

Correlation through Language Games*

Andreas Blume

Department of Economics

University of Arizona

Tucson, AZ 85721

August 29, 2025

Abstract

A *language game* is a finite complete-information game preceded by public pre-play communication with constraints on players' ability to produce and understand messages and on their knowledge of each other's constraints. The paper gives conditions under which it is possible to induce correlated equilibria outside the convex hull of the set of Nash equilibria through simple language games with only a few communication rounds. This is especially relevant for two-player games, in which unmediated communication is necessarily public. In the game of Chicken one can induce the entire set of correlated equilibria via a language game.

JEL: C70, C72, D82

Keywords: Correlated equilibrium; language types; pre-play communication; cheap-talk extension

*I am grateful to Inga Deimen, Dino Gerardi, Joel Sobel and conference participants at the TAMU Conference in Honor of Eric Maskin, the Paris Workshop on Games, Decisions, and Language, the Conference on Mechanism and Institution Design, and the Seventh World Congress of the Game Theory Society for their helpful comments on this paper.

To understand a sentence means to understand a language. To understand a language means to be master of a technique. (Wittgenstein, *Philosophical Investigations*, 199)

1 Introduction

A correlated equilibrium of a complete-information game is a joint distribution over action profiles with the property that a player who learns only her own action finds it optimal to take that action. Correlated equilibria were introduced by Aumann [2] [3]. The set of correlated equilibria is a (sometimes strict) superset of the set of Nash equilibria.

One way of obtaining a correlated equilibrium is via a mediator who sends action recommendations to the players prior to the complete-information base game. Alternatively, one can allow players to communicate directly with each other prior to the complete-information base game. In both cases, it matters whether communication is public or private. When private communication is allowed, the set of correlated equilibria is (sometimes strictly) larger than when communication is restricted to be public.

In this paper, I will be concerned with how much correlation can be achieved when player communicate publicly before taking actions in the base game. In that case, it is well known that with commonly known message spaces and faultless communication only the correlated equilibrium outcomes in the convex hull of the set of Nash equilibrium outcomes can be induced. I will show that if instead players may face constraints on which messages are available to them, which messages they can understand, and what they know about each other's constraints, one can induce correlated equilibrium outcomes outside of the convex hull of the set of Nash equilibrium outcomes.

I define a *language game* as a finite complete-information game preceded by (possibly multiple rounds of) public pre-play communication with constraints on players' ability to produce and understand messages and on their knowledge of each other's constraints. The terminology is not entirely acci-

dental. Wittgenstein [30] draws attention to the different ways in which we use language and how language use relates to the actions we take. He refers to an instance of this relation as “language game” (*Philosophical Investigations*, 7) and emphasizes the diversity of different language games. He also emphasizes that understanding of a language is the mastery of a technique. In that spirit, I treat language as a capacity, which may be imperfect, imperfectly shared, and may be private information. I am interested in how variations in this capacity interact with different incentive structures and communication patterns.¹

I refer to a player’s constraints on her ability to process messages in the communication game as her *language type*. When there is uncertainty about players’ language types, I assume that they are drawn independently from a common-knowledge distribution. Therefore, there is no correlation built into players’ private information prior to the communication game: players’ private information alone would not allow them to correlate their behavior.

Much of the paper is concerned with players’ difficulties with understanding messages. If there are two messages that a player cannot differentiate (if, say, she cannot distinguish the meaning of a statement about metaphysics from one about epistemology), I refer to this as a case of *non-understanding*. I will also make (limited) use of the possibility of *misunderstanding*, which occurs if a player does distinguish messages, but these distinctions are faulty. In addition, players may be constrained by limited message availability.

The paper highlights different pathways for achieving correlation via language games. One of these, illustrated with an example in Section 2 and generalized in Section 4, has players simultaneously send messages that are designed to induce a jointly controlled lottery (Aumann, Maschler and Stearns [1]) but sometimes are not understood. Whether or not a received message is understood depends on the receiving player’s language type. Thus this pathway is characterized by potential *non-understanding* of messages.

¹My use of the term “language game” differs from that of Hintikka [19] who formalizes Wittgenstein’s language games toward the end of assigning truth values to logical expressions.

An alternative pathway, explored in Section 5, does not leverage simultaneous message exchange: in any given communication round only one player’s message matters. I refer to equilibrium behavior that features this kind of de facto sequential communication as a *dialogue*. In the specific dialogue used, one player challenges another to match her message. Continuation play depends on the challenged player’s privately known ability to match the challenger’s messages, her language type. For this pathway, it is the potential *inarticulateness* of the challenged player that matters. Finally, when showing how to induce the entire set of correlated equilibria in Chicken in Section 6, I make (limited) use of *misunderstandings*: In the language game that supports the worst symmetric correlated equilibrium, the column player challenges the row player to match her message. The column player may misunderstand the row player’s reply and therefore be uncertain about whether the row player did or did not understand.²

Players in this paper communicate directly and publicly. In every round all players see all messages but may not understand or misunderstand them. There is no mediator who could send private correlated signals to the players. When there are at least four players, Bárány [5] shows how to implement the set of correlated equilibria through a “cheap-talk extension” of the game, in which the underlying game is preceded by multiple rounds of direct communication. In each round communication is private among subsets of players.³ This result cannot be extended to two-player games since there messages are necessary public.

Ben-Porath [7] obtains results for two-player games under the proviso that players can exchange urns in addition to cheap-talk messages. Dodis, Halevi and Rabin [18] and Urbano and Vila [28] show that it is possible to implement the entire set of correlated equilibria in two-player games when

²With misunderstanding there is positive probability that the received message is informative. The receiver of that message may act on that information but may be mistaken ex post. Non-understanding can be seen as the limit of informativeness going to zero.

³For incomplete information games Forges [15] shows that with four or more players every communication equilibrium distribution (Myerson [24], Forges [14]) is a Bayes Nash equilibrium distribution of an appropriate cheap-talk extension.

players are computationally constrained.

The equilibria that Bárány constructs require that players can sometimes verify prior communication and need not be sequentially rational. Gerardi [17] demonstrates that one can attain the entire set of correlated equilibria as sequential equilibria of a cheap-talk extension of the game that does not require players to be able to verify past messages when there are at least five players. Lehrer and Sorin [23], building on Lehrer [21] and [22], demonstrate that regardless of the number of players the entire set of rational correlated equilibrium distributions can be achieved with a mediator that receives one round of private messages from the players and sends public signals as a deterministic function of the players' messages. Forges [16] reviews this literature and provides additional references.

The literature considers various versions of constraints on players' language, including finite message spaces, Crémer, Garicano, and Prat [12] and Jäger, Metzger, and Riedel [20]; symmetry constraints on strategies, Crawford and Haller [11] and Blume [8]; limited sets of relations on partially nameless objects, Rubinstein [25]; and, clarification and comprehension costs, Dewatripont and Tirole [13]. Blume and Board [9] introduce the language type apparatus used here. Their focus is on common-interest games in which players have private payoff-relevant information in addition to their private information about their language types. In the present paper, if there is private information, it concerns only players ability to process messages. Blume and Board consider constraints on sending messages separately from constraints on understanding messages, whereas here I allow language types with both kinds of constraints.

2 Example: Non-understanding in Chicken

Consider pre-play communication in the game of Chicken, shown in the left panel of Figure 1. It is well known that if the two players, Pam and Leo, have no difficulty sending and understanding messages, the maximal symmetric

payoff they can achieve is 3.5. This can be done through a simple jointly controlled lottery: both players randomize uniformly over sending messages 1 and 2 in the communication stage, play (d, ℓ) in the action stage if the sum of their messages is even, and (u, r) otherwise. I will show that Pam and Leo can do better in a *language game*, in which they may be language constrained and are uncertain about their counterpart's constraints.

		Leo			
		ℓ	r		
Pam	u	4,4	2,5	u	$\frac{1}{3}, \frac{1}{3}$
	d	5,2	0,0	d	$\frac{1}{3}, 0$

Figure 1: Chicken

The right panel shows a symmetric correlated equilibrium distribution that achieves an expected payoff equal to $3.\bar{6}$ for each player. In the language game I construct, this correlated-equilibrium distribution can be achieved through a single round of simultaneous direct pre-play communication.

Suppose that Pam is a physician and Leo a lawyer. Pam can use and understand medical terms. Legal terms are not part of her active vocabulary and she may or may not be able to understand them. Leo can use and understand legal terms. Medical terms are not part of his active vocabulary and he may or may not be able to understand them.

The message space $M = \{*, \#, \&, \$\}$ contains all messages either Pam or Leo may be able to send.⁴ Which messages players can send and which

⁴The following construction would go through with any number of additional messages in M . All that would be needed would be to have any player who deviates to sending one of the additional messages be punished by playing that player's least favorite Nash equilibrium in the base game. For the same reason, the constraint that determines which messages players can send could be dispensed with in the example.

messages they can understand is captured by their *language type*. A player's language type is a pair consisting of a subset of the message space that contains the messages they can send and a partition of the message space that indicates which messages they can distinguish. Messages that belong to the same partition element of a player's language type the player cannot distinguish and therefore does not understand. A player's responses to messages they cannot understand have to be identical.

Suppose that in this example Pam can send messages $*$ and $\#$ and only those messages, independent of her language type. Likewise, Leo can send messages $\&$ and $\$$ and only those messages, independent of his language type. Pam's language types differ by whether or not she can understand legal language and Leo's language types differ by whether or not he can understand medical language.

Formally Pam has language type $\lambda_1^{\text{Pam}} = (\{*, \#\}; \{\{*\}, \{\#\}, \{\&\}, \{\$\}\})$ with probability $\frac{2}{3}$ and language type $\lambda_2^{\text{Pam}} = (\{*, \#\}; \{\{*\}, \{\#\}, \{\&, \$\}\})$ with probability $\frac{1}{3}$. Both of Pam's language types can only use/send the messages $*$ and $\#$; this can be thought of as capturing Pam's active (medical) vocabulary. In addition there may or may not be constraints on Pam's ability to understand messages: Her language type λ_1^{Pam} understands all messages, including legal messages. Her other language type λ_2^{Pam} understands only medical messages, and thus has a more limited passive vocabulary. Hence Pam is more likely to understand all message rather than only medical messages. Leo likewise has two language types, $\lambda_1^{\text{Leo}} = (\{\&, \$\}; \{\{*\}, \{\#\}, \{\&\}, \{\$\}\})$ with probability $\frac{2}{3}$ and $\lambda_2^{\text{Leo}} = (\{\&, \$\}; \{\{*, \#\}, \{\&\}, \{\$\}\})$ with probability $\frac{1}{3}$. Leo, regardless of language type, can use legal messages, and only those messages. In addition, he is more likely to understand all messages rather than only legal messages. Pam's and Leo's language types are drawn independently. This and the distribution from which they are drawn is commonly known.

Consider the following strategies in this language game: At the communication stage, both language types of Pam randomize uniformly over the

messages in $\{*, \#\}$. At the response stage, λ_1^{Pam} follows the rule $(*, \&) \mapsto d$, $(\#, \$) \mapsto d$, $(*, \$) \mapsto u$ and $(\#, \&) \mapsto u$ and λ_2^{Pam} takes action u . At the communication stage, both language types of Leo randomize uniformly over the messages in $\{\&, \$\}$. At the response stage, λ_1^{Leo} follows the rule $(*, \&) \mapsto \ell$, $(\#, \$) \mapsto \ell$, $(*, \$) \mapsto r$ and $(\#, \&) \mapsto r$ and λ_2^{Leo} takes action ℓ .

Intuitively, the language types who understand all messages use a jointly controlled lottery to coordinate on playing the two pure-strategy equilibria in the base game with equal probability. Language types who do not understand received messages are unable to evaluate the outcome of the jointly controlled lottery. These types assign equal probability to each of the actions of language types of their counterparts who understand all messages. Therefore, if the probability of players understanding all messages is sufficiently high, it is optimal for language types who do not understand messages to play u if they are Pam and ℓ if they Leo.

Consider the incentive constraints for Pam at the communication stage. Pam is indifferent between sending message $*$ and message $\#$, regardless of language type: Sending $*$ results in $(*, \&)$ and $(*, \$)$ with equal probability and sending $\#$ results in $(\#, \&)$ and $(\#, \$)$ with equal probability. The payoff consequences from $(*, \&)$ are the same as those from $(\#, \$)$ and the payoff consequences from $(*, \$)$ are the same as those from $(\#, \&)$.

To check the incentive constraints for language type λ_1^{Pam} of Pam at the response stage, first note that following message histories $(*, \&)$ and $(\#, \$)$ both language types of Leo take action ℓ . Therefore action d is (uniquely) optimal for λ_1^{Pam} following those histories. Following message histories $(*, \$)$ and $(\#, \&)$, Leo takes action ℓ if and only if his language type is λ_2^{Leo} . Since that language type has probability $\frac{1}{3}$ action u is (uniquely) optimal for λ_1^{Pam} following those message histories.

It remains to examine the incentive constraints for language type λ_2^{Pam} of Pam at the response stage. This language type expects Leo to take action ℓ with probability $\frac{2}{3}$, either because Leo's language type is λ_2^{Leo} , which occurs with probability $\frac{1}{3}$ or because Leo's language type is λ_1^{Leo} (probability $\frac{2}{3}$) and

the message history is $(*, \&)$ or $(\#, \$)$ (probability $\frac{1}{2}$). Hence language type λ_2^{Pam} is indifferent between actions u and d , and thus action u is optimal for her. This shows that Pam is using a best response. By symmetry, so is Leo.

Key elements of the example are that players communicate simultaneously in a single round; they sometimes do not understand messages; and, there is uncertainty about whether they understand messages. Simultaneity allows players to create a jointly controlled lottery. Lack of understanding – incomprehension – leads players to take actions that do not match the Nash equilibria that would otherwise be induced by the jointly controlled lottery. Uncertainty about incomprehension implies that players who do understand messages stick with the Nash equilibrium strategies of the base game that the jointly controlled lottery prescribes.

The remainder of the paper looks at the use of these devices to achieve correlated outcomes outside the convex hull of the set of Nash equilibria more generally, and introduces other devices and combinations thereof, including asynchronous communication and language types who suffer from misunderstanding rather than incomprehension.

3 Language Games

In a *language game* players who are possibly language constrained communicate directly and publicly prior to taking actions. Players may be limited in their ability to send and interpret messages, and these constraints may be private information. Following communication, players play a finite complete-information *base game* $G = \{I, \{A_i\}_{i \in I}, \{U_i\}_{i \in I}\}$, where I is the player set (I will use I to denote both the set and its cardinality), A_i player i 's action set, $A = \prod_{i=1}^I A_i$ the set of action profiles and $U_i : A \rightarrow \mathbb{R}$ player i 's payoff function, with the usual extension to mixed strategies.

Let M denote the universal message space, the set of all messages that are possibly available to some player in the game. Player i 's language constraint is captured through her *language type*, $\lambda_i = (M_i; \phi_i, \zeta_i)$. The first component,

M_i , describes the subset of messages of the universal message space M that player i can send. The function $\phi_i : M_i \rightarrow \Delta(M)$ maps *intended messages* into distributions over *sent messages*. Player i only knows the intended messages, not sent messages; she knows what she is trying to say but not what she says. The function $\zeta_i : M \rightarrow \Delta(M)$ maps *received messages* into distributions over *interpreted messages*. Player i only knows the interpreted messages, not the received messages; she only knows her interpretation of what has been said, not what has been said.

If player i has more than one possible language type λ_i , her language type is her private information. Denote player i 's set of language types by Λ_i , so that $\lambda_i \in \Lambda_i$. A profile of language types $\lambda = (\lambda_1, \dots, \lambda_I) \in \prod_{i=1}^I \Lambda_i$ is a *language state* and $\Lambda = \prod_{i=1}^I \Lambda_i$ denotes the *language state space*. I assume that language types are drawn independently from a common knowledge distribution q over the language state space Λ . Unless otherwise noted, the language state space will be assumed to be finite, so that $q(\lambda) = q_1(\lambda_1) \times \dots \times q_I(\lambda_I)$. The triple $\mathcal{L} = (M, \Lambda, q)$ is a *language structure*.⁵ Language structures include as a special case *degenerate language structures*, \mathcal{D}_M , in which there is certainty that all players can send and understand all messages in the universal message space M .

In most instances I use language types that can be described by the set of available messages, M_i and a partition \mathcal{Q}_i of the universal message space M , with typical element $Q_i \in \mathcal{Q}_i$. By doing so, I am tying my hands: these language type can be viewed as a restriction on a special case of the more general language types described above. The partition \mathcal{Q}_i specifies which messages player i can distinguish. A player can distinguish two messages if and only if they belong to different elements of her partition. Any two messages that belong to the same partition element Q_i , she has to treat identically,

⁵Evidently, the information players receive in a language game could alternatively be generated by a mediator. Therefore, since the set of outcomes that can be realized with a mediator coincides with the set of correlated equilibrium outcomes, the set of outcomes that can be achieved with language games is a subset of the set of correlated equilibrium outcomes.

both when sending messages and when interpreting received messages.

If a language type λ_i of player i is characterized through a partition \mathcal{Q}_i and a set of available messages M_i I represent it by $\lambda_i = (M_i; \mathcal{Q}_i)$. I assume throughout that $M_i \cap Q_i \neq \emptyset \Rightarrow Q_i \subseteq M_i$, i.e., a player can always distinguish messages that are available to her, M_i , from messages that are unavailable to her, $M \setminus M_i$. If all players' language types are of the form $\lambda_i = (M; \mathcal{Q}_i)$, i.e., $M_i = M$, the language structure is a *partitional language structure*.

Language types that can be expressed through partition relate to the more general ones that use intention and interpretation functions as follows: Given a partition \mathcal{Q}_i of the universal message space M , define intention and interpretation functions, ϕ_i and ζ_i by the property

$$\phi_i(m) = \zeta_i(m) = U[Q_i], \forall m \in Q_i, \forall Q_i \in \mathcal{Q}_i,$$

where $U[X]$ denote the uniform distribution over the set X . Then language type $\lambda_i = (M_i; \phi_i, \zeta_i)$ has no reason to treat messages belonging to the same partition element differently. This is slightly weaker than the requirement that language type $\lambda_i = (M_i; \mathcal{Q}_i)$ cannot treat messages belonging to the same partition element differently. This makes no difference to any of the results.

I assume that during a language game language constraints remain constant and unaffected by communication. This implies, for example, that with a partitional language structure players do not refine their partitions as a consequence of prior communication. There is no language learning: players do not become more articulate or more discerning during the course of the game.

When a player sends an intended message the realized sent message is the same for all other players and becomes their received message. Different players may have different interpretations of received messages. In a *language game* $\Gamma_n(G, \mathcal{L})$ the language structure is \mathcal{L} , and there are n rounds during which players send public messages prior to the base game G . A special case is the language game $\Gamma_n(G, \mathcal{D}_M)$, the n -round communication game in which

all players can send and understand all messages in M .

Player i 's interpretation of the (realized) message sent by player $j \neq i$ in period t is denoted by $m_{i,j}^t \in M$. Player i 's intended message in period t is $m_{i,i}^t \in M$. Hence, $\mathbf{m}_i^t = (m_{i,1}^t, \dots, m_{i,i}^t, \dots, m_{i,I}^t)$ is the period t message profile known to player i , $t = 1, 2, \dots$. It consists of player i 's intended message $m_{i,i}^t$ and player i 's interpretations $m_{i,j}^t$, of messages send to i by players $j \neq i$.

Players begin exchanging messages in period 1. At the end of period t , $t = 0, 1, 2, \dots$, player i knows her private history of messages $h_i^t = (\emptyset, \mathbf{m}_i^1, \dots, \mathbf{m}_i^t)$, with $h_i^0 = \emptyset$ denoting the null history. The set of all of player i 's private period- t histories is H_i^t and the set of all of player i 's private histories is $H_i = \bigcup_t H_i^t$.

Player i 's strategy in the language game $\Gamma_0(G, \mathcal{L})$ with zero rounds of communication is an action rule $\rho_i : \Lambda_i \rightarrow \Delta(A_i)$ that maps player i 's language types into distributions over player i 's actions. Player i 's strategy in the language game $\Gamma_n(G, \mathcal{L})$ with $n > 0$ rounds of communication is a pair $\sigma_i = (\gamma_i, \rho_i)$ that consists of a communication rule $\gamma_i : H_i^t \times \Lambda_i \rightarrow \Delta(M_i), t = 0, \dots, n-1$ that maps pairs of private message histories and language types into distributions over available intended messages and an action rule $\rho_i : H_i^n \times \Lambda_i \rightarrow \Delta(A_i)$ that maps pairs of period- n message histories and language types into distributions over player i 's actions in the base game.

I am interested in how the Bayesian Nash equilibria of the language games $\Gamma_n(G, \mathcal{L})$ relate to the correlated equilibria of the base games G . Among the Bayesian Nash equilibria of $\Gamma_n(G, \mathcal{L})$ I distinguish *polite equilibria*, in which in every round only the message of a single player matters, from *impolite equilibria*, in which players pay attention to multiple simultaneously sent messages. To formalize this, say that given a sequence $(i_\tau)_{\tau=1}^n$ of players, two private histories h_i^t and \hat{h}_i^t of player i are $(i_\tau)_{\tau=1}^n$ -equivalent if $m_{i,i_\tau}^\tau = \hat{m}_{i,i_\tau}^\tau$ for all $\tau \leq t$. The sequence $(i_\tau)_{\tau=1}^n$ of players singles out one player for every period of play, and two private histories of player i are $(i_\tau)_{\tau=1}^n$ -equivalent if the interpreted (intended if $i_\tau = i$) messages of the singled out players agree

in every period of play. Players who make no distinctions among $(i_\tau)_{\tau=1}^n$ -equivalent histories only pay attention to the communication of a designated player in each period. An equilibrium $\sigma = (\gamma, \rho)$ of a language game $\Gamma_n(G, \mathcal{L})$ is a *polite equilibrium* if there is a sequence of players $(i_\tau)_{\tau=1}^n$ such that

$$\gamma_i(h_i^t, \lambda_i) = \gamma_i(\hat{h}_i^t, \lambda_i)$$

and

$$\rho_i(h_i^n, \lambda_i) = \rho_i(\hat{h}_i^n, \lambda_i).$$

for every pair of $(i_\tau)_{\tau=1}^n$ -equivalent private histories h_i^t and \hat{h}_i^t , $t = 1, \dots, n$ every $i \in I$ and every $\lambda_i \in \Lambda_i$. The sequence of players used to define a polite equilibrium is a *sequence of effective players* for that equilibrium.

4 Non-understanding

In this section I generalize the observation about non-understanding from the introductory example. Players communicate simultaneously and publicly in a single communication round. The base game has an attribute that generalizes the property of having multiple strict Nash equilibria. In the communication game all but one of the players understand all messages. The remaining player understands all messages with high probability and with small strictly positive probability does not understand any messages.

Intuitively, if the base game has two strict equilibria, then there is an equilibrium in the communication game in which players use a jointly controlled lottery to approximately induce a distribution over the two strict equilibria. The players who understand messages play the actions of the strict equilibrium that is suggested by the outcome of the jointly controlled lottery. The player who sometimes does not understand has a unique best reply that only sometimes matches the action that would be required by that player in the strict equilibrium selected by the jointly controlled lottery. This induces play that is different from any convex combination of equilibria of the base game.

Call $\tilde{A} = \prod_{i=1}^I \tilde{A}_i$ with $\tilde{A}_i \subseteq A_i$ a *best response set* if for all $i \in I$ and every belief $\beta_i \in \Delta(\tilde{A}_{-i})$ player i has a best reply in \tilde{A}_i .⁶ Denote the base game in which each player i is restricted to strategies in $\Delta(\tilde{A}_i)$ by $G_{\tilde{A}}$.

Proposition 1 *For every base game G with a best response set \tilde{A} for which $G_{\tilde{A}}$ has two or more strict Nash equilibria, there is a language game $\Gamma_1(G, \mathcal{L})$ with a finite partitional language structure \mathcal{L} that has a Bayesian Nash equilibrium that induces an outcome of G outside of the convex hull of the set of Nash equilibrium outcomes of G .*

The construction that leads to this result has players send messages simultaneously to generate a jointly controlled lottery. Jointly controlled lotteries were introduced by Aumann, Maschler and Stearns [1]. Aumann and Hart [4] note their use as a way for players to compromise. Aumann and Hart refer to the simultaneous talk that is required for jointly controlled lotteries as “impolite talk.” Proposition 1 shows how correlated outcomes outside the convex hull of Nash outcomes can be achieved with a single round of simultaneous and therefore impolite talk. In the next section I examine how one can obtain correlated outcomes outside the convex hull of Nash outcomes with multiple rounds of polite talk.

Proof: Let s and s' be two strict Nash equilibria of $G_{\tilde{A}}$. There are at least two players whose actions differ in s and s' . Let player i be such a player, so that $s_i \neq s'_i$. Let M be a (large) finite message space of size $|M|$ and $\iota : M \rightarrow \{1, \dots, |M|\}$ a bijection from M to $\{1, \dots, |M|\}$. For $\eta \in (0, 1)$, let $\mathcal{L}(\eta, M)$ be the partitional language structure in which for all players $k \neq i$ their language type is $\lambda_k = (M, \{\{m\}_{m \in M}\})$ with probability 1 and in which

⁶Best response sets are the p -best response sets as defined by Tercieux [27] for $p = 1$. They differ from Voorneveld’s [29] *prep sets* by allowing correlated beliefs. They coincide with prep sets in two-player games. Tercieux’s [26] definition of p -best response sets is slightly different from the one given in Tercieux [27] and for $p = 1$ coincides with Basu and Weibull’s [6] definition of *curb sets*.

player i 's language type is $\lambda_i^1 = (M, \{\{m\}_{m \in M}\})$ with probability $1 - \eta$ and $\lambda_i^2 = (M, \{M\})$ otherwise.

Consider the auxiliary language games $\Gamma_1(G_{\tilde{A}}, \mathcal{L}(\eta, M))$ with $\eta \in (0, 1)$. Fix a player $i' \neq i$. Let $\sigma^M = (\gamma^M, \rho^M)$ in $\Gamma_1(G_{\tilde{A}}, \mathcal{L}(\eta, M))$ be the strategy profile defined by

$$\gamma_j^M(\emptyset, \lambda_j) = U[M], \forall j \in I, \forall \lambda_j \in \Lambda_j,$$

$$\rho_j^M((\emptyset, \mathbf{m}_j^1), \lambda_j) = \begin{cases} s'_j & \text{if } \iota(m_{j,i}^1) + \iota(m_{j,i'}^1) \equiv 1 \pmod{|M|} \\ s_j & \text{otherwise} \end{cases} \quad \forall j \neq i,$$

$$\rho_i^M((\emptyset, \mathbf{m}_i^1), \lambda_i^1) = \begin{cases} s'_i & \text{if } \iota(m_{i,i}^1) + \iota(m_{i,i'}^1) \equiv 1 \pmod{|M|} \\ s_i & \text{otherwise} \end{cases}$$

and

$$\rho_i^M((\emptyset, \mathbf{m}_i^1), \lambda_i^2) = s_i, \forall \mathbf{m}_i^1 \in \prod_{i=1}^I M$$

Note that no player can affect the probability that $\iota(m_{j,i}^1) + \iota(m_{j,i'}^1) \equiv 1 \pmod{|M|}$ (for any player j and any language type) by unilaterally deviating from γ^M . Furthermore, since s and s' are strict Nash equilibria of $G_{\tilde{A}}$, if $\eta > 0$ is sufficiently small, the strategies of all players other than player i are optimal and so is player i 's strategy whenever her language type is λ_i^1 . For every $\epsilon > 0$ there exists $|M|$ such that the probability that $\iota(m_{j,i}^1) + \iota(m_{j,i'}^1) \equiv 1 \pmod{|M|}$ is less than ϵ . Hence, for all $\epsilon > 0$ if we choose $|M|$ large enough, language type λ_i^2 assigns probability at least $1 - \epsilon$ to all other players j taking action s_j . Then, since s_j is a strict equilibrium, for sufficiently small ϵ , the action s_i is uniquely optimal for language type λ_i^2 of player i . This implies that for sufficiently small $\eta > 0$ and sufficiently large $|M|$, the strategy profile σ^M is a Bayesian Nash equilibrium of $\Gamma_1(G_{\tilde{A}}, \mathcal{L}(\eta, M))$. Since \tilde{A} is a best-response set of G , it follows immediately that σ^M is also a Bayesian Nash equilibrium of $\Gamma_1(G, \mathcal{L}(\eta, M))$.

It remains to show that the resulting outcome (distribution over action profiles) is outside of the convex hull of Nash equilibrium outcomes of the base game G . Note that the outcome associated with σ^M puts probability weight only on the three profiles s , s' and (s_i, s'_{-i}) . Suppose that this outcome is in the convex hull of the set of Nash equilibrium outcomes of G . In none of the equilibria that are in the support of this outcome more than one player can be mixing; otherwise additional profiles would be part of the support. In none of the equilibria that are in the support of this outcome only one player can be mixing, since s and s' are strict. Therefore the profile (s_i, s'_{-i}) must be a Nash equilibrium; but this contradicts s' being a strict Nash equilibrium. \square

Proposition 1 gives a sufficient condition for there to be a language game that induces correlated outcomes outside of the the convex hull of the set of Nash equilibrium outcomes. All games with multiple strict equilibria satisfy the required property automatically. In addition, base games that include a best-response set that satisfies the requirement of the proposition naturally arise from taking a game with multiple strict equilibria and adding prior moves such as pre-play communication, or a choice of which game to play, as in the following example.

	ℓ	r		L	R
u	$x, -x$	$-x, x$	U	$9, 9$	$0, 8$
d	$-x, x$	$x, -x$	D	$8, 0$	$7, 7$

Figure 2: Choosing games

Suppose the row player is given the choice between the matching pennies game, with $x > 0$, and the stag hunt game in Figure 2. The resulting game G does not have strict equilibria. For $x \leq 7$ the game has a prep set (which is also a best response set in this two player game) with two strict

equilibria, and thus satisfies the condition in Proposition 1. The condition is not necessary. For $x > 7$ it is not satisfied. Still, the language game $\Gamma_1(G, \mathcal{L})$ used in the proof of Proposition 1 has an equilibrium that induces an outcome of G outside of the convex hull of the set of Nash equilibrium outcomes of G .

The next observation uses Proposition 1 to offer conditions that are sufficient to ensure that for a base game G one can find a language game $\Gamma_1(G, \mathcal{L})$ with Bayesian Nash equilibrium payoffs outside the convex hull of the Nash equilibrium payoffs of G . In addition the observation provides sufficient conditions for being able to improve on the payoffs from the convex hull of Nash equilibria of base game G via a language game $\Gamma_1(G, \mathcal{L})$.

Let $V(G)$ denote the convex hull of the set of Nash equilibrium payoffs of the base game G . For any two strategy profiles s and s' in the game G and any player i let $V(G; s, s', i)$ denote the convex hull of the payoff profiles $U(s)$, $U(s')$ and $U(s_i, s'_{-i})$. Let $V^o(G; s, s', i)$ denote the relative interior of $V(G; s, s', i)$. For any $u, v \in \mathbb{R}^I$, let $u > v$ if and only if $u_\ell \geq v_\ell$ for all $\ell \in I$ and there is a player $i \in I$ with $u_i > v_i$. Define $F(G) := \{v \in V(G) | u > v \Rightarrow u \notin V(G)\}$, the set of efficient payoffs in the convex hull of Nash equilibrium payoffs of G . Let $D(G) := \{u \in \mathbb{R}^I | \exists v \in F(G) \text{ with } u > v\}$ denote the payoffs that dominate an efficient payoff in the convex hull of NE payoffs.

Corollary 1 *Suppose there is a best response set \tilde{A} of the base game G with two strict Nash equilibria s and s' of $G_{\tilde{A}}$ for which $V(G) \cap V^o(G; s, s', i) = \emptyset$. Then there is a language game $\Gamma_1(G, \mathcal{L})$ with a finite partitional language structure \mathcal{L} that has a Bayesian Nash equilibrium that induces an outcome of G with payoff profile outside of $V(G)$. If in addition $V^o(G; s, s', i) \subset D(G)$, then that payoff profile Pareto dominates a payoff profile in $F(G)$.*

Proof: The equilibrium of the language game constructed in the proposition induces an outcome with strictly positive weights on the profiles s , s' , (s_i, s'_{-i}) and only on those profiles. Hence, the payoff corresponding to that outcome

is in $V^o(G; s, s', i)$. □

One easily checks that the conditions in Corollary 1 are satisfied in the Chicken games in Figure 3 whenever $z > x > y > 0$, $2x > z + y$, and $z^2 + y^2 > zx + yx$.⁷ For each of those base games one can find a language game with a single communication round that has a Bayesian Nash equilibrium with an expected payoff profile that Pareto dominates a payoff profile that is efficient in the convex hull of Nash equilibrium payoff profiles of the base game. The intuition is simple: Have players use a jointly controlled lottery to coordinate on a mixture of the two pure-strategy equilibria (D, L) and (U, R) that places weight $1 - \epsilon$ on the profile (U, R) with $\epsilon > 0$ and small. Then, if the row player is language constrained with positive probability, she sometimes plays U when the jointly controlled lottery would prescribe D . Since $x - y > z - x$ the resulting expected payoff profile lies above one of the convex combinations of (z, y) and (y, z) .⁸ Since $z^2 + y^2 > zx + yx$, these convex combinations form the efficient frontier of the convex hull of the set of Nash equilibria of G .

	L	R
U	x, x	y, z
D	z, y	$0, 0$

Figure 3: General Chicken

⁷The third inequality ensures that the mixed strategy equilibrium is inefficient in the set of Nash equilibria of G .

⁸Note that Pareto dominance is with respect to *some* convex combinations of (z, y) and (y, z) , not with respect to the jointly controlled lottery that places weight $1 - \epsilon$ on the profile (U, R) .

5 Dialogues

It is not necessary for players to send messages simultaneously in order to induce correlation in the base game via a language game. In this section I show how to obtain correlated equilibrium outcomes outside the convex hull of the set of Nash equilibrium outcomes through a *dialogue*: Players engage in a dialogue when they play a polite equilibrium of a language game. In any given round, only the messages of a single player matter. The messages of the remaining players are ignored – as if they were silent during that round.

Before introducing the main results of this section, it will be useful to briefly remind ourselves of what dialogues can and cannot accomplish in degenerate language games, i.e., when players can send and understand all messages with certainty. When there are multiple Nash equilibria in the base game among which one of the players is indifferent, it is possible to convexify the set of those equilibria via polite equilibria of a language game with sufficiently many messages: Assign each of the equilibria among which the player is indifferent to a distinct message. Let the player randomize over those messages and have everyone play their part of the equilibrium assigned to the realized message. In contrast, the following observation shows that if the indifference condition fails, dialogue is ineffective in degenerate language games. To this end, let $E(G)$ denote the set Nash equilibria of the base game G .

Observation 1 *If $U_i(e) \neq U_i(e')$ for all $e, e' \in E(G)$ with $e \neq e'$ and all $i \in I$, then the set of polite Nash equilibrium outcomes of $\Gamma_n(G, \mathcal{D}_M)$ is the same as the set of Nash equilibrium outcomes of G .*

Proof: Since communication can always be ignored, it is clear that every Nash equilibrium outcome of G is a polite Nash equilibrium outcome of $\Gamma_n(G, \mathcal{D}_M)$. To establish the reverse inclusion, let σ be a polite Nash equilibrium of $\Gamma_n(G, \mathcal{D}_M)$. Since there is no private information and all communication is public, after every terminal message history h^n that is on the

equilibrium path of σ players must play a Nash equilibrium, $e(h^n)$, of G . Let $\{i_\tau\}_{\tau=1}^n$ be the sequence of effective players in σ . Since the last effective player i_n has a strict preference over equilibria in G , after every message history h^{n-1} that is on the equilibrium path of σ she will assign positive probability only to messages that result in the same equilibrium, $e(h^{n-1})$, of G being played. As a result every message m that player i_{n-1} sends in equilibrium following a history h^{n-2} that is on the equilibrium path of σ results in a single Nash equilibrium $e(h^{n-2}, m)$ of G being played after the last communication round. Since player i_{n-1} has a strict preference over equilibria, after every history h^{n-2} she will assign positive probability only to messages that result in the same equilibrium, $e(h^{n-2})$, of G being played. Iterating on n , it follows that after every history $h^{\tau-1}$ that an effective player i_τ may face on the equilibrium path there is a unique equilibrium $e(h^{\tau-1})$ of G that will be played once communication terminates. As special a case, following the null history, a unique equilibrium $e(\emptyset)$ of G will be played once the communication game terminates. Thus the outcome of the polite equilibrium σ of $\Gamma_n(G, \mathcal{D}_M)$ is the outcome of the Nash equilibrium $e(\emptyset)$ of the base game G . \square

Thus, generically polite equilibria of degenerate language games do not expand the set of equilibrium outcomes. The remainder of this section investigates what can be achieved with polite equilibria of language games that are not degenerate, i.e., with players who may be language constrained and with possible uncertainty about those constraints.

Both of the results in this section require, again, that the base-game contains a best-response set with two (or more) strict equilibria when players actions are restricted to that set. The second result requires additionally that there is a player who is not indifferent between these equilibria. The first result is achieved with partitional language structures and requires two communication rounds. The second result demonstrates that if one allows non-partitional language structures, one communication round suffices.

In the construction for the first result, there is a single player, the “con-

strained player,” who sometimes does not understand messages. With high probability she understands all messages and with small strictly positive probability she does not understand any messages. All other players always understand all messages.

When there are two strict equilibria in the base game, the intuition is as follows. One player other than the constrained player is singled out as the “designated player.” In the first round all players randomize uniformly over all messages (only the randomization of the designated player matters). In the second round the constrained player tries to match the message of the designated player. If she does not succeed, the equilibrium (weakly) less favorable to her is played. If she does succeed and understands all messages, her more favorable equilibrium is played. If she does not understand messages, then with a large message space she expects not being able to match and takes the action expected of her in the less favorable equilibrium. In the event that the constrained player does not understand messages there is a small probability that her message matches the message of the designated player. In that case a non-Nash-equilibrium profile will be played.

Proposition 2 *For every base game G with a best response set \tilde{A} for which $G_{\tilde{A}}$ has two or more strict Nash equilibria, there is a language game $\Gamma_2(G, \mathcal{L})$ with a finite partitional language structure \mathcal{L} that has a polite equilibrium that induces an outcome of G outside of the convex hull of the set of Nash equilibrium outcomes of G .*

Proof: Let s and s' be two strict Nash equilibria of $G_{\tilde{A}}$. There is a player i for whom $s_i \neq s'_i$. Without loss of generality suppose that $U_i(s') \geq U_i(s)$. Let M be a finite message space of size $|M|$. For $\eta \in (0, 1)$, let $\mathcal{L}(\eta, M)$ be the partitional language structure in which for all players $k \neq i$ their language type is $\lambda_k = (M, \{\{m\}_{m \in M}\})$ with probability 1 and in which player i 's language type is $\lambda_i^1 = (M, \{\{m\}_{m \in M}\})$ with probability $1 - \eta$ and $\lambda_i^2 = (M, \{M\})$ otherwise.

Consider the auxiliary language games $\Gamma_2(G_{\bar{A}}, \mathcal{L}(\eta, M))$ with $\eta \in (0, 1)$. Fix a player $i' \neq i$. Let $\sigma^M = (\gamma^M, \rho^M)$ in $\Gamma_2(G_{\bar{A}}, \mathcal{L}(\eta, M))$ be the strategy profile defined by

$$\begin{aligned}\gamma_j^M(\emptyset, \lambda_j) &= U[M], \forall j \in I, \forall \lambda_j \in \Lambda_j, \\ \gamma_j^M((\emptyset, \mathbf{m}_j^1), \lambda_j) &= U[M], \forall j \neq i, \forall \lambda_j \in \Lambda_j, \forall \mathbf{m}_j^1 \in M^I, \\ \gamma_i^M((\emptyset, \mathbf{m}_i^1), \lambda_i^1) &= m_{i,i'}^1, \forall \mathbf{m}_i^1 \in M^I, \\ \gamma_i^M((\emptyset, \mathbf{m}_i^1), \lambda_i^2) &= U[M], \forall \mathbf{m}_i^1 \in M^I, \\ \rho_j^M((\emptyset, \mathbf{m}_j^1, \mathbf{m}_j^2), \lambda_j) &= \begin{cases} s'_j & \text{if } m_{j,i}^2 = m_{j,i'}^1 \\ s_j & \text{otherwise} \end{cases} \quad \forall j \neq i, \forall \mathbf{m}_j^1, \mathbf{m}_j^2 \in M^I \\ \rho_i^M((\emptyset, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i^1) &= \begin{cases} s'_i & \text{if } m_{i,i}^2 = m_{i,i'}^1 \\ s_i & \text{otherwise} \end{cases}, \forall \mathbf{m}_i^1, \mathbf{m}_i^2 \in M^I\end{aligned}$$

and

$$\rho_i^M((\emptyset, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i^2) = s_i, \forall \mathbf{m}_i^1, \mathbf{m}_i^2 \in M^I$$

Since the choices of messages in period 1 have no impact on the probability that $m_{j,i}^2 = m_{j,i'}^1$ for any player j , and thus on actions taken, and since uniform randomization over all messages is feasible for all language types, the first period signaling rule $\gamma_j^M(\emptyset, \lambda_j) = U[M], \forall j \in I, \forall \lambda_j \in \Lambda_j$ is optimal for all players. The choices of messages in period 2 by players other than player i have no impact on the actions chosen. Therefore, and since uniform randomization is compatible with all language types, the second-period signaling rule $\gamma_j^M((\emptyset, \mathbf{m}_j^1), \lambda_j) = U[M], \forall j \neq i, \forall \lambda_j \in \Lambda_j$, is optimal for all players other than player i . Language type λ_i^1 of player i understands all messages and therefore can match player i' 's first period messages $m_{i,i'}^1$ (by choosing $m_{i,i}^2 = m_{i,i'}^1$). Player i 's payoff from matching player i' 's first period message is $U_i(s')$, which by construction is no less than the payoff from not matching, which is $U_i(s)$ and thus $\gamma_i^M((\emptyset, \mathbf{m}_i^1), \lambda_i^1) = m_{i,i'}^1$ is optimal. Language type λ_i^2 cannot distinguish any message and therefore the signaling

rule $\gamma_i^M((\emptyset, \mathbf{m}_i^1), \lambda_i^2) = U[M]$ is optimal.

If $m_{j,i}^2 \neq m_{j,i'}^1$, this is recognized by every player $j \neq i$ since their language types always understand all messages. They infer that player i must have language type λ_i^2 and therefore that all of their opponents, including player i , will play according to s_{-j} . Since s is a Nash equilibrium in $G_{\tilde{A}}$, j 's action s_j is a best reply, rendering $\rho_j^M((\emptyset, \mathbf{m}_j^1, \mathbf{m}_j^2), \lambda_j)$ optimal for all $j \neq i$ in the case that $m_{j,i}^2 \neq m_{j,i'}^1$.

If $m_{j,i}^2 = m_{j,i'}^1$, this is observed by player $j \neq i$ and player j expects all players $k \neq i, j$ to play s'_k . Conditional on $m_{j,i}^2 = m_{j,i'}^1$ the probability that $\lambda_i = \lambda_i^2$ and therefore that player i plays s_i is less than η . This and the fact that s' is a strict Nash equilibrium of $G_{\tilde{A}}$ implies that for sufficiently small $\eta > 0$ $\rho_j^M((\emptyset, \mathbf{m}_j^1, \mathbf{m}_j^2), \lambda_j)$ is also optimal for all $j \neq i$ whenever $m_{j,i}^2 = m_{j,i'}^1$.

If $m_{j,i}^2 \neq m_{j,i'}^1$, all players j other than player i will play s_j and if $m_{j,i}^2 = m_{j,i'}^1$ they will play s'_j . Since both s and s' are Nash equilibria of $G_{\tilde{A}}$ this implies that $\rho_i^M((\emptyset, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i^1)$ is optimal.

Language type λ_i^2 of player i expects that $m_{j,i}^2 = m_{j,i'}^1$ for players $j \neq i$ with probability $\frac{1}{|M|}$ regardless of the message she sent in period 2. Hence the probability that her opponents play s_{-i} is $1 - \frac{1}{|M|}$. Since s is a strict equilibrium of $G_{\tilde{A}}$, this implies that for sufficiently large $|M|$ the action rule $\rho_i^M((\emptyset, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i^2)$ of type λ_i^2 is optimal.

In summary, for sufficiently small $\eta > 0$ and sufficiently large $|M|$ the strategy profile σ^M is a Bayesian Nash equilibrium of the auxiliary language game $\Gamma_2(G_{\tilde{A}}, \mathcal{L}(\eta, M))$. Since \tilde{A} is a best-response set of G , it follows immediately that σ^M is also a Bayesian Nash equilibrium of $\Gamma_2(G, \mathcal{L}(\eta, M))$. The sequence of players (i', i) is a *sequence of effective players* for that equilibrium, and hence the equilibrium is polite.

The outcome associated with σ^M puts positive probability weight on the three profiles s , s' and (s_i, s'_{-i}) and only on those profiles. Suppose that this outcome is in the convex hull of the set of Nash equilibrium outcomes of G . In none of the equilibria that are in the support of this outcome more than one player can be mixing; otherwise additional profiles would be part of the

support. In none of the equilibria that are in the support of this outcome only one player can be mixing, since s and s' are strict. Therefore the profile (s_i, s'_{-i}) must be a Nash equilibrium; but this contradicts s' being a strict Nash equilibrium. \square

The second result regarding dialogues shows how to generate correlated equilibrium outcomes outside the convex hull of Nash outcomes with a polite equilibrium of a game that has only a single communication round. This is achieved with language structures that are non-partitional: There is one special message. One, inarticulate, player with positive probability is unable to send the special message; another, non-discerning, player sometimes does not understand messages; and, all other players are always unconstrained.

With two strict equilibria and one player who is not indifferent between them, the intuition is as follows: Let the inarticulate player be the one who is not indifferent between the two strict equilibria. In equilibrium, the inarticulate player sends the special message if it is available to her. All players who understand that message use actions consistent with the strict equilibrium that the inarticulate player prefers. If the inarticulate player is unable to send the special messages, all players with types who understand all messages take actions consistent with the inarticulate player's less preferred strict equilibrium. When the non-discerning player does not understand messages, she takes the action consistent with the inarticulate player's preferred strict equilibrium. This generates a distribution over three strategy profiles that is not in the convex hull of the set of Nash equilibrium profiles.

Proposition 3 *For every base game G with a best reply set \tilde{A} for which $G_{\tilde{A}}$ has two or more strict Nash equilibria and a player who is not indifferent among these equilibria there is a language game $\Gamma_1(G, \mathcal{L})$ that has a polite equilibrium that induces an outcome of G outside of the convex hull of the set of Nash equilibrium outcomes of G .*

Proof: Let s and s' be two strict Nash equilibria of $G_{\tilde{A}}$ and ℓ a player

for whom $U_\ell(s') > U_\ell(s)$. Let $i \neq \ell$ be a player for whom $s'_i \neq s_i$. Let M be a finite message space with $|M| \geq 2$, $m^* \in M$ and $\mathcal{L}(\epsilon, \eta)$ a language structure in which $\lambda_\ell^1 = (M, \{\{m\}_{m \in M}\})$ with probability $1 - \epsilon$, $\lambda_\ell^2 = (M \setminus \{m^*\}, \{\{m\}_{m \in M}\})$ with probability ϵ , $\lambda_i^1 = (M, \{\{m\}_{m \in M}\})$ with probability $1 - \eta$, $\lambda_i^2 = (M, \{M\})$ with probability η and $\lambda_k = (M, \{\{m\}_{m \in M}\})$ with probability 1 for all $k \neq i, \ell$. That is, player ℓ always understands all messages, but may or may not have message m^* available; player i either distinguishes all messages or none; and, the remaining players face no constraints.

Consider the auxiliary language games $\Gamma_1(G_{\bar{A}}, \mathcal{L}(\epsilon, \eta))$ with $\epsilon, \eta \in (0, 1)$. Let $\sigma = (\gamma, \rho)$ in $\Gamma_1(G_{\bar{A}}, \mathcal{L}(\epsilon, \eta))$ be the strategy profile defined by

$$\begin{aligned} \gamma_\ell(\emptyset, \lambda_\ell^1) &= m^* \\ \gamma_\ell(\emptyset, \lambda_\ell^2) &= U[M \setminus \{m^*\}] \\ \gamma_j(\emptyset, \lambda_j) &= U[M], \forall j \neq \ell, \forall \lambda_j \in \Lambda_j \\ \rho_j((\emptyset, \mathbf{m}^1), \lambda_j) &= \begin{cases} s'_j & \text{if } m_\ell^1 = m^* \\ s_j & \text{otherwise} \end{cases} \quad \forall j \neq i, \forall \lambda_j \in \Lambda_j \\ \rho_i((\emptyset, \mathbf{m}^1), \lambda_i^1) &= \begin{cases} s'_i & \text{if } m_\ell^1 = m^* \\ s_i & \text{otherwise} \end{cases} \\ \rho_i((\emptyset, \mathbf{m}^1), \lambda_i^2) &= s'_i \end{aligned}$$

By sending message m^* at the communication stage, language type λ_ℓ^1 of player ℓ guarantees that s' will be played at the action stage. Sending any other message instead results in action profile s being played with probability $1 - \eta$ at the action stage. Therefore, and since $U_\ell(s') > U_\ell(s)$, for sufficiently small η , $\gamma_\ell(\emptyset, \lambda_\ell^1) = m^*$ is optimal for language type λ_ℓ^1 of player ℓ .

Language type λ_ℓ^2 , being unable to send message m^* , has no ability to influence play at the action stage and thus, being indifferent, finds it optimal to use $\gamma_\ell(\emptyset, \lambda_\ell^2) = U[M \setminus \{m^*\}]$ at the communication stage.

The messages of all players j other than ℓ do not affect play at the ac-

tion stage and therefore these players find it optimal to use $\gamma_j(\emptyset, \lambda_j) = U[M], \forall \lambda_j \in \Lambda_j$ at the communication stage.

At the action stage, conditional on having observed message m^* all players j other than player i expect their opponents to play s'_{-j} and since s' is a strict equilibrium, the action s'_j is a (unique) best reply. Conditional on having observed a message other than m^* players j other than player i expect their opponents to play s_{-j} with probability $1 - \eta$ and since s is a strict equilibrium, for sufficiently small η the action s_j is a best reply. This confirms the optimality of

$$\rho_j((\emptyset, \mathbf{m}^1), \lambda_j) = \begin{cases} s'_j & \text{if } m_\ell^1 = m^* \\ s_j & \text{otherwise} \end{cases} \quad \forall j \neq i, \forall \lambda_j \in \Lambda_j$$

for sufficiently small η .

Language type λ_i^1 of player i , who understands all messages, after observing message m^* expects her opponents to play s'_{-i} and otherwise expects them to play s_{-i} . Since both s' and s are (strict) Nash equilibria of the auxiliary game, the optimality of

$$\rho_i((\emptyset, \mathbf{m}^1), \lambda_i^1) = \begin{cases} s'_i & \text{if } m_\ell^1 = m^* \\ s_i & \text{otherwise} \end{cases}$$

follows.

Language type λ_i^2 cannot differentiate messages but knows that message m^* is sent with probability $1 - \epsilon$. Since following m^* players j other than i will play s'_{-j} and since s' is a strict equilibrium, for sufficiently small ϵ the action rule $\rho_i((\emptyset, \mathbf{m}^1), \lambda_i^2) = s'_i$ is (uniquely) optimal.

In summary, for sufficiently small $\epsilon > 0$ and $\eta > 0$ the strategy profile σ is a Bayesian Nash equilibrium of the auxiliary language game $\Gamma_1(G_{\tilde{A}}, \mathcal{L}(\epsilon, \eta))$.

Since \tilde{A} is a best-reply set of G , it follows immediately that σ is also a Bayesian Nash equilibrium of the language game $\Gamma_1(G, \mathcal{L}(\epsilon, \eta))$.

Considering the (trivial) one element sequence of players (i_ℓ) , the equi-

librium is polite.

The outcome associated with σ puts positive probability weight on the three profiles s , s' and (s'_i, s_{-i}) and only on those profiles. Suppose that this outcome is in the convex hull of the set of Nash equilibrium outcomes of G . In none of the equilibria that are in the support of this outcome more than one player can be mixing; otherwise additional profiles would be part of the support. In none of the equilibria that are in the support of this outcome only one player can be mixing, since s and s' are strict. Therefore the profile (s'_i, s_{-i}) must be a Nash equilibrium; but this contradicts s being a strict Nash equilibrium. \square

6 Non-understanding, misunderstanding, and dialogue in Chicken

Consider the general class of Chicken games indicated in the left panel of Figure 4. The panel on the right, with $p + 2q + r = 1$ and $p, q, r \geq 0$, indicates a symmetric correlated distribution for that game.

	L	R			L	R
U	x, x	y, z	$z > x > y > 0$ $2x > z + y$	U	p	q
D	z, y	$0, 0$		D	q	r

Figure 4: General Chicken

The optimal symmetric correlated equilibrium solves the following program:

$$\begin{aligned}
& \max_{p,q,r} px + qy + qz \quad \text{s.t.} \\
& px + qy \geq pz \tag{1} \\
& qz \geq qx + ry \tag{2} \\
& p + 2q + r = 1 \tag{3} \\
& p, q, r \geq 0
\end{aligned}$$

If $r > 0$, we can increase p and q in equal proportion while maintaining condition (3). This maintains constraint (1), relaxes constraint (2), and increases the value of the objective. Hence, at any solution to the program we must have $r = 0$. With that and since $z > x$, constraint (2) is trivially satisfied and can thus be ignored. The program then simplifies to

$$\begin{aligned}
& \max_p px + \frac{1-p}{2}(y+z) \quad \text{s.t.} \\
& px + \frac{1-p}{2}y \geq pz \tag{4} \\
& 0 \leq p \leq 1 \tag{5}
\end{aligned}$$

and since $2x > y + z$, we want to increase p to the point where the remaining constraint is binding. Hence, the desired value of p is

$$\frac{y}{2(z-x) + y} =: \hat{p}$$

and therefore the optimal symmetric correlated equilibrium distribution is given by

The worst symmetric correlated equilibrium solves the following program:

	L	R
U	$\frac{y}{2(z-x)+y}$	$\frac{z-x}{2(z-x)+y}$
D	$\frac{z-x}{2(z-x)+y}$	0

Figure 5: Optimal symmetric correlated equilibrium distribution

$$\min_{p,q,r} px + qy + qz \quad \text{s.t.} \quad (6)$$

$$px + qy \geq pz \quad (6)$$

$$qz \geq qx + ry \quad (7)$$

$$p + 2q + r = 1 \quad (8)$$

$$p, q, r \geq 0$$

If constraint (7) is not binding, then $r < 1$ and therefore, we can reduce p and q in equal proportion while maintaining condition (8). This lowers the value of the objective while it maintains constraint (6). Hence, we have $qz = qx + ry$ or, equivalently, $q = r \frac{y}{z-x}$. Combining this with constraint (8), we can rewrite the objective as

$$\left(1 - 2\frac{ry}{z-x} - r\right)x + \frac{ry}{z-x}y + \frac{ry}{z-x}z,$$

which is decreasing in r since $2x > y + z$. Our ability to increase r is limited by the constraint that $1 - 2\frac{ry}{z-x} - r \geq 0$. If therefore we choose r such that $1 - 2\frac{ry}{z-x} - r = 0$, we get the values of r and (implicitly) q that minimize the payoff from a symmetric correlated equilibrium:

$$\hat{r} := \frac{z-x}{2y+z-x}$$

and hence

$$\hat{q} := \frac{y}{2y + z - x} = \frac{1 - \hat{r}}{2}.$$

Hence, the worst symmetric correlated equilibrium distribution is given by

	L	R
U	0	$\frac{y}{2y+z-x}$
D	$\frac{y}{2y+z-x}$	$\frac{z-x}{2y+z-x}$

Figure 6: Worst symmetric correlated equilibrium distribution

6.1 Inducing the optimal symmetric correlated equilibrium in Chicken through a language game

Consider a language game with message space $M = [0, 1]$ and players having partitional language types. The column player, C , has a single language type $\lambda_{\text{Col}} = (M; \{\{m\}_{m \in M}\})$, which is unconstrained. She has all messages available and can both correctly send and understand them. The row player has two language types: Language type $\lambda_{\text{Row}}^1 = (M; \{\{m\}_{m \in M}\})$ of the row player has access to all messages and sends and interprets them correctly. Language type $\lambda_{\text{Row}}^2 = (M; \{\{M\}\})$ of the row player also has access to all messages but can neither reliably send nor interpret them. Language type λ_{Row}^2 does not understand any of the messages in her repertoire. The probability that the row player has language type λ_{Row}^2 is commonly known to be $\psi \in (0, 1)$, where the value of ψ will be determined later.

The language game has two communication stages. In stage 1, the column player sends a message to the row player. In stage 2, the row player sends a message to the column player. After the two communication stages, Row and Column simultaneously take actions in the base game.

Players behave as follows: In the first communication stage, Column sends a message m_C that is generated by randomizing uniformly over M .

Row correctly observes message m_C when her language type is λ_{Row}^1 and observes a random uniform draw from M when her language type is λ_{Row}^2 . In the second communication stage, when Row's language type is λ_{Row}^1 , and she therefore correctly observes m_C , Row responds with a message $m_R \in [0, 1] \setminus [m_C, m_C + \chi]$, that is, Row avoids sending a message in the interval $[m_C, m_C + \chi]$ (where addition is modulo 1). The number $\chi \in (0, 1)$ will be determined later. When Row's language type is λ_{Row}^2 , Row responds with a message m_R that is a uniform draw from M , (which is implied by her language constraint). Conditional on Row having language type λ_{Row}^2 (and therefore randomizing uniformly over M) Row manages to avoid the interval $[m_C, m_C + \chi]$ with probability $1 - \chi$.

At the action stage, when Row's language type is λ_{Row}^1 and she therefore correctly observes m_C and can avoid the interval $[m_C, m_C + \chi]$ with probability 1, she takes action D . When, instead, Row's language type is λ_{Row}^2 , and she therefore can only avoid the interval $[m_C, m_C + \chi]$ with probability $1 - \chi$, Row takes action U . When Column observes $m_R \in [m_C, m_C + \chi]$, Column plays R . When Column observes $m_R \in [0, 1] \setminus [m_C, m_C + \chi]$, then Column plays L .

This behavior of players induces the distribution

	L	R
U	$\psi(1 - \chi)$	$\psi\chi$
D	$1 - \psi$	0

To ensure that the distribution is symmetric, require that $\psi\chi = 1 - \psi$.

$$\Rightarrow \chi = \frac{1 - \psi}{\psi} \quad \text{and therefore} \quad 1 - \chi = \frac{2\psi - 1}{\psi}.$$

With that requirement, the induced distribution is

	L	R
U	$2\psi - 1$	$1 - \psi$
D	$1 - \psi$	0

If we choose ψ so that $2\psi - 1 = \hat{p}$, we get the optimal symmetric correlated equilibrium distribution in Figure 5.

It remains to check the mutual optimality of players' strategies. First, consider the row player. At the message stage, language type λ_{Row}^1 anticipates obtaining the maximal achievable payoff z by avoiding to send messages in the interval $[m_C, m_C + \chi]$. Hence, her messaging rule is optimal. Language type λ_{Row}^2 cannot differentiate messages and is therefore indifferent between them.

At the action stage language type λ_{Row}^1 manages to avoid sending a message in the interval $[m_C, m_C + \chi]$ with probability one and therefore is certain that Column takes action L . Hence, action D , which the row player's strategy prescribes for her language type λ_{Row}^1 , is (uniquely) optimal for that language type.

Language type λ_{Row}^2 manages to avoid the interval $[m_C, m_C + \chi]$ with probability $1 - \chi = \frac{2\hat{p}}{1+\hat{p}}$. Hence, λ_{Row}^2 's payoff from action U equals $\frac{2\hat{p}}{1+\hat{p}}x + \frac{1-\hat{p}}{1+\hat{p}}y$ and the payoff from action D equals $\frac{2\hat{p}}{1+\hat{p}}z$. Thus, the payoffs from U and D are equal as long as $2\hat{p}x + (1 - \hat{p})y = 2\hat{p}z$. Recalling that $\hat{p} = \frac{y}{2(z-x)+y}$, this is equivalent to $2yx + 2(z-x)y = 2yz$, which is satisfied. This makes language type λ_{Row}^2 indifferent between actions U and D . Hence action U is optimal for language type λ_{Row}^2 .

Next, consider optimality of the column player's strategy. At the message stage, column is indifferent between all messages. Therefore it is optimal for column to randomize uniformly over all messages.

Conditional on observing $m_R \in [m_C, m_C + \chi]$, Column believes with probability one that Row's language type is λ_{Row}^2 and therefore that Row takes action U . Hence, action R , which is prescribed by Column's strategy after receiving message m_R , is (uniquely) optimal for Column.

Conditional on Column observing $m_R \in [0, 1] \setminus [m_C, m_C + \chi]$, Column's posterior probability of Row's language type being λ_{Row}^2 and hence Row taking action U equals $\frac{(1-\chi)\psi}{(1-\chi)\psi+1-\psi} = \hat{p}$. Hence, Column's payoff from action L equals $\frac{(1-\chi)\psi}{(1-\chi)\psi+1-\psi}x + \frac{1-\psi}{(1-\chi)\psi+1-\psi}y$. Column's expected payoff from R equals $\frac{(1-\chi)\psi}{(1-\chi)\psi+1-\psi}z$. These payoffs are the same provided $(1-\chi)\psi x + (1-\chi)y = (1-\chi)\psi z$. This condition is equivalent to $\frac{2\hat{p}}{1+\hat{p}} \frac{1+\hat{p}}{2}x + \frac{1-\hat{p}}{2}y = \frac{2\hat{p}}{1+\hat{p}} \frac{1+\hat{p}}{2}z$, which can be simplified to $2\hat{p}x + (1-\hat{p})y = 2\hat{p}z$. We verified before, when checking incentive compatibility for Row, that this equality holds. Hence, Column is indifferent between actions L and R following messages $m_R \in [0, 1] \setminus [m_C, m_C + \chi]$. Therefore taking action L is optimal for column after such messages.

6.2 Inducing the worst symmetric correlated equilibrium in Chicken

Consider again a language game with message space $M = [0, 1]$ in which players move sequentially. Unlike when inducing the optimal symmetric correlated equilibrium, here not all language types are partitional: there is a language type who may misunderstand messages as opposed to not understanding them.

The row player has two language types: Language type $\lambda_{\text{Row}}^1 = (M; \phi_R^1, \zeta_R^1)$, where $\phi_R^1(m) = \zeta_R^1(m) = m$ for all m , has all messages available and both the intention function and the interpretation function are the identity mapping. Language type $\lambda_{\text{Row}}^2 = (M; \phi_R^2, \zeta_R^2)$ likewise has all messages available but ϕ_R^1 and ζ_R^1 map all messages in M to the uniform distribution over M . That is regardless of intended and received messages, language type λ_{Row}^2 's sent and interpreted messages are drawn from the uniform distribution on M .⁹ The probability that the row player has language type λ_{Row}^2 is commonly known to be $\psi \in (0, 1)$ (where the value of ψ will be determined later). That is, with probability ψ the row player does not understand messages. With probabil-

⁹The row player's language types could also be expressed as partitional language types, as in the previous subsection. Using intention and interpretation functions highlights the contrast with the language constraints of the column player.

ity $1 - \psi$ the row player has language type λ_{Row}^1 , and therefore understands all messages.

The column player has a single, commonly known, language type $\lambda_{\text{Col}} = (M; \phi_C, \zeta_C)$. She has access to all messages and her intention function is the identity, i.e., $\phi_C(m) = m$ for all $m \in M$. Hence, the column player can reliably generate all intended messages. She interprets received messages m according to the interpretation function ζ_C , where

$$\zeta_C(m) \begin{cases} = m \text{ with probability } \chi \\ \sim U[M] \text{ otherwise} \end{cases}$$

Thus the column player correctly interprets received messages with probability $\chi \in (0, 1)$ (which will be determined later) and otherwise misunderstands them as some other message that is drawn uniformly from M .

The language game has two communication stages. In stage 1, the column player sends a message to the row player. In stage 2, the row player sends a message to the column player. Following the two communication stages, Row and Column simultaneously take actions in the base game.

Players behave as follows: In the first communication stage, Column sends a message m_C that is obtained by randomizing uniformly over M . Row correctly observes the message m_C when her language type is λ_{Row}^1 and observes a random uniform draw from M when her language type is λ_{Row}^2 . This is a case of non-understanding - row knows that she does not understand.

In the second communication stage, when Row's language type is λ_{Row}^1 , and she therefore correctly observes m_C , Row responds with a message $m_R = m_C$, i.e., Row "matches Column's message." When, instead, Row's language type is λ_{Row}^2 , Row responds with a message m_R that is a uniform draw from M , as dictated by her language type.

Column observes m_R with probability χ and otherwise observes a uniform draw from M . This is a case of misunderstanding; if Column's interpretation of the message received in the second communication stage does not match the message she sent in the first communication stage, Column is uncertain

whether the mismatch is due to Row's inability to match the message m_C or her own inability to properly interpret Row's reply. In the case of a mismatch of messages, column recognizes the mismatch but does not know whether to attribute the cause to the row player being language constrained or column's own misinterpretation.

At the action stage, language type λ_{Row}^2 of the row player takes action U ; language type λ_{Row}^1 of the row player takes action D ; when Column observes $m_R \neq m_C$, Column plays R ; and when Column observes $m_R = m_C$, Column plays L . The induced distribution is

	L	R
U	0	ψ
D	$(1 - \psi)\chi$	$(1 - \psi)(1 - \chi)$

Let $\chi = \frac{\psi}{1-\psi}$ to ensure symmetry and require that $(1 - \psi)(1 - \chi) = \hat{r}$ so that the distribution matches that of the worst symmetric equilibrium. Then $\psi = \frac{1-\hat{r}}{2} = \frac{y}{2y+z-x}$ and $\chi = \frac{1-\hat{r}}{1+\hat{r}} = \frac{y}{y+z-x}$.

Next, let us check the mutual optimality of players' strategies. First, consider optimality of Column's strategy. At the communication stage, Column's expected payoff is independent of the message she sends. Hence, her messaging rule is optimal. At the action stage, conditional on Column observing her message m_C being matched by Row's reply, Column believes that Row's language type is λ_{Row}^1 with probability 1 and therefore that Row takes action D . Hence, taking action L , which is prescribed by Column's strategy, is optimal.

Conditional on Column observing a mismatch between the message m_C she sent to Row and her interpretation of Row's reply, Column believes that Row was unable to match and therefore has language type λ_{Row}^2 with probability $\frac{\psi}{\psi+(1-\psi)(1-\chi)}$. Therefore, Column is indifferent between L and R provided $\psi z = \psi x + (1 - \psi)(1 - \chi)y$. This is equivalent to $\frac{1-\hat{r}}{2}z = \frac{1-\hat{r}}{2}x + (1 - \frac{1-\hat{r}}{2})(1 - \frac{1-\hat{r}}{1+\hat{r}})y$, which itself is equivalent to $\frac{1-\hat{r}}{2}z = \frac{1-\hat{r}}{2}x + \hat{r}y$. The

last equality holds from the definition of \hat{r} . Hence, it is optimal for Column to respond with action R to a mismatch, as is required by her strategy.

Next, consider the optimality of Row's strategy. At the messaging stage, Row's expected payoff from not matching Column's message is y . Her expected payoff from matching equals $\chi z = \frac{y}{y+z-x}z = y\frac{z}{z-(x-y)} > y$. That is, when possible Row prefers to match. Therefore Row's strategy is optimal at the messaging stage.

Conditional on Row having language type λ_{Row}^2 , and therefore being unable to match Column's message m_C , Row believes that Column takes action R with probability 1, and therefore finds it optimal to take action U .

Conditional on Row having language type λ_{Row}^1 and therefore being able to match m_C , Row expects Column to observe a match with probability χ and therefore to take action L with probability χ . Hence, provided that $\chi z = \chi x + (1 - \chi)y$, it is optimal for Row to take action D . That condition is equivalent to $\frac{y}{y+z-x}z = \frac{y}{y+z-x}x + \frac{z-x}{y+z-x}y$, which is satisfied.

6.3 Obtaining the entire set of correlated equilibrium outcomes in Chicken through a language game

We know from Calvó-Armengol [10] that the set of correlated equilibrium distributions of Chicken is a non-empty, convex and compact polytope in the 3-simplex with five vertices. The five vertices correspond to the two pure strategy equilibria, the mixed equilibrium, the optimal symmetric correlated equilibrium, and the worst symmetric correlated equilibrium.

Let the base game G be Chicken and o^k , $k = 1, \dots, 5$ the correlated equilibrium distributions that are the five vertices of the set of correlated equilibrium distributions of G . We saw that for each of those distributions, there is a language game $\Gamma_2(G, \mathcal{L}^k)$ that induces that distribution. For each language game $\Gamma_2(G, \mathcal{L}^k)$, let $\sigma^k = (\gamma^k, \rho^k)$ be an equilibrium that induces the distribution o^k . We will show that for any $\nu^k = 1, \dots, 5$ with $\nu^k \in [0, 1]$ and $\sum_{k=1}^5 \nu^k = 1$ there is a language structure \mathcal{L} and an equilibrium $\sigma = (\gamma, \rho)$ of

$\Gamma_3(G, \mathcal{L})$ that induces the correlated equilibrium distribution $o = \sum_{k=1}^5 \nu^k o^k$.

To describe the language structure \mathcal{L} , let $\mathcal{L}^k = (M^k, \Lambda^k, q^k)$, $k = 1, \dots, 5$ be language structures that make it possible to induce the five vertices of the set of correlated equilibria of Chicken with no more than two communication rounds. For the two pure-strategy equilibria and the mixed equilibrium of Chicken they can be trivial. For the other two vertices, we exhibited appropriate language structures above. Evidently, the message spaces M^k , $k = 1, \dots, 5$, can be chosen so that they are pairwise disjoint and do not include any element of the interval $[0, 1]$.

Define $M^0 := [0, 1]$ and $M := \bigcup_{k=0}^5 M^k$. The discussion in the previous two sections shows that it is without loss of generality to let $M_i^k = M^k$ for $i = \text{Row, Column}$ and $k = 1, \dots, 5$. For each k , let $\lambda_i^k = (M_i^k; \phi_i^k, \nu_i^k)$ be a typical language type of player i in the language structure \mathcal{L}^k . Given any 5-tuple $(\lambda_i^1, \dots, \lambda_i^5)$ of language types of player i (one for each of the five language structures), define a language type $\lambda_i = (M_i; \phi_i, \zeta_i)$ of player i as satisfying

$$\begin{aligned} M_i &:= [0, 1] \cup \bigcup_{k=1}^5 M_i^k, \\ \phi_i(m) &= \begin{cases} m & \text{if } m \in [0, 1] \\ \phi_i^k(m) & \text{if } m \in M_i^k \end{cases} \\ \zeta_i(m) &= \begin{cases} m & \text{if } m \in [0, 1] \\ \zeta_i^k(m) & \text{if } m \in M_i^k \end{cases} \end{aligned}$$

Denote the set of all these language types of player i by Λ_i , so that $\lambda_i \in \Lambda_i$. The probability that player i 's language type is λ_i then equals

$$q_i(\lambda_i) = q_i^1(\lambda_i^1) \times \dots \times q_i^5(\lambda_i^5).$$

This gives us a language structure $\mathcal{L} = (M, \Lambda, q)$ with language state space $\Lambda = \Lambda_1 \times \Lambda_2$, in which each language state $\lambda = (\lambda_1, \lambda_2) \in \Lambda_1 \times \Lambda_2$ has commonly known probability $q(\lambda) = q_1(\lambda_1) \times q_2(\lambda_2)$.

Partition the interval $[0, 1]$ into subintervals ι^k , each of which has length ν^k . Use $a_i(j)$ to denote the action of player i that is required in the Nash equilibrium that achieves player j 's lowest payoff from a Nash equilibrium of G . For the language game $\Gamma_3(G, \mathcal{L})$ with three communication rounds and language structure $\mathcal{L} = (M, \Lambda, q)$ consider the strategy profile $\sigma = (\gamma, \rho)$ defined by the following four rules: First,

$$\gamma_i(\emptyset, \lambda_i) = U[0, 1], \forall i \in I, \forall \lambda_i \in \Lambda_i.$$

That is, in the first communication round, both players randomize uniformly over all messages in $[0, 1]$. Second,

$$\gamma_i((\emptyset, \mathbf{m}_i^1), \lambda_i) = \begin{cases} \gamma_i^k(\emptyset, \lambda_i^k) \text{ if } m_{i,j}^1 \in [0, 1] \forall j \\ \text{and } m_{i,i}^1 + m_{i,-i}^1 \in \iota^k \pmod{1} \\ 0 \text{ otherwise} \end{cases}$$

Hence, in the second communication round, both players send messages as prescribed for the first round by the equilibrium σ^k of the language game $\Gamma_2(G, \mathcal{L}^k)$, provided both players sent messages in $[0, 1]$ in the first round and the sum of those messages (modulo 1) belongs to ι^k . Otherwise, they send a default message 0. Third,

$$\gamma_i((\emptyset, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i) = \begin{cases} \gamma_i^k(\emptyset, \mathbf{m}_i^2, \lambda_i^k) \text{ if } m_{i,j}^1 \in [0, 1] \forall j, m_{i,j}^2 \in M^k \forall j \\ \text{and } m_{i,i}^1 + m_{i,-i}^1 \in \iota^k \pmod{1} \\ 0 \text{ otherwise} \end{cases}$$

In the third communication round, both players send messages as prescribed for the second round by the equilibrium σ^k of the language game $\Gamma_2(G, \mathcal{L}^k)$, provided provided there have not been detectable deviations in the first and second round and the sum of first-round messages (modulo 1) belongs to ι^k . Otherwise, players send a default message 0. Fourth, and finally

$$\rho_i((\emptyset, \mathbf{m}_i^1, \mathbf{m}_i^2, \mathbf{m}_i^3), \lambda_i) = \begin{cases} \rho_i^k(\emptyset, \mathbf{m}_i^2, \mathbf{m}_i^3, \lambda_i^k) & \text{if } m_{i,j}^1 \in [0, 1] \forall j, m_{i,j}^2 \in M^k \forall j, \\ & m_{i,j}^3 \in M^k \forall j \text{ and } m_{i,i}^1 + m_{i,-i}^1 \in \iota^k \pmod{1}, \\ a_i(j') & \text{if } j' = \min_j \{j | m_{i,j}^{t'} \notin M^k\}, \\ & \text{where } t' = \min_t \{t \in \{1, 2, 3\} | \exists j \text{ s.t. } m_{i,j}^t \notin M^k\} \end{cases}$$

Hence, in the final (action) stage both players take actions as prescribed for the action stage by the equilibrium σ^k of the language game $\Gamma_2(G, \mathcal{L}^k)$, provided there have not been detectable deviations in any of the communication rounds and the sum of first-round messages (modulo 1) belongs to ι^k . Otherwise, they play the stage-game Nash equilibrium that minimizes the payoff of the first deviator with the lowest player index.

With these strategies players operate a jointly controlled lottery in the first communication round. Neither of them can individually influence the likelihood with which the intervals ι^k , $k = 1, \dots, 5$ are reached as long as they use messages in M^0 . Unilateral deviations by player i to sending a message not belonging to M^0 are detected by the other player and punished by that player taking the action that minimizes player i 's payoff from Nash equilibria in G . The punishment is effective because for each player every correlated equilibrium payoff of G is at least as large as that player's lowest payoff from a Nash equilibrium of G .

If both players send messages in M^0 in the first round and the sum of these messages modulo 1 belongs to the interval $\iota^{k'}$ the strategies of both players prescribe the same behavior as does $\sigma^{k'}$ as long as all subsequent messages belong to $M^{k'}$. Deviations by player i to messages not belonging to $M^{k'}$ are detected by the other player and punished by that player taking the action that minimizes player i 's payoff from Nash equilibria in G . Deviations by player i to messages in $M^{k'}$ that are not prescribed by $\sigma^{k'}$ are not profitable since $\sigma^{k'}$ is an equilibrium of $\Gamma_2(G, \mathcal{L}^{k'})$ by assumption. Finally, if all players only sent messages in $M^{k'}$ prior to playing the game

G , they are in the same situation as they would be in $\Gamma_2(G, \mathcal{L}^{k'})$ and hence the action prescribed by their strategy is optimal given their information. If some player deviated to sending a message not in $M^{k'}$, players' strategies prescribe playing a Nash equilibrium of G and this fact is commonly known since deviations to messages outside of $M^{k'}$ are publicly observable. Hence, there is no incentive to deviate at the action stage.

7 A game with two players, commonly known language constraints and a unique Nash equilibrium

All the results and examples up to this point make use of privately known language constraints and leverage the existence of best-reply sets with multiple strict equilibria in the base game. In this section I present a two-player example that drops both of these features.

	X	Y	Z
A	0,0	1,2	2,1
B	2,1	0,0	1,2
C	1,2	2,1	0,0

Figure 7: A base game with a unique Equilibrium

The game in Figure 7 has a unique Nash equilibrium in mixed strategies - players randomize uniformly over all of their actions. In addition, there is a correlated equilibrium with probability weight $1/6$ on each of the action profiles with positive payoffs.

The correlated equilibrium outcome can be realized as a Nash equilibrium of a language game with a partitional language structure: The set of messages available to both players is $M = \{*, \#, \&, \%, \$, \diamond\}$. Column's language

constraint is $\{\{*, \#\}, \{\&, \%\}, \{\$, \diamond\}\}$ and Row's is $\{\{\#, \&\}, \{\%, \$\}, \{\diamond, *\}\}$. These language constraints are commonly known. The language game has one communication round prior to the base game: Column sends a message to Row.

This language game has a Nash equilibrium in which Column randomizes uniformly over M and uses the action rule $\{*, \#\} \mapsto X$, $\{\&, \%\} \mapsto Y$, $\{\$, \diamond\} \mapsto Z$. Row's strategy in this equilibrium is $\{\#, \&\} \mapsto C$, $\{\%, \$\} \mapsto A$, $\{\diamond, *\} \mapsto B$: Having sent either message $*$ or message $\#$, and being unable to distinguish these two messages, Column expects Row to take actions B and C with equal probability. Hence, it is (uniquely) optimal for Column to take action X , after having sent either $*$ or $\#$. Having received either message $\#$ or message $\&$, and being unable to distinguish them, Row assigns equal probability to Column taking either action X or Y . Therefore responding with action C is (uniquely) optimal Row. Similar arguments apply to the other messages.

8 An aside on more than two players

The focus of this paper is on simple language games with no more than three communication rounds and equilibrium constructions that require no more than two players. With more than two players, language constraints can be used to model situations in which subsets of players can communicate among themselves without others understanding.¹⁰ This is analogous to having subsets of players communicate through channels that only they can access and therefore suggests that results from the literature on implementing correlated equilibria via direct communication through restricted channels (following Bárány [5]) may carry over to our environment.

An adaptation of the construction of Gerardi [17] can be used to show how to induce the entire set of rational correlated equilibria as Bayesian

¹⁰This is reminiscent of the roles played by Navajo code talkers, cants, argots, Cockney rhyming slang, and dog whistles (I am grateful to Jeffrey Shrader for pointing out this connection).

Nash equilibria of games with language constraints, provided that there are at least five players: whenever in Gerardi's construction only designated players can talk to each other, in the corresponding language game the same players communicate in a language only they can understand. Deviations to attempting to communicate using a language understandable to others are punished by minmaxing the deviator. Since the argument is straightforward, I state the result without proof.

Proposition 4 *For every base game G with five or more players and every rational correlated equilibrium outcome \mathcal{O} of G , there exists a finite language game that has a Bayesian Nash equilibrium that induces the outcome \mathcal{O} .*

By assuming that there are at least five players, Gerardi avoids requiring that players sometimes are able to verify past messages and achieves implementation via sequential equilibria. My adaptation to language games inherits the former feature, but not the latter. I make no attempt to ensure sequential rationality, which allows me to minmax deviators.

References

- [1] AUMANN, ROBERT J., MICHAEL MASCHLER AND RICHARD B. STEARNS [1968], "Repeated Games with Incomplete Information: An Approach to the Nonzero Sum Case. Reports to the US Arms Control and Disarmament Agency," ST-143, Chapter IV, 117-216, reprinted in AUMANN, ROBERT J. AND MICHAEL MASCHLER (with the collaboration of Richard B. Stearns) [1995], *Repeated Games with Incomplete Information* MIT Press: Cambridge.
- [2] AUMANN, ROBERT J. [1974], "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics* **1**, 67–96.
- [3] AUMANN, ROBERT J. [1987], "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica* **55**, 1–18.

- [4] AUMANN, ROBERT J., AND SERGIU HART [2003], “Long cheap talk,” *Econometrica* **71**, 1619–1660.
- [5] BÁRÁNY, IMRE [1992], “Fair Distribution Protocols or how the Players Replace Fortune,” *Mathematics of Operations Research* **17**, 327–340.
- [6] BASU, KAUSHIK, AND JÖRGEN W. WEIBULL [1991], “Strategy Subsets Closed Under Rational Behavior” *Economics Letters* **36**, 141–146.
- [7] BEN-PORATH, ELCHANAN [1998], “Correlation without Mediation: Expanding the Set of Equilibrium Outcomes by ‘Cheap’ Pre-Play Procedures,” *Journal of Economic Theory* **80**, 108–122.
- [8] BLUME, ANDREAS [2000], “Coordination and Learning with a Partial Language,” *Journal of Economic Theory* **95**, 1–36.
- [9] BLUME, ANDREAS AND OLIVER J. BOARD [2013], “Language Barriers,” *Econometrica* **81**, 781–812.
- [10] CALVÓ-ARMENGOL, ANTONIO [2006], “The Set of Correlated Equilibria of 2×2 Games,” Universitat Autònoma de Barcelona Working Paper.
- [11] CRAWFORD, VINCENT P. AND HANS HALLER [1990], “Learning how to Cooperate: Optimal Play in Repeated Coordination Games,” *Econometrica* **58**, 571–595.
- [12] CRÉMER, JACQUES, LUIS GARICANO AND ANDREA PRAT [2007], “Language and the Theory of the Firm,” *Quarterly Journal of Economics* **122**, 373–407.
- [13] DEWATRIPONT, MATHIAS AND JEAN TIROLE [2005], “Modes of Communication,” *Journal of Political Economy* **113**, 1217–1238.
- [14] FORGES, FRANÇOISE [1986], “An Approach to Communication Equilibria,” *Econometrica* **54**, 1375–1385.

- [15] FORGES, FRANÇOISE [1990], “Universal Mechanisms,” *Econometrica* **58**, 1341–1364.
- [16] FORGES, FRANÇOISE [2020] “Correlated Equilibria and Communication in Games,” in David Pérez-Castrillo (Editor), Marilda Sotomayor (Editor) and Filippo Castiglione (Editor), *Complex Social and Behavioral Systems: Game Theory and Agent-Based Models*, 107–118, Springer, New York.
- [17] GERARDI, DINO [2004], “Unmediated Communication in Games with Complete and Incomplete Information,” *Journal of Economic Theory* **114**, 104–131.
- [18] DODIS, YEVGENIY, SHAI HALEVI, AND TAL RABIN [2000] “A cryptographic solution to a game theoretic problem,” in *Advances in Cryptology—CRYPTO 2000: 20th Annual International Cryptology Conference Santa Barbara, California, USA, August 20–24, 2000 Proceedings 20*, 112–130, Springer, Berlin, Heidelberg.
- [19] HINTIKKA, JAAKKO [1979] “Language-games,” in Hintikka, Carlson, Peacocke, Rantala, and Saarinen (Editors), *Game-Theoretical Semantics: Essays on Semantics*, 1–26, Springer, Dordrecht, Netherlands.
- [20] JÄGER, GERHARD, LARS P. METZGER AND FRANK RIEDEL [2011], “Voronoi languages: Equilibria in cheap-talk games with high-dimensional types and few signals,” *Games and Economic Behavior* **73**, 517–537.
- [21] LEHRER, EHUD [1991], “Internal Correlation in Repeated Games,” *International Journal of Game Theory* **19**, 431–456.
- [22] LEHRER, EHUD [1996], “Mediated Talk,” *International Journal of Game Theory* **25**, 177–188.
- [23] LEHRER, EHUD, AND SYLVAIN SORIN [1997], “One-Shot Public Mediated Talk,” *Games and Economic Behavior* **20**, 131–148.

- [24] MYERSON, ROGER B [1986], “Multistage Games with Communication,” *Econometrica* **54**, 323–358.
- [25] RUBINSTEIN, ARIEL [1996], “Why Are Certain Properties of Binary Relations Relatively More Common in Natural Language?,” *Econometrica* **64**, 343–355.
- [26] TERCIEUX, OLIVIER [2006], “ p -Best response set,” *Journal of Economic Theory* **131**, 45–70.
- [27] TERCIEUX, OLIVIER [2006], “ p -Best response set and the robustness of equilibria to incomplete information,” *Games and Economic Behavior* **56**, 371–384.
- [28] URBANO, AMPARO, AND JOSE E. VILA [2002], “Computational complexity and communication: Coordination in two-player games,” *Econometrica* **70**, 1893–1927.
- [29] VOORNEVELD, MARK [2004], “Preparation,” *Games and Economic Behavior* **48**, 403–414.
- [30] WITTGENSTEIN, LUDWIG [1958], *Philosophical Investigations*, Prentice Hall, Upper Saddle River, NJ.