Automatically identify communicative acts in parent–child conversations

Anonymous CogSci Submission

Abstract

Over their first years of life, children learn not just the words of their native language, but how to use them to communicate. Becauese manual annotation of communicative acts does not scale to large corpora, our understanding of communicative development is limited to case studies of a few children at a few time points. We present a method for automatic identification of communicative acts using a Hidden Markov Topic Model, and apply it to all the entire CHILDES database. We first describe qualitative changes in parent-child interactions over development, and then use our method to demonstrate two large-scale features of communicative development: (1) children develop an adult-like repertoire of communicative acts months before they develop an adult-like diversity of words, and (2) children's mastery of communicative acts co-incides with the point of highest predictability between parents acts and children's, suggesting that structured interactions may play an important role in language development.

Keywords: language input, language acquisition, child-directed speech, corpus analysis

From their first utterances, children are not just producing language but *using* it to communicate. A child who can produce only one-word utterances can nonetheless convey several communicative intentions: using variations in pitch, she can use the word *mama* to identify a person, question possession of an object, or to call for someone's presence (Dore, 1975). From 14 to 30 months of age, children quickly branch out from communicative acts like *requesting*, *protesting*, and *marking an event* to *agreeing to an action*, *stating intent*, and asking and answering a variety of questions (Snow, Pan, Imbens-Bailey, & Herman, 1996). These close studies of children's conversations, using rich observations to infer intended meaning in context, show that much of the action of language acquisition happens at the level of what children *mean* to say.

SOMETHING ABOUT HOW WE HAVE LOTS OF MODELS OF LANGUAGE DEVELOPMENT BUT BASICALLY NONE ABOUT COMMUNICATIVE DEVELOPMENT. TO BUILD THESE MODELS, WE NEED TO HAVE SOME DATA TO EXPLAIN!

Describing children's communicative acts on a larger scale, however, is a challenging task. Without nuanced, context-sensitive human coding, communicative acts can be hard to identify. Words are amenable to identification, storage, and tabulation using common computational tools; perhaps due to their ease of use, models of language development have often approached language development at the level of words (i.e., vocabulary learning). The goals and intentions underlying those words are less amenable to such manipulation. We present an approach to modeling the children's communicative acts, working backwards from the words they produce: we model communicative acts as the latent sources from which words emerge, and characterize children's engagement in these acts across development.

Traditionally, studies of communicative acts among children have brought frameworks from adult communication, such as Speech Act theory (Austin, 1962; Searle & Searle, 1969) and Conversation Analysis (Sacks, Schegloff, & Jefferson, 1974), to bear on children's conversations. While these systems can be useful to characterize children's conversations, young children may not have the same communicative needs as adults—-these may not be relevant distinctions in children's communication. Studying children's communicative acts using a computational model allows us to extract communicative patterns across many children with less a priori specification of what those patterns are.

Here, we characterize children's growing repertoire of communicative acts using a Hidden Topic Markov Model. This model observes utterances produced by parents and children and attempts to infer common underlying processes—-topics-—that produced them. We first show that this model extracts several communicative acts analogous to those observed in close case studies of children's communication without specifying them top-down. We then show that use of these acts has a developmental trajectory in line with those studies: children's act usage quickly proliferates, and has parent-like diversity by XX months. Further, these acts have distinct developmental trajectories that are in line with prior research. Finally, we show that parents and children engage in these acts contingently within conversations, and that this sequential contingency seems to peak at a fairly young age, when children are engaging in predictable exchanges and social routines.

Method

Corpus

We used transcripts of conversations from the Child Language Data Exchange System (CHILDES), a database of child conversation corpora (MacWhinney, 2000). These corpora predominantly record spontaneous conversations be-

tween children and their family members, often in the home. We used transcripts from the North American English collection of CHILDES among children 6 months to 60 months old, and filtered these transcripts to include only utterances spoken by the target child or their parents. Overall, our training data included XXX transcripts from XXX children.

Model

We used a Hidden Topic Markov Model (HTMM) to extract communicative modes from parent—child conversations (Gruber, Weiss, & Rosen-Zvi, 2007). Topic models represent documents as mixtures of topics, and topics as mixtures of words. For instance, a simple topic model trained on news articles may extract a topic whose distinctive words are "fire," "flood," and "aid" and another whose distinctive words are "speech," "legislation," and "administration." Based on its distribution of words, an article about politicians' provision of disaster relief may be correctly inferred to feature these two topics, among others. Intuitively, the goal of a topic model is to recover the underlying sources—topics—from which the words in a document spring.

In contrast to a standard topic model, the Hidden Topic Markov Model takes into account the sequential utterance structure of a document, not just its static distribution of words. The HTMM assumes that words within an utterance are of the same topic, and that sequential utterances may be more likely to be of the same topic. It represents topic transitions between utterances in a coarse-grained way: either switch or stay. Gruber, Weiss, & Rosen-Zvi (2007) developed this model and used it to segment machine learning conference papers, showing that the model can distinguish instances of the word "support" in mathematical contexts (describing support vectors) from those in the context of acknowledgments.

We trained the HTMM on all the utterances in our corpus. Some markers for unintelligible or non-word speech were removed; when this resulted in empty utterances, a 'non-word utterance' token was included to preserve the temporal structure of the dialogue. Typically, function words are removed from corpora before training topic models to aid detection of thematic content. Here, we aim to classify communicative modes underlying utterances rather than thematic topics. We expected function words to be diagnostic of these modes, so included them in our training data.

Topic models require pre-specification of the number of topics. To determine the right number of topics, we trained the model several times with different numbers of topics— 5 to 30 topics, in intervals of 5—with Dirichlet parameters of $alpha = 1/k$, where k is the number of topics, and $beta = 0.01$. Each model produced a sequence of the most likely topic assigned to each utterance. Our selection metric was the proportion of other-topic transitions in this sequence: since we aimed to characterize the temporal structure of topic transitions, we wanted to choose a model that had many of transitions between topics rather than long stretches of utterances all assigned to the same topic. However, increasing the

number of topics almost necessarily increases the number of other-topic transitions, and may make results harder to interpret as topics proliferate; therefore, we balanced the proportion of other-topic transitions against number of topics. Plotting this proportion across number of topics, we judged 15 topics as an inflection point after which increasing the number of topics had diminishing effect on other-topic transitions; we therefore chose the 15-topic model.

Results and Discussion

We begin by we show that the model captures some aspects of communicative acts and explore the static distribution of these utterance types. We then examine trajectories of topic use across development among parents and children. Finally, we examine the temporal dynamics of topic use within discourse.

Topics and their static structure

After training, the model produces a set of topics with associated probability distributions over words. One can conceive of these topics as bags of words, in which some words will be highly likely to be produced and others will be unlikely to be produced. We extracted the most distinctive words from each topic by taking the difference between the likelihood of each word in a given topic and its average likelihood across all topics (Table 1).

Using this probability distribution of words within topics, the model also produces a probability distribution over topics for each utterance in the corpus. Note that this assignment of topic probabilities happens at the utterance level, but temporal structure between utterances is not taken into account. Since we aim to characterize communicative acts and not thematic content or conversational topics, the label 'topic' for these types can be misleading; from here on, we will refer to these types as communicative acts rather than topics.

The most distinctive words of each communicative act in the model, as measured by difference between a word's probability within a type of act and its average probability across all act types, are shown in Table 1. Although act type assignments are probabilistic, we examined these these most distinctive words and looking at utterances in each act type to produce subjective labels that capture the gist of each automatically identified topic for clarity of exposition.

The conversational turns it produces are reasonable! See Table 1

Use of communicative acts over development

Children's use of communicative acts changes drastically as they grow (Figure 2). Early on, children's utterances are predominantly backchannels, interjections, and affirmations or negations (e.g., "yeah," "no," "uh huh"). By around 20 months of age, social routines such as greetings and naming family members (e.g., "hi," "bye," "mommy") displace backchannels to become a substantial part of children's repertoire. Around the same age, an increasing proportion of children's utterances direct attention to or label things in the en-

	1	mental states:	{I, you, what, know, don't, it, do, that, did, think, like}
	2	labeling:	{a, that, what's, is, this, that's, what, the, look, who's, baby}
	3	counting:	{two, one, dis, three, four, de, dat, five, duh, six, eight}
	4	evaluation:	{not, it's, good, i'm, you're, a, that's, very, be, he, he's}
	5	proposed actions:	{you, it, me, to, can, want, okay, I, up, get, come}
	6	non-present events:	{to, we, go, and, you, going, did, when, at, school, the}
	7	miscellaneous:	{of, they, we, have, like, the, i, and, them, are, all}
	8	requests:	{some, you, want, i, more, have, eat, milk, juice, your, drink}
	9	knowledge questions:	{you, what, are, do, look, see, with, at, play, those, doing}
	10	location:	{the, in, there, put, here, it, go, on, right, this, over}
	11	social routines:	{mommy, daddy, hi, down, baby, bye, ball, n, tape, where's, where}
	12	backchannels:	{yeah, no, oh, okay, hm, uh, huh, ah, mhm, mm, yes}
	13	description:	{a, one, this, is, that's, blue, red, green, make, big, yellow}
	14	storytelling:	{the, and, he, was, his, said, in, she, of, to, they}
	15	body routines:	{your, on, my, put, his, her, hair, off, in, head, mouth}

Figure 1: The most distinctive words in each topic. Distinctiveness is measured by the difference between a word's likelihood in the given topic and its average likelihood across all topics.

act	23 months	39 months	57 months
5	C: put back	P: i'm going to turn off the tape recorder	C: i'm gonna fold this
12	P: oh thank_you	C: no no	P: okay
4	P: that is so nice	P: why not	P: why are you folding it
5	P: thank_you for putting that back	C: cause turn it on so i could reach	C: so it will fit me

Table 1: Three examples of the same communicative act sequence (5, 12, 4, 5) in conversations from different children at different ages. Though they involve different topical content, they follow a similar communicative pattern: a suggestion or request regarding action and location; an affirmative or negative response; an evaluative statement or question; and another suggestion regarding action and location.

vironment (e.g., "what's that?" "there's a..."). Parents also increasingly use these labeling utterances when their children are around 15-30 months old; after about 30 months, these utterances diminish in both parent and child speech.

Across our observed age range, both parents and children consistently increase the proportion of their statements about mental states and testimony (e.g., "know," "think," "said"). This is consistent with prior work showing an increase in mental state talk over the 3- to 5-year-old age range, which is correlated with theory of mind task performance (Hughes & Dunn, 1998), de-villiers2000]. It may also reflect growing mastery of the complex syntax required to produce embedded constructions about mental states (e.g., "She thinks it will rain.").

One way to characterize the diversity of communicative acts a person engages in is to measure the entropy of their communicative acts. Overall, parents' communicative acts had higher entropy than children's. Children's communicative act entropy increased drastically over development between 6 and around 24 months, and remained relatively stable across the rest of our age range, as shown in Figure 3.

Our automatically-identified communicative acts are therefore capturing some capacity that becomes more adult-like across development. [more description / perhaps some stats] [this paragraph could be moved to part 2.]

Beyond becoming more complex across development, the ability to engage in a variety of communicative acts might reasonably be expected to correlate with other measures of language ability. Indeed, Snow, Pan, Imbens-Bailey, & Herman (1996) found that the number of speech act types children use correlates highly with the number of word types they use, but does not consistently correlate with mean length of utterance (MLU). We replicate these findings here: children's number of communicative act types produced correlate with their word types produced (*stats*) and does not consistently correlate with MLU (*stats*). Children's diversity of communicative acts also correlates with their parents' diversity of communicative acts, even when controlling for age (*stats*)

Notably, children at the lowest end of our age range start with at least some representation of several communicative
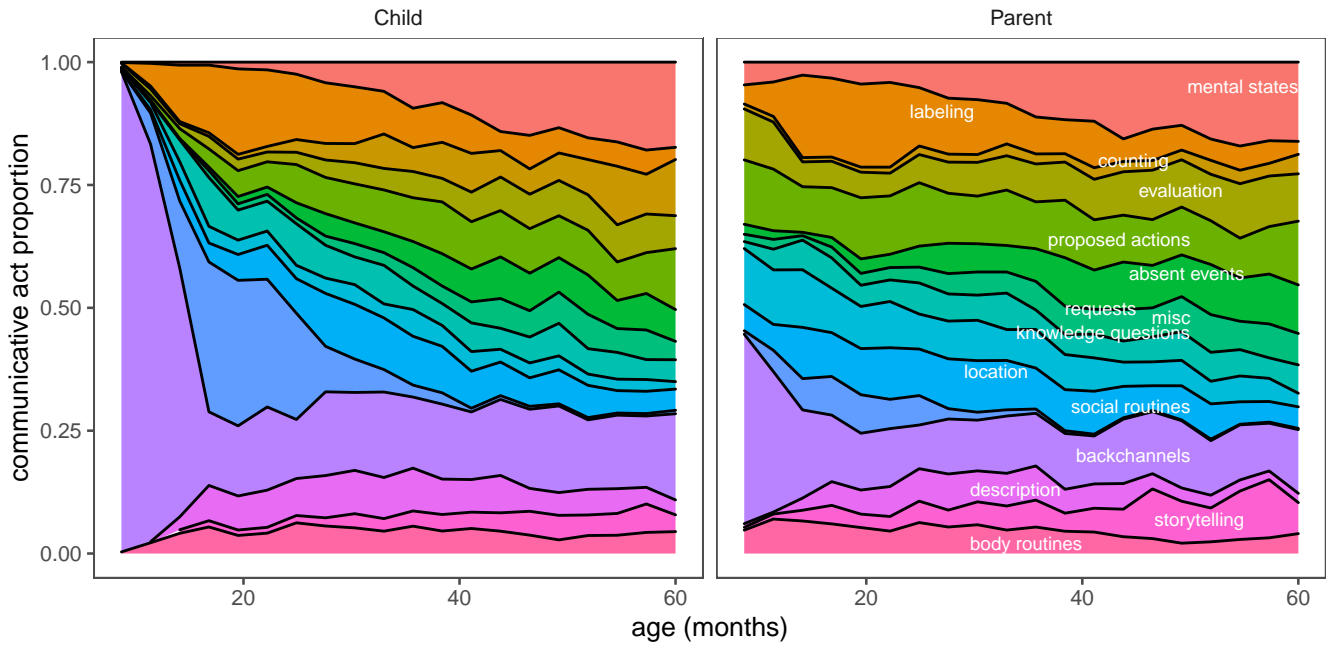
Figure 2: The proportion of communicative acts of each type produced by children and parents across development.
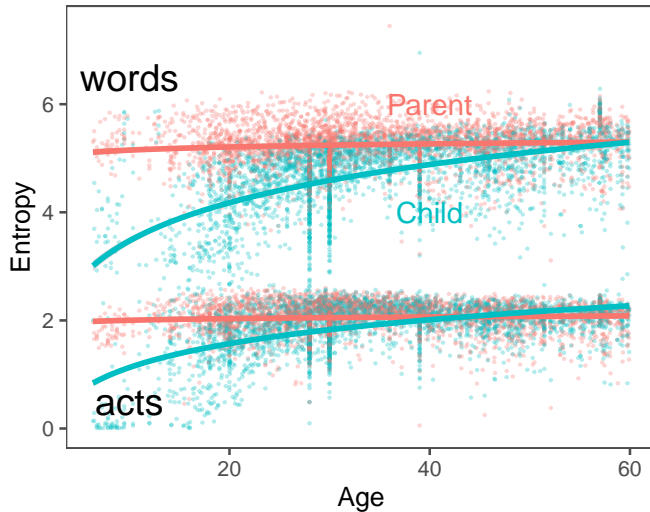

Figure 3: Entropy of topics produced by children, mothers and fathers over development.

act types, and expand on their repertoire rapidly. They also reach adult-like levels of diversity in communicative act production quickly, by around 36 months—before they reach comparably adult-like levels of vocabulary production, as shown in Figure 3.

Parents' and children's communicative usage patterns echo one another somewhat [. . .] [Potential alignment at young ages etc . . .]

Part 3: Dynamics of communicative acts in conversation

Over development, children gain the ability to engage in more structured dialogue, eventually becoming able to respond appropriately and contingently to their conversational partner [cite??]. To characterize the temporal dynamics of communicative acts in conversation, we turn our attention to the transitions between utterances in parent-child conversations. One way to measure how contingent parents' and children's utterances are on each other is to calculate their mutual information, a measure of how much knowing one piece of information (a prior utterance) reduces uncertainty about another (the current utterance). However, discourse with adult-like structure may not maximize mutual information: because discourse has recursive structure, raising topics and closing them in an embedded way, it may not be highly regular at the level of utterance-to-utterance transitions [cite??]. Mutual information between sequential utterances is maximized when utterances are highly predictable based only on the utterance prior, and can be dampened either because contingency is weak or because higher-order structure is at play.

To calculate the mutual information between parents' and children's utterances, we use the full distribution of communicative act probabilities for each utterance produced by the model. These probabilities are averaged within turns; that is, if a child produced four utterances in a row, the probabilities are averaged to produce one probability per act type per turn. We then compute conditional probabilities for each current act type given each prior act type. These conditional probabilities are used to calculate mutual information. We calculate mutual information values for the parent's utterance given the child's utterance, the child's utterance given the par-
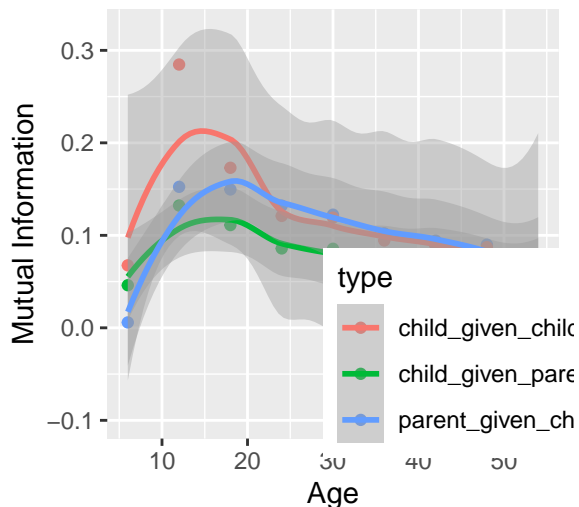
Figure 4: Mutual information between parents and children's communicative acts over development.

ent's utterance, and the child's utterance given their own prior utterance (their own last turn).

When children are young, mutual information is highest between the child's current utterance and their own last utterance [4]. One way to interpret this is that the child is on their own conversational track, steered less by what their conversational partner just said than what they themselves said prior. By around 24 months, this tendency lessens considerably, but children remain more predictable based on their own prior utterances than their parents' utterances across our age range.

Interestingly, there seems to be an increase in mutual information in all three of our measures—parents given children, children given parents, and children given children—from the beginning of our age range peaking at about 14-20 months, after which mutual information declines. Though this is an exploratory finding, it points to the interesting possibility that parents and children interact most contingently when children are still quite young. That is, at the youngest ages, there may be low mutual information between parents and children's sequential utterances because of disorganized interaction; mutual information may then increase as sequences become more orderly, but discourse is not necessarily structured on longer timescales; and mutual information may fall again as discourse structure emerges that is not captured in sequential utterance transitions.

Discussion

In this paper, we present one approach to characterizing children's communicative acts on a large scale. In doing so, we gain the ability to examine the communication of more children in more contexts and across a wider age range than afforded by hand coding, and are able to examine patterns of usage that only become clear across such a wide range of data.

We also lose nuance in the conversational context and nonverbal aspects of communication. Children can achieve com-

municative goals even before they can use language to do so: they can use gestures and vocalizations both to request a desired object and to call a person's attention to something in the environment (Bates, Camaioni, & Volterra, 1975). Studies of children's early one-word utterances demonstrate that they can use the same word to fulfill multiple communicative goals (Dore, 1975).

Nevertheless, the communicative acts extracted by the model show regularities at multiple levels that suggest they reflect communicative function. Children start off with a few act types at the beginning of our age range, and quickly branch out into others as they grow. Different act types show different developmental trajectories: children produce lots of interjections and backchannels early on, soon extending into social routines; parents and children increasingly engage in directing attention and labeling when children are 2 to 3 years old, and less so thereafter; talk about mental states and testimony increases among children and parents consistently across our age range. Children's diversity of communicative acts is correlated with their vocabulary, and with their parent's diversity of communicative acts. Finally, parents and children produce these acts contingently, depending on the prior utterance, and our findings suggest a period of heightened contingency before children reach an adult-like diversity of these acts. Overall, our approach offers a way to take advantage of large corpora of child conversations in the study of children's use of communicative acts across development.

References

10 Austin, J. L. (1962). *How to do things with words*. Cambridge: Harvard University Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The Acquisition of Performatives Prior to Speech. *Merrill-Palmer Quarterly of Behavior and Development*, *21*(3), 205–226.

Dore, J. (1975). Holophrases, speech acts and language universals*. *Journal of Child Language*, *2*(1), 21–40. http://doi.org/10.1017/S0305000900000878

Gruber, A., Weiss, Y., & Rosen-Zvi, M. (2007). Hidden Topic Markov Models. In *Artificial Intelligence and Statistics* (pp. 163–170). PMLR. Retrieved from http://proceedings.mlr.press/v2/gruber07a.html

Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology*, *34*(5), 1026.

MacWhinney, B. (2000). *The CHILDES Project: The database*. Psychology Press.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735. http://doi.org/10.2307/412243

Searle, J. R., & Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.

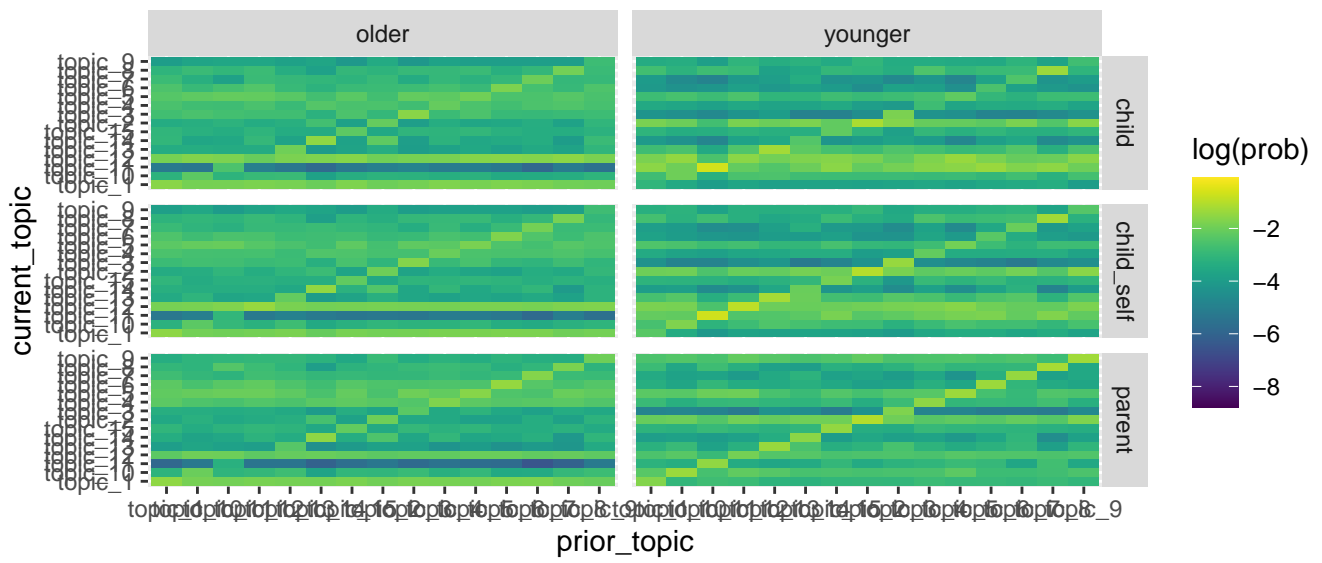Snow, C. E., Pan, B. A., Imbens-Bailey, A., & Herman, J. (1996). Learning How to Say What One Means: A Longi-

Figure 5: (A)

tudinal Study of Children's Speech Act Use*. *Social De-velopment*, *5*(1), 56–84. http://doi.org/https://doi.org/10.1111/j.1467-9507.1996.tb00072.x