Meaning in Communication Games*

Andreas Blume
Department of Economics
University of Arizona
Tucson, AZ 85721

May 22, 2022

Abstract

This paper addresses two related questions: How can we model the strategic use of a pre-existing language? And, how should we capture different degrees of sharing that language? The paper proposes an iterative procedure, interpreted as a mental process on part of the sender, that associates a set of equilibria, which we dub *language equilibria*, with every combination of a sender-receiver game and a pre-existing language. Every sender-receiver game has a language equilibrium. Language equilibrium makes sharp predictions about joint distributions over types and actions in common-interest games, in games with sender-preferred equilibria, and in games partial incentive alignment. This is the case when, as is frequently assumed, the language rich, but also when the language is impoverished. Predictions are sensitive to the degree to which language is shared. Importantly, unlike earlier suggestions for how to invoke the role of a pre-existing language in sender-receiver games, language equilibrium makes predictions about language use.

*I thank Joel Sobel, without meaning to implicate, for having provided inspiration. I am grateful for comments from Gerrit Bauch, Vince Crawford, Inga Deimen, Wojciech Olszewski, Marga Reimer, and Joel Sobel. I have benefitted from feedback received from seminar audiences at the University of Arizona and University of California – Riverside.

1 Introduction

The paper investigates the formation of *meanings in use* in communication games. We consider games in which a privately informed sender sends a message to a receiver who then takes an action. Payoffs depend only on the sender's private information and the receiver's action. Messages are cheap talk. Their meanings in use correspond to the information they convey and the actions they induce. If players have access to a pre-existing shared language, we expect these meanings in use to depend on both that shared language and on strategic considerations. The shared language provides the semantic meanings of messages, which, under the influence of incentives, are transformed into meanings in use.

We capture this transformation through a mental process on part of the sender, formalized as an iterative procedure. This mental process associates a set of equilibria with every combination of a sender-receiver game and a pre-existing language. Given the game and the pre-existing language the sender reasons until she reaches a rest point, an equilibrium, that rationalizes her message use.

We model the pre-existing language as a receiver strategy, interpreted as the receiver's nonstrategic interpretation of messages. It serves as a starting point for a sequence of best replies, with constraints on which best replies are admissible. This sequence converges to a limit set of strategies. The process then associates a set of equilibria with that limit set. Any such equilibrium is a *language equilibrium*, for the given pre-existing language. In the simplest case, when the mental process converges to a single strategy profile, it identifies a set of language equilibria that differ only in the receiver's responses to off-path messages.

Every sender-receiver game has a language equilibrium. Language equilibrium makes sharp predictions about joint distributions over types and actions in common-interest games, in games with sender-preferred equilibria, and in games partial incentive alignment. This is the case when, as is frequently assumed, the language rich, but also when the language is impoverished. Predictions are sensitive to the degree to which language is shared. Importantly, unlike earlier suggestions for how to invoke the role of a pre-existing language in sender-receiver games, language equilibrium makes predictions about language use.

'Meaning' has been given different meanings and some (Quine [27], Wittgenstein [34] in the interpretation of Kripke [20]) have expressed doubt about whether there are entities that are meanings at all. Is there something that is denoted by 'meaning,' and is that something shared? In this paper, we consider, inspired by Farrell [13], that there *is* a pre-existing language with well-defined meanings. We formalize this pre-existing language and different degrees of it being shared. We show that there is not necessarily a direct correspondence between a shared language and how it is used, that imperfectly sharing a language does not preclude it being useful, and that absence of a shared language rules out effective communication even when interests are perfectly aligned.

Part of what appears to be captured by meaning is reference.[1] We can, for example, think of a name for an object as referring to the named object, as in 'Wittgenstein' referring to (or denoting) Wittgenstein, the philosopher. Proper names, like 'Wittgenstein', 'John Nash', or

[1] Michaelson and Reimer [23] survey the topic of reference.

'Tucson', which refer to particular objects, appear to be simple enough. John Stuart Mill [24], for example, viewed a name's meaning straightforwardly as the object referred to by that name, its referent. Still, even something as seemingly simple as the meaning of proper names turns out to be problematic. Frege [14] observed that the statement "Hesperus is Phosphorus" ("The Evening Star is the Morning Star") would be uninformative if meaning were exhausted by reference – 'Hesperus' and 'Phosphorous' both denote Venus and therefore have the same reference. Frege therefore distinguished between the 'reference' and the 'sense' of a name. The sense of an expression is what we grasp when we hear it. It determines reference. Frege insisted on the objectivity of senses, thoughts held in common by mankind (see Miller [25]), as a prerequisite for communication being possible. This resonates with our finding that effective communication is incompatible with absence of a shared language.

There is a long tradition of tying meaning to truth.[2] In a theory of reference (Frege [14]) the reference of an expression in a sentence is its contribution to the truth value of that sentence: the sentence "Saguaros are green" is true, whereas "Saguaros are red" is false. Replacing "green" by "red" in the first sentence switches its truth value. The predicate "green" acts as a function that maps objects into truth values. Possible world semantics views the content/meaning of an expression as a function that indicates what that expression stands for in different states of the world. Carnap [7] refers to these functions as "intensions." Intensions map states of the world to truth values; the predicate "is rich" maps into 'true' in all worlds in which the person referred to is rich. Davidson [11] proposes a theory of semantics that is based on Tarski's [31] theory of truth. One question concerning communication games is whether and how meaning in communication games can be grounded in truth. A second related question is whether effective communication in games is necessarily truthful. Regarding the first question, the mental process we propose here is rooted in a language that lets the sender truthfully indicate her preferred receiver action. Regarding the second question, we find that outside of common-interest games message use in a language equilibrium is frequently systematically biased away from being truthful.

A widely accepted distinction is that between semantic meanings and meanings as mental entities, which are tied to the use of language. Semantic meanings relate expressions in a language to the world. In contrast, meanings as mental entities are psychological states that may be the speaker's intentions (Grice [17]) or beliefs (Lewis [21]). The distinction between semantic meanings and meanings as mental entities parallels somewhat the distinction we am interested in between the given meanings of a pre-existing language and the meanings in use that arising as equilibrium phenomena in a game.

From a purely game-theoretic perspective, the meaning of a message is captured by what players believe about each other's strategies: A receiver having a belief about the sender's strategy, after observing any message consistent with that strategy can form a belief about the sender's type. We can think of that belief as the meaning of that message to the receiver. Analogously, a sender having a belief about the receiver's strategy can anticipate how the receiver responds to any message and choose a message that induces an intended response. We can think of that intention as the meaning of that message to the sender. In this game-

[2]The following discussion relies heavily on Speaks [29].

theoretic account of meaning, there is no overt role for truth.

A problem with this purely game-theoretic conception of meaning is that, even if we commit to a solution concept (like Bayesian Nash equilibrium), it does not pin down the meanings of messages. Regardless of the incentive structure, for any solution we can find another game-theoretically equivalent solution by simply permuting messages. Anything that can be meant by one message can also be meant by any other message – messages are exchangeable.

A closely related problem with the purely game-theoretic conception of meaning is that in any communication situation in which interlocutors are given only a generic set of messages (which have no plausible association with states of the world or actions), we would not expect them to be able communicate. Having a large set of messages available is not enough to make effective communication possible if the messages do not already relate to the world in which the interlocutors interact. In terms of (Bayesian Nash) equilibria, in such an artificially constructed situation, regardless of the incentive structure (including those with perfectly aligned preferences), none of the equilibria in which the sender shares information are plausible.

To make effective communication plausible, an additional ingredient is needed. That ingredient is a pre-existing language. As Wittgenstein [34] (p.18) put it: "Can I say 'bububu' and mean 'If it doesn't rain I shall go for a walk'? – It is only in a language that I can mean something by something."

The role of a pre-existing language in making sense of behavior in communication games was first explored by Farrell [13]. Farrell posits that there is a rich language with commonly understood meanings. He appeals to richness to argue that for any equilibrium there are unused messages that can be activated to express any desired meaning. While messages are understood they are not necessarily believed. Farrell formulates a condition for an unused message to be credible relative to an equilibrium: a message is credible if the types indicated by that message gain relative to the equilibrium, and only those types gain. Farrell calls such a message a credible neologism, and proposes to reject equilibria for which there is a credible neologism. Equilibria that cannot be rejected are called neologism proof.

Neologism proofness predicts that there is effective communication in some games, including when interests are perfectly aligned. In that case, and others, it rejects the ever present "babbling equilibria" in which the sender's messages do not vary with the type and the receiver ignores messages. Rabin [28] gives a sui generis definition of when messages are credible, independent of a solution concept. According to his definition, for example, a message is credible for a set of types if all types in that set achieve their maximal payoff conditional on the message being believed and all other types receive their lowest payoff. Rabin's idea combines with both rationalizability and equilibrium.

Not all communication games have neologism proof equilibria. When such equilibria exist, neologism proofness places no constraints on message use *in* equilibrium. Rabin's credible message rationalizability makes sharp predictions about message use only for types who have credible messages. In contrast, the language equilibrium concept we propose guarantees existence and predicts message use.[3]

It is intuitive, and Blume, DeJong, Kim and Sprinkle [4] demonstrate experimentally, that there are regularities in the use of exogenously given message meanings if a meaningful language is available. Crawford [10] proposes to account for such regularities with a level-k model that is anchored in truthful behavior by senders. Level 0 senders are truthful; level 0 receivers best respond to level 0 senders; level 1 senders best respond to level 0 receivers; level 1 receivers best respond to level 1 senders; etc. Cai and Wang [6] and Wang, Spezio, and Camerer [33] conduct experiments on sender-receiver games and show that Crawford's level-k model applied to these games has explanatory power. It captures that senders are excessively truthful, receivers are excessively credulous, communication varies systematically with the bias, and senders inflate messages relative to truthfulness.

Truth matters, although in different ways, in the approaches of Farrell, Rabin, and Crawford. Farrell's credible neologisms are truthful statements. Rabin adds truth-telling as a behavioral assumption, capturing the idea that agents will tell the truth as long as that is consistent with incentives. Crawford anchors his level-k analysis in truth by assuming that level-0 senders are truthful. The approach taken here is similar to that taken by Crawford in that it anchors a mental process in truthful use of a pre-existing language and allows for systematic departures from truth at rest points of the mental process. A key difference with Crawford is the focus on making equilibrium predictions.

The goal of the present paper is to leverage the power of iterative reasoning to select equilibria and predict language use. It aims to tether strategic meaning, as expressed in message use, to semantic meaning, as given by a pre-established language. It proposes a general model that anchors meaning in a language, respects the strategic motives of interlocutors, predicts equilibrium behavior for rational players, satisfies existence in all games, and predicts message use for all types of the sender.

The paper employs four ideas to link (a pre-existing) language with its strategic use: (1) (*Anchoring*) iterating best replies from the language; (2) (*Non-proliferation*) containing the proliferation of best replies (by provisionally dropping unused messages, minimizing message use, and adjusting best replies only when necessary); (3) (*Expansion*) minimally expanding any limit set of strategies that is reached this way to the point where it includes a best reply to every belief concentrated on that set and focussing on equilibria belonging to that expansion; and, (4) (*Restoration*) restoring provisionally eliminated message in a way that extends the equilibria obtained to the entire game. A *language equilibrium* (relative to the pre-existing language in question) is any equilibrium identified in this manner.[4]

[3]Like Farrell and Rabin, Olszewski [26] explores the implications of a rich-language assumption: Given an equilibrium, there is a set of beliefs the sender can induce in that equilibrium. Consider adding messages that would permit the sender to induce additional beliefs. For any equilibrium there is a largest set of beliefs that one can generate in this manner without tempting the sender to deviate from her equilibrium strategy. If this set of beliefs is larger for equilibrium E than for equilibrium F, equilibrium E is said to have a *richer language*. The criterion of having a richer language selects among equilibria and equilibria with maximally rich languages exist. Note that language viewed this way is tied to an equilibrium. The question of how a pre-existing language with fixed semantic meanings is used does not arise.

[4]When the language is common knowledge restoration is always possible. When players do not share a

The proposed iterative procedure that links the pre-existing language to equilibria of the game is meant to capture the sender's deliberation: She contemplates what to say in a given situation. She latches on to what seems natural according to the pre-existing language, reflects on strategic implications, and stops when she has reached a point where everything coheres. The assumptions that unused messages get provisionally eliminated, that message use is minimized, and that message adjustments are only contemplated if they lead to strict improvements help ensure that the deliberation comes to a conclusion – the sender is aware that at some point she needs to speak and stop reflecting. Unlike in learning models, here the language and players' beliefs about language are fixed. The proposed iterative procedure is conceived as a mental process that determines the sender's message on a given occasion.

We model the pre-existing language as a function that assigns receiver interpretations to messages. In addition to addressing the question of how that function helps determine message use when it is common knowledge, the paper considers imperfectly shared languages. Players may be uncertain about the language or have private information about it. Languages being imperfectly shared imposes constraints on communication. The model uses translations as a device for capturing these constraints. A translation is a mapping from sender messages to receiver messages and is drawn from a set of possible translations. A translation limits communication options if it is not an injective function; there may be uncertainty about which translation has been drawn; and, sender and receiver may have private information about which translation has been drawn. The translation apparatus is flexible enough to accommodate (complete) absence of a common language, as in Crawford and Haller [9], gradations of language sharing, as well as uncertainty and private information about language constraints, as in Blume and Board [5], and Giovannoni and Xiong [15].

2 An informal introduction: examples

Consider a game, Game 1, between a sender and a receiver in which the sender's payoff type t belongs to the set $T = \{t_1, t_2\}$ and the receiver takes actions a in the set $A = \{a_1, a_2, a_3\}$. After privately observing her payoff type, the sender sends a message m from the message space $M = \{m_1, m_2, m_3\}$ to the receiver. In response to the sender's message, the receiver takes an action $a \in A$. Payoff types are equally likely and payoffs from any combination (t, a) of a payoff type t and an action a are given in Figure 1, with the first entry denoting the sender's payoff and the second entry denoting the receiver's payoff.

In addition to this standard structure of a sender-receiver game, assume that sender and receiver have a common language $\lambda : M \to A$, with $\lambda(m_i) = a_i$, $i = 1, 2, 3$. The language gives the semantic meaning of messages. It can be interpreted as the conventional way of referring to the receiver's actions, when incentives are of no concern.

The solution concept proposed in this paper, *language equilibrium*, is meant to capture the sender's reasoning about which message to send, given her type. The basic idea is that

common language restoration may be an issue. For that reason, we declare every equilibrium a language equilibrium if restoration is impossible for any candidate equilibrium obtained via anchoring, non-proliferation, and expansion.

	a_1	a_2	a_3
t_1	3,3	0,0	1,2
t_2	0,0	-1,3	1,2

Figure 1: Dropping messages

starting with the language λ, the sender iterates pure-strategy best replies until she reaches an equilibrium. In order to make this work, it is necessary to deal with a number of issues that may derail convergence: these include how to handle unused messages and how to deal with situations in which the iteration settles on a set instead of a single strategy profile, e.g., by reaching a cycle.

In Game 1 the sender's unique best reply against the receiver strategy $r_1 = \lambda$, defined by the language λ, is given by $s_1 = (t_1 \mapsto m_1, t_2 \mapsto m_3)$. Notice that s_1 does not use message m_2. The iterative procedure we will use to define *language equilibrium* provisionally drops unused messages. With message m_2 out of the picture, the receiver has a unique best reply $r_2 = (m_1 \mapsto a_1, m_3 \mapsto a_2)$ against s_1. The sender's unique best reply to r_2 is $s_2 = (t_1 \mapsto m_1, t_2 \mapsto m_1)$. The unused message m_3 is (provisionally) dropped, the receiver's unique best reply in the game without messages m_2 and m_3 is the pooling action a_3, and the iterative procedure has converged. At this point messages m_2 and m_3 are restored and a language equilibrium is defined as any equilibrium of the original game in which s_2 is the sender strategy and the receiver responds to message m_1 with action a_3.[5]

A few points are worth noting. First, in Game 1 Farrell's neologism-proofness test rejects all equilibria: Every equilibrium is a pooling equilibrium, with the receiver taking action a_3 on the equilibrium path, and given any such equilibrium the message "I am type t_1" is a credible neologism. Type t_1 prefers this message to be believed rather than receiving the pooling payoff and type t_2 prefers the pooling payoff to having this message believed. Second, there is no credible message profile, as defined by Rabin: Type t_1 would like to be identified, but type t_2 would have reason to mimic type t_1; type t_2 prefers not to be identified; and, type t_1 would not receive her maximal payoff if both types identified themselves as belonging to $\{t_1, t_2\}$. Third, level-k reasoning would reach the same conclusion as language equilibrium, for high enough levels and with suitable assumptions for how to deal with unused message. Finally, language equilibrium is consistent with equilibrium by construction and, in this game, makes a sharp prediction about message use: both types send message m_1.

What if the iterative procedure just described instead of converging enters a cycle? Game 2 with the payoff structure in Figure 2 illustrates this problem and how we address it.

As before, suppose that the two payoff types t_1 and t_2 are equally likely, that the message space is $M = \{m_1, m_2, m_3\}$, and players have a common language λ with $\lambda(m_i) = a_i$,

[5]In games with a common language this type of restoration of message is always possible. Simply have the receiver's off-equilibrium responses to messages coincide with one of the responses to a message that is sent in equilibrium.

	a_1	a_2	a_3
t_1	0,9	9,0	0,8
t_2	9,0	0,9	0,8

Figure 2: A role for Prep-sets

$i = 1, 2, 3$.

The sender's unique best reply against the receiver's strategy $r_1 = \lambda$, defined by the language λ, is the strategy $s_1 = (t_1 \mapsto m_2, t_2 \mapsto m_1)$. Consider the *reduced game*, in which the unused message m_3 is provisionally dropped. The receiver's unique best reply against the sender's strategy s_1 in the reduced game is the strategy $r_2 = (m_1 \mapsto a_2, m_2 \mapsto a_1)$. Iterating further generates a sequence of best replies that are unique at every step and form a cycle. Denote the set of pure strategies that support this cycle by $\mathsf{S}' \times \mathsf{R}'$. The set of strategies $\mathsf{S}' \times \mathsf{R}'$ does not support an equilibrium of the reduced game, in either pure or mixed strategies. To satisfy the desideratum of having the iterative procedure reach rest points that are equilibria, the procedure expands the set $\mathsf{S}' \times \mathsf{R}'$. Voorneveld [32] defines a *prep set* as a set of pure strategy profile that includes a best reply to every belief concentrated on that set. This inspires the definition of an $\mathsf{S}' \times \mathsf{R}'$-prep set as a set of pure strategy profiles in the reduced game that includes $\mathsf{S}' \times \mathsf{R}'$ as well as a best reply to every belief concentrated on that set. The procedure expands $\mathsf{S}' \times \mathsf{R}'$ to a *minimal* $\mathsf{S}' \times \mathsf{R}'$*-prep set.* Minimality is with respect to set inclusion. A *minimal* $\mathsf{S}' \times \mathsf{R}'$-*prep set* does not strictly contain another $\mathsf{S}' \times \mathsf{R}'$-prep set.

Given a receiver belief that assigns equal probability to all sender strategies in S' the receiver's unique best reply in the reduced game is the strategy $\tilde{r} = (m_1 \mapsto a_3, m_2 \mapsto a_3)$. Therefore, the strategy $\tilde{r}$ must be in any $\mathsf{S}' \times \mathsf{R}'$-prep set. Indeed, once that strategy is included we have a minimal $\mathsf{S}' \times \mathsf{R}'$-prep set and that set includes (a continuum of) equilibria of the reduced game. In any such equilibrium both messages m_1 and m_2 are used with positive probability and the receiver responds to both messages with the action a_3. Finally, we can restore the unused message m_3 to the game. Therefore, in every language equilibrium the sender uses both messages m_1 and m_2 and the receiver responds to all three messages with the pooling action a_3.

In Game 2 pooling, the only outcome supported by an equilibrium, passes Farrell's test. Neologism proofness does not, however, commit to which message or messages the sender uses. Likewise, since there is no credible message profile, Rabin's solution make no predictions about message use. A level-k analysis anchored at the language λ is inconclusive without additional commitments to how to treat unused messages and to the number of levels. Language equilibrium arrives at a sharp prediction: both message m_1 and m_2, and only those messages, will be used and the receiver responds to all messages with action a_3.

In both of our examples thus far the language has been common knowledge. The next example suggests a way of modeling lack of common knowledge of the language and explores the consequences for effective communication and message use.

	a_1	a_2	a_3
t_1	10,10	9,0	0,9
t_2	9,0	10,10	0,9

Figure 3: Uncertainty about language

Consider Game 3 with the payoff structure in Figure 3, two equally likely payoff types t_1 and t_2, and a message space $M = \{m_1, m_2\}$. Rabin [28] uses the example to demonstrate that the credibility of one message may depend on the credibility of other messages; Stalnaker [30] elaborates on this by raising the possibility of "ignorance or error about credibility." We want to use this example to investigate the possibility and consequences of language not being perfectly shared, in the sense that there may be uncertainty about how messages are interpreted.

Rabin points out that the message m_1, interpreted as "my type is t_1" is not credible, unless m_2 interpreted as "my type is t_2" is credible: if m_1 were credible but there was sufficient doubt about the credibility of m_2, then type t_2 would prefer to send m_1 (to receive the payoff 9 rather than the pooling payoff 0), undermining the credibility of m_1. In Rabin's case, with the assumption that the language is common knowledge, this ends up being unproblematic because the messages are jointly credible.

Suppose, as before, that there is a language λ with $\lambda(m_i) = a_i$, $i = 1, 2$, that corresponds to the receiver's interpretation of messages. Now, however, we want to capture the possibility that when sending a message the sender is uncertain about how it is interpreted. In order to have this be a material constraint, we add to the game a set of *translations* and a probability distribution over that set. A translation $\theta : M \to M$ maps intended messages into interpreted messages.

Specifically, suppose the sender has doubts about her ability to convey to the receiver her wish that action a_1 be taken. She believes that there is a small chance that whatever she says will be interpreted as asking for action a_2. Formally, there are two translations θ' and θ'', defined by $\theta'(m) = m$ and $\theta''(m) = m_2$ for both $m \in M$. Assume that there is common prior μ over translations with $\mu(\theta'') = p$, where p satisfies $\frac{1}{9} < p < 1$.

It may help to have in mind the following scenario: The sender wants to either express qualified skepticism about a scientific claim, t_1, or provide a qualified endorsement, t_2. The receiver interprets the sender's messages as either qualified skepticism, a_1, a qualified endorsement, a_2, or pays no attention, a_3. The appropriate qualifications in the possible statements the sender could make require careful wording and the sender may worry that despite her best effort her statements are misinterpreted. In addition, she worries that the receiver is aware of the possibility of misinterpretation and therefore pays no attention.

The iterative procedure singles out a language equilibrium as follows. The sender's unique best reply against the receiver's strategy $r_1 = \lambda$, defined by the language λ, is the strategy

$s_1 = (t_1 \mapsto m_1, t_2 \mapsto m_2)$. Given that the sender uses the strategy s_1, the receiver's posterior probability that the sender's type is t_2 after observing message m_2 equals $\frac{1}{p+1}$. Therefore, as long as $p > \frac{1}{9}$, the receiver has a unique best reply $r_2 = (m_1 \mapsto a_1, m_2 \mapsto a_3)$ to the sender's strategy s_1. Against r_2, the sender has a unique best reply $s_2 = (t_1 \mapsto m_1, t_2 \mapsto m_1)$. At that point message m_2 is dropped from the iteration. In any λ-equilibrium, the sender sends message m_1 exclusively. In order to have a λ-equilibrium, it is necessary that the receiver responds to message m_2 also with action a_3. This implies that there is a unique λ-equilibrium strategy profile $(\sigma, \rho) = ((t_1 \mapsto m_1, t_2 \mapsto m_1), (m_1 \mapsto a_3, m_2 \mapsto a_3))$.

Note that we get a sharp prediction about message use. The sender sends message m_1 in both states of the world. That message is natural for type t_1 to send. Type t_2 sends it out of concern for otherwise being ignored. In the end, the sender expects to be ignored regardless of the message send.

3 Setup

I consider sender-receiver games with a sender, S, who has private information about a payoff-relevant state, and a receiver, R, who takes an action that affects both players' payoffs. Prior to the receiver taking his action the sender sends a message to the receiver. There is a finite payoff type space T, a finite action space A, and a finite message space M.[6] For any (finite) set X, $\Delta(X)$ is the set of probability distributions over X. Players have a common prior $\pi \in \Delta(T)$ over the payoff type space, with $\pi(t) > 0$ for all $t \in T$. Players' payoffs $u^i(t, a)$, $i = S, R$, depend only on the sender's payoff type $t \in T$ and the receiver's action $a \in A$.

I refer to the structure described thus far as the *base game*. In the base game messages have no semantic meanings and are received as sent. The games considered modify the base game by adding a *language*, which endows messages with semantic meanings, and by introducing *translations*, which loosen the link between sent and received messages.

A *language* $\lambda : M \to A$ represents the receiver's non-strategic interpretation of messages. While the existence of such a language is assumed to be common knowledge, the sender may be uncertain about that language. She may, for example, know that for every action $a \in A$ the receiver has a term $m_a \in M$ that refers to that action but may have no knowledge of which term refers to which action. In that case every language $\lambda' = \lambda \circ \theta$, where $\theta : M \to M$ is a permutation of M, is just as likely as the language λ from the sender's perspective.

I assume that sent messages are subject to a *translation* $\theta : M \to M$. When the sender sends a message $m \in M$, the receiver observes the message $\theta(m) \in M$. Translations can but need not be permutations. Translations are drawn from a common prior distribution μ over a set of translations Θ, with $\mu(\theta) > 0$ for all $\theta \in \Theta$, and are not directly observed by either the sender or the receiver.[7]

[6] Finiteness of the message space is not essential; in fact, one of the merits of the approach here is that the cardinality of the message space is largely irrelevant. With a common language, for example, having any number of synonyms for messages or having the message space be unbounded does not present a problem.

[7] Since M is finite, so is Θ.

From the standpoint of an equilibrium analysis, adding a language to the base game has no consequence. The language can always be ignored. Subjecting messages to translations, in contrast, does impact the equilibrium structure whenever translations are not bijections or there is uncertainty about the translation. It is through the solution concept that we propose, which uses the language to select among equilibria, that language and translations become intertwined. Uncertainty about the translation becomes uncertainty about language. That way language and uncertainty about that language jointly determine the (set of) equilibria that are selected.

Each player i receives a private signal h^i about the translation from a finite set of signals H^i. The signal pair $h = (h^S, h^R)$ is generated by a conditional probability system $\eta : \Theta \to \Delta(H^S \times H^R)$ that assigns strictly positive probability to each pair of signals $(h^S, h^R) \in H^S \times H^R$, i.e., $\text{Prob}[(h^S, h^R) \times \Theta] > 0$ for all $(h^S, h^R) \in H^S \times H^R$. Denote player i's posterior probability of $\theta \in \Theta$ given his signal h^i by $\eta^i(\theta|h^i)$. We say that player i *learns the translation* if for each signal $h^i \in H^i$ there is a translation θ_{h^i} such that $\eta^i(\theta_{h^i}|h^i) = 1$; that is, player i's signal reveals the translation.

After obtaining her private information $\big(t, h^S\big) \in T \times H^S$ the sender sends a message $m \in M$ to the receiver. After observing $(\theta(m), h^R) \in M \times H^R$ the receiver takes an action $a \in A$. A pure strategy $s : T \times H^S \to M$ of the sender maps pairs of payoff states and sender signals about the translation into messages. A pure strategy $r : M \times H^R \to A$ of the receiver maps pairs of messages and receiver signals about the translation into actions. We denote the sender's set of pure strategies by S and the receiver's set of pure strategies by R. The corresponding sets of mixed strategies are $\Sigma_S = \Delta(\mathsf{S})$ and $\Sigma_R = \Delta(\mathsf{R})$, with typical elements $\sigma \in \Sigma_S$ and $\rho \in \Sigma_R$. Expected payoffs as a function of mixed strategy profiles (σ, ρ) are denoted by $U^i(\sigma, \rho)$.

If the language λ is surjective, then λ is a *rich language*. If Θ is a set of permutations of M, then the language λ is *accessible*.[8] If $\Theta = \{\theta\}$ is a singleton set, then the language λ is a *determinate language*. A language that is both accessible and determinate is a *shared language*. A shared language for which the translation θ is the identity mapping is a *common language*. A language that is rich, accessible and determinate is a *rich shared language*, and if θ is the identity, it is a *rich common language*. If Θ is the set of all permutations of M, μ is the uniform distribution on Θ, and $\eta(\theta') = \eta(\theta'')$ for all θ', θ'', then we have *absence of a shared language*.

4 Language Equilibrium

The key idea for what we propose is exceedingly simple: starting with the language λ, iterate pure-strategy best replies, changing strategies only when this increases payoffs, while at each step eliminating unused messages. This generates a sequence of strategy profiles in games with reduced message spaces. There is a limit game, with a reduced message space, and a

[8] The properties of Θ, μ and η^i, $i = S, R$, affect the use players can make of the language λ. In a broader sense they are part of the language. For that reason, and to save on notation, we refer to accessibility as an attribute of the language, and similarly for the attributes defined below.

limit set of strategies in that "reduced game" that recur infinitely often. If the limit set of strategies is a singleton, we have an equilibrium in the reduced game. Under some conditions (e.g., if there is no uncertainty about translations) the equilibrium in the reduced game can be extended to an equilibrium in the original game – by adding appropriate receiver responses to off-path messages. If this is the case, we have a *language equilibrium* for the language λ, or a λ-equilibrium.

Two issues have to be addressed. The procedure just described need not converge to a single strategy profile and, even if it does, need not generate a profile that can be extended to an equilibrium of the original game. To deal with the first issue, we take limit sets of pure strategy profiles that are reached under this procedure (in the reduced games), $\mathsf{S}' \times \mathsf{R}'$, which always exist, and consider minimal sets of strategy profiles P that contain $\mathsf{S}' \times \mathsf{R}'$ and a best reply for each player to every belief that is concentrated on P. Each P is a *prep set*, as defined by Voorneveld [32] that contains $\mathsf{S}' \times \mathsf{R}'$, and is minimal among all such sets. These sets always exist and contain an equilibrium of the reduced game. If such an equilibrium can be extended to an equilibrium in the original game, we designate the extension as a *language equilibrium* of the original game. Finally, if none of the equilibria of reduced games identified by this procedure can be extended to an equilibrium in the original game (which is only a potential issue when there is uncertainty about translations), this is taken to indicate that the language λ does not single out any of the equilibria of the original game. In that case, all of the equilibria of the original game are designated as *language equilibria* of the original game (for the language λ).

The proposed iterative procedure provisionally eliminates messages. This motivates introducing reduced games on subsets of the original message space. For any subset M^0 of the message space M, define $\Theta(M^0) = \{m' \in M | \exists m \in M^0, \theta \in \Theta \text{ such that } m' = \theta(m)\}$ as the set of all messages that are possible for the receiver to observe if the sender is restricted to sending messages in M^0. For each $M^0 \subseteq M$, use $\Gamma(M^0)$ to denote the game in which the sender is restricted to sending messages in M^0 and the receiver responds to messages in $\Theta(M^0)$. In the game $\Gamma(M^0)$, the sender's set of pure strategies is $\mathsf{S}(M^0)$ and the receiver's set of pure strategies is $\mathsf{R}(M^0)$. The corresponding sets of mixed strategies are $\Sigma_S(M^0)$ and $\Sigma_R(M^0)$.

For any game $\Gamma(M^0)$ and any set of strategy profiles $\mathsf{S}' \times \mathsf{R}' \subseteq \mathsf{S}(M^0) \times \mathsf{R}(M^0)$, a set $P = P_S \times P_R \subseteq \mathsf{S}(M^0) \times \mathsf{R}(M^0)$ is an $\boldsymbol{\mathsf{S}' \times \mathsf{R}'}$***-prep set*** if it satisfies:

1. $\mathsf{S}' \times \mathsf{R}' \subseteq P_S \times P_R$; and,

2. P_i contains a best reply in $\Gamma(M^0)$ to every belief concentrated on P_{-i} for $i = S, R$.

In our analysis, the sets $\mathsf{S}' \times \mathsf{R}'$ will be limit sets reached by iterating from the language λ. There may not be an equilibrium strategy profile of the game $\Gamma(M^0)$ that is supported on $\mathsf{S}' \times \mathsf{R}'$. This motivates considering $\mathsf{S}' \times \mathsf{R}'$-prep sets that are minimal with respect to set inclusion. Minimal $\mathsf{S}' \times \mathsf{R}'$-prep sets are the smallest expansions of the limit sets $\mathsf{S}' \times \mathsf{R}'$ to sets that satisfy a best-reply property and, as a result, support equilibrium strategy profiles of $\Gamma(M^0)$. For any game $\Gamma(M^0)$ and any set of strategy profile $\mathsf{S}' \times \mathsf{R}' \subseteq \mathsf{S}(M^0) \times \mathsf{R}(M^0)$, denote the collection of all minimal $\mathsf{S}' \times \mathsf{R}'$-prep sets P, by $\mathcal{P}(\mathsf{S}' \times \mathsf{R}')$.

Slightly abusing notation, we will use the same notation for sender strategies defined for different codomains (i.e., message spaces) but identical images. More formally, for any $M' \subseteq M$ and any sender strategy $s : T \times H^S \to M'$ in $\Gamma(M')$, if $s(T \times H^S) = M'' \subset M'$, we will also use s to denote the strategy $\tilde{s} : T \times H^S \to M''$ in $\Gamma(M'')$ that is defined by $\tilde{s}(t, h^S) = s(t, h^S)$ for all $(t, h^S) \in T \times H^S$.

Central to the definition of a language equilibrium is an iterative reasoning process anchored at the language $\lambda : M \to A$. To capture this reasoning process we will define a *λ-path*, which is a sequence $(M_k, s_k, r_k)_{k=1}^{\infty}$ of triples, each consisting of a message space, a pure sender strategy s_k, and a pure receiver strategy r_k. Each sender strategy s_k is a best response to the receiver strategy r_k; r_k is a best response to s_{k-1}; and, message space M_k is the set of messages used by s_{k-1}. In addition, $M_1 = M$ and $r_1 = \lambda$. In the formal definition of λ-paths we make use the following notation: For any pure receiver strategy r of the game $\Gamma(M')$, $M' \subseteq M$, define $\mathrm{BR}_S(r)$ as the set of pure-strategy best replies of the sender in $\Gamma(M')$. Likewise, define BR_R as the pure strategy best reply correspondence of the receiver (in the relevant game). For the sender, in addition, define $\underline{\mathrm{BR}}_S(r)$ as the set of sender best replies (in the relevant game) that are minimal with respect to the sets of messages used; that is $s \in \underline{\mathrm{BR}}_S(r)$ if $s \in \mathrm{BR}_S(r)$ and there is no strategy $s' \in \mathrm{BR}_S(r)$ that uses a strict subset of the set of messages used by s. We refer to the set $\underline{\mathrm{BR}}_S(r)$ as the sender's *minimal-message best replies* to strategy r of the receiver.

Best replies in the definition of a λ-path are "sticky": for either player, if a strategy from the previous iteration remains a best reply, it is retained in the current iteration. We also assume that the sender uses minimal-message best replies: when given a choice between two best replies whose message sets are nested, she picks the one with the smaller set of messages. The first of these properties rules out spurious iterations and helps minimize the sets of strategies reached in the limit. The second property rules out that the receiver makes spurious distinctions among sender types that have identical best replies.[9]

Definition 1 *A sequence $(M_k, s_k, r_k)_{k=1}^{\infty}$ with $M_k \subseteq M$, $s_k \in \mathsf{S}(M_k)$ and $r_k \in \mathsf{R}(M_k)$ for all $k \geq 1$ is a **$\boldsymbol{\lambda}$-path** if*

1. *$M_1 = M$ and $r_1(m, h^r) = \lambda(m)$ for all $m \in \Theta(M_1)$ and all $h^R \in H^R$;*
2. *for all k, $s_k \in \underline{BR}_S(r_k)$ in $\Gamma(M_k)$ – in addtion, if $k > 1$ and $s_{k-1} \in \underline{BR}_S(r_k)$ in $\Gamma(M_k)$, then $s_k = s_{k-1}$;*
3. *$M_{k+1} = s_k(T \times H^S)$; and,*
4. *for all k, $r_{k+1} \in BR_R(s_k)$ in $\Gamma(M_{k+1})$ – in addition, if $r_k \in BR_R(s_k)$ in $\Gamma(M_k)$, then $r_{k+1}(m, h^R) = r_k(m, h^R)$ for all $m \in \Theta(M_{k+1})$ and all $h^R \in H^R$.*

[9]In addition, one might want to define a λ-path in terms of the game in which actions that are dominated for the receiver have been eliminated. This would not affect any of our results, but does make a difference in examples. A sender who believes that the receiver is rational should not expect to be able to induce dominated receiver actions and should therefore never use such an expectation as a starting point for her deliberations.

To deal with cases in which λ-paths do not converge, we introduce λ-sets. A λ-set is a minimal prep set that contains the limit set of strategies reached by a λ path in the game $\Gamma(M^0)$, where M^0 is the limit message space reached through successive deletion of unused messages.

Definition 2 *A set of pure strategy profiles* $\tilde{\mathsf{S}}\times\tilde{\mathsf{R}}$ *in* $\Gamma(M^0)$ *is a* ***λ-set for $\Gamma(M^0)$*** *if there is a λ-path* $(M_k, s_k, r_k)_{k=1}^{\infty}$, $M^0 = \bigcap_{k=1}^{\infty} M_k$ *and* $\tilde{\mathsf{S}}\times\tilde{\mathsf{R}} \in \mathcal{P}\left(\{\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{s_k\}\} \times \{\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{r_k\}\}\right)$ *for the game* $\Gamma(M^0)$.

By construction, a λ-set for a game $\Gamma(M^0)$ contains the support of an equilibrium in $\Gamma(M^0)$. Any such equilibrium, we refer to as a λ-profile for $\Gamma(M^0)$.

Definition 3 *A strategy profile* (σ, ρ) *in* $\Gamma(M^0)$ *is a* ***λ-profile for $\Gamma(M^0)$*** *if there is a λ-set* $\tilde{\mathsf{S}} \times \tilde{\mathsf{R}}$ *for* $\Gamma(M^0)$ *with* $(\sigma, \rho) \in \Delta(\tilde{\mathsf{S}}) \times \Delta(\tilde{\mathsf{R}})$ *such that* (σ, ρ) *is an equilibrium strategy profile in* $\Gamma(M^0)$.

Once we have a λ-profile for some game $\Gamma(M^0)$, the question arises whether we can extend it to the original game by picking suitable receiver responses after the messages that have zero probability to be observed by the receiver under the λ-profile. Conversely, and equivalently, we can ask whether there is a way of reducing an equilibrium strategy profile of the original game to a λ-profile of a game with a reduced message space. A strategy profile (σ, ρ) in the original game $\Gamma(M)$ is a λ-equilibrium profile if it is an equilibrium profile in $\Gamma(M)$ and there is a message space $M^0 \subseteq M$ such that after restricting the receiver strategy to messages that can be observed in $\Gamma(M^0)$ it is a λ-profile in $\Gamma(M^0)$. For any receiver strategy ρ in $\Gamma(M)$ and any $M^0 \subseteq M$ let $\rho_{|M^0}$ denote the restriction of ρ to messages that can be received with positive probability in $\Gamma(M^0)$.

Definition 4 *An equilibrium strategy strategy profile* (σ, ρ) *in* $\Gamma(M)$ *is a* ***λ-equilibrium profile*** *if either*

1. *there exists* $M^0 \subseteq M$ *such that* $(\sigma, \rho_{|M^0})$ *is a λ-strategy profile in* $\Gamma(M^0)$*; or,*
2. *there is no equilibrium strategy profile* (σ', ρ') *in* $\Gamma(M)$ *and* $M' \subseteq M$ *such that* $(\sigma', \rho'_{|M'})$ *is a λ-strategy profile in* $\Gamma(M')$.

For each player, a strategy that is part of a λ-equilibrium profile is a λ-equilibrium strategy.

The first condition, which we refer to as *reducibility*, can always be satisfied for some equilibrium in a class of games that includes all games with a common language. When there is uncertainty about the language, however, it may be impossible to satisfy reducibility. In that case, the second condition ensures existence.

The following preliminary result establishes existence of λ-profiles. It is a simple consequence of the fact that a λ-path either converges or reaches non-singleton limit set and that any minimal prep set containing that limit set supports an equilibrium of the limit game $\Gamma(M^0)$.

Lemma 1 *For every game $\Gamma(M)$ and every language $\lambda : M \to A$ there exists $M^0 \subseteq M$ and a λ-profile for $\Gamma(M^0)$.*

Proof: Existence of a λ-path follows from the fact that all games $\Gamma(M')$ with $M' \subseteq M$ are finite: Given any $M_k \subseteq M$ and any receiver strategy in $\mathsf{R}(M_k)$ (sender strategy in $\mathsf{S}(M_k)$) there exists a pure-strategy best reply for the sender (receiver) since the set of pure strategies $\mathsf{S}(M_k)$ ($\mathsf{R}(M_k)$) is finite. Given any $M_k \subseteq M$ and any pure sender strategy s_k in $\Gamma(M_k)$ the set $M_{k+1} = s_k(T \times H^S)$ is well defined.

Given a λ-path $(M_k, s_k, r_k)_{k=1}^{\infty}$, the set $M^0 = \bigcap_{k=1}^{\infty}\{M_k\}$ is well defined and non-empty since each M_k is a finite non-empty subset of M_{k-1}. For sufficiently large k, each s_k is a pure strategy in $\Gamma(M^0)$. Since there are finitely many such strategies, at least one must appear infinitely often. Hence, the set $\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{s_k\}$ is well-defined and non-empty. Likewise, for sufficiently large k, each r_k is a pure strategy in $\Gamma(M^0)$. Since there are finitely many such strategies, at least one must appear infinitely often. Hence, the set $\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{r_k\}$ is well-defined and non-empty.

Trivially, the set $P_S \times P_R = \mathsf{S}(M^0) \times \mathsf{R}(M^0)$ satisfies $\{\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{s_k\}\} \times \{\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{r_k\}\} \subseteq P_S \times P_R$ and for every belief concentrated on P_{-i} contains a best reply in P_i, $i = S, R$. Finiteness then implies that there must be a minimal set with that property. Hence, there is a λ-set $\tilde{\mathsf{S}} \times \tilde{\mathsf{R}}$ for the game $\Gamma(M^0)$.

Since the λ-set $\tilde{\mathsf{S}} \times \tilde{\mathsf{R}}$ is a Prep set for the game $\Gamma(M^0)$, it contains an equilibrium in mixed strategies of $\Gamma(M^0)$. □

The next result shows that whenever the set of translations is a singleton, every λ-equilibrium profile can be obtained as an extension of a λ-profile to the entire game.

Lemma 2 *For every game $\Gamma(M)$ and every determinate language $\lambda : M \to A$, if $(\tilde{\sigma}, \tilde{\rho})$ is a λ-profile in $\Gamma(M^0)$, there is a λ-equilibrium profile (σ, ρ) in $\Gamma(M)$ with $(\tilde{\sigma}, \tilde{\rho}) = (\sigma, \rho_{|M^0})$.*

Proof: Recall that λ is a determinate language if the set of translations contains a single element, θ. Let $(\tilde{\sigma}, \tilde{\rho})$ be a λ-profile in $\Gamma(M^0)$. Since $(\tilde{\sigma}, \tilde{\rho})$ is an equilibrium profile in $\Gamma(M^0)$, the receiver strategy $\tilde{\rho}$ specifies a best reply to all messages in $\theta(M^0)$. Extend the receiver strategy $\tilde{\rho}$ from the game $\Gamma(M^0)$ to the game $\Gamma(M)$ by letting $\rho(m) = \tilde{\rho}(m^0)$ for all $m \in \theta(M \setminus M^0)$ and some $m^0 \in \theta(M^0)$. In the game $\Gamma(M)$, if the receiver uses the strategy ρ, then every action the sender can induce by sending a message in $M \setminus M^0$ she can also induce by sending a message in M^0. Hence, if we let $\sigma = \tilde{\sigma}$, (σ, ρ) is an equilibrium strategy profile for the game $\Gamma(M)$ with $(\sigma, \rho_{|M^0}) = (\tilde{\sigma}, \tilde{\rho})$. □

A key element of the definition of a λ-path and therefore of a λ-equilibrium is the requirement that at each iteration unused messages are provisionally dropped. Like the requirements that best replies are sticky and that the sender use minimal-message best replies, provisionally dropping unused messages serves to contain the proliferation of best replies.

The next example illustrates the role of provisionally dropping messages, while retaining sticky best replies and minimal-message best replies; we will discuss the impact of the latter two requirements later, in a more appropriate context.

Consider a sender-receiver game with the payoff structure in Figure 4, in which the two payoff types are equally likely, the message space is $M = \{m_1, m_2\}$, and there is a common language λ with $\lambda(m_i) = a_i$, $i = 1, 2$.

	a_1	a_2
t_1	3,3	0,2
t_2	3,-3	0,2

Figure 4: Not dropping messages

The game has a unique λ-equilibrium: the sender sends m_1 regardless of type and the receiver responds to both messages with action a_2.

Suppose, instead, that in the definition of a λ-path we did not prescribe to drop unused messages. Given the receiver strategy $r_1 = \lambda$ defined by $r(m_i) = a_i$ the following is a sequence of pure-strategy best replies starting with the sender's best reply s_1 to r_1:

$$s_1 = (t_1 \to m_1, t_2 \to m_1)$$

$$r_2 = (m_1 \to a_2, m_2 \to a_1)$$

$$s_2 = (t_1 \to m_2, t_2 \to m_2)$$

$$r_3 = (m_1 \to a_1, m_2 \to a_2)$$

$$s_3 = (t_1 \to m_1, t_2 \to m_1)$$

$$\dots$$

Since the sender uses only one message at every iteration, clearly the minimal message best reply condition is satisfied. Also, at every iteration each player's payoff from changing their strategy is strictly higher than from staying put and therefore the stickiness condition is satisfied. Unlike with dropping messages, however, we do not get a sharp prediction for language use. If we did not require that unused messages be dropped in the definition of a λ-path, there would be language equilibria in which either only one of the messages is used, as well as language equilibria in which both messages are used.

4.1 Common-interest games

The following result characterizes language use in games in which sender and receiver agree on which strategy profiles they prefer and they have a rich shared language. Following Blume, Kim and Sobel [2], say that a game is a *common-interest game* if in the corresponding base game there exists a strategy profile (σ^*, ρ^*) such that for any strategy profile (σ, ρ) either $U_i(\sigma, \rho) = U_i(\sigma^*, \rho^*)$ for $i = S, R$, or $U_i(\sigma, \rho) < U_i(\sigma^*, \rho^*)$ for $i = S, R$. That is, there is a unique efficient payoff pair.

Proposition 1 *In a common-interest game with a rich shared language λ every λ-equilibrium profile (σ, ρ) achieves the maximal payoff and satisfies $\rho(m) = \lambda(m)$ for all messages $m \in M$ that are received with positive probability.*

Proof: For every payoff type $t \in T$ let $a_t \in \arg\max_a u^S(a, t)$. Since the language λ is rich, for any $t \in T$ and any action $a_t \in A$, there is a message $m_t \in M_1$ with $\lambda(m_t) = a_t$. Because the language λ is shared and $r_1(m) = \lambda(m)$ for all $m \in M_1$, each payoff type $t \in T$ can achieve her maximal feasible payoff by sending the message $\theta^{-1}(m_t)$. Since the sender has a strategy that achieves her maximal feasible payoff against r_1 for each of her payoff types, for every λ-path the strategy s_1 of the sender must achieve the maximal feasible payoff $U^S(\sigma^*, \rho^*)$.

The common-interest assumption implies that a profile that achieves the sender's maximal payoff also achieves the receiver's maximal payoff. Hence s_1 and r_1 are mutual best replies in $\Gamma(M_1)$. Therefore r_2 agrees with r_1 in $\Gamma(M_2)$, where $M_2 = s_1(T)$. Since s_1 is a minimal message best reply to r_1, $s_2 = s_1$ uses all messages in M_2. For any λ-path, if s_k and r_k, are mutual best replies in $\Gamma(M_k)$ and s_k uses all messages in M_k, then $s_k = s_{k+1}$, $r_k = r_{k+1}$, and $M_{k+1} = M_k$ Hence, by induction, $(s_k, r_k, M_k) = (s_2, r_2, M_2) = (s_1, r_2, M_2)$ for all $k \geq 2$. This implies that (s_1, r_2) is a λ-profile for $\Gamma(M_2)$.

The receiver strategy r_2 in $\Gamma(M_2)$ agrees with r_1 for all messages received with positive probability given s_1 and therefore satisfies $r_2(m) = \lambda(m)$ for all messages $m \in M$ that are received with positive probability. The result follows from Definition 4 and Lemma 2. □

With absence of a shared language, communication is impossible. The following result confirms that in this case pooling is the only feasible outcome and shows in addition that in every λ-equilibrium all types of the sender send the same message.

Proposition 2 *For every game $\Gamma(M)$ with absence of a shared language, the set of λ-equilibrium strategies of the sender equals $\{s \in \mathsf{S}(M) | s(t') = s(t''), \forall t', t'' \in T\}$.*

This result differentiates the mental process that is captured through our iterative procedure from learning. Given that the translation, while uncertain, is fixed, repeated interaction would make it possible for sender and receiver to adjust their strategies toward effective communication, with sufficiently aligned preferences.

Proof: With absence of a shared language the sender assigns equal probability to every possible translation, regardless of her signal. As a result, all of her strategies have the same expected payoff against r_k for all $k \geq 1$, regardless of the receiver strategies r_k. Thus every strategy $s_1 \in \mathsf{S}(M)$ is a best reply to r_k for all $k \geq 1$, independent of the specification of r_k. Since the sender is using minimal message best replies, s_1 prescribes that all types use the same message. Since for $k > 1$ if $s_{k-1} \in \underline{\mathrm{BR}}_S(r_k)$ in $\Gamma(M_k)$ we have $s_k = s_{k-1}$, it follows that for every λ-path, $s_k = s_1$ for all $k \geq 1$. $\square$

Having the language be shared, or even common, is not necessary for achieving efficient communication in common-interest games. With a rich and accessible language, it suffices that the sender learns the translation. In contrast, as we will see later, it is not enough that the receiver learns the translation.

Proposition 3 *In a common-interest game with a rich and accessible language λ, if the sender learns the translation then every λ-equilibrium profile (σ, ρ) achieves the maximal payoff and satisfies $\rho(m, h^R) = \lambda(m)$ for all messages $m \in \Theta(M)$ that are received with positive probability and all receiver signals $h^R \in H^R$.*

Proof: Recall that for every payoff type $t \in T$, $a_t \in \arg\max_a u^S(a, t)$. For notational convenience, write θ for θ_{h^S}.

Since the language λ is rich and accessible and the sender learns the translation, for every $t \in T$, every $a_t \in A$, and every $h^S \in H^S$ there is a message $m_t \in M_1$ with $\lambda(\theta(m_t)) = a_t$. Denote that message by m_t^θ. Therefore, since the receiver strategy r_1 satisfies $r_1(m, h^R) = \lambda(m)$ for all $m \in \Theta(M_1)$ and all $h^R \in H^R$, each payoff type $t \in T$ can achieve her maximal feasible payoff against the strategy r_1 by sending the message m_t^θ. Hence the sender strategy $\hat{s}$ that is defined by $\hat{s}(t, \theta) = m_t^\theta$ for all $t \in T$ and all $\theta \in \Theta$ is a best reply to r_1 and achieves the sender's maximal feasible payoff $U^S(\sigma^*, \rho^*)$.

Since the sender has a strategy that achieves her maximal feasible payoff against r_1, for every λ-path the strategy s_1 of the sender (which may be different from $\hat{s}$) must achieve the maximal feasible payoff. The common-interest assumption implies that a strategy profile that achieves the sender's maximal payoff also achieves the receiver's maximal payoff. Hence s_1 and r_1 are mutual best replies in $\Gamma(M_1)$

Therefore r_2 agrees with r_1 on $M_2 = s_1(T \times H^S)$ and $s_2 = s_1$ uses all messages in M_2. For any λ-path, if s_k and r_k, are mutual best replies in $\Gamma(M_k)$ and s_k uses all messages in M_k, then $s_k = s_{k+1}$, $r_k = r_{k+1}$, and $M_{k+1} = M_k$. Hence, by induction $(s_k, r_k, M_k) = (s_2, r_2, M_2) = (s_1, r_2, M_2)$ for all $k \geq 2$ and r_2 agrees with r_1 for all messages received with positive probability given s_1. This implies that (s_1, r_2) is a λ-profile for $\Gamma(M_2)$. The receiver strategy r_2 in $\Gamma(M_2)$ agrees with r_1 for all messages received with positive probability given s_1 and therefore satisfies $r_2(m, h^R) = \lambda(m)$ for all $h^R \in H^R$ and for all messages $m \in \Theta(M)$ that are received with positive probability.

We can extend the receiver's strategy to $\Gamma(M)$ by letting $\rho(m, h^R) = \lambda(m)$ for all $\in M$. Hence there exists an equilibrium strategy strategy profile $(\sigma, \rho) = (s_1, \rho)$ in $\Gamma(M)$ that satisfies $(\sigma, \rho_{|M_2})$ is a λ-strategy profile in $\Gamma(M_2)$. This implies that every λ-equilibrium profile in $\Gamma(M)$ must satisfy condition 1 in Definition 4 and thus be reducible to a λ-profile for some $M^0 \subset M$.

The result follows by combining the facts that (1) every λ-strategy profile (s, r) for some M^0 achieves the maximal payoff and satisfies $r(m, h^R) = \lambda(m)$ for all $h^R \in H^R$ for all messages $m \in \Theta(M)$ that are received with positive probability and (2) every λ-equilibrium profile is reducible to a λ-strategy profile for some M^0. $\square$

With a rich and accessible language, as long as one of the players learns the translation, a common-interest game has multiple equilibria that achieve the maximal payoff: If the sender (or both players) learn the translation, any pure strategy profile (s, r) in which r is surjective and s is a best reply to r given the realized translation θ is an equilibrium profile that achieves the maximal payoff. Likewise, if only the receiver learns the translation, any pure strategy profile in which s is an arbitrary separating strategy and r is a best reply to s given the realized translation θ is an equilibrium profile that achieves the maximal payoff.

The language equilibrium selection, in contrast, differentiates among these cases. In the case in which the sender learns the translation Proposition 3 shows that any λ-equilibrium is efficient and satisfies that the receiver's strategy conforms with the pre-specified language. If, however, only the receiver learns the translation, language equilibria need not be either efficient or, if they are efficient, conform with the pre-specified language. The following example illustrates this.

	a_1	a_2
t_1	1,1	0,0
t_2	0,0	2,2

Figure 5: Common interest

Suppose that payoffs are the ones given in Figure 5 and that the two payoff types are equally likely; the message space is $M = \{m_1, m_2\}$; the set of translations is $\Theta = \{\theta_1, \theta_2\}$, with $\theta_1(m_i) = m_i$ and $\theta_2(m_i) = m_{3-i}$; $\mu(\theta_i) = \frac{1}{2}$, $i = 1, 2$, so that a priori both translations are equally likely; and, the language λ satisfies $\lambda(m_i) = a_i$ $i = 1, 2$. This is a common-interest game with a rich and accessible language. Let $H^R = \{h_1^R, h_2^R\}$, $H^S = \{h^S\}$, with $\eta^R(\theta_i | h_i^R) = 1$, so that the receiver signals fully reveal the translation and the set of sender signals is degenerate. Since the sender does not learn the translation and a priori both translations are equally likely, she expects the receiver to observe each message with equal probability independent of the message sent. Since λ is not a function of the translation, against $r_1 = \lambda$ the sender expects to induce each receiver action with equal probability. Hence, every sender strategy, including pooling on a single message, is a best reply against $r_1 = \lambda$. A minimal message best reply requires that s_1 pools on a single message. If s_1 is pooling on m_1 then $r_1 = \lambda$ is not a best reply to s_1 (and therefore stickiness does not prevent the receiver from updating his strategy), and for the receiver taking action a_2 independent of the message received and the signal observed is a best reply. Let this be r_2. s_1 and r_2 are mutual best replies and hence the profile (s_1, r_2) is a λ-equilibrium profile. This λ-equilibrium profile does not induce the common maximal payoff and the receiver strategy does not conform with the language λ.

A game is an *equilibrium-common-interest game* if in the base game there exists an equilibrium strategy profile (σ^*, ρ^*) such that for any equilibrium strategy profile (σ, ρ) either $U_i(\sigma, \rho) = U_i(\sigma^*, \rho^*)$ for $i = S, R$, or $U_i(\sigma, \rho) < U_i(\sigma^*, \rho^*)$ for $i = S, R$. That is, there is a unique payoff pair that is efficient in the set of equilibrium payoff pairs.

Propositions 1 and 3 do not extend to games with only equilibrium-common interest. Consider Game 6 with the payoff structure in Figure 6, two equally likely payoff types t_1 and t_2, a message space $M = \{m_1, m_2, m_3\}$, and a rich shared language λ with $\lambda(m_i) = a_i$, $i = 1, 2, 3$, where the single translation is the identity map.[10]

	a_1	a_2	a_3
t_1	2,3	1,-3	0,1
t_2	2,-3	1,3	0,1

Figure 6: State-independent preferences

The game has a continuum of equilibria, with payoffs ranging from 0 to 1 for the sender and from 1 to 1.5 for the receiver. There is a unique efficient equilibrium payoff pair with a payoff of 1 for the sender and 1.5 for the receiver. This payoff pair can be achieved by a strategy profile in which type t_1 sends message m_1, type t_2 sends messages m_1 and m_2 with probability 1/2 each, and the receiver responds to message m_1 with an equal-probability randomization over actions a_1 and a_3 and to both messages m_2 and m_3 with action a_2.

The set $M^0 = \{m_1\}$ is the unique subset of M with a λ-profile for M^0. For the sender, this λ-profile prescribes sending m_1 regardless of the payoff type. The receiver responds to m_1 with action a_3. The only way to extend this profile to all of $\Gamma(M)$ is to have the receiver respond to all messages with action a_3. Hence, for this game there is a unique λ-equilibrium profile. This equilibrium is inefficient and does not conform with the language.

Recall that in the definition of a λ-path best replies are "sticky": if a strategy from the previous iteration remains a best reply, it is retained in the current iteration. The next example demonstrates that Propositions 1 and 3 would fail if we dropped stickiness in the definition of a λ-path. Consider a sender-receiver game with the payoff structure in Figure 7, in which the three payoff types are equally likely, the message space is $M = \{m_1, m_2, m_3\}$, and there is a rich common language λ with $\lambda(m_i) = a_i$, $i = 1, 2, 3$.

The following is a sequence of best replies with elimination of unused messages and minimum-message best replies for the sender, starting with the sender's best reply s_1 to the receiver's strategy $r_1 = \lambda$:

[10]This is a game with state-independent sender-preferences, which are analyzed by Lipnowski and Ravid [22].

	a_1	a_2	a_3
t_1	1,1	0,0	1,1
t_2	1,1	1,1	0,0
t_3	0,0	1,1	1,1

Figure 7: Payoff ties

1. $s_1 = (t_1 \to m_1, t_2 \to m_1, t_3 \to m_3)$
2. $r_2 = (m_1 \to a_1, m_3 \to a_3)$
3. $s_2 = (t_1 \to m_3, t_2 \to m_1, t_3 \to m_3)$
4. $r_3 = (m_1 \to a_2, m_3 \to a_3)$
5. $s_3 = (t_1 \to m_3, t_2 \to m_1, t_3 \to m_1)$
6. $r_4 = (m_1 \to a_2, m_3 \to a_1)$
7. $s_4 = (t_1 \to m_3, t_2 \to m_3, t_3 \to m_1)$
8. $r_5 = (m_1 \to a_3, m_3 \to a_1)$

At this point the roles of the messages m_1 and m_3 have been exchanged. This means that there is a cycle in which the sender strategies s_1 and s_4 and the receiver strategies r_2 and r_5 appear infinitely often. The strategy profile in which the sender mixes with equal probability over s_1 and s_4 and the receiver mixes with equal probability over r_2 and r_5 is an equilibrium profile supported on this cycle. This equilibrium is inefficient and does not conform with the language.

With sticky best replies the inefficiency is removed. There are multiple λ-equilibria, but in all of these equilibria the receiver's interpretation conforms with the language λ. This seems plausible: the sender tells the receiver which action to take and the receiver, realizing that it is a common interest game, complies.

4.2 Block-aligned preferences

In this section we examine language use for a class of games in which preferences are imperfectly aligned. There is a partition of the payoff type space such that payoff types in every partition element strictly prefer to be thought of as belonging to their partition element rather than to any other partition element. For every set of types $T' \subseteq T$, define $B_R(T')$ as the set of receiver actions that are best replies to beliefs that assign positive probability only to types $t \in T'$.

Definition 5 *Players have* ***block-aligned preferences*** *for a nontrivial partition* $\mathcal{T} = \{T_1, \ldots, T_J\}$ *of the payoff type space* T *if*

1. $\arg\max_a u^S(t,a) \subseteq B_R(T_j)$, *and*

2. $\min_{a \in B_R(T_j)} u^S(t,a) > \max_{a \in B_R(T_\ell)} u^S(t,a)$

for all $t \in T_j$, *all* $j = 1, \ldots, J$, *and all* $\ell \neq j$.

Game 8 with the payoff structure in Figure 8, four equally likely payoff types, the message space $M = \{m_1, \ldots m_5\}$, and a common language λ with $\lambda(m_i) = a_i$, $i = 1, \ldots, 5$, has block-aligned preferences for the partition $\mathcal{T} = \{\{t_1, t_2\}, \{t_3, t_4\}\}$.

	a_1	a_2	a_3	a_4	a_5
t_1	5,2	1,6	-1,-1	-1,-1	4,3
t_2	1,5	5,2	-1,-1	-1,-1	4,3
t_3	-1,-1	-1,-1	5,2	1,6	4,3
t_4	-1,-1	-1,-1	1,5	5,2	4,3

Figure 8: Block-aligned preferences

In every λ-equilibrium, types t_1 and t_2 mix over messages m_1 and m_2 and types t_3 and t_4 mix over messages m_3 and m_4. In all of these equilibria, the receiver responds to messages m_1 and m_2 with action a_2 and to messages m_3 and m_4 with action a_4. Note that the sender *ex ante* prefers pooling to any λ-equilibrium, that there is no credible message profile, and that λ-equilibria are not neologism proof: the set of types t_1 and t_3 has a credible neologism.

In Game 8 language equilibrium does not pin down language use exactly. It does, however, place sensible constraints on language use that reflect the payoff structure. This holds more generally. To show this, given a partition of the payoff type space, we define what it means for language use to *block conform* with a language: Each sender type only induces received messages whose pre-specified meanings according to the language are best replies to beliefs concentrated on her partition element. The receiver responds to every received message whose pre-specified meaning is a best reply to beliefs concentrated on a partition element with an action that is a best reply to beliefs concentrated on the same partition element.

For any partition $\mathcal{T}$ of the payoff type space and every $t \in T$ denote the partition element that contains t by $T(t)$.

Definition 6 *Given a partition* $\mathcal{T}$ *of the payoff-type space, a set of strategy profiles* $\tilde{\Sigma}_S \times \tilde{\Sigma}_R \subseteq \Sigma_S \times \Sigma_R$ *in* $\Gamma(M^0)$ ***block conforms with the language*** λ *if*

1. *$[\sigma(m|t, h^S) > 0$ and $\eta^S(\theta|h^S) > 0] \Rightarrow \lambda(\theta(m)) \in B_R(T(t))$, $\forall \sigma \in \tilde{\Sigma}_S, t \in T, h^S \in H^S$.*

2. *$\lambda(m) \in B_R(T_j) \Rightarrow \rho(m, h^R) \in B_R(T_j)$, $\forall \rho \in \tilde{\Sigma}_R, j = 1, \ldots, J$, all messages $m \in \Theta(M^0)$ that are received with positive probability given some sender strategy $\sigma \in \tilde{\Sigma}_S$, and all receiver signals $h^R \in H^R$.*

The next result ensures that if a limit set of strategies reached by a λ-path block conforms with the language λ, then any minimal prep set containing that limit set also block conforms with the language λ.

Lemma 3 *Suppose that players have block aligned preferences for the partition $\mathcal{T}$, that $\mathsf{S}' \times \mathsf{R}' \subseteq \mathsf{S} \times \mathsf{R}$ in $\Gamma(M^0)$ block conforms with the language λ for the partition $\mathcal{T}$, and that for each $m \in \Theta(M^0)$ there is a strategy $s \in \mathsf{S}'$ such that message m is received with positive probability, then every minimal $\mathsf{S}' \times \mathsf{R}'$-Prep Set in $\Gamma(M^0)$ block conforms with the language λ.*

Proof: Suppose that $\mathsf{S}' \times \mathsf{R}'$ satisfies the conditions in the statement of the Lemma for the partition $\mathcal{T}$. Let $P_S \times P_R$ be an $\mathsf{S}' \times \mathsf{R}'$-Prep Set in $\Gamma(M^0)$.

Eliminate all sender strategies from P_S that do not satisfy Condition 1 for block conformity in Definition 6. Denote the resulting set by $\tilde{P}_S$ and observe that it is nonempty. Eliminate all receiver strategies from P_R that do not satisfy Condition 2 for block conformity in Definition 6, with $\tilde{\Sigma}_S$ the set of mixed strategy profiles of the sender supported on $\tilde{P}_S$. Denote the resulting set by $\tilde{P}_R$ and observe that it is nonempty.

Since $\tilde{P}_R$ satisfies Condition 2 for block conformity, since every message available to the sender induces a message in $\Theta(M^0)$ that is received with positive probability by the receiver for some sender strategy in $\mathsf{S}' \subseteq \tilde{P}_S$, and since preferences are block aligned for the partition $\mathcal{T}$, every sender best reply in $\Gamma(M^0)$ to beliefs concentrated on $\tilde{P}_R$ satisfies Condition 1 for block conformity. Since by assumption $P_S \times P_R$ is a Prep Set, P_S must contain a best reply for every belief that is concentrated on $\tilde{P}_R$. Since, as we saw, all such best replies satisfy Condition 1 for block conformity, they remain in $\tilde{P}_S$. Thus $\tilde{P}_S$ contains a best reply to every belief concentrated on $\tilde{P}_R$.

Let $\sigma' \in \Delta(\mathsf{S}')$ have full support on S', let $\tilde{\sigma} \in \Delta(\tilde{P}_S)$, and for all $\epsilon \in (0, 1)$, let $\sigma(\epsilon) = (1 - \epsilon)\tilde{\sigma} + \epsilon\sigma'$. Then $\sigma(\epsilon) \in \Delta(\tilde{P}_S)$ and $\sigma(\epsilon)$ induces every message in $\Theta(M^0)$ with positive probability. Since all strategies in $\tilde{P}_S$ satisfy Condition 1 for block conformity, for all $\epsilon > 0$ the strategy $\sigma(\epsilon)$ satisfies that condition. Since the strategy $\sigma(\epsilon)$ satisfies Condition 1 for block conformity and it induces every message in $\Theta(M^0)$ with positive probability, any receiver best reply to $\sigma(\epsilon)$ satisfies Condition 2 for block conformity. Since $P_S \times P_R$ is assumed to be a Prep Set, and $\tilde{P}_R$ is obtained from P_R by eliminating (only) strategies that do not satisfy Condition 2, the set $\tilde{P}_R$ contains a best reply to $\sigma(\epsilon)$ for all $\epsilon > 0$. Consider a sequence $(\epsilon_n, \rho(\epsilon_n))_{n=1}^{\infty}$ with $\lim_{n\to\infty} \epsilon_n = 0$, $\rho(\epsilon_n)$ a best reply to $\sigma(\epsilon_n)$ and $\rho(\epsilon_n) \in \tilde{P}_R$. Since $\tilde{P}_R$ is finite, there is a subsequence $(\epsilon_{n_j}, \rho(\epsilon_{n_j}))_{j=1}^{\infty}$ and $\tilde{\rho} \in \tilde{P}_R$ with $\rho(\epsilon_{n_j}) = \tilde{\rho}$ for all j. By continuity of the payoff function, $\tilde{\rho}$ is a best reply to $\tilde{\sigma}$. Hence $\tilde{P}_R$ contains a best reply to $\tilde{\sigma}$ for all $\tilde{\sigma} \in \Delta(\tilde{P}_S)$. Therefore, $\tilde{P}_S \times \tilde{P}_R$ is an $\mathsf{S}' \times \mathsf{R}'$-Prep Set. □

Using this observation, we now show that with a rich and accessible language and block-aligned preferences, all language equilibria block conform with the pre-specified language. Types belonging to an element of the partition for which there is block alignment send only messages whose pre-specified meanings are best replies to beliefs concentrated on that element. The proof proceeds by showing that for every λ-path block conformity is preserved at every iteration, then using Lemma 3 to establish that any set that is minimal in the class of prep sets that contain the limit set reached in this manner block conforms with the language, and finally to infer that any equlibrium supported on such a prep set must block conform with the language.

Proposition 4 *In games with block-aligned preferences and a rich and accessible language λ, if the sender learns the translation then every λ-equilibrium profile (σ, ρ) block conforms with the language λ.*

Proof: For every payoff type $t \in T$, let $a_t \in \arg\max_a u^S(a,t)$. For notational convenience, write θ for θ_{h^S}.

Since the language λ is rich and accessible and the sender learns the translation, for every $t \in T$, every $a_t \in A$, and every $h^S \in H^S$ there is a message $m_t \in M_1$ with $\lambda(\theta(m_t)) = a_t$. Denote that message by m_t^θ. Therefore, since the receiver strategy r_1 satisfies $r_1(m, h^R) = \lambda(m)$ for all $m \in \Theta(M_1)$ and all $h^R \in H^R$, each payoff type $t \in T$ can achieve her maximal feasible payoff against the strategy r_1 by sending the message m_t^θ.

Block alignment of preferences implies that

$$\max_a u^S(t,a) = \max_{a \in B_R(T(t))} u^S(t,a) \geq \min_{a \in B_R(T(t))} u^S(t,a) > \max_{a \in B_R(T_\ell)} u^S(t,a)$$

for all $T_\ell \neq T(t)$. Hence, for every λ-path the strategy s_1 satisfies Condition 1 in Definition 6. Given the strategy s_1, for any message m that the receiver observes with positive probability and that satisfies $\lambda(m) \in B_R(T_j)$, he knows that message was sent by a type in T_j. Hence r_2 satisfies Condition 2 in Definition 6.

If s_k satisfies Condition 1 in Definition 6, then for every message that has positive probability given s_k the receiver can infer the partition element containing the type who sent that message from the language λ. Therefore r_{k+1} satisfies Condition 2 in Definition 6 and for every $t \in T$ there is a message $m \in s_k(T \times H^S) = M_{k+1}$ with $r_{k+1}(\theta(m), h^R) \in B_R(T(t))$.

Since $M_k = s_{k-1}(T \times H^S)$ and $r_k \in \mathrm{BR}_R(s_{k-1})$ in $\Gamma(M_k)$, the strategy r_k specifies responses only for message in $\Theta(M_k)$, all of which are received with positive probability given s_{k-1}. Given s_{k-1}, if r_k satisfies Condition 2 in Definition 6 and for every $t \in T$ there is a message $m \in M_k$ with $r_k(\theta(m), h^R) \in B_R(T(t))$ then s_k satisfies Condition 1 in Definition 6.

Hence, by induction for every (s_k, r_{k+1}) and every $k \geq 1$, s_k satisfies Condition 1 in Definition 6 and given s_k, r_{k+1} satisfies Condition 2 in Definition 6.

Since M is finite and $M_{k+1} \subset M_k$ in the sequence $(M)_{k=1}^{\infty}$, there exists $K \geq 1$ such that $M_k = M^0$ for all $k \geq K$.

Let $\mathsf{S}' = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{s_k\}$ and $\mathsf{R}' = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{s_k\}$. Note that for every message $m \in M^0$ there is a sender strategy in S' for which m is sent with positive probability. Hence, the set $\mathsf{S}' \times \mathsf{R}' \subseteq \mathsf{S} \times \mathsf{R}$ in $\Gamma(M^0)$ block conforms with the language λ and for each $m \in \Theta(M^0)$ there is a strategy $s \in \mathsf{S}'$ such that message m is received with positive probability.

Let $P = P_S \times P_R \subseteq \mathsf{S}(M^0) \times \mathsf{R}(M^0)$ be a minimal $\mathsf{S}' \times \mathsf{R}'$-prep set in $\Gamma(M^0)$. Since $\mathsf{S}' \times \mathsf{R}'$ satisfies the conditions of Lemma 3, the set P block-conforms with the language λ. Hence, there there exists a λ-profile (σ^0, ρ^0) in $\Gamma(M^0)$ with support P, and every λ profile in $\Gamma(M^0)$ block conforms with the language λ.

We can extend the receiver's strategy to $\Gamma(M)$ by letting $\rho(m, h^R) = \rho^0(m^0, h^R)$ for all $m \in \Theta(M) \setminus \Theta(M^0)$ and some $m^0 \in \Theta(M^0)$ and $\rho(m, h^R) = \rho^0(m, h^R)$ for all $m \in \Theta(M^0)$ and all $h^R \in H^R$. The sender strategy σ^0 remains a best reply to ρ in $\Gamma(M)$.

Hence, there exists an equilibrium strategy strategy profile $(\sigma, \rho) = (\sigma^0, \rho)$ in $\Gamma(M)$ that satisfies $(\sigma, \rho_{|M^0})$ is a λ-strategy profile in $\Gamma(M^0)$. This implies that every λ-equilibrium profile in $\Gamma(M)$ must satisfy condition 1 in Definition 4 and thus be reducible to a λ-profile for some $M^0 \subset M$. The result follows since every every λ profile in $\Gamma(M^0)$ block conforms with the language λ. $\square$

4.3 Sender-preferred equilibria

One might suspect that if there is an equilibrium that maximizes the payoff of every type of the sender, the sender would be able to induce that equilibrium. In this section we show that this is the case with some qualifications.

A sender-receiver game is *generic* if $u^S(t, a') \neq u^S(t, a'')$ for $a', a'' \in A$ with $a' \neq a''$ and for each $T' \subseteq T$ the receiver has a unique best reply to the belief that equals the prior, π, restricted to T'. In the base game, an equilibrium is *sender ideal* if type t's payoff is $\max_{a \in A} u^S(t, a)$ for all $t \in T$.[11]

The proof of the main result in this section makes use of the following observation about generic games.

Lemma 4 *Suppose that a generic base game has a sender-ideal equilibrium. Then, if all types with the same ideal action $a \in A$ exclusively send message $m \in M$ and no other types send that message, action a is a best reply for the receiver to message m.*

The game in Figure 9 (with two equally likely types and at least two messages) illustrates the role of genericity in Lemma 4. The game has a sender-ideal equilibrium and both types agree on the set of actions that induce their ideal payoffs. If, however, the two types pool on a common message the receiver's best reply is not one of these actions.

[11] This matches the *sender's favorite equilibria* of Blume, Kim and Sobel [2].

	a_1	a_2	a_3
t_1	3,3	3,0	0,2
t_2	2,0	2,3	0,2

Figure 9: Perils of pooling

Proof: Let $A^* := \{a \in A | \exists t \in T \text{ s.t. } a = \arg\max_{a' \in A} u^S(t, a')\}$ be the set of receiver actions that maximize some type's payoff. For any action $a \in A^*$, define $T(a) := \{t \in T | a = \arg\max_{a' \in A} u^S(t, a')\}$ as the set of types for whom action a is the preferred action. This set is well defined by our genericity assumption. In a sender-ideal equilibrium each type $t \in T(a)$ sends only messages that induce action a.

Let σ be the strategy of the sender in a sender-ideal equilibrium e. For every $m \in M$ and $t \in T$, denote the receiver's posterior probability of type t given message m by $\beta(t|m)$. For any $a \in A^*$, let $M(a) := \{m \in M | \exists t \in T(a) \text{ s.t } \sigma(m|t) > 0\}$ be the set of messages that are sent with positive probability by some type in $T(a)$.

Then, in the presumed equilibrium, for any $a \in A^*$ and any $m \in M(a)$,

$$a \in \arg\max_{a'} \sum_{t \in T(a)} \beta(t|m) u^R(t, a').$$

For any $m \in M(a)$ let $p(m) = \sum_{t \in T(a)} \sigma(m|t)\pi(t)$.

Then the receiver's expected payoff from types in $T(a)$ equals

$$\begin{aligned}
\sum_{m \in M(a)} p(m) \sum_{t \in T(a)} \beta(t|m) u^R(t, a) &= \sum_{m \in M(a)} p(m) \sum_{t \in T(a)} \frac{\sigma(m|t)\pi(t)}{p(m)} u^R(t, a) \\
&= \sum_{m \in M(a)} \sum_{t \in T(a)} \sigma(m|t)\pi(t) u^R(t, a) \\
&= \sum_{t \in T(a)} \sum_{m \in M(a)} \sigma(m|t)\pi(t) u^R(t, a) \\
&= \sum_{t \in T(a)} \pi(t) u^R(t, a)
\end{aligned}$$

The action a must be a maximizer of $\sum_{t \in T(a)} \pi(t) u^R(t, a')$ since otherwise we could find at least one message $m \in M(a)$ for which a is not a maximizer of $\sum_{t \in T(a)} \beta(t|m) u^R(t, a')$, contradicting the assumption that we have an equilibrium. $\square$

The next result confirms the introductory conjecture for generic games with a rich and accessible language whose base games have a sender-ideal equilibrium and in which the sender learns the translation. Furthermore the receiver responds to every message that he receives with positive probability with an action that matches the pre-specified meaning of that message.

Proposition 5 *Suppose a generic game has a rich and accessible language λ and its base game has a sender-ideal equilibrium. Then, if the sender learns the translation, every λ-equilibrium profile (σ, ρ) achieves the sender's maximal payoff and satisfies $\rho(m, h^R) = \lambda(m)$ for all messages $m \in \Theta(M)$ that are received with positive probability and all receiver signals $h^R \in H^R$.*

Proof: In a generic game, for each type t there is a single receiver action $a^S(t) = \arg\max_a u^S(a, t)$ that maximizes that type's payoff. For notational convenience, write θ for θ_{h^S}. Since the language λ is rich and accessible and the sender learns the translation, for every $t \in T$, every $a^S(t)$, and every $h^S \in H^S$ there is a message $m_t \in M_1$ with $\lambda(\theta(m_t)) = a^S(t)$. Denote that message by m_t^θ.

Therefore, since the receiver strategy r_1 satisfies $r_1(m, h^R) = \lambda(m)$ for all $m \in \Theta(M_1)$ and all $h^R \in H^R$, each payoff type $t \in T$ can achieve her maximal feasible payoff against the strategy r_1 by sending the message m_t^θ.

Since in each iteration of a λ path the sender uses minimal-message best replies, we have that for all types t and t' with the same ideal action, $s_1(t, h^S) = s_1(t', h^S)$ for all $h^S \in H^S$. Let A^* be the set of all actions that are some type's ideal action, and for any action $a \in A^*$, let $T(a) := \{t \in T | a = \arg\max_{a' \in A} u^S(t, a')\}$.

By our minimal-message best reply assumption, for any message m that the receiver observes with positive probability given s_1 and that satisfies $\lambda(m) = a \in A^*$, his posterior belief is the prior concentrated on $T(a)$. Hence, by Lemma 4 and genericity, $r_2(m, h^R) = \lambda(m)$ for all $h^R \in H^R$ and all $m \in s_1(T \times H^S) = M_2$.

Since each type t induces her favorite action a_t, and since by our minimal-message best reply assumption all types with the same favorite action send the same message, for each $h^S \in H^S$ there is one and only one message in M_2 that induces type t's favorite action, given strategy r_2 of the receiver. This implies that s_2 agrees with s_1 on M_2, that s_2 and r_2 are unique best replies to each other in $\Gamma(M_2)$, and $r_2(m, h^R) = \lambda(m)$ for all messages $m \in \Theta(M_2)$ and all $h^R \in H^R$.

It follows that for all $k \geq 2$ we have $M_k = M_2$, $s_k = s_2$, $r_k = r_2$, s_k and r_k are unique best replies to each other in $\Gamma(M_k)$, and $r_k(m, h^R) = \lambda(m)$ for all messages $m \in \Theta(M_k)$ and all $h^R \in H^R$. Hence, there exists a set of message M^0 and a λ-profile (σ^0, ρ^0) in $\Gamma(M^0)$, and every λ profile (σ', ρ') achieves the sender's maximal payoff and satisfies $\rho'(m, h^R) = \lambda(m)$ for all messages $m \in \Theta(M)$ that are received with positive probability and all receiver signals $h^R \in H^R$. Since the sender achieves her maximal payoff for every λ-profile, each λ-profile can be trivially extended to a λ-equilibrium profile. $\square$

Proposition 5 applies to Game 10 with the payoff structure in Figure 10, two equally likely payoff types t_1 and t_2, a message space $M = \{m_1, m_2, m_3, m_3'\}$, and a rich shared

	a_1	a_2	a_3
t_1	1,3	0,0	2,2
t_2	0,0	1,3	2,2

Figure 10: Sender-preferred equilibrium

language λ with $\lambda(m_i) = a_i$, $i = 1, 2, 3$, and $\lambda(m'_3) = a_3$, where the single translation is the identity map. Notice that if we dropped the minimal-message reply assumption in the definition of λ-paths, there would be a λ-path with $s_1 = (t_1 \to m_3, t_2 \to m'_3)$ that would converge to a separating equilibrium with $\sigma = s_1$, in which the sender would not obtain her maximal payoff. It seems implausible, however, that the receiver would be able to tell the two types apart on the basis of which message exactly the sender chooses to indicate the desire that action a_3 be taken.[12]

For the result in Proposition 5 to hold, it is not enough that all sender types agree on their favorite equilibrium in the base game. Let Σ_R^{eqm} denote the set of all receiver strategies that are part of some equilibrium in the base game. An equilibrium of the base game is *sender optimal* if type t's payoff is

$$\max_{m \in M, \rho \in \Sigma_R^{\text{eqm}}} \sum u^S(t, a)\rho(a|m)$$

for all $t \in T$.

The base game of Game 6 in Figure 6 has a sender-optimal equilibrium and satisfies our genericity condition. The unique λ-equilibrium profile, however, results in a payoff 0 for both types, whereas their payoff at a sender-optimal equilibrium is 1.

4.4 Finite CS games

In this section we examine language use in a class of games that may be thought of as an adaptation of the setup of Crawford and Sobel [8] to a setting with finite type and action spaces.

For any linear ordering $\leq$ of the set of types T and any t', t'', refer to the set $[t', t''] := \{t \in T | t' \leq t \leq t''\}$ as an *interval* of types. The linear order $\leq$ of T induces a partial order $\precsim$ on the set of intervals of T defined by $[t'_1, t''_1] \precsim [t'_2, t''_2] \Leftrightarrow t'_1 \leq t'_2$ and $t''_1 \leq t''_2$. Observe that in a generic sender-receiver game for every state $t \in T$, each player i has a unique ideal point $a^i(t) = \arg\max_a u^i(a, t)$.

A generic sender-receiver game is a *finite CS game* if there exist orderings of types and actions such that:

[12] Assuming nominal message costs as in Blume, Kim and Sobel [2] has a similar effect as adopting the minimal-message best reply assumption. With nominal message costs, however, it would frequently not be possible to extend λ-profiles to λ-equilibria for the entire game – there would be a tension between the pre-specified meanings of messages and the incentive to use lower-cost messages.

1. The functions u^i are unimodal in a for all $t \in T$ and $i = 1, 2$; i.e., $a < a' \Rightarrow u^i(a,t) < u^i(a',t)$ for all $a' \leq a^i(t)$ and $a > a' \Rightarrow u^i(a,t) < u^i(a',t)$ for all $a' \geq a^i(t)$.

   That is, for each state and each player, the player's payoff is strictly increasing in the action below the player's ideal point and strictly decreasing above the player's ideal point.

2. The sender's preference has an upward bias relative to the receiver: $a^R(t) < a^S(t), \forall t \in T$.

3. The receiver's ideal point is responsive: $a^R(t') \neq a^R(t)$ for all $t, t' \in T$ with $t \neq t'$

4. Each player i's payoff function u^i satisfies the single crossing condition[13]:

$$t_2 > t_1 \text{ and } a_2 > a_1$$

   implies

$$u^i(a_2, t_1) - u^i(a_1, t_1) > 0 \Rightarrow u^i(a_2, t_2) - u^i(a_1, t_2) > 0.$$

In a finite CS game, the sender has an incentive to exaggerate her type. This suggests that in equilibrium the receiver may discount the pre-specified meaning of the messages that he receives: after every message sent in equilibrium the receiver takes an action that is lower than the action that matches the pre-specified meaning of the message. This is confirmed by the following result.

Proposition 6 *In any generic finite CS game with a rich and accessible language λ, if the sender learns the translation, then for every λ-equilibrium profile (σ, ρ) and all messages $m \in \Theta(M)$ that are received with positive probability,*

1. *$\lambda(m) = a^S(t)$ for some $t \in T$, and*

2. *$\rho(m, h^R) < \lambda(m)$ for all receiver signals $h^R \in H^R$.*

Every message that is observed with positive probability has a pre-specified meaning that matches some sender type's ideal point and is discounted by the receiver.

Proof: Recall that for each type t there is a single receiver action $a^S(t) = \arg\max_a u^S(a,t)$ that maximizes that type's payoff. By assumption, for each $h^S \in H^S$, the sender learns the translation θ_{h^S}. For notational convenience, suppress the explicit dependence of the realized translation on the sender's signal h^S and write θ for θ_{h^S}. Since the language λ is rich and accessible, for every $t \in T$ and every realized translation θ there is a message $m_t^\theta \in M_1$ with $\lambda(\theta(m_t^\theta)) = a^S(t)$.

Therefore, since the sender learns the translation θ, since the receiver strategy r_1 satisfies $r_1(m, h^R) = \lambda(m)$ for all $m \in \Theta(M_1)$ and all $h^R \in H^R$, each payoff type $t \in T$ has at least

[13] Genericity implies that we can ignore the possibility that $u^i(a_2, t_1) - u^i(a_1, t_1 = 0$ for $a_2 > a_1$.

one way of inducing her ideal action against the strategy r_1 by sending the message m_t^θ. Since at each iteration the sender uses minimal-message best replies, for all types t and t' with the same ideal action, $s_1(t, h^S) = s_1(t', h^S)$ for all $h^S \in H^S$.

Since each type t can induce her ideal action by sending message $m_t^\theta \in M_1$, every message $m \in \Theta(M_1)$ that is received when the sender uses strategy s_1 satisfies $\lambda(m) = a^S(t)$ for some $t \in T$. Hence, for all $m \in \Theta(M_2)$ where $M_2 = s_1(T \times H^S)$ there exists a type $t \in T$ such that $\lambda(m) = a^S(t)$. This establishes the first part of the proposition since $M_{k+1} \subseteq M_k$ for every λ-path.

Let A^* be the set of all actions that are some type's ideal action. For any action $a \in A^*$, let $T(a) := \{t \in T | a = a^S(t)\}$. This is the set of types whose ideal action is a. Given the single-crossing condition for the sender, for each $a \in A^*$ the set $T(a)$ is an interval.

Since at each iteration the sender uses minimal-message best replies, s_1 prescribes that all types with the same ideal point send the same message, for any message m that the receiver observes with positive probability given s_1 and that satisfies $\lambda(m) = a \in A^*$, his posterior belief is the prior concentrated on $T(a)$. This receiver inference is unaffected by the receiver's signal $h^R \in H^R$. Hence, by genericity, the receiver has a unique best reply $r_2(m, h^R)$ to all $m \in \Theta(M_2)$, which is independent of h^R for all $h^R \in H^R$.

For each $m \in \Theta(M_2)$ define $T_2(m)$ as the interval of types who induce (the received) message m. Each message in $m \in \Theta(M_2)$ induces a distinct sender ideal action. Therefore, for each realized translation $\theta \in \Theta$, the collection of intervals $\{T_2(m) | m \in \theta(M_2)\}$ forms a partition of T. Moreover, ignoring the indexing by messages, this partition is the same for all $\theta \in \Theta$. Denote this partition by $\mathcal{T}_2$. The elements of any partition of T into intervals are linearly ordered by $\prec$, the strict linear order associated with $\precsim$.

For each $m \in \Theta(M_2)$, $\lambda(m)$ is the common ideal point of types in $T_2(m)$. Hence, for each type $t \in T_2(m)$, $a^R(t) < a^S(t) = \lambda(m)$. Therefore the single-crossing condition for the receiver implies that for each $m \in \theta(M_2)$, $r_2(m, h^R) < \lambda(m)$ (this uses the fact that the distributions obtained by concentrating the support of the prior on intervals $[\underline{t}', \overline{t}']$ and $[\underline{t}, \overline{t}]$, with $\underline{t}' \leq \underline{t}$ and $\overline{t}' \leq \overline{t}$ are MLRP ranked).

We now proceed by induction. We have established that for (s_1, r_2) there is a partition $\mathcal{T}_2$ of the type space T such that for every $m \in \Theta(M_2)$ (where $M_2 = (s_1(T \times H^S))$, the partition element $T_2(m)$ is the set of types who induce message m, that this set is an interval, and that $r_2(m, h^R) < r_1(m, h^R) = \lambda(m)$

Assume that for (s_k, r_{k+1}) there is a partition $\mathcal{T}_{k+1}$ of the type space T such that for every $m \in \Theta(M_{k+1})$ (where $M_{k+1} = (s_k(T \times H^S))$, the partition element $T_{k+1}(m)$ is the set of types who induce message m, that this set is an interval, and that $r_{k+1}(m, h^R) \leq r_k(m, h^R) < \lambda(m)$ Genericity implies that for each type $t \in T$ there is a unique message in M_{k+1} that maximizes that type's payoff given the realized translation and receiver strategy r_{k+1}.

Consider two messages $m', m'' \in \Theta(M_{k+1})$ with $T_{k+1}(m') \prec T_{k+1}(m'')$. By the single-crossing condition for the sender, these messages satisfy $r_k(m', h^R) < r_k(m'', h^R)$. By the single-crossing condition for the receiver and since $t \neq t' \Rightarrow a^R(t') \neq a^R(t)$, these messages satisfy $r_{k+1}(m', h^R) < r_{k+1}(m'', h^R)$. By assumption, we also have $r_{k+1}(m', h^R) \leq r_k(m', h^R)$ and $r_{k+1}(m'', h^R) \leq r_k(m'', h^R)$. Hence, the unimodality of the sender's payoff function implies that given the receiver's strategy r_{k+1} any type $t \in T_{k+1}(m'')$ strictly prefers inducing the received message m'' to inducing the received message m'.

Recalling that $s_{k+1}(T \times H^S) = M_{k+2} \subseteq M_{k+1}$, and for each $m \in \theta(M_{k+2})$ defining $T_{k+2}(m)$ as the interval of types who induce (the received) message m with the the strategy s_{k+1}, this implies that for every message $m \in \Theta(M_{k+2})$, $T_{k+2}(m) \precsim T_{k+1}(m)$. Hence, from the single crossing condition for the receiver $r_{j+2}(m, h^R) \leq r_{k+1}(m, h^R) < \lambda(m)$ for all $m \in \Theta(M_{k+2})$. Therefore, for (s_{k+1}, r_{k+2}) there is a partition $\mathcal{T}_{k+2}$ of the type space T such that for every $m \in \Theta(M_{k+2})$ the partition element $T_{k+2}(m)$ is the set of types who induce message m, this set is an interval, and $r_{k+2}(m, h^R) \leq r_{k+1}(m, h^R) < \lambda(m)$.

Hence, for every message $m \in \Theta(M^0)$, where $M^0 = \bigcap_{k=1}^{\infty} M_k$, the sequence $(r_k(m, h^R))_{k=1}^{\infty}$ is monotonically decreasing on a finite set, with $r_k(m, h^R) < \lambda(m)$ for all $k > 1$. Clearly this sequence converges. Denote the limit by $\rho(m, h^R)$ and observe that $\rho(m, h^R) < \lambda(m)$. Hence the sequence $(r_k)_{k=1}^{\infty}$ restricted to M^0 converges to a function $\tilde{\rho} : M^0 \to A$. We can ignore the dependence of $\tilde{\rho}$ on H^R.

Let $(s_k)_{k=2}^{\infty}$ be the sequence of the sender's unique best replies s_k to r_k in $\Gamma(M_k)$ for $k = 2, \ldots, \infty$. Since $(r_k)_{k=2}^{\infty}$ restricted to M^0 converges, so does $(s_k)_{k=2}^{\infty}$ and the limit, $\tilde{\sigma} : T \times H^S \to M^0$, is the unique best reply in $\Gamma(M^0)$ to $\tilde{\rho}$. Likewise, $\tilde{\rho}$ is the unique best reply in $\Gamma(M^0)$ to $\tilde{\sigma}$.

Extend the receiver strategy $\tilde{\rho}$ from the game $\Gamma(M^0)$ to the game $\Gamma(M)$ by letting $\rho(m) = \tilde{\rho}(m^0)$ for all $m \in \theta(M \setminus M^0)$ and some $m^0 \in \theta(M^0)$. In the game $\Gamma(M)$, if the receiver uses the strategy ρ, then every action the sender can induce by sending a message in $M \setminus M^0$ she can also induce by sending a message in M^0. Hence, if we define $\sigma : T \times H^S \to M$ by letting $\sigma(t, h^S) = \tilde{\sigma}(t, h^S)$ for all $t \in T$ and $h^S \in H^S$, then (σ, ρ) is an equilibrium strategy profile for the game $\Gamma(M)$. $\square$

Game 11 with the payoff structure in Figure 11, four equally likely payoff types $t_1, \ldots, t_4$, a message space $M = \{m_1, \ldots, m_5\}$, and a common language λ with $\lambda(m_i) = a_i$, $i = 1, \ldots, 5$, is a finite CS game.

	a_1	a_2	a_3	a_4	a_5
t_1	4,1	3,5	1,2	-1,-1	-3,-3
t_2	0,0	4,1	3,5	1,2	-1,-1
t_3	-1,-1	0,0	4,1	3,5	1,2
t_4	-2,-2	-1,-1	0,0	4,1	3,5

Figure 11: The status of truth

The sender's unique best reply against the language λ is the strategy $s_1 = (t_1 \to m_1, t_2 \to m_2, t_3 \to m_3, t_4 \to m_4)$. The receiver's unique best reply against the sender's strategy

s_1 in the game in which message m_5 has been eliminated is the strategy $r_2 = (m_1 \to a_2, m_2 \to a_3, m_3 \to a_4, m_4 \to a_5)$. The sender's unique best reply against the receiver's strategy r_2 in the game in which message m_5 has been eliminated is the strategy $s_2 = (t_1 \to m_1, t_2 \to m_1, t_3 \to m_2, t_4 \to m_3)$. The receiver's unique best reply against the sender's strategy s_2 in the game in which messages m_4 and m_5 have been eliminated is the strategy $r_3 = (m_1 \to a_3, m_2 \to a_4, m_3 \to a_5)$. The sender's unique best reply against the receiver's strategy r_3 in the game in which messages m_4 and m_5 have been eliminated is the strategy $s_3 = (t_1 \to m_1, t_2 \to m_1, t_3 \to m_1, t_4 \to m_2)$. The receiver's unique best reply against the sender's strategy s_3 in the game in which messages m_3, m_4 and m_5 have been eliminated is the strategy $r_4 = (m_1 \to a_3, m_2 \to a_5)$. Iterating further leaves the remaining message space, $\{m_1, m_2\}$ unchanged. In the game with that reduced message space the strategies s_3 and r_4 are unique best replies to each other. Hence (s_3, r_5) is a λ-profile. Since best replies are unique at every step, it is the unique λ-profile. The λ-profile can be extended to an equilibrium of the entire game by having the receiver use on-path responses after off-path messages. In every λ-equilibrium (there is multiplicity because of different possible specifications of off-path responses) the sender uses the strategy $\sigma = (t_1 \to m_1, t_2 \to m_1, t_3 \to m_1, t_4 \to m_2)$.

Thus, in every λ-equilibrium types t_1, t_2 and t_3 send a common message whose pre-specified meaning is a request for action a_1 and type t_4 sends a message whose pre-specified meaning is a request for action t_2. Except for type t_1 none of the types request their favorite action, in terms of the language. They are all strategically distorting message meanings. The receiver takes none of the messages that he receives with positive probability at face value. Thus, while there is influential communication, message use is far from being a truthful expression of intentions on the part of the sender, and messages are not being taken as truthful by the receiver.

Neologism proofness rejects the pooling equlibrium outcome in Game 11, since type t_4 has a credible neologism. It does not reject the partial pooling equilibrium outcome that we observe in the language equilibria. It therefore agrees with the outcome prediction of language equlibrium in this game, while being silent about message use in equilibrium. Since there is no credible message profile, credible message rationalizability/equilibrium is equally silent about language use in Game 11.

The language equilibrium prediction in finite CS games is robust to enlarging the message space. One can add any number message, introduce any number of synonyms for messages, or have the message space become infinite. With such an enlargement there will be a proliferation of language equilibria, but they will only differ in terms of exchanging synonymous messages. In Game 11, if for example we added a message m_1' with $\lambda(m_1') = a_1$, in any λ-equilibrium types t_1, t_2 and t_3 would either send a common message m_1 or a common message m_1' and the receiver's equilibrium interpretations of these messages would be the same.

5 Reflections on uncertainty about language

Blume and Board [5] capture language constraints through limitations on the sender's ability to send messages and the receiver's ability to discriminate among messages. The translation apparatus employed here nests their constraints and links them to a language with pre-existing meanings. Language equilibrium imposes additional constraints on message use. Whereas Blume and Board analyze efficient equilibria of games with uncertainty about the ability to send and differentiate among messages, language equilibrium captures and emphasizes the requirement that message use be linked to the meanings in a pre-specified language.

	a_1	a_2	a_3
t_1	10,10	9,0	0,9
t_2	9,0	10,10	0,9

Uncertainty about language

To get a closer look at the connection, consider two variations on Game 3 from Section 2. For convenience, the figure above reproduces the payoff structure. Also, recall that there are two equally likely payoff types t_1 and t_2, a message space $M = \{m_1, m_2\}$, a language λ with $\lambda(m_i) = a_i$, $i = 1, 2$, and two translations θ' and θ'', defined by $\theta'(m) = m$ and $\theta''(m) = m_2$ for both $m \in M$ with a common prior μ over translations such that $\mu(\theta'') = p$, where p satisfies $\frac{1}{9} < p < 1$. With common knowledge of this structure, we found that there is a unique λ-equilibrium in which the sender sends message m_1 regardless of her payoff type.

Now suppose that instead of both players remaining uncertain about the translation, one of them receives a perfectly informative signal while the other remains uninformed. Assume that this fact is common knowledge.

First, suppose that it is the sender who becomes perfectly informed about the translation, while the receiver remains uninformed. This mirrors the situation in Blume and Board, where the sender is language constrained with probability p. In the event that she learns that the translation is θ', the identity mapping, she is unconstrained and can induce both messages on the receiver side. Otherwise, she is limited to inducing the received message m_2.

While the sender's strategy space is richer now (she can condition on both her payoff type and her language type), all that matters is her choice of message in the event the translation is the identity mapping. Focussing on that part of the sender's strategy, her unique (partial) best reply against the receiver's language is the (partial) strategy $s_1 = ((\theta', t_1) \to m_1, (\theta', t_2) \to m_2)$. As before, with this (partial) sender strategy, the receiver's posterior probability that the sender's type is t_2 after observing message m_2 equals $\frac{1}{p+1}$. Therefore, as long as $p > \frac{1}{9}$, the receiver has a unique best reply $r_2 = (m_1 \to a_1, m_2 \to a_3)$ to the sender's strategy s_1. Against r_2, the sender has a unique minimal-message best reply

$s_1 = ((\theta', t_1) \to m_1, (\theta', t_2) \to m_1, (\theta'', t_1) \to m_1, (\theta'', t_2) \to m_1)$. At that point message m_2 is dropped.[14]

With m_2 dropped, from here the iteration of best replies and elimination of messages proceeds as in the case where both players remain uninformed. There is a unique λ-equilibrium strategy profile with the sender using message m_1 regardless of type and the receiver responding with action a_3 after all messages.

Second, suppose that it is the receiver who becomes perfectly informed about the translation, while the sender remains uninformed. This mirrors the situation in Blume and Board, where the receiver is language constrained with probability p. In the event that he learns that the translation is θ', the identity mapping, he is unconstrained and can differentiate between both sent messages. Otherwise, he is limited to treating both sent messages identically.

The sender's unique best reply against the receiver's language is the strategy $s_1 = (t_1 \to m_1, t_2 \to m_2)$. Given the sender strategy s_1, the receiver's unique best reply is $r_2 = ((\theta', m_1) \to a_1, (\theta', m_2) \to a_2, (\theta'', m_2) \to a_3)$ (note that the receiver's language type θ'' never observes message m_1 and therefore does not have to condition on that message). Against r_2 the receiver has a unique best reply $s_2 = s_1$. Hence, there is a unique λ-strategy profile, in which the sender sends message m_i if her payoff type is t_i, the receiver responds with a_i to m_i if she can differentiate messages, and takes the action a_3 otherwise.

I conclude this section with an observation about dropping unused messages in games with language uncertainty. When there is language uncertainty, dropping a message may have no effect on which messages are observed with positive probability by the receiver. As a result, a message that that is not part of a sender's best reply σ to a strategy of the receiver, may become an indispensable part of the sender's best reply to a strategy ρ of the receiver that best responds to σ. It may be the case that a message become attractive as a result of not being used. One could make the case that such messages should not be provisionally eliminated.

Call a message m *redundant* given a sender strategy σ, if there exists a receiver best reply to σ such no type strictly prefers sending message m. In the definition of a λ-path, one might consider dropping a message only if it is redundant. This would make no difference for any of our results – in particular, with a common language or when the sender learns the translation, all unsent messages are redundant. The following example illustrates the impact of only dropping redundant messages on the language equilibrium prediction.

Consider Game 12 with the payoff structure in Figure 12, four equally likely payoff types $t_1, \ldots, t_4$, a message space $M = \{m_1, \ldots, m_4\}$, a language λ with $\lambda(m_i) = a_i$, $i = 1, \ldots, 4$, and a set of translations Θ that consists of all bijections $\theta : M \to M$. Use θ^* to denote the identity mapping, so that $\theta^*(m_i) = m_i$ for all $i = 1, \ldots, 4$. Assume that the common prior μ over translations satisfies $\mu(\theta^*) = 0.95$, and $\mu(\theta') = \mu(\theta'')$ for all $\theta', \theta'' \neq \theta^*$.

The sender's unique best reply against the language λ is the strategy $s_1 = (t_1 \to m_1, t_2 \to$

[14] If we did not restrict the sender to minimal message best replies, there would be other best replies for the sender, including $s_1 = ((\theta', t_1) \to m_1, (\theta', t_2) \to m_1, (\theta'', t_1) \to m_2, (\theta'', t_2) \to m_2)$. In that case m_2 would not be dropped and at the next step all sender strategies would be best replies. We would not get a sharp prediction for message use.

	a_1	a_2	a_3	a_4
t_1	1,3	0,0.1	0,0	0,0
t_2	2,1	1,3	0,0	0,0
t_3	0,0	2,1	1,3	0,0
t_4	0,0	0,0.1	2,1	1,3

Figure 12: Dropping only redundant messages

$m_1, t_3 \to m_2, t_4 \to m_3)$. Even though message m_4 is not sent by any type, it is received with positive probability since all bijections are translations that have positive probability.

Given the sender strategy s_1, the receiver's posterior after observing m_4 is the uniform distribution on T. Therefore, the receiver's unique best reply against the sender's strategy s_1 is the strategy $r_2 = (m_1 \to a_1, m_2 \to a_3, m_3 \to a_3, m_4 \to a_2)$. Since a_2 is the unique maximizer of type t_2's payoff and only m_4 induces that action, message m_4 fails to be redundant. Since message m_4 fails to be redundant, it is not dropped in this iteration (and similarly, for any other messages and iterations below).

The sender's unique best reply against the receiver's strategy r_2 is the strategy $s_2 = (t_1 \to m_1, t_2 \to m_1, t_3 \to m_4, t_4 \to m_2)$. From there, we get the following sequence of unique best replies: $r_3 = (m_1 \to a_1, m_2 \to a_4, m_3 \to a_2, m_4 \to a_3)$, $s_3 = (t_1 \to m_1, t_2 \to m_1, t_3 \to m_3, t_4 \to m_4)$, $r_4 = (m_1 \to a_1, m_2 \to a_2, m_3 \to a_3, m_4 \to a_4)$, $s_4 = (t_1 \to m_1, t_2 \to m_1, t_3 \to m_2, t_4 \to m_3)$, ... Since s_4 coincides with s_1, and best replies are unique, we have a cycle.

Denote the set of pure sender strategies in this game by S and the set of pure receiver strategies by R. Let $\mathsf{S}' = \{s_1, \ldots, s_4\}$ and $\mathsf{R}' = \{r_1, \ldots, r_4\}$. These are the sets of strategies that appear in the cycle that is generated by λ.

Let $\mathsf{S}'' = \{s \in \mathsf{S} | s(t) = m_1 \text{ if and only if } t \in \{t_1, t_2\}\}$ and $\mathsf{R}'' = \{r \in \mathsf{R} | r(m) = a_1 \text{ if and only if } m = m_1\}$. $\mathsf{S}' \times \mathsf{R}'$ is a (strict) subset of $\mathsf{S}'' \times \mathsf{R}''$. $\mathsf{S}'' \times \mathsf{R}''$ is an $\mathsf{S}' \times \mathsf{R}'$-curb set, i.e., it contains $\mathsf{S}' \times \mathsf{R}'$ and all best replies to beliefs concentrated on $\mathsf{S}'' \times \mathsf{R}''$.[15]

There is a unique minimal $\mathsf{S}' \times \mathsf{R}'$-curb set, and it is contained in $\mathsf{S}'' \times \mathsf{R}''$; this follows from the fact that the intersection of any two $\mathsf{S}' \times \mathsf{R}'$-curb sets is an $\mathsf{S}' \times \mathsf{R}'$-curb set. Every minimal $\mathsf{S}' \times \mathsf{R}'$-prep set is contained in the minimal $\mathsf{S}' \times \mathsf{R}'$-curb set and therefore in $\mathsf{S}'' \times \mathsf{R}''$; this follows from the fact that the intersection of any $\mathsf{S}' \times \mathsf{R}'$-curb set and any $\mathsf{S}' \times \mathsf{R}'$-prep set is an $\mathsf{S}' \times \mathsf{R}'$-prep set.

Hence, every λ-equilibrium has the property that types t_1 and t_2, and only those types, send message m_1. Furthermore, there is no pure-strategy λ-equilibrium. This follows, since for any candidate for such an equilibrium there would be an unused message, and type t_2 would strictly prefer sending that message.

[15] Basu and Weibull [1] define and discuss (minimal) curb sets.

If instead of dropping only redundant messages we dropped all unused messages, there would be a unique λ-profile with types t_1 and t_2 sending message m_1 and types t_3 and t_4 sending message m_2 (and the receiver best responding to those messages). This λ-profile, however, cannot be extended to an equilibrium in the entire game. Thus dropping only redundant messages gives us a sharper, and arguably more plausible, prediction in this game.

6 Discussion

Our analysis of meaning in sender-receiver games is rooted in truth, interpreted as the non-strategic meaning of messages, but does not require, or generally predict, that message use is truthful. In some cases, when players have common interests or there is a sender-ideal equilibrium, the theory predicts that the receiver responds to messages in accordance with their pre-specified meanings. When preferences are only imperfectly aligned this correspondence breaks down, although the pre-specified message meanings leave traces in players' behavior. With block-aligned preferences, the receiver responds to messages in accordance with pre-specified meanings that match beliefs concentrated on blocks of sender types. In finite CS games, the receiver systematically discounts the pre-specified meanings of messages received in equilibrium.

The approach is versatile. It yields predictions in a large class of games, including games in which there is uncertainty about language, games in which there is private information about language, games with message spaces of any size, games in which there is any number of synonyms for messages, etc. The analysis does not require but easily accommodates rich language assumptions that are customarily made in this literature.

Beside Farrell, Rabin, and the level-k approach of Crawford, a few others propose ways of giving a pre-existing language a role in the analysis of sender-receiver games. Blume [3] uses sender trembles to induce exogenous message meanings – the trembles govern message meanings for any message that is not used deliberately. Using Kalai and Samet's [18] persistence concept, he shows that in some classes of games message use is consistent with the tremble-induced exogenous meanings. Gordon, Kartik, Lo, Olszewski and Joel Sobel [16] impose the requirement that players use monotonic strategies in CS games. Monotonicity can be thought of as a mild condition on language use. When combining the monotonicity requirement with iterative deletion of dominated strategies, they find that with a finite message space only the maximal messages are used. Kartik, Ottaviani and Squintani [19] consider CS games with lying costs. The message space coincides with the type space and types pay a cost that is increasing in the distance of messages from the truth. They show that there are separating equilibria in which the sender exaggerates her type. Since the receiver can back out the truth, he discounts the stated message meanings. Analogous to Kartik et al, language equilibrium predicts language inflation in finite CS games, but without introducing lying costs.

It is fairly common in this literature to think of a pre-existing language in terms of subsets of the type space. Semantic meanings of messages are then of the form "my type belongs to the following set of types" or equivalently the prior restricted to the indicated set of types. We chose, instead, to have a language be a mapping from messages to receiver actions. There is no significant difference. One advantage of the approach chosen here is that there is a natural correspondence between the set of all beliefs and the set of all best replies to some belief. Thus modeling semantic meanings of messages in terms of receiver actions implicitly permits semantics meanings that are probabilistic statements about types like "my type is either s or t, but twice as likely to be s than t."

The framework proposed in this paper permits us to capture different degrees of sharing a language and to vary beliefs about what is shared. Davidson [12] is concerned with what it means to share a language. He states, somewhat provocatively, that "there is no such a thing as a language, not if a language is anything like what many philosophers and linguists have supposed." He proposes that what speaker and listener share on a give occasion is what he calls a "passing theory." Perhaps it is not too far off the mark to think of we call a language plus what the sender believes the translation to be as what Davidson would refer to as the sender's belief about the receiver's "prior theory" of interpretation. Likewise, a language equilibrium in this paper, which is reached upon reflection starting from a language, shares parallels with Davidson's "passing theory."

References

[1] Basu, Kaushik, and Jörgen W. Weibull [1991], "Strategy Subsets Closed Under Rational Behavior" *Economics Letters* **36**, 141–146.

[2] Blume, Andreas, Yong-Gwan Kim, and Joel Sobel [1993], "Evolutionary Stability in Games of Communication," *Games and Economic Behavior* **5**, 547–575.

[3] Blume, Andreas [1996], "Neighborhood Stability in Sender-Receiver Games," *Games and Economic Behavior* **13**, 2–25.

[4] Blume, Andreas, Douglas V. DeJong, Yong-Gwan Kim and Geoffrey B. Sprinkle [2001], "Evolution of Communication with Partial Common Interest," *Games and Economic Behavior* **37**, 79–120.

[5] Blume, Andreas and Oliver J. Board [2013], "Language Barriers," *Econometrica* **81**, 781–812.

[6] Cai, Hongbin, and Joseph Tao-Yi Wang [2006], "Over-Communication in Strategic Information Transmission Games," *Games and Economic Behavior* **56**, 7-36.

[7] Carnap, Rudolf [1947], *Meaning and Necessity: A Study in Semantics and Modal Logic*, Chicago: University of Chicago Press.

[8] Crawford, V.P.and J. Sobel [1982], "Strategic Information Transmission," *Econometrica* **50**, 1431–1451.

[9] Crawford, Vincent P. and Hans Haller [1990], "Learning how to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica* **58**, 571-595.

[10] Crawford, Vincent P. [2003], "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions," *American Economic Review* **93**, 133–149.

[11] Davidson, Donald [1967], "Truth and Meaning," in: Kulas J., Fetzer J.H., Rankin T.L. (eds) *Philosophy, Language, and Artificial Intelligence. Studies in Cognitive Systems, vol 2*, Springer, Dordrecht.

[12] Davidson, Donald [1986], "A Nice Derangement of Epitaphs," in: Grandy, Richard and Richard Warner (eds) *Philosophical Grounds of Rationality*, Clarendon Press, Oxford.

[13] Farrell, Joseph [1993], "Meaning and Credibility in Cheap-Talk Games," *Games and Economic Behavior* **5**, 514–531.

[14] Frege, Gottlob [1892] "Über Sinn und Bedeutung," *Zeitschrift für Philosophie und philosophische Kritik* **100**, 25–50; translated as "On Sense and Reference," in P.T. Geach and M.Black, (eds.), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford: Blackwell (1952), 56–78.

[15] Giovannoni, Francesco, and Siyang Xiong [2019], "Communication with Language Barriers," *Journal of Economic Theory* **180**, 274-303.

[16] Gordon, Sidartha, Navin Kartik, Melody Pei-yu Lo, Wojciech Olszewski, and Joel Sobel [2021], "Effective Communication in Cheap-Talk Games," Working Paper, University of California - San Diego.

[17] Grice, H. Paul [1957], "Meaning," *The Philosophical Review* **3**, 377-388.

[18] Kalai, Ehud and Dov Samet [1984], "Persistent Equilibria in Strategic Games," *International Journal of Game Theory* **13**, 129–144.

[19] Kartik, Navin, Marco Ottaviani and Francesco Squintani [2007], "Credulity, Lies, and Costly Talk," *Journal of Economic Theory* **134**, 93–116.

[20] Kripke, Saul A. [1982], *Wittgenstein on Rules and Private language: An Elementary Exposition*, Harvard University Press, Cambridge MA.

[21] Lewis, David [1969], *Convention: A Philosophical Study*, Harvard University Press, Cambridge, MA.

[22] Lipnowski, Elliot and Doron Ravid [2020], "Cheap Talk with Transparent Motives," *Econometrica* **88**, 1631–1660.

[23] Michaelson, Eliot and Marga Reimer [2019], "Reference," *The Stanford Encyclopedia of Philosophy (Spring 2019 Edition)*, Edward N. Zalta (ed.) URL = https://plato.stanford.edu/archives/spr2019/entries/reference/

[24] Mill, John Stuart [1969], *A System of Logic,*, Longmans, Green, and Co., London.

[25] Miller, Alexander [2018], *Philosophy of Language,*, Routledge, New York, NY.

[26] Olszewski, Wojciech [2006], "Rich Language and Refinements of Cheap-Talk Equilibria," *Journal of Economic Theory* **128**, 164-186.

[27] Quine, Willard van Orman [1960], *Word & Object*, MIT Press, Cambridge, MA.

[28] Rabin, Matthew [1990], "Communication Between Rational Agents," *Journal of Economic Theory* **51**, 144-170.

[29] Speaks, Jeff [2021], "Theories of Meaning," *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/spr2021/entries/meaning/

[30] Stalnaker, Robert [2006] “Saying and Meaning, Cheap Talk and Credibility,” In: Benz A., Jäger G., van Rooij R. (eds) *Game Theory and Pragmatics*, Palgrave Studies in Pragmatics, Language and Cognition. Palgrave Macmillan, London.

[31] Tarski, Alfred [1944], “The Semantic Conception of Truth: and the Foundations of Semantics,” *Philosophy and Phenomenological Research*, **4**, 341–376.

[32] Voorneveld, Mark [2004], “Preparation,” *Games and Economic Behavio* **48**, 403–414.

[33] Wang, Joseph Tao-yi, Michael Spezio, and Colin F. Camerer [2010], “Pinocchio’s Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games,” *American Economic Review* **100**, 984–1007.

[34] Wittgenstein, Ludwig [1958], *Philosophical Investigations*, Prentice Hall, Upper Saddle River, NJ.