

We are excited to chat with you all about grounded and distributional language learning in just a few weeks! This long email gives you some context on what we're hoping to do. It has two parts: 1. **Framing**, and 2. **Homework**.

1. Framing

Our original prompt is below; here's another (very opinionated) way of looking at the issue. Grounded models are beautiful and effective within their trained domain, but sometimes brittle and narrow. In contrast, distributional semantics models have broad coverage but are ungrounded and often relatively unrelated to the world. How can we put these two together? One way of looking at this question is what kind of distributional model is implied by the grounded models? From our (very pragmatics-heavy) perspective, one possible hypothesis is:

Hypothesis: Distributional semantics is what you get when you integrate the grounded use of language in pragmatic context across the distribution of communicative needs.

Can we identify evidence for or against this hypothesis? Can we write down models that instantiate this relationship? Is the hypothesis obviously wrong?

2. Homework

We recognize that everyone on this list is very busy! But the classic failure mode of interdisciplinary seminars is people talking past one another because they don't share basic concepts. Thus, in order to make our working group maximally productive, we would like to ask you to do two things, which should not take more than 2 hours total (including some plane/cab time):

I. Reading. Please familiarize yourself with the three broad traditions/styles of thinking about semantic representations that we are going to be discussing by reading 1-2 papers from each:

- Distributional semantics: [word2vec](#), [skip-thought](#)
- More grounded semantics: [semantic parsing](#), [image captioning](#)
- RSA-style models: [review](#)

We know that some of you are experts on one or more of these traditions. This reading list is so that those of us who haven't mastered all of them can be at least conversant with some main approaches.

II. Contribution. Please send us a one question, proposal, or statement you wish to consider in the context of this discussion, around a paragraph in length. We will collate these and use them as prompts for our discussion. If you would like to give a 10-15 minute talk to expand on your contribution, please include that request. We may not have time for everyone to do talk, though, and talks should be focused on describing

specific work or proposals that fit our question. If you happen to have an artifact that we can interrogate (e.g., a running model that people can poke at or feed examples to, that would be great!).

Please take the short amount of time necessary to accomplish these tasks, as they will make our ~10 hours of budgeted working time together much more productive, and might allow us to move towards the stage of products or proposals rather than simply working to achieve common ground.

We are very much looking forward to seeing you!