunds were included. Gerunds were included as they associate to verbs cognitively (Türkay, 2005). Within the counts for types and tokens for words in the 'other' category, onomatopoeic words, disfluencies, and other portions of speech which were difficult to categorize were excluded.

In counting word types, all the forms in which a particular word appeared in a child's vocabulary were counted as a separate type. For example, 'dada' and 'daddy' were counted as the single word type 'dada'. All inflected instantiations of the same lexical entry were all counted as a single word type matching the applicable lexical entry. For example, 'jumps', 'jumped', and 'jumping' were all counted as an instantiation of the word type 'jump'.

Single-use words were also examined. Those which arose spontaneously from the child were included in the counts while those elicited by other participants within the observational settings were eliminated from token and type counts.

For each of the three age levels examined, some corpora contained only one file for a child at a particular age level and the corresponding caregiver, while other corpora contained multiple files for the same child at a particular age level and the corresponding caregiver. In the case that there were multiple files corresponding to the same age level for one child and the corresponding caregiver, those files were examined individually and then averaged.

Given that the lengths of the files obtained from the CHILDES database varied from file to file, using counts for types and tokens within the noun, verb, and 'other' category across the three languages and the three age levels was not an appropriate method. Instead, counts were transformed into proportions to make the results obtained comparable.

Results

A strong noun bias is evident in English-speaking children's early vocabularies and a fairly strong noun bias is evident in Spanish-speaking children's early vocabularies. The number of noun types produced at each age level is significantly larger than the number of verb types and 'other' types produced. A Chi-square analysis performed on the number of noun types versus the number of verb types found at each age level confirmed that the children produced more nouns than verbs at each age level (English: χ^2 (1, N=3) = 75.577, $p < 0.001$; Spanish: (χ^2 (1, N=3) = 17.625, $p < 0.001$). A Chi-square analysis performed on the number of noun types versus the number of 'other' types found at each age level confirmed that the children produced more nouns than 'other' types at each age level (English: χ^2 (1, N=3) = 48.815, $p < 0.001$; Spanish: (χ^2 (1, N=3) = 6.512, $p < 0.05$). A noun-verb ratio, which is the proportion of nouns divided by the sum of the proportion of nouns and the proportion of verbs, indicates a noun bias when its value is greater than 0.50 and a verb bias when its value is less than 0.50. For the English-speaking children, the noun-verb ratio performed on types at each level also indicates a strong noun bias relative to verb types, as the values for the ratio are noticeably higher than 0.50 (Group 1: 0.771; Group 2: 0.677; Group 3: 0.643). The noun-verb type ratio at each level for the Spanish-speaking children indicates a fairly strong noun bias relative to verb types, as the values for the ratio are higher than 0.50 (Group 1: 0.700; Group 2: 0.649; Group 3: 0.597).

The results for the Mandarin-speaking children were quite different than those for the English- and Spanish-speaking children. While the English- and Spanish-speaking children exhibited a noun bias across all three age groups, the Mandarin-speaking children did so only in Group 1. Because the noun-verb type ratio is above 0.50 for Group 1, this is considered evidence in support of a noun bias (Group 1: 0.545; Group 2: 0.457; Group 3: 0.422). In addition, the difference between the number of noun types produced at each age level and the number of verb types produced at each age level was found to be statistically significant. A Chi-square analysis performed on the number of noun types versus the number of verb types found at each age level confirmed that there was a difference between the two categories (χ^2 (1, N=3) = 6.464, $p < 0.05$). A Chi-square

analysis performed on the number of noun types versus the number of 'other' types found at each age level did not confirm that there was a difference between the two categories (χ^2 (1, N=3) = 2.368, p = 0.306).

In examining the individual results for each of the children within each age group, it was found that, in Group 1, six children exhibited a mild noun bias and four exhibited a verb bias as defined by the value of the noun-verb type ratio. In Group 2, five children exhibited a mild noun bias and five exhibited a verb bias as defined by the value of the noun-verb type ratio. In Group 3, one child exhibited a noun bias with a noun-verb ratio of 0.571 for types and 0.509 for tokens and the other nine exhibited a verb bias. When examining only those children in Groups 1 and 2 exhibiting a noun bias, the noun-verb ratio for types was 0.608 for those children in Group 1 and 0.531 for those in Group 2. When examining only those children in Groups 1 and 2 exhibiting a verb bias, the noun-verb ratio for types was 0.451 for those children in Group 1 and 0.383 for those in Group 2.

For Groups 1 and 2, the number of noun types produced by the children with a noun bias is not significantly larger than the number of verb types produced. A Chi-square analysis performed on the number of noun types versus the number of verb types found at the first two age levels did not confirm that these children produced more nouns than verbs (χ^2 (1, N=2) = 3.799, p = 0.051). While the noun-verb type ratio indicated a noun bias for these children, noun dominance over verbs was not statistically proven in this case. Similarly, for Groups 1 and 2, the number of verb types produced by the children with a verb bias is not significantly larger than the number of noun types produced. A Chi-square analysis performed on the number of verb types versus the number of noun types found at the first two age levels did not confirm that these children produced more verbs than nouns (χ^2 (1, N=2) = 2.213, p = 0.137). While the noun-verb type ratio indicated a verb bias for these children, verb dominance over nouns was not statistically proven in this case. These results indicate that, for Groups 1 and 2, no clear noun or verb bias exists.

Regarding English- and Spanish-speaking caregivers, the noun-verb type ratio at each level indicates a noun bias, as the values for the ratio are higher than 0.50 (English: Group 1: 0.616; English: Group 2: 0.594; English: Group 3: 0.592; Spanish: Group 1: 0.585; Spanish: Group 2: 0.587; Spanish: Group 3: 0.561). The number of noun types produced at the corresponding age levels for the children is not significantly larger than the number of verb types for either language group. A Chi-square analysis performed on the number of noun types versus the number of verb types found at each age level did not confirm that the caregivers produced significantly more nouns than verbs (English: χ^2 (1, N=3) = 1.057, p = 0.589; Spanish: χ^2 (1, N=3) = 1.547, p = 0.461). Similarly, the number of noun types produced at the corresponding age levels for the children is not significantly larger than the number of 'other' types for either language group. A Chi-square analysis performed on the number of noun types versus the number of 'other' types found at each age level did not confirm that the caregivers produced significantly more nouns than 'other' types (English: χ^2 (1, N=3) = 2.687, p = 0.261; Spanish: χ^2 (1, N=3) = 1.733, p = 0.420). Even though the noun-verb type ratio points to a noun bias, noun dominance over verbs was not statistically proven for the English- or Spanish-speaking caregivers. In comparison to their caregivers, the English- and Spanish-speaking children exhibited a noun-verb type ratio greater than 0.50 and produced significantly more noun types than verb or 'other' types.

Regarding Mandarin-speaking caregivers, the noun-verb type ratio at each level indicates a verb bias, as the values for the ratio are less than 0.50 (Group 1: 0.436; Group 2: 0.417; Group 3: 0.441). The number of verb types produced at the corresponding age levels for the children is not significantly larger than the number of noun types. A Chi-square analysis performed on the number of verb types versus the number of noun types found at each age level did not confirm that the caregivers produced significantly more verbs than nouns (χ^2 (1, N=3) = 2.178, p = 0.337). Similarly, the number of verb types produced at the corresponding age levels for the children is not significantly larger than the number of 'other' types. A Chi-square analysis performed on the number of verb types versus the number of 'other' types found at each age level did not confirm that the caregivers produced significantly more verbs than 'other' types (χ^2 (1, N=3) = 1.153, p = 0.562). Even though the noun-verb type ratio points to a verb bias, verb dominance over nouns was not statistically proven for the Mandarin-speaking caregivers. In comparison to their caregivers, the collapsed data for the Mandarin-speaking children indicated that the difference between

the number of noun types and verb types produced was statistically significant while the stratified data for Groups 1 and 2 did not point to statistical significance.

In examining the individual results of the thirty total caregivers, all but four exhibited a verb bias. In Group 1, one caregiver exhibited a mild noun bias with a noun-verb type ratio of 0.531 and another caregiver exhibited neither a noun nor a verb bias, as the noun-verb type ratio for this particular caregiver was 0.500. In Group 3, two caregivers exhibited a mild noun bias with noun-verb type ratios of 0.541 and 0.546. Of these four caregivers who did not exhibit a verb bias, none of them corresponded to children who exhibited a noun bias.

In a cross-linguistic comparison, in which the goal is to see if the number of nouns is higher in one language when compared to another for the children, the difference between the number of nouns and verbs between each of the languages is statistically significant across all age groups (English vs. Spanish: Group 1: χ^2 (1, N=2) = 10.934, $p < 0.0005$; Group 2: χ^2 (1, N=2) = 6.576, $p < 0.05$; Group 3: χ^2 (1, N=2) = 9.545, $p < 0.0005$; English vs. Mandarin: Group 1: χ^2 (1, N=2) = 167.9634, $p < 0.0001$; Group 2: χ^2 (1, N=2) = 231.113, $p < 0.0001$; Group 3: χ^2 (1, N=2) = 110.292, $p < 0.0001$; Spanish vs. Mandarin: Group 1: χ^2 (1, N=2) = 9.545, $p < 0.0001$; Group 2: χ^2 (1, N=2) = 120.640, $p < 0.0001$; Group 3: χ^2 (1, N=2) = 42.822, $p < 0.0001$).

For instance, the difference in the number of nouns and verbs in English and Spanish for Group 1 is statistically significant, indicating that the proportions of noun and verb types produced by the English-speaking children in Group 1 differ from the proportions produced by the Spanish-speaking children in the same group. The same results were obtained for the caregivers.

Because the noun-verb type ratios for the English-speaking children are higher than those for the Spanish-speaking children and the Mandarin-speaking children, this information coupled with the results of the Chi-square analyses indicates that the English-speaking children produced a higher proportion of nouns than the Spanish-speaking children and the Mandarin-speaking children. Similarly, because the noun-verb type ratios for the Spanish-speaking children are higher than those for the Mandarin-speaking children, this information coupled with the results of the Chi-square analyses indicates that the Spanish-speaking children produced a higher proportion of nouns than the Mandarin-speaking children. The same generalizations also apply to the caregivers.

What is particularly notable here is that the above results clearly indicate that there is a difference between Spanish and Mandarin – specifically, this provides evidence for the classification of pro-drop versus topic-drop.

5 Discussion

Addressing the noun bias in a broad sense, the results indicate that the noun bias is not actually universal. While the noun-verb type ratios across age groups for the English- and Spanish-speaking children revealed a noun bias, the same did not occur for the Mandarin-speaking children. Instead, no clear noun or verb bias was apparent for Groups 1 and 2 of the Mandarin-speaking children, but a verb bias emerged in Group 3. Obviously, if the noun bias were indeed universal, this would require a clear noun bias being present within the results for the Mandarin-speaking children. As this did not occur, it can be concluded that, contrary to Gentner (1982) and Caselli et al. (1995), the noun bias is not universal.

In looking at the children's noun-verb type ratios across age groups, English-speaking children boast the highest ratios, followed by Spanish-speaking children, and Mandarin-speaking children have the lowest. What this reveals is that there is a difference in the emergence of nouns and verbs in children's early vocabularies in languages possessing different parameter settings with respect to the tolerance of null arguments. Specifically, there is a difference between languages prohibiting null arguments and pro-drop languages, between languages prohibiting null arguments and topic-drop languages, and between pro-drop languages and topic-drop languages. This finding is further confirmed by the results of the Chi-square analyses conducted on cross-linguistic data for children which revealed that the difference between the number of nouns and verbs between each of the languages is statistically significant across all age groups. What can be concluded from these results is that children acquiring a language such as English which prohibits null arguments will exhibit a large noun bias. Children acquiring a pro-drop language such as Spanish will also exhibit noun bias – but one which is smaller in comparison to that of children acquiring a language

which prohibits null arguments.

In arriving at a conclusion for topic-drop languages such as Mandarin, the situation becomes more complex. Turkish (Öztürk, 2002; Türkay, 2005) and Korean (Choi and Gopnik, 1995; Gopnik and Choi, 1995; Kim, 2000) are both topic-drop languages which are morphologically complex. Children who speak Korean have been found to exhibit a verb bias while children who speak Turkish have not been found to exhibit a noun or a verb bias. In comparison to Turkish, and Korean, which are morphologically complex and place morphological markings on nouns and verbs which can potentially assist the language-learning child in distinguishing between the two categories, Mandarin is morphologically transparent. As such, Mandarin lacks morphological markings which can serve to distinguish nouns from verbs, thus making the language-learning child's task of identifying nouns and verbs more complicated. While children acquiring languages prohibiting null arguments will exhibit a large noun bias and children acquiring pro-drop languages will exhibit a smaller noun bias, children acquiring topic-drop languages crucially will not exhibit a noun bias. Taken together, what is implied by the findings of the current study, as well the studies on Turkish (Türkay, 2005) and Korean (Choi and Gopnik, 1995; Gopnik and Choi, 1995), is that children acquiring morphologically complex topic-drop languages will either exhibit a verb bias or they will not initially exhibit a consistent verb or noun bias and children acquiring morphologically transparent topic-drop languages will not initially exhibit a consistent noun or verb bias but will eventually develop a verb bias. Why either a verb bias or a lack of a consistent noun or verb bias emerges in morphologically complex topic-drop languages may be due to variability in word order. Turkish (Türkay, 2005) permits a more variable word order than Korean, which allows some variation so long as the verb is always final (Lee and Ramsey, 2000). Because the verb is consistently final in Korean, it is easier for children to detect verbs based on their position within a sentence and so a verb bias emerges. In Turkish, based on its morphology and topic-drop nature, a verb bias would be expected. However, the high level of variability in word order interferes with the potential for a verb bias to emerge, leading to neither a noun nor a verb bias being present. In morphologically transparent topic-drop languages such as Mandarin, the lack of morphological markings on both nouns and verbs results in neither category having an advantage in terms of forming a bias until the child becomes more attuned to the verb-friendly nature of the language and forms a verb bias.

While the child data shows that languages prohibiting null arguments, pro-drop languages, and topic-drop languages reflect the noun bias differently, does the same occur in the caregiver data? In short, the answer to this question is yes. In examining the Chi-square analyses conducted on cross-linguistic data on caregivers, the difference between the number of nouns and verbs produced by the caregivers between each of the languages was statistically significant. While noun dominance over verbs or verb dominance over nouns was not statistically proven in the Chi-square analyses conducted on each individual language for the caregivers, the noun-verb type ratios were highest for the English-speaking caregivers, lowest for the Mandarin-speaking caregivers, and intermediate between the values for the English- and Mandarin-speaking caregivers for the Spanish-speaking caregivers. However, when looking at the noun-verb type ratios, unlike the Mandarin-speaking children, the Mandarin-speaking caregivers did produce a consistent verb bias. What can be concluded from these results is that, in the input, children hear a higher proportion of nouns than verbs in languages prohibiting null arguments like English, a higher proportion of nouns than verbs in pro-drop languages – but this proportion is lower than that of languages prohibiting null arguments, and a higher proportion of verbs than nouns in topic-drop languages. To reiterate, the input differs across language types. The input English-speaking children are exposed to is not the same as the input Mandarin-speaking or Spanish-speaking children hear. English-speaking caregivers use more nouns than verbs, so their children hear more nouns than verbs. Mandarin-speaking caregivers use more verbs than nouns, so their children hear more verbs than nouns.

What the caregiver results indicate is that the structural differences across the three language types examined are indeed present in the input. This finding begs the question of whether the input or the structure of a language is more dominant in determining the existence of a noun bias in children's early vocabularies. In looking at the noun-verb type ratios for English- and Spanish-speaking children and caregivers, while they all exhibited a noun bias, the children exhibited a larger noun bias than the caregivers in all cases. Additionally, noun dominance over verbs was statistically proven for the English- and Spanish-speaking children but not their respective care-

givers. For Mandarin, six children in Group 1, five children in Group 2, and one child in Group 3 exhibited a noun bias, as defined by the noun-verb type ratio. With respect to caregivers, one with a child in Group 1 and two with children in Group 3 exhibited a noun bias, and one with a child in Group 1 exhibited neither a noun nor a verb bias. More interestingly, none of the Mandarin-speaking children exhibiting a noun bias were the children of these caregivers. In fact, the caregivers who did not exhibit a verb bias all had children who did exhibit a verb bias. Because the results across Mandarin-speaking children and caregivers vary in this manner – in that the children's production does not match the caregiver's production, it can be concluded that it is the structure of a language which is more dominant in determining the existence of a noun bias than the input. While the input in Mandarin is generally more verb dominant, the morphological transparency of the language prevents the emergence of an initial verb bias in the children's vocabularies.

6 Conclusion

To summarize, these data allow for a number of general conclusions. It is now evident that the noun bias is not universal. There is a difference in the emergence of nouns and verbs in children's early vocabularies in languages possessing different parameter settings with respect to the tolerance of null arguments. While a difference in the emergence of nouns and verbs between languages prohibiting null arguments and topic-drop languages permitting null subjects and objects is expected, the current study shows that a difference in the emergence of nouns and verbs between these two types of languages and pro-drop languages should also be expected. Similarly, this difference is also reflected in the input when noun-verb type ratios are examined. In addition to these parameter settings, the level of variability in word order and the level of morphological transparency or complexity must be considered when anticipating a noun or verb bias. And finally, the structure of a language is more dominant than the input in terms of determining the existence of a noun bias in children's early vocabularies.

In a broader perspective, what is particularly interesting about the current study is that it justifies the interaction of language typology in language acquisition. By examining both children's and caregivers' speech in three typologically distinct languages, important insights on the role of structure versus input have been gained.

References

Au, Terry, Dapretto, Mirella, and You–Kyung Song. 1994. Input vs. constraints: early word acquisition in Korean and English. *Journal of Memory and Language* 35: 567–82.

Bassano, Dominique. 2000. Early development of nouns and verbs in French: exploring the interface between lexicon and grammar. *Journal of Child Language* 27: 521–59.

Bates, Elizabeth, Marchman, Virginia, Thal, Donna, Fenson, Larry, Dale, Philip, Reznick, J. Steven, Reilly, Judy, and Jeff Hartung. 1994. Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* 21: 85–121.

Bloom, Lois, Tinker, Erin, and Cheryl Margulis. 1993. The words children learn: evidence against a noun bias in early vocabularies. *Cognitive Development* 8: 431–50.

Bornstein, Marc H., Cote, Linda R., Maital, Sharone, Painter, Kathleen, Park, Sung–Yun, Pascual, Liliana, Pêcheux, Marie–Germaine, Ruel, Josette, Venuti, Paola, and Andre Vyt. 2004. Cross–linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, Vol. 75, Num. 4: 1115–39.

Caselli, Maria Cristina, Bates, Elizabeth, Casadio, Paola, Fenson, Judi, Fenson, Larry, Sanderl, Lisa, and Judy Weir. 1995. A cross–linguistic study of early lexical development. *Cognitive Development* 10: 159–99.

Choi, Soonja, and Alison Gopnik. 1995. Early acquisition of verbs in Korean: a cross–linguistic study. *Journal of Child Language* 22: 497–529.

De Houwer, Annick, and Steven Gillis. 1998. Dutch child language: an overview. In *The Acquisition of Dutch*, eds. S. Gillis and A. De Houwer, 1–100. Amsterdam: John Benjamins.

Gentner, Dedre. 1982. Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In *Language Development*, Vol. 2, ed. S. Kuczaj, 301–334. Hillsdale, NJ: Erlbaum.

Gopnik, Alison, and Soonja Choi. 1995. Names, relational words, and cognitive development in English– and Korean–speakers: nouns are not always learned before verbs. In *Beyond Names for Things: Young Children's Acquisition of Verbs*, eds. M. Tomasello and W. Merriman, 63–80. Hillsdale, NJ: Erlbaum.

Jackson–Maldonado, Donna, Thal, Donna, Marchman, Virginia, Bates, Elizabeth, and Vera Gutierrez–Clellen. 1993. Early lexical development in Spanish–speaking infants and toddlers. *Journal of Child Language* 20: 523–49.

Kauschke, Christina, and Christoph Hofmeister. 2002. Early lexical development in German: a study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language* 29: 735–57.

Kim, Mikyong, McGregor, Karla K., and Cynthia K. Thompson. 2000. Early lexical development in English–and Korean–speaking children: language–general and language–specific patterns. *Journal of Child Language* 20: 225–54.

Kim, Young–Joo. 2000. Subject/object drop in the acquisition of Korean: a cross–linguistic comparison. *Journal of East Asian Linguistics* 9: 325–351.

Lee, Iksop, and S. Robert Ramsey. 2000. *The Korean Language*. New York: SUNY Press.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Maital, Sharone L., Dromi, Esther, Sagi, Abraham, and Marc H. Bornstein. 2000. The Hebrew communicative development inventory: language specific properties and cross–linguistic generalizations. *Journal of Child Language* 27: 43–67.

Maratsos, Michael. 1991. How the acquisition of nouns may be different from that of verbs. In *Biological and Behavioral Determinants of Language Development*, eds. N. Krasnegor, D. Rumbaugh, R. Schiefelbusch and M. Studdert–Kennedy, 67–88. Hillsdale, NJ: Erlbaum.

Nelson, Katherine, Hampson, Juno, and Lea Kessler Shaw. 1993. Nouns in early lexicons: evidence, explanations and implications. *Journal of Child Language* 20: 61–84.

Öztürk, Balkız. 2002. Turkish as a non–pro–drop language. In *The Verb in Turkish*, ed. E. Erguvanlı, 239–259. Amsterdam: John Benjamins.

Pae, Soyeong. 1993. Early vocabulary in Korean: are nouns easier to learn than verbs? *Dissertation Abstracts International* 55, 1253A.

Parisse, Christophe, and Marie–Therese Le Normand. 2000. How children build their morphosyntax: the case of French. *Journal of Child Language* 27: 267–92.

Pine, Julian M., Lieven, Elena V.M., and Caroline Rowland. 1996. Observational and checklist measures of vocabulary composition: what do they mean? *Journal of Child Language* 23: 573–89.

Poulin–Dubois, Diane, Graham, Susan, and Lorrie Sippola. 1995. Early lexical development: the contribution of parental labeling and infants' categorization abilities. *Journal of Child Language* 22: 325–43.

Sakurai, Chikako. 1998. Why nouns before verbs? A case study of Japanese early lexical acquisition. In *CLS 34: The Panels*, 511–19.

Slobin, Dan I. 1973. Cognitive prerequisites for the development of grammar. In *Studies of Child Language Development*, eds. C. A. Ferguson and D. I. Slobin. New York: Holt, Reinhart and Winston.

Snyder, Lynn, Bates, Elizabeth, and Inge Bretherton. 1981. Content and context in early lexical development. *Journal of Child Language* 8: 565–82.

Tardif, Twila. 1996. Nouns are not always learned before verbs: evidence from Mandarin speakers' early vocabulary. *Developmental Psychology* 32: 492–504.

Tardif, Twila, Gelman, Susan A. and Fan Xu. 1999. Putting the 'noun bias' in context: a comparison of English and Mandarin. *Child Development* Vol. 70, Num. 3: 620–35.

Tardif, Twila, Shatz, Marilyn, and Naigles, Letitia. 1997. Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian, and Mandarin. *Journal of Child Language* 24: 535–65.

Türkay, Feyza A. 2005. *Children's Early Lexicon in Terms of Noun/Verb Dominance*. Unpublished doctoral dissertation, Çukurova University.

Verlinden, Ann, and Steven Gillis. 1988. Nouns and verbs in the input: Gentner (1982) reconsidered. In *First Language Acquisition*, ed. F. Van Besien, 163–187. Antwerp, Belgium: ABLA.

Wang, Qi, Lillo–Martin, Diane, Best, Catherine, and Andrea Levitt. 1992. Null subject versus null object: some evidence from the acquisition of Chinese and English. *Language Acquisition* 2: 221–254.

Yamashita, Yoshie. 1999. The acquisition of nouns and verbs in young Japanese children: why do verbal nouns emerge early? In *BUCLD 23 Proceedings*, 741–52.

Department of Linguistics
Yale University
370 Temple Street, Room 204
New Haven, CT 06511
rajdip.dhillon@yale.edu