WHAT 50 MILLION DRAWINGS CAN TELL US ABOUT SHARED MEANING

ANONYMOUS AUTHOR 1\*1,2 and ANONYMOUS AUTHOR 21

\*Corresponding Author: name@domain.com

1This Department, University X, City, Country

2That Department, University Y, City, Country

A foundational assumption of linguistic communication is that conversants have similar underlying concepts (Brennan & Clark, 1996; Wierzbicka, 2012). On this view, the ability of one person to understand another when she says “the tree” depends on the word activating the same concept in both people. One approach to verifying this assumption is to rely on definitions, but this reasoning is circular—how can we be sure the words in our definitions are the same? Here, we investigate the assumption of shared linguistic concepts by studying concepts represented in the visual modality—drawings—and examining predictors of their variability.

We analyzed a dataset of 50 million drawings (of mostly concrete artifacts such as “tree”) from 15 million participants (QuickDraw:quick-draw.withgoogle.com/data). Although all drawings were elicited in English, the participants spanned the globe and, we can assume, represent a wide variety of cultural and linguistic experience. Such drawings only capture a part of meaning—people know much more about trees than what they look like—and therefore offer a conservative estimate of diversity.

We quantified similarity of drawings in two ways: (1) Hausdorff Distance(HD), which quantifies image similarity as the minimum Euclidean distance be-tween two sets of points (Huttenlocher, Klanderman, & Rucklidge, 1993); (2) the internal weights (layer FC2) for each of our drawings from a neural net model trained on ImageNet (Deng, 2009), with similarity corresponding to the cosine distance (CD) on weights between pairs of images.

Initial analysis included 1500 image pairs of two categories—“bread” and “tree”—from participants located in 72 countries. We validated our similarity measures using human judgements. We selected 20 pairs from each HD decile for each item (Fig. 1), and asked participants to rate the similarity of the objects in the drawings using a 1 (almost identical) to 7 (completely different) Likert scale. Each participant (N= 100) rated 50 pairs of a single item.

Human judgements of similarity were highly correlated with HD (r= .29, p <.0001; Fig. 2) as well as CD (r= -0.20, p <.0001). In a mixed effect model with both HD and CD as fixed effects, the two measures were simultaneously predictive of human similarity judgements (HD:= .35;t= 12.39; CD:= -.26;t= -9.95) and thus appeared to capture different aspects of visual similarity.

With our automated similarity measures validated, we next examined predictors of variability in visual similarity. We hypothesized that speakers who were closer linguistically and geographically would also be closer in their conceptual space, as represented by their drawings. We quantified geographic distance as the distance in meters between the centroid of each pair of countries (GNS Database,2012), and linguistic distance using based on features values from the WALS typological database (Dediu, in press). *[WILL EDIT THIS and CHANGE FIGURE 2; Consistent with our prediction, in an additive model with all three predictors, only socio-political distance was a reliable predictor of ratings (= -.21;t=-2.29):Countries that have less socio-political distance (more events) have more similar drawings.]*

These data reveal systematic cross-cultural variability in semantics, and suggest that speakers’ physical and linguistic proximity lead to the convergence of shared semantics.