Confidential Review Copy. DO NOT DISTRIBUTE.

Instructions for EMNLP-IJCNLP 2019 Proceedings

Anonymous EMNLP-IJCNLP submission

Abstract

1 Introduction

The real-world objects that we interact with in our every-day life can be categorized into many thousands and maybe millions of categories. And even a single object can be member of many categories, i.e. at different taxonomical levels or in different parts of a taxonomy. For instance, both objects in Figure 1 are at once instances of CAKE, CHEESECAKE, DESSERT, SWEET, PASTRY, FOOD etc. Hence, when speakers name objects, e.g. when referring, they have to select a lexical item from a complex network of concepts and competing lexical alternatives.

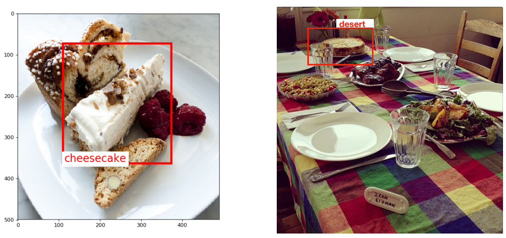

Figure 1: Two objects of the same type of cake, with different names in VisualGenome

To date, research in NLP has surprisingly little to say about object naming, despite the fact that there has been a recent explosion of interest in various, and even complex, language & vision tasks ranging from image captioning (**???**) to e.g. visual dialogue (**??**). In contrast, closely related areas, such as computer vision and cognitive science, have investigated very related tasks in quite some depth: object recognition systems developed in the area of computer vision are now able to classify images into thousands of different categories (e.g. **?**). Furthermore, work on concepts, following the seminal work by Rosch, suggests that objects are typically conceptualized at a preferred level of specificity called the **entry-level**. Psycho-linguistic studies have been able to support this theory based on collections of so-called object naming norms.

This paper aims at addressing the genuinely linguistic questions revolving around the phenomenon of object naming by (i) presenting a collection of high-quality, large-scale naming data, and (ii) analysis methods for this data and (iii) a first baseline model that accounts for the semantic flexibility of names for objects in real-world images. From computer vision, we borrow the idea of modeling realistic visual objects in realistic scenes (real-world images), but go beyond the simplistic assumption that object names correspond to unambiguous labels in a flat classification scheme (with no conceptual relations between the labels). From psycholinguistics, we borrow the idea of eliciting natural, representative naming data from many subjects, but go beyond using artificial, highly stylized objects.

2 Related Work

3 Analysis

3.1 Agreement

We compute the following agreement measures:

- **top %**: for each object, we calculate the relative frequency of the most common name, and then average over all objects
- **SD %**: for each object, we calculate the Snodgrass agreement measure, and then average over all objects
- **=VG**: the proportion of objects where the most frequent name coincides with the name annotated in VisualGenome

domain	all synsets % top	SD	=VG	max synset id	% top	SD	=VG	min synset id	% top	SD	=VG
person	0.52	2.14	0.50	professional.n.01	0.61	2.02	0.20	athlete.n.01	0.34	2.67	0.36
tableware	0.52	1.92	0.40	crockery.n.01	0.52	1.92	0.40	crockery.n.01	0.52	1.92	0.40
clothing	0.64	1.59	0.70	neckwear.n.01	0.79	0.91	0.77	footwear.n.01	0.47	2.55	0.40
instruments	0.66	1.52	0.79	furnishing.n.02	0.67	1.50	0.80	kitchen_utensil.n.01	0.60	1.85	0.56
solid food	0.67	1.43	0.56	baked_goods.n.01	0.67	1.43	0.56	baked_goods.n.01	0.67	1.43	0.56
structure	0.67	1.55	0.73	bridge.n.01	0.75	1.21	0.87	place_of_worship.n.01	0.46	2.26	0.08
vehicle	0.72	1.13	0.71	train.n.01	0.93	0.43	0.99	aircraft.n.01	0.52	1.50	0.41
food, nutrient	0.72	1.27	0.68	edible_fruit.n.01	0.80	0.89	0.79	vegetable.n.01	0.52	1.99	0.15
plants	0.79	0.86	0.73	flower.n.01	0.79	0.86	0.73	flower.n.01	0.79	0.86	0.73
ware	0.82	0.96	0.94	cutlery.n.02	0.82	0.96	0.94	cutlery.n.02	0.82	0.96	0.94
tool	0.86	0.73	0.94	tool.n.01	0.86	0.73	0.94	tool.n.01	0.86	0.73	0.94
animal	0.91	0.43	0.94	feline.n.01	0.95	0.29	0.99	fish.n.01	0.39	2.53	0.55
all	0.70	1.35	0.73								

Table 1: Agreement in object names for objects of different domains, if applicable, synsets with maximal and minimal agreement (top %) are shown

Table 1 shows that, overall, our annotators achieve a fair amount of agreement in the object naming choices. The domain where annotators agree most is the animal domain, which, interestingly, happens to be the domain that has been mostly discussed in the object naming literature. *//sz: ... much more to say//*

Why is naming more flexible in certain domains than in others?

3.2 Lexical relations

In this section, we take a closer look at the lexical variation we observe in our data set. We analyze the data points where participants attributed different names to the same object and extract a set of pairwise **naming variants**. These naming variants correspond to pairs of words that can be used interchangeably to name certain objects. For each object, we extract the set of naming variants $s = \{(w_{top}, w_2), (w_{top}, w_3), (w_{top}, w_4), ...\}$ where w_{top} is the most frequent name annotated for the object and $w_2 ... w_n$ constitute the less frequent alternatives of w_{top}. The **type frequency** of a naming variant (w_{top}, w_x) corresponds to the number of objects where this variant occurs. The **token frequency** of (w_{top}, w_x) corresponds the count of all annotations where w_x has been used instead of w_{top}. In Table 3, we show the the naming variants with the highest raw token frequency for each domain. *//sz: domains need to be updated//*

The naming variants can be grouped according to their lexical relation, as follows:

- **synonymy**: e.g. aircraft vs. airplane
- **hyponymy**: e.g. man vs. person

relation	% types	% tokens	av. depth
co-hyponymy (closure, max depth=10)	0.889	0.551	3.479
hyponymy (closure, max depth=10)	0.097	0.328	2.204
synonymy	0.015	0.121	1.000

Table 2: Lexical relations between naming variants according to WordNet, for the set of name pairs where both words can be found in WordNet and stand in a *//sz: should we produce this table for the different domains?//*

- **co-hyponymy**: e.g. chicken vs. dinner, swan vs. goose
- **no relation**: e.g. desk vs. apple

Research on object naming following the idea of entry-level categories has, essentially, exclusively looked at names that stand in a hierarchical relation (i.e. hyponymy/hypernymy).

We use WordNet to extract lexical relations between the naming variants in our data set. Unfortunately, this means that we have to exclude a certain portion of the data as either (i) one of the name is not covered in WordNet, (ii) we cannot find a lexical relation between the two names (see below). Also, we had to be relatively permissive with respect to the definition of hyponymy/co-hyponymy. For instance, to analyze *giraffe* as a hyponym of *animal* we have to look at the closure of the hyponyms of *animal* with a depth of 8 (in WordNet).

//sz: include Table that reports counts of the naming variants, coverage in WordNet etc.//

Table 2 shows the distribution of lexical relations for those naming variants that we were able to analyze with WordNet, for types and tokens of naming variants respectively. *//sz: Is it clear what types and tokens mean here? probably not...//*

3.3 Issues with WordNet

Can we tease these types of disagreements apart automatically, using WordNet?

category	most frequent naming variants
food, solid food	sandwich – food (492), hotdog – food (467), hotdog – sandwich (318), cake – food (265), bread – sandwich (257), bread – food (247), bread – bun (206), donut – food (188), hotdog – bun (184), sandwich – burger (163)
structure, construction	house – building (1160), building – house (511), bridge – train (326), bridge – overpass (235), house – window (161), house – home (123), tent – canopy (120), building – castle (101), bridge – building (98), bridge – pole (85)
tool	knife – pizza (27), knife – food (22), knife – cake (17), knife – banana (16), knife – plate (15), knife – apple (13), knife – pocket knife (12), knife – butter knife (12), knife – table (10), knife – carrot (8)
vehicle	airplane – plane (11194), plane – airplane (3829), motorcycle – bike (2624), airplane – jet (1319), boat – ship (1301), truck – car (1095), car – vehicle (874), motorcycle – wheel (861), truck – vehicle (718), truck – wheel (716)
plant, flora, plant life	flower – flowers (29), rose – flower (25), flower – rose (22), flower – vase (18), flowers – flower (14), flower – dandelion (14), plant – flower (12), flower – sunflower (11), plant – wheat (11), flower – flower pot (10)
food, nutrient	pizza – food (1796), sandwich – food (631), pizza – cheese (447), pizza – plate (423), salad – food (402), sandwich – burger (235), pizza – toppings (217), pizza – pizza slice (162), sandwich – bread (156), food – sandwich (147)
tableware	cup – mug (146), mug – cup (104), cup – coffee (93), cup – drink (73), bowl – food (60), cup – glass (39), bowl – cup (36), glass – wine glass (35), glass – cup (34), cup – coffee cup (31)
article of clothing	shirt – t-shirt (2914), jacket – coat (2396), jacket – shirt (1552), jacket – suit (1168), suit – jacket (1029), shirt – jacket (813), shirt – tie (723), shirt – man (487), shirt – dress (462), shirt – sweater (450)
animal	cow – bull (515), sheep – goat (486), cow – animal (445), giraffe – animal (380), bird – parrot (349), sheep – animal (294), sheep – lamb (282), horse – animal (269), cat – animal (237), bird – seagull (231)
ware	fork – spoon (71), fork – plate (44), spoon – food (32), fork – food (31), fork – cake (21), spoon – fork (18), spoon – wooden spoon (15), fork – broccoli (15), fork – silverware (10), spoon – vegetables (8)
person	woman – person (3594), man – person (3546), boy – child (3243), woman – girl (2328), girl – child (1985), woman – tennis player (1277), man – player (1273), man – boy (1214), skateboarder – skater (1194), man – t-shirt (1143)
instrumentality, instrumentation	couch – sofa (4090), desk – table (3435), carpet – floor (1697), bench – chair (1401), desk – keyboard (1380), counter – table (1201), table – desk (1135), counter – countertop (1101), table – counter (906), rug – carpet (895)

Table 3: Most frequent naming variants for each category