

World of Quantifiers

Judith Degen Michael Franke

February 5, 2015

General questions: how do quantifiers cover the space of meanings cross-linguistically?
Are any/all systems such that they exhibit an optimal partitioning of the space?

Goals: in analogy to the World Color Survey, to create a database of cross-linguistic quantifier production and interpretation data, freely available and with the potential to be extended to novel languages by researchers who would like to contribute, using a centralized set of stimuli/methods/annotation scheme/analysis scripts

1 The plan

- start with English, then branch out to other languages that one can easily get data on over the web
- conduct initial piloting that will include figuring out exactly which tasks/stimuli to use,
- figure out a general format that the data should be stored in (a csv file of some sort with certain mandatory column names and content presumably); create github repo where everything will be stored, including documentation, instructions, (anonymized) data, analysis scripts

2 Methodological considerations: production

2.1 Goal

Collect data about the quantifiers that people use to describe sets of objects of different cardinalities.

2.2 Method

2.2.1 Procedure

Participants will see an image of a set of objects (e.g. differently colored dots) and a sentence frame “_____ dots are blue” and will be asked to provide at least two possible expressions that can fill the blank. Instructions should emphasize that expressions need not be single words.

2.2.2 Materials

Participants will provide expressions for images of dots in a frame (?), where images differ from each other in two ways: cardinality of the reference set and cardinality of the subset with the property mentioned in the description (e.g. being blue). Some considerations: we probably want multiple reference sets and there should be both even and uneven reference sets. Proposed cardinalities of the reference sets: 5, 20, 100 (?)

Regardless of the actual reference sets we agree on, there are some issues with how to cover the space: for very small reference sets (e.g. of cardinality 5), it will be easy to ask for two descriptions of each of the states $t_0, t_1, t_2, t_3, t_4, t_5$. But for large reference sets (e.g. 100, but even 20), it will be more difficult to show each state, so we'll have to come up with a reasonable way to select images. One possibility is to specify certain proportions we want to cover (e.g. 0, .2, .4, .6, .8, 1) and show just those images (i.e. in the case of 100 dots, 0/20/40/60/80/100 blue dots). This has the advantage that we'll have direct comparability between reference sets, but the disadvantage that we'll never get things like "half" and we'll have gaps in the meaning space that some languages might have special terms for. The other option is to randomly sample cardinalities from pre-specified intervals by participant, e.g. for a size 100 reference set, each participant might get 10 trials, where each trial samples a cardinality from the intervals $[0, 9]$, $[10, 19]$, \dots , $[90, 100]$.

3 Methodological considerations: interpretation

3.1 Goal

Collect data about peoples' interpretation of the top n quantifiers obtained from the production study. (It will be important to determine how to choose this set of quantifiers in practice.)

3.2 Method

3.2.1 Procedure

Good question. Generally: participants will see some utterance with a quantifier, e.g. "Few of the dots are blue". Then we want to assess their interpretation. There are different ways of doing this. Here are some:

- ask them to write down a point estimate of how many dots (out of the reference set cardinalities decided upon for the production experiment) are blue
- give them a slider from 0 - n and show the number they chose (e.g. "42")
- give them a slider from 0 - n and show the number of dots they chose to be blue
- give them n uncolored dots and have them 'build' the interpretation by coloring dots

- give them three (two? four?) different images to choose from (will involve harrowing selection of images)
- give them 5 different images with cardinalities randomly selected from certain intervals (comes with its own set of problems, e.g. you would definitely want to include the n image for “all”)
- sentence verification about various individual images (randomly selected?) for each quantifier (but we don’t know what sentence verification actually measures)

We should talk about which subset of these (and possibly others?) we would want to test.