Method & Results

Xiuyuan (Flora) Zhang
5/8/2018

A Basic Model

In this section, I will talk about our basic model and mathematical dervations that lead us to our current theorectical construct. On a theorectical level, we are interested in learning about how people learn new knowledge about categories that they were just introduced to. If a listener hears something that is unknown to her, what is the next step of cognitive process that would happen for her? We hypothesize that the listener will not completely generate a new understanding of this new category out of nothing, which is against both literatures in developmental psychology and linguistics. Since the language structure human beings have is one that is systematic and have mappings with cognitive concepts and our experiences, we think that a listener, upon hearing a novel category and a familiar feature, will infer a prevalence rate of that novel category given the feature based on her prior experience of the world knowledge that contains sets of entries. This inference is thus conditional on one's prior knwoledge, which is why we think that building a model using Bayes Theorem would be appropriate. A listener will search within her reportoire the set of objects that have a prevalence rate that is of certain threshold θ given the feature. Based on Grice's Maxim of Quality, a listener would assume that if a speaker utters the generic containing a novel generic, it would be a piece of information that is informative and valuable. Thus, as our initial results show, listners would infer the prevalence rate of the novel category and the given feature based on some prior understanding of the comparison set.

While we manipulated the background knowledge in our first experiment, we are aiming at using participants' demographic differences to get at an estimation of prevalence for certain feature and categories that better describe certain probability terms from the equations below.

Using a naive Bayes rule,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

We are interested in learning about $P(Feature)$ and $P(Category)$. $P(Feature)$,the probability of certain feature in the world knowledge can be expressed in the following way, where X is a set of all categories that has this $Feature$ as a charactertistic feature.

$$\begin{aligned} P(Feature) &= \sum_{x_i \in X} P(Feature|X = x_i)\, P(X = x_i) \\ &\sim \sum_{x_j \in X'} P(Feature|X' = x_j)\, P(X' = x_j) \end{aligned}$$

We first make a qualitative prediction that for a listener who hears a generic statement about a novel category and its feature, she would infer that the prevalence rate of this $Feature$ given that novel category as

$$\begin{aligned} P(Feature|Novel\ Category) &\gg P(Feature) \\ &\gtrapprox P(Feature|X = x_i) \end{aligned}$$

where different x_i, considered as comparison sets (background knowledge), are explicitly provided to participants by us during our experiment. Following M.H. Tessler and M.C. Frank, we also hypothesize that a listener would infer that the prevalence rate of this feature given novel category is related and higher than the comparison set they have to compare against this novel generic with.

$$\frac{P(Feature|Category\ x_i)}{\mathbf{E}(\sum_{x_i \in X} P(Feature|X = x_i))} \gg \theta$$

$$\frac{P(Feature|Novel\ Category)}{\mathbf{E}(\sum_{x_i \in X} P(Feature|X = x_i))} \gg \theta$$

where θ is an unknown threshold which the listener thinks that the speaker's surpass for the speaker to utter the generic statement 'Categories are Feature'. At the moment, we are getting $P(Heavy|X = x_i)$ from survey data.

Thus, using the above conditional probability conidtions, we derive a speaker and listener model. In the model below, I have chosen one example from our survey as events for the probablity calculation so that one can get more contextual intuition on how this model works. In both models, we apply a softmax since the output of the softmax function can be used to represent a categorical distribution, which is what we have here, a probability distribution over different possible feature and category sets.

For a speaker,

$$
\begin{aligned}
P(to\ utter\ ''Feps\ are\ friendly'') &= f(P(Friendly|Feps), P(Friendly)) \\
&= \alpha \frac{P(Friendly|Feps)}{P(Friendly|Feps) + P(Friendly)} \\
&= \frac{e^{\alpha P(Friendly|Feps)}}{e^{\alpha P(Friendly|Feps)} + e^{\alpha P(Friendly)}}
\end{aligned}
$$

For a listener,

$$
\begin{aligned}
P(Feps\ are\ friendly|speaker\ uttered\ ''Feps\ are\ friendly'') &= \frac{P(speaker\ uttered\ ''Feps\ are\ friendly''|Feps\ are\ friendly)P(Feps\ are\ friendly)}{P(speaker\ uttered\ ''Feps\ are\ friendly'')} \\
&= \mathbf{E}(P(speaker\ uttered\ ''Feps\ are\ friendly''|Feps\ are\ friendly)P(Feps\ are\ friendly)) \\
&= \mathbf{E}(P(speaker\ uttered\ ''Feps\ are\ friendly''|Feps\ are\ friendly)) \cdot 0.5 \\
&= \mathbf{E}(\int_{P(Friendly|Feps)} \frac{e^{\alpha P(Friendly|Feps)}}{e^{\alpha P(Friendly|Feps)} + e^{\alpha P(Friendly)}} \delta P(Friendly|Feps)) \cdot 0.5
\end{aligned}
$$

where $P(speaker\ uttered\ ''Feps\ are\ friendly'') = 1$ is assumed since, by the time the participant(the listener) does the survey, the event of speaker uttering an utterance already has happened. Further, $P(Feps\ are\ friendly)$ is drawn from uniform distribution, and thus have its expected value to be 0.5.

$$
P(speaker\ uttered\ ''Daxes\ are\ heavy''|Daxes\ are\ heavy) = \int_{P(Heavy|Daxes)} \frac{P(Heavy|Daxes)}{P(Heavy|Daxes) + P(Heavy)} \delta P(Heavy|Daxes)
$$

Method

Hypothesis

In a speaker-listener interaction scenario, where the speaker utters a true generic statement 'C(category) are F(feature),' we hypothesize that, when C is a novel category, the listner uses her own background knowledge to infer the prevalence rate of feature F in category C upon hearing the generic. In our current paradigm, we provide participants a familiar category $C_{familiar}$, $C_{familiar}$ serving as background knowledge.

Conditions

The table below includes our feature selection, its corresponding comparison categories (chosen based on our estimate of their low, medium, and high prevalence rate) and novel categories.

Feature	Alternative Comparison Categories	Novel Category
friendly	Puppies (H), Goats (M), Squirrels (L)	Feps
tasty	Pizzas (H), Fruits (M), Vegetables (L)	Kobas
heavy	Trucks (H), Stones (M), Bikes (L)	Dands

(H = high prevalence, M = medium prevalence, L = low prevalence)

Setup

We run three separate survey studies on Amazon Mechanical Turk. All three surveys provide participants a narrative that introduces them to an imaginary country, Akar. Sample questions from all three surveys are provided in the section below. The first survey recorded participants's evaluation of a given generic statement as True or False as well as their prevalence rate estimates for all abovementioned 9 categories (3 per feature.) The second survey introduced a novel category C_{novel} along with a familiar comparison category $C_{familiar}$, then asked participants to estimate the prevalence rate of the feature F in novel category C_{novel}. The third survey is similar to the second, with the difference that we only stated 'C_{novel} are like $C_{familiar}$.' before asking participants to estimate the prevalance rate of feature F for C_{novel}. This survey serves as a sanity check, checking whether participants treat C_{novel} as equivalent to $C_{familiar}$ and provide a smiliar rather than higher estimate for C_{novel} comparing to estimates for $C_{familiar}$.

Ongoing Experiment

Demographic	Level	Conditions the Respective Level would Estimate Higher (Major Source: Statista.com)
Gender	Female	Like Desserts
Gender	Male	Watch Professional Sports
Age	< mean(age for all participants) , cut off = 60	Like online shopping
Age	>= mean(age for all participants), cut off = 60	Watch nightly news on tv

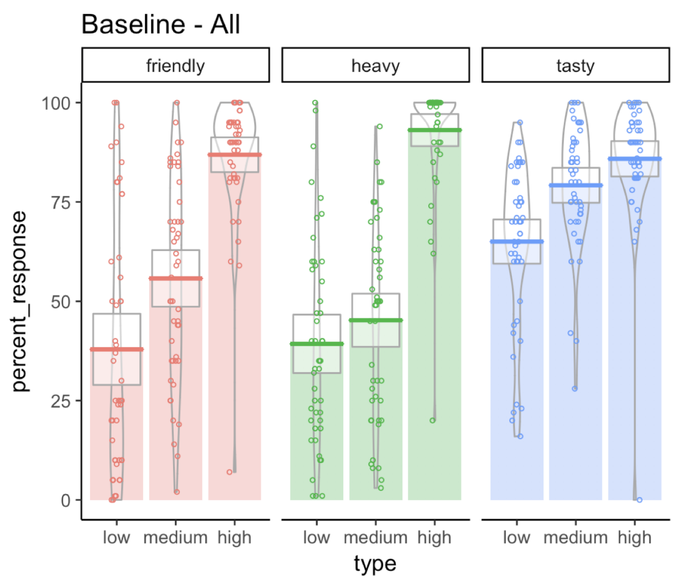
Demographic	Level	Conditions the Respective Level would Estimate Higher (Major Source: Statista.com)
Non	Control	Drive to work
Non	Control	Like the window seat on the plane

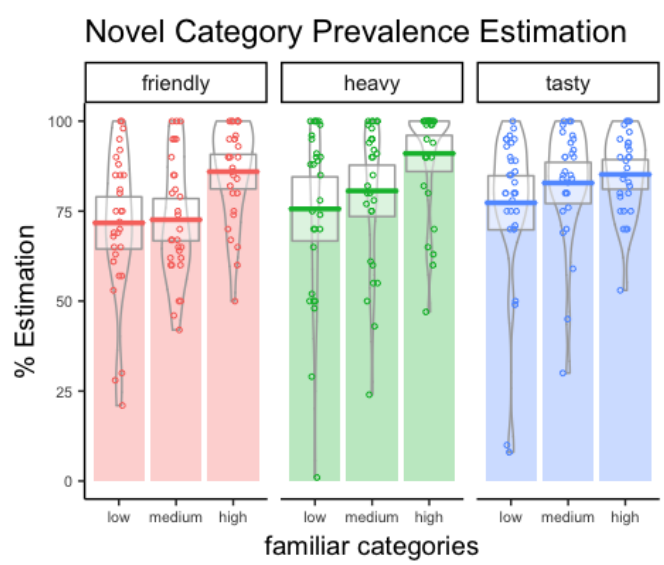
Continuing the generic statement survey project, this time we adjusted our paradigm so that instead of explicitly providing participants alternative sets, we plan to see whether there are some differences in prevalence rate estimations among participants from different demographic backgrounds. In this experiment, we are specifically looking at any potential similarity and differences in response controlling for gender, and age. We are running 200 participants in total on Amazon Mechanical Turk for this experiment.

After running the previous survey, we will conduct another survey asking people to estimate the prevalence rate of the same activities from previous survey among people in their own country(U.S.). The objective of this survey is to provide a baseline measure for comparison between one's estimation of novel generics and familiar generics. We will use Amazon Mechanical Turk to run 200 participants. The format of the survey as well as the questions can be found in the OSF repo.

Initial Results

The two graphs below show a comparison between our baseline prevalence estimation for participants by different features and novel category prevalence estimation for participants. There is a general trend that follows our hypothesis, which is that, provided with a comparison set (our operationalized construct for a listener's background knowledge,) the listener's estimation of a novel category is influenced by this comparison set.

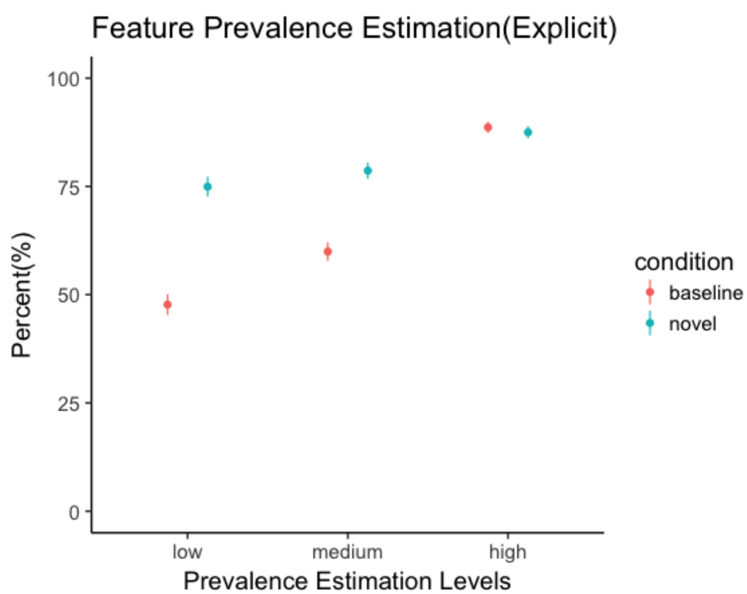




From the table and the graph below, one can see that there is a consistency that confrims our hypothesis that the prevalence rate estimation of a feature and a given novel category is higher than its corresponding baseline familiar category and the same feature. Yet, as one can also see, when level = high, the prevalence estimations are very close to each other. This is not unexpected, since level = high is a condition where people reach more agreement on the high prevalence rate for the given feature and category. In the table below, one can see that we have two major conditions: baseline, novel, and each has three levels: low, medium, and high. This correspond to the first table shown during the Method sections, where, instead of grouping by different features, we are focusing on the effects between different levels for the baseline (background knowledge) and the novel (newly introduced category.) This table provides some summary statistics on the data.

Condition	**Level**	**Mean**	**Standard Error**	**Total Number of Participants**	
1	baseline	low	47.69655	2.399619	145
2	baseline	medium	59.95172	2.163931	145
3	baseline	high	88.64138	1.283504	145
4	novel	low	74.93407	2.310289	91
5	novel	medium	78.62637	1.883457	91
6	novel	high	87.50549	1.395035	91

In the figure below, one can see that first there is a steady increase from level low to medium to high in terms of prevalence rate estimatons for both the baseline and the novel conditions. Without applying our model yet(still collecting more data), we can at least see that there is a strong correlation between one's background knowledge and one's estimation of a novel category based on this background knowledge once the connection is established.



We are still in the process of gathering more data before integrating our empirical data to our theorectical hypothesis. We plan to run logistic regerssion on demographic conditions based on participants' responses as well as a linear regession model that takes in to consideration the possible interactions between the demographic conditions that we have.