Signaling Under Uncertainty: Interpretative Alignment Without a Common Prior

Abstract

Communication involves a great deal of uncertainty. Prima facie, it is therefore surprising that biological communication systems – from cellular to human – exhibit a high degree of ambiguity and often leave its resolution to contextual cues. This puzzle deepens once we consider that contextual information may diverge between individuals. In the following we lay out a model to analyze ambiguous communication in iterated interactions between subjectively rational agents lacking a common contextual prior. We argue ambiguity's justification to lie in endowing interlocutors with means to flexibly adapt language use to each other and the context of their interaction in order to serve their communicative preferences best. Linguistic alignment is shown to play an important role in this process as it foments convergence of contextual expectations, thereby influencing the use and interpretation of ambiguous messages. We conclude that ambiguity is ecologically rational when (i) interlocutors' (beliefs about) contextual expectations are generally in line or (ii) they interact multiple times in an informative context, enabling for the alignment of their expectations. In light of these results, we argue that meaning multiplicity can be understood as an opportunistic outcome enabled and shaped by linguistic adaptation and contextual information.

1 Meaning multiplicity in communication

In principle, speakers can draw from a large and virtually inexhaustive pool of alternatives to convey a state of affairs. We can refer to an entity as *Donald Trump*, as *the* 45^{th} *president of the United States*, or simply as *that guy*. Similarly, we may say *bank* rather than *financial institute*, *bat* rather than *baseball club*, *superfluous hair remover* rather than *remover of superfluous* *hair*, or *thing* instead of any of the aforementioned. When the primary goal is information transfer, the linguistic choices of speakers are chiefly constrained by whether their interlocutors will be able to infer this information from a message as intended. Why and when, then, would a speaker opt for a more ambiguous expression over one that is less ambiguous?

The diverse nature of the examples above illustrates the overall issue to be addressed: From cellular signals to those employed by meerkats and baboons, biological signaling is rife with meaning multiplicity (Greenough et al. 1998, Arnold and Zuberbühler 2006, Santana 2014). Natural languages are no exception. Prima facie, this fact may be qualified as puzzling, if not as downright indicative for a lack of communicative efficiency in their design (Chomsky 2002; 2008). Ambiguity avoidance has an intuitive appeal as the association of multiple meanings with a single expression can give rise to uncertainty in interpretation. Consequently, unambiguous languages may be argued to be better suited for communication. This idea has prominently figured in investigations on the emergence of signaling systems, where a communication system is standardly evaluated against the ideal of one-to-one form-meaning mappings (e.g., Lewis 1969, Steels 1998; see Spike et al. 2016 for a recent review). Notwithstanding, a number of investigations have argued for functional advantages of meaning multiplicity. It allows for smaller vocabularies (Santana 2014), greater signal compression (Juba et al. 2011), for the reuse of forms that are easier to produce or parse (Piantadosi et al. 2012, Dautriche 2015, van Rooij and Sevenster 2006), the partition of large semantic spaces (O'Connor 2015), coordination on non-lexicalized meaning (Brochhagen 2015), and deception in non-cooperative communication (Crawford and Sobel 1982). In most of these accounts the exploitation of contextual information plays a central role. The argument is simple: The information provided by a context needs not be codified in a signal. As a consequence, ambiguous languages are more compressed or enable for a more optimal reuse of their inventory than unambiguous counterparts while transmitting information as faithfully.

A complication ignored by this justification is that the gain attained from contextual information is not necessarily cashed out in situations in which it varies across agents. Once divergence of contextual expectations is admitted it remains to be shown what the consequences of meaning multiplicity are for cooperative communication without imposing a bottleneck on a language's inventory size (cf. Crawford and Sobel 1982, O'Connor 2015). Furthermore, even if the assumption of a common contextual prior were justified, it is not clear how such aligned expectations come about nor how they relate to the context itself. In the following we take up these challenges by analyzing ambiguous communication in (iterated) interactions without a common contextual prior. To this end, we put forward a game theoretic model that couches simple adaptive dynamics in rational language use (Frank and Goodman 2012, Franke and Jäger 2014), combining mutual reasoning with pragmatic uncertainty. Our main goal is twofold. First, we set out to investigate the conditions under which meaning multiplicity is advantageous by going beyond static approaches and by decoupling the context from its subjective perception. Second, we seek to further our understanding of the consequences of linguistic alignment by analyzing how the interplay of context, contextual expectations and iterated interactions shapes (un)ambiguous language use.

We proceed as follows. Section 2 lays out our main assumptions together with the core model we employ to characterize communication under pragmatic uncertainty. Section 3 showcases the model's main predictions and explores the consequences of possible refinements, as well as those that follow from environmental constraints. We discuss and review our main findings in section 4 and conclude in section 5.

2 Ambiguous signaling through pragmatic inference

In linguistics it is common to distinguish different types of meaning multiplicity based on syntactic, phonetic, graphemic, or semantic criteria. We take a decision-theoretic point of view under which the relevant distinction instead concerns whether or not communicative success hinges on the discrimination of interpretations associated with the same form. That is, whether an expression requires its addressee to settle for a particular interpretation over others – be it simplex or complex, and irrespective of the locus of its meaning multiplicity. While this conception is broad, it excludes notions such as vagueness, where an expression may have multiple precifications but (at least partially) successful interpretation does not hinge in teasing them apart (see, e.g., Jaegher and van Rooij 2011, Franke et al. 2011, O'Connor 2014). For the sake of brevity we will call this property ambiguity, as tacitly done so far.

We model language use to the effect that a speaker decides whether to send an ambiguous signal based on her assumptions about her interlocutor's likely interpretation. That is, the speaker gauges whether her addressee will be able to infer the intended meaning from an ambiguous message. If not, she may opt for a less ambiguous one to minimize the risk of misunderstanding. In turn, the hearer's interpretation will depend on her subjective expectations in a given context, i.e., the relative saliency of interpretations that are truthconditionally compatible with the message used by the speaker. We construe such expectations as any source of information beyond an expression's literal meaning. These expectations make up an interlocutor's subjective prior over meanings. Among others, a prior may draw from the context in which communication takes place, expectations of language use, or perceptual information. In short, it represents condensed information of the association strength with which an interpretation comes to mind (Franke 2009).

The choice of an expression is further modulated by a speaker's preferences. For instance, she may prefer a polite but less clear expression over one that is more explicit, have particular stylistic predispositions, prefer shorter over longer expressions, or have preferences over matters such as the relative cognitive load of the retrieval of expressions. We will illustrate our predictions by assuming a speaker preference for brevity. As argued in the following brevity is a plausible candidate for a preference shared across individuals and additionally has a bearing on domains central to our purpose: linguistic choice, dialogal adaptation, ambiguity and contextual predictability.

Brevity is often argued to be a rational speaker-oriented principle, posited, e.g., in Grice's (1975) maxim of manner, Horn's (1984) R-principle, and Zipf's (1949) principle of least effort. The tension between ambiguity and brevity is explicit in the interaction between Grice's maxims of quantity – to be as but not more informative than required – and his manner (sub)maxims – to be brief but to avoid ambiguity. In dialog, message brevity has been reported to increase incrementally in iterated tasks (Clark and Wilkes-Gibbs 1986, Motamedi et al. 2016). This provides some indirect evidence for the claim that speakers seek to increase message compression when possible. In the case of word length, Piantadosi et al. (2011) report cross-linguistic evidence for its predictability based on contextual information, a prediction that was subsequently corroborated by Mahowald et al. (2013) in a behavioral study suggesting that this relation is a consequence of speaker choices (instead of a statistical effect of language use or word classes). That is, there is some support for the assumption that brevity interacts with contextual information and influences linguistic behavior. Furthermore, there is a wealth of evidence for a negative correlation between contextual predictability and the pronunciation length of phones and words (see Brennan and Hanna 2009:§2.1, Piantadosi et al. 2011, and references within).

The claim that ambiguity's risk is assessed through contextual rather than language internal factors has also received some empirical support (see Ferreira 2008 and Wasow 2015 for recent psycholinguistic overviews). Two main findings are relevant here. First, there is little evidence for the idea that ambiguity influences linguistic behavior to the extent that speakers always prefer unambiguous expressions over ambiguous ones. This is contrary to the claim that ambiguity avoidance exerts a strong influence on speakers' choices, as phrased in the introduction. Second, while no conclusive evidence for this kind of avoidance has been found, Ferreira et al. 2005 do report a tendency for the avoidance of "extra-linguistic" ambiguity in naming tasks in which a disambiguated reading would still apply to multiple objects. For instance, when subjects were presented with multiple baseball bats of different sizes, the plain label *bat* was avoided to name one of them. The same degree of avoidance was not registered when the naming target was a baseball bat but a bat of the zoological kind was also present. A possible rationalization of this difference is that speakers may expect meaning multiplicity rooted in language to be manageable, whereas more information is supplied when context typically would not lend sufficient support to a single interpretation – as is the case when a label applies to multiple objects of the same kind. As we argue in the following: for ambiguity to be advantageous meanings attached to a single form should generally appear in contrasting contexts to safeguard understanding, i.e., in contexts in which priors sufficiently favor one interpretation over the other. Such expectation of disambiguation may not be warranted when the risk of misunderstanding stems from atypical or language external factors. Particularly when the addressee is unknown to the speaker as in Ferreira et al.'s task.

We make three main assumptions drawing from the above. First, speakers have preferences over messages. Second, priors are representations of subjective a priori meaning saliency. Third, interlocutors engage in mutual reasoning about rational language use. In particular, speakers reason about their interlocutors' contextual expectations in order to gauge whether a message will be understood as intended.

Preliminaries. Communication is represented as a signaling game between two players; a sender and a receiver (Lewis 1969). The sender's aim is to communicate a state of affairs $s \in S$ by sending a message $m \in M$. In turn, the receiver's task is to interpret the message. An interaction's outcome accordingly depends on whether the state was conveyed successfully.

A sender's inclinations over ways of expressing states of affairs is codified in a cost vector $\vec{c}$. Cost is inversely related to preference, and $\vec{c}_i$ is the cost of m_i. For convenience, we sometimes write $c(m_i)$ for the cost of message i. Subjective expectations in a given context are represented by a prior over states $pr \in \Delta(S)$, where pr^x is player x's prior. Importantly, players are uncertain about their interlocutor's prior. This uncertainty is captured by a distribution over priors over states, $\mathscr{P} \in \Delta(pr)$. Accordingly, $\mathscr{P}(pr)$ represents a player's certainty about pr being her interlocutor's prior.

Signaling behavior. Sender and receiver reason about each other to inform their linguistic choices. Following previous models of rational language use this process of mutual reasoning is captured by a hierarchy over reasoning types (Frank and Goodman 2012, Franke and Jäger 2014). The bottom of the hierarchy, level 0, corresponds to literal signaling behavior. Such players do not reason about their interlocutors but simply produce/comprehend according to their preferences/expectations and their language's semantics. Player x's literal receiver and sender behavior is given by ρ^0 and σ^0, respectively:

$$\begin{aligned} \rho^0(s|m, pr^x) &\propto L(s,m)pr^x(s); \\ \sigma^0(m|s) &\propto L(s,m) - c(m), \end{aligned}$$

where L is a lexicon that maps a state-message pair of a language to the Boolean truth-value of the message in that state. A minimal lexicon fragment that makes the choice between ambiguous messages over unambiguous ones precise is one with three messages and two states, where $L(s_1, m_1) = 1 = L(s_2, m_2)$, $L(s_1, m_3) = 1 = L(s_2, m_3)$ and all other state-message pairings are false. Put into words, message m_1 is exclusively true of state s_1, m_2 only of s_2 and m_3 is ambiguous between these two states. Speakers of L therefore need not use ambiguous m_3 but may nevertheless choose to do so. Furthermore, let us assume that m_3 is shorter and thereby preferred over m_1 and m_2; $c(m_3) < c(m_1)$ and $c(m_3) < c(m_2)$. Crucially, ambiguous m_3 is more risky to use than either alternative because it is semantically associated with both states. Our initial question can therefore be recast as asking under which conditions the risk incurred by the use of preferred m_3 undercuts the benefit of safe unambiguous communication using only m_1 and m_2.

The tension of a sender wanting to uphold her preferences as much as possible while taking the, possibly diverging, expectations of her interlocutor into consideration arises when higher reasoning types of level $n + 1$ are considered. As in other models of rational language use these types behave rationally according to the expected behavior of a level n interlocutor.[1] In the following, we restrict our analysis to level 1 senders and receivers. This

[1] A more flexible alternative is for players to have beliefs about their interlocutor's level of sophistication and for them to derive choice probabilities from these beliefs (see, e.g., Camerer et al. 2004). Our assumption of players who believe their interlocutor to be exactly one level less sophisticated than them is first and foremost made for simplicity, but has also been shown to succeed in predicting attested linguistic patterns (see Goodman and Frank 2016 for an overview of some predictions and Franke and Degen 2016 for more detailed discussion). Ultimately, this issue is empirical. For the time being, we opt for a simpler and to our mind more perspicuous model while bounding agents' reasoning to a low degree of sophistication.

minimal degree of mutual reasoning suffices to associate m_3 with a salient state under suitable conditions (when the sender's beliefs and the receiver prior are informative enough). Our departure from the standard approach concerns the behavior of the sender, who, instead of using her own prior to anticipate the receiver's behavior, employs her beliefs about the receiver's prior $\mathscr{P}$. The level 1 behavior of player x is accordingly as follows:

$$\rho^1(s|m, pr^x) \propto \exp(\lambda \underbrace{N(\sigma^0(m|s))}_{\propto \sigma^0(m|s)pr^x(s)});$$
$$\sigma^1(m|s, \mathscr{P}) \propto \exp(\lambda((\sum_{pr} \mathscr{P}(pr)\rho^0(s|m, pr)) - c(m))),$$

where λ is a rationality parameter, $\lambda \geq 0$ (Luce 1959, Sutton and Barto 1998). As λ increases so does the agents' tendency to maximize expected utility (see below for details). For a sender this means that messages judged to have a high probability of being understood that are of low cost are increasingly prioritized over low success and/or high cost ones. In the case of receivers, states true of a message that are favored by their prior are more likely to be inferred over less expected or false ones.

The behavior of speakers of level 1 corresponds to the quantal best response to a belief-weighted level 0 hearer. The latter is derived from the priors in the domain of $\mathscr{P}$, the speaker's beliefs about the receiver's prior, with weights according to the sender's belief in them as corresponding to the prior of the hearer. This proposal is conservative in that it retains the predictions made by previous models of rational language use (Frank and Goodman 2012, Franke and Jäger 2014) when the prior is (believed to be) common, i.e., when $\mathscr{P}$ is degenerate, ruling out all priors except for the speaker's. Importantly, it can additionally capture situations in which the sender is either uncertain about her interlocutor's expectations, or is certain but assumes that it differs from her own. Such situations can arise in a number of ways but boil down to a speaker's increased tendency to use safer messages when uncertain, or to use messages in a way that might go against her own prior but be in line with her beliefs about her addressee's, respectively. Reasoning beyond level 1 allows for further variability in receiver behavior depending on her beliefs about the sender's beliefs, and vice-versa for the sender. As noted above, we do not make use of such additional layers of complexity here.

For illustratory purposes, let us assume that there are only two distributions in the support of $\mathscr{P}$. For example, $pr_i(s_1) = 0.9 = pr_j(s_2)$. The prior pr_i strongly favors state s_1 over s_2 and vice-versa for pr_j. Furthermore, let interlocutors tend to maximize expected utility, rendering their behavior more deterministic, and assume that the lexicon and the cost-induced order over messages is as above. While there is a gamut of possible speaker behaviors that arise from an interaction between $\mathscr{P}$ and the values assigned to λ and $\vec{c}$, there are three general cases of interest. The first is given by $\mathscr{P}$ assigning high probability to pr_i. In this case, ambiguous m_3 is sent in s_1 to maximize expected utility. Since the receiver is assumed to expect s_1, m_3 is judged to be risky in s_2 and unambiguous m_2 is sent in this state instead. The second case, in which high probability is assigned to pr_j, is the opposite of the first: m_3 is sent in s_2 but not in s_1, where m_1 is sent instead. Lastly, the sender may be uncertain about the receiver's prior, reflected, e.g., by uniform $\mathscr{P}$. In this case, the speaker will opt for the safe strategy of sending m_1 in s_1 and m_2 in s_2.

Communicative success. After an interaction players receive a payoff depending on whether communication succeeded. Payoffs differ between sender and receiver as the former incurs some cost depending on the message chosen. When a sender sends m in s and a receiver interprets m as s', the sender's payoff is $u^S(s, m, s') = \delta(s, s') - c(m)$ and the receiver's is $u^R(s, m, s') = \delta(s, s')$, where $\delta(s, s') = 1$ iff $s = s'$ and otherwise 0. The expected utility of sender x using strategy σ and receiver y using strategy ρ is $U^S(x, y) = \sum_s P^*(s) \sum_m \sigma(m|s, \mathscr{P}) \sum_{s'} \rho(s'|m, pr^y) u^S(s, m, s')$ and $U^R(x, y) = \sum_s P^*(s) \sum_m \sigma(m|s, \mathscr{P}) \sum_{s'} \rho(s'|m, pr^y) u^R(s, m, s')$, where $P^* \in \Delta(S)$ is the true distribution over states.

The difference between P^* and the players' subjective expectations, $pr \in \Delta(S)$, is that the former gives the true frequency of states, which can be thought of as the actual context in which communication takes place, whereas the latter corresponds to subjective expectations entertained by agents in this context. As stressed above, contextual expectations may diverge from each other as well as from the context itself. As reflected by $\delta(\cdot)$, we restrict our attention to cooperative communication. Interlocutors strive to understand each other and senders do not gain from deceiving their addressees through the use of ambiguous messages.

Single interactions already provide us with the means to quantify how well a pairing of signaling strategies fares in a context. However, the degree to which agents succeed chiefly depends on their (beliefs about their interlocutors') priors, and on how well these match the context. Furthermore, a crucial component missing from such an analysis is the possibility of players to interact with each other in a given environment. Clearly, if they know nothing about each other the best a player can do is to make a guess and hope for the best. In contrast, iterated interactions allow the sender to change her beliefs according to information obtained from her interlocutor's linguistic behavior, as well as for expectations over states to adapt to the context itself.

Iterated interactions. More often than not communication involves iterated rather than single interactions. This allows interlocutors to adapt to each other. In dialog, linguistic alignment is evinced on many levels; from phonetic (Kim et al. 2011) or syntactic (Pickering and Ferreira 2008) to lexical and referential (Brennan and Clark 1996, Clark and Wilkes-Gibbs 1986). Here, we are concerned with the relation between subjective contextual expectations, beliefs about them, and the information provided by the context in which interactions take place. The latter is codified in $P^*(\cdot)$, which interlocutors are indirectly exposed to while they interact. What is missing, then, are means for priors and beliefs about them to change over time.

Communication ensues as before: the sender wants to convey a state and sends a message, the receiver interprets it, and both players receive a payoff. However, now the players' own subjective priors and their beliefs about their interlocutor's prior are updated based on information gained from the interaction.

In the following we assume priors to be updated based on a player's accumulated propensity for each state s at interaction t, $ap_t(s)$. Accumulated propensity can be likened to a record of the states that the sender intended to communicate, or the receiver interpreted, in previous interactions. Before any interaction player x's propensity is simply proportional to her prior, $ap_0^x(s_i) \propto pr^x(s_i)$. Subsequently, the prior for interaction $t+1$ is derived from a player's amassed propensity up to interaction t, $pr_{t+1}^x(s_i) \propto ap_t^x(s_i)$. After each interaction the propensity for the state in play is updated by a value r, positive in case of communicative success and negative otherwise. That is, for sender x that sent m in s with receiver y interpreting this message as s':

$$ap_{t+1}^x(s) = ap_t^x(s) + f(r), \text{ where } \begin{array}{ll} f(r) = r & \text{if } \delta(s, s') = 1, \\ f(r) = -r & \text{otherwise.} \end{array}$$

The receiver's prior is updated analogously via $ap_{t+1}^y(s')$. The motivation behind this rather simple learning mechanism is to (ideally) obtain high rationality outcomes from low rationality behavior (Huttegger et al. 2013). Additionally, it allows us to maintain an analogy to simple biological learning processes (Thorndike 1898, Herrnstein 1970). In human terms this process is akin to priming in that a state's saliency increases as interlocutors are exposed to it (Pickering and Garrod 2004, Reitter and Moore 2014).

The value by which propensities change controls how fast the initial prior is overridden. Small r gives the initial prior more weight whereas larger values lead players to abandon their preconceptions faster. Negative reinforcement is not required for the iterated convergence of priors but speeds up the process.[2] The value of r is dissociated from payoff values. This diverges from most previous signaling models with adaptive learning dynamics (e.g., Barrett and Zollman 2009, Franke 2015). Notwithstanding, this assumption is warranted here as there is no reason to relate a speaker's prior over states to incurred production cost. In fact, a direct association of payoffs to updates would have undesirable consequences in cases where messages true of less frequent states are less costly than those true of more frequent ones. This would lead players to expect the former to be more salient than the latter. In informal terms: having a preference to talk about something in a particular fashion should not make it a priori more probable.

In contrast to the somewhat mechanistic fashion in which priors are updated, we assume the change of a sender's beliefs about her addressee's prior to involve an inferential component, here modeled as an update of $\mathscr{P}$ that consolidates old with new information using Bayes' rule. This reflects the sender's primary goal to (actively) reach understanding by correctly anticipating her addressee's interpretation, a motivation already rooted in the agents' engagement in mutual reasoning.

The evidence witnessed by the sender is whether communication succeeded. However, she receives no information about the receiver's interpretation if communication failed – beyond the fact that it failed. More precisely, in an interaction in which the speaker wanted to convey s_i with message m, interpreted as s_k by the receiver, she witnesses $w(s_i)$, where $w(s_i) = \{s_i\}$ if $\delta(s_i, s_k) = 1$ and $S \setminus \{s_i\}$, otherwise. Based on $w(s_i)$, the sender adjusts her beliefs about her interlocutor's prior based on the likelihood of a prior leading to the witnessed receiver behavior. Accordingly, $\mathscr{P}$ is updated as

[2]Other possibilities include the addition of recency effects, by weighting recent states higher than less recent ones, or learning with suppression, by decreasing the association strength of states that were not in play (Erev and Roth 1998, Franke and Jäger 2011). Alternatively, interlocutors could use more sophisticated mechanisms to update their priors. As with our previous choices, we decide for a simple mechanism that serves our purpose. The contribution of reinforcement learning to our predictions is straightforward and can be achieved in a number of ways; a player's expectations of a state should grow with increased exposure to it.

follows:

$$\mathscr{P}_{t+1}(pr \mid w(s_i), m) \ \propto (\sum_{s \in w(s_i)} \rho^0(s \mid m, pr))\mathscr{P}_t(pr).$$

When interacting again linguistic choice is computed as before with updated priors and updated beliefs over them.

3 Predictions for single and iterated interactions

Based on the preceding discussion, a straightforward first prediction is that ambiguous communication is at least functionally equivalent, in terms of information transfer and fulfillment of speaker preferences, to unambiguous counterparts provided that (i) the speaker's beliefs about the receiver's prior correctly anticipates her actual behavior, and that (ii) signaling behavior is relatively deterministic, particularly to ensure that receivers have a tendency to associate an ambiguous message with a single state in a given context. More importantly, under these conditions ambiguity is functionally advantageous when there are at least two contexts in which the true distribution over states assigns a non-zero probability to distinct states associated with a preferred ambiguous message and the speaker uses this message in both contexts.[3] Lastly, it is maximally advantageous in a context if the most frequent state in it is associated with the least costly message in $\vec{c}$. Put differently, the most frequent state(s) in a context ought ideally be associated with the most preferred form(s) when speaker economy is at stake.

The adoption of an ambiguous strategy ultimately hinges on the sender's beliefs about the receiver. This advantage therefore depends on factors that would lead agents to have similar expectations (over expectations). This also means that ambiguous signaling is more risky in a world in which contextual expectations greatly vary across agents. In the case of humans, behavioral experiments suggest that they generally succeed, at least significantly beyond chance, in matching their expectations with those of others when it is known that the other party is trying to do the same (Schelling 1980, Mehta et al. 1994). However, from previous accounts and our analysis so far, it is unclear how agents may come to entertain such aligned expectations.

[3] For every single context there is an unambiguous language that fares at least as well as an ambiguous one. For example, either one in which m_3 is only true of s_1 or one in which this message is true only of s_2. An advantage for ambiguity can therefore only manifest when there are multiple contexts. We return to this matter in §4.

To recapitulate, ambiguity can be advantageous in single interactions as long as sender beliefs anticipate receiver behavior. Subjective priors therefore need not match for ambiguity to be exploited. No common prior is required. On a general level, this characterization is nevertheless in the spirit of previous justifications of ambiguity, with a shift of the explanatory burden from a common prior to sufficiently accurate beliefs about the receiver's prior. Crucially, this shift highlights that the conditions for safe ambiguity exploitation may not always be given and allows us to ask when and how they can be reached. Whether an ambiguous signal is understood depends on the receiver's own expectations, whether it is sent depends on the sender's beliefs about these expectations, and the expected utility of conveying a particular state by an ambiguous form will depend on the true distribution over states. We now turn to iterated communication to tease apart the interaction between these factors and to elucidate how and under which conditions this advantage crystallizes.

Simulations. In order to illustrate the model's predictions a sender's initial beliefs about the receiver's prior need to be set. Here, we assume sender x's initial $\mathscr{P}$ to be Dirichlet distributed, with weights for state s set to $k \times pr^x(s)+1$.[4] High k corresponds to the speaker believing that the receiver's expectations are close to her own with $k \rightarrow \infty$ approaching the belief of a common prior. Lower values correspond to more divergence and uncertainty. In the extreme, $k = 0$ leads to complete uncertainty about the receiver's prior – initially deeming every prior equally likely.

We use L as above, $\lambda = 20$ and $\vec{c} = \langle 0.4, 0.4, 0.1 \rangle$. That is, players are subjectively rational but might occasionally make mistakes and ambiguous m_3 is preferred over either unambiguous message, each of equal cost. To inspect the average outcome of interactions between two players, including best- and worse-case scenarios, priors are randomly sampled at the onset of a first interaction. The value of k is sampled from $[0, 20]$ at the onset as well.

The mean development of players' subjective prior in a context is illustrated in Figure 1 for two values of r, showing that priors approach the true distribution of the context as players interact in it. When there are only two states under consideration this simple learning process is particularly fast because negative reinforcement in one state leads to the prominence of the other. This figure also showcases the role of the reinforcement value r in controlling the speed by which priors converge to a context's distribution.

In the following we focus on results obtained from an r-value of 0.5 after 50 interactions. The choice of the latter ensures that the reported outcomes

[4]We thank an anonymous reviewer for this suggestion.

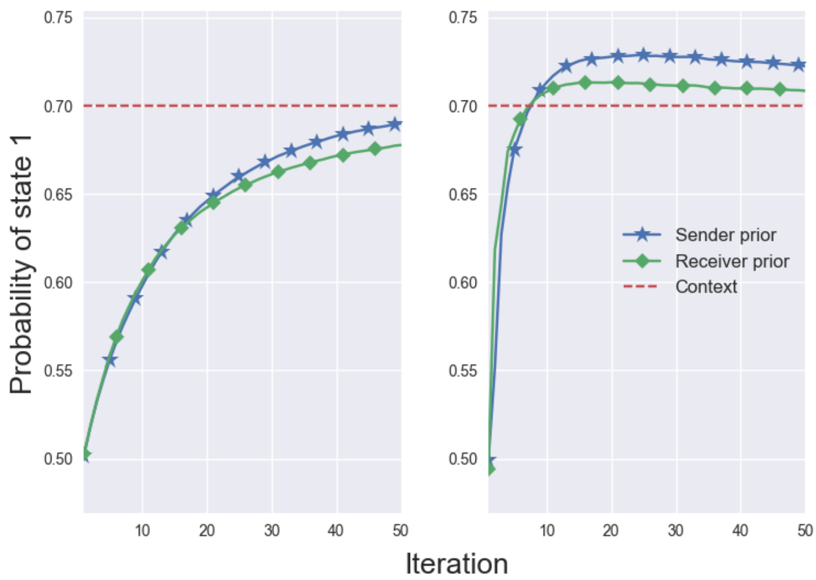

Figure 1: Mean subjective prior development in 10^4 independent simulations with $P^*(s_1) = 0.7$ for $r = 0.1$ (left) and $r = 1$ (right).

approximate endpoints of the dynamics but should not be taken as indicative of the minimal number of interactions required to reach them. Supplementary results obtained from less interactions and different r-values are provided in Appendix A.

A more central interaction is that between P^* and a sender's beliefs about her interlocutor's prior, as well as their bearing on the choice of an ambiguous message. Figure 2 showcases how the context influences sender beliefs. Recall that $\mathscr{P}$ is updated based on what can be inferred about the receiver's prior from her behavior. The only interactions that are informative about this matter, and therefore influence a sender's beliefs, are the receiver's interpretations of ambiguous messages. In turn, the receiver's interpretation of an ambiguous message may change over time due to her exposure to the context (cf. Figure 1). In particular, contexts that are not very informative can lead to fluctuations in the receiver's expectations, making them more difficult to predict for the sender. Consequently, as showcased by the left plot in Figure 2, senders grow uncertain about their interlocutor's expectations in such an uninformative context. More importantly, the uninformative prior that receivers converge to in such contexts does not lend itself for the safe exploitation of ambiguity either. Uncertainty about expectations centered around uninformative priors therefore often lead to the avoidance of risky signals. By contrast, receiver expectations in informative contexts are

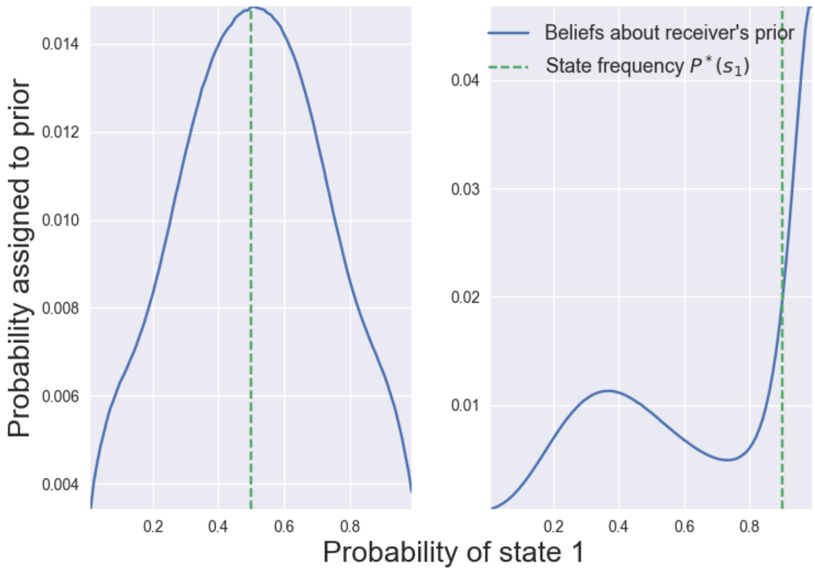

Figure 2: Mean beliefs about receiver expectations in 10^4 independent simulations with $P^*(s_1) = 0.5$ (left) and $P^*(s_1) = 0.9$ (right).

quite predictable after a few interactions, a fact senders picks up on once they employ an ambiguous signal. In fact, senders tend to overestimate their interlocutor's prior in such contexts because the likelihood of a correct interpretation of m_3 is higher the more degenerate subjective priors are. Such overestimation decreases as mutual reasoning levels increase but the predictions about the use or avoidance of an ambiguous message do not hinge on the shape of the sender's belief but on the range of priors it concentrates on. That is, a false belief about an addressee's prior is not detrimental to communication if it correctly predicts behavior.

The amount of senders that adopt an ambiguous strategy in a context is reflected most clearly by their expected utility. An excerpt of the mean expected sender utility together with the mean Jensen-Shannon divergence (JSD) between the interlocutors' priors is given in Table 1.[5] Note first that even in a non-informative context the mean expected utility of senders is higher than 0.6, the value guaranteed by the use of unambiguous messages m_1 and m_2. It is also higher than the mean utility of approximately 0.57 expected in the first interaction. The latter value is lower than the safe guaranteed value of 0.6 because priors and k were sampled randomly, inevitably leading

[5] Informally, JSD measures the closeness of two distributions as a divergence to their average. More precisely, $\text{JSD}(pr_i, pr_j) = {}^1\!/\!{}_2 D(pr_i||M) + {}^1\!/\!{}_2 D(pr_j||M)$, where $M = {}^1\!/\!{}_2(pr_i + pr_j)$ and $D(P||Q) = \sum_s P(s) \log {}^{P(s)}\!/\!{}_{Q(s)}$.

$P^*(s_1)$	U^S (σ)	JSD	$U^S_{\max}$
0.5	.61 (.06)	.004	.75
0.7	.68 (.1)	.002	.81
0.9	.72 (.13)	.002	.87

Table 1: Mean sender expected utility and JSD of interlocutors' priors after 50 iterations in 10^4 independent games. $U_{\max}$ indicates the maximum expected utility reachable for a given P^*.

to the failure of some initial attempts to exploit ambiguous m_3. As suggested by Figure 1, iterated interactions also strongly improve upon the mean initial JSD of approximately 0.15. The increase in standard deviation of expected utility with the informativity of a context is a consequence of the increasing difference between the utility gained from an ambiguous signaling strategy against that of adopting an unambiguous one.

In sum, these results (i) generalize past analyzes of ambiguity by relaxing the assumption of a common prior, (ii) show how agents may come to entertain (beliefs about) contextual expectations that allow for the safe exploitation of ambiguity, (iii) highlight the role of context informativity in enabling or preventing such exploitation, and (iv) connect this research with claims in the alignment literature about its role in dialog optimization, providing interlocutors with means to establish patterns of language use better tailored to the context of their interaction and their preferences (Clark and Wilkes-Gibbs 1986, Reitter and Moore 2014). More broadly, this interactive perspective also highlights the function of ambiguity as an opportunistic adaptive device that endows agents with the ability to mold language use to their interlocutors and the environment, and links this opportunism to the information provided by a context. Contexts of high informativity are particularly conducive to ambiguity exploitation over less informative ones as they (i) foment less fluctuations in the receiver's interpretation of ambiguous messages, and, consequently, (ii) lead to less uncertainty in the sender's beliefs about her interlocutors prior. Lastly, the expected utility of senders increases with context informativity as these contexts more often lead to the association of frequent states with preferred but ambiguous messages. This association is not explicitly sought after by senders but rather is a byproduct of a receiver's association of an ambiguous message with its most salient interpretation. The relationship between saliency, frequency and interpretation will therefore often lead players to adopt Pareto optimal signaling strategies.

Notwithstanding, as shown in Table 1, not even informative contexts guarantee that an ambiguous strategy is always adopted. There are two intertwined reasons for this. First, we allowed the priors of interlocutors to vary freely before engaging in communication. This may cause a speaker with an uninformative prior and high k to believe her interlocutor's prior to be uninformative as well. Consequently, such a speaker will never try to use an ambiguous message even after exploring the context (which may turn out to be informative). Similarly, initial uncertainty from low k may lead speakers to not use risky signals, meaning that they never learn anything about the receiver's expectations. Second, a great number of interactions started with opposing contextual expectations. This may lead to an early communicative failure when using an ambiguous message. As in the other cases mentioned this may deter the sender from using risky messages in the future. We briefly explore two assumptions that can address these issues.

3.1 Exploration and past experience

Communication often draws from past experience. In particular, agents may often find themselves in similar contexts. This enables visitors of zoos and baseball courts alike to use plain *bat* without first probing whether their interlocutor is attentive to the same meaning. They have experience in these contexts and assume that their interlocutors have had some too – at least to a degree to which one interpretation of ambiguous *bat* is markedly more expected than the other. Once we allow for the exploration of a context, either by taking into consideration previous interactions with other agents or non-linguistic exposure to it, the issue of strongly diverging initial priors is reduced. That is, a shared cultural background and experience in an environment suggest themselves as partial answers to the question how linguistic coordination with ambiguous messages can succeed prior to multiple interactions.[6]

The question how the speaker's initial confidence k is determined remains, however. While a detailed treatment is outside the scope of this investigation, one possibility is to make it sensitive to the informativity of a context in tandem with assumptions about the receiver's experience in similar contexts. In broad strokes; high k may come about because the context is assumed to be well known. Either because this is known about the receiver or because this context is common enough that members of a population are taken to be familiar with it. An informative context that is assumed to have been encountered frequently enough may then lead to an optimistic speaker

[6]There are many ways in which this idea could be implemented. For example, by deriving initial priors from samples of the true distribution or from past interactions with other agents. We chose not to do so as we hope its positive effect on coordination with ambiguous messages is clear from the preceding discussion.

$P^*(s_1)$	U^S (σ)	JSD	$U^S_{\max}$
0.5	.58 (.114)	.033	.75
0.7	.8 (.022)	.001	.81
0.9	.87 (0)	0	.87

Table 2: Mean sender expected utility and JSD of interlocutors' priors after 50 iterations in 10^4 independent games using "preemptive" belief updates. $U_{\max}$ indicates the maximum expected utility reachable for a given P^*.

strategy in which ambiguity is believed to be (usually) resolvable (cf. Clark and Schober's (1992) *presumption of interpretability*). In such cases adaptive dynamics still play a role in unknown or infrequent contexts, as well as as corrective devices when a presumption of interpretatibility turns out to be misplaced.

3.2 Preemptive adaptation

Next, we turn to the issue of senders who, due to early communicative failure or initial uncertainty about their interlocutors' expectations, remain averse to ambiguity even in informative contexts. A reason for senders occasionally "locking-in" on an unambiguous strategy even if they could safely exploit ambiguity is that the update of $\mathscr{P}$ is not sensitive to the information gained from the context nor to the fact that interlocutors adapt to it as they interact. One possibility to stimulate the exploration of ambiguous strategies after learning about the context is therefore for the beliefs about a prior being the interlocutor's to be affected by the probability of the current state s_i under that prior:

$$\mathscr{P}_{t+1}(pr \mid w(s_i), m, s_i) \propto (\sum_{s \in w(s_i)} \rho^{n-1}(s \mid m, pr))\mathscr{P}_t(pr)pr(s_i).$$

This operationalizes a sender that changes her beliefs on the assumption that her interlocutor adapts to the context – "preemptively" exploiting the relative saliency of states that were relevant before. Table 2 shows how this inference mechanism affects the outcome of interactions. As with the previous update rule, supplementary results obtained from less interactions and different r-values are provided in Appendix A.

In a nutshell, this less conservative update fares well in informative contexts but less so in uninformative ones. In the former the proportion of dyads that adopt the Pareto optimal strategy of associating m_3 with the most frequent state is markedly higher than under the simpler update mechanism. However, as shown for $P^*(s_1) = 0.5$, this may come at a cost in less informative contexts because this update favors priors that are informative about the current state in play. Consequently, while in such a context the receiver converges to a prior that is not well-suited for ambiguity exploitation, speakers instead tend to infer more informative priors, attempt the use of risky signals, and often fail. By contrast, our main proposal for updating $\mathscr{P}$ leads to more cautious behavior that may not always result in ambiguity exploitation but generally ensures that communication succeeds.

4 Discussion

We proposed a conservative generalization of models of rational language use and combined it with simple adaptive dynamics to generate predictions about ambiguous communication between players lacking a common prior. The model decouples interlocutors' subjective contextual expectations from each other, as well as from the environment itself. This weakens the assumptions of past investigations by neither assuming a common prior (Piantadosi et al. 2012, Santana 2014) nor shared randomness in a language's forms (Juba et al. 2011). Beyond their separation, these components were argued to iteratively feed into each other. A sender's beliefs about her interlocutor play a central role in her linguistic behavior, change according to the receiver's actions, and, at the same time, interlocutors' communicative intentions and expectations are indirectly shaped by the environment and the outcome of interactions.

In single interactions ambiguity is predicted to be advantageous when (beliefs about) priors are sufficiently aligned relative to the truth-conditions of a language (cf. Juba et al. 2011, Piantadosi et al. 2011, Santana 2014). We further showed that these conditions can often be reached when iterated interactions and adaptive mechanisms are considered even if players' priors are allowed to initially vary freely. In a nutshell, the more speakers interact, the closer their (beliefs about) contextual expectations grow, and the riskier their communication can be. Crucially, whether (beliefs about) expectations facilitate the safe exploitation of ambiguity is influenced by how informative the context of interaction is. More informative contexts allow interlocutors to reach an implicit agreement on salient meanings faster and more reliably than less informative ones. A byproduct of this interaction is the association of preferred forms with frequent meanings. Overall, these results add to the growing list of realms in which ambiguity has been argued to be justified; e.g., non-cooperative communication (Crawford and Sobel 1982), unaligned preferences (Jaegher and van Rooij 2013), and when a language's form inventory is restricted in size (O'Connor 2015).

The model also establishes a connection between models of rational language use, usually confined to single interactions, and linguistic alignment. In analogy to experimental findings with human subjects, it predicts increased signal compression as interlocutors interact (Fowler and Housum 1987, Clark and Wilkes-Gibbs 1986, Bard et al. 2000, Motamedi et al. 2016, Kim et al. 2011, Pickering and Ferreira 2008, Brennan and Clark 1996), a strong connection between linguistic adaptation and task success (Fusaroli et al. 2012), and audience and interaction dependent adaptation (Branigan et al. 2010, Garrod and Doherty 1994, Brennan and Clark 1996, Metzing and Brennan 2003). These parallels should however not be taken to suggest the model to be a comprehensive model of dialog adaptation. Our main aim was to add to the general understanding of the conditions under which ambiguity may be justified in cooperative communication, as well as how these conditions can be reached and how they interact. The model is therefore best viewed as an informed but idealized abstraction of communication. It is at this level that it makes predictions about ambiguous communication under the assumption that interlocutors (i) have preferences over messages, (ii) engage in mutual reasoning, (iii) are influenced by information acquired from (iiia) context and (iiib) their interlocutor, and that they (iv) have private contextual expectations. The specifics of these assumptions depend on the situation at hand. For instance, interactions in which linguistic feedback from addressees is limited – such as speeches, lectures or meetings – may require higher degrees of reasoning. Particularly from addressees. On the other extreme, other cases of biological signaling may involve less rather than more sophistication. Particularly, assumptions (ii) and (iiib) may seem contentious when applied to communication of non-human organisms. Along the lines of our and previous accounts, whether ambiguity is justified in such cases would instead depend on whether priors are generally aligned, dissipating the need for mutual reasoning. An important contributing factor to successful ambiguous communication without (ii) and (iiib) may be that other organisms have been argued to lack or only show very limited degrees of displacement, i.e., the ability to communicate about things that are not spatio-temporally present (Hockett 1960). By contrast, in the case of human communication, nothing prevents two zoologists at a baseball court to discuss their work on bats. In short, we laid out a general model of rational language use and analyzed its predictions under assumptions that draw from insights of previous research, but make no claim to have exhausted the diverse conditions under which biological signaling takes place.

We assumed that players accurately recognize the context they are in and that they approximate subjectively rational behavior (albeit bounded in mutual reasoning depth and allowing for occasional mistakes). These simplifying assumptions do not have a strong bearing on our main argument.

A weakening of either is tantamount to the introduction of a higher error rate when using ambiguous signals. It follows that if this rate exceeds the benefit of the use of preferred but ambiguous messages, then unambiguous communication is predicted to be more advantageous and consequently to be adopted. This is well in line with our argument that the benefit of meaning multiplicity is enabled by particular conditions rather than being a property that benefits language users across the board.

This investigation focused on analyzing the conditions under which ambiguous signals can be used without incurring communicative disadvantages in a single context. As noted earlier, one may therefore contend that, for any given context an unambiguous language that semantically associates the most frequent state with the most preferred message can be constructed. We agree. Were the world such that language users would always find themselves in exactly the same context, there would be little use to associating multiple meanings to a single form because contextual information would be invariant. In such a case speakers would do better if they avoided the risk of ambiguous communication altogether and opted for unambiguous expressions instead. It should therefore be stressed that the advantage of expressions that are true of more than one state lies in their ability to fulfill speaker preferences in multiple contexts simultaneously. This is something unambiguous language cannot do. Unambiguous alternatives are nevertheless important – at least for communication that allows for displacement. They come into play either when speakers need to signal a state that is not in line with (beliefs about) contextual expectations, or when these are not sufficiently informative.

To summarize, ambiguity endows agents with the ability to adapt their linguistic resources to an environment without incurring too great a risk of misunderstanding. This may involve an adaptation process between interlocutors in a particular situation, but can also draw from general knowledge about commonly experienced domains in single interactions. The more varied the world but more shared the experience the better ambiguous language users fare.

5 Conclusion

We argued that the risk of ambiguity lies not in the meaning multiplicity of expressions but rather in uncertainty about contextual expectations. In turn, its advantage lies in the reuse of preferred forms, leaving coordination on meaning to be partially resolved by the context of interaction. We have shown under which conditions this justification holds without a common contextual prior and characterized how language users may come to successfully communicate even when these conditions are initially not given, as well as when they fail to materialize. Linguistic alignment was shown to play a pivotal role in this process by having a bearing on coordination and convergence of (beliefs about) expectations over meaning, and thereby influencing linguistic choice. In more general terms, we argued that meaning multiplicity is an adaptive tool that enables agents to fit language to their needs, their interlocutors, and the environment, through an exploitation of shared pragmatic principles and (partially) shared contextual information.

Ambiguity is not inevitable. However, when the conditions for its exploitation are given it is likely to emerge through interaction. In functional terms our analysis echoes the sentiment already expressed by Miller (1951:111). Ambiguity is not the unruly creature it often is branded to be. Instead, its qualification as disruptive or suboptimal is an artifact of theoretical idealization – a product of expressions' isolated inspection instead of in the naturally richer contexts they are produced.

A Supplementary results

The following tables supplement the simulation results reported in §3. All outcomes correspond to a mean of 10^4 independent simulations with $\lambda = 20$ and $\vec{c} = \langle 0.4, 0.4, 0.1 \rangle$. The results in Table 3 were obtained from the simple update mechanism of $\mathscr{P}$ presented in §2 and those in Table 4 from its "preemptive" refinement in §3.

		10 iterations		30 iterations		50 iterations		
r	$P^*(s_1)$	U^S (σ)	JSD	U^S (σ)	JSD	U^S (σ)	JSD	U^S_{max}
0.1	0.5	.61 (.08)	.03	.62 (.07)	.00	.62 (.06)	.00	.75
	0.7	.62 (.09)	.03	.64 (.09)	.01	.67 (.10)	.00	.81
	0.9	.64 (.10)	.03	.68 (.12)	.01	.71 (.13)	.00	.87
0.5	0.5	.59 (.08)	.01	.61 (.06)	.01	.61 (.06)	.00	.75
	0.7	.62 (.10)	.01	.66 (.10)	.00	.68 (.10)	.00	.81
	0.9	.66 (.12)	.00	.70 (.13)	.00	.72 (.13)	.00	.87
1	0.5	.59 (.08)	.02	.61 (.06)	.01	.61 (.06)	.00	.75
	0.7	.62 (.10)	.01	.66 (.10)	.00	.68 (.10)	.00	.81
	0.9	.67 (.12)	.01	.70 (.13)	.00	.72 (.13)	.00	.87

Table 3: Mean sender expected utility and JSD of interlocutors' priors in 10^4 independent games. U_{max} indicates the maximum expected utility reachable for a given P^*.

r	$P^*(s_1)$	10 iterations U^S (σ)	JSD	30 iterations U^S (σ)	JSD	50 iterations U^S (σ)	JSD	U^S_{max}
0.1	0.5	.59 (.12)	.03	.59 (.12)	.01	.58 (.11)	.02	.75
	0.7	.67 (.14)	.03	.76 (.09)	.01	.79 (.05)	.00	.81
	0.9	.77 (.12)	.03	.86 (.01)	.01	.87 (.00)	.00	.87
0.5	0.5	.58 (.12)	.02	.58 (.12)	.03	.58 (.11)	.03	.75
	0.7	.72 (.11)	.01	.79 (.05)	.00	.80 (.02)	.00	.81
	0.9	.83 (.05)	.01	.87 (.00)	.00	.87 (.00)	.00	.87
1	0.5	.58 (.12)	.02	.58 (.12)	.03	.58 (.12)	.03	.75
	0.7	.73 (.11)	.01	.79 (.05)	.00	.80 (.02)	.00	.81
	0.9	.84 (.04)	.00	.87 (.00)	.00	.87 (.00)	.00	.87

Table 4: Mean sender expected utility and JSD of interlocutors' priors in 10^4 independent games using "preemptive" belief updates. U_{max} indicates the maximum expected utility reachable for a given P^*.

References

Arnold, K. and Zuberbühler, K. (2006). The alarm-calling system of adult male putty-nosed monkeys, cercopithecus nictitans martini. *Animal Behaviour*, 72(3):643–653.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., and Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42(1):1–22.

Barrett, J. and Zollman, K. J. (2009). The role of forgetting in the evolution and learning of language. *Journal of Experimental & Theoretical Artificial Intelligence*, 21(4):293–309.

Branigan, H. P., Pickering, M. J., Pearson, J., and McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, 42(9):2355–2368.

Brennan, S. E. and Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6):1482–1493.

Brennan, S. E. and Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, 1(2):274–291.

Brochhagen, T. (2015). Improving coordination on novel meaning through context and semantic structure. In *Proceedings of the Sixth Workshop on Cognitive Aspects of Computational Language Learning*, pages 74–82.

Camerer, C. F., Ho, T.-H., and Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898.

Chomsky, N. (2002). *On Nature and Language*. Cambridge University Press.

Chomsky, N. (2008). On phases. In *Foundational Issues in Linguistic Theory*, pages 132–166. MIT Press.

Clark, H. H. and Schober, M. F. (1992). Asking questions and influencing answers. In *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*, pages 14–48.

Clark, H. H. and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1):1–39.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431.

Dautriche, I. (2015). *Weaving an Ambiguous Lexicon*. PhD thesis, École Normale Supérieure.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, 88(4):848–881.

Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. In *Psychology of Learning and Motivation: Advances in Research and Theory*, pages 209–246. Elsevier BV.

Ferreira, V. S., Slevc, L. R., and Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, 96(3):263–284.

Fowler, C. A. and Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26(5):489 – 504.

Frank, M. C. and Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.

Franke, M. (2009). *Signal to Act: Game Theoretic Pragmatics*. PhD thesis, University of Amsterdam.

Franke, M. (2015). The evolution of compositionality and proto-syntax in signaling games. *Journal of Logic, Language and Information.*

Franke, M. and Degen, J. (2016). Reasoning in reference games: Individual- vs. population-level probabilistic modeling. *PloS one*, 11(5).

Franke, M. and Jäger, G. (2011). Bidirectional optimization from reasoning and learning in games. *Journal of Logic, Language and Information*, 21(1):117–139.

Franke, M. and Jäger, G. (2014). Pragmatic back-and-forth reasoning. In *Pragmatics, Semantics and the Case of Scalar Implicatures*. Nature Publishing Group.

Franke, M., Jäger, G., and van Rooij, R. (2011). Vagueness, signaling and bounded rationality. In *New Frontiers in Artificial Intelligence*, pages 45–59. Springer Science+Business Media.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., and Tylen, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, 23(8):931–939.

Garrod, S. and Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, 53(3):181–215.

Goodman, N. D. and Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11):818–829.

Greenough, A., Cole, G., Lewis, J., Lockton, A., and Blundell, J. (1998). Untangling the effects of hunger, anxiety, and nausea on energy intake during intravenous cholecystokinin octapeptide (CCK-8) infusion. *Physiology & Behavior*, 65(2):303–310.

Grice, P. (1975). Logic and conversation. In *Studies in the Ways of Words*, chapter 2, pages 22–40. Harvard University Press, Cambridge, MA.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the experimental analysis of behavior*, 13(2):243–266.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, (203):89–97.

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In Schiffrin, D., editor, *Meaning, Form and Use in Context*, pages 11 – 42. Georgetown University Press.

Huttegger, S. M., Skyrms, B., and Zollman, K. J. S. (2013). Probe and adjust in information transfer games. *Erkenntnis*, 79(S4):835–853.

Jaegher, K. D. and van Rooij, R. (2011). Strategic vagueness, and appropriate contexts. In *Language, Games, and Evolution*, pages 40–59. Springer Science+Business Media.

Jaegher, K. D. and van Rooij, R. (2013). Game-theoretic pragmatics under conflicting and common interests. *Erkenntnis*.

Juba, B., Kalai, A. T., Khanna, S., and Sudan, M. (2011). Compression without a common prior: An information-theoretic justification for ambiguity in language. In *Proceedings of the 2nd Symposium on innovations in computer science*.

Kim, M., Horton, W. S., and Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, 2(1).

Lewis, D. (1969). *Convention: A Philosophical Study*. Harvard University Press, Cambridge.

Luce, D. R. (1959). *Individual choice behavior: a theoretical analysis*. Wiley.

Mahowald, K., Fedorenko, E., Piantadosi, S. T., and Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, 126(2):313–318.

Mehta, J., Starmer, C., and Sugden, R. (1994). Focal points in pure coordination games: An experimental investigation. *Theory and Decision*, 36(2):163–185.

Metzing, C. and Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49(2):201–213.

Miller, G. A. (1951). *Language and Communication*. McGraw-Hill.

Motamedi, Y., Schouwstra, M., Smith, K., and Kirby, S. (2016). Linguistic structure emerges in the cultural evolution of artificial sign languages. In *The Evolution of Language: Proceedings of the 11th International Conference*.

O'Connor, C. (2014). The evolution of vagueness. *Erkenntnis*, 79(S4):707–727.

O'Connor, C. (2015). Ambiguity is kinda good sometimes. *Philosophy of Science*, 82(1):pp. 110–121.

Piantadosi, S. T., Tily, H., and Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9):3526–3529.

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

Pickering, M. J. and Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3):427–459.

Pickering, M. J. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02).

Reitter, D. and Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, 76:29–46.

Santana, C. (2014). Ambiguity in cooperative signaling. *Philosophy of Science*, 81(3):398–422.

Schelling, T. C. (1980). *The Strategy of Conflict*. Harvard University Press.

Spike, M., Stadler, K., Kirby, S., and Smith, K. (2016). Minimal requirements for the emergence of learned signaling. *Cognitive Science*, pages 1–36.

Steels, L. (1998). The origins of ontologies and communication conventions in multi-agent systems. *Agents and Multi-Agent Systems*, 1(2):169–194.

Sutton, R. S. and Barto, A. G. (1998). *Introduction to Reinforcement Learning*. MIT Press, Cambridge, MA, USA.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs: General and Applied*, 2(4):i–109.

van Rooij, R. and Sevenster, M. (2006). Different faces of risky speech. In *Game Theory and Pragmatics*. Springer.

Wasow, T. (2015). Ambiguity avoidance is overrated. In *Ambiguity*. Walter de Gruyter GmbH.

Zipf, G. (1949). *Human behavior and the principle of least effort*. Addison-Wesley Press.