RUNNING HEAD: OBJECT AND SUBSTANCE IN TSOTSIL

Concepts of object and substance kinds: A comparison of speakers of English and of Tsotsil Mayan

Word Count: 10,182

Peggy Li[1], Frankie Chen[1], David Barner[2], Susan Carey[1]
[1] Harvard University
[2] University of California, San Diego

Please address correspondence to:
Peggy Li
Laboratory for Developmental Studies
Harvard University · 25 Francis Avenue · Cambridge, MA 02138
pegs@wjh.harvard.edu

Acknowledgements: We would like to thank Jesse Snedeker, Linda Abarbanell, David Bolles, Luis Cárcamo-Huechante, Gene Heyman, and Marc Zende for suggestions and help, and to Daniel Friel, Elisabeth Michaud, and Evan Favermann for help in collecting English data. We are indebted to the people in Chiapas, and our special thanks goes to Enrique Pérez López, Ana Catalina Pérez López, and Agustín Gutiérrez Ruíz. This research was supported by grants from the Alex G. Booth Fellowship and the Harvard Mind, Brain & Behavior Initiative to F.C.

Abstract

Three experiments tested the claim that grammatical differences between classifier languages and mass-count languages influence nonlinguistic representation of material entities. Participants were speakers of Tsotsil Mayan, a classifier language that contains numeral classifiers and optional, little used, plural markers. In keeping with previous Whorfian claims, Experiment 1 found that Tsotsil speakers were more likely to match entities by substance kind than by object kind relative to English speakers in word extension and similarity judgment tasks. However, using a quantity judgment task, Experiment 2 found that Tsotsil speakers did not differ from speakers of English in their quantification of material entities. Participants based judgments on number when the nouns were translations of English count or object-mass nouns (e.g., Who has more shoes? Who has more clothing?), and on mass for words that labeled non-solid substances in English (e.g., Who has more butter?). Finally, Experiment 3 asked Tsotsil speakers' to rate their perception of novel entities as objects or substances. Although Tsotsil speakers used a narrower range of numbers, they were not more likely than English speakers to rate entities as a kind of substance. These results support the hypothesis that nouns have similar meanings across these two languages, and that speaking a mass-count language does not lead to the features relevant to object individuation being weighted more heavily in representations of the material world, relative to the weightings of speakers of a classifier language. Rather, we suggest that the differences on novel word extension tasks arise from syntactic inferences that are mediated by the obligatory use of mass-count syntax in language comprehension, rather than by general differences in how entities in the world are construed.

Keywords: linguistic relativity, mass-count syntax, classifiers, Tsotsil language

Introduction

Languages vary in how they express thought, a fact that has led some researchers to propose that speakers of different language perceive and think about the world differently. The linguistic relativity of thought was most famously defended by Benjamin Lee Whorf, who argued that differences between grammars lead language users to make "different types of observations and different evaluations of externally similar acts of observation, and hence … arrive at some different view of the world" (p. 221). In contrast, others have defended a universalist view, whereby language does not fundamentally change how humans perceive the world, but instead offers a window into thought (e.g., Barner, Li & Snedeker, 2010; Pinker, 2007). Differences in how languages express thought, by this view, offer different perspectives into its nature, but do not fundamentally change how people to think about or perceive the world.

One domain that has received particular attention in the study of language and thought is the representation of discrete countable individuals. The linguistic representation of individuals differs substantially from one language to the next, causing significant differences in how people talk about objects and other countable things. In English, for example, an obligatory syntactic distinction is made between nouns that appear in mass and count syntax. Count nouns (e.g., *cup*) can be used with singular-plural marking (*a cup; some cups*), with quasi-cardinal quantifiers and determiners (*these/those/many/several cups*), and can occur directly with numerals (*five cups*). Mass nouns (e.g., *water*) can be used in none of these contexts, and require a unitizer or measure word to be used with numerals (e.g., *five cups of water*). Critically, languages differ with respect to whether they make this distinction between mass and count quantification, and many do not. Classifier languages like Chinese, Japanese, and Mayan languages, for example, do not have obligatory plural marking and do not allow nouns to occur directly with number words.

Somewhat like English mass nouns they require a classifier when a numeral is used, even if the noun already denotes a type of object (e.g., 'three oranges' in Tsotsil Mayan requires the numeral plus a classifier *p'ej* that is used to count round things -- *oxp'ej xas* or 'three CL-round orange').

According to a strong form of the linguistic relativity hypothesis, mastering this grammatical difference between mass and count syntax is instrumental to the development of concepts like 'object' and 'individual', such that languages with different quantificational structures express different construals of the world. Quine (1960), for example, argued that the first evidence that human infants perceive entities as objects is when they begin to use count syntax. Count nouns, but not mass nouns "possess built in modes, however arbitrary, of dividing their reference" (p. 91). Count nouns like *cup* delineate the boundary between one cup and another, and thus provide criteria of individuation. Mass nouns like *water* provide no such boundaries. Thus, Quine predicted that, prior to learning count syntax, children should not make an object-substance distinction. His hypothesis, which was also independently articulated by Whorf (1956), predicted that children who have learned mass-count languages like English should arrive at different construals of the world than those who have learned classifiers languages like Tsotsil or Japanese.

Evidence against this strong version of the linguistic relativity hypothesis comes from studies of children who are in the first stages of language acquisition. Soja, Carey, and Spelke (1991) addressed Quine's hypothesis by testing whether 2-year-olds learning English make an object-substance distinction prior to learning count syntax. In their study, Soja et al. taught children labels for both solid entities (e.g., a honey dipper) and non-solid entities (e.g., hand gel) that had particular shapes. After hearing a label for a referent that lacked mass-count syntax – e.g., "Look at the dax!", children were asked to extend the word to one of two additional entities – one that

matched the original in shape, but not material (i.e., object-kind match), and one that matched in material, but not shape (i.e., substance-kind match). Contrary to the prediction of Quine, children extended labels for solid objects according to shape, but extended labels for non-solid referents according to their substance. Thus, despite failing to produce count syntax in their productive speech, English-speaking 2-year-olds made a clear distinction between objects and substances.

While Quine's and Whorf's strong version of the linguistic relativity hypothesis is ruled out by Soja et al.'s developmental data, similar studies comparing speakers of classifier languages (Japanese, Mandarin, and Korean) with mass-count languages (English, Spanish, and Czech) raise the possibility of weaker effects of language on object representation (Imai & Gentner, 1997; Gathercole & Min, 1997; Crane, 2003; Li, Dunham, & Carey, 2009; Lucy, 1992). These studies raise the possibility that acquiring count syntax may make speakers more likely to perceive entities as objects. For example, Imai and Gentner (1997) who tested Japanese 2-year-olds using the methods of Soja et al. Although Imai and Gentner replicated Soja et al.'s finding that young children make an object-substance distinction, they also found that Japanese-learning children were significantly less likely than English-learning children to extend novel words according to shape. This finding suggests that English-speaking 2-year-olds may in fact have some subtle knowledge of mass-count syntax (see also Soja, 1992), and also that this knowledge may make them more likely to perceive entities as discrete physical objects, relative to their Japanese peers. Because speakers of a mass-count language must choose, for every noun, whether to explicitly mark it as referring to an individuated, countable entity, features that are correlated with count syntax (e.g., solidity, complexity of shape, functionality) might become more salient, and thus become more strongly weighted over time. According to this *feature weighting* account, learning count syntax shifts the ontological boundary between objects and

substances, leading English speakers to attend more strongly to the status of things as discrete individuals relative to speakers of Japanese (see also Gentner & Boroditsky, 2003; Imai & Mazuka, 2007; Yoshida & Smith, 2003).

Imai and Gentner's analysis built on the earlier work of John Lucy (1992), who investigated Yucatec Mayan, a classifier language spoken in the Yucatan peninsula in Mexico. Lucy was the first to demonstrate that speakers of a classifier language are more likely than speakers of English to choose material over shape when making similarity judgments. Indeed the magnitude of the effect Lucy found was more than twice as great as that found by Imai and Gentner and in subsequent studies. However, his explanation of the greater reliance on material matches by Mayan speakers than by English speakers differed from that of Imai and Gentner's. Instead of arguing that acquiring mass-count syntax shifts the ontological boundary between objects and substances, he argued that Yucatec Mayan and English differ in their lexical semantics. He reasoned that since Yucatec Mayan is a classifier language, its nouns do not individuate. Instead, "Yucatec nouns, lacking such a specification of unit, simply refer to the substance or material composition of an object" (p. 89). In support of this, Lucy provided examples of how Yucatec nouns combine with classifiers to generate a series of distinct meanings, suggesting that nouns have highly unspecified meanings. For example, he argued that the Yucatec word *kib'* (typically glossed as *candle*) does not actually mean 'candle,' but rather, a meaning more like 'wax'. The claim is that unlike the English noun *candle*, *kib'* does not specify a unit of individuation as part of its lexical semantics. Similarly, Lucy argues that the Yucatec word *há'as*, typically glossed as *banana*, does not denote a kind of individual. To illustrate this, he provided the following list of contexts in which the word can be used, each containing a different classifier that radically changes the way *há'as* is interpreted (p. 74):

’un-¢’íit	há’as	‘one/a 1-dimens. banana (i.e., the fruit)’
’un-wáal	há’as	‘one/a 2-dimens. banana (i.e., the leaf)’
’un-p’éel	há’as	‘one/a 3-dimens. banana (e.g., the fruit)’
’un-kúul	há’as	‘one/a planted banana (i.e., the tree)’
’un- kúuch	há’as	‘one/a load banana (i.e., the bunch)’
’um-p’íit	há’as	‘a-little-bit-of/some banana’

Based on this analysis, Lucy hypothesized that Yucatec routinely draws attention to the unformed material or essence that its nouns denote. Hence, he predicted that English speakers should weight shape or object kind more heavily in their similarity judgments than would Yucatec Mayans. To test this, he used a triad task like that of Soja et al. (1991), but which differed from their paradigm in two ways. First, he elicited similarity judgments rather than novel noun projections. Second, instead of using novel objects, he asked participants to make judgments for familiar entities. Each triad included an original standard item—such as a cardboard box—and two alternatives: one that had the same shape as the standard—such as a plastic box—and one that had the same material as the standard—such as a piece of cardboard. Participants were asked to select which of the two alternatives was more similar to the standard. Consistent with his hypothesis, he found that Yucatec speakers were more likely to choose the substance match than were English speakers. Later studies extended these results to a wider range of items that controlled for the functionality of the standard and choices, and pinpointed 9 years old as the age at which the effect of language emerges on this task (Lucy & Gaskins, 2001, 2003).

Recently, several studies have questioned whether the cross-linguistic differences reported by Lucy, Imai and Gentner, and others should be taken as evidence for linguistic relativity. First, several studies have questioned Lucy’s claim that nouns in classifier languages have systematically different meanings relative to nouns in mass-count languages and that count syntax is required for individuation, even in mass-count languages like English. These studies

show that although count syntax does specify individuation, mass nouns are also capable of denoting individuals. For example, in one study Barner and Snedeker (2005) asked English speakers (both 4-year-old children and adults) to make quantity judgments for entities described by familiar mass and count nouns (Who has more peanut butter? Who has more shoes?). Participants of both ages judged that six small shoes were *more shoes* than two very large ones, but that two very large portions of peanut butter were *more peanut butter* than six tiny portions. Also, subjects made judgments based on number when flexible words like *string* were used in count syntax (Who has more strings?), but on mass when the same words were used in mass syntax (Who has more string?). These results are consistent with the idea that mass and count syntax are related to different modes of quantification. However, in a final condition, participants based quantity judgments for object-denoting mass nouns like *jewelry, furniture,* and *clothing*, on number, suggesting that, like count nouns, these words quantify on the basis of individuals, in spite of being mass nouns. Thus, words used in mass syntax can denote individuals, leaving open the possibility that the nouns in classifier languages (which behave syntactically like mass nouns) may denote individuals also.

In support of this idea, Barner, Inagaki, and Li (2009) found that speakers of Japanese and English do not interpret common nouns differently, contrary to the idea that count syntax makes words more likely to individuate. Using the quantity judgment task described above, Barner et al. tested nouns that for English speakers were either count nouns, mass nouns, or mass-count flexible nouns (words like “string” or “rock” which individuate in count syntax or quantify as substances in mass syntax). Overall, Japanese speakers were no less likely to base quantity judgments on number than English speakers, even though words were presented with explicit mass-count syntax to English-speaking subjects. For words used as count nouns in English, and

for words that are object-mass nouns in English, Japanese speakers based quantity judgments systematically on number. For substance mass nouns like *ketchup*, they based judgments on mass or volume. For flexible words, Japanese judgments fell exactly between the mass and count judgments of English speakers for these same words, as would be expected if both groups found these nouns ambiguous between object and substance interpretations (e.g., a rock vs. some rock). Barner et al. concluded that nouns in Japanese and English have the same meanings and the same quantificational implications, in spite of the fact that Japanese lacks mass-count syntax. Echoing this, Cheung, Barner, and Li (2010) found parallel results for another classifier language, Mandarin Chinese (see Sandhofer, Smith, & Luo, 2000, Colunga & Smith, 2005, for additional evidence that the nouns of Mandarin and Japanese have the same meanings as those of English). However, left open by these studies is the possibility that these Asian classifier languages differ from Mayan languages, and that nouns in Mayan languages do not provide criteria for individuation and denote unindividuated essences instead. Experiments 1 and 2 explore this possibility.

A second issue raised by Lucy's findings is that, besides differing linguistically, speakers of English and Yucatec Mayan differ in countless other ways. For example, in a study using Lucy's (1992) triad similarity judgment task, Mazuka and Friedman (2000) failed to find any difference between Japanese and English speakers. Both groups judged a plastic box to be more similar to a cardboard box than to a piece of cardboard, and did so to the same extent. Thus, some have argued that Lucy's findings are attributable to sociocultural factors, like the fact that the Mayan participants were subsistence farmers whereas the English and Japanese participants were college students immersed in a world of manufactured products and artifacts. Consequently, Mayans may appreciate and place greater emphasis on the materials from which agricultural

food products or artifacts are created relative to Americans or Japanese. Take for example Lucy's triad "grains of corn, beans, and tortilla". Although tortilla is derived from corn, the relation between grains of corn and tortillas might only be transparent to Mayans for which corn is a staple. Relatedly, Lucy used familiar and nameable items in almost all of his experiments. As a result, participants could use naming conventions of their community as the basis resolving sameness in a triad task. Of the stimuli tested, the same-shape choices typically shared the same lexical label in Japanese and English whereas the same-material choices were more likely to share the same lexical label in Yucatec. For example, for the triad "sheet of paper, sheet of plastic, and book", paper and book share the same morphological roots in Yucatec. Thus, Yucatec Mayans might more often select the material choice relative to Japanese or Americans on the basis of naming conventions.

One way to circumvent pre-existing sociocultural norms of categorization and naming is to test participants on novel stimuli, as did Imai and Gentner (1997). Based on this idea, Experiment 1 tested Tsotsil Mayan speakers using the triad task previously used in studies that compared Japanese, Mandarin, English, and Czech (see Imai & Gentner, 1997; Crane, 2003; Li, Dunham, & Carey, 2009). In these triad tasks, small but replicable differences are repeatedly found between classifier languages, on the one hand, and mass-count language, on the other.

Finally, although some of the cross-linguistic effects reported in previous studies are robust and easily replicable across different languages (Imai & Gentner, 1997; Li et al., 2009; Lucy, 1992; Gathercole & Min, 1997), there is disagreement about how to interpret such effects and whether they reflect true differences in non-linguistic representation. For example, Li, Dunham, and Carey report that cross-linguistic differences disappear when participants are simply asked to rate whether they find it more natural to construe a given entity as a kind of object (a table) or a

kind of substance (some wood). In their study, Li et al. (2009) asked Japanese, Mandarin, and English speakers to rate a series of novel entities which varied in solidity, complexity of shape, material, and the extent to which they appeared to have a function. If language truly shifts the categorical boundaries between object and substance kinds then Mandarin and Japanese speakers should be less likely to view novel entities as instances of object kinds than would English speakers. What Li et al. found, however, was that Mandarin, Japanese, and English speakers made identical judgments on the task.

Based on these findings, Li et al. argued that previously reported differences on word extension tasks do not reflect cross-linguistic difference in ontological commitments, as argued by Whorf or Lucy, or in ontological boundaries, as argued by the feature weighting account. Instead, they reflect non-Whorfian language-on-language effects, and in particular the use of language-specific *lexical statistics* when interpreting sentences (see also Gleitman & Papafragou, 2005; Fisher & Gleitman, 2002; Barner, Inagaki, & Li, 2009). According to this view, the cross-linguistic differences in word extension can be explained by the fact that English requires its speakers to assign mass or count syntax to its noun phrases, and that the frequency of count nouns is greater than that of mass nouns in English. As Gleitman and Papafragou (2005) put it, "any English speaker equipped with even a rough subjective probability counter should take into account the massive preponderance of count nouns over mass nouns in English and conclude that a new word, *blicket,* used to refer to some indeterminate display, is probably a new count noun rather than a new mass noun. Count nouns, in turn, tend to denote individuals rather than stuff and so have shape predictivity." (p. 644)

Consistent with this analysis, Barner et al. (2009) found that participants who speak both English and Mandarin Chinese make different judgments in a word extension task depending on

which language they are tested in. When tested in English with mass-count ambiguous syntax (e.g., Look at the blicket!), participants showed a shape bias relative to when they were tested in Mandarin, which lacks a mass-count distinction. This result suggests that effects were due to the use of language online, rather than on underlying differences in how speakers think about or perceive objects.

Although this analysis can explain reported differences between English and Asian classifier languages like Japanese and Mandarin, it is currently unknown how speakers of Mayan languages perform on equivalent tasks, and thus whether Mayan languages are qualitatively different from other classifier languages. The present study addressed this issue by using the same methods and materials that have been used in the Mandarin/Japanese/English comparisons with speakers of a Mayan language—Tsotsil. Tsotsil is part of a different branch of Mayan than is Yucatec Mayan, so we first establish that it resembles Yucatec Mayan in every respect relevant to exploring the effects of the contrast between classifier and mass-count status on thought.

Like Yucatec Mayan, Tsotsil is an indigenous Mexican language. Tsotsil is spoken in Chenalhó, in the Highlands of Chiapas, Mexico, about 25 kilometers north of San Cristóbal de las Casas. Also, it features a rich system of numeral classifiers much like that of Yucatec Mayan. In Appendix A, we describe the extensive work of linguists and cultural anthropologists who have studied Tsotsil, including their characterization of the Tsotsil system of numeral classifiers. This work makes clear that Tsotsil Mayan, like its cousin, Yucatec, is a classifier language that lacks a mass-count distinction. As in Japanese and Mandarin, Tsotsil typically requires the use of classifiers for counting, lacks a distinction between mass and count quantifiers, and lacks obligatory plural marking (despite making an optional plural marker available). In sum, linguistic

and cultural similarities between Tsotsil and Yucatec make Tsotsil a relevant test case to examine noun meanings of Mayan languages, and whether these languages differ qualitatively from other classifier languages, like Japanese and Mandarin.

Experiment 1: Word Extension and Similarity Judgment Tasks in Tsotsil

Experiment 1 tested speakers of Tsotsil Mayan using the triad task that has previously been used to evaluate the linguistic relativity hypothesis in English, Japanese, and Mandarin Chinese (Imai & Gentner, 1997; Li et al., 2009). Experiment 1 sought to establish whether Tsotsil speakers, like those of Asian classifier languages, are more likely to select substance-matched entities than English speakers.

In the present study, we presented participants with stimuli that were relatively unfamiliar and unlikely to have conventional labels in the languages tested. Previous studies that tested subjects with novel stimuli using a triad task have found that both Japanese and Mandarin speakers choose the material match about 20% more often than English speakers. By comparison, using familiar and nameable stimuli, Lucy found that Yucatec Mayans chose the material match about 40% more than English speakers (thus showing twice the difference that is found in comparisons of English and other classifier languages using novel entities). If Mayan languages have a larger effect on entity construal than Asian languages (e.g., due to different lexical structures), or if cultural differences between subsistence farmers and college students make substances more salient than artifact kinds, then Tsotsil speakers should be more likely to choose the substance match than are Japanese and Mandarin speakers. However, if Mayan and Asian classifier languages are similar, then Tsotsil speakers should pattern like Japanese / Mandarin speakers and differ from English speakers to the same degree.

Lucy's triad task was a similarity judgment task, whereas Imai and Gentner, like Soja et al.,

used a word extension task. However, in a subsequent study, Imai and Mazuka (2003) tested how Japanese and English speakers judged similarity among the items used in the word extension tasks and found no differences between the two methods. That is, Japanese speakers differed from English speakers in their judgments as to which of two stimuli was more similar to a standard to the same extent that they were likely to differ on which entity is likely to be named by the same word. Li et al. (2009) replicated this pattern of findings in a Japanese / English comparison, and extended it to a Mandarin / English comparison. Experiment 1 explored whether similarity judgments, as well as word extension, differ between Tsotsil speakers and English speakers. In particular, we asked whether these materials generate effects with a magnitude comparable to that found by Lucy in his studies, or comparable to those found between speakers of Asian classifier languages and English.

Previous studies also suggest that the physical properties of the entities affect how they are construed across all languages. Japanese, Mandarin, and English speakers are more likely to choose on the basis of shape when the initial entity is a solid object than when it is a non-solid. They are also more likely to choose on the basis of shape when the initial entity appears to have a shape-dependent function. The present study tested whether Tsotsil Mayans are also affected by these factors in their construal of entities. Although Lucy's study with familiar entities provided some indication that solidity and functionality modulated Yucatec Mayan's choices of shape or material (see Lucy & Gaskins, 2001; 2003), no previous study has tested Mayan speakers with novel stimuli, to compare with speakers of other languages. Experiment 1 tested speakers of Tsotsil Mayan using stimuli previously used in Mandarin, Japanese, and English, to directly compare their relative sensitivity to these conceptual factors.

Method

Participants. Participants were 45 native monolingual Tsotsil-speaking adults (mean age=38, SD = 16, ranging from 17 to 80) recruited in Chenalhó, Chiapas. Although some people in Chenalhó interact with the world outside the town and some individuals speak some Spanish, the present study only included participants identified as monolingual Tsotsil speakers who spoke very little to no Spanish. Because men are more likely to have contact with Spanish-speaking individuals outside the community, the participants in this study were disproportionately female; across the experiments 14% of participants were male and the remaining 86% were female. However, it should be noted that prior research using these methodologies has given no indication that gender is a significant factor (e.g., Imai & Gentner, 1997; Li, Dunham, & Carey, 2009; Barner & Snedeker, 2006). Participants received 20 pesos for their participation. They were tested individually in a house in the center of town. The participants were randomly assigned to one of two conditions: a Label condition (n = 25) or a No-Label condition (n = 20).

Stimuli. The stimuli used in this study were borrowed from Li et al. (2009). The stimuli consisted of three entity types (Complex Solids, Simple Solids, and Non-Solids), normed to be comparable to Imai and Gentner's (1997) stimuli in terms of both complexity and perceived functionality. Norms were obtained from the ratings of ten naïve native English speakers on a 7-point Likert scale, as reported by Li et al. (2009). Complex Solids were rated as more complex in shape than either the Simple Solids or the Non-Solid Substances, and that the latter two classes did not differ from each other in rated shape complexity (Complex: 4.31 > Simple: 2.57 = Non-solids: 2.78). Also, the Complex Solids were rated as having significantly more function-relevant shape than the Simple Solids, which were in turn rated as having significantly more function-relevant shape than the Non-Solids (Complex: 5.53 > Simple: 3.45 > Non-solid: 2.23). There

were four test triads for each entity type, with each triad consisting of the standard, a shape-matched alternative, and a material-matched alternative. See Table 1 for a complete list of stimuli.

TABLE 1. Stimuli for Experiment 1.

	Standard	Shape-Match	Material-Match
Complex Solids	Clear plastic clip	Metal clip	Clear plastic piece
	Copper T	Ivory plastic T	Copper piece
	Wooden reamer	Rubber reamer	Wooden pieces
	Metal whisk	Nylon whisk	Metal piece
Simple Solids	Cork stopper	Plaster stopper	Cork piece
	Terracotta wedge	Marble wedge	Terracotta pieces
	Red Sculpey ½ egg	Styrofoam ½ egg	Sculpey pieces
	Pink wax kidney	Yellow clay kidney	Pink wax piece
Non-Solids	Nivea spiral	Sparkle gel spiral	Nivea piles
	Orange sand trapezoid	Blue crystal trapezoid	Orange sand piles
	Shaving-foam gamma	Brown sugar gamma	Shaving-foam pile
	Sawdust omega	Flour omega	Sawdust pile

Procedure. Participants were tested individually. The experimenter introduced the participants in both the Label and the No-Label conditions to the standard entity with a verbal cue, “Quelo li’i” (“Look at this”), accompanied by pointing. Participants in the Label condition were then taught a novel word for the entity: “Li’i ja sbi blicket” (“This is called blicket”). Participants in the No-Label condition did not receive this additional instruction. Following the presentation of the standard entity, all participants were presented with the shape- and material-match entities, side-by-side. In the Label condition, they were asked, “¿Cusi sbi blicket ek?” (“Which is also called blicket?”). In the No-Label condition, they were asked, “¿Cusi ja chi’il?” (“Which is the same?). All participants indicated their responses by pointing. There were two item orders, one the reverse of the other. The left and right positions of the shape-match and the material-match were pseudo-randomized and counterbalanced so that for half of the trials the

shape choice was on the right side.

Practice Trials. Participants were presented with two unambiguous practice triads before the twelve test triads. In one practice triad, a metal fork was used as the standard and the two comparison entities were a plastic fork and a plastic spoon. Thus, in this practice triad, there was no material-match available and the only logical choice was the shape-match. In the other practice triad, a rectangular portion of jam was used as the standard and the two comparison entities were two round portions of jam and a triangular portion of Play-Doh. In this practice triad, there was no shape-match available and the only logical choice was the material-match. All participants chose as expected (extending by shape in the first case and substance in the second), suggesting that they understood the task.

[Figure 1 about here]

Results

The data patterned as in previous studies with Japanese and Mandarin speakers. Entity type affected participants' choice of the shape over material whereas the label/no label manipulation did not (see Figure 1). These results were confirmed with a 3 Entity Type (Complex Solid, Simple Solid, and Non-solid) x 2 (Label, No-Label) ANOVA, which found a significant effect of entity type ($F1(2,86)=59.973$, $p<.001$, $\eta_p^2=.582$; $F2(2,9)=35.84$, $p<.001$, $\eta_p^2=.89$), but no effect of Label (Label: 42% vs. No-Label: 40%; $F1(1,43)=.11$, $p=.74$; $F2(1,9)=1.03$, $p=.34$). Also, there was no interaction between Entity Type and Label ($F1(2,86)=0.2$, $p > .95$). Pair-wise t-tests showed that the means for Complex Solid entities (74.4%) differed from both Simple Solid (30%) and Non-Solid entities (18.3%) at $p<.001$. The means for the Simple and Non-Solid entities (30% vs. 18.3%) were significantly different in an analysis that treated subjects as a random factor ($t1(44) = 2.49$, $p=.043$), but not in analysis with items as a random factor

($t2(6)=1.80$, p=.12).[1]

We next compared the data from the Tsotsil speakers with those from our previous studies of Mandarin and English speakers using the same methods and stimuli. Lucy's finding that Mayan speakers weight material more highly than shape in similarity judgments than do English speakers was replicated. Overall, English participants' similarity judgments were based on shape 60%, compared to 40 % for Tsotsil participants. Similarly, English speakers were more likely to generalize the novel noun on the basis of shape than were Tsotsil speakers (57% vs. 42% by shape, respectively), replicating the standard finding in comparisons between speakers of mass-count languages and classifier languages.

Figure 2 plots the results from the Tsotsil speakers alongside the results from the English and Mandarin speakers from Li et al (2009), collapsing across the Label and No-Label conditions (since there was no effect of this manipulation in any language). Two results are immediately apparent. First, just like the English and Mandarin, Tsotsil speakers speakers differentiated the three entity types the order of Complex > Simple > Non-solids for percentage shape-match choice. Second, there is an effect of language, such that English speakers preferred the shape-match choice more often than either the Mandarin or the Tsotsil speakers.

[Figure 2 about here]

A 3 Entity Type (Complex Solid, Simple Solid, Non-Solid) x 2 Instruction (Label, No-Label) x 3 Language (English, Mandarin, Tsotsil) ANOVA on percentage of shape-match choice revealed no main effect or interaction involving Instruction (*p*s>.10). In contrast, Entity Type

[1] Although all of the participants were monolingual, it is possible that the Tsotsil spoken by older participants might differ in some ways from that spoken by younger participants who are likely to have come into increased contact with Spanish. To explore this, participants were divided into three age groups, consisting of Young Adults (younger than 30 years of age; n = 16), Middle-aged Adults (between 30 and 45; n = 11) and Older Adults (older than 45; n = 18). Adding Age Group (Young, Middle, Old) as a factor to the above ANOVA model revealed no significant Age Group effect ($F(2, 39) = .60$, $p = .56$) or interactions due to Age (all p's > .44). An Entity x Label ANCOVA with Age as a covariate also revealed no effect of age (p=.58) or age related interaction (p=.19).

(*F1*(2,206)=209.48, p<.001, η_p^2=.67; *F2*(2,9)=146.22, *p*<.001, η_p^2=.97), Language (*F1*(2, 103)=13.60, *p*<.001, η_p^2=.21; *F2*(2,18)=39.20, *p*<.001, η_p^2=.81), and the interaction between Entity Type and Language (*F1*(4,206)=4.12, *p*<.005, η_p^2=.07; *F2*(4,18)=4.71, *p*<.01, η_p^2=.51) produced significant effects.

The main effect of Entity Type was due to the fact that Complex Solids elicited more shape-match choices than the Simple Solids (81.0% vs. 35.2%), and the fact that Simple Solids received more shape-match choices than the Non-Solids (35.2% vs. 14.6%, all paired t-tests p's <.001).

The effect of Language was due to the fact that English speakers were more likely than Mandarin and Tsotsil speakers to select the shape match (58.6% shape-match vs. 31.3% and 41.0% respectively; t-tests, p's < .001). Tsotsil speakers patterned more like Mandarin speakers; t-tests found that Tsotsil speakers did not differ from Mandarin speakers by subject-analysis (*t1*(75)=1.95, *p*=.06), though the difference reached significance by items-analysis (*t2*(11)=2.51, *p*=.03), with Tsotsil speakers being slightly *more* inclined to select the shape-match than Mandarin speakers.[2]

The Entity x Language effect indicates that the difference in the percentage of shape-match choices across the three languages varies as a function of entity type. A difference score for each entity was computed for each entity type: % shape-match choice by English speakers *minus* % shape-match choice by Mandarin / Tsotsil speakers. The difference scores among English and Mandarin speakers for Complex Solids, Simple Solids, and Non-Solids were 21.1%, 43.0%, and 18.0% respectively. The corresponding difference scores among English and Tsotsil speakers

[2] However, with a more conservative measure, using bonferroni corrections, the p-values comparing Mandarin and Tsotsil speakers' choice for shape-match were insignificant for both subject-analysis (p = .17) and items analysis (p = .08).

were 19.6%, 29.5%, and 3.7%. As predicted by both Imai and Gentner (1997) and Yoshida and Smith (2003)'s feature weighting hypothesis, the biggest difference between English speakers and Mandarin or Tsotsil speakers was found for the Simple Solids. However, as the difference scores indicate, even though the effect of language is greatest for the Simple Solids, it is not restricted to that particular entity type.

Discussion

Despite the striking cultural differences between the monolingual, unschooled, Tsotsil-speaking subsistence farmers who participated in Experiment 1, and the urban, college educated participants in previous studies using these stimuli, the Tsotsil speakers' performance was virtually identical to that of speakers of other classifier languages (Mandarin and Japanese). These results strongly suggest that grammatical differences between mass-count and classifier languages, rather than other cultural factors, lead to differences on these triad tasks. Further, they suggest that Tsotsil Mayan does not differ from other classifier languages in how it influences similarity judgments and word extension.

Three results support these conclusions. First, we found a cross-linguistic difference between speakers of English and Tsotsil for both similarity judgment and word extension which was similar in magnitude to previously reported differences between English and Mandarin Chinese. Second, the word extension task generated effects similar in magnitude to a similarity judgment task that resembled the task previously used by Lucy (1992), but which used novel stimuli instead of familiar things.And third, speakers of Tsotsil and of English were affected by the same stimulus factors when making judgments for entities.

As in previous studies of classifier languages like Japanese and Chinese, Tsotsil speakers systematically differentiated between the three types of entities, providing evidence for cognitive

universals in entity construal. The fact that the Tsotsil speakers extended names for complex things according to object kind and names for non-solid substances according to their material also suggests that Tsotsil nouns do not differ radically in meaning from nouns in English or Asian classifier languages like Mandarin, contrary to Lucy's (1992) analysis of Yucatec Mayan nouns.

Still, although we did not find evidence for Lucy's strong Whorfian view, we did nonetheless find that speakers of Tsotsil, like speakers of Mandarin and Japanese, were less likely to choose shape matched stimuli, relative to speakers of English. As noted in the Introduction, there are two possible interpretations of the differences between speakers of classifier and mass-count languages. On one interpretation, favored by Imai and colleagues (see also Yoshida & Smith, 2003; Colunga & Smith, 2005), the obligatory marking of mass-count syntax leads to greater attention to the features that specify kinds of individuals, including shape and function, which in turn leads to a greater likelihood of inferring that a novel noun refers to a kind of individual object. This feature weighting interpretation explains the fact that the effects of language group are as great on the similarity judgment task as on the word extension task. However, this interpretation cannot naturally explain the finding that the same participants, if bilingual, show the Mandarin pattern of judgments if tested in Mandarin and the English pattern if tested in English (Barner et al., 2009).

The alternative explanation is the lexical statistics hypothesis. Given the greater frequency of count nouns than mass nouns in mass-count languages such as English, a novel noun used in neutral syntax is more likely to be a count noun. If speakers use lexical statistics to infer the syntactic status of noun phrases, they should be more likely to conclude that an ambiguous NP is count, and refers to a kind of individual. No such statistical evidence is available in a classifier

language, and a syntactic inference regarding count vs. mass is not required.

Some researchers have taken the similarity judgment task to be a non-linguistic task (e.g., Lucy 1992; Imai & Mazuka, 2003). If so, then the difference between speakers of classifier languages and speakers of count / mass languages are evidence for linguistic relativity, because a non-Whorfian account would predict a difference only on linguistic tasks. However, others have suggested that the similarity judgment task *is* linguistic, and requires inferences based on lexical statistics (e.g., Li, Dunham, & Carey, 2009; Barner, Inagaki, & Li, 2009). This is because in the so-called "no label" condition, the novel entities are in fact labeled by a deictic pronoun ("Look at *this*. *Which* is the same?"). On standard syntactic analyses, the word "this" is either the head of a determiner phrase that contains an elided noun, or the head of a noun phrase, such that "this" is a noun itself. On any account, sentences like this contain an NP, and thus require a syntactic decision about the mass-count status of the head noun and its agreement properties (e.g., in English, whether "this" or "these" is appropriate when more than one portion or object is present). Consequently, similar performance on the two triad tasks cannot easily adjudicate between different accounts of how language and thought interact

In the next two Experiments, we attempt to adjudicate between the Whorfian accounts of these cross-linguistic differences and the non-Whorfian lexical statistics view. Following the logic of Barner, Inagaki, & Li (2009), Experiment 2 tested how Tsotstil Mayan speakers interpret actual Tsotsil nouns, and thus removed the need for word learning inferences entirely. Using a quantity judgment task that is known to be a sensitive measure of individuation, we asked whether these real words encode individuals, as in English, or unindividuated essences, as claimed by Lucy (1992). In Experiment 3, we asked Tsotsil Mayan speakers to perform object-substance ratings for a variety of novel entities, to test whether speaking this language shifted

their perception of novel entities, as predicted by the feature weighting hypothesis (Imai & Gentner, 1997; Yoshida & Smith, 2003), or whether they exhibit no difference in perception, as predicted by the lexical statistics hypothesis.

Experiment 2: Quantity Judgment in Tsotsil

In Experiment 2, we used a quantity judgment task to explore whether the nouns of Tsotsil, unlike the nouns of Mandarin and Japanese, differ quantificationally from nouns in English. If lexical nouns in Tsotsil encode unindividuated essences, as Lucy (1992) suggests, then Tsotsil speakers should be more likely than English speakers to base quantity judgments on mass or volume, rather than number, when classifiers are omitted. They should judge, for example, that two large shirts are "more clothing" than six tiny shirts that have a less combined volume. On the strongest interpretation of Lucy's lexical semantics hypotheses about noun meanings, these judgments should not differ significantly from judgments for non-solid substances like butter. However, if Tsotsil nouns individuate like nouns in English, then Tsotsil speakers should base quantity judgments on number when tested with the translations of English count nouns object-mass nouns like "clothing" and "furniture" and should exhibit a sharp difference in their judgments for these words compared to words that are translations of substance-mass nouns in English.

In addition, we also explored how the optional Tsotsil plural marker affects quantification. For nouns that are count nouns in English (e.g., shoes, *sxonobtak* in Tsotsil), we added the plural marker when asking "Of these two men, who has more shoes" "Li chib viniketike boch'o oy ep *sxonobtak*". For nouns that are object-mass nouns in English (e.g., clothing), we used the bare noun in Tsotsil. Recall that English object-mass nouns provide criteria of individuation comparably to English count nouns. If bare nouns in Tsotsil do not provide criteria of

individuation, but rather need plural markers or classifiers to do so, Tsotsil speakers should be more likely to quantify by number when the nouns are pluralized.

Methods

Participants. Participants were 24 new monolingual Tsotsil-speaking adults (mean age=39, SD = 15, ranging from 23 to 75) recruited in Chenalhó, Chiapas. As in Experiment 1, participants received 20 pesos for their participation. They were also tested individually in a house in the center of town.

Procedure. Participants were shown photos of two characters and, via a native speaker of Tsotsil who acted as a translator, they were asked to choose which of the two characters had more, i.e., "Li chib viniketike boch'o oy ep *sxonobtak*" ("Of these two men, who has more *shoes*?"). In these photos, borrowed from Barner and colleagues (Barner & Snedeker, 2005; Inagaki & Barner, 2009), one character always had two large objects or portions of stuff, while the other character had six small objects or portions of the same kind (see Figure 3 for sample stimuli). The six small objects or portions always had a smaller combined volume and surface area than the two large objects or portions. In this way, responses based on number could be distinguished from those based on mass or volume. The character who had the greater number of items was randomized across trials.

[Figure 3 about here]

Each participant was tested on all words from the following categories: words that are count nouns in English (*sxonobtak*-shoes, *skantelatak*-candles, *svasotak*-cups, *sets'tak*-plates), words that are substance-denoting mass nouns in English (*ya'lel schichol*-ketchup, *sjuch'bil kaxlam chenek'*-peanut butter, *smostaza*-mustard, *skrema*-toothpaste), and words that are object-denoting

mass nouns in English (*schotleval*-furniture, *sk'u'spok*-clothing, *takvilsvun*-mail, *nats*-jewelry[3]). As mentioned above, we presented the translations of the English count nouns with a plural marker (i.e., the possessive –tak suffix) and the translations of the object-mass and substance mass words were without the plural marker. There were three pseudo-random item orders, each of which was given to 8 participants.

Results

Figure 4 depicts the quantity judgments made by Tsotsil speakers for each of the entities, indicating how often they made judgments based on number rather than on volume. Data were submitted to a 3 x 2 ANOVA with Word Type (English count vs. English substance-mass vs. English object-mass) as a within-subjects factor. There was a main effect of word type[4] ($F1(2,46)=47.8$, $p<.001$, $\eta_p^2 = .68$; $F2(2,9) = 301.4$, $p<.001$, , $\eta_p^2=.99$). Overall, Tsotsil speakers based quantity judgments on number significantly more for the translations of English count nouns (94.8% of the time) than substance-mass nouns (40.6% of the time), $t(23)=7.11$, $p<.001$, and more for object-mass nouns (90.6% of the time) than substance-mass nouns, $t(23)=7.08$, $p<.001$. However, there was no significant difference between translations of count nouns and object-mass nouns, $t(23)=1.70$, $p=.10$. Thus, the use of a plural marker on nouns that refer to object kinds (i.e., translations of English count nouns and object-mass nouns) did not affect the likelihood that participants' would base judgments on number.

[Figure 4 about here]

Discussion

If bare nouns in Tsotsil refer to unindividuated essences, Tsotsil speakers should quantify by

[3] No satisfactory Tsotsil translation could be obtained for "jewelry", so the word "nats", meaning "necklace", was used instead. Thus, in this trial a bare noun corresponding to an English count noun was used.

[4] For all ANOVAs, both the subject analysis (F1) and the items analysis (F2) are reported. We also ran additional analyses in which the data was arcsine-square-root transformed, and the analyses yielded the same pattern of significance using $p < .05$ cut-off.

volume rather than number when classifiers and plural markers are omitted. Two types of bare nouns were tested: English object-mass nouns and substance-mass nouns. Tsotsil speakers almost never quantified volume for those nouns that are object-mass nouns in English; they quantified by number almost 100% of the time, and did so significantly more often than for substance-mass nouns. A third type of noun, translations of English count nouns, was tested with the Tsotsil plural marker. If plural morphology, like the classifier, combines with the noun to make the individual unit salient, one might expect speakers to quantify by number more frequently when the plural marker is present. However, the presence of the possessive *-tak* plural marker did not affect quantification for nouns that refer to individuals. Tsotsil speakers were equally likely to quantify by number for both English count nouns and English object-mass nouns. Indeed, the translation of one English count noun ("necklace") was used as a bare noun, and Tsotsil participants quantified by number to an equal extent that they did for pluralized nouns.

These results fail to support Lucy's (2000) claim that the English distinction between names for objects and substances is simply "another reflex of our own mass-count distinction" (p. 333). Instead, our results suggest that the distinction between nouns that denote discrete, individuated entities and those that denote non-discrete, unindividuated entities does not depend on a language having obligatory mass-count syntax.

If anything, we found that Tsotsil speakers were more likely to treat referents as discrete individuals, relative to speakers of English. When asked to make quantity judgments for non-solid substances, like ketchup, Tsotsil speakers based judgments on number around 40% of the

time, whereas Japanese and English speakers almost never did so.[5] Although it is possible that Tsotsil speakers are indeed more likely to construe entities as individuated, a more likely possibility is that they are relatively less familiar with experimental tasks. In our study, there were twice as many nouns that denoted object kinds vs. nouns that denoted substance kinds. Thus, it is possible that Tsotsil speakers were relatively uncertain of what was required by the experiment, and adopted a strategy early in the study which they deployed for subsequent trials. Consistent with this, nine participants (38%) quantified by number for at least eleven out of twelve trials. Removing these participants (N = 9) and participants who always chose by volume at least eleven out of twelve trials (N = 1) from the analysis, the remaining twelve participants quantified by number 98% of the time for count noun trials, 91% of the time for object-mass noun trials, and 18% of the time for substance-mass noun trials, more closely matching the pattern of performance of both English and Japanese speakers.

Critically, these data fail to support Lucy's lexical semantics claim that classifier languages like Tsotsil de-emphasize the status of entities as individuals, relative to languages like English. If Tsotsil speakers were predisposed to perceive referents as portions of unindividuated essence, then they should have based judgments on number less frequently than English speakers, not more. Moreover, Tsotsil speakers clearly exhibited a systematic, robust and highly significant difference between names for objects and substances. In sum, Experiment 2 provides no evidence that Mayan languages differ from other classifier languages such as Japanese or Chinese with respect to whether its nouns provide criteria of individuation. Bare nouns that are count nouns in English ("necklace") or object-mass nouns in English ("clothing") pick out

[5] For the three category of words, count nouns, object-mass nouns, and substance-mass nouns, English speakers tested by Barner et al. (2009) quantified by number 98%, 94%, and 0% of the time respectively. Japanese speakers quantified by number 92%, 90%, and 2% of the time respectively.

individuals; "more necklace" is taken to mean the same thing as "more necklaces," not "more necklace essence."

Experiment 3: Object-Substance Rating in Tsotsil

Instead of asking participants to interpret nouns or judge similarity among entities, Experiment 3 asked them to rate, on a scale of 1-7, whether they perceive individual entities as objects or substances. The items rated were the standards from Experiment 2. If linguistic structure influences how speakers represent entities non-linguistically then Tsotsil speakers should be less likely to rate novel entities as examples of kinds of objects relative to speakers of English. On the other hand, if language does not affect how objects are construed non-linguistically, then the ratings of Tsotsil speakers should be similar to those of English speakers.

Methods

Participants. Twenty-six native monolingual Tsotsil speakers (mean age=36, SD=17, range 17 to 80) were recruited. Because of the difficulty recruiting participants, 21 individuals had previously participated in Experiment 1. Because Li et al. (2009) asked participants to rate a different set of stimuli than those used here, a group of native English speakers was also recruited for comparison. The comparison group of native English speakers consisted of 26 American college students (mean age=21, SD=2; ranging from 19 to 24), matched to the Tsotsil speaking participants in number by whether they were first tested on the triad task of Experiment 1 and on which versions (similarity judgment or word extension). This English comparison group performed identically to the English speakers reported in Experiment 1.[6]

Stimuli. The stimuli consisted of the 12 standards used in Experiment 1.

[6] This group extended by shape 95% of the time for complex solids, 65% for simple solids, and 33% for non-solids, which did not differ statistically from English speakers in Experiment 1 (95% for complex solids, 59% for simple solids, and 22% for non-solids).

Procedure. Following Li et al. (2009), participants were asked to rate entities on a scale of 1 to 7, with 1 being a "good object" and 7 being a "good substance". However, a satisfactory Tsotsil translation for the words "object" and "substance" could not be determined. Instead, Tsotsil speakers were told that some entities are like tables and shirts and can vary in what they are made of, while other entities are like ice cream and wool can vary in shape. Because the Tsotsil participants were inexperienced with the use of a Likert scale, the experimenter also introduced them to the scale by explaining that it could be used to represent age, with infants at one end and elderly people at the other. After this explanation, participants' understanding of the Likert scale was tested by asking them to select numbers that correspond to an adolescent, a middle-aged person, and a grandmother. All participants selected numbers at the low end of the scale to represent an adolescent, a number at the middle of the scale to represent a middle-aged person, and a number at the high end of the scale to represent an elderly grandmother.

For the English participants, 15 were tested using the original instructions of Li et al. and 11 were tested using these new instructions. This allowed us to confirm that the new instructions produced results comparable to the instructions used by Li et al. (2009).[7]

Next, participants were given practice trials using the two practice items from Experiment 1 (jam and fork). All of the English participants and the majority of the Tsotsil participants spontaneously selected a number higher than 4 for the jam (rating it as an instance of a substance kind) and a number lower than 4 for the fork (rating it as an instance of an object kind). Participants who did not do so received a second explanation, and were told again that things like tables and shirts should be assigned a number at the lower end of the scale, and things like ice cream and wool should be assigned a number at the higher end of the scale.

[7] An 2 Instruction (Original, New) x 3 Entity Type (Complex, Simple, Non-solids) ANOVA found no effects of Instruction ($F(1, 20) = 1.56$, $p = .23$, ns) and Instruction x Entity Type ($F(1, 20) = .588$, $p = .45$, ns).

Following these practice trials, the standard entities were placed into 12 open containers (semi-transparent Tupperware) and were presented to the participant one at a time for their ratings. Participants were tested individually, and were asked to verbally state their rating for each standard, one at a time.

Results

The mean object-substance ratings of Tsotsil speakers did not differ from those of English participants. Figure 5 shows English and Tsotsil ratings on each of three categories of entities. A 3 Entity type (Complex Solids, Simple Solids, and Non-solids) x 2 Language (Tsotsil, English) ANOVA, with Language as a between-subjects factor, found a main effect of Entity type (*F1*(2, 100)=160.94, $p<.001$, $\eta_p^2 = .76$; *F2*(2, 9) = 295.30, $p < .001$, $\eta_p^2 = .99$), but no main effect of language (*F1*(1, 50)=2.56, $p=.12$, *F2*(1, 9) = 1.75, p = .22). Thus, English speakers were no more likely to rate entities as members of object kinds, relative to speakers of Tsotsil. However, we did find an interaction of Entity Type and Language (*F1*(2,100)=35.9, $p<.001$, $\eta_p^2=.42$; *F2*(2,8)=45.5, $p<.001$; $\eta_p^2=.91$). English ratings were higher than Tsotsil ratings for Complex Solids (English: 6.62 vs. Tsotsil: 5.0; $t(50)=8.4$, $p<.0001$), but significantly *lower* for Non-Solids (English: 1.73 vs. Tsotsil: 3.25; $t(50) = 7.5$, $p<.0001$). There was no difference between English and Tsotsil speakers for the Simple Solids (English: 4.80, Tsotsil: 4.36; $t(50)=1.40$, $p=.18$). In other words, the ratings of Tsotsil speakers were relatively compressed, and were less likely to include numbers near the poles of the scales (e.g., 1,2,6,7). Despite this compression, the ratings of the two groups were very highly correlated (Pearson's $r=.903$, $p<.0001$). This is shown in Figure 6, where the average rating of each item by English speakers is plotted as a function of the average ratings of that item from Tsotsil speakers. Paired t-test of the items ratings indicate that the two groups did not differ in their likelihood to rate entities as a kind of object, $t(11)=.70$,

p=.50, n.s..

Discussion

Experiment 3 found no evidence that Mayan speakers perceive entities in a qualitatively different way from speakers of a mass-count language like English. Although Tsotsil speakers used a more compressed range of the scale than their English-speaking counterparts, they clearly recognized that the Complex Solids were the most object-like and the Non-Solids were the most substance-like. They also recognized that the Simple Solids lay in between the other two categories of entities. Here, their ratings were no different from those of English speakers, a result that is clearly at odds with the Whorfian feature weighting hypothesis, which predict that simple solid objects should be most subject to cross-linguistic differences due to the obligatory use of count syntax shifting the ontological boundaries of English speakers relative to Tsotsil speakers. Thus, like previous studies that find no difference in object-substance ratings between classifier and mass-count languages (e.g., Li et al., 2009; Barner, Inagaki, & Li, 2009), we find little support for the idea that differences in syntactic structure cause differences in how objects in the world are perceived.

General Discussion

Three experiments compared Tsotsil speakers' and English speakers' construals of material entities as kinds of substances or kinds of individuals, using materials identical to those used in previous studies that compared speakers of Asian classifier languages (Japanese and Mandarin) and English. Across all three experiments Tsotsil speakers' patterned closely with those of the Asian classifier languages. This is a striking result, given the large cultural differences between college students in Taiwan or Tokyo, on the one hand, and monolingual subsistence farmers in the highlands of Chiapas, on the other. The results of these studies provide strong evidence that

nouns in Tsotsil Mayan do not differ fundamentally from nouns in other languages, and that speakers of Tsotsil represent objects and substances much like speakers of English, Japanese, and Mandarin Chinese.

Experiment 1 confirmed Lucy's and Imai and Gentner's findings that speakers of classifier languages weight similarity of substance more heavily than do speakers of count / mass languages like English when judging similarity among material entities, or when assigning meanings to newly heard nouns. The magnitude of the effect was just the same in the Tsotsil / English comparison as in previous Japanese / English and Mandarin / English comparisons using the same materials, about half as great as in Lucy's original studies. Lucy's larger effect most probably reflected the use of familiar entities for which labels were known, and shared labels for many of the material matches. Still, a key component of Lucy's hypothesis is supported by our data: speakers of classifier languages like Tsotsil are less likely to base similarity judgments on object kind than are speakers of count / mass languages. The fact that Japanese, Mandarin, and Tsotsil speakers differ from English speakers to the same degree suggests that linguistic factors drive the cross-linguistic effects, rather than other cultural differences between educated urban participants and subsistence farmers.

The next two experiments asked whether the effects in Experiment 1 were best explained by a Whorfian account, or by lexical statistics. First, Experiment 2 tested Lucy's claim that nouns in Tsotsil Mayan fail to individuate entities, thus leading speakers of Mayan to focus on material essences. First, when asked "Who has more peanut-butter?" Tsotsil speakers based quantity judgments on mass or volume, as expected by any account. When asked to judge, "Who has more clothing?", or "Who has more shoes?" Tsotsil speakers overwhelmingly quantified in terms of the number, as do English and Mandarin and Japanese speakers. Adding or omitting plural

marking across these items had no effect on judgments, and thus did not alter the likelihood of quantifying by number. Together, these results suggest that nouns in Tsotsil Mayan individuate even when plural marking and classifiers are absent, consistent with previous findings for English count nouns and object-mass nouns and their translation equivalents in Japanese nouns (Barner, Inagaki, & Li, 2009). Thus, against the Whorfian hypothesis, learning mass-count syntax does not make speakers more likely to encode referents as individuated when learning real words. Nouns across grammatically distinct languages encode individuation identically, regardless of how number is marked syntactically.

Finally, Experiment 3 tested Tsotsil Mayan speakers with an object-substance rating task. The fact that Tsotsil speakers were affected by the same variables (solidity, complexity, shape-dependent function) in their entity construal in Experiments 1 and 3 as were English speakers speaks against Lucy's hypothesis that speakers of Mayan languages conceptualize material entities differently from speakers of mass-count languages like English. However, the findings of Experiment 1 are nonetheless consistent with the Whorfian view that speaking a mass-count language leads to a greater weighting of object kinds. Experiment 3 provided evidence against this interpretation. When explicitly asked whether entities were kinds of objects or kinds of substances, Tsostil speakers were no more likely than English-speaking subjects to rate ambiguous entities as kinds of substances. These results resemble those of Li, Dunham, and Carey (2009) who found no difference in object-substance ratings between English, Japanese, and Mandarin Chinese.

Together, the findings from Tsotsil Mayan, Mandarin, Japanese, and English suggest that cross-linguistic variation in number marking has little effect on how material entities are construed. Although reliable differences between languages are found for word extension and

similarity judgment tasks in Experiment 1, the lack of similar effects for known words in Experiment 2, and the lack of effect in object-substance rating in Experiment 3 suggest that lexical statistics, rather than Whorfian effects on object perception, drive the cross-linguistic effects on triad tasks.

In his discussion of the relation between language and thought, Whorf (1958) suggested that the world is presented to us as "a kaleidoscopic flux of impressions which has to be organized … by the linguistic systems in our minds" (p. 5). The experiments presented in this study, however, are consistent with a different conclusion – that the meanings of words in language are organized in large part by non-linguistic ontological knowledge. This appears to be true not only for classifier languages like Japanese and Chinese, but also for unrelated Mayan languages like Tsotsil. Despite lacking count syntax, nouns in all of these languages readily permit individuation. Also, when they do individuate, their interpretation does not appear to differ significantly from that of equivalent English count nouns. Although mass-count syntax can be used in languages like English to disambiguate between multiple construals of an entity, it does not appear to transform how things are perceived non-linguistically, nor to be necessary for individuation. Nouns in classifier languages can also individuate, whether or not classifiers are used.

In summary, this study of Tsotsil Mayan, placed in the context of other classifier languages like Mandarin and Japanese, indicates that speakers of classifier languages respect the same ontological distinctions as speakers of mass-count languages like English. Cross-linguistic differences in number marking do not cause speakers of different languages to perceive objects differently. Rather, all languages, whether they have count syntax or not, express a common ontology of individuals, and weight the same features identically in construing a given entity as a

exemplar of a kind of substance or a kind of individual.

References

Bale, A. & Barner, D. (2009). The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics, 26*, 217-252.

Barner, D., Inagaki, S., & Li, P. (2009). Language, thought, and real nouns. *Cognition*, 111, 329-344.

Barner, D., Li, P., & Snedeker, J. (2010). Words as windows to thought: The case of object representation. *Current Directions in Psychological Science*.

Barner, D., & Snedeker, J. (2006). Children's early understanding of mass-count syntax: individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development, 2*, 163-194.

Barner, D. & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition, 97*, 1-26.

Barner, D., & Wagner, L., & Snedeker, J. (2008). Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition, 106*, 805-832.

Bolles, D. & de Bolles, A.K. (1973).*A grammar of Yucatecan Mayan*. Komchhen, Yucatán.

Colunga, E., & Smith, L. B. (2005) From the Lexicon to Expectations About Kinds: A Role for Associative Learning. *Psychological Review*, 112(2), 347-382.Cowan, M.M. (1969). *Tzotzil grammar*. Norman, OK: Summer Institute of Linguistics.

Crane, S. (2003). Effects of grammar on entity construals and word leaning among speakers of English and Czech. Undergraduate honors thesis at Swarthmore College.

De Delgaty, A.H. & Ruiz Sanchez, A. (1978). Diccionario tzotzil de San Andrés con variaciones dialectales. México, DF: Instituto Lingüístico de Verano.

De León Pasquel, M. (1988). *Noun and numeral classifiers in Mixtec and Tsotsil*. Unpublished

doctoral dissertation, University of Sussex.

Gathercole, V.M. & Min, H. (1997). Word meaning biases or language-specific effects? Evidence from English, Spanish, and Korean. *First Language*, 17, 31-56.

Gutiérrez Ruíz, A. (2007). *On Tsotsil classifiers and pluralizers*. [Address to Frankie Chen]. San Cristóbal, Chiapas.

Haviland, J.B. (1981). Sk'op sotz'leb: el tsotsil de San Lorenzo Zinacantán. Mexico: Universidad Nacional Autónoma de México.

Imai, M. & Gentner, D. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence. *Cognition, 62*, 169-200.

Imai, M. & Mazuka, R. (2003). Re-evaluation of linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. In D. Gentner & S. Goldin-Meadow (Eds*.), Language in Mind: Advances in the issues of language and thought*. Cambridge, MA: MIT Press.

Inagaki, S., & Barner, D. (2009). Countability in absence of count nouns: Evidence from Japanese quantity judgments. In S. Inagaki & M. Hirakawa (Eds.), *Studies in Language Sciences, 8*. Tokyo: Kurosio.

Laughlin, R.M. & Haviland, J.B. (1988). *The great Tzotzil dictionary of Santo Domingo Zinacantán*. Washington, DC: Smithsonian Institution Press.

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal, *Cognitive Psychology, 58*, 487-524.

Lucy, J.A. & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In M. Bowerman and S. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press.

Lucy, J.A. (1992). *Grammatical categories and cognition*. Cambridge, UK: Cambridge University Press.

Lucy, J., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 257–283). Cambridge, UK: Cambridge University Press.

Lucy, J. A., & Gaskins, S. (2003). Interaction of language type and referent type in the development of nonverbal classification preferences. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the issues of language and thought* (pp. 465–492). Cambridge, MA: MIT.

Lyons, J. (1977). *Semantics*. Cambridge, UK: Cambridge University Press.

Martínez, A. (1818). *Gramática del tzotzil*. San Cristóbal, Chiapas.

Mazuka, R. & Friedman, R. (2000). Linguistic relativity in Japanese and English: Is language the primary determinant in object classification? *Journal of East Asian Linguistics, 9*, 353-377.

Pinker, S. (2007). *The stuff of thought: Human ideas and where they come from*. New York, NY: Viking Adult.

Quine, W.V.O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Sandhofer, C. & Smith, L. B.& Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language*, 27, 561-585.

Schuller, R. (1925). La lengua Ts'ots'il. *International Journal of American Linguistics, 3*(2-4), 193-218.

Soja, N. (1992). Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development*, 7, 29–46

Soja, N.N., Carey, S., & Spelke, E.S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*, *38*(2), 179-211.

Tozzer, A.M. (1921). *A Maya grammar*. Cambridge, MA: Peabody Museum of American Archaeology and Ethnology, Harvard University.

Vázquez López, M.R. (2004). *Chano bats'i k'op: Aprenda Tsotsil. Chiapas: Centro Estatal de Lenguas, Arte y Literatura Indígenas*.

Whorf, B.L. (1958). Science and linguistics. In E.E. Maccoby, T.M. Newcomb, and E.L. Hartley (Eds.), *Readings in social psychology*. New York, NY: Henry Holt and Company.

Whorf, B.L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Yoshida, H. & Smith, L. (2003). Shifting ontological boundaries: How Japanese- and English-speaking children generalize names for animals and artifacts, *Developmental Science, 6*(1), 1-34.

Zavala, M. (1896). *Gramática maya*. Mérida, Yucatán: Impr. de la Ermita.

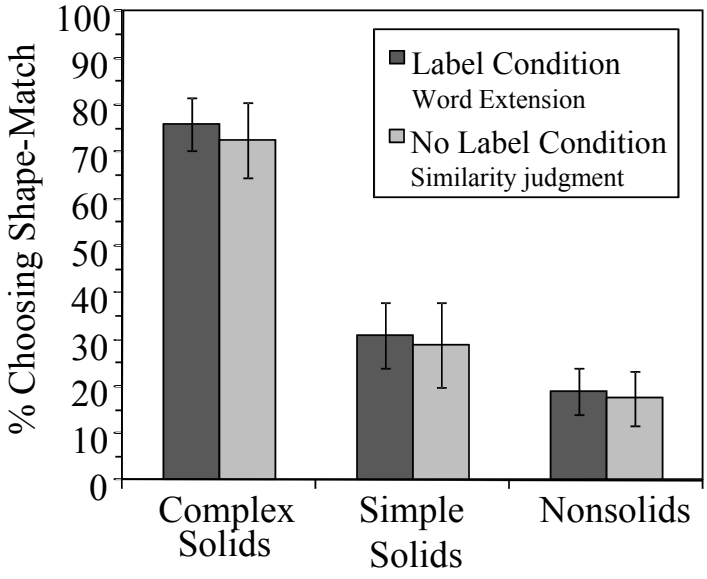

Figure 1. Tendency to choose shape-match across the three entity types broken-down by condition (Exp. 1).

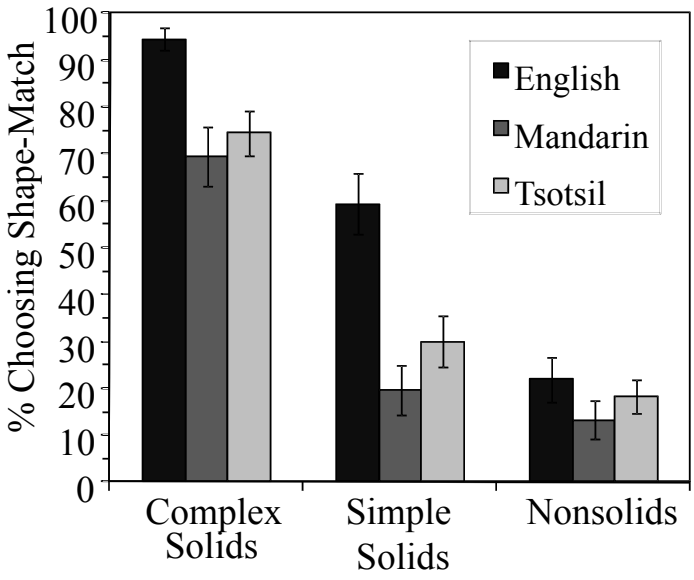

Figure 2. Comparison across language groups for the three entity types, collapsing across condition (Exp.1).

Figure 3. Quantity Judgment Sample Stimuli (Exp. 2).

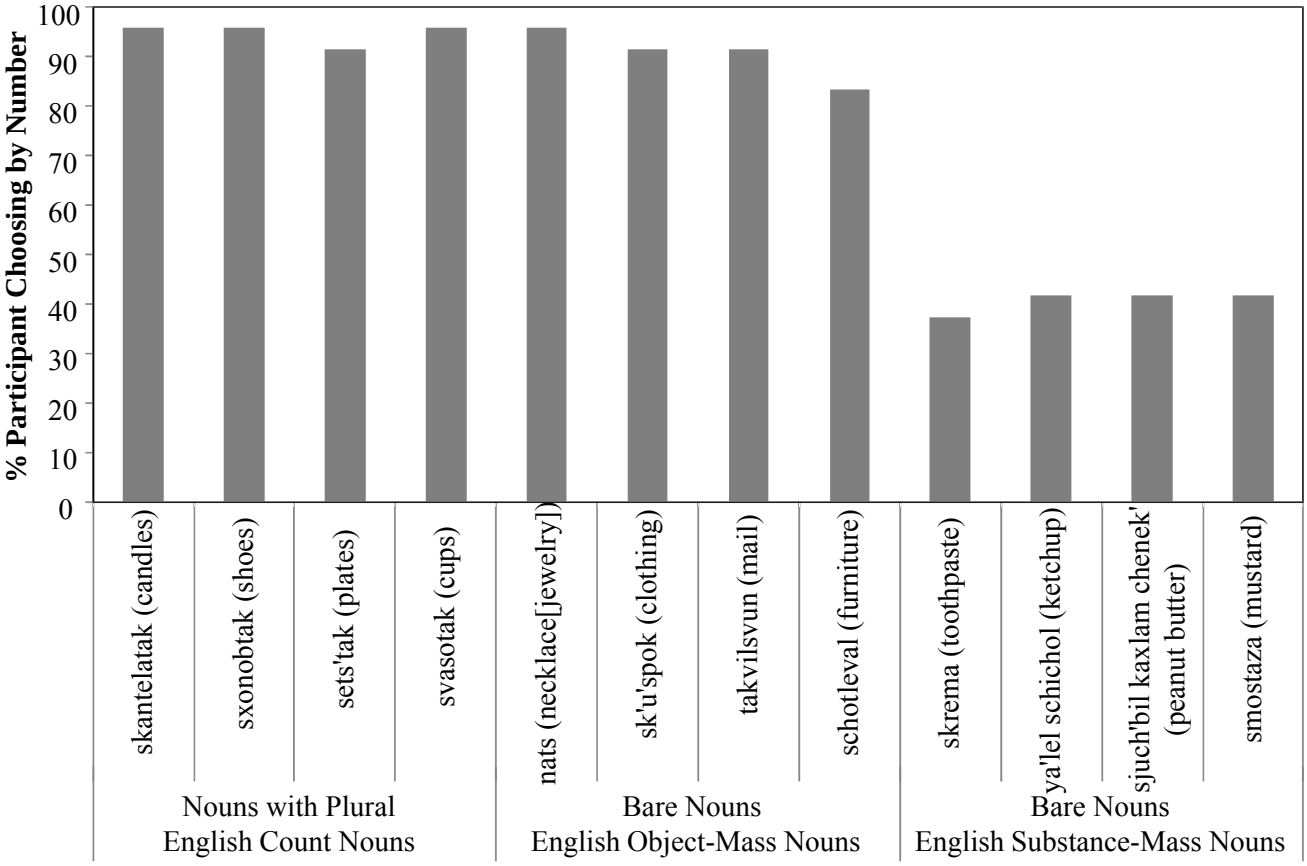

Figure 4. Quantity Judgment Results (Exp. 2). Percentage of Tsotsil speakers choosing by number for each stimuli tested.

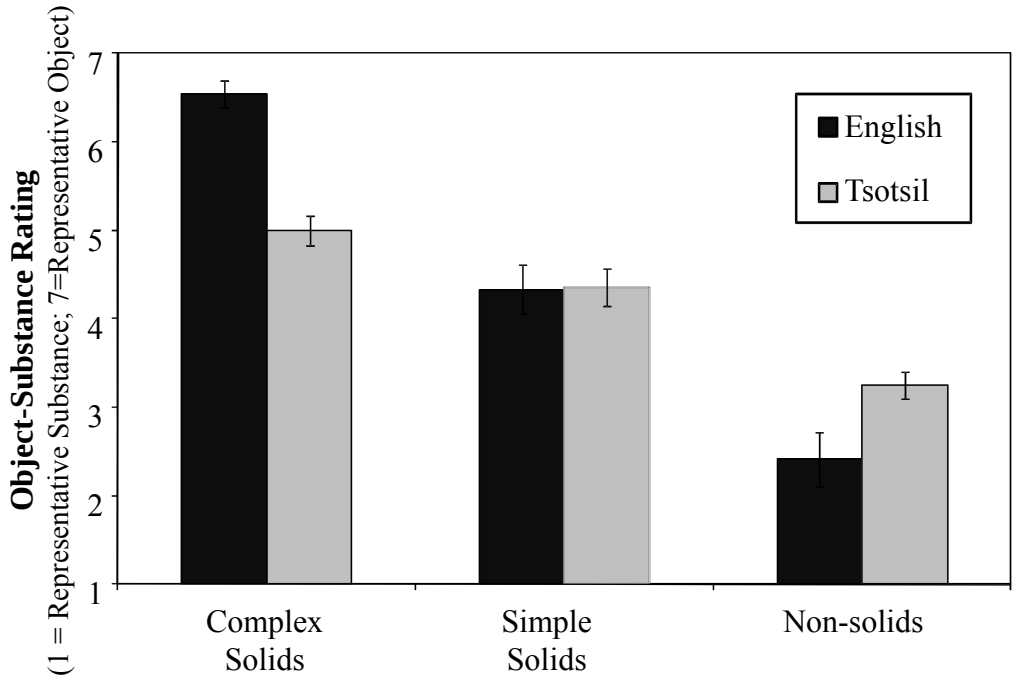

Figure 5. Comparison between English and Tsotsil speakers in entity construal rating (Exp. 3).

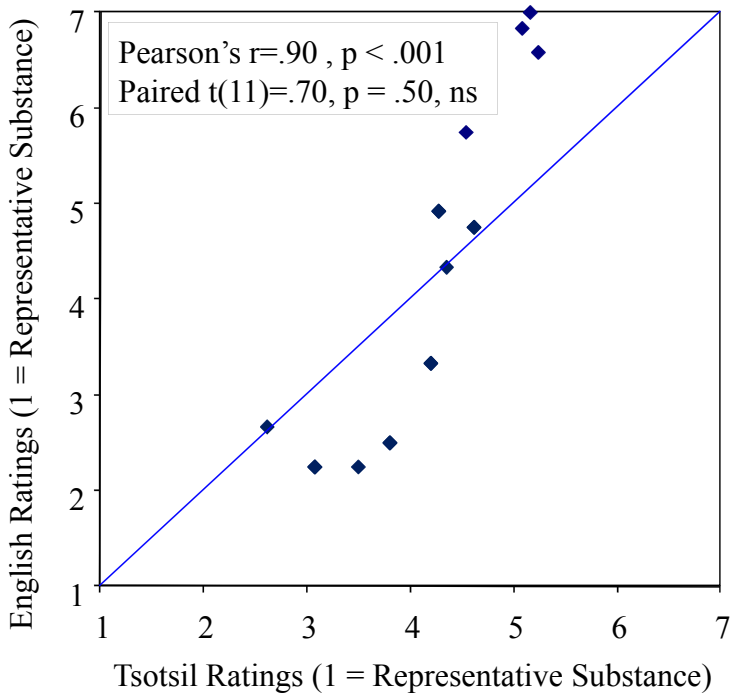

Figure 6. By-item comparisons between English and Tsotsil speakers on their entity construal ratings (Exp. 3).

Appendix A

In one early Tsotsil grammar, Schuller (1925) notes that classifiers are required when counting:

> To count irrational beings, the suffix *–kot* is used. To count the number of whips, they use the suffix *–mox*; to count round things, like balls, watermelons, melons, eggs, etc. the suffix *–pex*; to count bananas, strips of leather, or other long things, they use the suffix *–lik* (p. 206).

Similarly, in their *Great Tsotsil Dictionary of Santo Domingo Zinacantán*, widely considered to be the most authoritative dictionary of the language, Laughlin and Haviland (1988) note:

> When counting things, colonial Tsotsil, like modern Zinacantec Tsotsil, uses numeral classifiers to form quantifying expressions. That is, a numeral expression takes the form: {numeral + classifier} noun, where the classifier specifies aspects of the shape, position, size or form of the enumerated entities" (p. 105).

In fact, a more precise description of numeral expressions in Tsotsil might be {numeral *root* + classifier} noun. Unlike Mandarin, which adds a classifier directly to the cardinal numeral without modifying it, in Tsotsil, the classifier is added to a numeral root, which is formed by dropping the last syllable of the cardinal number (De Delgaty & Ruíz Sánchez, p. 461):

Table A1. Tsotsil Cardinal Numbers and Numeral Roots

Quantity	Numeral	Numeral Root
1	jun	j-
2	chib	cha'-
3	oxib	ox-
4	chanib	chan-
5	jo'ob	jo'-
6	vaquib	vac-
7	jucub	juc-
8	vaxaquib	vaxac-
9	baluneb	balun-
10	lajuneb	lajun-
11	buluchib	buluch-
12	lajchaeb	lajcha'-

Tsotsil features approximately 120 numeral classifiers. These include words like *–cot*, used for animals, birds, and entities with four legs; *–tec*, used for bird nests; *–ts'uts'unel*, used for

kisses; *–p'ej*, used for round things; and *–vo'*, used for people. Other examples, akin to English measure words, include: *–chuc*, for a handful of something; *–co'-c'abal*, for a unit of measure from the tip of the finger to the first joint; *–nulel*, which is a mouthful of corn; *–pich*, which is a small ball of something malleable like dough or mud; and *–poj*, which is a bunch or cluster. Although these and other classifiers can be used, speakers may also rely frequently on a more general default classifier *–Vb*, which can appear on all numerals except *one* (Laughlin & Haviland, 1988; for further discussion of the Tsotsil classifier system, see also Vázquez López, 2004; De León Pasquel, 1988).

Tsotsil also lacks an obligatory plural marker, although much like Mandarin Chinese and Japanese, a non-obligatory and highly infrequent plural, *–etik,* is available (see Bolles & Bolles, 1973; Cowan, 1969; De León Pasquel, 1988; De Delgaty & Ruíz Sánchez, 1978; Haviland, 1981; Martínez, 1818; Schuller, 1925; Zavala, 1896).[8] For example, the noun *nichim* (flower) can be pluralized as *nichimetik* as in (1):

(1) a. Oy ep nichim. Hay muchas flores. 'There are a lot of flower.'
b. Oy ep nichimetik. Hay muchas flores. 'There are a lot of flowers.'

This marker is similar to that found in Yucatec, which features the optional plural *–oob* (or *–ob* if the noun ends with a vowel). For example: *pek* (dog)/*pek<u>oob</u>* (dogs), *be* (road)/*be<u>ob</u>* (roads), and *na* (house)/*na<u>ob</u>* (houses). In both languages the optional plural can be used with nouns that denote both animate and inanimate entities (Tozzer, 1921; Zavala, 1896; Bolles & Bolles, 1973). Also, in both languages, nouns used without the plural are ambiguous, and can be used to refer to either singular or plural sets, as shown by the Tsotsil example in (2), which can be used to describe the purchase of either one flower or a plural set of flowers.

[8] We confirmed the non-obligatory status of the plural in an elicitation task. We tested ten adult monolingual speakers of Tsotsil (mean age=42, ranging from 23 to 72) who were shown and asked to describe a total of six pictures which depicted: three tables, three bananas, three bunches of bananas, three branches of bananas, three trees, and eight wooden tablets. None of the Tsotsil speakers used the *–etik* plural marker to describe the images.

(2) Lajmankutik nichim. Compramos flores. ‘We bought flower(s).’

Finally, similar to Mandarin, Japanese, and other classifier languages, Tsotsil quantifiers do not exhibit a distinction between those that can denote countable individuals vs. those that cannot.[9] This is shown by the examples in (3) which show that the same quantifier, translated as *many/much*, can be used for sand and cows in Tsotsil (see 3a and 3b), just as the English equivalent of *few/little* can be used for both small numbers of chickens and small amounts of salt (3c and 3d):

(3)	a. ep vacax	muchas vacas	many cows
	b. ep yi’	mucha arena	a large amount of sand
	c. jutuk alak’	pocos pollos	few chickens
	d. jutuk ats’am	poca sal	a small amount of salt

[9] Note there are mass-count languages in which there are often no distinct quantifiers used with nouns for object and substance kinds. For example, in Czech, a language with obligatory singular-plural morphology, the quantifier *moč* means both “many” (as in “many books”) and “much” (as in “much water”). Similarly, the quantifier *málo* means both “few” (as in “few books”) and "little” (as in “little water”).