Language Games: Correlation through Non-Understanding, Dialogue, Inarticulateness, and Misunderstanding*

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

January 8, 2024 (first draft); this version: January 22, 2024

Abstract

A *language game* is a finite complete-information game preceded by pre-play communication with explicit constraints on players' ability to produce and understand messages and on their knowledge of each other's constraints. Players communicate directly and publicly but may not understand or may misunderstand each other's messages. The paper gives conditions under which it is possible to implement correlated equilibria outside the convex hull of the set of Nash equilibria through language games. These conditions can be satisfied in games with any numbers of players, including two. In the game of Chicken it is possible to induce the entire set of correlated equilibria via a language game.

*I have am grateful to Joel Sobel for his 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 can 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 it is only possible to induce correlated equilibrium outcomes that are in the convex hull of the set of Nash equilibrium outcomes. 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) pre-play communication with explicit constraints on players' ability to produce and understand messages and on their knowledge of each other's constraints. The terminology is not entirely accidental. Wittgenstein [27] 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.

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 "hypertension" from "hypotension"), 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 (e.g., if she is likely to interpret "hypertension" correctly but possibly interprets it as "hypotension"). In addition, players may be constrained by limited message availability (e.g., the word "hypotension" may not be part of their active vocabulary).

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 3, 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. An alternative pathway, explored in Section 5, does not leverage simultaneous message exchange: in any given communication round only one player's message matter. 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, what matters is the potential *inarticulateness* of the challenged player. Finally, when showing how to implement the entire set of correlated equilibria in Chicken in Section 6, I make (limited) use of *misunderstandings*: In the implementation of 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 row did or did not understand.

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.[1] This result cannot be extended to two-player games since there message 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. 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 player to be able to verify past messages when there are at least five players. Lehrer and Sorin [21], building on Lehrer [19] and [20], 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 has considered various versions of constraints on players' language, including finite message spaces, Crémer, Garicano, and Prat [12] and Jäger, Metzger, and Riedel [18]; symmetry constraints on strategies, Crawford and Haller [11] and Blume [8]; limited sets of relations on partially nameless objects, Rubinstein [23]; 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 consider constraints on sending messages separately from constraints on understanding messages, whereas here I allow language types with both kinds of constraints.

[1] For incomplete information games Forges [15] shows that with four or more players every communication equilibrium distribution (Myerson [22], Forges [14]) is a Bayes Nash equilibrium distribution of an appropriate cheap-talk extension.

2 Example: Non-understanding in Chicken

In the game of Chicken, shown in the left panel of Figure 1, the maximal symmetric payoff players can achieve through direct pre-play communication (using a jointly controlled lottery) when they have no difficulty sending and understanding messages is 3.5. The right panel shows a symmetric correlated equilibrium distribution that achieves an expected payoff equal to $3.\overline{6}$ for each player. I will show that there is a *language game*, i.e., a game in which players may be language constrained and may have private information about those constraints, in which this correlated-equilibrium distribution can be achieved through a single round of simultaneous direct pre-play communication.

	L	R
U	4,4	2,5
D	5,2	0,0

	L	R
U	$\frac{1}{3}$	$\frac{1}{3}$
D	$\frac{1}{3}$	0

Figure 1: Chicken

Consider a language game in which players simultaneously send messages from the message space $M = \{*, \#\}$ prior to playing the base game in the left panel of Figure 1.[2] Each player may be constrained by being unable to distinguish the two messages. A player who cannot distinguish the messages has to treat them identically when sending and responding to them. A player who can distinguish the two messages is said to "understand" them. Otherwise, she suffers from non-understanding.

In this example, a player's *language type* corresponds to her ability or inability to understand messages. We can represent these language types as partitions of the message space. Given a language type, each partition element contains those messages the player cannot distinguish from each other. Each player has two possible language types, one corresponding to the trivial partition and the other to the finest partition.

[2] The 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.

Let the row player have language type $\lambda^1_{\text{Row}} = \{\{*\}, \{\#\}\}$ with probability $\frac{2}{3}$ and language type $\lambda^2_{\text{Row}} = \{\{*, \#\}\}$ with probability $\frac{1}{3}$. Language type λ^1_{Row} understands both messages, while her other language type λ^2_{Row} does not understand the two messages. The column player, likewise, has two language types, $\lambda^1_{\text{Col}} = \{\{*\}, \{\#\}\}$ with probability $\frac{2}{3}$ and $\lambda^2_{\text{Col}} = \{\{*, \#\}\}$ with probability $\frac{1}{3}$. 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 the row player randomize uniformly over the messages in $\{*, \#\}$.[3] At the response stage, λ^1_{Row} follows the rule $(*, *) \mapsto D$, $(\#, \#) \mapsto D$, $(*, \#) \mapsto U$ and $(\#, *) \mapsto U$ and λ^2_{Row} takes action U. At the communication stage, both language types of the column player randomize uniformly over the messages in $\{*, \#\}$. At the response stage, λ^1_{Col} follows the rule $(*, *) \mapsto L$, $(\#, \#) \mapsto L$, $(*, \#) \mapsto R$ and $(\#, *) \mapsto R$ and λ^2_{Col} 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 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 a row player and L if they are a column player.

Consider the incentive constraints for the row player at the communication stage. She 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_{Row} of the row player at the response stage, first note that following message histories $(*, *)$ and $(\#, \#)$ both language types of the column player take action L. Therefore action D is (uniquely) optimal for λ^1_{Row} following those histories. Following message histories $(*, \#)$ and $(\#, *)$, the column player takes action

[3] For language type λ^2_{Row} this is dictated by the requirement that both messages have to be treated identically.

L if and only if her language type is λ^2_{Col}. Since that language type has probability $\frac{1}{3}$ action U is (uniquely) optimal for λ^1_{Row} following those message histories.

It remains to examine the incentive constraints for language type λ^2_{Row} of the row player at the response stage. This language type expects the column player to take action L with probability $\frac{2}{3}$, either because the column player's language type is λ^2_{Col}, which occurs with probability $\frac{1}{3}$ or because it is λ^1_{Col} (which has probability $\frac{2}{3}$) and the message history is $(*, *)$ or $(\#, \#)$ (which has probability $\frac{1}{2}$). Hence language type λ^2_{Row} is indifferent between actions U and D, and thus action U is optimal for her. This shows that the row player is using a best response. By symmetry, so is the column player.

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. Non-understanding leads players to take actions that do not match the Nash equilibria that would otherwise be induced by the jointly controlled lottery. Uncertainty about non-undertanding 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, a significant role of constraints on sending messages, and language types who suffer from misunderstanding, rather than non-understanding.

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 (we 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 $U_i : A \to \mathbb{R}$ player i's payoff function, with the usual extension to mixed strategies.

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 set of messages, M, that player i can send. The function $\phi_i : M_i \to \Delta(M)$ 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 $\zeta_i : M \to \Delta(M)$ 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 $q(\lambda) = q_1(\lambda_1) \times \cdots \times q_I(\lambda_I)$. The triple triple $\mathcal{L} = (M, \Lambda, q)$ is a *language structure.*[4] Language structures include as a special case *degenerate language structures*, $\mathcal{D}_M$, in which the universe of messages is M and there is certainty that all players can send and understand all messages in M.

In most of the cases I consider, the intention and interpretation functions ϕ_i and ζ_i can be expressed in terms of a partition $\mathcal{Q}_i$ of the universal message space M, with typical element $Q_i \in \mathcal{Q}_i$. The partition $\mathcal{Q}_i$ describes 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, both when sending messages and when interpreting received messages.

Given a partition $\mathcal{Q}_i$ of the universal message space M, the intention intention and interpretation functions, ϕ_i and ζ_i can be derived from the partition as follows:

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

where $\mathrm{U}[X]$ denote the uniform distribution over the set X. Throughout, I assume that $M_i \cap Q_i \neq \varnothing \Rightarrow Q_i \subseteq M_i$, i.e., a player can always distinguish messages in her active

[4]Evidently, the information players receive in a language game could alternatively be generated by a mediator. Therefore, since the 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.

vocabulary, M_i, from messages in her passive vocabulary, $M \setminus M_i$. If intention functions and interpretation functions can be expresses through a partition, then a typical language type takes the form $\lambda_i = (M_i; \mathcal{Q}_i)$. If, in addition, $M_i = M$ for all i, and therefore language types can be expressed as $\lambda_i = (M; \mathcal{Q}_i)$, I say that the language type structure is *partitional*.

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

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

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

I will be 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 of 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 $\hat{h}_i^t$, $t = 1, \ldots, n$ 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 result 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 $\tilde{A}_i$.[5] 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 $\iota : M \to \{1, \ldots, |M|\}$ a bijection from M to $\{1, \ldots, |M|\}$. 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

[5] Best response sets are the p-best response sets as defined by Tercieux [25] for $p = 1$. They differ from Voorneveld's [26] *prep sets* by allowing correlated beliefs. They coincide with prep sets in two-player games. Tercieux's [24] definition of p-best response sets is slightly different from the one given in Tercieux [25] 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

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

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

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

and

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

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
u	x,-x	-x,x
d	-x,x	x,-x

	L	R
U	9,9	0,8
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 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 $E(G) := \{v \in V(G) | u > v \Rightarrow u \notin V(G)\}$, the set of efficient payoffs in the convex hull of Nash equilibrium payoffs of G. Let $D(G) := \{u \in \mathbb{R}^I | \exists v \in E(G) \text{ with } u > v\}$ 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 $V(G) \cap V^o(G; s, s', i) = \varnothing$. 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 $E(G)$.*

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 $z^2 + y^2 > zx + yx$.[6] 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 large weight on the profile (U, R). 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 the convex combinations of (z, y) and (y, z) and 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.

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

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

Figure 3: General Chicken

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 different 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 $e, e' \in E(G)$ 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 outcome. The remainder of this section investigates what can be done 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 "constrained 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.

The intuition for that result when there are two strict equilibria in the base game 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. The constrained player tries to match the message of the designated player. If she does not succeed, the equilibrium less favorable to her player 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 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 $\Gamma_2(G_{\tilde{A}}, \mathcal{L}(\eta, M))$ with $\eta \in (0,1)$. Fix a player $i' \neq i$. Let $\sigma^M = (\gamma^M, \rho^M)$ in $\Gamma_2(G_{\tilde{A}}, \mathcal{L}(\eta, M))$ be the strategy profile defined by

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

$$\gamma_j^M((\varnothing, \mathbf{m}_j^1), \lambda_j) = U[M], \forall j \neq i, \forall \lambda_j \in \Lambda_j, \forall \mathbf{m}_j^1 \in M^I,$$

$$\gamma_i^M((\varnothing, \mathbf{m}_i^1), \lambda_i^1) = m_{i,i'}^1, \forall \mathbf{m}_i^1 \in M^I,$$

$$\gamma_i^M((\varnothing, \mathbf{m}_i^1), \lambda_i^2) = U[M], \forall \mathbf{m}_i^1 \in M^I,$$

$$\rho_j^M((\varnothing, \mathbf{m}_j^1, \mathbf{m}_j^2), \lambda_j) = \begin{cases} s'_j & \text{if } m_{j,i}^2 = m_{j,i'}^1 \\ s_j & \text{otherwise} \end{cases} \quad \forall j \neq i, \forall \mathbf{m}_j^1, \mathbf{m}_j^2 \in M^I$$

$$\rho_i^M((\varnothing, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i^1) = \begin{cases} s'_i & \text{if } m_{i,i}^2 = m_{i,i'}^1 \\ s_i & \text{otherwise} \end{cases} \quad , \forall \mathbf{m}_i^1, \mathbf{m}_i^2 \in M^I$$

and

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

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 $\gamma_j^M(\varnothing, \lambda_j) = \mathrm{U}[M], \forall j \in I, \forall \lambda_j \in \Lambda_j$ 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 $\gamma_j^M((\varnothing, \mathbf{m}_j^1), \lambda_j) = \mathrm{U}[M], \forall j \neq i, \forall \lambda_j \in \Lambda_j$, 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 $\gamma_i^M((\varnothing, \mathbf{m}_i^1), \lambda_i^1) = m_{i,i'}^1$ is optimal. Language type λ_i^2 cannot distinguish any message and therefore the signaling rule $\gamma_i^M((\varnothing, \mathbf{m}_i^1), \lambda_i^2) = U[M]$ 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 $\rho_j^M((\varnothing, \mathbf{m}_j^1, \mathbf{m}_j^2), \lambda_j)$ 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 $\eta > 0$ $\rho_j^M((\varnothing, \mathbf{m}_j^1, \mathbf{m}_j^2), \lambda_j)$ 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 $\rho_i^M((\varnothing, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i^1)$

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 $\rho_i^M((\varnothing, \mathbf{m}_i^1, \mathbf{m}_i^2), \lambda_i^2)$ 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 $\Gamma_1(G_{\tilde{A}}, \mathcal{L}(\epsilon, \eta))$ with $\epsilon, \eta \in (0, 1)$. Let $\sigma = (\gamma, \rho)$ in $\Gamma_1(G_{\tilde{A}}, \mathcal{L}(\epsilon, \eta))$ be the strategy profile defined by

$$\gamma_\ell(\emptyset, \lambda_\ell^1) = m^*$$

$$\gamma_\ell(\emptyset, \lambda_\ell^2) = U[M \setminus \{m^*\}]$$

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

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

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

$$\rho_i((\emptyset, \mathbf{m}^1), \lambda_i^2) = s'_i$$

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 that ℓ do no affect play at the action 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 that 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 that 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'_{-i} 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 equilibrium 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 again the general game of Chicken. The left panel of Figure 4 reproduces the payoff structure for convenience. 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
U	x, x	y, z
D	z, y	0,0

$z > x > y > 0$
$2x > z + y$

	L	R
U	p	q
D	q	r

Figure 4: General Chicken

The optimal symmetric correlated equilibrium solves the following program:

$$\max_{p,q,r} \; px + qy + qz \quad \text{s.t.}$$

$$px + qy \geq pz \tag{1}$$

$$qz \geq qx + ry \tag{2}$$

$$p + 2q + r = 1 \tag{3}$$

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

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

$$\max_{p} \; px + \frac{1-p}{2}(y + z) \quad \text{s.t.}$$

$$px + \frac{1-p}{2} y \geq pz \tag{4}$$

$$0 \leq p \leq 1 \tag{5}$$

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

	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

The worst symmetric correlated equilibrium solves the following program:

$$
\begin{aligned}
&\min_{p,q,r}\ px + qy + qz \quad \text{s.t.} \\
&px + qy \geq pz \qquad (6)\\
&qz \geq qx + ry \qquad (7)\\
&p + 2q + r = 1 \qquad (8)\\
&p, q, r \geq 0
\end{aligned}
$$

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 Implementing 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 $\lambda^1_{\text{Row}} = (M; \{\{m\}_{m \in M}\})$ of the row player has access to all messages and sends and interprets them correctly. Language type $\lambda^2_{\text{Row}} = (M; \{\{M\}\})$ of the row player also has access to all messages but can neither reliably send nor interpret them. Language type λ^2_{Row} does not understand any of the messages in her repertoire. The probability that the row player has language type λ^2_{Row} 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^1_{Row} and observes a random uniform draw from M when her language type is λ^2_{Row}. In the second communication stage, when Row's language type is λ^1_{Row}, 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^2_{Row}, 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^2_{Row} (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^1_{Row} 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^2_{Row}, 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^1_{Row} 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^2_{Row} cannot differentiate messages and is therefore indifferent between them.

At the action stage language type λ^1_{Row} 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^1_{Row}, is (uniquely) optimal for that language type.

Language type λ^2_{Row} manages to avoid the interval $[m_C, m_C+\chi]$ with probability $1-\chi = \frac{2\hat{p}}{1+\hat{p}}$. Hence, λ^2_{Row}'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^2_{Row} indifferent between actions U and D. Hence action U is optimal for language type λ^2_{Row}.

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^2_{Row} 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^2_{Row} 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 Implementing 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 implementing the optimal symmetric correlated equilibrium, here not all language types are partitional. This captures the possibility of misunderstanding messages as opposed to not understanding them.

The row player has two language types: Language type $\lambda^1_{\text{Row}} = (M; \phi^1_R, \zeta^1_R)$, where $\phi^1_R(m) = \zeta^1_R(m) = m$ for all m, has all messages available and both the intention function and the interpretation function are the identity mapping. Language type $\lambda^2_{\text{Row}} = (M; \phi^2_R, \zeta^2_R)$ likewise has all messages available but ϕ^1_R and ζ^1_R map all messages in M to the uniform distribution over M. That is regardless of intended and received messages, language type λ^2_{Row}'s sent and interpreted messages are drawn from the uniform distribution on M.[7] The

[7] The row player's language types could also be expressed as partitional language types, as in the the previous subsection. Using intention and interpretation functions highlights the contrast with the language

probability that the row player has language type λ^2_{Row} 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 probability $1-\psi$ the row player has language type λ^1_{Row}, 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^1_{Row} and observes a random uniform draw from M when her language type is λ^2_{Row}. 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^1_{Row}, 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^2_{Row}, 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

constraints of the column player.

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^2_{Row} of the row player takes action U; language type λ^1_{Row} 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^1_{Row} 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^2_{Row} 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^2_{Row}, 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^1_{Row} 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 U. 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, $k = 1, \ldots, 5$ 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 implements that distribution. For each language game $\Gamma_2(G, \mathcal{L}^k)$, let $\sigma^k = (\gamma^k, \rho^k)$ be an equilibrium that implements the distribution o^k. We will show that for any $\nu^k = 1, \ldots, 5$ 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 implements the correlated equilibrium distribution $o = \sum_{k=1}^n \nu^k o^k$.

To describe the language structure $\mathcal{L}$, let $\mathcal{L}^k = (M^k, \Lambda^k, q^k)$, $k = 1, \ldots, 5$ be language structures that make it possible to implement 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 the other two vertices, we exhibited appropriate language structures above. Evidently, the message spaces M^k, $k = 1, \ldots, 5$, 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 $M^k_i = M^k$ for $i =$ Row, Column and $k = 1, \ldots, 5$. For each k, let $\lambda^k_i = (M^k_i; \phi^k_i, \nu^k_i)$ be a typical language type of player i in the language structure $\mathcal{L}^k$. Given any 5-tuple $(\lambda^1_i, \ldots, \lambda^5_i)$ 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

$M_i := [0,1] \cup \bigcup_{k=1}^{5} M_i^k$,

$$\phi_i(m) = \begin{cases} m \text{ if } m \in [0,1] \\ \phi_i^k(m) \text{ if } m \in M_i^k \end{cases}$$

$$\zeta_i(m) = \begin{cases} m \text{ if } m \in [0,1] \\ \zeta_i^k(m) \text{ if } m \in M_i^k \end{cases}$$

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

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

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,

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

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

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

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,

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

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

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

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, $k = 1, \ldots, 5$ 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 equiilibrium

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 equals 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 Five or more 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.[8] 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.

In this section, I outline how the construction of Gerardi [17] can be adapted to show how to implement the entire set of rational correlated equilibria as Bayesian Nash equilibria of games with language constraints, provided that there are at least five players. 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.

Suppose that there are at least five players and the base game G has rational parameters. Gerardi considers *plain cheap-talk extensions* $\Gamma^{ext}(G)$ of the base game G, in which players

[8]This 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).

communicate in a finite number of stages prior to taking actions in the base game G. A plain cheap-talk extension specifies for each stage which players can send messages, the set of messages available to them, and who receives those messages. Communication is direct (that is unmediated), may be public (addressed to all players), private (addressed to a single player), or semi-public (addressed to a non-trivial subset of players). Gerardi only uses plain cheap-talk extensions in which players who are designated as senders at some stage send a single message at that stage.[9]

Suppose that $\Gamma^{ext}(G)$ is a plain cheap-talk extension of G with n communication stages and σ^{ext} a Bayesian Nash equilibrium of $\Gamma^{ext}(G)$ that implements a rational correlated equilibrium outcome $\mathcal{O}$ of G in the manner proposed by Gerardi. I will construct a language game $\Gamma_{n+1}(G, \mathcal{L})$ that has a Bayesian Nash equilibrium σ that also implements the rational correlated equilibrium outcome $\mathcal{O}$ of G.

This language game has the following language structure $\mathcal{L}$: For each subset J of the set of players I, there is a set of messages M_J that only players in J have available and only they can understand. Players in J can reliably send and interpret messages in M_J – in this set their sent messages match their intended messages and their interpreted messages match their received messages. Players in $I \setminus J$ cannot send messages in M_J and cannot make distinctions among messages in M_J when they observe them. All players can differentiate messages in M_J from messages not in M_J.[10] Assume the message spaces M_J to be finite and to be at least as large as the minimum of the number of action profiles in G and any set of messages available to any player at any stage in $\Gamma^{ext}(G)$.[11]

For every player j, let $a^k_{-j} \in A_{-j}$, $k = 1, \ldots, |A_{-j}|$ be a typical partial action profile in G that excludes player j. For each set of players $I \setminus \{j\}$, $j = 1, \ldots, I$, single out $|A_{-j}|$ messages $m_{j,k} \in M_{I\setminus\{j\}}$. All players except player j can send and understand these messages; they

[9] With more than two players one can imagine plain cheap-talk extensions as well as language games in which a player sends multiple messages simultaneously, each directed at different subsets of players. A sender fluent in both Basque and Spanish might want to simultaneously send public messages to two receivers, each of whom understands only one of those languages. The way we have defined language games does no allow for this possibility – it would require letting players send and interpret sets of messages. The monolingual Spanish speaker in our example would observe the message in Basque as well as the message in Spanish, but only understand one of the messages. Since the plain cheap talk extensions employed by Gerardi do not make use of this feature, we have no need for it here and stick with our formulation of language games.

[10] In the language structure $\mathcal{L}$ the language constraints of all players are commonly known – the language type spaces are singletons.

[11] In every plain cheap talk extension considered by Gerardi all message spaces are finite.

will later be used to coordinate the strategies used by players other that player j to minmax player j in the event of a deviation by player j.

The language game $\Gamma_{n+1}(G, \mathcal{L})$ has one more communication stage than the plain cheap-talk extension $\Gamma^{ext}(G)$. The additional communication stage is used to enable minmaxing players who deviate during the first n communication stages. In $\Gamma^{ext}(G)$, in any given round only designated senders and receivers can communicate with each other. In the language game all communication is public, even if not necessarily understood by all players, and every player sends a message in every communication stage.

The language game $\Gamma_{n+1}(G, \mathcal{L})$ includes histories that, for the first n communication rounds, mirror the communication constraints imposed by the plain cheap-talk extension $\Gamma^{ext}(G)$ as well as histories that do not mirror those constraints: Divide the sets of private histories H_i of each player i in the language game $\Gamma_{n+1}(G, \mathcal{L})$ into *conforming histories* H_i^C and *nonconforming histories* H_i^{NC}. A private history $h_i^t = (\varnothing, \mathbf{m}_i^1, \ldots, \mathbf{m}_i^t)$ of player i in the language game $\Gamma_{n+1}(G, \mathcal{L})$ is a conforming history if $m_{i,j}^\tau \in M_{J \cup \{j\}}$ for all players $j \in I$ and all periods $\tau \leq \min\{t, n\}$ in which player j sends a message that is received by players in the set J of players in the cheap-talk extension $\Gamma^{ext}(G)$. All other histories are non-conforming histories. In conforming histories of the language game $\Gamma_{n+1}(G, \mathcal{L})$, players j who in $\Gamma^{ext}(G)$ send messages in period t to players in the set J, send messages at time t that are understood by players in $J \cup \{j\}$ and only by those players. Messages in the language game sent by players at a time when they are not sending messages in the cheap-talk extension are ignored for the purpose of determining whether a history is conforming or not. Notice that since all players can differentiate messages in M_J from messages not in M_J for all $J \subset I$, after every history it is common knowledge whether a history is conforming or not.

In $\Gamma^{ext}(G)$, let M_{it} denote the set of messages available to player i in period t, and J_{it} the set of players to whom player i sends a message in period t. Identify every message m in the set of messages M_{it} with a distinct subset $\Phi_{it}(m)$ of messages in $M_{J_{it} \cup \{i\}}$, so that those subsets form a partition of $M_{J_{it} \cup \{i\}}$.

For every conforming private history h_i^t in the language game $\Gamma_{n+1}(G, \mathcal{L})$ let $\xi(h_i^t)$ denote player i's private history in $\Gamma^{ext}(G)$ that agrees (up to the identification of messages m in M_{it} with sets of messages $\Phi_{it}(m)$ in $M_{J_{it} \cup \{i\}}$) with h_i for all messages sent in periods $\tau \leq t$ by players designated to send messages in those periods in $\Gamma^{ext}(G)$.

Construct a strategy profile σ in $\Gamma_{n+1}(G, \mathcal{L})$ that mimics the strategy profile σ^{ext} in

$\Gamma^{ext}(G)$ as follows: For all histories $h_i^t \in H_i^C$ with $t < n$, if σ^{ext} prescribes that for history $\xi(h_i^t)$ player i in stage $t+1$ send a message that is addressed to players in the set $J_{i,t+1} \subset I$, have player i send a message from $M_{J_{i,t+1}\cup\{i\}}$ in the language game. Specifically, have every player i for whom σ^{ext} prescribes to send message m in period $t+1$ in $\Gamma^{ext}(G)$ randomize uniformly over the messages in $\Phi_{i,t+1}(m)$ in period $t+1$ of the language game. For histories $h_i^t \in H_i^{NC}$ with $t < n$, if player i sends a message in $\Gamma^{ext}(G)$, have that player randomize uniformly over all messages available to player i. For all histories h_i^t with $t < n$, all players in $\Gamma_n(G, \mathcal{L})$ who in $\Gamma^{ext}(G)$ are not sending messages in period $t+1$, randomize uniformly over all of their messages in that period.

For all histories $h_i^n \in H_i^C$, have all players randomize uniformly over all of their messages in period $n+1$. For histories $h_i^n \in H_i^{NC}$ players' period $(n+1)$–messages are determined as follows: For each player j, let $\alpha_{-j} \in \Delta\left(\prod_{j'\neq j} A_i\right)$ be a (correlated) action profile of players other than player j that minmaxes player j in the base game G. Use α_{-j}^k to denote the probability of the partial profile $a_{-j}^k \in A_{-j}$ according to α_{-j}, $k = 1, \ldots, |A_{-j}|$. Suppose that in history $h_i^n \in H_i^{NC}$, period $t \leq n$ is the first period in which $h_i^t \in H_i^{NC}$ and that player $j \neq i$ is the lowest index player for whom the set J_{jt} is nonempty, and $m_{i,j}^t \notin M_{J_{jt}\cup\{j\}}$. Denote the lowest index player other than player j by $\ell(j)$ and have player $\ell(j)$ randomize over the messages $m_{j,k} \in M_{I\setminus\{j\}}$, $k = 1, \ldots, |A_{-j}|$, assigning probability α_{-j}^k to message $m_{j,k} \in M_{I\setminus\{j\}}$ and (publicly) send the realized message. The player j, who deviated in period t, randomizes uniformly over all of her messages.

For all histories $h_i^{n+1} \in H_i^C$, if σ^{ext} prescribes that following history $\xi(h_i^n)$ in the cheap-talk extension of G player i take action $a_i \in A_i$, then have player i take action a_i following $h_i^{n+1} \in H_i^C$ in the language game (notice that for conforming histories the messages sent in period $n+1$ are simply ignored in the action phase of the language game). Consider now histories $h_i^{n+1} \in H_i^{NC}$ and use $a_{-j}^k(i)$ to denote player i's component in the partial profile $a_{-j}^k \in A_{-j}$, where $i \neq j$. For any such history, suppose that $t \leq n$ is the first period in which $h_i^t \in H_i^{NC}$ and player $j \neq i$ is the lowest index player for whom the set J_{jt} is nonempty, and $m_{i,j}^t \notin M_{J_{jt}\cup\{j\}}$. Then, have player $i \neq j$ take action $a_{-j}^k(i)$ following message $m_{j,k} \in M_{I\setminus\{j\}}$ from player $\ell(j)$ provided that player $\ell(j)$ sent one of the messages $m_{j,k} \in M_{I\setminus\{j\}}$, $k = 1, \ldots, |A_{-j}|$ in period $n+1$. If player $\ell(j)$ did not sent one of the messages $m_{j,k} \in M_{I\setminus\{j\}}$, $k = 1, \ldots, |A_{-j}|$ in period $n+1$, have player i randomize uniformly over each of her actions.

To verify that the strategy profile σ is a Bayesian Nash equilibrium of the language game $\Gamma_{n+1}(G, \mathcal{L})$, two kinds of deviations need to be considered. The first are *on-schedule-deviations*, where a player i who is meant to be sending a message in $M_{J\cup\{i\}}$ in period $t \leq n$ sends a message in $\Phi_{it}(m) \subset M_{J\cup\{i\}}$, but not one that is prescribed by the equilibrium strategy. Following such deviations messages sent in period $n+1$ are ignored. By construction, given the strategy profile σ, the message in $\Phi_{it}(m)$ sent in periods $t \leq n$ has the same consequence in the action phase of $\Gamma_{n+1}(G, \mathcal{L})$ as does the message m in the action phase $\Gamma^{ext}(G)$ according to the profile σ^{ext}. Hence, since by assumption the strategy profile σ^{ext} is a Bayesian Nash equilibrium of the plain cheap-talk extension of $\Gamma^{ext}(G)$, the assumed deviation from σ_i by player i in period t of the language game $\Gamma_{n+1}(G, \mathcal{L})$ is not profitable.

The other type of deviations, off-schedule deviations, are those where a player i who in period $t \leq n$ is meant to be sending a message in $M_{J\cup\{i\}}$ for some $J \subset I$ instead sends a message $m \notin M_{J\cup\{i\}}$. Any such deviation will be common knowledge among players. According to our specification of strategy profiles, the first player to do so (if there are multiple simultaneous deviations of this kind, the player with the lowest index among them) receives their correlated minmax payoff and therefore has no incentive to deviate in this manner. Subsequent deviations at either a messaging stage or when minmaxing a player during the action stage are off the equilibrium path and therefore can be ignore for the purpose of checking for Bayesian Nash equilibrium.

The following observation summarizes this discussion.

Proposition 4 *For every base game G with five are 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 $\mathcal{O}$.*

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] 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.

[19] Lehrer, Ehud [1991], "Internal Correlation in Repeated Games," *International Journal of Game Theory* **19**, 431–456.

[20] Lehrer, Ehud [1996], "Mediated Talk," *International Journal of Game Theory* **25**, 177–188.

[21] Lehrer, Ehud, and Sylvain Sorin [1997], "One-Shot Public Mediated Talk," *Games and Economic Behavior* **20**, 131–148.

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

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

[24] Tercieux, Olivier [2006], "p-Best response set," *Journal of Economic Theory* **131**, 45–70.

[25] Tercieux, Olivier [2006], "p-Best response set and the robustness of equilibria to incomplete information," *Games and Economic Behavior* **56**, 371–384.

[26] Voorneveld, Mark [2004], "Preparation," *Games and Economic Behavior* **48**, 403–414.

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