Language Barriers*

Andreas Blume
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260

Oliver Board
New York University School of Law
40 Washington Square South
New York, NY 10012

First draft: October, 2009
This version: Preliminary revision, February, 2011

Abstract

Different people use language in different ways. Private information about language competence can be used to reflect the idea that language is imperfectly shared. In optimal equilibria of common interest games there will generally be some benefit from communication with an imperfectly shared language, but the efficiency losses from private information about language competence in excess of those from limited competence itself may be significant. In optimal equilibria of common-interest sender-receiver games, private information about language competence distorts and drives a wedge between the indicative meanings of messages (the decision-relevant information indicated by those messages) and their imperative meanings (the actions induced by those messages). Indicative meanings are distorted because information about decision relevant information becomes confounded with information about the sender's language competence. Imperative meanings of messages become distorted because of the uncertain ability of the receiver to decode them. We show that distortions of meanings persist with higher-order failures of knowledge of language competence. In a richer class of games, where both senders and receivers move at the action stage and where payoffs violate a self-signaling condition, these distortions may result in complete communication failure for any finite-order knowledge of language competence.

*We are grateful for comments received from Christopher Gertz, Jiwoong Lee, Joel Sobel and from seminar audiences at the Harvard/MIT Theory seminar, the Trimester Program on Mechanism Design at the Hausdorff Institute for Research in Mathematics (Universität Bonn), the European University Institute, Lehigh University, Maastricht University, Rice University, the University of Bielefeld, the University of Pittsburgh, the University of Texas-Austin, the University of Toronto, the University of Western Ontario, the Econometric Society World Congress (Shanghai), Workshop on Decentralized Mechanism Design, Distributed Computing, and Cryptography sponsored by the Institute for Advanced Study and DIMACS (Princeton) and the 2nd Brazilian Workshop of the Game Theory Society (São Paulo).

1 Introduction

Individuals differ in their language use, express themselves more clearly in some domains than others and do not always agree on the meanings of utterances. The very notion of an *organizational code* (Arrow [1]) presumes a privileged understanding of the code by the members of the organization. If meanings were always clear, there would be no need for *statutory interpretation* of laws by courts (Posner [31], Eskridge, Frickey and Garrett [13]), or to create *trading zones* to mediate communication across subcultures in science (Galison [15]). The following quote from a white paper on electric power transmission in the US illustrates the problem:

> *One of the many difficulties with discussing who should pay for transmission expansion is the surprising lack of a common language for conveying the critical underlying concepts. Important words such as "benefits" and "beneficiaries," and phrases such as "economic upgrades" and "participant funding" are too often used in radically different ways by different parties. At best, the meanings intended by some speakers are not transparent, and different meanings are inferred by different listeners. At worst, the same words have opposite meanings to different people.* (Baldick *et al.* [4], page 12.)

Our aim here is to present a simple, portable, formal framework for expressing the idea that language is imperfectly shared, that some individuals are better equipped to use language and that there can be disagreements about meaning. To this end, we introduce privately known *language competence* into standard communication games. Our approach lets us express that agents are language constrained, that their constraints differ, the different degrees to which agents know a language, the different degrees to which language is shared among agents, how players reason and form beliefs about the language use of others and how they talk about language.[1]

In the examples that Lewis [23] uses to illustrate conventional meaning, meaning is clear: Each state of the world is indicated by one and only one message and each message induces one and only one action. In that case the meaning of a message can be equivalently expressed as the state in which that message is appropriate (its *indicative meaning*) or as the action which is appropriate for that message (its *imperative meaning*). Unlike Lewis, Crawford

[1] Language constraints also appear in Crémer et al [7] and Jäger et al [19], who limit players to finite numbers of messages, Crawford and Haller [10] and Blume [8], who impose symmetry requirements on players' strategies and Rubinstein [33], who deals with agents for whom some objects are nameless and who have access to a limited set binary relations on the set of objects. None of these focusses on communication with an imperfectly shared language.

and Sobel [11] (henceforth CS) concentrate on the case where there is conflict between the communicating parties. They show that when there is not too much conflict, communication is possible but there is a necessary coarsening of meaning. The indicative meaning of a message is now a nontrivial set of states and its imperative meaning the action that is induced by beliefs concentrated on that set of states. The same coarsening that CS derive by introducing conflict is generated even in the common interest case if the language is restricted by limiting the set of available messages, as is done by Crémer et al [7] in their work on optimal organizational codes and Jäger et al [19] on convex categories in optimal (natural) languages.

In our model players have private information about which messages they can send and understand — their language competence — in addition to their decision-relevant private information.[2] This information about language competence is purely instrumental in that each player's payoffs from any decision are completely determined by decision-relevant information. This suggests that it is sensible to continue to think of the indicative meaning of a message as the decision-relevant information conveyed by the message and the imperative meaning as the action induced by that message. With this interpretation, in addition to being coarse, meaning becomes uncertain. The sets of decision-relevant states indicated by a message — its indicative meaning — may vary with the messages available to the player sending the message. The action induced by a message — its imperative meaning — may depend on whether or not the recipient of the message understands it. The indicative meaning and the imperative meaning of a message become imperfectly correlated random variables, thus severing the rigid link between the two.

We are interested in how this uncertainty about message meaning affects the ability of players to communicate and the manner of their communication. In his account of language as a convention, Lewis [23] makes common knowledge an integral part of a convention. We want to know whether and how language can function if it is not common knowledge. Is it possible to communicate at all? Do players avoid messages whose meanings are uncertain? Does the uncertainty about meaning persist if language competence is known but not common knowledge?

Our main focus is on common interest games. Here we find that generally there is some benefit from using an imperfectly shared language, while the efficiency losses from making language competence private information can be severe. We isolate the effects of uncer-

[2] There are sensible ways of expressing imperfectly shared meanings other than through private information about language competence, e.g. communication through noisy channels (Blume, Board and Kawamura [9]), correlated equilibria (De Jaegher [12]) and local interaction (Zollman [34]). We view these as complementary. What our approach adds is a natural way to express different degrees of knowing a language and reasoning about other players' knowledge of language in communication games.

tainty about the sender's language competence from those about the receiver's competence in sender-receiver games. In optimal equilibria of common-interest sender-receiver games where only the sender's language competence is an issue, we show that the sender will always make effective use of all the messages available to her. In these equilibria indicative meaning will generally be distorted in the sense that decision relevant information gets confounded with instrumental information about language.[3] Similarly, if only the receiver's competence is an issue, then imperative meaning will be distorted. Private information about language competence drives a wedge between the indicative meanings of messages and their imperative meanings. We show that these distortions of meanings persist with higher-order failures of knowledge of language competence although they diminish with higher orders of knowledge. In contrast, in a richer class of games, where both senders and receivers move at the action stage and where payoffs violate a self-signaling condition, these distortions may result in complete communication failure for any finite-order knowledge of language competence.

2 Private Knowledge of Language Competence

We begin by setting up a framework that incorporates privately known language competence in a class of two-stage games in which players simultaneously and publicly communicate in the first stage and simultaneously take actions in the second stage.

Players $i = 1, \dots, I$ interact in two stages (we will use I to indicate both the player set and its cardinality). In the communication stage each player sends a message m from a finite set M that is observed by all other players. In the action stage each player takes an action $a_i \in A_i$. At the beginning of the game each player $i \in I$ is privately informed of her *decision type* $t_i \in T_i$ and her *language type* $\lambda_i \subseteq M$. We assume that there is one message, m_0, that is always available to all players and define the set of language types as $\Lambda := \{\lambda \in 2^M | m_0 \in \lambda\}$. Each t_i is drawn from a distribution F_i on T_i and the language type profile $\boldsymbol{\lambda}$ is drawn

[3]Similar difficulties of separating different dimensions of private information from each other arise when all of these dimensions directly impact payoffs. Morgan and Stocken [25] show in a variant of the CS model where the sender is privately informed about her preferences in addition to the state that in equilibrium the sender cannot fully reveal the state even if preferences are fully aligned at the interim stage. It is impossible to completely separate the two dimensions of private sender information. Unlike in our setup, both dimensions directly impact payoffs and the confounding of information about the state with information about preferences is driven by *ex ante* conflict between sender and receiver. Morris [26] demonstrates how a reputational dimension may prevent full revelation of the state even if interim preferences are fully aligned. Levy and Razin [22] show that communication in one (common interest) dimension may be hindered by conflict in another dimension because they are linked through the prior. Sometimes there is a benefit from private information being multidimensional, either because it permits the expert to trade off incentives across dimensions [6] or because uncertainty about the expert's bias leads to a bias that is diminished in expectation [21].

from a distribution π on Λ^I. The distributions $F_1, \ldots, F_I, \pi$ are independent and common knowledge. The profile of decision decision types $\mathbf{t} \in T = \times_{i \in I} T_i$ and the profile of actions $\mathbf{a} \in A = \times_{i \in I} A_i$ determine player i's payoff $U_i(\mathbf{a}, \mathbf{t})$. Player i's language type λ_i is the set of messages that she can send and understand. Messages that she does not understand, she has to treat identically. To express this formally, for any λ_i introduce an equivalence relation $\sim_{\lambda_i}$ on the set of all profiles of messages that is defined by the property that $\mathbf{m} \sim_{\lambda_i} \mathbf{m}'$ if and only if for all $j \in I$ it is the case that $m_j \neq m'_j \Rightarrow m_j, m'_j \in M \setminus \lambda_i$; i.e. player i does not distinguish message profiles that at any give component differ only in messages she does not understand.[4] Each player i's strategy is a pair (σ_i, ρ_i) of a signaling rule $\sigma_i : T_i \times \Lambda_i \rightarrow \Delta(M)$ at the communication stage and a decision rule $\rho_i : T_i \times \Lambda_i \times M^I \rightarrow \Delta(A_i)$ at the action stage. The signaling rule σ_i must satisfy the condition that $\sigma_i(t_i, \lambda_i) \in \Delta(\lambda_i)$ for all $t_i \in T_i$ and $\lambda_i \in \Lambda_i$ and the decision rule the condition $\rho_i(t_i, \lambda_i, \mathbf{m}) = \rho_i(t_i, \lambda_i, \mathbf{m}')$ for all $t_i \in T_i$, $\lambda_i \in \Lambda_i$ and for all $\mathbf{m}, \mathbf{m}' \in M^I$ with $\mathbf{m} \sim_{\lambda_i} \mathbf{m}'$. We refer to these two conditions as player i's language constraints.

3 Universal Private Language Constraints

Our initial focus is on the case where all players communicate, take actions and face privately known language constraints. Here we make two observations: (1) There is generally a role for communication even with privately known language competence – for language to be useful, it does not have to be common knowledge. (2) Universal private information about language constraints may imply a significant efficiency loss in environments where the loss from partial private information is negligible.

[4]There are many natural ways enrich this framework: (1) One can allow players to make some, albeit coarse, distinctions among messages that they do not understand by letting a language type be a pair $(\lambda_i, \mathcal{P}_i)$, where $\lambda_i \subseteq M$ is the set of messages that player i can send and $\mathcal{P}_i$ is a partition of M that satisfies $m \in \lambda_i \Rightarrow \{m\} \in \mathcal{P}_i$ and indicates which distinctions the player can make among messages. This would allow one to capture the phenomenon that the agent does not understand the difference between "metaphysics" and "dialectics" but places both in philosophy. (2) Instead of letting players respond to unknown messages strategically, as we do, one could introduce a nonstrategic default interpretation of unknown messages. (3) One could permit the sender to send some messages that she does not understand, by letting a language type be a pair $(\mathcal{Q}_i, \mathcal{P}_i)$ of (possibly identical) partitions of M, where as before $\mathcal{P}_i$ captures the distinctions she can make among received messages and where she has to treat messages in any element of $\mathcal{Q}_i$ identically by randomizing uniformly over those messages. This would capture the player being able to use the terms "dialectics" and "metaphysics", but without being able to differentiate their meanings.

3.1 A Role for Communication

Assume that A_i and T_i are finite for all i and that all players have common interests, i.e. there is a function $U : A \times T \to \mathbb{R}$ such that $U_i = U$ for all $i \in I$. All distributions, $F_1, \ldots, F_I$ and π have full support. Suppose that U has a unique maximizer $\mathbf{a}(\mathbf{t})$ for every $\mathbf{t}$ and that M contains at least three messages, m_0, m' and m''. We call such a common interest game *information responsive* if it satisfies the condition

$$\exists \mathbf{t}', \ell, t''_\ell \text{ such that } \mathbf{a}_j(\mathbf{t}') \neq \mathbf{a}_j(\mathbf{t}'_{-\ell}, t''_\ell) \ \forall j \in I.$$

Proposition 1 *In information-responsive common-interest games, there is an optimal equilibrium with communication that is strictly superior to any equilibrium without communication.*

Proof: Since we have a finite game, the problem of finding a profile of strategies $\alpha_i : T_i \to A_i$ in the game without communication that maximize joint payoffs,

$$\max_{\alpha_1,\ldots,\alpha_I} \sum_{t \in T} U(\alpha_1(t_1), \ldots, \alpha_I(t_I), t) F(t),$$

has a solution, $\hat{\alpha}$. Evidently, there is no loss in restricting attention to pure strategies and given that we have a common-interest game, the profile $\hat{\alpha}$ forms an equilibrium.

For every player i let $\tilde{\lambda}_i$ be a language type with $m', m'' \in \tilde{\lambda}_i$ and recall that by our full-support assumption on π any such language type has positive probability. In the communication game, consider the strategy profile (σ, ρ) that prescribes for all players $i \neq \ell$ the signaling rule $\sigma_i(t'_i, \tilde{\lambda}_i) = m'$ and $\sigma_i(t_i, \lambda_i) = m_0$ for all $(t_i, \lambda_i) \neq (t'_i, \tilde{\lambda}_i)$ and for player ℓ the signaling rule $\sigma_\ell(t'_\ell, \tilde{\lambda}_\ell) = m'$, $\sigma_\ell(t''_\ell, \tilde{\lambda}_\ell) = m''$ and $\sigma_\ell(t_\ell, \lambda_\ell) = m_0$ for all $(t_\ell, \lambda_\ell) \neq (t'_\ell, \tilde{\lambda}_\ell), (t''_\ell, \tilde{\lambda}_\ell)$ at the communication stage. Define $\mathbf{t}'' := (\mathbf{t}'_{-\ell}, t''_\ell)$, $\mathbf{m}' := (m', m', \ldots, m')$ and $\mathbf{m}'' := (m', m', \ldots, m', m'', m', \ldots, m')$ (with m'' in the ℓth component).

At the action stage, let the strategy profile prescribe the action rule $\rho_i(t_i, \tilde{\lambda}_i, \mathbf{m}') = \mathbf{a}_i(\mathbf{t}')$, $\rho_i(t_i, \tilde{\lambda}_i, \mathbf{m}'') = \mathbf{a}_i(\mathbf{t}'')$ and $\rho_i(t_i, \lambda_i, \mathbf{m}) = \hat{\alpha}_i(t_i)$ otherwise.

Then, for any decision type profile $\mathbf{t} \neq \mathbf{t}', \mathbf{t}''$, the *ex post* payoff in the communication game is the same as in the game without communication. For any decision type profile $\mathbf{t} = \mathbf{t}', \mathbf{t}''$ an *ex post* optimal action profile is chosen whenever the language state $\tilde{\lambda}$ is realized, which occurs with positive probability, and otherwise the *ex post* payoff is the same as in the game without communication. Therefore the *ex post* payoff in the game with communication is never less than the *ex post* payoff in the game without communication. If there is $i \in I$ with

$\hat{\alpha}(t_i') \neq \rho_i(t_i', \tilde{\lambda}_i, \mathbf{m}')$ then in state $\mathbf{t}'$ the *ex post* payoff in the communication game strictly exceeds the *ex post* payoff in the no-communication game. If, however, $\hat{\alpha}(t_i') = \rho_i(t_i', \tilde{\lambda}_i, \mathbf{m}')$ for all $i \in I$, then it must be the case that $\hat{\alpha}_i(t_i') \neq \rho_i(t_i', \tilde{\lambda}_i, \mathbf{m}'')$ for all $i \neq \ell$, in which case in state $\mathbf{t}''$ the *ex post* payoff in the communication game strictly exceeds the no-communication payoff.

Therefore the *ex ante* payoff from the profile (σ, ρ) in the communication game strictly exceeds the payoff from any optimal profile $\hat{\alpha}$ in the game without communication. While (σ, ρ) itself need not be an optimal profile in the communication game, since the game is finite and has the common interest property an optimal profile (σ^*, ρ^*) exists. Using the fact that we have a common-interest game once more, it follow that (σ^*, ρ^*) is an equilibrium. Therefore the communication game has an optimal equilibrium and this equilibrium has a strictly higher payoff than the optimal equilibrium of the game without communication. □

Note that this result tells us nothing about the form of the optimal equilibria. It simply utilizes the fact that when the distribution of language types has full support, there will be instances in which it is possible to accurately reveal the profile of decision types, to signal universal comprehension of the relevant messages, and to take the corresponding optimal profile of actions. It may, however, not be optimal to ever fully reveal the state, and the manner in which decision types pool on messages may vary with their language types, thus confounding message meanings. Message meaning may be further confounded because players may be unable to signal message comprehension. These questions, which are of central interest to us, we will investigate later in environments with more structure.

3.2 The interaction of uncertainties about language competence

Our purpose in this section is to show, via a two-player example, that universal uncertainty about language constraints can lead to significant efficiency losses even when the constraints themselves or one-sided uncertainty imply no substantial loss.

For the next result assume that there are two players. At the action stage each player has the choice among going to one of $2n$ locations, i.e. $A_i = \{1, 2, \dots, 2n\}$, where we assume that $n \geq 2$. Locations are either good or bad. If both players choose the same location and that location is good, then both receive a payoff of 1. Otherwise their common payoff is 0. Exactly one of the first n locations is good and exactly one of the second n locations is good. Good locations are drawn independently from uniform distributions on the sets $\{1, \dots n\}$ and $\{n+1, \dots, 2n\}$ respectively. Player 1 privately knows which of the first n locations is good. Thus 1's decision type set is $T_1 = \{1, 2, \dots, n\}$ with $t_1 \in T_1$ indicating the good location. Similarly, player 2 privately knows which of the second n locations is good. Her decision type set is $T_2 = \{n+1, \dots, 2n\}$ with $t_2 \in T_2$ indicating the good location. For each $\ell = 1, 2, \dots$ the finite message space M has at least $4 \times \ell$ elements (for notational purposes we do not explicitly index M by ℓ). $\#(\lambda_i) = 2\ell$ with probability one for both individuals and $\#(\lambda_1 \cap \lambda_2) = \ell$ with probability one. Pairs of language types (λ_1, λ_2) are drawn from a uniform distribution on the set $\{(\lambda_1, \lambda_2) \in 2^M \times 2^M | \#(\lambda_1 \cap \lambda_2) = \ell, \#(\lambda_i) = 2\ell, i = 1, 2\}$.[5]

To accommodate all cases of players having or not having access to information about their counterpart's language competence, we write strategies as functions of the entire language state and make the appropriate restrictions when part of the state is private information. Then a (behaviorally mixed) strategy for player i is a pair (σ_i, ρ_i) consisting of a signaling rule $\sigma_i : \Lambda_i \times \Lambda_{-i} \times T_i \to \Delta M$ where $\sigma_i(\lambda_i, \lambda_{-i}, t_i) = \sigma_i(\lambda_i, \lambda'_{-i}, t_i)$ $\forall \lambda_i \in \Lambda_i, \lambda_{-i}, \lambda'_{-i} \in \Lambda_{-i}, t_i \in T_i$ if i does not know $-i$'s language competence, and an action rule $\rho_i : \Lambda_i \times \Lambda_{-i} \times T_i \times M \to \Delta A_i$, so that $\rho_i(\lambda_i, \lambda_{-i}, t_i, m_{-i})$ is player i's action as a function of player $-i$'s message $m_{-i} \in M$, where $\rho_i(\lambda_i, \lambda_{-i}, t_i, m_{-i}) = \rho_i(\lambda_i, \lambda'_{-i}, t_i, m_{-i})$ $\forall \lambda_i \in \Lambda_i, \lambda_{-i}, \lambda'_{-i} \in \Lambda_{-i}, t_i \in T_i, m_{-i} \in M \in$ if i does not know $-i$'s language competence, with the understanding that both message and action rules respect player i's language constraints.

Observation 1 *In the two-player location choice game: (1) If the language competence of both players is private information, then the common efficient payoff is bounded from above by a value $\overline{v} < 1$ for all ℓ. (2) If at least one player knows the language competence of the other, the common efficient equilibrium payoff converges to one as $\ell \to \infty$.*

Proof: For (1) observe that for any message in λ_i that player i may send, that message does not belong to λ_{-i} with probability 1/2, in which case player $-i$'s action does not depend on the message sent by player i and the probability off being able to coordinate on location t_i is at most a $1/n$. Since the probability that neither player sends a message that the other player understands is 1/4, the optimal payoff when both language competences are private information is bounded from above by $\frac{3}{4} + \frac{1}{4}\frac{1}{n} < 1$. To show (2), without loss of generality, assume that it is player 1 who knows player 2's language competence λ_2. For any $\ell \geq n$ there are $\lfloor \frac{2\ell}{n} \rfloor$ mutually exclusive subsets of λ_2 of size n, $S_1(\lambda_2), S_2(\lambda_2), \dots, S_{\lfloor \frac{\ell}{n} \rfloor}(\lambda_2)$. For any λ_2 define a function $\phi_{\lambda_2} : \lambda_2 \to T_1$ with the property that the restriction of ϕ_{λ_2} to any set $S_i(\lambda_2)$ is a bijection denoted ϕ_{i,λ_2}. At the communication stage let player 1 use the message rule $\sigma_1 : \Lambda_1 \times \Lambda_2 \times T_1 \to M$ that is defined by

$$\sigma_1(\lambda_1, \lambda_2, t_1) = \begin{cases} \phi^{-1}_{i^*,\lambda_2}(t_1) & \text{if } i^* = \min\{i | \phi^{-1}_{i,\lambda_2}(t_1) \in \lambda_1\} \\ m_{(\lambda_1,\lambda_2)} & \text{otherwise} \end{cases}$$

[5]Note that in this example the language types are not drawn independently from each other.

where $m_{(\lambda_1,\lambda_2)}$ is an arbitrary element of $\lambda_1 \cap \lambda_2$ and let player 2 use an arbitrary message rule σ_2. At the action stage let player 1 use the action rule ρ_1 that is defined by $\rho_1(\lambda_1, \lambda_2, t_1, m_2) = t_1$ for all $(\lambda_1, \lambda_2, t_1, m_2)$ and let player 2 use the action rule ρ_2 that is defined by

$$\rho_2(\lambda_2, \lambda_1, t_2, m_1) = \begin{cases} \phi_{\lambda_2}(m_1) & \text{if } \ m_1 \in \lambda_2 \\ t_2 & \text{otherwise} \end{cases}$$

for all $(\lambda_2, \lambda_1, t_2, m_1)$.

Notice that whenever the set $\{i|\phi^{-1}_{i,\lambda_2}(t_1) \in \lambda_1\}$ is nonempty, the above strategy profile guarantees both players a payoff of 1. Conditional on λ_2, the probability that the set $\{i|\phi^{-1}_{i,\lambda_2}(t_1) \in \lambda_1\}$ is empty is less than $\left(\frac{1}{2}\right)^{\lfloor \frac{\ell}{n} \rfloor}$, which converges to zero as $\ell \to \infty$. Therefore, as $\ell \to \infty$ we have a sequence of games and corresponding strategy profiles with payoffs converging to 1. The result follows from the fact that in finite common interest games, optimal profiles are equilibrium profiles, an optimal profile exists and the payoff from an optimal profile is bounded below by the payoff from the profile we constructed. □

4 Sender-Receiver Games

In this section we restrict attention to sender-receiver games, and separately analyze the cases where only the language competence of the sender is the issue and where only the language competence of the receiver is the issue.

4.1 Language Competence of the Sender

A privately informed sender, S, communicates with a receiver, R, by sending one of a finite number of messages $m \in M$, where $\#(M) \geq 2$. The payoffs $U^S(a,t)$ and $U^R(a,t)$ of the sender and the receiver depend on the receiver's action, $a \in A = \mathbb{R}^\ell$, and the sender's payoff-relevant information $t \in T$, her *decision type*; we assume that T is a convex and compact subset of $\mathbb{R}^\ell$ that has a nonempty interior. It is common knowledge that the sender's decision type is drawn from a distribution F with density f that is everywhere positive on T. The function U^S is differentiable and strictly concave in a for every $t \in T$. Denote the set of distributions over T by $\Delta(T)$ and assume that the receiver has a unique best reply $\hat{\rho}(\mu)$ to any belief $\mu \in \Delta(T)$, and for any measurable set $\Theta \subset T$, sightly abusing notation, denote by $\hat{\rho}(\Theta)$ his optimal response to his prior belief concentrated on Θ. Assume that for all $t' \neq t$, $\hat{\rho}(t') \neq \hat{\rho}(t)$. Note that for any set $\Theta \subset T$ that has positive probability and any set Θ^0 that has zero probability,

$$\hat{\rho}(\Theta) = \hat{\rho}(\Theta \setminus \Theta^0).$$

For any $\Theta \subset T$ and any two actions $a_1 \in A$ and $a_2 \in A$ define

$$\Theta_{a_1 \succsim a_2} := \{t \in \Theta | U^S(t, a_1) \geq U^S(t, a_2)\},$$

the set of types in Θ who prefer action a_1 to action a_2, and similarly define $\Theta_{a_1 \succ a_2}$ for strict preference, and $\Theta_{a_1 \sim a_2}$ for indifference. Note that for any measurable set $\Theta \subset T$ and for any pair $a_1, a_2 \in A$ with $a_1 \neq a_2$, the continuity of the sender's payoff function implies that the sets $\Theta_{a_1 \succ a_2}, \Theta_{a_2 \succ a_1}$ and $\Theta_{a_1 \sim a_2}$ are measurable. Assume that for any two $a_1, a_2 \in A$ with $a_1 \neq a_2$, $\text{Prob}(T_{a_1 \sim a_2}) = 0$. This implies that $\text{Prob}(\Theta) = \text{Prob}(\Theta_{a_1 \succ a_2} \cup \Theta_{a_2 \succ a_1})$. For any finite set of K actions $\{a_1, \ldots, a_K\}$ with $2 \leq K \leq M$ define $\Theta_{a_1 \succsim a_2, \ldots, a_K} := \cap_{n=2}^{K} \Theta_{a_1 \succsim a_n}$, the set of sender types who prefer action a_1 over actions $a_2, \ldots, a_K$, and use Ω to denote the collection of all such sets.

Assumption 1 *(A) For any $\Theta \in \Omega$ and any pair of actions a_1, $a_2 \subset A$ such that $\Theta_{a_1 \succ a_2}$ and $\Theta_{a_2 \succ a_1}$ both have positive probability, $\hat{\rho}(\Theta_{a_1 \succ a_2}) \neq \hat{\rho}(\Theta)$. (B) For any belief μ, there exists a type $t(\mu)$ such that $\hat{\rho}(\mu) = \hat{\rho}(t(\mu))$.*

Part (A) of Assumption 1 formalizes the idea that the optimal receiver response is sufficiently sensitive to beliefs. This is the key assumption that ensures that the receiver responds differently to a message, depending on knowing whether or not the sender has alternative attractive messages available. Part (B) requires that any best response to some belief is also the receiver's ideal point for some state of the world. Essentially it says that there are no gaps in the type space.

We will assume that not every message $m \in M$ may be available to the sender. Instead the sender privately learns a set $\lambda \subset M$ of available messages, her *language type*.[6] One message, $m_0 \in M$ is assumed to be always available. Thus the sender's language type λ is drawn, independently from her decision type t, from a commonly known distribution π on $\Lambda = \{\lambda \in 2^M | m_0 \in \lambda\}$, the set of all subsets of M that contain the message m_0. As usual, we assume that this entire structure is common knowledge. Having the message space be common knowledge but not necessarily all messages available can be interpreted as the sender having a description of the message space that is common for some messages and

[6]Our distinction between decision types and language types is a convenient terminological device. Of course, one could follow Harsanyi [17] and express the inability of the sender to send a particular message by assigning an arbitrarily large negative payoff to doing so. This would not affect our results but would, in our view, obscure the fact that ultimately both parties are interested in communicating information about t. Any information transmission about language competence is merely instrumental. Finally, note that we will leave the analysis of a still more general model in which different messages are available at different privately known costs for later work.

private for others; this is the situation when a speaker of one natural language knows that another natural language is as expressive as her own but is not proficient in that language.

A sender strategy is a mapping $\sigma : T \times \Lambda \to \Delta(M)$ that satisfies the condition $\sigma(t, \lambda) \in \Delta(\lambda)$. A receiver strategy is a mapping $\rho : M \to A$.[7] We study perfect Bayesian Nash equilibria (σ, ρ, β) where β is a belief system that is derived from the sender's strategy σ by Bayes' rule whenever possible, the sender's strategy σ is a best reply to the receiver's strategy ρ, and ρ is a best reply after every message, given the belief system β.

4.1.1 Example

The following example illustrates how the indicative meanings of messages may be compromised when there is private information about language competence. The focus is on the language competence of the sender.

The receiver is uncertain about the indicative meaning of equilibrium messages because he is uncertain about the sender's language competence. There will be a message for which he is unable to determine whether the sender sent this message because no other message was available or because she preferred to send it in lieu of another available message. His equilibrium response to that message will be a compromise that averages over the possibility that the sender pooled over all decision types and the alternative that the sender used the message to indicate a strict subset of the set of decision types.

Example 1 *Assume that the sender's decision type is drawn from a uniform distribution on the interval [0,1]. Sender and receiver have common interests and receive identical payoffs* $-(t-a)^2$ *when the sender's decision type is* t *and the receiver takes action* a*. The message space is* $M = \{m_0, m_1\}$ *and the language type distribution* π *assigns positive probability to two language types,* $\lambda_0 = \{m_0\}$ *and* $\lambda_1 = \{m_0, m_1\}$*, where* $\pi(\lambda_1) = p$ *and* $\pi(\lambda_0) = 1 - p$*. Consider an equilibrium in which the sender adopts a strategy of the following form:*

- *if the sender's language type is* λ_0*, send message* m_0 *for all* $t \in [0, 1]$
- *if the sender's language type is* λ_1*, send message* m_0 *for* $t \in [0, \theta_1)$ *and message* m_1 *for* $t \in [\theta_1, 1]$

The receiver's best response to this strategy is to choose action a_0 *if he received message* m_0

[7] The restriction to pure strategies for the receiver is without loss of generality because of our assumption that the receiver has a unique best reply given any belief.

and a_1 if m_1, where a_0 and a_1 are given by

$$a_0 = \frac{(1-p)\frac{1}{2} + p\theta_1\frac{\theta_1}{2}}{(1-p)+p\theta_1}, \text{ and}$$
$$a_1 = \frac{\theta_1+1}{2}.$$

(Note that these actions are equal to the receiver's expectation of t conditional on the message received.) We have an equilibrium if the sender of type θ_1 is indifferent between a_0 and a_1, i.e.

$$\theta_1 = \frac{a_0+a_1}{2}$$
$$\Rightarrow \quad \theta_1 = \frac{4p+\sqrt{9-8p}-3}{4p}$$

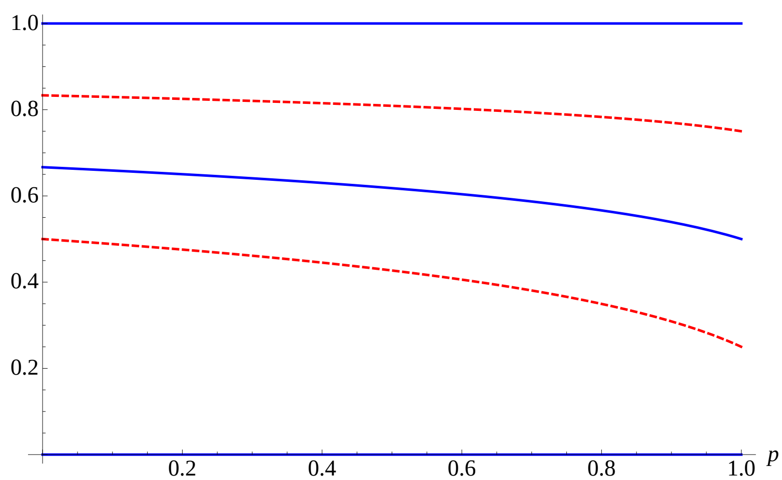

Figure 1: a_0, θ_1, and a_1

Figure 3 plots the equilibrium actions a_0 and a_1 chosen by the receiver (dashed red), and the cutoff type θ_1 for the sender (solid blue) as functions of p, the probability that the second message available. Notice that for low values of p, there is considerable distortion *in the choice of a_0 compared with what it would be if the receiver knew that both messages were available ($\theta_1/2$); similarly, for high values of p, there is significant distortion compared with what the receiver would choose if he knew that only one message were available (1/2). There is no such distortion in the choice of a_1, because if m_1 is observed, the receiver knows that both message were available.*

Despite the fact that in the previous example the sender is language constrained and the exact nature of the constraint is private information, the sender can reliably induce all equilibrium actions. In that sense, while the indicative meaning of messages is imprecise, their imperative meaning remains precise. We will later (Section 4.2) see how to obtain imprecision of imperative meaning by introducing uncertainty about the language competence of the receiver. An alternative route to imprecision of imperative meaning is through higher-order uncertainty about the sender's language competence. Then the sender finds herself unable to reliably induce specific equilibrium actions because she does not know the receiver's belief about her language competence. This we will examine in detail in Section 5.

4.1.2 Distortion of Indicative Meaning

Inspired by Lewis [23], we refer to the *indicative meaning* of a message as the (decision-relevant) information about the sender that is conveyed by the message. Distortions of indicative meaning arise when the receiver's strategy fails to be optimal given the sender's language competence.

Definition 1 *There is* distortion of indicative meaning *in equilibrium* (σ, ρ, β) *if there exists an language type* λ *and* $m \in \lambda$ *that is used with positive probability by* λ *such that* $\rho(m)$ *is not optimal for the receiver conditional on the language type* λ *being revealed.*

Distortions of indicative meaning need not arise if only a few actions are induced in equilibrium and, given the equilibrium strategy of the receiver, the sender is never constrained by her language ability so that for every action that can be induced she always has a message that induces that actions. This is, trivially, the case in pooling equilibria.[8] Intuitively, however, the more information is transmitted and the more actions are induced in equilibrium the more likely it is that there will be distortions of indicative meaning. Those language types of the sender who have access to fewer message will sometimes find themselves language constrained and forced to send messages that they would prefer not to send if they had access

[8] It is also possible to find games with equilibria in which there is positive probability that the sender is unable to induce some of the equilibrium actions of the receiver, but there is no distortion of indicative meaning. A simple example is this: decision types are uniformly distributed on $[0, 1]$, sender and receiver have identical payoffs $-(t-a)^2$, the sender's set of available messages is $\lambda_1 = \{m_0, m_1, m_2\}$ with probability p and $\lambda_0 = \{m_0\}$ otherwise. Regardless of the value of p, there is an equilibrium in which language type λ_1 divides the decision type space into three equal-length intervals, sends message m_1 for decision types in the interval $\left(0, \frac{1}{3}\right)$, sends message m_0 for decision types in the interval $\left(\frac{1}{3}, \frac{2}{3}\right)$, and sends message m_2 for decision types in the interval $\left(\frac{2}{3}, 1\right)$. This equilibrium is optimal and there is no distortion of indicative meaning: Conditional on observing either message m_1 of m_2, the receiver knows the sender's language type and after message m_0, the sender's language type is irrelevant to him. Note that this example is non-generic because it depends on the fact that the receiver's pooling action coincides with one of the actions in a three-step equilibrium.

to a larger set of messages. Thus different language types will pool on the same message for different sets of decision types. When receiving such messages the receiver best responds by averaging over these sets of decision types and will generally take an action that differs from the action he would take if he knew the sender's language type and therefore did not have to average. The following proposition formalizes this observation.

Proposition 2 *There will be distortion of indicative meaning in any equilibrium* (σ, ρ, β) *for which there is a message* $m^* \in M$ *and a pair of language types* $\lambda^* \neq \tilde{\lambda}$ *such that* $\lambda^* = \tilde{\lambda} \cup \{m^*\}$, $\pi(\tilde{\lambda}) \neq 0$, $\pi(\lambda^*) \neq 0$, λ^* *uses all of her available messages with positive probability and all those messages induce distinct actions.*[9]

Proof: Since m_0 is always available, the set $\tilde{\lambda}$ is not empty. The fact that λ^* uses all of her messages with positive probability and all of those messages induce distinct actions, and using the fact that for any two $a_1, a_2 \in A$ with $a_1 \neq a_2$ we have $\text{Prob}(T_{a_1 \sim a_2}) = 0$, implies that language type $\tilde{\lambda}$ also uses all her messages with positive probability. Hence, there must be a set of decision types, that has positive probability, who use m^* when their language type is λ^* and use a message $\tilde{m} \neq m^*$ when their language type is $\tilde{\lambda}$. Use a^* to denote the action that is induced by m^* and $\tilde{a}$ the action that is induced by $\tilde{m}$. Let $\tilde{\Theta}$ denote the set of decision types who use message $\tilde{m}$ when their language type is $\tilde{\lambda}$. Since $\tilde{\lambda}$ uses all of its messages with positive probability, the set $\tilde{\Theta}$ has positive probability. Similarly, since λ^* uses all of its message with positive probability the set $\tilde{\Theta}_{a^* \succ \tilde{a}}$ of types who switch to message m^* and the set $\tilde{\Theta}_{\tilde{a} \succ a^*}$ of types who continue to send $\tilde{m}$ both have positive probability. The set $\tilde{\Theta}_{\tilde{a} \succ a^*}$ differs at most by a set that has probability zero from the set of decision types who send message $\tilde{m}$ when their language type is λ^*. Hence, if there is no distortion in the equilibrium (σ, ρ, μ), then $\rho(\tilde{m}) = \hat{\rho}(\tilde{\Theta}_{\tilde{a} \succ a^*})$. Also, in the equilibrium (σ, ρ, μ) by assumption $\tilde{\Theta}$ is the set of decision types who send message $\tilde{m}$ when their language type is $\tilde{\lambda}$. Therefore, if there is no distortion, then $\rho(\tilde{m}) = \hat{\rho}(\tilde{\Theta})$. By Assumption 1 however,

$$\hat{\rho}(\tilde{\Theta}_{\tilde{a} \succ a^*}) \neq \hat{\rho}(\tilde{\Theta}),$$

[9] The result is stated in terms of a one-message difference between language types in order to avoid counterexamples like the one in footnote 8. While the one-message difference is sufficient for our purposes and always satisfied if we impose a full support condition on the distribution of language types, it is clearly not necessary for distortion of indicative meaning to arise in equilibrium. After all, if different language types have access to a common message, they have different alternatives to using that message and therefore are likely to use that message for different sets of decision types. Only rarely will the receiver's best responses to beliefs concentrated on these sets of decision type coincide with each other and thus satisfy a necessary condition for absence of distortion of indicative meaning. As an illustration, in the example of footnote 8 any small positive sender bias would resurrect distortion of indicative meaning.

which is inconsistent with having no distortion. □

We observe next that Proposition 2 holds in the setup of Crawford and Sobel [11] (CS). Recall that in the CS model the sender's decision type t is drawn from a differentiable distribution F on $[0,1]$ with a density f that is everywhere positive on $[0,1]$. The receiver takes an action $a \in \mathbb{R}$. It is assumed that the functions U^S and U^R are twice continuously differentiable and, using subscripts to denote partial derivatives, the remaining assumptions are that for each realization of t there exist an action a_t^* such that $U_1^S(a_t^*, t) = 0$; for each t there exists an action a_t' such that $U_1^R(a_t', t) = 0$; $U_{11}^S(a,t) < 0 < U_{12}^S(a,t)$ for all a, t; and, $U_{11}^R(a,t) < 0 < U_{12}^R(a,t)$ for all a and t. For the the next result we mean by 'CS model' the combination of these assumptions about preferences and the decision type distribution with our assumption about the messages space and the distribution of language types.

Corollary 1 *Proposition 2 holds for the CS model.*

Proof: CS preferences satisfy all the conditions we have imposed on sender and receiver utilities. Specifically, Assumption 1 is satisfied because sender and receiver preferences satisfy the single-crossing condition, $U_{12}^S, U_{12}^R > 0$: Single-crossing for the sender implies that for any positive-probability set $\Theta \subset T$ the set $\Theta_{a_1 \succ a_2}$ is of the form $\Theta \cap T_{a_1 \succ a_2}$ where $T_{a_1 \succ a_2}$ is an interval that is either of the form $(-\infty, t)$ or of the form (t, ∞). Hence, the distribution that is the prior probability concentrated on $\Theta \cap T_{a_1 \succ a_2}$ either stochastically dominates or is stochastically dominated by the distribution that is the prior probability concentrated on Θ. Therefore the single-crossing condition for the receiver implies that $\hat{\rho}(\Theta_{a_1 \succ a_2}) \neq \hat{\rho}(\Theta)$. □

Another environment in which Proposition 2 holds is one where payoffs can be expressed in terms of convex loss functions and the sender's decision type space T is permitted to be multi-dimensional. Suppose the sender's and receiver's payoffs are given by $U^S(a,t) = \nu_S(||t + b - a||)$ and $U^R(a,t) = \nu_R(||t - a||)$ respectively, where $||\ ||$ is the Euclidean norm and $-\nu_S$ and $-\nu_R$ are strictly increasing convex functions.[10]

[10] Jäger et al [19] have examined the optimal equilibria of this environment, without uncertainty about language competence, for the common-interest case, where $b = 0$. There are well-defined indicative meanings ("categories" in their terminology). In any optimal equilibrium categories are shown to be convex giving rise to a Voronoi tessellation of the type space, and all messages are used with positive probability and induce distinct actions. In the present paper the indicative meanings of messages become more fluid: While it is still the case that in equilibrium each language type partitions the set of decision types into convex sets, at the same time for a given message these sets will generally differ for different language types and it is no longer the case that the set of decision types is partitioned into categories with fixed boundaries. The receiver's posterior distributions after different messages will generally have overlapping supports. For an

Corollary 2 *Proposition 2 holds when sender and receiver have convex loss functions.*

Proof: With convex loss functions every set Θ in Ω will be convex. For any pair of distinct actions a_1 and a_2, the set $T_{a_1 \succsim a_2}$ is a halfspace and thus if $\Theta_{a_1 \succsim a_2} = \Theta \cap T_{a_1 \succsim a_2}$ and $\Theta_{a_2 \succsim a_1} = \Theta \cap T_{a_2 \succsim a_1}$ have positive probability, they are convex and have a nonempty interior. If we denote the interior of a set X by $\text{int}(X)$ then $\hat{\rho}(\Theta_{a_1 \succsim a_2}) \in \text{int}(\Theta_{a_1 \succsim a_2})$ and $\hat{\rho}(\Theta_{a_2 \succsim a_1}) \in \text{int}(\Theta_{a_2 \succsim a_1})$. To see this, let

$$V(a, K) = \int_K \nu_R(||t - a||) f(t) dt$$

for a convex set K and consider a point $\overline{a}$ on the boundary of K. By the supporting hyperplane theorem, there exists a vector $c \neq 0$ with $c \cdot t \geq c \cdot \overline{a}$ $\forall t \in K$. Furthermore, $c \cdot t > c \cdot \overline{a}$ $\forall t \in \text{int}(K)$. The derivative of $V(\cdot, K)$ at $\overline{a}$ in the direction c satisfies

$$\nabla V(\overline{a}, K) \cdot \frac{c}{||c||} = \int_K \nu_R'(||t - \overline{a}||) \frac{1}{2} ||t - \overline{a}||^{-\frac{1}{2}} (\overline{a} - t) \cdot \frac{c}{||c||} f(t) dt > 0$$

because ν_R is increasing and $(\overline{a} - t) \cdot \frac{c}{||c||} > 0$ for almost all $t \in K$. Use a_{12} to denote $\hat{\rho}(\Theta_{a_1 \succsim a_2})$ and a_{21} to denote $\hat{\rho}(\Theta_{a_2 \succsim a_1})$. Since $a_{12} \notin \Theta_{a_2 \succsim a_1}$, there exists a vector $d \neq 0$ with $d \cdot t \geq d \cdot a_{12}$ $\forall t \in \Theta_{a_2 \succsim a_1}$ (and $>$ for all $t \in \text{int}(\Theta_{a_2 \succsim a_1})$). Consider the derivative of $V(\cdot, \Theta)$ at a_{12} in the direction d:

$$\begin{aligned} \nabla V(a_{12}, \Theta) \cdot \frac{d}{||d||} &= \nabla V(a_{12}, \Theta_{a_1 \succ a_2}) \cdot \frac{d}{||d||} + \nabla V(a_{12}, \Theta_{a_2 \succ a_1}) \cdot \frac{d}{||d||} \\ &= \nabla V(a_{12}, \Theta_{a_2 \succ a_1}) \cdot \frac{d}{||d||} \\ &= \int_{\Theta_{a_2 \succ a_1}} \nu_R'(||t - a_{12}||) \frac{1}{2} ||t - a_{12}||^{-\frac{1}{2}} (a_{12} - t) \cdot \frac{d}{||d||} f(t) dt > 0, \end{aligned}$$

which shows that $\hat{\rho}(\Theta_{a_1 \succsim a_2}) \neq \hat{\rho}(\Theta)$. □

extreme example, if instead of always permitting silence, we required the availability distribution to have full support on the power set of M, then trivially in any equilibrium the receiver's posterior would have full support on T after every message. We will show below that in our setting with common interests it remains true that all messages (that are available to some language type) will be used and that therefore by Proposition 2 there will be distortion of indicative meaning in optimal equilibria.

4.1.3 Common-Interest Games

In this section we consider the case where sender and receiver have identical preferences, $U^S \equiv U^R \equiv U$. We show that an optimal equilibrium exists. Furthermore, in any optimal equilibrium all language types use all their messages with positive probability and all available messages induce distinct actions. It is interesting that this holds despite the fact that, as we showed above, different language types using all their messages may lead to distortion of indicative meaning.

First-order intuition for why every language type uses all of her messages is simple: unused messages can be introduced to refine the information that the sender transmits. A complication arises because other language types may already use that message and may see their payoffs reduced as the action induced by that message changes. We will show, however, that the magnitude of such losses is of second order in comparison to the gains of the language type who begins using that message.

We proceed by first establishing existence of an optimal strategy profile. Here we argue in terms of the receiver's strategy ρ which, as we will see, can be viewed as a point in the compact set T^M.[11] We construct a function that assigns to each strategy of the receiver the payoff that results from the sender using a best response to that strategy. Under our assumptions this function is continuous. Hence, we face the problem of maximizing a continuous function over a compact set, which has a solution. Therefore an optimal strategy profile exists and since we have a common interest game, this profile must be part of an equilibrium profile.

For each language type and any optimal receiver strategy, one can partition the set of decision types into subsets for whom the same message is optimal. We will show that each language type induces every action that she can achieve with her repertoire of messages on a set of decision types that has positive probability. Hence, if she does not use one of her messages, it must be because one of her other messages induces the same action. Then, if there is an language type who does not use all of her messages, we can take a pair of messages that induce the same action a, one of which is used by the language type under consideration and one of which is not. Split the subset of decision types who induce action a into two positive-probability subsets and have one of these subsets continue to use the message they used before and while the other subset switches to the formerly unused message m. Other language types may already have been using message m, but note that since we are considering an optimal strategy profile the receiver's response to message m was itself optimal. Therefore an infinitesimal change in the response to m results in a first-order common loss that is zero when the expectation is taken over the types who used message

[11] This result generalizes the corresponding one of Jäger et al [19] to environments with private information about language competence.

m to begin with. At the same time there is a positive first-order gain for the language type who starts using message m because she transmits useful information to the receiver. The following results formalize this intuition.

Lemma 1 *With common interests, there exists an optimal strategy profile.*

Proof: Without loss of generality we can confine attention to receiver strategies for which each action is a best response to some belief. Then, by Assumption 1 each receiver strategy prescribes only actions that are optimal for some type. Thus receiver strategies can be thought of as associating with each message m the type for whom the action $\rho(m)$ is optimal, i.e. it suffices to think of receiver strategies as elements of T^M. Suppose that for any given strategy ρ of the receiver, the sender uses a best reply; that best reply exists because given the receiver's strategy each sender type maximizes her payoff over a finite set of alternatives. Then the resulting payoff for type (t, λ) equals

$$\max_{m \in \lambda} \{U(\rho(m), t)\}.$$

Given this behavior of the sender, we can assign the following expected payoff to the receiver's strategy ρ:

$$Q(\rho) = \sum_{\lambda \in \Lambda} \pi(\lambda) \int_T \max_{m \in \lambda} \{U(\rho(m), t)\} f(t) dt.$$

Since U and the max operator are continuous functions, the integrand is continuous and therefore by the Lebesgue dominated convergence theorem, Q is continuous. Therefore, by Weierstrass's theorem, Q achieves a maximum on the compact set T^M. $\square$

Note that in an optimal profile the receiver's response after unsent messages is entirely arbitrary and therefore it is without loss of generality to assume that it is the same as after one of the sent messages; if it were not arbitrary, then for some specification the sender would have a profitable deviation which would contradict optimality.

Lemma 2 *In an optimal profile, each language type induces every action a' for which she has a message m' with $\rho(m') = a'$ on a set of decision types that contains an open set and therefore has positive probability.*

Proof: By Assumption 1 and common interest, $\rho(m)$ is some type's ideal point for all $m \in M$. Hence, a' is the ideal action of some type t'. Strict concavity implies that type t'

strictly prefers a' to any of the finitely many other actions she can induce. By continuity this remains true for an open set of types $\mathcal{O}(t')$ containing t' and since f is everywhere positive the set $\mathcal{O}(t')$ has positive probability. $\square$

For CS preferences, the single-crossing condition implies that the set of actions that are optimal for some type is of the form $[\underline{a}, \overline{a}]$ and that with common interest for any belief μ of the sender we have $\hat{\rho}(\mu) \in [\underline{a}, \overline{a}]$. Therefore $\rho(m) \in [\underline{a}, \overline{a}]$ for all $m \in M$, as required by Assumption 1. The assumption also holds for convex loss functions.

Lemma 3 *In an optimal profile all messages of an availability type λ with $\pi(\lambda) > 0$ induce distinct actions.*

Proof: In order to derive a contradiction, suppose not, i.e. there is an language type λ^* with $\pi(\lambda^*) > 0$ with two or more messages that induce the same action. It is without loss of generality to consider an optimal strategy profile in which the sender of any given language type uses only one out of any set of available messages that induce identical actions. Thus, suppose that $m^0, m^1 \in \lambda^*$, $\rho(m^1) = \rho(m^0)$, and λ^* uses m^0, but not m^1. The common *ex ante* payoff from the optimal strategy profile (σ, ρ) equals

$$\sum_{m \in M} \sum_{\lambda \in \Lambda} \pi(\lambda) \int_T U(\rho(m), t) \sigma(m|t, \lambda) f(t) dt.$$

Since all messages that type λ^* uses induce distinct actions, Lemma 2 implies that each of those messages is sent by an open set of types that has positive probability. Let Θ_0 be the set of decision types for which language type λ^* sends message m^0. Recall that different types have different best replies. Therefore we can find a type t_1 that is an element of an open subset of Θ_0 and that satisfies $\hat{\rho}(t_1) \neq \rho(m^1)$. By continuity, for a sufficiently small open ball Θ_1 containing t_1 and satisfying $\Theta_1 \subset \Theta_0$, we have $\hat{\rho}(\Theta_1) \neq \rho(m^1)$. Now alter (only) type λ^*'s behavior by having her split the set Θ_0 on which she sends m^0 into two subsets so that she sends m^1 on Θ_1 and continues to send m^0 on $\Theta_0 \setminus \Theta_1$. Denote the resulting sender strategy by $\tilde{\sigma}$ to distinguish it from the original strategy σ. Note that as long as we do not also modify the receiver strategy, this change in the sender strategy has no effect on the common *ex ante* payoff. If we use a^1 to denote the action that is induced by message m^1, we can define the contribution to the expected payoff from message m^1 as

$$\begin{aligned} W(m^1, a^1) &:= \sum_{\lambda \in \Lambda} \pi(\lambda) \int_T U(a^1, t)\tilde{\sigma}(m^1|t, \lambda) f(t) dt \\ &= \pi(\lambda^*) \int_T U(a^1, t)\tilde{\sigma}(m^1|t, \lambda^*) f(t) dt \\ &\quad + \sum_{\lambda \in \Lambda \setminus \lambda^*} \pi(\lambda) \int_T U(a^1, t)\sigma(m^1|t, \lambda) f(t) dt \\ &= \pi(\lambda^*) \int_T U(a^1, t)\tilde{\sigma}(m^1|t, \lambda^*) f(t) dt \\ &\quad + \sum_{\lambda \in \Lambda} \pi(\lambda) \int_T U(a^1, t)\sigma(m^1|t, \lambda) f(t) dt. \end{aligned}$$

Observe that when we change a^1 we affect the contribution to the *ex ante* payoff from message m^1 only. Also, since a^1 was optimal for m^1 given the original sender strategy, we have

$$\nabla_a W(m^1, a^1) = \pi(\lambda^*) \int_T \nabla_a U(a^1, t)\tilde{\sigma}(m^1|t, \lambda^*) f(t) dt.$$

It follows from our choice of Θ_1 that $\nabla_a W(m^1, a^1) \neq 0$. This implies that the original profile (σ, ρ) was not optimal. □

The following result summarizes our findings and connects them to distortion of indicative meaning.

Proposition 3 *In any common interest game, there exists an optimal equilibrium; in any such equilibrium all messages of an language type that has positive probability induce distinct actions; all such language types use each of their messages with positive probability; and, if the language type distribution π has full support on Λ, there will be distortion of indicative meaning.*[12]

Proof: The first three parts of the proposition summarize Lemmas 1-3. This sets the stage for invoking Proposition 2, which proves the fourth part of the proposition: If the availability distribution π has full support on Λ, there will be pairs of language types both of which have

[12]While all messages that are in the repertoire of an language type induce distinct actions in an optimal equilibrium, it need not be the case that all message in M induce distinct actions in an optimal equilibrium. For example, with two language types $\{m_0, m_1\}$ and $\{m_0, m_2\}$, identical quadratic loss functions and a uniform distribution of decision types, in any optimal equilibrium m_1 and m_2 are synonyms, while there is a non-optimal equilibrium with $\rho(m_0) = \frac{1}{2}$, $\rho(m_1) = \frac{1}{6}$ and $\rho(m_1) = \frac{5}{6}$.

positive probability and which differ only by one available message and by Lemmas 1-3 all of these messages are used by both language types and induce distinct actions. □

Proposition 3 is our key result. It demonstrates the ubiquity of distortion of indicative meaning that results from combining private information about language competence with closely aligned incentives. With congruence of incentives, optimality requires that a large variety of messages will be used; private information about language competence then implies that the receiver cannot always be sure whether a message was sent out of necessity, because more preferable message were not available, or out of a desire to communicate payoff-relevant information.

Note that repeated talk by the sender alone, i.e. replacing the set of messages M by the set M^T of strings of length T that can be formed with the elements of M, is no guarantee for absence of meaning distortions. In particular, the intuition that it may be optimal to first talk about language and then about payoff states is frequently incorrect. This is easiest to see if the availability distribution on the expanded message space M^T that results from letting the sender talk repeatedly is subject to the full support assumption that is used in Proposition 3, in which case this result implies that there is distortion of indicative meaning. Even if the perhaps more natural assumption is made that there is an availability distribution on the set of elementary messages M and any concatenation of a given length of available elements of M is itself available, the logic of Proposition 3 applies: It is generally optimal that all messages in the expanded message space induce distinct receiver replies and language types use all messages in their repertoire. Then it is impossible for any language type λ_0 that is a strict subset of an language type λ_1 to send a message that identifies her language λ_0 because λ_1 would want to send the same message for a positive probability set of decision types.

It should be clear that while the common-interest case is emblematic for what can go wrong with private information about language competence, the insight that there will be distortion of indicative meaning generally also holds when there is conflict of interest, as long as there is not so much conflict as to rule out all communication in equilibrium. There are, however, other more subtle interactions between conflicts of interest and private information about language competence. These we turn to next.

4.2 Language Competence of the Receiver

The recipient of a message is as likely limited by his language competence as the sender is by hers. In this section we propose a simple model in which the receiver's language competence is private information. We show that in general this gives rise to distortion of the imperative meanings of messages. When the receiver's language competence is his private information, then even if he uses a pure strategy and there is no randomness in the transmission channel, the sender can no longer be sure how her message will be interpreted; messages typically induce non-degenerate distributions over receiver actions; and, the sender's strategy is generally not optimal given the receiver's language competence.

For simplicity, in this section we focus exclusively on the receiver's language competence and assume that the sender's language competence is not an issue. We model the receiver's language competence as a partition $\mathcal{P}$ of the message set M, with the interpretation that the receiver cannot distinguish messages that belong to the same partition element $P \in \mathcal{P}$. Formally, we require the receiver's strategy to be measurable with respect to $\mathcal{P}$. The receiver's partition type $\mathcal{P}$ is private information and is drawn from a common-knowledge distribution π_R on the set $\mathbf{P}$ of partitions of M. We restrict attention to CS preferences. In this environment a sender strategy is a mapping $\sigma : T \to \Delta(M)$ and it is convenient to represent a receiver strategy as a mapping $\rho : 2^M \to A$. With CS preferences this is without loss of generality because the receiver has a unique best reply to any belief and therefore his best response to observing a partition element P, which we denote by $\rho(P)$, is the same regardless of the partition (type) to which the element P belongs. The following example illustrates how distortion of imperative meaning arises in this environment.

Example 2 *Suppose sender and receiver have identical payoffs* $-(t-a)^2$ *from action* a *when the sender's decision type is* t*, the sender's decision types are uniformly distributed on the interval* $[0, 1]$*, there are three messages* m_1*,* m_2*, and* m_3 *and the receiver has two possible partition types, the type* $\{\{m_1\}, \{m_2, m_3\}\}$ *with probability* p *and the type* $\{\{m_1\}, \{m_2\}, \{m_3\}\}$ *with probability* $1 - p$*. Then there is a three-step equilibrium in which the lowest interval* $[0, \theta_1]$ *uses message* m_1 *and the critical type* θ_1 *increases monotonically from* $\frac{1}{3}$ *to* $\frac{1}{2}$ *as* p *increases from* 0 *to* 1*. In this equilibrium, the language constrained type, with partition element* $\{m_2, m_3\}$*, does not understand the meaning of messages* m_2 *and* m_3*. Generally, a non-trivial partition element like this one can either represent a catch-all set of messages the receiver does not understand or, if there are multiple, non-trivial partition elements as a category of terms that he can identify as such but within which he cannot further discriminate, as when someone can associate impressionism and expressionism with art but not distinguish between the two.*

In this equilibrium there is distortion of imperative meaning: The sender would want to change her strategy conditional upon learning the receiver's language type. For example, as p*, the probability of the receiver having a limited ability to discriminate among messages* m_2 *and* m_3 *converges to one, the action* a_2 *that the receiver takes if he receives and identifies* m_2*, converges to* $\frac{5}{8}$*, whereas* a_1*, the action he takes in response to* m_1*, converges to* $\frac{1}{4}$*. Thus,*

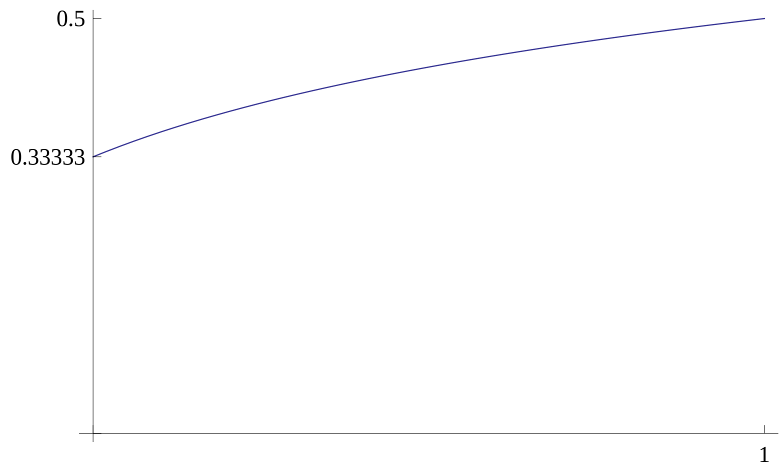

Figure 2: The critical type $\theta_1 = \frac{1}{3}\left(-1 + \frac{4p}{-1+p} - \frac{2\sqrt{1+3p}}{-1+p}\right)$

in the limit the type who would be indifferent between sending messages m_1 and m_2 if she knew the receiver's partition type to be $\{\{m_1\},\{m_2\},\{m_3\}\}$ is $\frac{7}{16}$, while in equilibrium the critical type is $\frac{1}{2}$. Types in the interval $\left(\frac{7}{16},\frac{1}{2}\right)$ would want to switch from their equilibrium message m_1 to sending m_2 if they learned that the receiver can distinguish all messages.

There is also another equilibrium in which m_1 is used on the middle interval $(\frac{1}{3},\frac{2}{3})$. In this equilibrium there is no distortion of imperative meaning. Note, however, that in this equilibrium useful information is transmitted only if the receiver can distinguish all three messages.

For our next result we first formally define distortion of imperative meaning. Then we introduce the notion of a *varied receiver response* that lets us distinguish between the two equilibria in the above example that will allow us to give a sufficient condition for distortion of imperative meaning.

Definition 2 *There is* **distortion of imperative meaning** *in equilibrium (σ,ρ,β) if there exists a set of decision types $\Theta \subset T$ that has positive probability, a message $m \in M$ with $\sigma(m|t) > 0$ for all $t \in \Theta$ and a partition type $\mathcal{P}$ of the receiver that has positive probability such that message m fails to be optimal for decision types in Θ conditional on the receiver's partition type $\mathcal{P}$.*

For the case where the sender's language-competence is privately known we showed that it is sufficient for distortion of indicative meaning to occur that there is variety in the use of messages and in the support of the language type distribution, i.e. when there are language types that differ in just one message, who use all their messages and all of their messages induce distinct actions. In Definition 3 we introduce an analogous condition that requires the existence of multiple receiver types each of which responds differently to each of its partition elements and that suffices for distortion of imperative meaning when the receiver's language competence is the issue.

Definition 3 *There is a* **varied receiver response** *in equilibrium $\mathcal{E} = (\sigma, \rho, \beta)$ if there is a pair of partition types $\mathcal{P}^* \neq \tilde{\mathcal{P}}$ of the receiver with a common element P_0 such that $\pi_R(\tilde{\mathcal{P}}) \neq 0$, $\pi_R(\mathcal{P}^*) \neq 0$ and for every $P \in \mathcal{P}^* \cup \tilde{\mathcal{P}}$ the set $\{t \in T | U^S(\rho(P), t) > U^S(\rho(P'), t), \forall P' \neq P, P' \in \mathcal{P}^* \cup \tilde{\mathcal{P}}\}$ has positive probability.*

With a varied receiver response it becomes important for the sender to know exactly what the partition type of the sender is. The reason is that it guarantees that there will be at least one pair of receiver types for which a positive probability set of sender types would want to induce the action associated with a common partition element for one receiver type and another action for the other receiver type.

Proposition 4 *There will be distortion of imperative meaning in any equilibrium $\mathcal{E} = (\sigma, \rho, \beta)$ with a varied receiver response.*

Proof: Call two elements P_i and P_j of the set $\mathcal{P}^* \cup \tilde{\mathcal{P}}$ adjacent for equilibrium $\mathcal{E}$ if $\rho(P_i) \neq \rho(P_j)$ and there does not exist $P_k \in \mathcal{P}^* \cup \tilde{\mathcal{P}}$ with $\rho(P_k) \in (\rho(P_i), \rho(P_j))$. Since $\mathcal{P}^*$ and $\tilde{\mathcal{P}}$ have a common element and because $\mathcal{P}^* \neq \tilde{\mathcal{P}}$, there is (at least) one common element, P_C, that is adjacent to a non-common element, P_{NC}. With CS preferences, the sender's single-crossing condition implies that there is a unique type who is indifferent between the actions $\rho(P_C)$ and $\rho(P_{NC})$. Without loss of generality, let $\rho(P_C) < \rho(P_{NC})$ and $\rho(P_{NC}) \in \tilde{\mathcal{P}}$. Define $P_+ := \arg\min\{\rho(P) | P \in \mathcal{P}^* \text{ and } \rho(P) > \rho(P_C)\}$ if there exists $P \in \mathcal{P}^*$ with $\rho(P) > \rho(P_C)$ and define $P_+ := P_C$ otherwise. Suppose that $P_+ = P_C$. Since P_C is common to both partitions, we have $P_C \cap P_{NC} = \emptyset$. From the sender's single-crossing condition, it follows that those types who would want to induce $\rho(P_{NC})$ when learning $\tilde{\mathcal{P}}$, would want to induce $\rho(P_C)$ when learning $\mathcal{P}^*$. Since $P_C \cap P_{NC} = \emptyset$, they would want to send different message in both cases. Thus in one of the cases the message they would want to send differs from their equilibrium message, which establishes our claim. Now consider the case where $P_+ \neq P_C$. Since P_C and P_{NC} are adjacent, it must be the case that $\rho(P_+) > \rho(P_{NC})$. Since $\rho(P_C) < \rho(P_{NC}) < \rho(P_+)$, the sender's single crossing condition implies that there is a positive probability set of types (near the type who is indifferent between $\rho(P_C)$ and $\rho(P_+)$,) who would want to induce $\rho(P_{NC})$ when learning $\tilde{\mathcal{P}}$ and would want to induce $\rho(P_C)$ when learning $\mathcal{P}^*$. Thus, as before in one of these two cases the message these types would want to send differs from their equilibrium message, which establishes our claim. □

At this point one might be tempted to proceed as in the case where the language competence of the sender is the issue and to try to show that with common interests all messages will be used and that this in turn leads to having the varied-receiver-response condition satisfied in optimal equilibria. The following example, however, demonstrates that there is an interesting asymmetry in the effects of making the sender's language competence private information versus doing the same for the receiver. It shows that in the latter case optimality sometimes requires that there are messages that will never be used.

Example 3 *Suppose the sender's type is drawn from a uniform distribution on* $[0, 1]$ *and both players receive identical payoffs* $-(t-a)^2$ *when the receiver takes action* a *in state* t. *Let* $M = \{m_1, m_2, m_3, m_4\}$. *For any* $\epsilon \in [0, 1)$, *define a game* Γ^ϵ *by the property that each of the receiver types* $\{\{m_1, m_4\}, \{m_2\}, \{m_3\}\}$, $\{\{m_1\}, \{m_2, m_4\}, \{m_3\}\}$ *and* $\{\{m_1\}, \{m_2\}, \{m_3, m_4\}\}$ *has probability* $\frac{1-\epsilon}{3}$ *and the remaining receiver types are equally likely. Note that if* $\epsilon \in (0, 1)$, *the partition-type distribution* π_R *has full support.*

If $\epsilon = 0$, *then in any optimal equilibrium, the type space is partitioned into three equal-length intervals and the actions that are induced in equilibrium are* $\frac{1}{6}$, $\frac{1}{2}$ *and* $\frac{5}{6}$. *To see this, observe first that this holds if for the moment we make the receiver type common knowledge. This provides an upper bound. Then note that the same outcome that is optimal when the receiver type is common knowledge can be realized when the receiver type is private information. Denote the corresponding* ex ante *payoff by* $v^0_{\max}$.

With positive small ϵ, *the messages* m_1, m_2 *and* m_3 *must approximately induce the same set of actions in an optimal equilibrium as they do in an optimal equilibrium for* $\epsilon = 0$. *Otherwise, the* ex ante *payoff from optimal equilibria,* $v^\epsilon_{\max}$, *would remain bounded away from* v^0_{max}, *and we know that cannot be the case because the strategy profile that results in* v^0_{max} *when* $\epsilon = 0$ *yields approximately* v^0_{max} *when* $\epsilon > 0$ *and since we have a common-interest game the optimal equilibrium strategy must do even better.*

For any ϵ, *let* $\mathcal{E}(\epsilon)$ *be an optimal equilibrium for the game* Γ^ϵ. *We will argue that for sufficiently small* $\epsilon > 0$ *no type* $t \in [0, 1]$ *of the sender sends message* m_4 *in the equilibrium* $\mathcal{E}(\epsilon)$. *For any* $\delta > 0$, *there exists* $\epsilon(\delta) > 0$ *such that for all* $\epsilon \in (0, \epsilon(\delta))$ *type* t*'s payoff from sending message* m_4 *is bounded from above by*

$$\overline{v}^\epsilon(t) = \left(\frac{1-\epsilon}{3}\right)\left(-\left(t-\frac{1}{6}\right)^2 - \left(t-\frac{1}{2}\right)^2 - \left(t-\frac{5}{6}\right)^2\right) + \epsilon \cdot 0 + \delta$$

while at the same time the payoff to t from sending the optimal message from the set $\{m_1, m_2, m_3\}$ is bounded from below by

$$\underline{v}^{\epsilon}(t) = (1-\epsilon)\left(-\min\left\{\left(t-\frac{1}{6}\right)^2, \left(t-\frac{1}{2}\right)^2, \left(t-\frac{5}{6}\right)^2\right\}\right) - \epsilon \cdot 1 - \delta.$$

For sufficiently small ϵ and δ, we have $\underline{v}^{\epsilon}(t) > \overline{v}^{\epsilon}(t)$ for all $t \in [0,1]$, which shows that there is no type of the sender who would be willing to send message m_4 in any optimal equilibrium of the game Γ^{ϵ} for sufficiently small $\epsilon \in (0,1)$.

The example shows that unlike in the case where only sender competence is the issue, when there is uncertainty about receiver competence, there may be instances when the sender may not want to use all messages in an optimal equilibrium. This will be the case when there are messages for which the probability is high that the receiver does not understand them. Therefore only a few of the receiver's partition types may be relevant. This undermines the varied-response condition from the previous proposition. On the other hand, in an optimal equilibrium of a common interest game, the sender will want to communicate some information. Thus, an optimal equilibrium will not be a pooling equilibrium and for the communicated information to have an impact, there will be receiver messages that induce distinct actions.

For the following result we adopt a slightly different perspective. Denote by $\mathcal{P}^f$ the finest partition of M, i.e. the type of the receiver who understands all messages. We will show that in any optimal equilibrium of a game that is near an optimal equilibrium of the game in which $\mathcal{P}^f$ has probability one but where π_R has full support there is distortion of imperative meaning.

Proposition 5 *With common interests, an optimal equilibrium exists. For any class of games that differ only in the distributions π_R, if there are finitely many optimal equilibria in the game with $\pi_R(\mathcal{P}^f) = 1$ (e.g. if CS's condition M holds), then there exists an $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$ and for every π_R that has full support and satisfies $\pi_R(\mathcal{P}^f) = 1 - \epsilon$, there will be distortion of imperative meaning in any optimal equilibrium.*

Proof: We begin by proving existence. Without loss of generality we can confine attention to receiver strategies for which each action is a best response to some belief. Then, by Assumption 1 each receiver strategy prescribes only actions that are optimal for some type of the sender. Thus receiver strategies can be thought of as associating with each receiver message P the type for whom the action $\rho(P)$ is optimal, i.e. it suffices to think of receiver strategies as elements of T^{2^M}, the set of functions from the powerset of M into the sender's type space. Suppose that for any given strategy ρ of the receiver, the sender uses a best reply; that best reply exists because given the receiver's strategy each sender type maximizes her payoff over a finite set of alternatives, the set of distributions over actions that are induced by each message. Then the resulting payoff for a sender of type t equals

$$\max_{m\in M} \sum_{\mathcal{P}\in\mathbf{P}} \pi_R(\mathcal{P}) \sum_{P\in\mathcal{P}} U\left(\rho(P), t\right) 1_{\{m\in P\}}.$$

Given this behavior of the sender, we can assign the following expected payoff to the receiver's strategy ρ:

$$Q(\rho) = \int_T \max_{m\in M} \left\{ \sum_{\mathcal{P}\in\mathbf{P}} \pi_R(\mathcal{P}) \sum_{P\in\mathcal{P}} U\left(\rho(P), t\right) 1_{\{m\in P\}} \right\} f(t)dt.$$

Since U and the max operator are continuous functions, the integrand is continuous and therefore by the Lebesgue dominated convergence theorem, Q is continuous. Therefore, by Weierstrass's theorem, Q achieves a maximum on the compact set T^{2^M}.

It remains to show that there is distortion of imperative meaning for sufficiently small positive ϵ. If the receiver's language competence is not an issue, which corresponds to $\epsilon = 0$, then any optimal equilibrium partitions T into M nonempty intervals I_m, $m \in M$, with types belonging to the same interval sending the same message and the receiver's optimal actions following any two messages $m \neq m'$ satisfying $a_m \neq a_{m'}$. For sufficiently small positive ϵ any optimal equilibrium $\mathcal{E}^\epsilon$ of a game in which π_R has full support must approximately induce the same set of actions in the event that messages are understood as in one of the optimal equilibria $\mathcal{E}^0$of the game where message are always understood. Without loss of generality, we can name the messages in ascending order of the actions they induce in $\mathcal{E}^0$. Now consider two receiver types, $\mathcal{P}^f$ and $\mathcal{P}^p$ who only differ in that the latter type cannot distinguish messages m_1 and m_2. With ϵ sufficiently small, the sets of type who send messages m_1 and m_2 respectively are approximately the same in $\mathcal{E}^0$ and $\mathcal{E}^\epsilon$ and the receiver responds in $\mathcal{E}^\epsilon$ to $\{m_1\}$, $\{m_1, m_2\}$ and $\{m_2\}$ with actions $a_1 < a_{12} < a_2$. Hence, the varied-response condition is satisfied. The result then follows from Proposition 4. □

4.3 Conflict of Interest

According to Propositions 2 and 4 in equilibria of sender-receiver games with private information about language constraints where a large number of messages are sent and induce distinct actions, there will be distortion of meaning. This observation applies equally to games with and without conflict of interest. However, while such distortions indicate an efficiency loss with common interests, with conflict of interest the additional vagueness afforded by privately known language constraints may be efficiency enhancing.

The potential benefit from privately known language constraints in divergent-interest sender-receiver games closely mirrors that from communication through faulty channels (e.g. Myerson [30], Blume, Board and Kawamura (BBK) [9]), repeated simultaneous message exchange that implements jointly controlled lotteries (e.g. Aumann and Hart [3], Krishna and Morgan [20]) and from communication through strategic biased mediators (Ivanov [18]). Consider the leading example of CS where t is drawn from a uniform distribution on $[0, 1]$ and sender and receiver preferences are given by quadratic loss functions, i.e.

$$\begin{aligned} U^S(a, t, b) &= -(t + b - a)^2, \\ U^R(a, t) &= -(t - a)^2, \end{aligned}$$

with $b > 0$ a positive parameter that measures the sender's bias relative to the receiver. For this environment Goltsman, Hörner, Pavlov and Squintani (GHPS) [16] identified an efficiency bound for communication equilibria, i.e. for equilibria from games in which players communicate through an incentiveless mediator (as considered by Forges [14] and Myerson [29]). BBK showed how to implement the GHPS bound for any value of the bias b in games in which players communicate through a noisy channel that passes the sender's message on with probability $1 - \epsilon$ and otherwise replaces it with a draw from a full support distribution on the message space. The methods of BBK can be used to show (both for the case of sender uncertainty and receiver uncertainty) that *there exists a language-type distribution such that for any bias $b > 0$ there exists an equilibrium that attains the GHPS bound.* The intuition is simple: Take an optimal equilibrium from BBK for the game with communication through a noisy channel and construct an analogous equilibrium for the game with private information about language competence where the noise is simulated through uncertainty about language competence.[13]

5 Higher-order knowledge failures

David Lewis [23] places common knowledge at center stage in his account of language as a convention. Our framework permits us to study the effect of a failure of higher-order knowledge on language use.

In this section we specifically address the three questions of whether we continue to see distortions of meaning when players' knowledge only begins to fail at a level higher than

[13]The details are available in the appendix.

first-order knowledge, whether there is a sense in which these distortions diminish with increasing knowledge order, and, leaving the realm of sender-receiver games, whether there are circumstances where lack of higher-order knowledge of language competence leads to complete communication failure.

5.1 Higher-order knowledge failures in sender-receiver games

To represent players' higher-order knowledge about the sender's language competence, we use an **information structure** $I = \langle \Omega, \lambda, \mathcal{O}^S, \mathcal{O}^R, q \rangle$.

- $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ is a countable state space;
- $\mathcal{L} : \Omega \to 2^M$ specifies the set of messages available to the sender at each state (her *language type*).;
- $\mathcal{O}^S$ is a partition of Ω, the sender's information partition;
- $\mathcal{O}^R$ is the receiver's information partition;
- q is the (common) prior on Ω.

To streamline the notation, let $\mathcal{L}(\omega) = \lambda_\omega$ and let $q(\omega) = q_\omega$. The information partitions describe the knowledge of the players: at state ω, the sender knows that the true state is in $\mathcal{O}^S(\omega)$ but no more (where $\mathcal{O}^S(\omega)$ is the element of $\mathcal{O}^S$ containing ω); and similarly for the receiver. We assume that the sender knows her own language type: if $\omega' \in \mathcal{O}^S(\omega)$, then $\lambda_\omega = \lambda_{\omega'}$.

Information structures encode uncertainty only about the sender's language type, not about the decision-relevant information t (the sender's *decision type*). We assume that the distribution from which t is drawn is independent of q. Given that the sender is fully informed about t, and the receiver knows nothing about t, it would be straightforward to extend the partitions and common prior over the full space of uncertainty, $T \times \Omega$, but to do so would unnecessarily complicate the notation.

In the resulting game a sender strategy is a function $\sigma : T \times \Omega \to \Delta(M)$ that satisfies $\sigma(t, \omega) \in \lambda_\omega$ for all $t \in T$ and all $\omega \in \Omega$ and is measurable with respect to $\mathcal{O}^S$. A receiver strategy is a function $\rho : M \times \Omega \to \mathbb{R}$ that is measurable with respect to $\mathcal{O}^R$. Thus for any strategy strategy pair (σ, ρ), $\rho(m, \omega)$ denotes the receiver's response to the message m at state ω and $\sigma(t, \omega)$ the distribution over messages if the sender's decision type is t at state ω.

At any state ω a sender strategy σ induces a mapping $\sigma_\omega : T \to \Delta(\lambda_\omega)$, where $\sigma_\omega(t) = \sigma(t, \omega)$ for all $t \in T$ and all $\omega \in \Omega$. We will refer to this mapping as the sender's *language at*

ω. Similar, we can define the receiver's language at state ω, $\rho_\omega : M \to \mathbb{R}$, via the property that $\rho_\omega(m) = \rho(m, \omega)$. A language $\hat{\sigma}_\omega$ of the sender is *optimal at* ω if together with a best response $\hat{\rho}_\omega$ by the receiver it maximizes the sender's payoff at ω over all sender languages that are feasible at ω. The receiver's language at ω is optimal if it is a best response to an optimal language of sender at ω.

Note that typically in informative equilibria in this setting both sender and receiver will be uncertain about each other's language. The receiver cannot associate a given message with a fixed set of types and the sender will be uncertain about which action will be induced by a message. Players cannot agree on either sender or receiver meanings of messages, despite the fact that useful information is transmitted in equilibrium.

We begin our investigation of distortion of meaning that arise with higher-order knowledge failure with a simple example.

Example 4 *Consider a sender-receiver game in which both players have identical quadratic payoff functions* $-(a-t)^2$ *and the sender's decision type,* t*, is uniformly distributed on the interval* $[0, 1]$. *Suppose that* $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ *and that the information partitions are given by*

$$\begin{aligned} \textit{Sender}: \quad & \mathcal{O}^S = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4, \omega_5\}\} \\ \textit{Receiver}: \quad & \mathcal{O}^R = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5\}\}. \end{aligned}$$

In addition, assume that at ω_1 *the sender's language type is* $\lambda_{\omega_1} = \{m_1\}$, *(i.e. the sender has only message* m_1 *available), and that at every other state* $\omega \in \Omega$ *the sender's language type is* $\lambda_\omega = \{m_1, m_2\}$, *(i.e. the sender has both messages* m_1 *and* m_2 *available). The common prior,* q*, is uniform on* Ω.

Notice that $\{\omega_2, \omega_3, \omega_4, \omega_5\}$ *is the set of information states at which the sender has all messages available;* $\{\omega_3, \omega_4, \omega_5\}$ *is the set of states at which the receiver knows that the sender has all messages available;* $\{\omega_4, \omega_5\}$ *is the set of states at which the sender knows that the receiver knows that the sender has all messages available; and* $\{\omega_5\}$ *is the set of states in which the receiver knows that the sender knows that the receiver knows that the sender has all messages available. At no state does any player have higher than third-order knowledge of the sender having all messages available and in particular at no state is there common knowledge of this fact.*

We look for an equilibrium (σ, ρ) *in which at element* $\{\omega_i, \omega_{i+1}\}$, *with* $i \in \{2, 4\}$, *of the sender's information partition there is a critical type* θ_i *such that decision types* $t < \theta_i$ *send message* m_1 *and decision types* $t > \theta_i$ *send message* m_2. *Let* a_1^j *denote the receiver's equilibrium response at* ω_j *(*$j \in \{1, 3, 5\}$*) to message* m_1, *and let* a_1^j *denote the response* *to message m_2. In equilibrium θ_i, $i = 2, 4$, and a_k^j, $k = 1, 2$, $j = 1, 3, 5$ must satisfy the conditions:*

$$\begin{aligned}
&a_1^1 = \frac{\frac{1}{2} + \theta_2 \frac{\theta_2}{2}}{1 + \theta_2}; a_2^1 = \frac{1 + \theta_2}{2}; a_1^3 = \frac{\theta_2 \frac{\theta_2}{2} + \theta_4 \frac{\theta_4}{2}}{\theta_2 + \theta_4};\\
&a_2^3 = \frac{(1 - \theta_2)\frac{1+\theta_2}{2} + (1 - \theta_4)\frac{1+\theta_4}{2}}{(1 - \theta_2) + (1 - \theta_4)}; a_1^5 = \frac{\theta_4}{2}; a_2^5 = \frac{1 + \theta_4}{2};\\
&(\theta_2 - a_1^1)^2 + (\theta_2 - a_1^3)^2 = (\theta_2 - a_2^1)^2 + (\theta_2 - a_2^3)^2; \textit{and},\\
&(\theta_4 - a_1^3)^2 + (\theta_4 - a_1^5)^2 = (\theta_4 - a_2^3)^2 + (\theta_4 - a_2^5)^2.
\end{aligned}$$

This system of equations has a unique solution satisfying the constraints that $0 < \theta_2 < 1$ and $0 < \theta_4 < 1$: $\theta_2 = 0.54896, \theta_4 = 0.509768, a_1^1 = 0.420074, a_1^2 = 0.77448, a_1^3 = 0.265045, a_2^3 = 0.764274, a_1^5 = 0.254884$ and $a_2^5 = 0.754884$. Thus, at every state where the sender has a choice of which message to send, each message induces a non-degenerate lottery over receiver actions. Hence, not only is the receiver uncertain about the sender's use of messages, but the sender is also uncertain about the receiver's interpretation of messages. There is no state in which either sender-meaning or receiver-meaning is known by both players.

Importantly, even though at state ω_5 players both have at least second-order knowledge of the fact that the sender has both messages available, they are not making optimal use of the available messages, which would require that $\theta_4 = \frac{1}{2}$. Notice also that at ω_3, where the receiver has only first-order knowledge of the fact that the sender has both messages, there is a larger distortion in the sender's strategy, i.e $\theta_2 - \frac{1}{2} > \theta_4 - \frac{1}{2}$ and therefore players appear to make better use of the available messages with a higher order of knowledge of message availability.

In the example the distortions that arise from failures of higher-order knowledge of language competence diminish with increasing knowledge order. Consider then a variant of Example 4 where we increase the number of states, states remain equally likely and information sets are (primarily) intersecting pairs as in

$$\begin{aligned}
\text{Sender :} \quad & \mathcal{O}^S = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4, \omega_5\}, \ldots\ldots\ldots\ldots\ldots, \{\omega_{2K}, \omega_{2K+1}\}\}\\
\text{Receiver :} \quad & \mathcal{O}^R = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5, \omega_6\}, \ldots, \{\omega_{2K-1}, \omega_{2K}\}, \{\omega_{2K+1}\}\}.
\end{aligned}$$

It is easy to check that at ω_3 regardless of the value of K there must be a substantial distortion in any equilibrium; i.e. it cannot be the case that the sender's and the receiver's language at ω_3 are close to optimal languages. Is it the case though that at states with higher-order knowledge of the sender's language competence this distortion is reduced? The next result examines this question in a general class of common-interest sender-receiver games.

For this purpose consider the CS model with common payoff function U. For $N = 1, 2, \ldots$, let $\mathcal{I}_N$ denote the set of information structures satisfying the following conditions:

- $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$;
- $\lambda_{\omega_1} = \tilde{\lambda}$ and $\lambda_{\omega_n} = \lambda$ for all $n \neq 1$, where λ and $\tilde{\lambda}$ are finite sets of messages with $\tilde{\lambda} \subset \lambda$;
- $\mathcal{O}^S(\omega_1) = \{\omega_1\}$
- q is the uniform distribution on Ω.

Let U_I denote the (*ex ante*) payoff from an optimal equilibrium in the game with information structure I. Define $U_N := \min_{I \in \mathcal{I}_N} U_I$. Let U_λ denote the payoff from an optimal equilibrium given an information structure where λ is common knowledge at every state (i.e. where $\lambda_\omega = \lambda$ for all $\omega \in \Omega$. Then:

Proposition 6 $\lim_{N \to \infty} U_N = U_\lambda$.

In particular this means that in the variant of Example 4 where we increase the number of states, the invariable distortion must be largely limited to lower-order knowledge states.

Proof: For any decision type t use a_t to denote the ideal action for that type. The single crossing condition implies that it is never optimal for the receiver to take an action outside of the set $[a_0, a_1]$. Define $\underline{U} := \min_{t \in [0,1], a \in [a_0, a_1]} U(a, t)$. At every state ω_n with $n \neq 1$ let the sender use a strategy that would be optimal if her language competence were common knowledge. Such an optimum exists by standard compactness arguments. Have the receiver best respond. The payoff from this profile is greater than or equal to $\frac{N-1}{N} U_\lambda + \frac{1}{N} \underline{U}$. Hence the payoff from an optimal profile is greater than or equal to $\frac{N-1}{N} U_\lambda + \frac{1}{N} \underline{U}$. Since we have a common interest game, the optimal profile is an equilibrium profile. Hence the optimal equilibrium profile has a payoff no less than $\frac{N-1}{N} U_\lambda + \frac{1}{N} \underline{U}$. The claim follows. $\square$

It is also instructive to examine the change in behavior from increasing the order of knowledge of language competence of the sender in a given equilibrium of a fixed game. Unfortunately, finding interesting informative equilibria in a game with a large information state space is generally not straightforward. For that reason, in the following example we depart from the usual practice of deriving player's interim beliefs about the probability of information states from a common knowledge distribution over the state space and instead take these beliefs as primitives.

Example 5 *Consider a uniform-quadratic sender-receiver environment. Suppose that* Ω *is the infinite set* $\{\omega_1, \omega_2, \ldots\}$ *and that the information structure is given by*

$$\begin{aligned} \textit{Sender}: &\quad \mathcal{O}^S = \{\{\omega_1\}, \{\omega_2, \omega_3\} \ldots \{\omega_{k-1}, \omega_k\}, \{\omega_{k+1}, \omega_{k+2}\} \ldots\} \\ \textit{Receiver}\ : &\quad \mathcal{O}^R = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\} \ldots\ldots \{\omega_k, \omega_{k+1}\} \ldots\ldots\ldots\}. \end{aligned}$$

In addition, assume that at ω_1 *the sender's language type is* $\lambda_{\omega_1} = \{m_1\}$*, i.e. the sender has only message* m_1 *available, and that at any other state* $\omega \in \Omega$ *the sender's language type is* $\lambda_\omega = \{m_1, m_2\}$*, i.e. the sender has both messages available. Finally, assume that the sender conditional on learning that the state is in* $\{\omega_i, \omega_{i+1}\}$ *assigns interim probability one to state* ω_i *for all* $i \geq 2$ *and similarly for the receiver for all* $i \geq 3$ *and that conditional on* $\{\omega_1, \omega_2\}$ *the receiver assigns probability* $1 - \xi$ *to state* ω_1 *and probability* $\xi \in (0, 1)$ *to state* ω_2*.*[14]

We will consider the class of equilibria (σ, ρ) *in which at every element* $\{\omega_i, \omega_{i+1}\}$ *of the sender's information partition there is a critical type* θ_i *such that decision types* $t < \theta_i$ *send message* m_1 *and decision types* $t > \theta_i$ *send message* m_2*. Denote the receiver's equilibrium response at* $\{\omega_1, \omega_2\}$ *to message* m_1 *by* $a_1^1 = \rho(m_1, \{\omega_1, \omega_2\})$ *and the response to message* m_2 *by* a_2^1*. In equilibrium* θ_2*,* a_1^1 *and* a_2^1 *have to satisfy the following conditions:*

$$\begin{aligned} \theta_2 &= \frac{a_1^1 + a_2^1}{2} \\ a_1^1 &= (1 - \xi)\frac{1}{2} + \xi\frac{\theta_2}{2} \\ a_2^1 &= \frac{1 + \theta_2}{2}, \end{aligned}$$

which implies that $\theta_2 = \frac{2-\xi}{3-\xi}$*, i.e.* $\theta_2 > \frac{1}{2}$ *for all* $\theta \in (0, 1)$ *so that the sender is not using the messages that are available to her optimally at* $\{\omega_2, \omega_3\}$ *If we let* θ_k *denote the critical type at* $\{\omega_{2k}, \omega_{2k+1}\}$ *and* (a_1^k, a_2^k) *the equilibrium actions at* $\{\omega_{2k-1}, \omega_{2k}\}$ *we can iterate from these initial values of* θ_2*,* a_1^1 *and* a_2^1 *according to the rule:*

[14] The reason for adopting this somewhat nonstandard specification of each type's belief regarding the information state is that it helps us avoid a great deal of simultaneity that would otherwise make the derivation of an equilibrium intractable. Later, when we consider general results, we will have beliefs derived from a common knowledge distribution on the state space. For now, note that one can think of these beliefs as the limits as $\eta \to 0$ of posteriors derived from common knowledge distributions that assigns probability $1 - \xi$ to state ω_1 and probability $\xi(1 - \eta)\eta^{i-2}$ to all other states, where $\eta \in (0, 1)$.

$$\begin{aligned} a_1^k &= \frac{\theta_k}{2} \\ a_2^k &= \frac{1+\theta_k}{2} \\ \theta_{k+1} &= \frac{a_1^k + a_2^k}{2} \end{aligned}$$

or

$$\theta_{k+1} = \frac{1}{4} + \frac{1}{2}\theta_k.$$

Notice that although θ_k converges to $\frac{1}{2}$, for any given k we have $\theta_k > \frac{1}{2}$, i.e. for any finite n despite having nth order knowledge of the messages available to the sender, players are not making optimal use of those messages.

In Example 5 while distortions disappear in the limit as the order of knowledge increases, they are pervasive; there is distortion for any finite order of knowledge of the sender's language competence. In the remainder of our discussion of sender receiver games we will show that at least in terms of the information structure this is a general observation.

In both of our examples we focused on equilibria that satisfied a plausible monotonicity property. We will continue to do so and begin by formalizing this condition. Let $\Theta(m_i, \omega_k) := \{t \in T \mid \sigma(t, \omega_k)(m_i) > 0\}$ denote the set of all decision types who send message m_i with strictly positive probability at state ω_k. For any two sets $T_1 \subset T$ and $T_2 \subset T$ that have positive probability we say that $T_1 > T_2$ if $\inf T_1 \geq \sup T_2$.

Definition 4 *An equilibrium is* **order preserving** *if it is interval-partitional and $\Theta(m', \omega_k) > \Theta(m, \omega_k)$ at some state ω_k implies that $\Theta(m', \omega_{k'}) > \Theta(m, \omega_{k'})$ at all states $\omega_{k'}$ at which m' and m are used with positive probability.*

Since our intent is to identify a characteristic of informative order-preserving equilibria, it is useful to know that they always exist. The following result establishes existence of informative order-preserving equilibria for the uniform-quadratic CS model with two messages and for arbitrary information structures. From now on assume that $M = \{m_1, m_2\}$, that there is at least one information state ω with $\lambda_\omega = \{m_1, m_2\}$, and that the distribution of decision types is uniform on $[0, 1]$ and that for every action a and decision type t sender and receiver have identical payoffs $-(a-t)^2$.

Lemma 4 *In the uniform-quadratic CS game with two messages and an arbitrary information structure an informative order-preserving equilibrium exists.*

Lemma 4 establishes not only that communication is possible in this environment (compare Proposition 1), but that there is some degree of common meaning, in the the sense that the sender and the receiver commonly agree on which message means "low" and which means "high."

Proof: Let σ_1 be an arbitrary informative order-preserving strategy of the sender. Assume without loss of generality that at every ω where m_2 is available and sent, $\Theta(m_2, \omega_k) > \Theta(m_1, \omega_k)$. Let $\Omega_1^S, \dots, \Omega_L^S$ be an enumeration of the elements of $\mathcal{O}^S$, and $\Omega_1^R, \dots, \Omega_M^R$ an enumeration of the elements of $\mathcal{O}^R$. Since $\sigma_1(t, \omega)$ is constant across Ω_ℓ^S for each $t \in T$, for $\ell = 1, \dots, L$, we can write $\Theta(m, \Omega_\ell^S)$ and denote the $\sup \Theta(m_1, \Omega_\ell^S) = \inf \Theta(m_2, \Omega_\ell^S)$ by θ_ℓ. For all $\omega \in \Omega_j^R$, let $q_{j\ell}$ denote the receiver's posterior belief that $\omega \in \Omega_\ell^R$. Then the receiver's best reply to message m_1 is

$$a_j^1 = \frac{\sum_{\ell=1}^L q_{j\ell}\theta_\ell \frac{\theta_\ell}{2}}{\sum_{\ell=1}^L q_{j\ell}\theta_\ell}$$

and his best reply to message m_2 is

$$a_j^2 = \frac{\sum_{\ell=1}^L q_{j\ell}(1-\theta_\ell)\frac{1+\theta_\ell}{2}}{\sum_{\ell=1}^L q_{j\ell}(1-\theta_\ell)}$$

as long as the denominators are well defined.

Notice that for all ℓ, $\frac{\theta_\ell}{2} \leq \frac{1}{2}$ and $\frac{1+\theta_\ell}{2} \geq \frac{1}{2}$, and since σ_1 is informative, there is at least one ℓ' for which $\frac{\theta'_\ell}{2} < \frac{1}{2}$ and $\frac{1+\theta'_\ell}{2} > \frac{1}{2}$. Therefore at every Ω_j^R at which the receiver expects to receive both messages with positive probability any best reply by the receiver satisfies $a_j^1 < a_j^2$. For every other $\Omega_{j'}^R$ one of the actions is equal to $\frac{1}{2}$ and we are free to choose the other action so that the $a_j'^1 < a_j'^2$ holds.

Hence there exists be a best reply ρ_1 of the receiver to σ_1 that satisfies the property that $a_j^1 < a_j^2$ all $j = 1, \dots, M$. Call any receiver strategy with this property *order preserving*. Note that the payoff from (σ_1, ρ_1) exceeds the payoff from pooling.

At any element Ω_ℓ^S of her information partition the sender has a posterior belief $\phi_{\ell j}$ that the receiver's information is given by Ω_j^R. Therefore, for a sender with decision type t and information Ω_ℓ^S, the payoff difference between sending message m_2 and m_1 is given by

$$\begin{aligned} &-\sum_{j=1}^J (a_j^2 - t)^2 \phi_{\ell j} + \sum_{j=1}^J (a_j^1 - t)^2 \phi_{\ell j} \\ &= \mathbb{E}[a^2|m_1] - \mathbb{E}[a^2|m_2] + 2t(\mathbb{E}[a|m_2] - \mathbb{E}[a|m_1]), \end{aligned}$$

Since the receiver strategy ρ_1 is order-preserving, it follows that $\mathbb{E}[a|m_2] > \mathbb{E}[a|m_1]$ and therefore the sender's best reply σ_2 to ρ_1 is order-preserving.

Continuing in this manner we can construct a sequence of order-preserving strategy pairs $\{(\sigma_n, \rho_n)\}$. Note that each strategy pair (σ_n, ρ_n) can be viewed as an element of a compact Euclidean space. Therefore the sequence $\{(\sigma_n, \rho_n)\}$ has a convergent subsequence. Reindex, so that now $\{(\sigma_n, \rho_n)\}$ stands for that subsequence. Denote the limit of that subsequence by $(\overline{\sigma}, \overline{\rho})$. Since payoffs are continuous, the limit is an equilibrium. Since the payoff from (σ_1, ρ_1) exceeds the payoff from pooling and since payoffs are nondecreasing along the sequence $\{(\sigma_n, \rho_n)\}$, the payoff from $(\overline{\sigma}, \overline{\rho})$ exceeds the payoff from pooling. Hence the equilibrium $(\overline{\sigma}, \overline{\rho})$ is informative. □

Having established the existence of informative order-preserving equilibria for general information structures, we now show that, regardless of the information structure in informative order-preserving equilibria, distortions of meaning are pervasive.

Proposition 7 *Suppose that at any state ω either $\lambda_\omega = \{m_1\}$ or $\lambda_\omega = \{m_1, m_2\}$. Then for any information structure and for any state ω^* with $\lambda_{\omega^*} = \{m_1, m_2\}$ unless the language type is common knowledge at ω^*, in any order-preserving equilibrium the sender does not use an optimal language at ω^*.*

Proof: Given that attention is restricted to order-preserving equilibria, it is without loss of generality to focus on equilibria in which for every sender with information Ω^S_ℓ there exists $\theta_\ell \in [0, 1]$ such that every decision type $t < \theta_\ell$ sends message m_1 and every decision type $t > \theta_\ell$ sends message m_2.

Since Ω is finite, we can define $\underline{\theta} := \min\{\theta_\ell | \theta_\ell > 0\}$. Note that the set $\{\theta_\ell | \theta_\ell > 0\}$ is nonempty because there is at least one state at which it is not common knowledge that message m_2 is available (hence there must be a state at which only m_1 is available).[15]

At every information state ω_i at which message m_1 is sent with positive probability all decision types $t < \underline{\theta}$ (and possibly others) send message m_1 with probability one. Hence, for every receiver type Ω^R_j who expects to receive both messages with positive probability the response a^1_j to receiving message m_1 satisfies $a^1_j \geq \underline{a}_1 = \frac{\underline{\theta}}{2}$. Since m_1 is always available, for every receiver type Ω^S_j who expects to receive both messages with positive probability a message m_2 indicates that the sender's decision type is in a set of the form $(\theta_j, 1]$ with

[15]The reason for restricting attention to this set is that for example there may be an isolated information state at which both messages are available and only message m_2 is used with positive probability.

$\theta_j \geq \underline{\theta}$. Hence, for every receiver type Ω^R_j who expects to receive both messages with positive probability a^2_j satisfies $a^2_j \geq \underline{a}_2 = \frac{1+\underline{\theta}}{2}$.

For every information type Ω^S_ℓ of the sender who sends message m_1 with positive probability, $\theta_\ell > 0$. Thus, $\underline{\theta}$ either equals one, or is realized at an information state of the sender where she sends both messages with positive probability. Assume that Ω^S_ℓ is such an information state, i.e. the sender of type $(\underline{\theta}, \Omega^S_\ell)$ is indifferent between the lottery over actions induced by message m_1, with payoff $-\sum_{j=1}^J (a^1_j - \underline{\theta})^2 \phi_{\ell j}$, and the lottery over actions induced by m_2, with payoff $-\sum_{j=1}^J (a^2_j - \underline{\theta})^2 \phi_{\ell j}$. Note that for any j with $\phi_{\ell j} > 0$ it is the case that $a^2_j > a^1_j$ and $a^2_j > \underline{\theta}$. Consider two cases: If $a^1_j \geq \underline{\theta}$, then $-(a^1_j - \underline{\theta})^2 > -(a^2_j - \underline{\theta})^2$. If $a^1_j < \underline{\theta}$ and $\underline{\theta} < \frac{1}{2}$, then $-(a^1_j - \underline{\theta})^2 \geq -(\frac{\underline{\theta}}{2} - \underline{\theta})^2 > -(\frac{1+\underline{\theta}}{2} - \underline{\theta})^2 \geq -(a^2_j - \underline{\theta})^2$. Therefore, if we had $\underline{\theta} < \frac{1}{2}$, type $\underline{\theta}$ would strictly prefer to send message m_2, which contradicts our assumption that type $(\underline{\theta}, \Omega^S_\ell)$ is indifferent. Therefore, we conclude that $\underline{\theta} \geq \frac{1}{2}$.

Suppose there is an information state ω_i with $\theta_i = \frac{1}{2}$ (i.e. the sender is using an optimal language at ω_i) where it is not common knowledge that $\lambda_{\omega_i} = \{m_1, m_2\}$. Then from above $\theta_i = \underline{\theta} = \frac{1}{2}$. Observe that in order for $\theta_i = \underline{\theta} = \frac{1}{2}$ the receiver's response $a^1_{j'}$ at $\Omega^R(\omega_i)$ to m_1 must be $\frac{\underline{\theta}}{2}$ and the response $a^2_{j'}$ to m_2 must be $\frac{1+\underline{\theta}}{2}$. Otherwise, since for all j, $a^1_j \in [\frac{\underline{\theta}}{2}, \frac{1}{2}]$ and $a^2_j \in [\frac{1+\underline{\theta}}{2}, 1]$ we would have $-(a^1_j - \underline{\theta})^2 \geq -(\frac{\underline{\theta}}{2} - \underline{\theta})^2 = -(\frac{1+\underline{\theta}}{2} - \underline{\theta})^2 \geq -(a^2_j - \underline{\theta})^2$ for all j and at least one of the two inequalities strict for j' and therefore sender type $(\Omega^S(\omega_i), \underline{\theta})$ would strictly prefer to send message m_1.

Call an information state ω_j **adjacent** to ω_i if there exists $\omega_l \in \Omega^R(\omega_i)$ such that $\omega_j \in \Omega^S(\omega_l.)$ At every state ω_j that is adjacent to ω_i, it must be the case that $\theta_j = \frac{1}{2}$. Otherwise the receiver with type $\Omega^R(\omega_i)$ will take actions $a^1_{j'} > \frac{\underline{\theta}}{2}$ and $a^1_{j'} > \frac{1+\underline{\theta}}{2}$, which would be inconsistent with $\theta_i = \underline{\theta} = \frac{1}{2}$. If it is not common knowledge at ω_i that $\lambda_{\omega_i} = \{m_1, m_2\}$, then there exists a chain of states $(\omega_1, \ldots, \omega_i)$ with the property that any two consecutive elements in the chain are adjacent, $\lambda_{\omega_l} = \{m_1, m_2\}$ for all $l \neq 1$ and $\lambda_{\omega_1} = \{m_1\}$. By induction, at every information state in the chain we must have decision types $t > \frac{1}{2}$ sending message m_1 and decision types $t < \frac{1}{2}$ sending message 2. But this contradicts $\lambda_{\omega_1} = \{m_1\}$. □

The following example shows that the restriction to order-preserving equilibria is needed for Proposition 7.

Example 6 *Consider the information structure with partitions*

$$\textit{Sender:} \quad \mathcal{O}^S = \{\{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4, \omega_5\}\}$$
$$\textit{Receiver:} \quad \mathcal{O}^R = \{\{\omega_1\}, \{\omega_2, \omega_3, \omega_4\}, \{\omega_5\}\}$$

Assume that all states are equally likely and that the set of available messages is $\{m_1\}$ *at*

ω_3 and $\{m_1, m_2\}$ otherwise. One easily checks that the following strategy pair, (σ, ρ), is an equilibrium: At $\{\omega_1, \omega_2\}$ the sender sends m_1 for $t \in \left[0, \frac{1}{2}\right)$ and m_2 otherwise. At $\{\omega_4, \omega_5\}$ the sender sends m_2 for $t \in \left[0, \frac{1}{2}\right)$ and m_1 otherwise. At both $\{\omega_1\}$ and $\{\omega_5\}$ the receiver knows that the sender is using an optimal language, despite the fact that the set of available messages is not common knowledge.

It is worth noting that there is a better equilibrium in which the sender never uses an optimal language when she has two messages available. To see this, modify the above strategy profile so that at $\{\omega_4, \omega_5\}$ the sender sends m_1 for $t \in \left[0, \frac{1}{2}\right)$ and m_2 otherwise and the receiver uses a best reply at $\{\omega_5\}$. The resulting strategy profile, $(\tilde{\sigma}, \tilde{\rho})$ has a strictly higher ex ante *payoff than (σ, ρ). Therefore, an optimal strategy profile for this game also must have a higher payoff than (σ, ρ) and since this is a common-interest game any optimal equilibrium must have a higher payoff than (σ, ρ).*

It is an open question whether order-preservation is a necessary condition for optimality in the general case.

5.2 Communication collapse with higher-order knowledge failures

Our analysis thus far has shown that in sender-receiver games lack of common knowledge of the sender's language competence leads to a pervasive distortion of meaning, while increasing knowledge order tends to be associated with a diminishing distortion. We finish this section with showing that once we allow the informed player also to take actions following the communication stage, lack of common knowledge of language competence can entail complete communication breakdown regardless of finite knowledge order, in situations where communication could be put to good use with common knowledge of language competence.

The following example constructs such a scenario by building on insights of Rubinstein [32], Baliga and Morris [5] and Aumann [2].

Example 7 *Two players play a two-stage game with one-sided private information represented by two equally likely payoff states t_1 and t_2 (so the decision type space for the sender is $T = \{t_1, t_2\}$). In the communication stage the privately-informed sender sends a message to the receiver. In the action stage both players simultaneously take their actions which determine payoffs according to the tables in Figure 3.*

It is easily verified that if it is common knowledge that the sender has two messages, m_α and m_β, available, then there is an equilibrium in which the sender sends message m_α in payoff state t_1, message m_β in payoff state t_2 and each player i takes action α_i if and only if message m_α has been sent.

t_1	α_R	β_R
α_S	$3, 3$	$-10, 2$
β_S	$2, -10$	$1, 1$

t_2	α_R	β_R
α_S	$-10, -10$	$-10, -9$
β_S	$-9, -10$	$1, 1$

Figure 3: Payoff States

Suppose instead that it is not common knowledge which messages are available to the sender. Consider an information structure with state space $\Omega = \{\omega_1, \omega_2, \ldots\}$ *and some common prior* q *with the property that* $q_k > q_{k+1}$ *for all* $k = 1, 2 \ldots$. *The players' information partitions are given by*

$$\begin{aligned} \textit{Sender}: \quad & \mathcal{O}^S = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5, \omega_6\}, \{\omega_7, \omega_8\} \ldots\} \\ \textit{Receiver}: \quad & \mathcal{O}^R = \{\{\omega_1, \omega_2, \omega_3\}, \{\omega_4, \omega_5\}, \{\omega_6, \omega_7\}, \{\omega_8, \omega_9\}, \ldots\}. \end{aligned}$$

Finally, assume that $\lambda_{\omega_1} = \{m_\alpha\}$, $\lambda_{\omega_2} = \{m_\beta\}$, *and* $\lambda_{\omega_k} = \{m_\alpha, m_\beta\}$ *for all* $k = 3, 4, \ldots$ *(i.e. the sender has only* m_α *available at* ω_1*, she has only* m_β *available at* ω_2*, and she has both messages available at every other state).*

Then it follows from Proposition 8 below that in any equilibrium of the game only actions β_S *and* β_R *are taken, regardless of payoff state and the information state. In particular for any finite order of knowledge of the fact that both messages are available to the sender, they remain ineffective in equilibrium.*

The following is a sketch of the argument for the game under consideration:

1. *At* ω_1 *the receiver believes it is more likely that the sender's message is uninformative than informative, and hence the relatively safe action* β_R *is uniquely optimal for the receiver.*

2. *At* ω_3 *and* ω_4*, the sender considers* ω_3 *more likely than* ω_4*, and therefore she believes that the receiver takes action* β_R *with at least probability one half; it follows that action* β_S *is uniquely optimal for the sender.*

3. *If at* ω_4 *the sender had a message that would induce the receiver to take action* α_R *with positive probability, she would send such a message in payoff state* t_1 *despite (as we*

showed) taking action β_S herself. This is a consequence of a violation of Aumann's [2] self-signaling condition at state t_1: at state t_1 the sender wants to persuade the receiver to take action α_R regardless of her own intended action.[16]

4. *At ω_4 and ω_5, the receiver considers ω_4 more likely than ω_5; thus (3) implies that, regardless of the message, he believes that the sender uses action β_S with probability greater than one half.*

5. *Given (4), it is uniquely optimal for the receiver to take action β_R at ω_4 and at ω_5.*

6. *Steps (2)-(5) can be turned into an induction argument that shows that for all ω_k the sender uses action β_S regardless of the payoff state and the receiver uses action β_R regardless of the message.*

The following result identifies characteristics of communication games with one-sided private information that lead to communication breakdown for any finite order of knowledge of language competence (it also verifies the details of the example). For this purpose we consider a class of games with two players, a sender (S) and a receiver (R). The sender privately observes her decision type t from a finite set T and sends a message m from a finite set M to the receiver. Each $t \in T$ has strictly positive prior probability $\pi(t)$. The sender's message has to satisfy the constraint that $m \in \lambda$ where $\lambda \subset M$ is her privately known language type. Each player $i = S, R$ has a finite set of actions A_i. Following the communication stage, both players simultaneously take actions $a_S \in A_S$ and $a_R \in A_R$. Given these actions and the sender's decision type, each player i receives a payoff $U_i(a_S, a_R, t)$. As before, the players' knowledge about the sender's language competence is represented by an information structure $I = \langle \Omega, \lambda, \mathcal{O}^S, \mathcal{O}^R, q \rangle$. Call any game of this form a *sender-receiver game with sender actions.*

We are interested in a subclass of such games in which (i) the receiver has a preferred "safe" action that is uniquely optimal if there is sufficient uncertainty about either the sender's action or her payoff type; (ii) the sender has a unique "safe" best reply for sufficiently strong beliefs that the receiver will use his safe action, and (iii) where it is difficult for the sender credibly to communicate an intent to take an action other than her safe best reply.

An action a_R^0 for the receiver is "safe" if is uniquely optimal regardless of the sender's (rational) action rule for any belief that does not assign more than probability $\frac{2\pi(t)}{1+\pi(t)}$ to any

[16] Baliga and Morris [5] demonstrated how failure of the self-signaling condition can render communication ineffective in games with one-sided private information. Morris [27] connects this to Rubinstein's [32] electronic mail game.

type t, i.e.

$$\sum_{t \in T} U_R(\alpha_S(t), a_R^0, t)\mu(t) > \sum_{t \in T} U_R(\alpha_S(t), a_R, t)\mu(t)$$

for all $a_R \neq a_R^0$, for all $\alpha_s : T \to A_S$ that are best responses to some (mixed) receiver action and for all $\mu \in \Delta(T)$ with $\mu(t) < \frac{2\pi(t)}{1+\pi(t)}$ $\forall t$. We say that the game satisfies the **safe-action condition** if the receiver has a safe action.

For a game that satisfies the safe-action condition, we call a sender action a_S^0 "safe" if independent of the payoff type, it is a unique best reply against beliefs that assign at least probability one half to the receiver taking action a_R^0, i.e.

$$U_S(a_S^0, pa_R^0 + (1-p)\alpha_R, t) > U_S(a_S, pa_R^0 + (1-p)\alpha_R, t)$$

$\forall \alpha_R \in \Delta(A_R), \forall a_S \neq a_S^0, \forall p \geq 1/2, \forall t \neq T$. We say that the game satisfies the **sender safe-response condition** if the sender has a safe action.

A game that satisfies the sender-safe response condition satisfies the **receiver safe-response condition** if at every payoff state t, provided the sender uses her safe response a_S^0 with at least probability one half, the receiver's safe action a_R^0 is a unique best reply, i.e.,

$$U_R(pa_S^0 + (1-p)\alpha_S, a_R^0, t) > U_R(pa_S^0 + (1-p)\alpha_S, a_R, t)$$

$\forall \alpha_S \in \Delta(A_S), \forall a_R \neq a_R^0, \forall p \geq 1/2, \forall t \neq T.$

A game that satisfies the sender- and receiver-best response conditions satisfies the **no-self-signaling condition** if in every state t in which a_S^0 is not dominant for the sender, conditional on taking action a_S^0 herself, the sender prefers that the receiver does not take action a_R^0, i.e. for all t such that there exist $a_S \neq a_S^0$ and a_R with $U_S(a_S, a_R, t) \geq U_S(a_S^0, a_R, t)$ it is the case that

$$U_S(a_S^0, a_R^0, t) < U_S(a_S^0, a_R, t) \; \forall a_R \neq a_R^0.$$

Proposition 8 *In any sender-receiver game with sender actions that satisfies the safe-action, sender-safe-response, receiver-safe-response and no-self-signaling conditions, with information partitions*

$$\begin{aligned} \mathcal{O}^S &= \{\{\omega_1\}, \ldots, \{\omega_\nu\}, \{\omega_{\nu+1}, \omega_{\nu+2}\}, \{\omega_{\nu+3}, \omega_{\nu+4}\}, \{\omega_{\nu+5}, \omega_{\nu+6}\}, \ldots\ldots.\} \\ \mathcal{O}^R &= \{\{\omega_1, \ldots, \omega_\nu, \omega_{\nu+1}\}, \{\omega_{\nu+2}, \omega_{\nu+3}\}, \{\omega_{\nu+4}, \omega_{\nu+5}\}, \{\omega_{\nu+6}, \omega_{\nu+7}\}, \ldots\}, \end{aligned}$$

where $\nu = \#(M)$, $\lambda_{\omega_i} = \{m_i\}$ for $i = 1, \ldots, \nu$, $\lambda_{\omega_i} = M$ for $i > \nu$ and $q_i \geq q_{i+1}$, only the safe actions a_S^0 and a_R^0 are taken in equilibrium.

Proof: Given an equilibrium strategy pair (σ, ρ), and somewhat economizing on notation, use $\mathrm{P}(t|m)$ to denote the receiver's posterior probability of type t conditional on having observed message m at information set $\Omega^R(\omega_1)$. For any two events E and F, use $E \wedge F$ to denote the joint event that both E and F occurred. Sightly abusing notation write ω_m for the event that $\lambda_\omega = \{m\}$, i.e. the sender only has message m available. Then

$$\begin{aligned}
\mathrm{P}(t|m) &= \frac{\mathrm{P}(m|t \wedge \omega_{\nu+1})\mathrm{P}(t \wedge \omega_{\nu+1}) + \mathrm{P}(m|t \wedge \omega_m)\mathrm{P}(t \wedge \omega_m)}{\sum_\tau \mathrm{P}(m|\tau \wedge \omega_{\nu+1})\mathrm{P}(\tau \wedge \omega_{\nu+1}) + \sum_\tau \mathrm{P}(m|\tau \wedge \omega_m)\mathrm{P}(\tau \wedge \omega_m)} \\
&= \frac{\mathrm{P}(m|t \wedge \omega_{\nu+1})\pi(t)q_{\nu+1} + \mathrm{P}(m|t \wedge \omega_m)\pi(t)q_m}{\sum_\tau \mathrm{P}(m|\tau \wedge \omega_{\nu+1})\pi(\tau)q_{\nu+1} + \sum_\tau \mathrm{P}(m|\tau \wedge \omega_m)\pi(\tau)q_m} \\
&= \frac{\mathrm{P}(m|t \wedge \omega_{\nu+1})\pi(t)q_{\nu+1} + \pi(t)q_m}{\sum_\tau \mathrm{P}(m|\tau \wedge \omega_{\nu+1})\pi(\tau)q_{\nu+1} + q_m} \\
&\leq \frac{\pi(t)q_{\nu+1} + \pi(t)q_m}{\pi(t)q_{\nu+1} + q_m} \\
&\leq \frac{\pi(t)q_m + \pi(t)q_m}{\pi(t)q_m + q_m} \\
&= \frac{2\pi(t)}{1 + \pi(t)}
\end{aligned}$$

Hence the safe-action condition implies that for all $\omega \in \mathcal{O}^R(\omega_1)$, regardless of the message observed, the receiver's unique optimal reply is the safe action a_0^R.

At $\omega_{\nu+1}$ and $\omega_{\nu+2}$ the sender assigns posterior probability at least $1/2$ to state $\omega_{\nu+1}$. Therefore, and since we just showed that at $\omega_{\nu+1}$ the receiver uses action a_0^R exclusively, by the sender-safe-response condition, at $\omega_{\nu+1}$ and $\omega_{\nu+2}$ the sender will use action a_S^0 regardless of her decision type t.

Suppose there exists a message m' such that following m' at $\omega \in \mathcal{O}^R(\omega_{\nu+2})$ the receiver takes an action other than a_R^0 with positive probability, i.e. $\rho(a_R^0 \mid m', \omega_{\nu+2})) < 1$. Let $T_D \subset T$ denote the sent of decision types for whom a_S^0 is dominant and $T_N = T \setminus T_D$. Let $\tilde{M} \subset M$ be the set of messages that induce receiver actions other than a_R^0 with positive probability at $\omega \in \mathcal{O}^R(\omega_{\nu+2})$. Then the no-self-signaling condition implies that at $\omega \in \mathcal{O}^S(\omega_{\nu+1})$ all types in T_N send messages that induce actions other than a_R^0 with positive probability, i.e.

$$\sum_{m \in \tilde{M}} \mathrm{Prob}(m|T_N \wedge \mathcal{O}^S(\omega_{\nu+1})) = 1.$$

Thus, since

$$\sum_{m \in \tilde{M}} \text{Prob}(m|T_N \wedge \mathcal{O}^S(\omega_{\nu+3})) \leq 1,$$

there exists $\tilde{m} \in \tilde{M}$ such that

$$\text{Prob}(\tilde{m}|T_N \wedge \mathcal{O}^S(\omega_{\nu+1})) \geq \text{Prob}(\tilde{m}|T_N \wedge \mathcal{O}^S(\omega_{\nu+3})).$$

Together with $q_i \geq q_{i+1}$ for all i this implies that

$$\begin{aligned}
\text{Prob}(\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+1})) &= \text{Prob}(\tilde{m}|T_N \wedge \mathcal{O}^S(\omega_{\nu+1}))\text{Prob}(T_N \wedge \mathcal{O}^S(\omega_{\nu+1})) \\
&\geq \text{Prob}(\tilde{m}|T_N \wedge \mathcal{O}^S(\omega_{\nu+3}))\text{Prob}(T_N \wedge \mathcal{O}^S(\omega_{\nu+3})) \\
&= \text{Prob}(\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+3}))
\end{aligned}$$

Therefore, again economizing on notation, if we let $\text{Prob}(a_S^0|\tilde{m})$ denote the receiver's posterior probability of the sender taking action a_S^0 conditional on having observed message $\tilde{m}$ at $\omega \in \mathcal{O}^R(\omega_{\nu+1})$, then

$$\begin{aligned}
\text{Prob}(a_S^0|\tilde{m}) &= \text{Prob}(a_S^0|\tilde{m} \wedge T_D)\frac{\text{Prob}(\tilde{m} \wedge T_D)}{\text{Prob}(\tilde{m})} \\
&+ \text{Prob}(a_S^0|\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+1}))\frac{\text{Prob}(\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+1}))}{\text{Prob}(\tilde{m})} \\
&+ \text{Prob}(a_S^0|\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+3}))\frac{\text{Prob}(\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+3}))}{\text{Prob}(\tilde{m})} \\
&= \text{Prob}(a_S^0|\tilde{m} \wedge T_D)\frac{\text{Prob}(\tilde{m} \wedge T_D)}{\text{Prob}(\tilde{m})} \\
&+ \left(1 - \frac{\text{Prob}(\tilde{m} \wedge T_D)}{\text{Prob}(\tilde{m})}\right) \times \\
&\quad \left\{\text{Prob}(a_S^0|\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+1}))\frac{\text{Prob}(\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+1}))}{\text{Prob}(\tilde{m}) - \text{Prob}(\tilde{m} \wedge T_D)}\right. \\
&\quad \left. + \text{Prob}(a_S^0|\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+3}))\frac{\text{Prob}(\tilde{m} \wedge T_N \wedge \mathcal{O}^S(\omega_{\nu+3}))}{\text{Prob}(\tilde{m}) - \text{Prob}(\tilde{m} \wedge T_D)}\right\} \\
&\geq \frac{1}{2}
\end{aligned}$$

This, however, implies by the receiver-safe-response condition that following message $\tilde{m}$ at $\omega \in \mathcal{O}^R(\omega_{\nu+2})$ the receiver takes action a_R^0 with probability one, contradicting the fact that $\rho(a_R^0|\tilde{m}, \mathcal{O}^R(\omega_{\nu+2})) < 1$.

Suppose that for $k \geq 1$ we have $\rho_R(a_R^0|m, \mathcal{O}^R(\omega_{\nu+2k})) = 1$ for all $m \in M$. Then, using the same logic as above, $\rho_S(a_S^0|\mathcal{O}^S(\omega_{\nu+2k+1})) = 1$ by the sender-safe-response condition, from which we get $\rho_R(a_R^0|m, \mathcal{O}^R(\omega_{\nu+2k+2})) = 1$ for all $m \in M$ by the no-self-signaling and receiver-safe-response conditions. Therefore, by induction, $\rho_R(a_R^0|m, \mathcal{O}^R(\omega_i)) = 1$ for all $m \in M$ and all i and $\rho_S(a_S^0|\mathcal{O}^S(\omega_i)) = 1$ for all i. □

6 Conclusion and Discussion

We have proposed and explored a simple portable framework for expressing the idea that language is imperfectly shared. We have found that common knowledge of language is not necessary for language to be useful but that lack of common knowledge of language distorts message meaning and renders it uncertain. The distortions are preserved when knowledge of language fails at high finite orders, and in a class of games where both senders and receivers move at the action stage may result in complete communication failure for any finite-order knowledge of language competence.

In a very stimulating paper, Lipman [24] has asked "Why is Language Vague?" Perhaps privately known language competence is part of the answer. If we interpret the indicative meaning of a message as the decision-relevant information conveyed by that message, then we showed that in optimal equilibria of common-interest games meaning will generally be confounded by auxiliary information about the speaker's language competence.[17] It is unlikely that standard game theory can do more. After all, it is inherent in the notion of equilibrium that players know each others' strategies, which implies that in a communication game the receiver of a message always precisely knows the rule by which a message is generated. Our interest is in exploring the boundaries of what can be said about imprecise languages with the precise tools of game theory.

One noteworthy feature of our model with privately known language competence of either sender or receiver is that there is a sense in which there can be misunderstandings. It is tempting to speculate about the consequences of the ensuing disagreements regarding the meaning of verbal agreements and contracts. Generally, one would expect that if language is imperfectly shared in our sense, there will be different perceptions of which obligations are entailed by agreements or contracts and it may be advantageous to have a dispute resolution mechanism in place to resolve conflicts once these differences become apparent.

[17] We are grateful to Benny Moldovanu for reminding us that this vagueness is a result of projecting two-dimensional private information into one dimension. The language dimension however is purely a nuisance in common-interest games.

References

[1] Arrow, Kenneth J. [1974], *The Limits of Organization*, Norton, New York NY.

[2] Aumann, Robert J. [1990], "Nash Equilibria are not Self-Enforcing," in *Economic Decision-Making: Games, Econometrics and Optimization* (J. J. Gabszewicz, J.-F. Richard, and L. A. Wolsey, Eds.) Elsevier, Amsterdam.

[3] Aumann, Robert J. and Sergiu Hart [2003], "Long Cheap Talk," *Econometrica* **71**, 1619–1660.

[4] Baldick, Ross Ashley Brown, James Bushnell, Susan Tierney and Terry Winter [2007], "A National Perspective on Allocating the Costs of New Transmission Investment: Practice and Principles," White paper prepared for WIRES (working group for investment in reliable and economic electric systems).

[5] Baliga, Sandeep and Stephen Morris [2002], "Co-ordination, Spillovers, and Cheap Talk," *Journal of Economic Theory* **105**, 450-468.

[6] Chakraborty, Archishman and Rick Harbaugh [2010], "Persuasion by Cheap Talk," *American Economic Review* **100**, 2361-2382.

[7] Crémer, Jacques, Luis Garicano and Andrea Prat [2007], "Language and the Theory of the Firm," *Quarterly Journal of Economics* **122**, 373-407.

[8] Blume, Andreas [2000], "Coordination and Learning with Partial Language," *Journal of Economic Theory* **95**, 1–36.

[9] Blume, Andreas, Oliver J. Board and Kohei Kawamura [2007], "Noisy Talk," *Theoretical Economics* **2**, 395–440.

[10] Crawford, V.P. and H. Haller [1990], "Learning how to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica* **58**, 571–595.

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

[12] De Jaegher, K. [2003]: "A Game-Theoretic Rationale for Vagueness," *Linguistics and Philosophy* **26**, 637–659.

[13] Eskridge, William N., Phillip P. Frickey and Elizabeth Garrett [2006], *Legislation and Statutory Interpretation*, Foundation Press: New York, NY.

[14] Forges, F. [1986], "An Approach to Communication Equilibria," *Econometrica* **54**, 1375–1385.

[15] Galison, Peter [1997], *Image & Logic: A Material Culture of Microphysics*, The University of Chicago Press, Chicago, IL.

[16] Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani [2007], "Mediation, Arbitration and Negotiation," *Journal of Economic Theory*, forthcoming.

[17] Harsanyi, J.C. [1967-8], "Games of Incomplete Information Played by Bayesian Players, I, II, and III," *Management Science* **14**, 159-182, 320-334, 486-502.

[18] Ivanov, M. [2009], "Communication via a Strategic Mediator," *Journal of Economic Theory*, forthcoming.

[19] Jäger, Gerhard, Lars Koch-Metzger and Frank Riedel [2009], "Voronoi Languages," Working paper, University of Tübingen and University Bielefeld.

[20] Krishna, V. and J. Morgan [2004], "The Art of Conversation: Eliciting Information from Experts Through Multi-Stage Communication," *Journal of Economic Theory*, **117**, 147-179.

[21] Li, Ming and Kristóf Madarász [2008], "When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests," *Journal of Economic Theory*, **139**, 47-74.

[22] Levy, Gilat and Ronny Razin [2007], "On the Limits of Communication in Multidimensional Cheap Talk," *Econometrica* **75**, 885–893.

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

[24] Lipman, Barton L. [2009], "Why is Language Vague?," Working paper, Boston University.

[25] Morgan, John and Phillip C. Stocken [2003], "An Analysis of Stock Recommendations," *The Rand Journal of Economics* **34**, 183–203.

[26] Morris, Stephen [2001], "Political Correctness," *Journal of Political Economy* **109**, 231–265.

[27] Morris, Stephen [2002], "Coordination, Comunication and Common Knowledge: A Retrospective on the Electronic-Mail Game," *Oxford Review of Economic Policy* **18**, 433-445.

[28] Myerson, R.B. [1982], "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics* **10**, 67–81.

[29] Myerson, R.B. [1986], "Multistage Games with Communication," *Econometrica* **54**, 323–358.

[30] Myerson, R.B. [1991], *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, MA.

[31] Posner, Richard A. (1987): "Legal Formalism, Legal Realism, and the Interpretation of Statutes and the Constitution," *Case Western Reserve Law Review* **37**, 179–217.

[32] Rubinstein, Ariel [1989], "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *American Economic Review* **79**, 385–391.

[33] Rubinstein, Ariel [1996], "Why Are Certain Properties of Binary Relations Relatively More Common in Natural Language?," *Econometrica* **64**, 343–355.

[34] Zollman, Kevin J.S. [2005], "Talking to Neighbors: The Evolution of Regional Meaning," *Philosophy of Science* **72**, 69–85.

A Sender-Receiver Games with Conflict of Interest

A.1 Uncertainty about Sender Competence

It is well-known that when there is conflict of interest, access to a noisy channel or, more generally, a nonstrategic mediator can improve communication outcomes in sender-receiver games. In this section we show that private information about message availability can substitute for communication through a nonstrategic mediator. Specifically, in the leading example of the CS model, with a uniform payoff-type distribution and quadratic payoff functions, the efficiency gains from mediated communication can be fully replicated through direct communication when there is private information about message availability.

Myerson [30] gives an example in which there is no communicative equilibrium when the communication technology is perfect, but there is one when agents have to rely on sending a carrier pigeon that gets lost with positive probability. Blume, Board and Kawamura [9] (henceforth BBK) consider communication through a noisy channel that lets the sender's message pass through with probability ϵ and otherwise transmits a random draw from a distribution G on the interval $[0, 1]$. They show that with quadratic preferences, i.e.

$$\begin{aligned} U^S(a, t, b) &= -(t + b - a)^2, \\ U^R(a, t) &= -(t - a)^2, \end{aligned}$$

and a uniform type distribution on the interval $[0, 1]$ (the "uniform quadratic model") for almost all values of the sender's bias $b \in (0, \frac{1}{2})$ there exists a value of the error probability ϵ and an equilibrium with higher *ex ante* payoffs than from the most efficient equilibrium in the model without noise. Goltsman, Hörner, Pavlov and Squintani (GHPS) [16], also in the uniform quadratic model, investigate the limits from mediated communication; that is, they permit agents to send messages to a correlation device and to receive instructions from the device. This amounts to finding the payoffs from optimal *communication equilibria*, as defined by Forges [14] and Myerson [29]. Using the revelation principle (Myerson [28]) one can characterize the set of communication equilibria in the CS model as corresponding to a family of conditional distributions on $\mathbb{R}$, $\{p(\cdot|t)\}_{t \in T}$, that satisfies:

$$\begin{aligned} & t = \arg\max_{t' \in T} \left[- \int_{\mathbb{R}} (t + b - a)^2 dp(a|t') \right], \ \forall t \in T \\ & a = \mathrm{E}_t[t|a] \ \forall a \in A. \end{aligned}$$

Goltsman, Hörner, Pavlov and Squintani (GHPS) [16] use this characterization to show that the receiver's *ex ante* payoff in any communication equilibrium of the CS model is bounded above by $-\frac{1}{3}b(1-b)$. Since the *ex ante* payoffs of the receiver, V_R and the sender, V_S, are related through $V_R = V_S + b^2$, this is also the efficiency bound for communication equilibria in the CS model.

BBK provide a mechanism that attains this efficiency bound.[18] For any b there exists a noise level $\epsilon(b)$ and an equilibrium of the corresponding $\epsilon(b)$-noise game, $\Gamma(\epsilon(b))$, that achieves the GHPS bound. We will show that this bound can also be attained with private information about message availability. In that case, all communication between the players is direct and misunderstandings arise exclusively because of receiver uncertainty about the sender's repertoire of messages: When receiving a message, the receiver does not know to what degree the sender was forced to use that message rather than some other message that she would have preferred had it been in her repertoire. Our proof strategy is to show that for any so-called "front-loading equilibrium" of BBK that achieves the efficiency bound there exists an outcome-equivalent equilibrium in the model with private information about message availability.

As background it is useful briefly to recall the key elements of the construction of the front-loading equilibria in BBK. In such an equilibrium the type set, $[0,1]$, is partitioned into a finite number K of intervals Θ_k (with left endpoint θ_{k-1} and right endpoint θ_k) that are indexed from left to right; for any partition element Θ_k with $k>1$ there is a single message m_k that is sent by types in that partition element; and, types in the leftmost partition element, Θ_1, uniformly randomize over all the remaining messages. As a result, when the receiver observes one of the messages m_k he believes with probability one that there was no transmission error, that the sender's type belongs to the interval Θ_k and takes action

$$a_k = \frac{\theta_{k-1}+\theta_k}{2}.$$

When the receiver observes any of other messages, his posterior probability of an error having occurred is

$$\frac{\epsilon}{\epsilon+\theta_1(1-\epsilon)}$$

and he takes action

$$a_1 = \frac{\theta_1(1-\epsilon)\frac{\theta_1}{2}+\epsilon\frac{1}{2}}{\epsilon+\theta_1(1-\epsilon)},$$

which is the average of the actions he would have taken with and without error weighted by the posterior probabilities of error and no error respectively.

Proposition A1 below is proven by translating this BBK front-loading construction into

[18] Ivanov [18] has recently demonstrated how to attain this bound through a strategic mediator.

the present environment through substituting private information for transmission errors. For example, in the BBK equilibrium, when the receiver observes a message that is voluntarily sent by the lowest interval of decision types, he must average over the two possibilities that the message was sent in error and that it was sent intentionally. In the present environment, analogously, we have the receiver be uncertain between the possibility that a type from the lowest interval deliberately sent the message that is always available and the possibility that another decision type sent the message because no other message was available to her.

Proposition A1 *With a uniform type distribution, quadratic preferences and sender bias $b > 0$, there exists a message space M, an availability distribution π on $\Lambda = \{\lambda \in 2^M | m_0 \in \lambda\}$ and an equilibrium in the corresponding game that attains the efficiency bound for communication equilibria.*

Proof: Suppose that the optimal BBK-front-loading equilibrium $\mathcal{E}(b)$ has K steps, $\Theta_1, \ldots, \Theta_K$. Pick any message space that satisfies $\#(M) \geq K$. Let there be an language type $\tilde{\lambda} \subset M$ with $\#(\tilde{\lambda}) \geq K$ and choose an availability distribution π that satisfies the conditions $\pi(\tilde{\lambda}) = 1 - \epsilon(b)$ and $\pi(\lambda) > 0 \Rightarrow \lambda \cap \tilde{\lambda} = \{m_0\}\ \forall \lambda \neq \tilde{\lambda}$. Then we can induce the outcome of the optimal BBK-front-loading equilibrium $\mathcal{E}(b)$ in our environment by prescribing the following sender strategy. Whenever the realized language type is $\tilde{\lambda}$, decision types in the interval Θ_1 pool on the message m_0 (which is always available) and for each interval Θ_k, $k = 2, \ldots, K$, there is a message in $\tilde{\lambda}$ that is sent by decision types in that interval, and only by those decision types. All decision types send the message m_0 whenever their language type λ is not equal to $\tilde{\lambda}$. The receiver chooses his best response given this sender strategy for any of the messages that are sent with positive probability. Following any of the messages that are sent with probability zero by the sender the receiver's posterior is assumed to be the same as following m_0, and he takes the corresponding optimal action. □

A.1.1 A Universal Availability Structure

To establish our last result we chose Λ and π as a function of the sender's bias b. Tying Λ and π to the sender's bias is not necessary, if we allow infinite message spaces. Specifically, it is possible to find Λ and π that are universal in the sense that for any $b > 0$ there is an equilibrium that achieves the efficiency bound for communication equilibria.

For our next result, we will assume M to be infinite. We consider an *availability structure*, which is a 4-tuple $(M; (\Lambda, \mathcal{F}, \pi))$ that consists of a set of potential messages, M, a set of language types $\Lambda \subset 2^M$, a sigma-algebra $\mathcal{F}$ of subsets of Λ, and a probability measure π on

$(\Lambda, \mathcal{F})$. We are interested in a class of availability structures where language types λ can be ordered in such a way that all messages available to a given type are also available to all lower types and the probability distribution μ can be described in terms of this ordering; in this availability structure there is a natural sense of the degree to which the sender knows the language.

Definition A1 *An availability structure* $(M; (\Lambda, \mathcal{F}, \mu))$ *is* **nested** *if for each* $\alpha \in [0,1]$ *there exists an infinite set* $M_\alpha \subset M$ *such that* $M_\alpha \cap M_{\alpha'} = \emptyset$, $\cup_{\alpha \in [0,1]} M_\alpha = M$, $\Lambda = \{\lambda_\alpha \subset M | \lambda_\alpha = \cup_{\alpha' \leq \alpha} M_{\alpha'}\}$, $\mathcal{F} = \{F \subset 2^\Lambda | F = \cup_{\alpha \in B} \lambda_\alpha \text{and} B \in \mathcal{B}\}$ *(where* $\mathcal{B}$ *denotes the set of Borel subsets of the interval* $[0,1]$*) and there exists an atomless distribution* G *on* $[0,1]$ *with density* g *such that* $g(\alpha) > 0$ *for all* $\alpha \in [0,1]$ *and*

$$\mu(\{\lambda_{\alpha'} | \alpha' \leq \alpha\}) = G(\alpha) \; \forall \alpha \in [0,1].$$

Example A1 *Let* M *be the unit square,* $M_\alpha = \{(x,y) \in M | x = \alpha\}$ *and* G *the uniform distribution. Then, given a draw* α *from* G*, the set of available messages is the rectangle* $[0,\alpha] \times [0,1]$*. The probability that the messages in* $M_{\alpha*}$ *are available equals the probability that* $\alpha \geq \alpha^*$*, i.e.* $1 - \alpha^*$*.*

Proposition A2 *With quadratic preferences, uniform type distribution and a nested availability structure there exists an equilibrium that achieves the efficiency bound for communication equilibria.*

Proof: For any $\epsilon \in [0,1]$, define $\alpha(\epsilon)$ as the (unique) solution of the equation $G(\alpha(\epsilon)) = \epsilon$. Thus, the probability that the messages in $M_{\alpha(\epsilon)}$ are not available is ϵ. Define $\epsilon(b)$ as the noise level for the BBK front-loading equilibrium that attains the GHPS efficiency bound when the sender's bias is b. Suppose that the optimal BBK-front-loading equilibrium has K steps, $\Theta_1, \ldots, \Theta_K$. Then we can replicate this outcome in our environment with a nested availability structure by prescribing the following sender strategy. Whenever the messages in $M_{\alpha(\epsilon(b))}$ are available, types in interval Θ_1 pool on one of the messages $m_0 \in M_0$ (which are always available) and for each interval Θ_k, $k = 2, \ldots, K$, there is a message in $M_{\alpha(\epsilon(b))}$ that is sent by types in that interval. All types send a message m_0 when the messages in $M_{\alpha(\epsilon(b))}$ are not available. The receiver chooses his best response given this sender strategy for any of the messages that are sent with positive probability. Following any of the messages that are sent with probability zero by the sender the receiver's posterior is assumed to be the same as following m_0, and he takes the corresponding optimal action. □

A.2 Uncertainty About Receiver Competence

We will conclude by showing that as in the case where the language competence of the sender is private information, private information about the language competence of the receiver can substitute for mediated communication. A particularly simple way of utilizing private information about the receiver's language competence replicates an equilibrium outcome from Krishna and Morgan's (KM) [20] study of multi-stage communication in the CS environment. This is the subject of the following observation.

Observation. *With a uniform type distribution, quadratic preferences and sender bias* $b \in \left(0, \frac{1}{8}\right)$*, there exists a finite message space* M*, an availability distribution* π_R *of the receiver on the set* $\mathbf{P}$ *of partitions of* M *that assigns positive probability to exactly two elements of* $\mathbf{P}$ *and an equilibrium in the corresponding game that attains the efficiency bound for communication equilibria.*

Proof: To verify the observation, first recall that KM showed that for $b \in \left(0, \frac{1}{8}\right)$ there is a class of equilibria that achieve an *ex ante* payoff of $-\frac{1}{3}b(1-b)$ for the receiver and that GHPS's showed that this is the efficiency bound for communication equilibria in this environment. It remains to show how to replicate KM's construction with private information about the receiver's language competence. For this we briefly summarize the key aspects of their construction: Communication proceeds in two stages. In the first stage the sender reveals whether her type t is less than some quantity x, or not. In the second stage, if $t < x$ then a partition equilibrium is played on the interval $[0, x]$; otherwise with probability p (that is generated by a jointly controlled lottery) the sender sends a message to indicate whether $t \in (x, z)$ or $t \in [z, 1]$, and with probability $1 - p$ no further message is sent.

The outcome of any such equilibrium can be induced with private information about the receiver's language as follows: If the partition equilibrium on the interval $[0, x]$ has $K - 2$ steps, let M contain K messages. The receiver's partition type is either the finest partition of M, denoted $\mathcal{P}^f$, or it is the finest partition that contains the element $\{m_{K-1}, m_K\}$, denoted $\mathcal{P}^c$. The receiver's type distribution π_R is given by $\pi_R(\mathcal{P}^f) = p$ and $\pi_R(\mathcal{P}^c) = 1 - p$. Sender types in the kth interval of the partition equilibrium on $[0, x]$ send message m_k, sender types in (x, z) send message m_{K-1}, and sender types in the interval $[z, 1]$ send message m_K. The key observation is that the distinction between the messages m_{K-1} and m_K is activated only if the receiver's type is $\mathcal{P}^f$, which happens with probability p. Otherwise, the receiver uses an action in response to these message that is optimal against prior beliefs concentrated on the interval $[z, 1]$. It is now easy to see that sender and receiver face the exact same incentives as in the KM construction. □

It is also possible, as in the case of private information about language competence of the sender, to translate the BBK construction into an outcome equivalent equilibrium when there is private information about language competence of the receiver. The construction works for all biases $b > 0$ and is universal in the sense that the message space and language type distribution are independent of the bias, but comes at the cost of requiring an infinite message space. Before proving this result, we will show by way of examples how one can use the BBK construction in simple settings.

We begin by constructing the analog to a two-step front-loading equilibrium of BBK. Suppose that $M = \{m_0, m_1, m_2\}$ and that the receiver's partition types are either $\mathcal{P}_1 = \{\{m_1\}, \{M \setminus \{m_1\}\}\}$ or $\mathcal{P}_2 = \{\{m_2\}, \{M \setminus \{m_2\}\}\}$, with equal probability. Types in the low step always send m_0. Types in the high step randomize uniformly over m_1 and m_2. Thus when a $\mathcal{P}_i$ type observes $\{M \setminus \{m_i\}\}$ he does not know whether this is the result of a low-step sender having sent m_0 or a high-step sender's randomization having failed to result in m_i. This scheme is analogous to having an error probability of $\frac{1}{2}$ in BBK. The role of the sender's randomization in implementing the BBK equilibrium in the present framework is to ensure that the receiver responds identically to m_0 regardless of his partition type. We will use the same type of construction when proving Proposition A3 below.

Higher error probabilities can be simulated as follows: Let $M = \{m_0, m_1, m_2, m_3\}$. There are three equally likely receiver types $\mathcal{P}_i = \{\{m_i\}, \{M \setminus \{m_i\}\}\}$ $i = 1, 2, 3$. Types in the low step always send m_0. Types in the high step randomize uniformly over m_1, m_2 and m_3. This scheme is analogous to having an error probability of $\frac{2}{3}$ in BBK. Finally, lower error probabilities can be simulated as follows: Let $M = \{m_0, m_1, m_2, m_3\}$. There are three equally likely receiver types $\mathcal{P}_i = \{\{m_i, m_{i+1}\}, \{M \setminus \{m_i, m_{i+1}\}\}\}$ $i = 1, 2, 3$, where addition is mod3. If the high-step sender randomizes uniformly over m_1, m_2 and m_3, this scheme is analogous to having an error probability of $\frac{1}{3}$ BBK. It should be clear now how all two-step equilibria with rational error probabilities in BBK can be simulated with private information about the receiver's language ability.

The following result extends these ideas to all error probabilities and all BBK front-loading equilibria.

Proposition A3 *With a uniform type distribution and quadratic preferences, there exists a message space M and an availability distribution π_R on the set of partition types of the receiver such that for every $b > 0$ there is an equilibrium in the corresponding game that attains the efficiency bound for communication equilibria.*

Proof: We will first show that there exists a message space M such that for every b there is a π_R and corresponding equilibrium with the desired property.

Suppose that the optimal BBK front-loading equilibrium has K steps, $\Theta_1, \ldots, \Theta_K$ and that the associated error rate is $\epsilon(b)$. Denote the finest partition of a set S by $\mathcal{F}(S)$. Let $\widetilde{M}$ be the unit square, and for any $\alpha \in [0,1]$ define

$$\widetilde{M}_\alpha := \left\{(x,y) \in \widetilde{M} \middle| x \in \left(\left(\alpha - \frac{\epsilon(b)}{2}\right)(\text{mod}1), \left(\alpha + \frac{\epsilon(b)}{2}\right)(\text{mod}1)\right)\right\},$$

$M = \widetilde{M} \cup \{m_0\}$, $M_\alpha = \widetilde{M}_\alpha \cup \{m_0\}$, $\mathcal{P}_\alpha = \mathcal{F}(M \setminus M_\alpha) \cup \{M_\alpha\}$, and let α be drawn from a uniform distribution on $[0,1]$. The realization of α determines the receiver's partition type $\mathcal{P}_\alpha$.

To replicate the outcome from the optimal BBK front-loading equilibrium, consider the following sender strategy: Select $K-1$ distinct values $y_2, \ldots, y_k \in [0,1]$. Before sending a message let the sender randomize uniformly over the interval $[0,1]$ and denote the realization of this randomization by x. Let types in the lowest step send message m_0 and types in step $k > 1$ send message (x, y_k). Sender types in the lowest step, Θ_1, send message m_0 and types in step Θ_k with $k > 1$ send message (x, y_k). Since all α are equally likely, the sender cannot foresee or control which pairs of messages (x, y) the receiver can distinguish from m_0, because they belong to $\widetilde{M} \setminus \widetilde{M}_\alpha$, and which ones he cannot distinguish from m_0, because they belong to $\widetilde{M}_\alpha$. Observe that given this strategy of the sender regardless of the value α, a receiver with partition type $\mathcal{P}_\alpha$ will receive a message $\{M_\alpha\}$ with probability $\epsilon(b)$ when a message other than m_0 is sent. Messages sent by any step Θ_k with $k > 1$ are observed unchanged by the receiver with probability $1 - \epsilon(b)$ and otherwise the receiver cannot distinguish these messages from m_0. Thus, exactly as in the optimal BBK front-loading equilibrium, messages sent by types in the lowest step induce the intended action with probability one, and for any $k > 1$ the message sent by types in step k is correctly identified as coming from that set of types with probability $1 - \epsilon(b)$ and otherwise pooled with the message sent by the step Θ_1. Therefore, for all messages sent and received in the candidate equilibrium, both sender and receiver face the exact same incentives as in the optimal BBK front-loading equilibrium. Finally, assume that the receiver believes that any other (off-equilibrium) message was sent by type $t = 0$. Then no type will want to send this message because in any equilibrium that implements the efficiency bound for communication equilibrium (including the BBK equilibrium) type $t = 0$ induces her ideal action. Therefore, we have an equilibrium that induces the same outcome as the optimal front-loading BBK equilibrium.

Finally, we can make both the messages space M and the receiver's availability distribution π_R independent of b by replicating the above construction for every b, thus adding a dimension to the message space, making it the union of the unit cube and the always available message m_0. □