Early Language Learning in Social Context

Michael C. Frank

1 Introduction

How do speechless, wordless infants develop so quickly into toddlers who can use language to communicate their needs and their thoughts to others? As this question suggests, language learning is both an important topic in its own right and a case study of the mechanisms of human learning. In addition, the foundations of language are important for many later life outcomes including academic success.

My work focuses on the intersection of language learning and social development. Qualitative, informal descriptions of language acquisition emphasize the social nature of children's input but provide little detail about exactly how social information facilitates learning. In contrast, my research aims to create data-driven models of language learning from early speech segmentation all the way to sophisticated pragmatic inferences in context. I begin from the assumption that language is embedded in a social context and investigate both the nature of this context and how it facilitates the learning process.

My research strategy is to begin by characterizing children's basic learning mechanisms (§2) and then to investigate how these mechanisms are applied in social word learning (§3) and pragmatic reasoning (§4). To address questions about the relationship of language learning to cognition more broadly, I also study the role of language in representing numbers (§5). In each of these areas, I examine theoretical hypotheses using computational models in the Bayesian tradition, which make assumptions explicit and help to summarize complex patterns of data (Frank, 2013, *Cognition*). These computational models are informed by experimental and observational work, using techniques from eye-tracking and head-mounted cameras to massive web-based experiments. The aim of this multi-method approach is to reveal a more detailed, quantitative picture of children's early language learning.

2 Characterizing statistical learning

My first goal is to understand the abilities (often known collectively as "statistical learning") that children bring to bear on the problems of early language learning. To characterize these abilities, we evaluate novel and pre-existing computational models on large experimental datasets. Contrary to expectations, the models that fit human performance best are not the ones that compute specific statistics like transitional probability. Instead, they are based on the chunking mechanisms that help overcome capacity limitations in human memory (Frank et al., 2010, *Cognition*; Kurumada et al., 2013, *Cognition*).

Examining restrictions on memory can also help us understand the generalization ("rule learning") abilities displayed by human infants and adults (Frank et al., 2011, *LL&D*). Learners in these experiments recover the regularities that best compress the input data (Frank & Tenenbaum, 2011, *Cognition*). Across statistical and rule learning, our work in this area shows the commonalities between disparate phenomena: All learning is "statistical," and sophisticated inferences often emerge from simple, general memory mechanisms.

3 Word learning in social and discourse context

How do these inferential abilities help learners link words to their meanings in rich, complex environments? Our work suggests that statistical learning abilities play a role, but that they operate over the output of social and communicative processes, rather than over

the raw perceptual signal. In our initial investigation, we modeled word learning as a joint inference at two timescales: about speakers' communicative intentions in each individual moment and about the stable, long-term correspondences between words and their meanings. A model that performed this joint inference fit developmental data better and learned words more accurately than models that only looked for associative correspondences (Frank et al., 2009, *Psych Sci*). Our followup work pursues three primary questions raised by this result: (1) what social cues signal communicative intentions, (2) how do children use these cues, and (3) how do children learn from incomplete or noisy social cue information?

Although understanding speakers' communicative intentions is a critical part of language use and language learning, no one perceptual cue always reveals those intentions. Individual cues on their own—even eye-gaze and pointing—are weak indicators of reference (Frank et al., 2013, *LL&D*). Pooling social information across many cues together can contribute to word learning performance, however (Johnson et al., *ACL*). To better understand the structure of cue use "in the wild," in our ongoing work we are using wearable, head-mounted cameras in combination with tools from computer vision (Frank et al, 2013, *Proc Cog Sci*).

In order to characterize children's uptake of social cues during word learning, we have also developed eye-tracking methods for analyzing looking at complex social scenes. Our initial work focused on infants' social attention (Frank et al., 2009, *Cognition*; Frank et al., 2012, *Infancy*). More recently we have measured both children's sensitivity to different social cues and their learning of words that are disambiguated using these cues. Just as young children must build skill in processing spoken language in real time, our results suggest that they also must become proficient at allocating their social attention; this proficiency is a strong predictor of word learning (Yurovsky et al., 2013, *Proc Cog Sci*; in prep). Work in progress applies these methods to children with autism, a disorder which provides both an important application area and a valuable case study of the consequences of social impairment for language learning.

Finally, given that social cues are individually noisy, children may learn by aggregating information across broader discourses (sets of utterances with the same topic). In an ultra-dense corpus describing a single child's experience, words that appeared in consistent topics were learned earlier (Roy, Frank, & Roy, 2012, *Proc Cog Sci*; in prep). And an unsupervised learning model that assumed topic continuity across utterances learned word meanings better as a consequence (Luong et al., in press, *TACL*). These studies together make a strong prediction, confirmed in a recent experiment: When novel labels are embedded in a coherent discourse context, young children are able to able to learn words even if social cues like pointing are not coincident with labels (Horowitz & Frank, 2013, *Proc Cog Sci*; under review).

Taken together, this body of work describes a synthetic view of word learning as statistical learning over communicative contexts and provides new tools to measure both children's social input and the mechanisms by which they use this input to learn.

4 Pragmatic reasoning and learning from others

If word learning takes place through communication, then it is critical to understand the processes underlying language use in context. We focus on pragmatic inferences about meaning, considering how the words speakers say relate to their underlying intended meaning. The aims of our work in this area are (1) to construct quantitative theories of pragmatic reasoning and (2) to understand how children's pragmatic abilities can help them learn.

Although linguistic pragmatics has primarily been characterized using informal, verbal theories, we have recently created a probabilistic model of general pragmatic reasoning that provides a precise quantitative fit to adults' judgments (Frank & Goodman, 2012, *Science*). Unlike verbal theories, our model explains the way linguistic inferences integrate prior knowledge about the world (Stiller, Goodman, & Frank, 2011, *Proc Cog Sci*). In addition, an iterated pragmatics model can help describe how communicators converge on a shared vocabulary to talk about new situations (Smith, Goodman, & Frank, under review).

Pragmatic reasoning is a valuable method for learning, both about language and about the world (Shafto, Goodman, & Frank, 2012, *Pers Psych Sci*). Congruent with our model, both adults and preschoolers can infer the meaning of an ambiguous word by presuming the speaker has a pragmatic goal of being informative in context (Frank & Goodman, under review). And if learners already know word meanings, they can even make inferences about the structure of the world beyond language (Horowitz & Frank, 2012, *Proc Cog Sci*).

Overall, this work provides a formal account of pragmatic reasoning as rational inference in a social context, and—like the work on social word learning described above—points the way forward towards integrating the joint tasks of communication and language learning.

5 Language and cognition: The case of number

What are the cognitive consequences of learning a language? Previous work on so-called "Whorfian" effects—effects of cross-linguistic differences on cognition—produced small but reliable findings. In the domain of color, for example, knowing a language with different words for "light blue" and "dark blue" slightly facilitates discrimination of these colors (Winawer er al., 2007, *PNAS*). Exact number is a case where language does not just facilitate an ability that is already present, however: It plays a critical role. The Pirahã, an Amazonian indigenous group, have no words for number and instead rely on their approximate number system for manipulating large quantities (Frank et al., 2008, *Cognition*), and when MIT undergraduates are prevented from using numbers via verbal interference, their matching performance looks similar to that of the Pirahã (Frank et al., 2012, *Cognit Psych*).

But language is not the only method for mental representation of exact quantities. "Mental abacus" (MA) is a technique in which users visualize a picture of an abacus. Unlike standard mental math, MA computations rely on representations in visuospatial working memory, not on language (Frank & Barner, 2012, *JEP:General*), and learning MA representations helps improve children's mathematics skills, especially for children with the working memory resources to use it appropriately (Frank et al., in prep). This work suggests that language can be a powerful "cognitive technology" for using our limited mental resources to represent complex concepts, but it is not the only method for doing so.

6 Conclusion

Language learning is fundamentally tied to social context. My research program aims to create a computational characterization of children's basic learning abilities and how these abilities intersect with the social context of learning. Running throughout this work is an emphasis on bringing large-scale data sources to bear on difficult empirical problems. Although this work addresses fundamental theoretical questions in psychology and psycholinguistics, it also has substantial application potential in fields from autism research to natural language processing. By better characterizing the social context of language learning, we will be better positioned to understand both our learning abilities and language itself.