2000. In J. Tai and Y. Chang (eds.), *Proceedings of the Seventh International Symposium on Chinese Languages and Linguistics*.

The Limits of Observation: Can the Situations in Which Words Occur Account For Cross-Linguistic Variation in Vocabulary Composition?

Jesse Snedeker and Peggy Li
University of Pennsylvania

Abstract

Crosslinguistic variation in vocabulary composition is usually attributed to differences in the structural properties of languages and correlated differences in the frequency and salience of different word classes in the input. Here we explore whether another property of the input, the extralinguistic contexts of word use, can account for the similarities and differences in the early vocabularies of English and Mandarin learners. Early word learning is limited by the child's initial representation of the input. Since novice language learners know few words and little syntax, they must initially learn words from the extralinguistic contexts in which they appear. By simulating this learning procedure with adults we can determine whether the information in word-to-world pairs is sufficient to account for the composition of early vocabularies in both linguistic environments.

We assessed the potency of word-to-world mapping in each language by asking adult native-speakers to identify nouns and verbs from their extralinguistic contexts in maternal speech. We found that verbs were identified more often in Mandarin ($p < .001$) while nouns were identified more often in English ($p < .001$). Even after the data were corrected to compensate for potential differences in response bias, there was still a large and persistent difference in verb identification. Subjects in both groups performed well on action verbs and poorly on verbs referring to unobservable relations or mental states. Since Mandarin mothers used more action verbs than the Americans, the Mandarin subjects were better able to identify their target verbs.

These results indicate that the information available in extralinguistic contexts can account both for the early acquisition of verbs in Mandarin and for the predominantly nominal initial vocabulary of English learners. Cognitively mature adults, forced to rely on the evidence of their senses, were able to identify only those words that can be learned by children who are similarly impaired.

Introduction

A child's vocabulary is the product of a constantly evolving interaction between child and her experiences with language. One of the primary challenges for research on lexical development is unraveling the respective contributions of the learner and the input and tracing how this interaction changes over time. Crosslinguistic investigation of vocabulary composition is critical to this endeavor. By observing children whose language experiences vary, we can map out both how the input shapes the lexicon and the ways in which children are resistant to variation.

For twenty years research in this area has centered on Gentner's noun dominance hypothesis (1982). Gentner compared parental reports of children's early vocabularies in a variety of languages. She found that nouns always made up the majority of the child's first words, even when the child was learning a language that placed the main verb in a salient location or allowed arguments to be omitted. She argued that nouns are over-represented, relative to their frequency in input, because they refer to categories that are conceptually and perceptually simpler. This analysis is supported by studies showing that verb meanings are more variable, within and across languages, and less resilient to interference in a variety of cognitive tasks (Talmy, 1975; Maratsos, 1990; Gentner, 1981; Gentner & France, 1988).

This claim has been challenged by recent work, showing that children learning Mandarin (Tardif, 1996), Korean (Choi & Gopnik, 1995), and Tzeltal (Brown, 1999) have early vocabularies in which the number of verbs equals or surpasses the number of nouns. These authors attribute crosslinguistic differences in vocabulary composition to variation in the structural properties of languages and correlated differences in the frequency and salience of nouns and verbs.

Our goal is not to judge whether language plays a role in shaping vocabulary composition. Nor is it to determine whether person and object labels are more easily learned regardless of language. These two positions are both well-evidenced and compatible: Studies which have examined the speech of children and adults have generally found both an effect of language and a difference between the composition of the input and the child's vocabulary (Tardiff, Shatz & Naigles, 1997; Tardif, Gelman & Xu, 1999; for a review see Gentner & Boroditsky, in press). Our goal instead is to explore one facet of the input that might favor or hinder words from different classes.

Previous research on crosslinguistic differences in vocabulary composition has focused on two ways in which the input can shape children's lexicons. First, many researchers have examined the frequency of different types of words in the input, making the plausible assumption that the number of learning opportunities should effect the likelihood that the child acquires a word. Mandarin, Korean and Tzeltal, the languages argued to show early verb dominance, all allow subject and object omission. Consequently, casual speech in these languages includes many utterances that contain a main verb but no lexical nouns. Children learning a language of this kind are exposed to more verb tokens and fewer noun tokens than children learning English (Tardif et al., 1997; Choi, 2000).

The second type of explanation places the explanatory weight on the perceptual salience of nouns and verbs in connected speech. Presumably a child can only learn a word if she can segment the word-form from the ongoing speech stream. Previous work suggests that this task is easier when the word appears at the beginning or end of the utterance (Fernald, McRoberts & Herrera, 1992; Newport, Gleitman & Gleitman, 1977). Tardiff and colleagues (1997) analyzed child-directed speech in Mandarin, Italian, and English to explore whether differences in the typical position of nouns and verbs in each language might account for the observed variation in vocabulary composition. They found that in Italian and English, languages where nouns dominated the early lexicon, verbs were generally buried away in the center of the utterance while nouns grabbed the salient utterance final position. In contrast, Mandarin speaking caregivers were more

likely to produce sentences that ended with a verb (for parallel findings in Korean see Au, Dapretto & Song, 1994; Choi, 2000).

This work has been useful in understanding crosslinguistic variation in vocabulary composition. However, by focusing on the frequency and distribution of word classes, researchers have overlooked another way in which language could shape word learning. Learning a word does not consist of merely of hearing the word and isolating it from speech. To learn a word you also must pair the word-form with its meaning. Factors that influence the child's ability to perform this mapping, should also affect vocabulary composition and are therefore a potential source of crosslinguistic variation.

Initially children must identify the meanings of words from the real-world contexts in which they occur. A novice language learner, who knows few words and little syntax, has no other information source to draw upon. We would expect then, that children's early vocabularies would be limited to words whose meanings can be identified from the situational concomitants of their use. To the extent that languages vary in their stock of common nouns and verbs or in the pragmatic and situational cues that are available for word learning, early vocabularies should also vary.

The Human Simulations

Gleitman and colleagues have conducted a series of experiments which illustrate the potential importance of the mapping problem for explanations of vocabulary composition (see Snedeker, Brent, and Gleitman, 1999; Gillette, Gleitman, Gleitman, and Lederer, 1999; and Snedeker, 2000). They find that developmental changes in vocabulary composition within a single language (English) can be explained by changes in the child's ability to use linguistic context to solve the mapping problem. In these studies adult subjects try to identify words from partial information about the contexts in which they occur in infant-directed speech. In some conditions the adults are given only the extralinguistic context, information that is available to novice language learners. In other cases they are provided with linguistic information, like syntactic context or co-occurrence, which would only be available to savvy learners.

Conceptually, these experiments are analogous to computer simulations in which a device, endowed with whatever ("innate") ideas and learning procedures its makers deem desirable to program into it, is exposed to data of the kind naturally received by the target learner it is simulating. The measure of success of the simulation is how faithfully it reproduces the learning function for that target using these authentic data. These experiments provide us with an estimate of the psychological potency of various cues to word meaning that are available in the real learning situation.

The results of Snedeker, Gleitman and Brent's study of extralinguistic context are particularly relevant to the current debate about crosslinguistic vocabulary composition (1999, henceforth SGB). Because this study will be used both as a model and comparison for the work that follows, we will describe it in some detail.

The stimuli for these experiments were generated by Gillette et al. (1999) who videotaped mothers interacting with their 18 to 24-month old children in an unstructured situation. The maternal speech was transcribed to find the 24 most frequent nouns and the 24 most frequent verbs that these mothers uttered during these taped sessions. Table 1 lists each target and its frequency in the original videotapes. To simulate a condition

under which learners were presumed able only to identify recurrences of the same word in the speech stream and to match these with their extralinguistic contexts of use, Gillette and colleagues selected 6 video clips during which the mother was uttering each of these words. Each video clip started about 30 seconds before the mother uttered the word, and ended about 10 seconds afterwards. The audio was removed from each of the clips, to remove the linguistic context and a beep was inserted at the very instant during the depicted event when the mother had actually uttered the mystery word

Table 1: The target words used in Snedeker, Gleitman and Brent (1999).

Noun	Frequency	% Correct
plane	30	78.6
drum	19	71.4
elephant	33	71.4
pig	119	71.4
swing	7	67.9
ball	53	67.9
bag	24	46.4
hammer	8	25.0
hole	16	25.0
nose	17	17.9
tail	9	14.3
hand	10	14.3
music	12	10.7
people	18	10.7
peg	7	7.1
toy	21	7.1
pilot	7	3.6
kiss	22	3.6
hat	36	3.6
camera	7	0.0
shoes	7	0.0
things	8	0.0
daddy	14	0.0
mommy	16	0.0

Verb	Frequency	% Correct
throw	24	53.6
come	65	46.4
push	28	42.9
play	32	28.6
hammer	30	21.4
look	51	21.4
wait	13	17.9
see	54	17.9
put	69	10.7
love	11	7.1
fell	11	3.6
stand	17	3.6
like	20	3.6
go	87	3.6
know	13	0.0
make	13	0.0
think	16	0.0
pop	23	0.0
say	27	0.0
have	30	0.0
turn	35	0.0
get	46	0.0
want	59	0.0
do	74	0.0

Before the experiment began, subjects were told that half the words were nouns and half were verbs to ensure that they had accurate expectations about the type of words that would be presented. Subjects viewed the videos and wrote down their guess after viewing each of the 6 clips for a word; that is, as cross-situational evidence accumulated. Then they were asked to think back over the full set of 6, and to make a final conjecture based on all the evidence. Thus for each subject there were 7 sequenced conjectures. This procedure continued for all the test words.

The results of this procedure have been both dramatic and consistent. When subjects were limited to the information of their senses, that is, when they had to acquire the mappings of sounds (here, beeps) onto meanings solely via inspection of the

extralinguistic contingencies for their use, they were able to identify 26% of the nouns but only 11% of the verbs, a difference which was both reliable and replicable (see Experiments 2 and 3, Snedeker, Brent & Gleitman, 2000). For many of the verbs the information provided by the scenes is not only inadequate, but actually misleading. As a group, the subjects settled upon the same incorrect hypotheses for many of the target verbs. Like infants learning their first words, our adult subjects "acquired" a vocabulary dominated by nouns, and under-populated by verbs.

Why should cross-situational observation have been so much more useful for noun identification than it was for verb identification? Gleitman and colleagues found that the noun dominance effect is itself an artifact of another property which was closely correlated with the noun/verb distinction in these stimulus materials (Gillette et al., 1999; SGB). Mothers frequently say words like *think* and *like* to their toddlers but they rarely say *thought* or *liking*. SGB found a strong correlation between how often a target was correctly identified and its concreteness rating in a separate study (see Table 1). Abstract nouns like *music* were less often identified than object labels like *drum*; and mental state verbs like *want* were identified less often than action verbs like *throw*. What these subjects' responses seem to be telling us is that, because they were limited to extralinguistic observation, they could only identify words whose real-world concomitants were stable and observable. The verb repertoire in English infant-directed speech is systematically less concrete than the noun repertoire, consequently, it is less amenable to acquisition by word-to-world pairing.

These studies suggest that the limitations of word-to-world mapping strongly shape the content of early vocabularies. If words of a particular class can be easily identified from observation alone, then children will tend to learn these words quite early. If words of a given category cannot be mapped to their meanings without aid of linguistic information, then these words should be relatively rare in the novice lexicon. Given cross-linguistic differences in categorial composition, we might therefore expect to find underlying differences in the difficulty of the mapping problem. The differences between the initial vocabularies of Mandarin learners and English learners could simply reflect the relative difficulty of learning nouns and verbs from observation in each of these linguistic environments.

We set out to explore this by conducting an experiment in Mandarin that directly paralleled SGB's study of English. The target nouns and verbs were selected from videotapes of Mandarin speaking mothers playing with their children. Mandarin speaking adults were asked to identify these words from the situations in which they were used. If vocabulary composition is strongly shaped by the information available for word-to-world mapping, then we would expect that the noun advantage should be smaller in this study than it was in the English study. This could occur either because the most common verbs in Mandarin are more easily mapped to their meaning than common verbs in English or because the most common nouns in Mandarin are more difficult to map to their meanings than common nouns in English.

Methods

Subjects

Subjects consisted of 16 students from the University of Pennsylvania and 26 students from the Su-Te Institute of Technology and the Chia Yi National University in Taiwan.

All of the students were native Mandarin speakers, who had learned Mandarin in their home before age five and had been speaking it daily since that time. Each subject received either ten dollars or partial class credit in Introductory Psychology for their participation.

Stimuli

The stimuli were drawn from videotapes of mothers playing with their 18-24 month old children. The six tapes, each approximately one-hour long, recorded 2 boys and 4 girls with their respective mothers playing with a set of toys provided by the experimenter. The mother was asked to "play naturally" with her child, using the toys if and when convenient. The situations in which the children were taped, the toys that were provided and the instructions that were given to the mother were similar to those used in the SGB experiment.

The procedure for selecting the target words was identical to that used by SGB. The videotapes were transcribed and the 24 most common nouns and verbs used by the mothers were chosen as targets. In choosing nouns and verbs, we followed the following criteria: 1) Different forms of a noun, e.g. mao4 (hat,) mao4 zi (little hat,) mao4 mao4 (hat hat), were considered equivalent. 2) Auxiliary verbs were not counted as verbs. 3) We counted a verb as a single type across all of the contexts in which it appeared, e.g. da3 kai1 (open,) kai1 men2 (open door,) kai1 che1 (drive a car) were all counted as kai1 (open.) Refer to Table 2 for a complete list of target nouns and verbs and their frequencies in the transcripts. We divided both the noun and verb targets into 4 frequency groups: high, medium-high, medium-low, and low. We randomly assigned two words from each frequency group to each list. One word from each category and each frequency group appeared in the first half of the list while the other word appeared in the second half.

For each target word, six instances in which the word was used were selected for inclusion in the study. Uses of a target word were excluded if the referent of the word was visible to the child but off camera, or if the mothers lips were visible and might provide information about the word form. When there were more than six instances of the word that met these criteria, instances were selected at random. For each instance, a video clip was constructed that began 30 seconds before the target word was used and ended 10 seconds after it was said. In many cases, the mother said the target word at another time during this 40-second period. In these cases, the clip was expanded to include 30 seconds before the first use and 10 seconds after the last. Each of the uses of the word in a single clip counted as one of the six stimuli. The audio was removed from each video clip and a beep was inserted exactly where the word had occurred.

Procedure

Subjects were tested in small groups of one to three**.** Subjects were told the truth: that we were studying how well they could identify a word based on contextual observation. They were instructed to write down their best guess of the word the mother was saying each time they heard a beep. They were told that the target words were approximately half nouns and half verbs. Then, they were shown the silent video of mothers playing with their child. They knew that the target word remained constant for all six beeps. After hearing six beeps corresponding to six maternal utterances of the same word, the subject was asked to reconsider all the input and to offer a final guess. We informed subjects that guesses could vary from instance to instance and that the final guess could

differ from the previous guesses, so they must not leave out any answers. This procedure was repeated for 16 different words, half of them nouns and half verbs. The written instructions and debriefing were in Mandarin.

Table 2: The target words for the Mandarin study.

Noun	English Meaning	Frequency	% Correct
qiu2	ball	100	78.6
dian4hua4	telephone	69	57.1
che1	car	46	42.9
mao4	hat	33	35.7
xiang4	elephant	82	35.7
gou3	dog	56	28.6
yan3jing1	eye	44	21.4
quan1quan1	circle	65	21.4
tou2	head	41	14.3
ba4ba	father	79	14.3
wan3ju4	toy	29	7.1
bei1	cup	36	7.1
ya1ya1	duck	73	7.1
shou3	hand	99	7.1
ming2zi4	name	19	0.0
bing3gan1	cookie	22	0.0
shui3	water	36	0.0
dong1xi1	thing	47	0.0
ren2	people	53	0.0
mei4mei	younger sister	89	0.0
jie3jie	older sister	98	0.0
wa1wa1	doll	114	0.0
fei1ji1	plane	147	0.0
ma1ma	mother	331	0.0

Verb	English Meaning	Frequency	% Correct
kan4	look	261	100.0
lai2	come	457	71.4
chui1	blow	76	64.3
tui1	push	80	57.1
che1	eat	77	50.0
fang4	place	179	50.0
qu4	go	282	42.9
zuan3	turn	66	35.7
na2	take	188	35.7
guan1	close	49	28.6
kai1	open	70	21.4
gei3	give	228	14.3
fei1	fly	49	7.1
zuo4	sit	81	7.1
wan2	play	125	7.1
he1	drink	58	0.0
jiao4	call	59	0.0
jiang3	say	80	0.0
hui4	will	95	0.0
chang4	sing	96	0.0
zou3	walk	99	0.0
yao4	want	387	0.0
shuo1	speak	391	0.0
you3	have	393	0.0

Coding

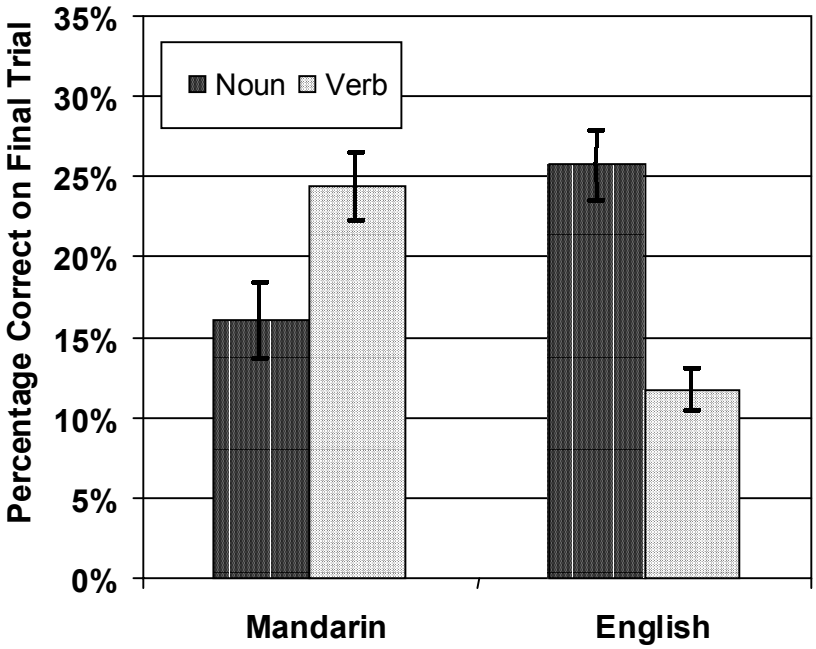
Responses were scored correct only if they contained the target word. If a response contained a phrase, only the head of the phrase was considered. For example, in the phrase "eat food," the answer could only be considered correct if the target word was "eat" not "food." Subjects who gave six or more final answers that were phrases or words that were not either nouns or verbs, were removed from the design and replaced. As were subjects who left more than one final answer blank.

Results and Discussion

Identifying the Target

As Figure 1 illustrates, on the final trial Mandarin-speaking subjects correctly identified more verbs than nouns (M = 24.4% and M= 16.1% respectively). This difference was evaluated by calculating the number correct on the final trial for both noun and verb targets and entering them into a subjects ANOVA with two between subjects factors (List, Order) and one within subjects factor (Syntactic Category). There was a reliable effect of Syntactic Category ($F = 8.522$, $p < 0.006$). There was also a significant List effect ($F = 11.426$, $p < 0.001$), indicating that some sets of targets were more difficult to identify than others. There were no other reliable main effects or interactions.

Figure 1: Correct Identification in Mandarin and English Experiments



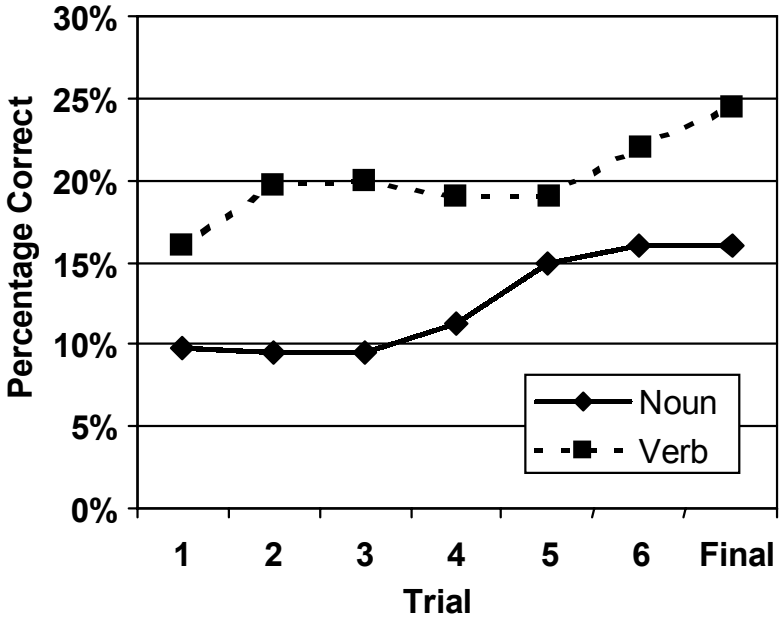
Learning across Trials

Figure 2 shows the percentage correct for each target type across trial. Performance on both nouns and verbs appeared to improve as subjects were shown more word-scene pairs. On the first trial, 9.82% nouns and 16.1% verbs were correctly identified. By the final trial 16.1% nouns and 24.4% verbs were correctly identified, an increase of approximately 50%. The reliability of this trend was evaluated by submitting the percentage correct on every trial to a subjects ANOVA with two between subjects factors (List, Order) and two within subjects factors (Trial, Syntactic Category). There was a significant effect of Trial ($F = 10.601$, $p < 0.001$); subjects performed better on the later trials than they did on the earlier ones. There were also significant effects for List ($F = 14.591$, $p < 0.001$) and Syntactic Category ($F = 17.139$, $p < 0.001$).

Effects of language: a Comparison of Word-to-World Mapping in Mandarin and English

The results of this experiment were compared with a parallel study conducted in English (Snedeker, Gleitman, & Brent, 1999). As Figure 1 demonstrates, there was a clear difference in noun and verb identification in the two studies. As noted above, verbs were identified more often than nouns in the Mandarin study. In contrast, nouns were identified more than twice as often as verbs in the English study (M = 25.7% and 11.8% respectively). The reliability of these differences was tested by conducting a subjects ANOVA with one between subject factor (Experiment) and one within subject factor (Syntactic Category). There was a significant interaction between Experiment and Syntactic Category (F = 26.61, p < 0.001). The main effect of experiment did not approach significance (F = 0.51, p < 0.475). Planned comparisons indicated that the subjects in the Mandarin study identified more verbs (p < .001) but fewer nouns (p < .001) than those in the English study.

Figure 2: Performance in Mandarin Experiment as a Function of Trial



Correcting for Differences in Response Bias

The subjects in both these studies were told that the targets were evenly divided between nouns and verbs. Despite this warning, they tended to make far more verb responses than noun responses. In the Mandarin study, 65.8% of the final responses were verbs while just 23.6% were nouns (the remaining 10.6% were from another syntactic category, e.g., proper nouns or pronouns or social words). In the English study subjects were somewhat less biased towards verbs; 56.3% of the final responses were verbs and 35.3% were nouns. The difference in bias between the two studies were reliable across subjects (t = 4.19, p < .001 for noun percentage; t = 3.32, p = .001 for verb percentage).

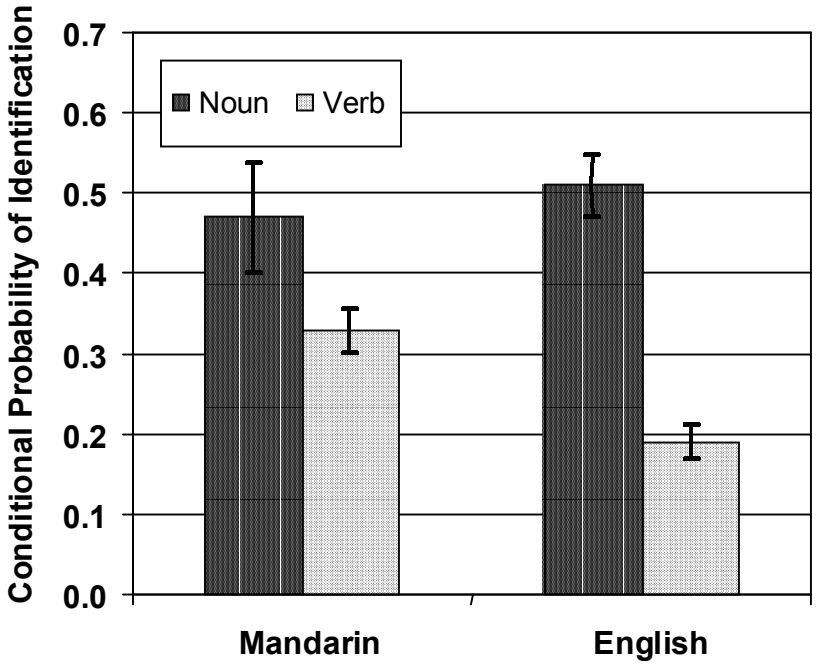
This difference in response bias could spring from two possible sources. First, it could be caused by differences in the two samples of maternal speech, which could

potentially shape word-to-world mapping in young children. For example, it is possible that the Mandarin videotapes were more action oriented and therefore led the subjects in that study to consider verb meanings more often than those in the English study. If this were the case, then the difference in response bias might be an accurate reflection of crosslinguistic differences in the extralinguistic input, rather than an experimental artifact.

However, we must also consider the possibility that these differences in response bias are attributable to differences between the subjects rather than differences in the input. Mandarin subjects may have been more likely to make verb responses because they realize, perhaps only implicitly, that verbs should be about twice as frequent as nouns in casual speech of this kind. While verbs also dominated the responses of the English speakers, this bias may have been held in check by the subjects' knowledge that nouns were equally frequent in their language. These biases, built up over a lifetime of language use, may have been more powerful than the explicit instructions which we provided.

Unfortunately, such a difference in response bias could account for the cross-linguistic pattern of identification performance found in these studies. If the subjects in the Mandarin study entered the experiment with the belief that the targets were verbs, then we would expect that they would tend to look for evidence suggesting that the target was a verb but ignore evidence that suggested it was a noun. Such a bias would lead the Mandarin subjects to better verb performance and poorer noun performance, even if the evidence provided by the scenes was equivalent in the two language samples.

Figure 3: Probability of Correctly Identifying Target Given that the Response is from the Correct Syntactic Category



To determine whether the strength of the bias difference was strong enough to account for the variation in performance, we focused in on those trials in which the

subject made a response of the correct syntactic category. Performance on this subset of the trials should be uninfluenced by a simple bias towards noun or verb responses. If subjects in the Mandarin study were performing more poorly on noun targets simply because of a verb bias, then we would expect that when they did make a noun response their rate of success would be just as high as that of the subjects in the English study. Similarly, if the smaller verb bias of the subjects in the English study is the only factor leading to lower verb performance, then the success rate for verb responses to verb targets should be identical in the two studies.

For each subject and for each type of target (noun or verb) we calculated the probability that the subject was correct given that she chose a word of the correct syntactic category. These probabilities were submitted to a subjects ANOVA with one between subjects factor (Experiment) and one within subjects factor (Syntactic Category). As Figure 3 suggests, there was a significant main effect of Syntactic Category, which was superseded by an interaction between Syntactic Category and Experiment. To further explore the source of this interaction we conducted planned comparisons of the two experiments for noun and verb targets. In both experiments, subjects who correctly guessed that a target was a noun identified the correct noun about half of the time ($M = .51$, $M = .47$ for English and Mandarin respectively, $F = .35$, $p > .5$). But performance on verb targets differed sharply. Subjects in the Mandarin study who correctly guessed that the target was a verb were able to identify almost twice as often as those in the English study ($M = .19$, $M = .33$ for English and Mandarin respectively, $F(1,124) = 15.84$, $p < .001$).

These results demonstrate that the possible difference in response bias could account for the differences in performance for noun targets. They by no means prove that this is the correct explanation of these differences. The bias correction that is described above assumes that the decision to make a noun or verb response is independent of the subjects ability to identify the correct response. To the degree to which this is false, the analysis overestimates the effect of bias on performance. Untangling response bias from differences in the input will require additional experimental work. However, the analysis of contingent probabilities demonstrates conclusively that the differences in performance for verb targets **cannot** be solely attributed to a difference in response bias. Even when we examine only those cases in which the subject correctly guessed that the target was a verb, we find that the subjects in the Mandarin study were more accurate than those in the English study.

Which Words can be Learned from Observation

Our comparison of the English and Mandarin studies uncovered crosslinguistic differences in the mapping cues that are available in the input. But they do not tell us what the sources of these differences are. One possible source of variation could be a difference in the target words that were chosen for analysis. The criterion for target selection was the same in both of the studies but the targets themselves were not matched in any way. Perhaps the uncovered differences are attributable to differences in the types of words that the mothers said. Not all nouns and verbs are created equal. The earlier studies of English input found that the noun-verb difference was actually an artifact of the concreteness of the word (Gillette, et al, 1999; Snedeker, et al., 1999).

Our analysis of the noun targets focused on identifying the targets that referred to categories of people or objects at the so-called basic level (BLOCs). It has been argued

that categorization at the basic level is perceptually and conceptually priviledged, in part because BLOCs typically share a large number of perceptual features (Rosch, 1988; Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976). SGB found that the noun advantage in the English word-to-world mapping study was solely attributable to the subjects success in identifying BLOCs. Performance on nouns of other types was no better than performance on verbs.

To compare the composition of the targets in the two experiments, we classified each one as a BLOC or a non-BLOC. The composition of the target set was nearly identical in the two studies. The English noun set was comprised of 12 BLOCs and 12 nonBLOCSs (4 partitives, 3 superordinates, 3 relational nouns, and 2 abstractions). The Mandarin noun set was made up of 11 BLOCs and 13 nonBLOCs (3 partitives, 3 superordinates, 4 relational nouns, 2 abstractions, and 1 mass noun). In both studies, performance on BLOCs was considerably higher than performance on nonBLOCs (M = 42.6% and M = 8.9% for English, and M = 26.6% and M = 6.6% for Mandarin). Thus the differences in mapping performance across the two studies clearly cannot be attributed to differences in the type of nouns that appeared in the two samples of maternal speech.

In contrast, the target verbs in the two studies differed considerably. Previous analyses of English human simulations have found that subjects typically have a reasonable amount of success identifying verbs that refer to observable actions. But they tend to perform abysmally when the target referred to a mental state. We compare the verbs in the two studies by classifying each one as either "observable" or "unobservable". All verbs referring to actions whose presence or absence could typically be seen or heard were classified as "observable". The "unobservable" verbs were further divided into mental state verbs and all others. Thirteen of the twenty-four English targets were classified as unobservable verb. Seven of these were mental state verbs. In contrast there were just 6 unobservable verbs in the Mandarin target set and only 2 mental verbs. Unsurprisingly, subjects in both studies performed much better when the target verb referred to an observable action (M = 19.5%, M = 5% for English observable and unobservable verbs respectively and M = 27% and M = 16.7% for Mandarin).

Critically, this difference in type of verbs that the mothers used appears to be typical of child-directed speech in each of these linguistic environments. Recently, Sandhofer, Smith, & Luo conducted a large-scale corpus analysis of speech to children learning either American English or Mandarin (in press). They found that in both languages a fairly small number of verbs accounted for a high percentage of all verb tokens. In each language this set of verbs overlaps considerably with the set that we identified from our admittedly smaller corpus. We selected the 24 most common verbs in English and in Mandarin from Sandhofer et al. sample and coded them as described above. The semantic composition of the verb sets in the two studies is almost identical. Their English set also contains many unobservable verbs and mental state verbs (13 and 7 respectively) while their Mandarin set is dominated by action verbs (6 unobservable and 3 mental state verbs). We conclude that child-directed speech in Mandarin, for reasons unknown, contains more verbs that refer to observable concrete actions. Since observable verbs are easier to identify from extralinguistic context, this pattern of verb use contributed to the high performance of the adults in the Mandarin study. Similarly, young children, who are limited to the evidence of their senses, should also profit by the frequent appearance of verbs that can be learned from extralinguistic context alone.

Summary

Our comparison of word-to-world mapping in Mandarin and English unearthed a pattern of performance across syntactic class that is similar to what we see early vocabularies of young children. Subjects in the English were able to identify more nouns and fewer verbs than those in the Mandarin study. Our analyses indicated that the difference in noun performance might be attributable to difference in the response bias of the two subject populations. The difference in verb performance, however could not be accounted for by bias alone. Instead it appeared to be due, in part at least, to a substantial difference in the types of verbs that the mothers in each group chose to use. A more fine-grained analysis of the targets, revealed similarities in the subjects' performance for targets of several semantic types. In both experiments subjects were more likely to identify a noun when it was a concrete object labels, and more likely to identify a verb when it referred to an observable action.

Our findings cannot be attributed to differences in the frequency or salience of the noun and verb targets. We solved the problem of segmentation for our subjects by presenting each word in isolation. We held the number of presentations of a word constant across language and syntactic class. Instead our results show that linguistic environments also vary in the potency of the cues that they provide for linking word forms to word meanings. These differences are consistent with observed differences in early vocabularies. The next challenge for research on vocabulary composition is to demonstrate that differences in frequency, perceptual salience, and/or situational cue strength actually play a causal role in shaping vocabulary composition

Acknowledgements

Carol Chan and Renee Lam helped us prepare the videotaped stimuli and tested the subjects from the University of Pennsylvania. We thank them for their invaluable assistance. This work was supported by an National Science Foundation Center Grant to the Institute for Research in Cognitive Science (SBR-89-20230), by an National Institute for Health Grant to John Trueswell and Lila Gleitman (1-R01-HD3750707-01), and a Predoctoral Fellowship from the Institute for Research in Cognitive Science to Peggy Li.

References

Au, Terry; Dapretto, Mirella & Song, You-Kyung (1994). Input versus constraints: Early word acquisition in Korean and English. *Journal of Memory and Language,* **33**, 567-582.

Brown, P. (1998). Children's first verbs in Tzeltal: Evidence for an early verb category. *Linguistics*, **36**, 713-753..

Choi, S. (2000) Caregiver input in English and Korean: Use of nouns and verbs in book-reading and toy-play contexts. *Journal of Child Language*, **27**, 69-96.

Choi, S. & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, **22**, 497-529.

Fernald, A., McRoberts, G. & Herrera, C. (1992). Prosodic features and early word recognition. Paper presented at the 8th International Conference on Infant Studies, Miami, FL.

Gentner, D. (1978) On relational meaning: The acquisition of verb meaning. *Child Development*, **49**, 988-998.

Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory*, **4**, 161-178.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S.A. Kuczaj, II (Ed.), *Language development: Vol 2. language, thought, and culture* (pp. 31-53). Hillsdale, NJ: Erlbaum.

Gentner, D., & Boroditsky, L. (in press). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development.* England: Cambridge University Press.

Gentner, D. & France, I. (1988). The verb mutability effect: Studies of the combinatorial semantics of nouns and verbs. In S. Small, G. Cottrell & M. Tannenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence* (pp. 343-382). San Mateo, CA: Kaufmann.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulation of vocabulary learning. *Cognition*, **73**, 135-151.

Maratsos, M. (1990). Are actions to verbs as objects are to nouns? On the differential semantic bases of form class category. *Linguistics*, **28**, 1351-1379.

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and noneffects of maternal speech stale. In C. Snow & C. Ferguson (Eds) *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.

Rosch, E. (1988). Principles of categorization. In A. Collins & E. Smith (Eds), *Readings in cognitive science: A perspective from psychology and artificial intelligence* (pp. 312-322), San Mateo, CA, USA: Morgan Kaufmann, Inc.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976) Basic objects in natural categories. *Cognitive Psychology*, **8**, 382-439.

Sandhoffer, C., Smith, L., & Luo, J. (in press). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? To appear in *Journal of Child Language.*

Snedeker, J. & Gleitman, L. (in press). Why it is so hard to label our concepts. To appear in Hall, G. & Waxman, S. (Eds), *Weaving a Lexicon*. Cambridge, MA: MIT Press.

Snedeker, J, Gleitman, L & Brent, M. (1999), The Successes and Failures of Word-to-World Mapping. In A. Greenhill, M. Hughs, H. Littlefield, & H. Walsh (Eds.), *Proceedings of the Twenty-third Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Talmy, L. (1975). Semantics and syntax of motion. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 181-238). New York: Academic Press.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, **32**, 492-504.

Tardif, T., Gelman, S. & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, **70**, 620-635.

Tardif, T., Shatz, M. & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, **24**, 535-565.