SEMANTIC APPROXIMATION AND ITS EFFECT ON THE DEVELOPMENT OF LEXICAL CONVENTIONS

BILL NOBLE, RAQUEL FERNÁNDEZ

Institute for Logic, Language and Computation
University of Amsterdam

We define a signaling games setting for investigating how short- and long-term conventions are established in a community of interacting speakers. Using simulations, we model a particular type of non-literal use of linguistic expressions, *semantic approximation*, and investigate its effects on lexical alignment, ambiguity, polysemy, and communicative success. Critically, in our approach agents do not only keep track of a lexicon reflecting conventions at the level of the community, but also of a *discourse lexicon* that stores information agreed upon by the participants in a specific dialogue. We find that semantic approximation creates opportunities for discourse-level lexicalization, which boosts the expected utility of the discourse lexicon, and that it can have a profound effect on the evolution of community-level lexical resources.

1. Introduction

One of the most striking features of human language is its flexibility—despite the presence of ambiguity and possibly mis-aligned linguistic resources, speakers manage (most of the time) to communicate effectively with each other. We argue that the use of *semantic approximation* is a key mechanism for achieving communicative success. Notable examples of semantic approximation are found in the context of language acquisition, as illustrated in (1) from Chaix, Barry, and Duvignau (2012), where young children extend the meaning of a known-to-them term (in this case the verbs *'undress'* and *'turn on'*) to approximate the concept they want to express, for which presumably they lack a more suitable word:

(1) a. Undress the potato? [*age 2; context: mother peels a potato*]
     b. Go on mum, turn on your eyes [*age 3; context: mother has her eyes closed*]

Semantic approximation, however, is also common in regular language use amongst adults, often signalled by hedging (e.g., *'kind of'*, *'-ish'*). For instance, consider the following dialogue excerpt from Brennan and Clark (1996):

(2) A: A docksider.
     B: A what? [...] Is that a kind of dog?
     A: No, it's a kind of um leather shoe, kinda pennyloafer.
     B: Okay, okay, got it.

In (2), after having used an unsuccessful term (*'docksider'*) to jointly refer to a particular object, speaker A proposes an alternative expression via semantic approximation (*'kinda pennyloafer'*). From then onwards in the conversation, these speakers successfully use the (unhedged) term *'pennyloafer'*.

This illustrates two key features of language use. First, by resorting to semantic approximation (rather than to coining a completely new expression), speakers exploit the potential for polysemy present in natural language, which arguably can have a beneficial effect on efficient communication (Piantadosi, Tily, & Gibson, 2012; Juba, Kalai, Khanna, & Sudan, 2011). Second, speakers can make use of semantic approximation to establish *ad-hoc* conventions over the course of a dialogue. These factors raise several intriguing questions, which we set out to investigate in this paper. How do novel conventions emerge in one-on-one conversation? What effect do semantic approximation and ambiguity have on the lexicon of interacting agents? To what extent does local interaction shape the lexicon of a linguistic community?

2. Formal Model

We use signaling games (Lewis, 1969) to model semantic approximation and investigate its effect on the communicative success of a linguistic community. Our model consists of a community of agents, A, who share a set of atomic expressions (or messages) M, and a conceptual domain C. A similarity measure is defined over pairs of concepts in C. A simulation consists of a series of dialogues between pairs of agents in A. In a dialogue, the interlocutors take turns producing and interpreting single-expression utterances from M. For each utterance, the speaker chooses an expression based on a *speaker intention* modeled as a concept chosen at random from C. The addressee then guesses the speaker's intention, selecting an interpretation for the uttered expression from C as well. Communicative success is determined by the degree of similarity between the speaker's intention and the addressee's interpretation.

Lexicons. A lexicon is a linguistic resource that agents use to keep track of the agreed-upon meanings of words. In natural language there is no centralized objective lexicon. Instead, a lexicon is a subjective representation of joint information—it is what each agent takes to be the lexical common ground (Stalnaker, 1978). We follow Clark (1996) in the distinction between different types of common ground: Personal common ground is built up between individuals based on shared experience, whereas communal common ground holds based simply on mutual community membership. In this model, the two kinds of common ground correspond to two different lexicons that agents $a \in A$ keep: the community lexicon (CL_a) and the discourse lexicon (DL_a). The discourse lexicon is based on the community lexicon, but also incorporates lexical information agreed upon by the participants of a particular dialogue (i.e., the interlocutors' personal common
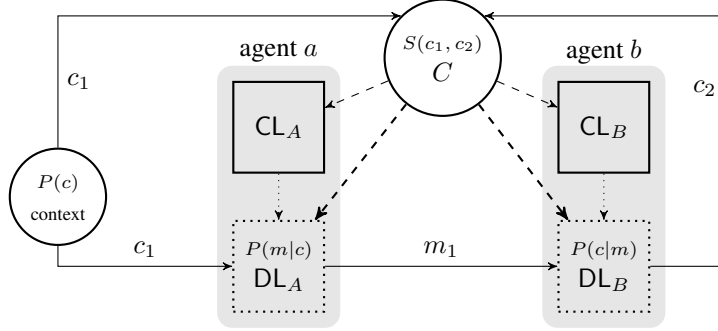
Figure 1. At the beginning of a dialogue, agents construct a discourse lexicon based on their representations of the community lexicon (dotted arrows). In this diagram, agent a expresses c_1 to b using expression m_1, and agent b interprets m_1 as c_2. The interlocutors update their lexicons (dashed arrows) based on their communicative success as determined by the similarity of c_1 and c_2. The discourse lexicons are more radically influenced by this update.

ground based on the discourse itself as a shared experience).

For each lexicon (C|D)L (community or discourse), agents $a \in A$ keep a $|C| \times |M|$ matrix L_a, where $L_a[c, m] \in \mathbb{R}^+$ represents the sum of a's evidence that it is common ground that m means c. From this matrix, agents may derive conditional probability distributions L_a^p over messages given concepts (for production) and L_a^i over concepts given messages (for interpretation).[a] An agent's discourse lexicon is initially just a normalized copy of her community lexicon—the initial discourse lexicon replicates the community lexicon in its semantic content, but disregards the magnitude of evidence from previous discourses.

After each turn, DL_a and CL_a are updated in parallel by Roth-Erev reinforcement learning (Erev & Roth, 1998) according to the production/interpretation choices made by a and how successful the communication was. Although the same learning scheme is applied to both lexicons, the discourse lexicon is more sensitive to these updates since the magnitude of evidence for each concept-expression pair tends to be smaller. The evolution of the discourse lexicon is thus sensitive to the communicative successes of a *particular* dialogue, while the accumulation of utility across dialogues causes the community lexicon to become more stable with each conversation. In our simulations we compare linguistic communities that make use of a discourse lexicon to those that only keep track of a community lexicon.

---

[a]Conditional probabilities are represented by row-normalized matrices. L_a^p is a $|C| \times |M|$ matrix where row $L_a^p[c]$ is the distribution of messages given concept c. Likewise, $L_a^i[m]$, a $|M| \times |C|$ matrix, is a distribution of interpretations (concepts) given message m

Alignment. Alignment measures how much agents' representations of a lexicon agree (i.e., the degree of actual, rather than subjective, lexical common ground). For a set of agents $A' \subseteq A$ and lexicon $(C|D)L$, alignment is defined by the Jenson-Shannon divergence (JSD) of the agents' respective representations:[b]

$$\text{Alignment}(A') = 1 - \underset{a \in A'}{\text{JSD}}(\mathsf{L}_a) = 1 - H\Big[\sum_{a \in A'} \frac{\mathsf{L}_a}{|A'|}\Big] + \frac{1}{A'} \sum_{a \in A'} H[\mathsf{L}_a] \quad [1]$$

Under this definition, a community in perfect semantic alignment (i.e., every agent's representation of the given lexicon is identical) has an alignment of 1. Alignment is 0 when no two agents ever interpret any expression in the same way.

Ambiguity and Polysemy. In our model, the meaning of an expression m in a lexicon $(C|D)L_a$ is a discrete probability distribution over concepts; that is $\mathsf{L}_a^i[m]$. We define ambiguity of m in L_a as the probability that two independent draws from $\mathsf{L}_a^i[m]$ will result in distinct concepts:

$$\text{Ambiguity}(m, \mathsf{L}_a) = 1 - \sum_{c \in C} \mathsf{L}_a^i[m, c]^2 \quad [2]$$

The ambiguity of an agent's lexicon is defined as the average ambiguity of its expressions. The ambiguity of a lexicon as an abstract joint entity is simply the average ambiguity of all the relevant agents' representations of the lexicon.

An expression (or lexicon) that is perfectly deterministic has ambiguity 0. Maximum ambiguity is bounded by the size of the conceptual domain (but is always less than 1). An expression whose meaning is a perfectly flat distribution over all concepts has ambiguity $1 - \frac{1}{|C|^2}$.

A word is *polysemous* if its meaning is ambiguous among conceptually similar interpretations. We define polysemy as the probability that two different interpretations of m are similar (weighted by how similar they are):

$$\text{Polisemy}(m, \mathsf{L}_a) = \sum_{c_1 \neq c_2 \in C} \mathsf{L}^i[m, c_1] \cdot \mathsf{L}_a^i[m, c_2] \cdot \text{sim}(c_1, c_2) \quad [3]$$

By definition, $0 \leq \text{Polisemy}(m, \mathsf{L}_a) \leq \text{Ambiguity}(m, \mathsf{L}_a) \leq 1$. One may think of polysemy as measuring how much ambiguity is due to ambiguity among similar concepts. Polysemy of an agent's lexicon is likewise defined as the average polysemy of its expressions and polysemy of the abstract joint lexicon is the average polysemy of all agents' representations.

---

[b]Here H is Shannon entropy defined with the base $|A'|$ logarithm. This ensures that the Jenson-Shannon divergence (Lin, 1991) is bounded by 1 for communities of any size.

Communication Strategies. Agents may employ a *literal* strategy (LL), simply producing expressions and selecting interpretations according to the conditional probabilities encoded in their respective lexicons: $P(m \mid c, \text{LL}) = \mathsf{L}^p[c, m]$ for production and $P(c \mid m, \text{LL}) = \mathsf{L}^i[m, c]$ for interpretation. However, we also endow our agents with the possibility to employ a *semantic approximation* (SA) strategy. With this, we attempt to model situations such as those exemplified in the Introduction, where, for example, an agent wants to communicate a concept c, but there are no lexical items that express that concept especially well, so she instead utters some expression m because there is a concept c' similar to c such that $\mathsf{L}^p[c', m]$ is high.

When choosing a message to express a concept c with semantic approximation, the speaker considers all the concepts each message m expresses and how similar they are to the intended concept.

$$P(m \mid c, \text{SA}) = \frac{\sum_{c'} \text{sim}(c, c') \cdot \mathsf{L}^p[c', m]}{\sum_{m', c'} \text{sim}(c, c') \cdot \mathsf{L}^p[c', m']} \tag{4}$$

Interpretation according to SA is defined similarly:

$$P(c \mid m, \text{SA}) = \frac{\sum_{c'} \text{sim}(c, c') \cdot \mathsf{L}^i[m, c']}{\sum_{m', c'} \text{sim}(c, c') \cdot \mathsf{L}^i[m', c']} \tag{5}$$

Which communicative strategy is being used may change from utterance to utterance within a single dialogue. We assume, however, that the current speaker always chooses the agents' communicative strategy.[c] In our model, how likely a speaker is to use semantic approximation is proportional to how much worse than average her lexicon is at expressing her intention:

$$P(\text{SA}) = \min\left(0, 1 - |C| \sum_{m \in M} \mathsf{L}[c, m]\right) \tag{6}$$

3. Experimental Setup & Results

We are interested in investigating the interplay between semantic approximation and the local common ground built up by a pair of agents over the course of a dialogue (encoded in the transient discourse lexicon), as well as how this interplay may shape the lexical resources of an overall linguistic community in the long term. To this end, simulations were run varying two parameters: communication strategy (whether agents were restricted to the literal strategy or could also use semantic approximation) and discourse lexicon (whether agents only kept track of the community lexicon or also made use of a local discourse lexicon).

---

[c]How coordination on a given strategy is achieved is beyond the scope of this paper, but recall (§1) that, for example, speakers may explicitly indicate their use of semantic approximation by hedging.
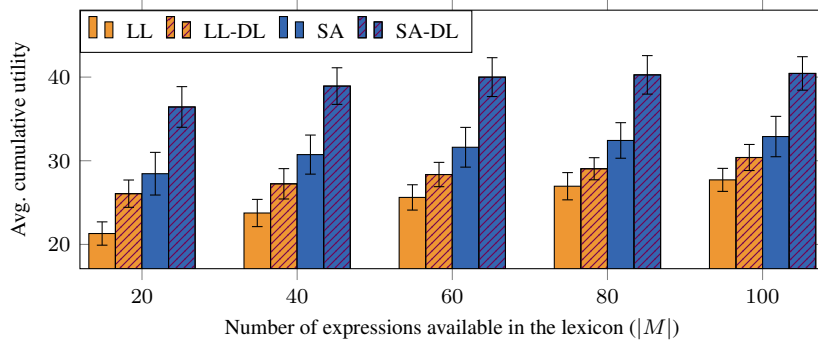
Figure 2. Average cumulative utility (as a percentage of total available utility) across 50 random concept domains after 1000 dialogues.

Settings. All four combinations (SA, LL, SA-DL and LL-DL) were simulated with $|M|$ ranging from 20 to 100 expressions (intervals of 20) on 50 random concept domains. Communities of 5 agents and concept domains of 100 concepts were used. We generate concept domains using the Holme-Kim growing random graph algorithm, which encourages clustering (Holme & Kim, 2002). That is, if c_1 and c_2 are similar and c_2 and c_3 are similar, it is more likely than average that c_1 and c_3 are similar as well. All simulations lasted for 1000 dialogues of 50 utterances each. Agents employed Roth-Erev reinforcement learning with a discounting rate of $\lambda = 0.99$.

The concepts speakers intend to express are typically non-uniformly distributed in a given conversation—concepts that have appeared once are likely to come up again in the same discourse than are concepts that have not been expressed. To model this intuition, each discourse is accompanied by a "context"[d] (a probability distribution over concepts) which is itself drawn from a symmetric Dirichlet prior ($\alpha = 0.1$). Because the prior distribution is symmetric, no concept is more likely than any other to come up accross discourses, but since the Dirichlet prior tends to produce (more so with lower α), in any given discourse, some concepts will tend to be expressed more than others.

Results. To measure communicative success, we consider a simulation's cumulative utility; that is, the sum of the similarity between speaker intention and addressee interpretation over all turns in all dialogues (Figure 2). We find that the option to use semantic approximation consistently improves communicative success, but that this improvement is more pronounced when agents keep a discourse

---

[d]We assume that interlocutors do not exploit any knowledge of the prior probability of an intention in their interpretation and production strategies. In this way, the distribution over concepts that accompanies a discourse is different than in other models.
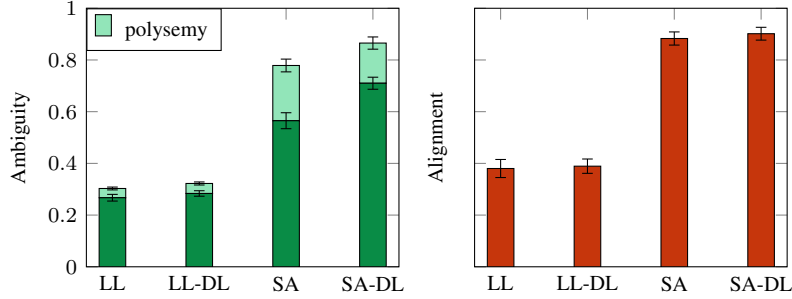
Figure 3. Characteristics of agents' joint community lexicon after 1000 dialogues (with $|M| = 40$). Polysemy is shown as the top portion of ambiguity (represented by the whole bar).

lexicon ($d > 5.01$ for all lexicon sizes with DL vs. $d < 3.47$ without DL).[e] Furthermore, the discourse lexicon improves communicative success whether or not semantic approximation is used ($d > 1.41$ for all lexicon sizes and both strategies). We also note that using semantic approximation mitigates the disadvantage of having a smaller lexicon. For example, the effect on cumulative utility of going from 100 to 20 expressions is smaller if SA is used ($d = -1.79$ and -1.80, for SA and SA-DL versus $d = -4.63$ and -2.70, for LL and LL-DL respectively).

Figure 3 shows the effect of semantic approximation on community-level lexical alignment, ambiguity and polysemy. The use of SA results in higher alignment amongst agents, and it also leads to higher levels of ambiguity and polysemy in the joint (abstract) community lexicon. In both cases this effect is exaggerated for larger lexicons. However, when SA is used, the presence of a discourse lexicon reduces the proportion of polysemy ($p < 0.01$).

Finally, Figure 4 shows how agents' discourse lexicon compares to their community lexicon after the former has evolved over the course of a dialogue in terms of expected utility—the utility yielded by an additional hypothetical utterance by one of the interlocutors. At the end of each dialogue (after 50 utterances), we computed the expected utility of the interlocutors' discourse and community lexicons. We observe that overall the discourse lexicon has higher expected utility than the community lexicon ($p < 0.01$ for all dialogues after the 250th), but this advantage is greater when semantic approximation is allowed.

4. Discussion

Discourse-specific conventions allow agents to express concepts that have poor coverage in the community lexicon. We see, however, that the discourse lexicon is most helpful when semantic approximation is an option (Figure 2). In order for the discourse lexicon to tune its expressivity to a particular discourse, there

---

[e]All reported effect sizes (Cohen's d) accompany a dependent student's t-test with $p < 0.01$.
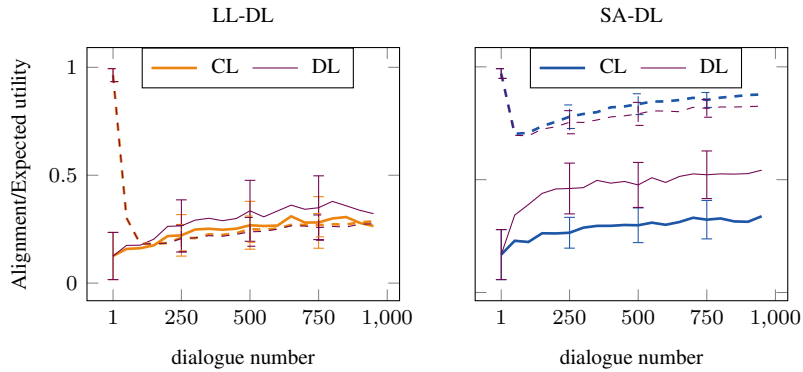
Figure 4. Expected utility (solid lines) and alignment (dashed lines) of the interlocutors' discourse lexicons (after a just-completed dialogue) versus their community lexicons.

must be some extra-lexical way of establishing those conventions not present in the community lexicon. Non-literal communication strategies such as semantic approximation fill this role. As in example (2) in the Introduction, a message-concept pairing is first suggested through semantic approximation, and then, when it is found to be successful, conventionalized for use in the remainder of the discourse, and possibly beyond.

In addition to creating opportunities for discourse-level lexicalization, semantic approximation has a profound effect on the long-term character of community-level lexical resources. For example, when semantic approximation is used, community lexicons are more ambiguous (Figure 3). Ambiguity is a paradox that has long puzzled linguists—it is commonplace in natural language despite the fact that it is intuitively detrimental to communicative success. One explanation for lexical ambiguity (Piantadosi et al., 2012) is that the re-use of expressions improves communicative efficiency, assuming that context is informative about meaning. But this answer doesn't explain some of the specific features of natural language ambiguity—particularly the prevalence of polysemy. The results of these simulations suggest yet another explanation for ambiguity in natural language: the use of semantic approximation (which is communicatively beneficial on a discourse-level) leads to polysemous *ad hoc* conventions, which, with enough use, are lexicalized at the community level.

Semantic approximation is not the only method of non-literal semantic co-ordination in natural language, however. In addition to further investigation of semantic approximation, future work should investigate how other local communication strategies such as metaphor contribute to discourse-level communicative success and influence the development of community-level semantic consensus.

References

Brennan, S., & Clark, H. (1996). Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology*, *22*(6), 1482–1493.

Chaix, P. Y., Barry, I., & Duvignau, K. (2012). Semantic approximation in SLI and normal development. *Revue Française de Linguistique Appliquée*, *XVII*(2),

Clark, H. H. (1996). *Using language.* Cambridge University Press.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 848–881.

Holme, P., & Kim, B. J. (2002). Growing scale-free networks with tunable clustering. *Physical Review E*, *65*.

Juba, B., Kalai, A. T., Khanna, S., & Sudan, M. (2011). Compression without a common prior: An information-theoretic justification for ambiguity in language. In *Innovations in Computer Science* (pp. 79–86).

Lewis, D. (1969). *Convention: A philosophical study.* Harvard University Press.

Lin, J. (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, *37*(1), 145–151.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*(3), 280–291.

Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), *Pragmatics* (Vol. 9).