Part-whole categorization is culture-specific

Timothy J. Tilbe (timothyt@buffalo.edu)
Department of Linguistics, University at Buffalo, 609 Baldy Hall
Buffalo, NY 14260 USA

Jürgen Bohnemeyer (jb77@buffalo.edu)
Department of Linguistics, University at Buffalo, 609 Baldy Hall
Buffalo, NY 14260 USA

Gabriela Pérez Báez (PerezBaezG@si.edu)
Department of Anthropology, National Museum of Natural History,
P.O. Box 37012 Smithsonian Institution, Washington D.C., 20013-7012 USA

Dedre Gentner (gentner@northwestern.edu)
Department of Psychology, Northwestern University, Swift Hall 102, 2029 Sheridan Road
Evanston, IL 60208 USA

Abstract
We present two experiments on the role of culture in the categorization of object part-whole structures. A triadic categorization task pitted shape against function as factors driving similarity judgments on selected parts of different types of objects. Speakers of American English were significantly more likely than speakers of two indigenous languages of Mexico, Tzeltal Maya and Isthmus Zapotec, to choose categorization by function, even when familiarity of the various stimulus objects was factored in. In the second study, members of the two indigenous groups matched parts of a doll to parts of novel objects of unfamiliar shape. The Tzeltal participants were significantly more likely to match according to a shape-analytical algorithm rather than global analogy, consistent with predictions based on prevalent strategies in verbal part labeling in the two languages. We conclude that while cognition of object parts undoubtedly has a strong biological basis, there are also robust cultural effects.

Keywords: object meronymy; meronym; shape perception; function; cross-cultural research

Introduction
The ability to categorize objects crucially involves identifying their parts. In general, merology – the conceptualization of parts, and of how they relate to each other and to the wholes they form – is deeply involved in how humans make sense of their physical world.

A strong case can be made that the segmentation of physical objects into parts has a basis in shape recognition (Biederman 1987; Marr 1982; Palmer 1977; Tversky & Hemenway 1984; inter alia) and is thus likely biologically grounded. At the same time, function plays a key role in the categorization of both body parts of living things and object parts of artifacts (e.g., Croft & Cruse 2004: 153-156; Rose & Schaffer 2015; Svorou 1994: 78-79, 91-92; Tversky 1989). Ears and lids can come in a great many distinct shapes; what unites these diverse manifestations is the function they play in the whole of which they are a part. But the attribution of functions depends on knowledge, beliefs, and assumptions that are at least to a large extent learned. Function dependence thus creates an opening for cultural effects in cognition of parts.

In addition, there is some evidence suggesting that the role of shape, and geometry more generally, in merological cognition may be to some extent subject to cultural variation as well. This evidence comes from meronomy, the nomenclature for object parts. In-depth studies of the meronomy of non-Indo-European languages are few and far between. But the few available reports present evidence of striking differences vis-à-vis the terminology familiar from English and other European languages.

Our interest in this complex was aroused by descriptions of the meronomies of two indigenous languages of southern Mexico, Ayoquesco Zapotec (MacLaury 1989) and Tzeltal Maya (Levinson 1994). Both languages belong to the Mesoamerican linguistic and cultural area, the members of which have been in contact with one another for millennia (Campbell, Kaufman & Smith-Stark 1986).

A feature that the accounts of MacLaury (1989) and Levinson (1994) converge on is a core set of meronyms that are assigned both to body parts of humans and animals and to the parts of inanimate objects (though not generally to plants). They claim that labeling of object parts with general-purpose body part terms is pervasive, and based largely on shape and geometry. In contrast, function-based meronyms for inanimate objects are largely absent.

Where the two systems appear to diverge is in the strategies used to assign the generalized meronyms to object parts. MacLaury (1989) describes a strategy strictly based on a global analogical mapping from the human body in canonical erect position to the object in its actual orientation at the time the assignment pertains to. (See Figure 1A for the part labelings that this account predicts for a novel object.) The parts that are named in this manner, the ‘head,’ ‘face,’ ‘sides,’ ‘back,’ and ‘buttocks,’ have fixed spatial
relationships to each other in any object. For instance, if one knows which part of the object is called its ‘face,’ and the vertical axis is determined, one can correctly predict the locations of its ‘head,’ ‘sides,’ ‘back,’ and ‘buttocks.’ The orientation of the object with respect to gravity is crucial; for example, the topmost part is the ‘head,’ no matter what the structure of that part happens to be, and changing the orientation of the object causes the labels to be reassigned.

In contrast, Levinson (1994) discusses the process by which meronyms are assigned to objects in Tseltal as an algorithm that takes a visual segmentation of the object as its input. (See Figure 1B for an example this account’s predicted labelings for a novel object.) Axes of generalized cones are identified for the main volume and any secondary volumes of the object, as well as axes of symmetry. A sense of direction is assigned to each axis, and meronyms are assigned to the ends of the axes, in some instances taking the shape of the part into account. So, for instance, the default meronym for the head (in the vector sense) of the main axis is the word for ‘head,’ but if that region of the object is pointy, ‘nose’ is used, and if it is a negative space, ‘mouth’ is used. This system does not take into account the orientation of the object with regard to gravity, and because the axes are in certain ways independent of each other, meronyms do not occur in a fixed spatial schema.1

The question we wish to address in this paper is whether these putatively distinctive properties of meronym are restricted to language, or whether they are associated with deeper cognitive differences - between Mesoamericans and English speakers on the one hand, and between Zapotec and Tseltal speakers on the other. We present two studies. Experiment 1 explores the respective role of shape and function in object part categorization, comparing data from speakers of Tseltal, Zapotec, and American English in a three-population design. Based on the available descriptions of verbal behavior, we predict function to play a greater role in the morphological categorization of Americans than in those of either Tseltal or Zapotec participants.

Experiment 2 compares Tseltal and Zapotec participants in terms of their preference for categorizing the parts of unfamiliar objects by comparing them globally to the human body vs. by doing so based on the shape-analytical algorithm even when it is not licensed by a global mapping. If the differences in part categorization strategies go beyond language, we predict that the Tseltal participants should be more likely to prefer mappings that are at odds with global analogies.

**Experiment 1: shape vs. function**

Speakers of English, Tseltal Maya, and Isthmus Zapotec compared images (with one exception, photographs were used) of part-whole configurations. Isthmus Zapotec is closely related to Ayoquesco Zapotec as described by MacLaury (1989); ongoing field research by the third author suggests that MacLaury’s analysis of Ayoquesco meronymy applies to Isthmus Zapotec as well as far as the predictions of the present study are concerned, though with additional complications (Pérez Báez 2011). Each trial involved a triad of images. The participants selected the configuration that was least like the other two. The triads were composed so as to trade off functional against shape-based similarity. Since function attribution likely depends on the participants’ knowledge of the object, a norming study was carried out to assess their familiarity with the stimuli.

**Method**

**Participants.** 27 participants of each population were recruited at field sites in La Ventosa, Oaxaca, Mexico (Isthmus Zapotec – 16 women, 11 men; 14 young adult, 9 middle-aged, 4 elderly) and Tencajapa, Chiapas, Mexico (Tseltal Maya – 18 women, 7 men; 16 young adults, 8 middle-aged, and 1 elderly, with 2 participants’ demographic information missing), the University at Buffalo and in Raynham, Massachusetts (English – 19 women, 8 men; 17 young adults, 6 middle-aged, and 4 elderly). Recruitment was conducted by word of mouth and at the University at Buffalo by flyer. Participants completed the tasks in about 30 minutes and were compensated 100 Pesos (approximately $5) and $10, respectively.

**Materials.** Each participant was given 12 test trials interspersed with 17 fillers, for a total of 29 trials, preceded by one practice trial of the same design as the fillers. The three pictures for each trial were printed on a single sheet, and the sheets placed into a binder so that one triad could be displayed at a time. The placement of the images on each page was pseudo-randomized, in order to reduce any possible bias toward choosing the picture in any particular one of the three positions.

Each trial consists of showing a participant three pictures of artifacts or plants that are presumably familiar to all three populations, with certain parts highlighted in red, and asking them to choose the one whose highlighted part they judge to

---

1 MacLaury (1989) characterizes the application of Zapotec body part terms to parts of inanimate objects as analogical or metaphorical mappings from human body parts. An alternate view, taken by Levinson (1994) in his work on Tseltal, is that the terms are general abstractions and not metaphorical.
be most different from the other two. The experimental triads are designed to pit shape against function, in that there is a pivot object part and two alternates: the pivot shares its shape with one alternate, and its function with the other. An experimental triad is shown in Figure 2.

The filler trials use the same visual layout as the experimental trials, and the action the participant is expected to perform is the same. However, instead of a single pivot, there are two pictures in which the indicated parts of the object are similar in both shape and function, while the remaining picture’s indicated part is the odd one out in terms of both shape and function. Therefore, a shape-based strategy and a function-based strategy would tend to produce the same response. This provides a check on the participants’ attention to and comprehension of the task.

**Procedure.** The participants were instructed in their native languages by the first author and, in the case of the Tseltal and Zapotec participants, by bilingual assistants with experience in linguistic field research, to pick out the part-whole configuration in each triad they considered least like the others. The following standardized instructions were used: “In this game, I am going to show you some drawings and photos of various things – three at a time. And if you are not sure of what some of the objects are, please ask me. One part of each thing is red. Two of the parts are more similar, and the other is different. I want you to look at those parts, and find the different one. You should compare only the parts, not the whole objects. When you decide which is different, circle it using this marker. For example, let’s look at these three [the practice triad]. Here there’s a dog, with its leg red. And here is a cat, with its head red. And a pig, with its leg red. So, which part is different? Correct, the head. And so, do you have any questions before beginning?”

**Norming.** As part of the follow-up task ‘Shape-Function Norming,’ participants rated the familiarity of each individual picture used in the Shape-Function Triads. This serves the purpose of checking their interpretation of the pictures and providing additional factors for statistically modeling the experiment’s results. Participants rated each of the three pictures in each experimental trial for familiarity, using a five-point Likert scale in their native language.

**Results**

**Participant exclusions.** The responses of two Zapotec participants were excluded from the analysis because their performance on the filler trials was below the pre-established 80% threshold, indicating that they did not sufficiently comprehend the task.

**Trial exclusions.** Trials were excluded from the analysis in two situations. In one of these, the participant chose the pivot as the odd one out in the trial in question, which is not interpretable as classifying by shape or function (11 trials in the Zapotec data set, 12 in the Tseltal data set, and 8 in the English data set). In the other situation, they did not give an answer to that trial at all (1 trial in the Zapotec data, 0 in the Tseltal data, and 0 in the English data). Approximately 4% of the data points were missing or excluded.

**Analysis.** Figure 3 shows the breakdown of responses. The responses to the norming scale, interpreted as numerals, were centered so that -2 corresponded to “completely unfamiliar” and 2 corresponded to “completely familiar.” (The standard deviations of the ratings for each object ranged from 0.37 to 1.26 for the Zapotec speakers, and from 0.23 to 1.49 for the Tseltal speakers.) Within each language population, the mean of these familiarity scores was calculated for each picture in order to generalize across participants. Any picture whose average rating was equal to or greater than 1 was counted as “familiar,” and any whose average rating was less than 1 was counted as “unfamiliar.” The overall familiarity of each triad was coded by the number of familiar and unfamiliar pictures it contained, from “A” for three familiar pictures through “D” for none, and this code was treated as an ordinal factor in the regression models of the responses.

A binomial logistic mixed-effects regression model was fitted with population (identified by language) and code of triad familiarity as fixed factors, and random intercepts for participant and trial. Tseltal and Zapotec participants proved

---

2 Since the instruction for the practice trial contained the meronyms ‘head’ and ‘leg’, a subvocal rehearsal effect on the test trials cannot be ruled out.
significantly different from American participants at the p<.001 level and from one another at the p<.05 level. No significant effect of familiarity was obtained.

Discussion
As predicted, the Mesoamerican participants were more likely to categorize parts by shape than the American participants. Shape in fact strongly dominated the Tseltal and Zapotec speakers' categorizations (at 65.4% and 76.8% of responses), whereas function strongly dominated among the English speakers (67.7%).3 Familiarity did not appear to significantly affect these ratings, suggesting that the selection of the stimuli from among objects of everyday interactions for all three cultures was successful. It is possible, however, that the similarity judgments were influenced by subvocal use of meronyms.

Experiment 2: global vs. shape-analytical mapping
The purpose of this experiment was to test for population differences in merological cognition, by obtaining responses of nonverbal mapping from body parts to the parts of novel objects. This was done with the researcher indicating various parts of a humanoid doll, and asking the participants to find as many corresponding parts as they could on the Novel Objects. In order to test for effects of subvocal rehearsal, this task had two conditions: In one condition, the participants did a verbal elicitation task before the experiment, and in the other condition, they did not.

Method
Participants. 44 Tseltal speakers (29 women, 16 men; mean age 39.7, SD 12.6) and 45 Isthmus Zapotec speakers (33 women, 12 men; mean age 33.2, SD 13.6) were recruited and tested in La Ventosa, Oaxaca, Mexico (Isthmus Zapotec) and San Cristóbal de las Casas, Chiapas, Mexico (Tseltal), relying again on word of mouth. The task took about an hour to complete and the participants were compensated 100 Pesos (approximately $5).

Materials. The tasks used as stimuli a set of six solid plastic forms, part of the MesoSpace Novel Objects stimuli.4 These are abstract forms that, as far as possible, bear little resemblance to any item familiar to the participants, so that the objects are not biased toward any particular meronym assignment strategy, and the participants have to fall back on their general principles for construing merological structure. Figure 1 features an example. For the practice trial, instead of a Novel Object, a blobby humanoid figure made of Sculpey modeling clay was used. A plastic action figure doll representing a young adult male, fairly realistic in proportions, was used to represent a target body part in each trial - either the head, face, side (flank), back, or buttocks. These five parts were used because both MacLaury (1989) and Levinson (1994) had identified the meronyms for them as belonging to the languages' most productive meronymic systems, and these Zapotec and Tseltal terms are rough translation equivalents with regard to the human body. The participants used bits of Play-Doh to mark the parts of the Novel Objects that they judged as corresponding with these target parts. This doll was in a standing position in each trial, which, if an orientation-dependent mapping strategy is used, would favor the choice of the uppermost part of the Novel Object as corresponding with the doll’s head, the part(s) of the Novel Object to the participant’s left or right as corresponding to the doll’s side, etc.

Procedure. Conditions: In order to detect possible subvocal rehearsal effects, half of the participants carried out a verbal labeling task prior to the experiment; the other half did not do this task. Participants were randomly assigned to one condition or the other. The verbal elicitation task was administered as follows: the participant was handed each of the Novel Objects, one at a time (in an opaque bag to avoid imposing any orientation on the object), asked to take it out and inspect it from all sides, and then prompted to delineate and label its parts. The experimenter recorded the delineated areas on two-dimensional images of the objects.

Setup: In each trial, the experimenter set one of the Novel Objects on the tabletop directly in front of participant, sticking it into a base of Play-Doh if it would not stay in the desired orientation unsupported. Each Novel Object was therefore relatively in front of the participant and absolutely to the north (the participants were seated to face north). Trials: In each trial, one of the target parts of the doll was manually indicated by the experimenter. In view of the importance of vertical orientation in Zapotec meronymy, the trials were administered in two orientation variants: “aligned” (that is, the gravitationally-defined vertical axis coincides with the Novel Object’s algorithmically-defined ‘model axis’, i.e., the axis from which the central volume of the object is generated), or “unaligned” (these axes are orthogonal). This yielded a total of 60 test trials: 5 doll parts x 6 Novel Objects x 2 orientations. In addition, there was one practice trial employing instead of a Novel Object a blobby humanoid figure made of Sculpey, designed to abstractly resemble the doll.

Instructions: The participants were instructed that when the experimenter indicated a part on the doll by delineating it with a finger, they should mark with the Play-Doh as many parts on the Novel Object as they thought corresponded to the doll’s part, whether this resulted in no part being marked, just one part, or multiple parts. The

3 An anonymous reviewer points out that, since the instructions indicated there was a correct answer for each trial, the English speakers' responses may have been motivated by trying to find counterintuitive answers, as in an IQ test.
4 A total of nine Novel Objects were originally designed by the third author and produced for the project Spatial language and cognition in Mesoamerica ('MesoSpace'; NSF Award BCS-0723694) directed by the second author.
participants were instructed in their native languages by bilingual research assistants working with the first author as follows: “I’m going to give you some objects, one at a time. This first object is an example. You can turn it around to see how it is. And I’m going to show you some part of this doll, and you should decide if the object has a part of the same kind. It’s possible that it doesn’t have any. In that case, simply tell me that it isn’t there. It’s also possible that it has a part like that, or more than one at a time. In that case, take a bit of this Play-Doh and stick it to the part or parts that are similar. And when it’s finished, I want you to lift the object and turn it slowly to show what you have done.”

**Recording.** The responses, in terms of the landing sites on the Novel Objects the participants marked with Play-Doh, were recorded by the first author verbally in English on a coding sheet. The sessions were videotaped in their entirety.

**Results**

**Participant exclusions.** One Zapotec participant’s responses had to be excluded because the participant appeared unable to grasp the instructions.

**Trial exclusions.** 16 of the 60 test trials were excluded from the analysis because the algorithm Levinson (1994) proposed for meronymic labeling in Tseltal predicted that the particular Novel Object lacked a corresponding part, and the Tseltal participants nevertheless in almost all cases identified some part of it despite having been given the option not to select a mapping. These responses therefore could not be evaluated for whether they fulfilled the predictions of the algorithm. One Tseltal trial was not completed. The analysis was thus performed on 1,936 trials with Zapotec participants and 1,979 trials with Tseltal participants.

**Coding.** In order to code the responses for whether they fit the global mapping account, simplifying assumptions were adopted. When the ‘head’ of the doll had been indicated, the global prediction was considered fulfilled if and only if the Play-Doh was placed somewhere on the upper region of the object (as defined by the vector of gravity). Globally, the ‘buttocks’ had to be on the lower region, the ‘face’ on the region toward the participant, the ‘back’ on the region away from them, and the ‘sides’ on the regions to their relative left or right. These regions were interpreted as both surfaces and volumes. So, for example, a placement that was both on the top surface of the object, and also displaced from the center of that surface in the direction toward the participant, would satisfy the prediction for ‘head’ by virtue of being on the upper part of the object, and would also satisfy the prediction for ‘face’ by virtue of being on a volume part that is toward the participant. The volume interpretation of parts was also followed for the algorithmic predictions.

**Response types.** Four response types were distinguished, based on whether the proposed match was predicted solely by global mapping (‘Global only’), solely by Levinson’s (1994) shape-analytical algorithm (‘Algorithm only’), by both, or by neither. From MacLaury’s (1989) global mapping account, two key predictions were derived: the ‘face’ of any object faces toward the observer, and the ‘head’ of any object points up against the pull of gravity. Since Levinson’s (1994) algorithm is orientation-independent, it was assumed that given a part of the doll and a Novel Object, the intrinsic location of the matched part on the Novel Object will be constant across varied orientations.

**Analysis.** Figure 4 shows the breakdown of the four response types across the two populations.

![Figure 1. Experiment 2 responses by population and type.](image)

Using Begg-Gray approximation of multinomial logistic regression (Begg & Gray 1984), four binomial logistic mixed effects regression models were fitted, one for each response category, with population identified in terms of language, condition, alignment, and trial as fixed factors, and random intercepts for the stimulus doll part, the Novel Object, and the participant. The Algorithm-only model showed significant effects of population and alignment at the p<.001 level. The Global-only model showed effects of alignment and trial at the p<.001 level. There was a significant interaction between Zapotec and alignment at the p<.05 level. The Both model yielded a significant effect of population at the p<.01 level and alignment and trial at the p<.001 level. The Neither model yielded no significant effects. None of the models produced a significant effect of condition.

**Discussion**

As predicted, the Zapotec participants were significantly more likely to propose matches that agreed with global analogical mapping, but violated Levinson’s (1994) shape-analytical algorithm. Also in line with predictions, the orientation of the Novel Object had a significant effect on the Zapotec participants’ matches, but not on those proposed by the Tseltal participants. Also as predicted, the “aligned” trials favored responses in the ‘both’ category. There was no effect of condition; we take this to suggest that subvocal rehearsal played no major part in the results. Subvocal rehearsal would have predicted that the two populations should have performed significantly more
different from one another in the Verbal-priming condition, contrary to fact.  

General discussion
In both experiments, language proved a significant predictor of nonverbal mereological categorization: English speakers significantly preferred categorizing parts in terms of function, whereas Tzeltal and Zapotec speakers were significantly more likely to categorize parts by shape (Experiment 1). And Zapotec speakers proved significantly more likely than Tzeltal speakers to adhere to global analogy in mapping the parts of the human body to those of inanimate objects of unfamiliar shape, and were also significantly more likely to factor the orientation of the objects into their matches (Experiment 2).

We cannot exclude the possibility of subvocal rehearsal effects in Experiment 1. It is possible that the participants used their native languages for guidance in deciding between function-based and shape-based categorization. Future research will have to determine to what extent our results are truly representative of the nonverbal cognition of these groups. However, our findings are in line with previous research suggesting that geometry, as opposed to function, plays a relatively greater role among Mesoamericans compared to Westerners (Lucy & Gaskins 2001). Meanwhile, in Experiment 2, we plausibly ruled out a significant contribution from language as a direct resource, suggesting robust differences in nonverbal cognition.

The findings presented here are also in line with Whorfian interpretations according to which language use may habituate speech communities to particular biases in mereological cognition and serve as a conduit of their cultural transmission (Bohnemeyer et al 2015). Here, too, we must defer to future research for ascertaining whether language merely reflects mereological cognition or is a causal factor in it.

Conclusions
We have provided evidence of the existence of significant cross-cultural differences in the mereology of physical objects. This is hardly surprising, as the categorization of objects and their parts clearly depends on acquired knowledge. What is surprising, in our view, is that research into such cultural effects is still in its infancy. We hope to have made a small contribution towards rectifying this.

Acknowledgments
This material is based upon work supported by the National Science Foundation under Grant No. BCS-1354277, ‘Doctoral Dissertation Research: Parts and Wholes in Mesoamerican Language and Cognition,’ awarded to the first author.

We gratefully acknowledge the contributions of Jean-Pierre Koenig, Peter Pfordresher, Gilles Polian, Reyna López López, Rosaura López Cartas, Antonia Sáñez Girón, María Gómez Sánchez, and John Pate.

References

5 However, this result is also consistent with subvocal rehearsal being already at ceiling even in the non-primed condition.