

QTUNA: A Corpus for Understanding How Speakers Use Quantification

Guanyi Chen[♣], Kees van Deemter^{♣♥}, Silvia Pagliaro[♣], Louk Smalbil[♣], Chenghua Lin[♣]

[♣]Department of Information and Computing Sciences, Utrecht University

[♥]Department of Computing Science, University of Aberdeen

[♣]Department of Computer Science, University of Sheffield

{g.chen, c.j.vandeemter}@uu.nl, s.pagliaro@students.uu.nl

l.smalbil@students.uu.nl, c.lin@sheffield.ac.uk

A Instruction

Here, we show the instruction for the $n = 4$ experiment, the instructions are the other two experiments are quite similar with minor differences on wording.

Each participant in $n = 4$ experiment will see:

We're interested in understanding how people describe sets of objects. To find out, we're doing a small experiment, in which we'll show you a number of situations in which some objects (circles and squares, as it happens) are displayed on a screen.

We'd like you to describe each situation in one or more grammatically correct English sentences. The experiment should take you little over 15 minutes. Finally we'll ask you what you thought of the experiment (1 minute). Please note:

- 1. Based on your description, a reader will try to "reconstruct" the situation. We use the word "reconstruct" loosely here, because the only thing that matters is the different types of objects that the sheet contains. Therefore, please do not say where in the grid a particular object is located (e.g., "top left", "in the middle", "on the diagonal");*
- 2. Each object is a circle or a square, and either red or blue. Your reader knows this;*
- 3. Please do not "enumerate" the different types of objects. For example, do not say "There is a red circle, two blue circles, and ...";*
- 4. Every situation contain four objects. Your reader knows this in advance, and he/she will take this information into account when interpreting your description;*
- 5. Please describe each situation separately, without referring to other situations.*

Here is an Example of what you will do: [...]

B List of Vague Quantifiers in QTUNA

When annotating the corpus, we treat the following quantifiers we found in the corpus as vague quantifiers: *many, some, a lot of, lots of, most, few, a few, slightly more, slightly more than half, a small amount of, majority, minority, about half, roughly the same amount, almost all, almost half, many more, almost a quarter, and several.*