UTS Business School –
Department of Economics
Dr Chau Chak Wing Building
14 – 28 Ultimo Road, Ultimo NSW 2007

T: +61 2 9514 7713
economics@uts.edu.au

Department of Economics
Working Paper Series

'Pragmatic Ambiguity and Rational Miscommunication'

Toru Suzuki [1]

[1] University of Technology Sydney

No. 4 / 2021 **June 2021**

Pragmatic Ambiguity and Rational Miscommunication

Toru Suzuki[*,†]

June 1, 2021

Abstract

This paper provides a model of miscommunication in a common-interest setting. The speaker describes the state with a preexisting language to the decision-maker, whereas using a longer description is more costly. It is shown that, given any non-zero communication cost, any reasonably efficient equilibrium exhibits miscommunication caused by ambiguous descriptions whenever agents communicate across various occasions and their perceptions of occasions are imperfect but sufficiently accurate. Equilibrium miscommunication disappears when agents' perceptions of occasions are too noisy, suggesting more accurate perceptions do not always reduce miscommunication. The model also provides insight into the miscommunication that triggered a well-known aircraft crash.

Keywords: Miscommunication, common-interest communication, pragmatic ambiguity, economics and language

JEL Codes: D83

[*]University of Technology Sydney, PO Box 123, Broadway NSW 2007, Australia. Email address: toru.suzuki@uts.edu.au

[†]I am grateful to the seminar participants at the University of Arizona and Keio University for their valuable comments. I appreciate Andreas Blume and Bart Lipman for their helpful comments.

1 Introduction

Successful communication is essential for cooperative economic activities. At first glance, it seems easy to communicate perfectly when agents share a common interest since there is no incentive to lie. Nevertheless, in practice, miscommunication occasionally occurs even if agents share the same goal. In fact, miscommunication is a common cause of fatal incidences such as aircraft crashes and medical errors. The purpose of this paper is to provide a formal framework to analyze miscommunication in a common-interest setting.

There can be various types of miscommunication. The current paper focuses on miscommunication caused by an expression that can refer to different states depending on the occasion of communication, i.e., "pragmatic ambiguity." To illustrate the idea, suppose that agent 1 and 2 work for a Japanese company where ideal candidates are those who have an MD, i.e., Doctor of medicine, and can speak English. Agent 1 has already read all applications, whereas agent 2 has not. Suppose that they encounter Ken, one of the applicants. Agent 1 then says "Ken has an MD." Note that the natural interpretation of agent 1's expression depends on the occasion, i.e., the expression is pragmatically ambiguous. If the expression is uttered when they encounter Ken at a US university, it is natural to interpret that Ken has an MD (and he can speak English); otherwise, agent 1 would add "but he does not speak English" to clarify the unusual case. In contrast, if the expression is uttered when they meet Ken in a small Japanese town, it is natural to interpret that Ken has an MD (but he cannot speak English); otherwise, agent 1 would add "Ken speaks English" to emphasize his special skill. To see how the pragmatically ambiguous expression causes miscommunication, suppose that "Ken has an MD" is uttered when agents encounter him in Japan but they have already met Ken at a US university before. Then, if agent 1 says "Ken has an MD," believing that agent 2 remembers the past encounter in the US, agent 2, who forgot the past encounter, could incorrectly interpret that the candidate does not speak English. Note that this miscommunication can be a rational outcome if agent 1 wishes to communicate economically, and agent 2 usually remembers the past encounters well.

Section 2 introduces the model. To formalize pragmatic ambiguity, we need a model of language that incorporates descriptions. That is, a state is *described* by a combination of statements with preexisting meaning rather than directly referred to by an abstract message. The current paper then uses propositional logic to model the language. With propositional logic, the literal meaning of a description is determined by the preexisting meaning of each statement in the description. Then, the pragmatic meaning of a description that stems from the literal meaning is determined by the equilibrium use of the description given the occasion. Moreover, the cost of description can be defined by the number of statements in the description.

This paper then analyzes the following common-interest communication game. There is a set of elementary events, and only the speaker observes whether each event occurred. The state is then determined by which events do and do not occur. An important component of this model is "occasion of communication," which is payoff-irrelevant but can affect the distribution of states. Each player has one's perception of the occasion, which is imperfect and modeled as a noisy private signal about the occasion. A description of a state is a combination of statements about elementary events. The speaker then chooses a description given the state and her perception of the occasion. Since this paper is interested in the use of a preexisting language in common-interest communication, it is assumed that the set of feasible descriptions at a state consists of descriptions that do not contradict the state. The speaker's strategy is then defined as the probability of using a feasible description given the state and her perception of the occasion. The listener chooses an action given the speaker's description and his perception of the occasion. This paper then analyzes the perfect Bayesian equilibrium of the game. Since the set of feasible descriptions depends on the state, we focus on equilibria in which the off-equilibrium belief is consistent with the feasibility of descriptions. If the communication cost is too high, any equilibrium communication can be imperfect by the setting. Thus, it is assumed that the communication cost is small enough that it does not preclude fully precise communication.

Section 3 provides the equilibrium analysis. First, it is shown that any equilibrium is informative enough that the listener chooses the state-optimal action with some probability. Moreover, any equilibrium uses a separating strategy with respect to a state given a signal about the occasion. The most important property of equilibrium is the trade-off between simplicity-ambiguity; whenever the equilibrium description of a state given a signal is more economical than that of the same state given a different signal, the more economical description is always pragmatically ambiguous referring to completely different states depending on the signal. It is shown that an equilibrium in which the description does not depend on a signal about occasions, "occasion-free equilibrium," always exists. It is also found that an equilibrium in which the description depends on a signal, "occasion-sensitive equilibrium," exists whenever the signal is sufficiently accurate. The basic idea is that if two agents share the same perceptions of the occasion with high probability, they can utilize the situational information to economize their communication across occasions.

The main interest of this paper is miscommunication. An equilibrium exhibits rational miscommunication when the listener chooses a suboptimal action with positive probability. First, it is shown that rational miscommunication occurs if and only if the speaker uses a pragmatically ambiguous expression in equilibrium. Thus, whether rational miscommunication occurs or not in the current model boils down to the use of a pragmatically ambiguous expression in equilibrium. An equilibrium is language-comparable to another equilibrium if the use of descriptions in the latter is preserved in the former while adding some new uses. It is shown that given any non-zero communication cost, any equilibrium without miscommunication is Pareto-dominated by a language-comparable equilibrium with miscommunication when the distribution of states has non-constant modes across occasions, and agents' perceptions of occasions are imperfect but sufficiently accurate. That is, miscommunication is a natural consequence when agents try to communicate efficiently across various occasions by utilizing situational information. Moreover, it is also shown that the equilibrium miscommunication disappears if at least one agent's perception of occasions is too noisy. Thus, a higher ability to recognize occasions can increase the equilibrium miscommunication in the current model. Since a higher ability to recognize occasions reduces the probability of miscommunication once the ability reaches a certain level, the effect of a higher ability to recognize occasions on miscommunication can be non-monotonic.

Some organizations demand workers to follow a communication rule; for example, in aviation, an emergency needs to be explicitly stated rather than implicated by an incidence such as fuel exhaustion; medical practitioners often need to follow a communication format to describe patients. Section 4 analyzes the effect of such communication rules on miscommunication based on the current framework. It is shown that if a communication rule demands a more explicit expression for some state compared to the natural expression, any equilibrium that respects the rule exhibits miscommunication caused by pragmatic ambiguity. This paper then shows how this result can help us to comprehend the miscommunication that triggered a well-known aircraft crash.

Section 5 provides some discussions, and the paper is concluded by Section 6.

Related literature: This paper contributes to the growing literature on imperfect communication in common-interest settings.[1] There are three major approaches to modeling imperfect communication in common-interest settings. The oldest approach is to introduce an exogenous noise to the communication channel, e.g., Shannon (1948), Marschak and Radner (1972), which makes communication inevitably noisy. Another approach is to impose a restriction on the richness of language. For example, Cremer, Garicano, and Prat (2007) and Jäger, Metzger, and Riedel

[1]Since Crawford and Sobel (1982), there is vast literature on imperfect communication in the setting with conflict of interest. However, the logic behind imperfect communication is fundamentally different from that in the common-interest setting. There is also literature on pre-play communication. While some pre-play communication can be considered as common-interest communication, it is not the subject of the current paper.

(2011) consider models in which the set of messages is coarser than the state space. In Blume and Board (2013), an agent's set of available messages, which may or may not be coarser than the state space, is given as private information, i.e., "language type." The third approach does not restrict language but takes into account the cost of communication, e.g., Dewatripont and Tirole (2005) and Sobel (2012).[2] Since the current paper does not assume either a noisy communication channel or a limited language but a communication cost, it can be categorized into the third approach. However, since the payoff-irrelevant private information, i.e., a noisy perception of occasions, plays an essential role in imperfect communication, the current paper also shares some spirit with Blume and Board (2013), where the unobservability of a payoff-irrelevant language type makes communication noisy. A distinct feature of the current paper is that, on the contrary to the existing literature where the friction is large enough to preclude perfect communication in equilibrium, whether the efficient equilibrium exhibits imperfect communication or not depends on the environment. In fact, if an agent's perception of occasions is sufficiently noisy, no equilibrium exhibits imperfect communication in the current paper.

This paper also contributes to the literature on "Economics and Language," specifically, the economics literature that utilizes concepts from pragmatics.[3] Glazer and Rubinstein (2001) and Glazer and Rubinstein (2006) study a persuasion rule, a non-cooperative analog of Gricean pragmatics, in a partial disclosure game. Suzuki (2017) analyzes the principal-agent model with moral hazard by utilizing the concepts of "directives" and "expressives" from speech act theory. Sobel (2020) introduces the definitions of lying, deception, and damage in the standard cheap talk game based on the distinction between locution, illocution, and perlocution introduced by Austin (1975). Suzuki (2020) shows the efficient use of silence in cooperative communication exhibits the defining property of indexicals. The current paper analyzes imperfect communication by modeling the language explicitly. The explicit model of language allows us to analyze the pragmatic meaning of a description that stems from the literal meaning given an occasion. As a result, we can analyze how a specific expression used at a certain state could be misunderstood given a certain communication environment.

[2]Sobel (2012) points out that "complexity plays a role in communication, but economic models that treat conflict of interest as the driving force in strategic interaction may be paying insufficient attention to complexity... It would be valuable to study the implications of richer models of communication costs on the nature of communication."

[3]Pragmatics is a branch of linguistics that studies meaning that depends on the use of language. There is also pragmatics literature that utilizes game theory. See Benz, Jäger, and Van Rooij (2005) for a comprehensive survey.

2 Model

2.1 Basics

There are two agents, $i = 1, 2$, who have a common interest. There is a finite set of actions A, and a finite set of elementary events X.[4] The payoff from $a \in A$ depends on which elementary events do and do not occur. Agent 1 (she) can observe whether each $x \in X$ occurs or not, whereas agent 2 (he) has no information about any x. Agent 1 then describes what she observed to agent 2, who is in charge of choosing an action $a \in A$.

A state ω is founded on elementary events. Specifically, each ω is defined by the set of elementary events that occurred, denoted by $X_\omega \subset X$, and the set of elementary events that did not occur, denoted by $X_\omega^- \subset X$. Let Ω be the set of possible states. Note that since $X_\omega^- = X \backslash X_\omega$, the state ω can be determined once X_ω is specified. Moreover, since some elementary events can be mutually exclusive, $|\Omega| \leq |P(X)|$.[5]

Example 1. Consider the illustrative example in the introduction. The set of elementary events X is $\{x_1, x_2\}$ where x_1 is "Ken has an MD," and x_2 is "Ken can speak English." The set of possible states Ω is then $\{\omega_1, \omega_2, \omega_3, \omega_4\}$ where $X_{\omega_1} = \{x_1, x_2\}, X_{\omega_2} = \{x_1\}, X_{\omega_3} = \{x_2\}$, and $X_{\omega_4} = \emptyset$. That is, ω_1 is the state at which Ken has an MD and can speak English; ω_2 is the state at which Ken has an MD but does not speak English, etc.

In the literature of sender-receiver games, the state is often defined as a value of a numerical variable. The current model can also accommodate such a setting.

Example 2. Suppose $X = \{x_n\}_{n \in 1,2,,,N}$ where $x_n \in \Re$ is this year's annual profit of a company. Then, since $x_{n'}$ and $x_{n''}$ are mutually exclusive events, each profit level directly determines the state, i.e., $\Omega = \{\omega_1, \omega_2, .., \omega_N\}$ where $X_{\omega_n} = \{x_n\}$.

An important component of the current model is an **occasion** θ. An occasion of communication is determined by the background of communication such as location, season, time, surroundings, etc. Let Θ be a finite set of occasions where $|\Theta| > 1$. An occasion θ does not affect the payoff from an action a given ω, but it can affect the distribution of ω. Then, let $\pi_\theta(\omega)$ be the probability of ω given θ.

[4] An elementary event can literally be an event, e.g., "the company's annual profit is doubled," but it can be a property, e.g., "Ken has an MD."

[5] For example, "Ken worked until 5 pm" and "Ken worked until 6 pm" are mutually exclusive.

Each agent's perception of an occasion θ is assumed to be imperfect; specifically, agent i observes a noisy signal $s_i \in S_i$ about θ where S_i is finite and $|S_i| > 1$. Let $g(s_1, s_2, \theta)$ be a joint probability distribution. Assume that $supp(g) = S_1 \times S_2 \times \Theta$ and $S_1 = S_2$. Moreover, assume (s_1, s_2, θ) and ω are independent, i.e., the probability of $(s_1, s_2, \theta, \omega)$ is $g(s_1, s_2, \theta)\pi_\theta(\omega)$. In this paper, the marginal distribution of s_i and that of (s_i, θ) are simply written as $g(s_i)$ and $g(s_i, \theta)$. The distribution g reflects both environmental and human factors. For example, g may depend on the complexity of communication background; if θ is determined by a larger amount of situational information, g may have a nosier s_i. Another factor can be the agent's ability to identify the current occasion, i.e., "pragmatic competence." If an agent has higher pragmatic competence, her signal about θ can be more accurate.[6]

2.2 Language

Assume that agents communicate with a language that has no name for ω but elementary statements for each $x \in X$.[7] Formally, let ϕ_x be the elementary statement "x occurs," and let $\Phi = \{\phi_x\}_{x \in X}$ be the set of available elementary statements. Agent 1 can also use negation $\neg$ "not," disjunction $\vee$ "or," and conjunction $\wedge$ "and" to combine elementary statements. A **description** ψ is then a statement that can be produced from Φ with the help of $\{\wedge, \vee, \neg\}$. Let Ψ be the set of all descriptions, including that without any statement, denoted by $\varnothing$.

A statement ϕ_x is true at ω if $x \in X_\omega$, i.e., x occurs at ω. For any $\psi \neq \varnothing$, whether ψ is true or not at ω is determined deductively.[8] The **literal meaning** of ψ is then defined as the set of ω at which ψ is true, denoted by Ω_ψ. Intuitively, the literal meaning of ψ is "ω is in Ω_ψ." A description ψ is a **fully precise description** of ω if $\Omega_\psi = \{\omega\}$. Note that Ψ is rich in the sense that, given any $Z \subsetneq \Omega$, there exists ψ such that the literal meaning is Z.

Example 3. Consider Example 1. The description $\psi' = \phi_{x_1}$ corresponds to the English expression "Ken has an MD." Then, ϕ_{x_1} is true at both ω_1 and ω_2, i.e., $\Omega_{\psi'} = \{\omega_1, \omega_2\}$. The description

[6]There are at least two kinds of pragmatic competence; one is the ability to identify the current occasion or context, and another is the ability to speak or interpret a message according to an equilibrium strategy. In the current paper, agents are always pragmatically competent in the latter sense, but their competence can vary in the former sense.

[7]Consider Example 1. Most natural languages do not have a name or a single word that exactly refers to "Ken has an MD and speaks English." As Rubinstein (1996) points out, one of the major functions of natural language is to describe an object that has no mutually recognized name.

[8]Note that (i) $\neg\phi'$ is true iff ϕ is not true; (ii) $\phi' \vee \phi''$ is true iff at least one of them is true; (iii) $\phi' \wedge \phi''$ is true iff both ϕ' and ϕ'' are true. Once whether each ϕ in ψ is true or not at ω is known, we can determine whether ψ is true or not at ω according to (i)-(iii).

$\psi'' = \phi_{x_1} \wedge \neg\phi_{x_2}$ corresponds to the English expression "Ken has an MD and does not speak English." Note that ψ'' is true only at ω_2, i.e., $\Omega_{\psi''} = \{\omega_2\}$. Thus, ψ'' is a fully precise description of ω_2.

Describing a state and receiving a description are both costly. Specifically, a longer description is more costly for the speaker and the listener. Then, assume that, for both agents, the cost of using ψ is increasing in the number of elementary statements in ψ.[9] Moreover, for simplicity, assume that the cost of using ψ is the same for both agents. Formally, let $n(\psi)$ be the number of statements in ψ. Then, for each agent, the cost of communication with ψ is $c(\psi) = k(n(\psi))$ where $k(n)$ is strictly increasing in n and $k(0) = 0$.[10]

Comment 1. We can also consider an alternative model of language where a description is directly modeled as $m \subset \Omega$, and the cost of description is decreasing in the cardinality of m. However, such a model does not always capture the nature of description costs. To see this, consider Example 1. In the alternative approach, $m' = \{\omega_1, \omega_2\}$ and $m'' = \{\omega_1, \omega_4\}$ are just coarse descriptions of ω_1 with the same description cost. However, once we consider corresponding English expressions, they are quite different kinds; $\{\omega_1, \omega_2\}$ corresponds to "Ken has an MD," whereas $\{\omega_1, \omega_4\}$ is "Ken has an MD and speaks English, or he does not have an MD and does not speak English." That is, $\{\omega_1, \omega_2\}$ can be described concisely because of the common event x_1, whereas $\{\omega_1, \omega_4\}$ needs two fully precise descriptions since they have no common event. The current model precisely captures the difference; the simplest description with $\Omega_\psi = \{\omega_1, \omega_2\}$ is ϕ_x, whereas the simplest description with $\Omega_\psi = \{\omega_1, \omega_4\}$ is $(\phi_{x_1} \wedge \phi_{x_2}) \vee (\neg\phi_{x_1} \wedge \neg\phi_{x_2})$.[11]

2.3 Description game

This paper analyzes the following common-interest communication game. At period 1, agent 1 observes the state ω and a private signal s_1. She then chooses a description ψ. Since this paper is interested in the use of a preexisting language in common-interest communication, assume that

[9] A similar assumption can be found in Battigalli and Maggi (2002) in the context of the writing cost of contracts.

[10] The results of this paper can be preserved even if two descriptions with the same number of statements do not always have the same cost as long as a description with more statements is more costly. Furthermore, we can also consider description costs that are based not only on the number of statements but also on that of connectives.

[11] In Section 5-2, we discuss a set-theoretical model of language that captures description costs.

agent 1 only uses a description ψ that does not contradict the state ω.[12] Formally, let

$$\Psi(\omega) = \{\psi \in \Psi : \omega \in \Omega_\psi\} \cup \{\varnothing\}$$

be the set of feasible descriptions at ω. Then, agent 1's communication strategy $\sigma(\psi|\omega, s_1)$ specifies the probability that she uses $\psi \in \Psi(\omega)$ given (ω, s_1). Let $\Sigma(\omega)$ be the set of feasible communication strategies at ω. At period 2, agent 2 receives the description ψ and observes a private signal s_2. He then chooses an action $a \in A$. Formally, agent 2's decision-strategy $f(\psi, s_2)$ specifies an action $a \in A$ given (ψ, s_2). Let F be the set of decision strategies.

Turning to the payoff function, if agent 2 chooses an action a at ω, each agent receives the reward $u(a, \omega)$ where $u : A \times \Omega \to \Re_+$. As mentioned earlier, when agent 1 uses ψ, it costs $c(\psi)$ to both agents. Then, the payoff from (a, ω, ψ) is $u(a, \omega) - c(\psi)$ for both agents.[13] There is a couple of assumptions on the payoff function.

The first assumption states that each ω has a distinct optimal action.

Assumption 1. For each ω, there exists $a_\omega \in A$ such that $\{a_\omega\} = \arg\max_{a \in A} u(a, \omega)$.

If communication is highly costly, imperfect communication is an immediate consequence. Thus, the current paper focuses on the setting where the communication cost is small.

Assumption 2. For any ω, $u(a_\omega, \omega) - k(|X|) > \frac{1}{2}u(a_\omega, \omega) + \frac{1}{2}\max_{a \neq a_\omega} u(a, \omega)$.

Note that any state can be described fully precisely with $|X|$ statements. Thus, Assumption 2 states that the communication cost is small enough that the payoff from the optimal action with the cost of a fully precise description is higher than the expected payoff from 50-50 lottery over the optimal action and a suboptimal action without a description cost.

This paper employs perfect Bayesian equilibrium (PBE) to analyze the game. To define PBE,

[12] For example, the description "Ken does not have an MD" is not feasible when Ken has an MD and speaks English. This type of consistency restriction can also be found in the literature of verifiable communication, e.g., Milgrom (1981) and Grossman (1981). In this literature, communication is interpreted as (partial) disclosure, which cannot contradict the state. In the current paper, the restriction can be motivated by the complexity cost; if agents have limited cognitive capacity, it can be too costly to follow and memorize a use of ψ that is radically different from the preexisting meaning. Since there is no incentive to lie in common-interest communication, the current restriction is innocuous; while it allows us to obtain sharper results, the equilibria in the current paper still exist without the restriction.

[13] Unlike the speaker's cost of description, the listener's cost of receiving a description does not affect equilibrium behaviors. However, this paper still takes into account the listener's cost since it simplifies the efficiency comparison of equilibria.

let

$$\Psi_\sigma = \{\psi \in \Psi : \exists(\omega, s_1), \psi \in supp(\sigma(.|\omega, s_1))\}.$$

That is, Ψ_σ is the set of on-path descriptions in σ. Then, a tuple (σ^*, f^*, μ^*) is a PBE if the following conditions are satisfied.

1. For any (ψ, s_2) where $\psi \in \Psi_{\sigma^*}$, agent 2 forms his belief μ^* based on Bayes' rule to be consistent with σ^*. That is,

$$\mu^*(\omega|\psi, s_2) = \frac{\sum_\theta \sum_{s_1} \sigma^*(\psi|s_1, \omega)g(s_1, s_2, \theta)\pi_\theta(\omega)}{\sum_{\omega'} \sum_\theta \sum_{s_1} \sigma^*(\psi|s_1, \omega')g(s_1, s_2, \theta)\pi_\theta(\omega')}.$$

2. For any (ψ, s_2), agent 2's decision strategy f^* given μ^* is optimal. That is, for all $a \in A$,

$$\begin{aligned}&\sum_\omega u(f^*(\psi, s_2), \omega)\mu^*(\omega|\psi, s_2) - c(\psi)\\ \geq &\sum_\omega u(a, \omega)\mu^*(\omega|\psi, s_2) - c(\psi)\end{aligned}$$

3. For any (ω, s_1), agent 1's communication strategy σ^* is optimal given f^*. That is, for all $\psi' \in \Psi(\omega)$,

$$\begin{aligned}&\sum_\psi \sum_{s_2} \sum_\theta [u(f^*(\psi, s_2), \omega)g(s_2, \theta|s_1) - c(\psi)]\sigma^*(\psi|\omega, s_1)\\ \geq &\sum_{s_2} \sum_\theta u(f^*(\psi', s_2), \omega)g(s_2, \theta|s_1) - c(\psi')\end{aligned}$$

Recall that a description ψ can be chosen at ω only if $\psi \in \Psi(\omega)$. Thus, the current paper imposes the following off-path belief restriction. Assume that if ψ is an off-path description, agent 2 updates his belief based on Bayes' rule given $\{\omega : \psi \in \Psi(\omega)\}$ and s_2. Formally, if $\psi \notin \Psi_{\sigma^*}$, then

$$\mu^*(\omega|\psi, s_2) = \begin{cases} \frac{\sum_{s_1} \sum_\theta g(s_1,s_2,\theta)\pi_\theta(\omega)}{\sum_{\omega' \in \{\tilde{\omega}:\psi\in\Psi(\tilde{\omega})\}} \sum_{s_1} \sum_\theta g(s_1,s_2,\theta)\pi_\theta(\omega')} & \text{if } \omega \in \{\tilde{\omega} : \psi \in \Psi(\tilde{\omega})\} \\ 0 & \text{if } \omega \notin \{\tilde{\omega} : \psi \in \Psi(\tilde{\omega})\} \end{cases}.$$

Note that the above off-path belief restriction can also be founded on the following equilibrium stability argument. Suppose that, given σ, agent 1 at ω mistakenly uses an off-path description

ψ with small probability $\epsilon_\psi > 0$, which is state-independent. Then, since an off-path description ψ can be used only at a state in $\{\tilde{\omega} : \psi \in \Psi(\tilde{\omega})\}$, the belief that is consistent with the error distribution is the above off-path belief. Hence, we can consider that the current paper focuses on PBE that are stable under a small $\epsilon_\psi > 0$ for all $\psi \notin \Psi_{\sigma^*}$. Henceforth, a PBE with the above off-path belief is simply called "equilibrium."

Before starting the analysis, there are some comments on the current model.

Comment 2. In the current model, the speaker describes a state ω rather than recommends an action a directly. This setting is reasonable when an action a consists of "elementary actions," which is also costly to describe. Another rationale for the setting can be that, while it is not in the current setting, the speaker does not know exactly which action is currently available for the decision-maker.[14] In the current setting, the speaker does not describe her perception of the occasion s_1. This setting can also be natural when her perception consists of payoff-irrelevant elementary events and is costly to describe. Note that since a signal s_1 is payoff-irrelevant, it is more effective to describe ω rather than s_1.

Comment 3. The common-interest communication literature often focuses on the most efficient equilibria without imposing any restriction on the strategy space and off-equilibrium beliefs.[15] We can still preserve the basic insight of this paper even if we follow the same approach. However, this paper does not follow the approach since the most efficient equilibrium with no restriction is not always the most convincing equilibrium in the current setting; the equilibrium meaning of a description can contradict the literal meaning; for example, the equilibrium meaning of "Ken does not have an MD" can be "Ken has an MD and can speak English" in the most efficient equilibrium.[16] This fact is not paradoxical; the most efficient equilibrium with no restriction does not take into account the cognitive cost of following and memorizing a use of ψ that ignores the preexisting meaning. This paper thus analyzes the properties of reasonable equilibria under restrictions mo-

[14]Even if A is agent 2's private information, the results of this paper can be preserved.

[15]For example, Sobel (2012) and Blume and Board (2013) focus on the most efficient equilibrium. The most efficient equilibrium in the common-interest sender-receiver game has been shown to be evolutionary stable by Blume, Kim, and Sobel (1993) and perturbed message persistence by Blume (1996).

[16]When the pragmatic meaning of an expression contradicts the literal meaning, it violates a property of Gricean pragmatics called cancellability; the pragmatic meaning of an expression can be overridden naturally by adding some clarification. For example, suppose that the pragmatic meaning of "Ken has an MD," uttered in the US, is "Ken has an MD and speaks English." Then, if the speaker wants to override the pragmatic meaning, he can just add "but Ken does not speak English" to "Ken has an MD." However, the property is violated when "Ken does not have an MD" means "Ken has an MD and speaks English." If the speaker wants to cancel the pragmatic meaning by adding "Ken does not have an MD" to "Ken has an MD," the expression is non-sensical. For more details, see Huang (2007).

tivated by the use of a preexisting language rather than focusing on the most efficient equilibrium with no restriction.

3 Analysis

3.1 Properties of equilibria

An equilibrium is **effectively informative** if, for any (ω, s_1), there exists $\psi \in supp(\sigma^*(.|\omega, s_1))$ such that $f^*(\psi, s_2) = a_\omega$ for some s_2. That is, in an effectively informative equilibrium, the state-optimal action is induced at every state with strictly positive probability. An equilibrium is **conditionally separating** if whenever $\psi' \in supp(\sigma^*(.|\omega, s_1'))$ and $\omega' \neq \omega$, then $\psi' \notin supp(\sigma^*(.|\omega', s_1'))$. That is, given s_1, the same description is never used at more than one state.

To introduce the next, let

$$\Omega_\sigma(\psi, s_1) = \{\omega : \psi \in supp(\sigma(.|\omega, s_1))\}.$$

That is, $\Omega_\sigma(\psi, s_1)$ is the set of states at which σ uses ψ with positive probability given s_1. A description ψ is **pragmatically ambiguous** in σ if there exists s_1' and s_1'' such that $\Omega_\sigma(\psi, s_1') \neq \emptyset$, $\Omega_\sigma(\psi, s_1'') \neq \emptyset$, and $\Omega_\sigma(\psi, s_1') \cap \Omega_\sigma(\psi, s_1'') = \emptyset$. In other words, a description ψ is pragmatically ambiguous if ψ can refer to two mutually exclusive sets of states under different s_1.[17] An equilibrium is **ambiguity-monotonic** if whenever $\psi' \in supp(\sigma^*(.|\omega', s_1'))$, $\psi'' \in supp(\sigma^*(.|\omega', s_1''))$, and $c(\psi') > c(\psi'')$ in the equilibrium, then ψ'' is pragmatically ambiguous. That is, if an equilibrium is ambiguity-monotonic, then whenever an equilibrium description of ω' at s_1' is more economical than that of ω' at s_1'' in σ^*, the more economical description is pragmatically ambiguous.

Proposition 1. *Any equilibrium is effectively informative, conditionally separating, and ambiguity-monotonic.*

Communication games usually have various kinds of equilibria, including the one that conveys no information. In contrast, any equilibrium communication in this paper is reasonably informative

[17]Pragmatic ambiguity is different from lexical ambiguity, which is caused by a word that has two different meanings, e.g., "Ken is near the *bank*," and syntactic ambiguity, which is caused by an ambiguous sentential structure, e.g., "the chicken is ready to eat." The current model of language, which is based on a formal language, precludes both kinds of ambiguity. Pragmatic ambiguity is also different from vagueness as a formal semantic concept. If ψ is pragmatically ambiguous, the meaning of ψ is completely different across some occasions, i.e., $\Omega_\sigma(\psi, s_1') \cap \Omega_\sigma(\psi, s_1'') = \emptyset$. In contrast, if ψ is vague, there must have borderline cases, that is, the meaning of ψ depends on s_1, but there are some overlaps across s_1, i.e., $\Omega_\sigma(\psi, s_1') \neq \Omega_\sigma(\psi, s_1'')$ but $\Omega_\sigma(\psi, s_1') \cap \Omega_\sigma(\psi, s_1'') \neq \emptyset$.

because of the restriction on feasible descriptions and the consistent off-path belief together with a small communication cost. Moreover, there is a trade-off between simplicity and ambiguity in equilibrium; whenever the equilibrium description of ω is simpler at one occasion than another, it must be pragmatically ambiguous. Note that the trade-off here is not between simplicity and coarseness, i.e., a partial pooling with respect to ω, since the conditional separation property precludes coarseness in equilibrium.

3.2 Occassion-free and occasion-sensitive equilibria

It is useful to categorize equilibria based on occasion sensitivity. A communication strategy σ is **occasion-free** if $\sigma(\psi|\omega, s_1)$ is constant in s_1 given any ω. An equilibrium is occasion-free if the equilibrium communication strategy is occasion-free. In contrast, a communication strategy is **occasion-sensitive** if $\sigma(\psi|\omega, s_1)$ is not occasion-free. An equilibrium is occasion-sensitive if the equilibrium communication strategy is occasion-sensitive.

To state the next result, let

$$\rho(s_i) = \sum_{\theta} g(s_j = s_i, \theta|s_i).$$

That is, $\rho(s_i)$ is the probability that agent j observes $s_j = s_i$ when agent i observes s_i. Let $\rho = \min_{i\in\{1,2\}} \min_{s_i\in S_i} \rho(s_i)$.

Proposition 2. *An occasion-free equilibrium exists. Moreover, there exists an occasion-sensitive equilibrium if $\rho \in (0, 1)$ is sufficiently large.*

Note that since agent 2 interprets off-path descriptions based on the literal meaning, his optimal action can depend on s_2. Thus, to prove the first part, we construct an occasion-free strategy so that the most economical descriptions whose literal meaning can induce the optimal action at ω under some s_1 are not left out as off-path descriptions. For the second part, we construct an occasion-sensitive strategy from the occasion-free strategy constructed earlier so that the most economical description in the strategy is pragmatically ambiguous. The pragmatically ambiguous description can mislead agent 2 when agents do not share their perception of the occasion. However, if ρ is sufficiently high, agents share their perception with high probability, and the ambiguous description is reliable enough to use in equilibrium.

For further insight, consider the illustrative example in the introduction (and Example 1). Let $\Theta = \{\theta_1, \theta_2\}$ where θ_1 is the occasion where agents met Ken in the US, whereas θ_2 is the occasion where agents met only in Japan. Assume that it is likely that Ken can speak English if agents met him in the US, whereas it is much less likely if they met him only in Japan, i.e., $\Pr(x_2|\theta_1) > 0.5 > \Pr(x_2|\theta_2)$. Since people do not have an MD in general, assume $\Pr(x_1|\theta) < 0.5$ for both θ. Let $S_i = \{s_i', s_i''\}$ where s_i' and s_i'' are signals that indicate that θ is likely to be θ_1 and θ_2 respectively.

The left-hand side in Figure 1 is an occasion-free strategy in which agent 1 only mentions x in X_ω at ω. The right-hand side in Figure 1 is the occasion-sensitive strategy in which agent 1 only mentions "unexpected information" given s_1. For example, agent 1 does not mention that "Ken speaks English" if agents met Ken in the US. In contrast, if Ken does not speak English even if agents met in the US, agent 1 mentions the unexpected information "Ken does not speak English." If agents met Ken in the US, agent 1 mentions nothing at ω_3 since most people in the US speak English but do not have an MD.[18]

While the occasion-sensitive strategy in Figure 1 seems more natural than the occasion-free strategy in Figure 1, both can be equilibrium strategies. To see this, first, consider the occasion-free strategy. Suppose agent 1 is at ω_1. Clearly, agent 1 has no reason to use other on-path descriptions as it induces a suboptimal action. Moreover, she also has no off-path description that is feasible at ω_1. At ω_2 and ω_3, the equilibrium description induces the optimal action with one statement. Then, since all off-path descriptions are at least as costly as them, agent 1 has no incentive to deviate from the strategy. At ω_4, there is no need to check potential deviations since agent 1 gets the highest possible payoff from the costless description $\varnothing$.

Turning to the occasion-sensitive strategy in Figure 1, when agent 1 uses the pragmatically ambiguous description ϕ_{x_1} at (ω_1, s_1'), it could be misinterpreted as ϕ_{x_1} at (ω_2, s_1''). However, if ρ is high, the probability of misinterpretation is low since the chance that agents share their perceptions is high. Then, if ρ is sufficiently high, agent 1 at (ω_1, s_1') has no incentive to use the fully precise description $\phi_{x_1} \wedge \phi_{x_2}$. Agent 1 also has no incentive to use the less costly description ϕ_{x_1} at (ω_1, s_1'') if ρ is high. To see this, note that if agent 1 uses ϕ_{x_1} and ρ is high, agent 2 who observes $s_2 = s_1''$ chooses the suboptimal action a_{ω_2}, misinterpreting the state is ω_2. Thus, if ρ is sufficiently high, agent 1 prefers to use $\phi_{x_1} \wedge \phi_{x_2}$ to ϕ_{x_1} at (ω_1, s_1''). By similar reasoning, agent 1 has no incentive to deviate from the use of another pragmatically ambiguous description $\varnothing$ if ρ is sufficiently high. Finally, note that the off-path description $\neg\phi_{x_1}$ is feasible only at ω_3 or ω_4. Then, since $\neg\phi_{x_1}$ is at

[18] The occasion-sensitive strategy in Figure 1 is consistent with conversational maxims in Grice (1975). The most relevant maxim is that of quantity "Make your contribution as informative as required." For example, consider ω_1. If agents met Ken in the US, it is natural to assume Ken can speak English. Then, since adding "Ken can speak English" is more than required, one can simply say "Ken has an MD."

ω	X_ω	$supp(\sigma(.\vert\omega,s_1'))$	$supp(\sigma(.\vert\omega,s_1''))$	ω	X_ω	$supp(\sigma(.\vert\omega,s_1'))$	$supp(\sigma(.\vert\omega,s_1''))$
ω_1	$\{x_1,x_2\}$	$\phi_{x_1}\wedge\phi_{x_2}$	$\phi_{x_1}\wedge\phi_{x_2}$	ω_1	$\{x_1,x_2\}$	ϕ_{x_1}	$\phi_{x_1}\wedge\phi_{x_2}$
ω_2	$\{x_1\}$	ϕ_{x_1}	ϕ_{x_1}	ω_2	$\{x_1\}$	$\phi_{x_1}\wedge\neg\phi_{x_2}$	ϕ_{x_1}
ω_3	$\{x_2\}$	ϕ_{x_2}	ϕ_{x_2}	ω_3	$\{x_2\}$	$\varnothing$	ϕ_{x_2}
ω_4	$\emptyset$	$\varnothing$	$\varnothing$	ω_4	$\emptyset$	$\neg\phi_{x_2}$	$\varnothing$

Figure 1: Occasion-free strategy (left) and occasion-sensitive strategy (right)

least as costly as any equilibrium description at ω_3 or ω_4, agent 1 has no incentive to use $\neg\phi_{x_1}$.

3.3 Rational miscommunication

An equilibrium exhibits **rational miscommunication** if there exists (ω, s_1, s_2) such that $f^*(\psi, s_2) \neq a_\omega$ for some $\psi \in supp(\sigma^*(.|\omega, s_1))$. That is, an equilibrium exhibits rational miscommunication if an equilibrium description induces a suboptimal action with positive probability at some state. The following lemma states that rational miscommunication is solely attributed to pragmatic ambiguity in the current model.

Lemma 1. *An equilibrium exhibits rational miscommunication if and only if the equilibrium communication strategy uses a pragmatically ambiguous description at some* (ω, s_1).

In the occasion-sensitive equilibrium strategy in Figure 1, ϕ_{x_1} and $\varnothing$ are both pragmatically ambiguous descriptions. If agent 1 says "Ken has an MD" at (ω_1, s_1'), agent 2 incorrectly interprets the state is ω_2 if $s_2 \neq s_1'$ and ρ is large. Similarly, when agent 1 mentions nothing at (ω_3, s_1'), agent 2 misinterprets the state is ω_4 if $s_2 \neq s_1'$ and ρ is large.

Note that an occasion-free equilibrium does not use any pragmatically ambiguous description. Thus, from Lemma 1 and Proposition 2, an equilibrium without rational miscommunication always exists. However, an equilibrium without rational miscommunication is not necessarily the most reasonable equilibrium in the current paper. In fact, it will be shown that any reasonably efficient equilibrium exhibits rational miscommunication in certain environments.

To formally state the result, an environment (g, π) has **various modes** if there exists s_1' and s_1'' such that

$$\arg\max_\omega \sum_\theta \sum_{s_2} \pi_\theta(\omega) g(\theta, s_2|s_1') \cap \arg\max_\omega \sum_\theta \sum_{s_2} \pi_\theta(\omega) g(\theta, s_2|s_1'') = \emptyset.$$

In short, an environment has various modes if the most common state conditional on s_1 is not constant in s_1. In our example in Section 3.2, the most common state is ω_3, i.e., Ken does not have an MD and speaks English, if $\theta = \theta_1$, whereas it is ω_4, i.e., Ken does not have an MD and does not speak English," if $\theta = \theta_2$. Thus, if s_1 is sufficiently accurate, the environment has various modes.

Given the combinatorial nature of language, it is not obvious whether a natural use of language needs to be the most efficient one. For example, suppose ψ' refers to ω' under some s_1 but never refers to ω'' in an equilibrium, whereas ψ' refers to ω'' under some s_1 but never refers to ω' in the most efficient equilibrium. Then, since the use of ψ' is completely different across equilibria, people could be stuck to the less efficient use of ψ'.[19] In this sense, such equilibria are not "language-comparable." Thus, instead of focusing on the most efficient equilibrium, the current paper states the main result on equilibria that are not Pareto-dominated by any "language-comparable" equilibrium. Formally, an equilibrium with σ' is **language-comparable** to another equilibrium with σ if (i) $\Psi_\sigma \subset \Psi_{\sigma'}$; (ii) for any s_1 and $\psi \in \Psi_\sigma$, there exists s_1' such that $\Omega_\sigma(\psi, s_1) = \Omega_{\sigma'}(\psi, s_1')$. That is, σ' uses all on-path ψ in σ, and whenever ψ refers to ω under some s_1 in σ, ψ refers to ω under some s_1 in σ'.

Proposition 3. *Suppose (g, π) has various modes. If ρ is sufficiently high, any equilibrium that does not exhibit rational miscommunication is Pareto-dominated by an equilibrium that is language-comparable and exhibits rational miscommunication.*

Proposition 3 suggests that if agents communicate across various environments and their perceptions of occasions are imperfect but sufficiently accurate, any reasonably efficient communication exhibits rational miscommunication.

Corollary 1. *Suppose (g, π) has various modes. If ρ is sufficiently high, the most efficient equilibria exhibit rational miscommunication.*

The basic idea behind the proof of Proposition 3 is as follows. First, it is shown that, given any occasion-free equilibrium, we can construct a language-comparable occasion-sensitive equilibrium with miscommunication if the conditions in Proposition 3 are satisfied. The occasion-sensitive strategy is constructed so that the most economical description in the occasion-free strategy is used as the description of the most common state given s_1. Clearly, the constructed strategy is

[19] In principle, the use of language can evolve to be the most efficient one in the long run. However, when the environment changes over time, there may not be enough time to be fully adjusted. Given the enormous number of possible sentences and occasions in reality, it is not obvious whether the current pragmatics of, say, English is fully efficient.

occasion-sensitive whenever the environment has various modes. Moreover, the most economical description in the strategy is pragmatically ambiguous. Then, it is shown that the constructed strategy can be supported in equilibrium when ρ is sufficiently high. In the second step, it is shown that the constructed occasion-sensitive equilibrium is more efficient than the original occasion-free equilibrium. Note that since agents use the most economical description for the most common state given a signal, and the communication cost at other states is at least as low as in the occasion-free strategy, the expected communication cost of the constructed strategy is strictly lower than that of the occasion-free strategy. In contrast, since the most economical description is pragmatically ambiguous in the constructed strategy, the equilibrium exhibits rational miscommunication from Lemma 1. However, it can be shown that if ρ is sufficiently low, the gain from the saved expected communication cost surpasses the expected loss from miscommunication. Finally, an equilibrium that does not exhibit rational miscommunication can be occasion-sensitive. However, it can be shown that any occasion-sensitive equilibrium without miscommunication has a language-comparable occasion-free equilibrium that is payoff-equivalent to the original occasion-sensitive equilibrium. Then, we can apply the earlier argument.

In the current paper, whether a reasonable equilibrium exhibits miscommunication or not depends on the environment.

Fact 1. *If π_θ is constant in θ, then the most efficient equilibria do not exhibit rational miscommunication.*

If the distribution of ω does not depend on θ, there is no way to save the communication cost by utilizing a pragmatically ambiguous description. Then, since an ambiguous description creates miscommunication with positive probability, an equilibrium that uses a pragmatically ambiguous description is Pareto-dominated by an occasion-free equilibrium.

In the existing literature, larger communication friction such as a coarser message space can make the equilibrium communication noisier. On the contrary, a highly noisy perception of occasions eliminates the equilibrium miscommunication in the current paper.

Proposition 4. *If $g(s_1, s_2, \theta) = g(s_j, \theta)g(s_i)$ for some i, then no equilibrium exhibits rational miscommunication.*

To see the idea, first, note that if agent 2's perception is pure noise, his belief about ω does not depend on s_2. Thus, whenever an ambiguous description induces the state-optimal action at one state, the description always induces the sub-optimal action at another state. Then, from

Assumption 2, agent 1 prefers to use a fully precise description that ensures the optimal action. Second, if agent 1's perception is pure noise, agent 2's belief can depend on s_2. However, since agent 2 cannot infer anything about s_1 from s_2, whether an ambiguous description induces the state-optimal action or not depends solely on s_2. Then, whenever an ambiguous description induces the state-optimal action at more than one state, the probability of inducing the state-optimal action has to be low at one of the other states. Agent 1 then prefers to use a fully precise description to induce the state-optimal action at the state.

Proposition 4 suggests that the effect of a higher ability to recognize occasions on miscommunication can be non-monotonic. If both agents have poor ability to recognize occasions, there is no equilibrium communication as they communicate precisely. However, from Proposition 3, when their ability to recognize occasions reaches a certain level, the efficient equilibrium exhibits rational miscommunication if the environment has various modes. When agents have a higher ability, the probability of miscommunication in the efficient equilibrium decreases as it reduces the misinterpretation of ambiguous descriptions.

4 Miscommunication and communication rule

In the last section, it is shown that any reasonably efficient equilibrium can exhibit miscommunication caused by pragmatic ambiguity when agents communicate economically across various occasions. In this section, it is shown that the use of a communication rule in organizations can also create miscommunication caused by pragmatic ambiguity.

In organizations that demand precise communication, professional communication is often regulated by some rules. One of the typical rules is to use a more explicit expression than a natural expression when describing a certain situation. For example, in aviation, an emergency needs to be explicitly stated rather than implicated by a stated incidence.[20] This is contrary to a natural expression with which stating "fuel is running out" is enough to convey the sense of emergency. In hospitals, medical practitioners often need to describe patients based on a format such as SBAR.[21] Such a rule can also make a description more explicit than a natural expression; when two medical professionals communicate casually, some diagnoses could naturally indicate standard treatments unless emphasized to be unnecessary.

[20]See Estival, Farris, and Molesworth (2016) for more details about aviation English.

[21]SBAR is a communication format that is designed to help medical practitioners to formulate and comprehend the description of patients.

To investigate how such a restriction affects equilibrium communication, suppose that a communication rule is introduced. Let $\hat{\theta} \in \Theta$ be the occasion at which agents are expected to follow the rule, and let $\hat{s}_1 \in S_1$ be the signal that indicates the occasion is likely to be $\hat{\theta}$. Then, given a communication strategy σ, $\hat{\sigma}$ is a **rule-adjusted strategy** of σ if (i) there exists (ω', s_1') such that $n(\psi') < n(\psi'')$ where $\psi' \in supp(\sigma(.|\omega', s_1'))$ and $\psi'' \in supp(\hat{\sigma}(.|\omega', \hat{s}_1))$; (ii) given any ω, $\hat{\sigma}(\omega, s_1) = \sigma(\omega, s_1)$ for all $s_1 \neq \hat{s}_1$. Note that (i) captures that there is a state at which the description that respects a rule is more explicit than the original description in σ; (ii) states that the use of language in σ is preserved in $\hat{\sigma}$ at any state as long as $s_1 \neq \hat{s}_1$.

If an organization can enforce agents to follow any communication rule, we can easily eliminate miscommunication by prohibiting the use of pragmatically ambiguous descriptions. However, since it is costly to monitor and regulate the use of language in practice, the current paper focuses on self-enforceable communication rules. The next result states that if an initial equilibrium strategy is adjusted to another equilibrium strategy that respects a typical communication rule, the adjusted equilibrium exhibits rational miscommunication.

Proposition 5. *Given an equilibrium strategy σ, any equilibrium with a rule-adjusted strategy of σ exhibits rational miscommunication.*

Proposition 5 is almost immediate from Proposition 1 and Lemma 1.

Proof. Suppose $\hat{\sigma}$ is an equilibrium rule-adjusted strategy of σ. Since $\hat{\sigma}$ is a rule adjusted strategy of σ, there exists (ω', s_1') such that $n(\psi') < n(\psi'')$ where $\psi' \in supp(\sigma(.|\omega', s_1'))$ and $\psi'' \in supp(\hat{\sigma}(.|\omega', \hat{s}_1))$. Thus, $c(\psi') < c(\psi'')$. Since any equilibrium is ambiguity monotonic from Proposition 1, ψ' must be pragmatically ambiguous in $\hat{\sigma}$. Then, from Lemma 1, the equilibrium with $\hat{\sigma}$ must exhibit rational miscommunication. $\square$

To provide further insight into Proposition 5, consider aviation communication between a pilot and an air-traffic controller (ATC). For simplicity, suppose there are only two possible causes of emergency in flights: fuel exhaustion or engine failure. Then, let $X = \{em, fr, ef\}$ where em is "Emergency," fr is "Fuel is running out," and ef is "Engine failure." For the probability distribution of events, assume that ef and fr are rare events, but when either ef or fr occurs, it triggers em with a high probability. Suppose that $\Theta = \{\theta', \theta''\}$ and $S_i = \{s_i', s_i''\}$ where s_i' and s_i'' are noisy signals that indicate the occasion is most likely to be θ' and θ'' respectively. Moreover, for simplicity, assume that π_θ is constant in θ.

Figure 2 illustrates "conversational English." When fuel is running out, it is natural to infer

ω	X_ω	$supp(\sigma(.\|\omega, s_1))$
ω_1	$X_{\omega_1} = \{em, fr, ef\}$	$\phi_{fr} \wedge \phi_{ef}$
ω_2	$X_{\omega_2} = \{em, ef\}$	ϕ_{ef}
ω_3	$X_{\omega_3} = \{em, fr\}$	ϕ_{fr}
ω_4	$X_{\omega_4} = \{fr, ef\}$	$\neg\phi_{em} \wedge \phi_{fr} \wedge \phi_{ef}$
ω_5	$X_{\omega_5} = \{ef\}$	$\neg\phi_{em} \wedge \phi_{ef}$
ω_6	$X_{\omega_6} = \{fr\}$	$\neg\phi_{em} \wedge \phi_{fr}$
ω_7	$X_{\omega_7} = \emptyset$	$\varnothing$

Figure 2: Conversational English

ω	X_ω	$supp(\hat{\sigma}(.\|\omega, s_1'))$	$supp(\hat{\sigma}(.\|\omega, s_1''))$
ω_1	$X_{\omega_1} = \{em, fr, ef\}$	$\phi_{fr} \wedge \phi_{ef}$	$\phi_{em} \wedge \phi_{fr} \wedge \phi_{ef}$
ω_2	$X_{\omega_2} = \{em, ef\}$	ϕ_{ef}	$\phi_{em} \wedge \phi_{ef}$
ω_3	$X_{\omega_3} = \{em, fr\}$	ϕ_{fr}	$\phi_{em} \wedge \phi_{fr}$
ω_4	$X_{\omega_4} = \{fr, ef\}$	$\neg\phi_{em} \wedge \phi_{fr} \wedge \phi_{ef}$	$\phi_{fr} \wedge \phi_{ef}$
ω_5	$X_{\omega_5} = \{ef\}$	$\neg\phi_{em} \wedge \phi_{ef}$	ϕ_{ef}
ω_6	$X_{\omega_6} = \{fr\}$	$\neg\phi_{em} \wedge \phi_{fr}$	ϕ_{fr}
ω_7	$X_{\omega_7} = \emptyset$	$\varnothing$	$\varnothing$

Figure 3: A rule adjusted strategy

that the situation is an emergency. Thus, a speaker simply says "fuel is running out" when fuel is running out and in an emergency. In contrast, when fuel is running out, but somehow it is not an emergency (maybe because the airport is close enough), a speaker says "fuel is running out but not an emergency" to clarify the unusual situation.[22] Moreover, a speaker does not say "fuel is not running out" since if it is not mentioned, it is natural to assume it's not.[23] If Assumption 1 and 2 are satisfied, there is an occasion-free equilibrium with conversational English.

Now, suppose θ' is the occasion at which the pilot is expected to use conversational English whereas θ'' is the occasion at which the pilot is expected to mention "emergency" explicitly.[24] Then, consider the rule adjusted strategy $\hat{\sigma}$ in Figure 3. Agent 1 uses conversational English if $s_1 = s_1'$, whereas if $s_1 = s_1''$, agent 1 uses descriptions that ϕ_{em} "emergency" is added to conversational expressions at ω_1, ω_2, and ω_3.

[22]This pragmatic inference is consistent with Gricean maxims.

[23]In reality, there are many possible events that can cause an emergency. If a speaker needs to mention "x is not happening" for all events that are not happening, the speaker needs to use impractically long expressions.

[24]One might wonder that if the pilot speaks from an aircraft to an ATC at an airport, the pilot should not be expected to use plain English. However, according to Estival, Farris, and Molesworth (2016), aviation communication allows plain English for flexibility.

The rule adjusted strategy in Figure 3 can be an equilibrium strategy if Assumption 1 and 2 are satisfied and $\rho \in (0, 1)$ is sufficiently high. Moreover, the equilibrium exhibits rational miscommunication; in fact, the miscommunication at ω_3 captures the one that triggered the crash of Avianca flight 054 in 1990.[25] If the pilot's perception of θ was s_1', whereas the ATC's perception of θ was s_2'', the pilot would report "we are running out of fuel," indicating ω_3, whereas the ATC would believe the state is ω_6, misinterpreting the statement. The model suggests that similar miscommunication could happen when a pilot reports "engine failures" without adding "emergency," i.e., at ω_2. Moreover, there can be a different type of miscommunication that is less harmful; the pilot reports "fuel is running out" at ω_6 according to s_1'', but the ATC interprets the report as emergency according to s_2', creating overreaction.

While the above observation is based on the specific rule-adjusted strategy, the main insight can be preserved under any rule-adjusted strategy that uses "emergency" explicitly.

Fact 2. *Let σ' be conversational English in Figure 2. Any equilibrium with a rule-adjusted strategy where ϕ_{em} is added to the conversational expression at $(\omega_1, s_1''), (\omega_2, s_1'')$ and (ω_3, s_1'') exhibits rational miscommunication.*

Comment 4. In the above analysis, we do not see the benefit of a communication rule but the pitfall. This is simply because the current paper focuses on miscommunication caused by pragmatic ambiguity. In practice, there is a benefit of using explicit expressions since people could fail to infer the equilibrium meaning of implicit expressions.

5 Discussion

5.1 Organizational codes

Since miscommunication between workers can create a significant negative externality to an organization, it is important for organizations to manage miscommunication. In the current paper, rational miscommunication is caused by the use of pragmatically ambiguous expressions. One way to eliminate pragmatic ambiguity is to create an artificial language that can refer to some states directly, i.e., organizational codes. Specifically, if an organization creates a code ξ_ω for ω, which directly refers to ω, agents do not need to describe ω according to the equilibrium communication

[25] In this incident, the pilot reported "we are running out of fuel," indicating an emergency, whereas the ATC, who did not interpret the report as an emergency, did not provide a quick instruction, resulting in the crash. For more details, see Estival, Farris, and Molesworth (2016).

strategy σ. Then, since each code is designated to a specific state, there is no room for pragmatic ambiguity.

However, the use of organizational codes has some limitations. First, as Arrow (1974) points out, it can be quite costly to learn and use organizational codes.[26] To illustrate the cost of using organizational codes, note that n mutually exclusive events can generate 2^n states; for example, if there are ten elementary events, they can generate more than one thousand states. Given our cognitive capacity, it would be more practical to describe states by combining some of ten statements than handling one thousand codes. Another potential issue of complex codes is self-enforceability; if there are many states, codes could consist of a long string of numbers or/and alphabets. Once the string gets too long, and the description cost of an English expression becomes cheaper than that of the string, she might not have an incentive to keep using codes. Thus, unless the environment is simple or codes are used only for a small number of recurring states, the net benefit of a communication rule that is built on a natural language such as aviation English can be larger than that of organizational codes.

5.2 A set-theoretical model of description

As discussed in Comment 1, it is not so effective to model a language based on subsets of Ω if we consider description costs. Since descriptions are based on elementary events, the set-theoretical model that captures description costs needs to be based on subsets of elementary events. Thus, suppose that a description is modeled as a pair (Z^+, Z^-) where Z^+ is the set of x such that the description mentions x occurs whereas Z^- is the set of x such that the description mentions x does not occur. For example, the description $(\{x', x''\}, \{x'''\})$ corresponds to the English expression "x' and x'' happen, and x''' does not happen." With this model, $\{\omega_1, \omega_2\}$ in Example 1 is simply described by $(\{x_1\}, \emptyset)$, i.e., "Ken has an MD."

However, this model is still not rich enough to describe $\{\omega_1, \omega_4\}$ in Example 1. As mentioned in Comment 1, since there is no common event between ω_1 and ω_4, the description of $\{\omega_1, \omega_4\}$ has to be a disjunction of the description of ω_1 and ω_4, i.e., "it is either Ken has an MD and speaks English or he has no MD and does not speak English." Clearly, (Z^+, Z^-) cannot accommodate such a disjunctive description. A fully rich set-theoretical model of descriptions is then $\{(Z^+_n, Z^-_n)\}_{n=1,2,..,N}$, which is read as "$(Z^+_1, Z^-_1)$ or (Z^+_2, Z^-_2) or ... or (Z^+_N, Z^-_N)." For example,

[26] Cremer, Garicano, and Prat (2007) provide a formal analysis of the optimal organizational codes by taking into account the cost of using organizational codes.

$\{\omega_1, \omega_4\}$ is described by $\{(\{x_1, x_2\}, \emptyset), (\emptyset, \{x_1, x_2\})\}$ with the extended model. The number of statements in $\{(Z_n^+, Z_n^-)\}_{n=1,2,..,N}$ is then given by $\sum_{n=1,2,..,N} |Z_n^+| + |Z_n^-|$.

6 Conclusion

This paper provided a model of equilibrium miscommunication in the common-interest setting. A language is formally modeled to define the literal meaning and the cost of descriptions. The equilibrium use of a description then determines the pragmatic meaning of a description that stems from the literal meaning. It is shown that given a non-zero communication cost, any reasonably efficient equilibrium exhibits miscommunication caused by pragmatic ambiguity when agents communicate across various occasions and their perceptions of occasions are imperfect but sufficiently accurate. That is, miscommunication is a natural consequence when agents try to communicate efficiently across various occasions by utilizing situational information. On the contrary to the existing literature in which larger friction enhances imperfect communication, the equilibrium miscommunication disappears if at least one agent's perception of occasions is too noisy. That is, having a higher ability to recognize occasions can increase the chance of miscommunication in the current paper. Since a higher ability to recognize occasions reduces the probability of miscommunication once the ability reaches a certain level, the effect of a higher ability to recognize occasions on miscommunication can be non-monotonic.

The current framework can also analyze how a communication rule in organizations can create rational miscommunication. It is shown that if an organization introduces a communication rule that demands a more explicit expression than a natural expression at some state, any equilibrium that respects the rule exhibits rational miscommunication. It is then demonstrated how this framework can explain the miscommunication that triggered a well-known aircraft crash. One way to prevent rational miscommunication is to use organizational codes. However, since learning and using organizational codes are also costly, organizations need to balance the expected cost of rational miscommunication and that of organizational codes when designing the optimal communication rule.

7 Appendix

7.1 Proof of Proposition 1

To prove Proposition 1, I establish the following claim.

Claim. If $\psi' \in supp(\sigma^*(.|\omega', s_1'))$, then ψ' induces $a_{\omega'}$ with positive probability.

Suppose $\psi' \in supp(\sigma^*(.|\omega', s_1'))$ never induces $a_{\omega'}$. Then, agent 1's payoff at (ω', s_1') is at most $\max_{a \neq a_{\omega'}} u(a, \omega')$. Suppose agent 1 uses the following description of ω',

$$\tilde{\psi} = \bigwedge_{x \in X_{\omega'}} \phi_x \wedge \bigwedge_{x \in X \setminus X_{\omega'}} \neg \phi_x.$$

Note that since $\Omega_{\tilde{\psi}} = \{\omega'\}$, $\tilde{\psi}$ is feasible only at ω'. Hence, $\tilde{\psi}$ always induces $a_{\omega'}$. That is, agent 1's payoff from $\tilde{\psi}$ is $u(a_{\omega'}, \omega') - c(\tilde{\psi})$. Note that $c(\tilde{\psi}) = k(|X|)$. Then, $u(a_{\omega'}, \omega') - c(\tilde{\psi}) > \max_{a \neq a_{\omega'}} u(a, \omega')$ from Assumption 2. That is, agent 1 has an incentive to use $\tilde{\psi}$ at (ω', s_1'), a contradiction.

From Claim, there exists s_2 such that $f^*(\psi, s_2) = a_\omega$ for any $\psi \in supp(\sigma^*(.|\omega, s_1))$. Then, any equilibrium is effectively informative. For the second part, suppose $\psi' \in \bigcap_{\omega \in O} supp(\sigma^*(.|\omega, s_1'))$ for some $O \subset \Omega$ where $|O| > 1$. From Claim, ψ' induces a_ω at $\omega \in O$ with positive probability. Hence, for each $\omega \in O$, there exists a non-empty set

$$S_\omega^{\psi'} = \{s_2 \in S_2 : f_{\sigma^*}(\psi', s_2) = a_\omega\}.$$

Then, the probability of inducing the optimal action $a_{\omega'}$ at $\omega' \in O$ with ψ' conditional on s_1' is $\sum_\theta \sum_{s_2 \in S_{\omega'}^{\psi'}} g(s_2, \theta | s_1')$. From Assumption 2, if $\sum_\theta \sum_{s_2 \in S_{\omega'}^{\psi'}} g(s_2, \theta | s_1') \leq 0.5$, the agent 1 prefers to use a fully precise description of ω', which induces $a_{\omega'}$ for sure. Thus, we must have $\sum_\theta \sum_{s_2 \in S_{\omega'}^{\psi'}} g(s_2, \theta | s_1') > 0.5$. But then, since $S_2 \setminus S_{\omega'}^{\psi'} \supset S_{\omega''}^{\psi'}$, the probability of inducing the optimal action $a_{\omega''}$ at $\omega'' \in O$ with ψ' conditional on s_1', i.e., $\sum_\theta \sum_{s_2 \in S_{\omega''}^{\psi'}} g(s_2, \theta | s_1')$, has to be at most 0.5. Then, from Assumption 2, agent 1 has an incentive to use a fully precise description of ω'' at (ω'', s_1'), a contradiction.

Finally, I show that any equilibrium is ambiguity monotonic. Suppose $\psi' \in supp(\sigma^*(.|\omega', s_1'))$, $\psi'' \in supp(\sigma^*(.|\omega', s_1''))$, and $n(\psi') > n(\psi'')$ but ψ'' is not ambiguous. From the equilibrium condition,

$$\sum_{\theta}\sum_{s_2} u(f(\psi', s_2), \omega')g(s_2, \theta|s_1') - c(\psi')$$
$$\geq \sum_{\theta}\sum_{s_2} u(f(\psi'', s_2), \omega')g(s_2, \theta|s_1') - c(\psi'').$$

Since $n(\psi') > n(\psi'')$,

$$\sum_{\theta}\sum_{s_2} u(f(\psi', s_2), \omega')g(s_2, \theta|s_1') > \sum_{\theta}\sum_{s_2} u(f(\psi'', s_2), \omega')g(s_2, \theta|s_1') \tag{1}$$

As I showed earlier, any equilibrium is conditionally separating. Thus, $|\Omega_\sigma(\psi'', s_1)| = 1$ for any s_1. Thus, if ψ'' is not ambiguous, $\Omega_\sigma(\psi'', s_1) = \{\omega'\}$ for all s_1. Then, $\mu(\omega'|\psi'', s_2) = 1$ for all s_2 and thus $f(\psi'', s_2) = a_{\omega'}$ for all s_2. It follows that

$$\sum_{\theta}\sum_{s_2} u(f(\psi'', s_2), \omega')g(s_2, \theta|s_1') = u(a_{\omega'}, \omega').$$

But then, since $u(a_{\omega'}, \omega')$ is the highest possible payoff at ω', Inequality (1) cannot be satisfied, a contradiction.

7.2 Proof of Proposition 2

To prove the existence of an occasion-free equilibrium, let

$$\Psi_\pi(\omega|s_2) = \left\{\psi \in \Psi(\omega) : a_\omega \in \arg\max_{a\in A} \frac{\sum_{s_1}\sum_\theta g(s_1, s_2, \theta)\pi_\theta(\omega)}{\sum_{\omega'\in\Omega_\psi}\sum_{s_1}\sum_\theta g(s_1, s_2, \theta)\pi_\theta(\omega')} u(a, \omega)\right\}$$

That is, this is the set of feasible descriptions at ω that induces a_ω if agent 2's belief respects the literal meaning. Then, define

$$\Psi_1^{\min}(\omega) = \arg\min_{\psi\in\cup_{s_2\in S_2}\Psi_\pi(\omega|s_2)} c(\psi).$$

That is, this is the set of the most economical descriptions that can induce a_ω at ω under some s_2. Choose $\omega_0 \in \Omega$. Then, consider an assignment function $\alpha_1(\psi)$, which specifies the state at which

ψ is used. Formally, α_1 is such that

$$\alpha_1 : \bigcup_{\omega \neq \omega_0} \Psi_1^{\min}(\omega) \to \Omega$$

$$\alpha_1(\psi) \in \{\omega \neq \omega_0 : \psi \in \Psi_1^{\min}(\omega)\}.$$

Then, specify $\tilde{\sigma}$ such that $\tilde{\sigma}(\psi|\omega, s_1)$ is uniform over $\{\psi : \alpha_1(\psi) = \omega\}$ and $supp(\tilde{\sigma}(.|\omega_0, s_1)) = \{\varnothing\}$.

Note that there is no guarantee that there exists $\alpha_1(\psi)$ that assigns some $\psi \in \bigcup_{\omega \neq \omega_0} \Psi_1^{\min}(\omega)$ to every state in Ω. Then, to complete construction of $\tilde{\sigma}$, let $\Omega_1^{\alpha_1}$ be the set of $\omega \neq \omega_0$ to which α_1 assigns no ψ. Let $\Psi_1^{\alpha_1}$ be the set of ψ such that $\alpha_1(\psi) = \omega$ for some ω. Then, define

$$\Psi_2^{\min}(\omega) = \arg \min_{\psi \in \cup_{s_2 \in S_2} \Psi_\pi(\omega|s_2) \setminus \Psi_1^{\alpha_1}} c(\psi).$$

This is the set of the most economical descriptions in the unassigned descriptions that can induce a_ω at ω. Then, consider an assignment function $\alpha_2(\psi)$ such that

$$\alpha_2 : \bigcup_{\omega \in \Omega_1^{\alpha_1}} \Psi_2^{\min}(\omega) \to \Omega$$

$$\alpha_2(\psi) \in \{\omega \in \Omega_1^{\alpha_1} : \psi \in \Psi_2^{\min}(\omega)\}$$

Then, for $\omega \in \Omega_1^{\alpha_1}$ such that $\alpha_2(\psi) = \omega$, specify $\tilde{\sigma}$ such that $\tilde{\sigma}(\psi|\omega, s_1)$ is uniform over $\{\psi : \alpha_2(\psi) = \omega\}$.

Again, there is no guarantee that there exists $\alpha_2(\psi)$ that assigns some $\psi \in \bigcup_{\omega \neq \omega_0} \Psi_2^{\min}(\omega)$ to every state in $\Omega_1^{\alpha_1}$. Thus, we need to iterate the same procedure until some descriptions are assigned to every $\omega \in \Omega$. To illustrate $n+1$-th round, let $\Psi_n^{\alpha^n} = \cup_{n'=1}^{n} \Psi_{n'}^{\alpha^{n'}}$ be the set of descriptions that are already assigned to some ω by $\alpha_{n'}$ where $n' \leq n$. Then, define

$$\Psi_{n+1}^{\min}(\omega) = \arg \min_{\psi \in \cup_{s_2 \in S_2} \Psi_\pi(\omega|s_2) \setminus \Psi_n^{\alpha^n}} c(\psi)$$

Moreover, let $\Omega_n^{\alpha^n} = \bigcap_{n'=1}^{n} \Omega_n^{\alpha_n}$, that is, this is the set of $\omega \neq \omega_0$ to which $\alpha_1, \alpha_2, .., \alpha_n$ assign no ψ. Consider a function $\alpha_{n+1}(\psi)$ such that

$$\alpha_{n+1} : \bigcup_{\omega \in \Omega_n^{\alpha^n}} \Psi_{n+1}^{\min}(\omega) \to \Omega$$

$$\alpha_{n+1}(\psi) \in \{\omega \in \Omega_n^{\alpha^n} : \psi \in \Psi_{n+1}^{\min}(\omega)\}.$$

Then, for $\omega \in \Omega_n^{\alpha^n}$ such that $\alpha_{n+1}(\psi) = \omega$, specify $\tilde{\sigma}$ such that $\tilde{\sigma}(\psi|\omega, s_1)$ is uniform over $\{\psi : \alpha_{n+1}(\psi) = \omega\}$. Clearly, the construction of $\tilde{\sigma}$ completes within $|\Omega|$ rounds. Moreover, by construction, $\tilde{\sigma}$ is occasion-free and conditionally separating.

Now, I claim that $\tilde{\sigma}$ is an occasion-free equilibrium strategy. First, by construction, if $\psi \in supp(\tilde{\sigma}(.|\omega, s_1))$, then $\mu_{\tilde{\sigma}}(\omega|\psi, s_2) = 1$ for all s_2 and induces a_ω for sure. Since the cost of $\psi', \psi'' \in supp(\tilde{\sigma}(.|\omega, s_1))$ are the same by construction, agent 1 is indifferent between ψ' and ψ''. Moreover, at ω', any $\tilde{\psi} \in \Psi_{\tilde{\sigma}} \backslash supp(\tilde{\sigma}(.|\omega', s_1))$ induces $a_\omega \neq a_{\omega'}$. Then, by Assumption 2, there is no incentive to use such $\tilde{\psi}$. Finally, by construction, $c(\tilde{\sigma}(\omega, s_1)) < c(\tilde{\psi})$ for any off-path description $\tilde{\psi}$ that can induces a_ω. Then, since $\tilde{\sigma}(\omega, s_1)$ induces a_ω for sure, there is no incentive to use such $\tilde{\psi}$. It follows that $\tilde{\sigma}$ is an occasion-free equilibrium strategy.

Turning to the second part of Proposition 2, I claim that we can always construct an occasion-sensitive equilibrium strategy $\hat{\sigma}$ that uses an pragmatically ambiguous description from the occasion-free strategy $\tilde{\sigma}$.

Let $\tilde{\omega}$ be any $\omega \neq \omega_0$. Then, given $\tilde{\sigma}$, construct $\hat{\sigma}$ as follows.

(i) $\hat{\sigma}(\psi|\tilde{\omega}, s_1') = \tilde{\sigma}(\psi|\omega_0, s_1')$ and $\hat{\sigma}(\psi|\tilde{\omega}, s_1) = \tilde{\sigma}(\psi|\tilde{\omega}, s_1)$ if $s_1 \neq s_1'$;

(ii) $\hat{\sigma}(\psi|\omega_0, s_1')$ is uniform over $\arg\min_{\psi \in \cup_s \Psi_\pi(\omega_0|s_2') \backslash \Psi_{\tilde{\sigma}}} c(\psi)$ where $s_2' = s_1'$;

(iii) For the rest, $\hat{\sigma}(\psi|\omega, s_1) = \tilde{\sigma}(\psi|\omega, s_1)$.

Note that $\hat{\sigma}$ is the same as $\tilde{\sigma}$ except for the descriptions at (ω_0, s_1') and $(\tilde{\omega}, s_1')$. At (ω_0, s_1'), agent 1 uses the most economical unused descriptions in $\tilde{\sigma}$ that can induce a_{ω_0} under some s_1; At $(\tilde{\omega}, s_1')$, agent 1 uses $\varnothing$. By construction, $\hat{\sigma}$ is a conditionally separating occasion-sensitive strategy. Now I claim that $\hat{\sigma}$ is an equilibrium strategy if ρ is sufficiently large.

Step 1. If $\rho \in (0, 1)$ is sufficiently large, then $f_{\hat{\sigma}}(\psi, s_2) = a_\omega$ for $\psi \in supp(\hat{\sigma}(.|\omega, s_1))$ and $s_2 = s_1$

Suppose $\psi' \in supp(\hat{\sigma}(.|\omega', s_1'))$. Agent 2's belief about ω' given ψ' and s_2 is

$$\mu_{\hat{\sigma}}(\omega'|\psi', s_2) = \frac{\sum_\theta \sum_{s_1} \hat{\sigma}(\psi'|\omega', s_1) g(s_1, s_2, \theta) \pi_\theta(\omega')}{\sum_\omega \sum_\theta \sum_{s_1} \hat{\sigma}(\psi'|\omega, s_1) g(s_1, s_2, \theta) \pi_\theta(\omega)}.$$

Since $\hat{\sigma}$ is conditionally separating, $\psi' \notin supp(\hat{\sigma}(.|\omega, s_1'))$ for all $\omega \neq \omega'$. Thus, given ψ' and $s_2 = s_1'$,

$$\mu_{\hat{\sigma}}(\omega'|\psi', s_2) \geq \min_\theta \frac{\rho(s_2)\pi_\theta(\omega')}{\rho(s_2)\pi_\theta(\omega') + (1 - \rho(s_2))}.$$

Let $\zeta_{\omega'}(s_2;\rho)$ be the RHS of the above inequality. If agent 2 chooses $a_{\omega'}$ given ψ' and s_2, his expected payoff is at least $\zeta_{\omega'}(s_2;\rho)u(a_{\omega'},\omega')$. If agent 2 chooses $a \neq a_{\omega'}$, then his expected payoff from a suboptimal action a at ω' is at most $\zeta_{\omega'}(s_2;\rho)u(a,\omega')+(1-\zeta_{\omega'}(s_2;\rho))\max_{(a,\omega)} u(a,\omega)$. Thus, agent 2 chooses $a_{\omega'}$ given ψ' and $s_2 = s_1'$ if

$$\zeta_{\omega'}(s_2;\rho)u(a_{\omega'},\omega') \geq \zeta_{\omega'}(s_2;\rho)\max_{a\neq a_{\omega'}} u(a,\omega') + (1-\zeta_{\omega'}(s_2;\rho))\max_{(a,\omega)} u(a,\omega)$$

Note that we can make $\zeta_{\omega'}(s_2;\rho)$ arbitrarily close to 1 by choosing large $\rho \in (0,1)$. Then, since $u(a_{\omega'},\omega') - \max_{a\neq a_{\omega'}} u(a,\omega') > 0$, agent 2 chooses $a_{\omega'}$ if $\rho \in (0,1)$ is sufficiently large.

Step 2. Agent 1 has no incentive to deviate from $\hat{\sigma}$ if $\rho \in (0,1)$ is sufficiently large.

Case 1. At $\omega \neq \omega_0, \tilde{\omega}$

By construction, $\hat{\sigma}$ uses $\psi' \in supp(\hat{\sigma}(.|\omega, s_1))$ only at ω. Thus, $\mu_{\hat{\sigma}}(\omega|\psi', s_2) = 1$ for all s_2, and agent 1's expected payoff from $\hat{\sigma}$ at (ω, s_1) is $u(a_\omega,\omega) - c(\psi')$. First, if agent 1 uses an on-path description $\tilde{\psi} \notin supp(\hat{\sigma}(.|\omega, s_1))$, then her expected payoff from $\tilde{\psi}$ is at most $\max_{a\neq a_\omega} u(a,\omega)$. Then, from Assumption 2, agent 1 has no incentive to use $\tilde{\psi}$. Second, note that the payoff from ψ' is the same as the equilibrium payoff at ω with $\tilde{\sigma}$. Then, since $\Psi\backslash\Psi_{\hat{\sigma}} \subset \Psi\backslash\Psi_{\tilde{\sigma}}$, agent 1 has no incentive to use any off-path description $\tilde{\psi} \in \Psi\backslash\Psi_{\hat{\sigma}}$.

Case 2. At (ω_0, s_1')

By construction, $\hat{\sigma}$ uses $\psi' \in supp(\hat{\sigma}(.|\omega_0, s_1'))$ only at ω_0. Thus, agent 1's expected payoff from ψ' is $u(a_{\omega_0},\omega) - c(\psi')$.

First, suppose agent 1 uses $\varnothing$, which is pragmatically ambiguous by construction. From Step 1, if $\rho \in (0,1)$ is sufficiently high, $f_{\hat{\sigma}}(\varnothing, s_2) = a_{\tilde{\omega}}$ if $s_2 = s_1'$. Thus, the expected payoff from $\varnothing$ at (ω_0, s_1') is at most $\rho(s_1')u(a_{\tilde{\omega}},\omega_0) + (1-\rho(s_1'))u(a_{\omega_0},\omega_0)$. Thus, agent 1 prefers ψ' to $\tilde{\psi}$ if

$$u(a_{\omega_0},\omega_0) - c(\psi') \geq \rho(s_1')u(a_{\tilde{\omega}},\omega_0) + (1-\rho(s_1'))u(a_{\omega_0},\omega_0).$$

That is,

$$\rho(s_1') \geq \frac{c(\psi')}{u(a_{\omega_0},\omega_0) - u(a_{\tilde{\omega}},\omega_0)}.$$

From Assumption 2, the RHS is strictly smaller than 1. Thus, if $\rho \in (0,1)$ is sufficiently large, the above condition is satisfied.

Second, suppose agent 1 uses $\tilde{\psi} \in supp(\hat{\sigma}(.|\tilde{\omega}, s_1''))$ where $s_1'' \neq s_1'$. Then, by construction, it always induces $a_{\tilde{\omega}}$. Thus, there is no incentive to use $\tilde{\psi}$.

Finally, if agent 1 uses an off-path description $\tilde{\psi}$, agent 1's expected payoff is at most $u(a_{\omega_0}, \omega_0) - c(\tilde{\psi})$. By construction, if $\tilde{\psi} \in \Psi_\pi(\omega_0|s_2) \backslash \Psi_{\hat{\sigma}}$ for some s_2, then $c(\tilde{\psi}) > c(\psi')$. Thus, there is no incentive to use $\tilde{\psi}$.

Case 3. At (ω_0, s_1'') where $s_1'' \neq s_1'$

Agent 1 uses $\varnothing$ at (ω_0, s_1'') in $\hat{\sigma}$. From Step 1, her expected payoff from $\varnothing$ is at least $\rho(s_1'')u(a_{\omega_0}, \omega_0)$ if $\rho \in (0,1)$ is sufficiently large. Agent 1's expected payoff from $\tilde{\psi} \neq \varnothing$ is at most $u(a_{\omega_0}, \omega_0) - c(\tilde{\psi})$. Hence, agent 1 has no incentive to deviate if $\rho(s_1'')u(a_{\omega_0}, \omega_0) \geq u(a_{\omega_0}, \omega_0) - c(\tilde{\psi})$. Since $c(\tilde{\psi}) > 0$, agent 1 has no incentive to deviate if $\rho \in (0,1)$ is sufficiently large.

Case 4. At $(\tilde{\omega}, s_1')$

The strategy $\hat{\sigma}$ uses $\varnothing$ at $(\tilde{\omega}, s_1')$. From Step 1, her expected payoff from $\varnothing$ is at least $\rho(s_1')u(a_{\tilde{\omega}}, \tilde{\omega})$ if $\rho \in (0,1)$ is sufficiently large, whereas her expected payoff from any deviation $\tilde{\psi}$ is at most $u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\tilde{\psi})$. Then, since $c(\tilde{\psi}) > 0$, agent 1 has no incentive to deviate as long as $\rho \in (0,1)$ is sufficiently large.

Case 5. At $(\tilde{\omega}, s_1'')$ where $s_1'' \neq s_1'$

Agent 1 uses $\psi' \in supp(\hat{\sigma}(.|\tilde{\omega}, s_1''))$ only at $\tilde{\omega}$ in $\hat{\sigma}$. Thus, agent 1's expected payoff from ψ' is $u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\psi')$. Note that the payoff is the same as that in the occasion free equilibrium with $\tilde{\sigma}$. Then, since $\tilde{\sigma}$ is an equilibrium strategy and $\Psi_{\hat{\sigma}} \supset \Psi_{\tilde{\sigma}}$, agent 1 has no incentive to use any $\tilde{\psi} \notin \{\varnothing\} \cup supp(\hat{\sigma}(.|\omega_0, s_1'))$.

Suppose agent 1 uses $\varnothing$ at $(\tilde{\omega}, s_1'')$, then the expected payoff is at most $\rho(s_1'')u(a_{\omega_0}, \tilde{\omega}) + (1 - \rho(s_1''))u(a_{\tilde{\omega}}, \tilde{\omega})$. Thus, agent 1 prefers ψ' to $\varnothing$ if

$$\rho(s_1'') \geq \frac{c(\psi')}{u(a_{\tilde{\omega}}, \tilde{\omega}) - u(a_{\omega_0}, \tilde{\omega})}.$$

From Assumption 2, the RHS is strictly smaller than 1. Thus, if $\rho \in (0,1)$ is sufficiently large, the above condition is satisfied. Finally, if $\tilde{\psi} \in supp(\hat{\sigma}(.|\omega_0, s_1'))$, then it induces a_{ω_0} for sure. Then, from Assumption 2, there is no incentive to use $\tilde{\psi}$.

From Case 1-5, agent 1 has no incentive to deviate from $\hat{\sigma}$ if $\rho \in (0,1)$ is sufficiently large.

7.3 Proof of Lemma 1

Let σ^* be an equilibrium strategy. First, to prove "Only if" part, suppose every $\psi \in \Psi_{\sigma^*}$ is not pragmatically ambiguous. Then, since σ^* is conditionally-separating from Proposition 1, $\mu_{\sigma^*}(\omega|\psi, s_2) = 1$ for all s_2 if ψ is an on-path description in σ^*. Hence, $f_{\sigma^*}(\psi, s_2) = a_\omega$ for all s_2 if ψ is an on-path description in σ^*. That is, the equilibrium exhibits no rational miscommunication.

To prove "If" part, suppose $\psi' \in supp(\sigma^*(.|\omega', s_1'))$ is pragmatically ambiguous. Then, since σ^* is conditionally-separating from Proposition 1, there exists $\omega'' \neq \omega'$ and $s_1'' \neq s_1'$ such that $\psi' \in supp(\sigma^*(.|\omega'', s_1''))$. Then, from Claim in the proof of Proposition 1, $f_{\sigma^*}(\psi', s_2) = a_{\omega'}$ for some s_2 whereas $f_{\sigma^*}(\psi', s_2) = a_{\omega''}$ for some s_2. That is, the equilibrium exhibits rational miscommunication.

7.4 Proof of Proposition 3

Any occasion-free equilibrium does not use a pragmatically ambiguous description. Then, from Lemma 1, no occasion-free equilibrium exhibits rational miscommunication. First, I show that given any occasion-free equilibrium, we can construct a language-comparable occasion-sensitive equilibrium that Pareto-dominates the occasion-free equilibrium. Let σ_{OF} be an occasion-free equilibrium strategy.

Claim 1. If $\psi', \psi'' \in supp(\sigma_{OF}(.|\omega', s_1))$, then $c(\psi') = c(\psi'')$.

Suppose $\psi', \psi'' \in supp(\sigma_{OF}(.|\omega', s_1))$ but $c(\psi') > c(\psi'')$. From Proposition 1, σ is conditional-separating, i.e., $\psi', \psi'' \notin supp(\sigma_{OF}(.|\omega, s_1))$ at any $\omega \neq \omega'$. Thus, both ψ' and ψ'' induce $a_{\omega'}$ for sure. Then, agent 1 strictly prefers ψ'' to ψ' if $c(\psi') > c(\psi'')$, a contradiction.

Now, we construct an occasion-sensitive strategy σ_{OS} from σ_{OF}. Depending on the use of $\varnothing$ in σ_{OF}, we construct two kinds of σ_{OS}. Let ω_{s_1} be an element of $\arg\max_\omega \sum_\theta \sum_{s_2} \pi_\theta(\omega) g(\theta, s_2|s_1)$.

Case A: σ_{OF} has $\tilde{\omega}$ such that $\sigma_{OF}(\varnothing|\tilde{\omega}, s_1) > 0$. That is, $\varnothing$ is an on-path description in σ_{OF}.

Let s_1' be s_1 such that $\sum_\theta \sum_{s_2} \pi_\theta(\omega) g(\theta, s_2, |s_1') > \sum_\theta \sum_{s_2} \pi_\theta(\tilde{\omega}) g(\theta, s_2|s_1')$ for some ω. Note that such s_1' exists whenever (g, π) has various modes. Clearly, $\omega_{s_1'} \neq \tilde{\omega}$. Then, construct σ_{OS} as follows:

(i) $\sigma_{OS}(\varnothing|\omega_{s_1'}, s_1) = 1$ for $s_1 = s_1'$;

(ii) if $s_1 \neq s_1'$, $\sigma_{OS}(\psi|\omega_{s_1'}, s_1)$ is uniform over

$$supp(\sigma_{OF}(.|\omega_{s_1'}, s_1)) \cup \left\{ \psi : \psi \in \bigcup_{s_2} \Psi_\pi(\omega_{s_1'}|s_2) \backslash \Psi_{\sigma_{OF}}, c(\psi) = c(\psi'') \right\}$$

where ψ'' is any description in $supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))$.

(iii) If

$$\min_{\psi \in supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))} c(\psi) < \min_{\psi \in \bigcup_{s_2} \Psi_\pi(\tilde{\omega}|s_2) \backslash \Psi_{\sigma_{OF}}} c(\psi),$$

then $\sigma_{OS}(\psi|\tilde{\omega}, s_1')$ is uniform over $supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))$;

If

$$\min_{\psi \in supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))} c(\psi) = \min_{\psi \in \bigcup_{s_2} \Psi_\pi(\tilde{\omega}|s_2) \backslash \Psi_{\sigma_{OF}}} c(\psi),$$

then $\sigma_{OS}(\psi|\tilde{\omega}, s_1')$ is uniform over $supp(\sigma_{OF}(.|\omega_{s_1'}, s_1)) \cup \bigcup_{s_2} \Psi_\pi(\tilde{\omega}|s_2) \backslash \Psi_{\sigma_{OF}}$;

If

$$\min_{\psi \in supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))} c(\psi) > \min_{\psi \in \bigcup_{s_2} \Psi_\pi(\tilde{\omega}|s_2) \backslash \Psi_{\sigma_{OF}}} c(\psi),$$

then $\sigma_{OS}(\psi|\tilde{\omega}, s_1')$ is uniform over $\arg\min_{\psi \in \bigcup_{s_2} \Psi_\pi(\tilde{\omega}|s_2) \backslash \Psi_{\sigma_{OF}}} c(\psi)$.

(iv) For the rest, $\sigma_{OS}(\psi|\omega, s_1) = \sigma_{OF}(\psi|\omega, s_1)$.

Note that even though the use of $\varnothing$ and $supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))$ in σ_{OF} are altered, they are still used at $\tilde{\omega}$ and $\omega_{s_1'}$ respectively. Then, since $\Psi_{\sigma_{OF}} \subset \Psi_{\sigma_{OS}}$, σ_{OS} is language-comparable to σ_{OF}.

Case B: $\sigma_{OF}(\varnothing|\omega, s_1) = 0$ for all ω. That is, $\varnothing$ is an off-path description in σ_{OF}.

(i) For all s_1, $\sigma_{OS}(\varnothing|\omega_{s_1}, s_1) = 1$;

(ii) For the rest, $\sigma_{OS}(\psi|\omega, s_1) = \sigma_{OF}(\psi|\omega, s_1)$.

Clearly, σ_{OS} is language-comparable to σ_{OF}.

Claim 2. σ_{OS} is an occasion-sensitive equilibrium strategy if $\rho \in (0, 1)$ is sufficiently large.

Case A: σ_{OF} has $\tilde{\omega}$ such that $\sigma_{OF}(\varnothing|\tilde{\omega}, s_1) > 0$.

Case A-1. At $(\omega_{s_1'}, s_1')$

By construction, σ_{OF} is conditionally separating. Thus, by a similar argument to Claim 1 in Proposition 2, if $\rho \in (0, 1)$ is sufficiently high, $\varnothing$ induces $a_{\omega_{s_1'}}$ if $s_2 = s_1'$. Thus, agent 1's expected payoff from $\varnothing$ is at least $\rho(s_1')u(a_{\omega_{s_1}}, \omega_{s_1})$ if $\rho \in (0, 1)$ is sufficiently high. If agent 1 uses $\tilde{\psi} \in supp(\sigma_{OS}(.|\omega_{s_1'}, s_1''))$ by deviating from σ_{OS}, then the expected payoff is at most $u(a_{\omega_{s_1}}, \omega_{s_1}) - c(\tilde{\psi})$. Then, since $c(\tilde{\psi}) > 0$, there is no incentive to use $\tilde{\psi}$ if $\rho \in (0, 1)$ is sufficiently high. If agent 1 uses $\tilde{\psi} \in \Psi_{\sigma_{OS}} \backslash supp(\sigma_{OS}(.|\omega_{s_1'}, s_1''))$ by deviating, it always induces some $a \neq a_{\omega_{s_1}}$. Then, from Assumption 2, there is no incentive to use such $\tilde{\psi}$. Finally if agent 1 uses $\tilde{\psi} \in \Psi \backslash \Psi_{\sigma_{OS}}$, her expected payoff from $\tilde{\psi}$ is at most $u(a_{\omega_{s_1}}, \omega_{s_1}) - c(\tilde{\psi})$. Since $c(\tilde{\psi}) > 0$, if $\rho \in (0,1)$ is sufficiently high, there is no incentive to use $\tilde{\psi}$.

Case A-2. At $(\tilde{\omega}, s_1')$

If agent 1 uses $\psi' \in supp(\sigma_{OS}(.|\tilde{\omega}, s_1'))$ at $(\tilde{\omega}, s_1')$, then it can induce $a_{\tilde{\omega}}$ if $s_2 = s_1'$ and $\rho \in (0,1)$ is sufficiently high. Then, agent 1's expected payoff from ψ' is at least $\rho(s_1')u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\psi')$ if $\rho \in (0,1)$ is sufficiently high. If agent 1 uses $\varnothing$ by deviating from σ_{OS}, then it induces $a_{\omega_{s_1'}}$ at $s_2 = s_1'$ if $\rho \in (0,1)$ is sufficiently high. Thus, the expected payoff from $\tilde{\psi}$ is at most $\rho(s_1')u(a_{\omega_{s_1'}}, \tilde{\omega}) + (1 - \rho(s_1'))u(a_{\tilde{\omega}}, \tilde{\omega})$. Then, agent 1 prefers ψ' to $\tilde{\psi}$ if

$$\rho(s_1') \geq \frac{u(a_{\tilde{\omega}}, \tilde{\omega}) + c(\psi')}{2u(a_{\tilde{\omega}}, \tilde{\omega}) - u(a_{\omega_{s_1'}}, \tilde{\omega})}.$$

From Assumption 2, the RHS is strictly smaller than 1. Thus, if $\rho \in (0,1)$ is sufficiently large, the above inequality is satisfied.

If agent 1 uses some on-path description $\tilde{\psi} \neq \varnothing$ by deviating from σ_{OS}, it always induces some $a \neq a_{\tilde{\omega}}$. Thus, from Assumption 2, if $\rho \in (0,1)$ is sufficiently large, there is no incentive to use such $\tilde{\psi}$.

Finally, suppose agent 1 uses $\tilde{\psi} \in \Psi \backslash \Psi_{\sigma_{OS}}$. Then, agent 1's payoff from $\tilde{\psi}$ is at most $u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\tilde{\psi})$. Note that $\Psi_{\sigma_{OF}} \subset \Psi_{\sigma_{OS}}$. Thus, if $\tilde{\psi} \in \Psi_\pi(\tilde{\omega}|s_2) \backslash \Psi_{\sigma_{OS}}$ for some s_2, then $\tilde{\psi} \in \Psi_\pi(\tilde{\omega}|s_2) \backslash \Psi_{\sigma_{OF}}$. Then, by construction, $c(\tilde{\psi}) > c(\psi')$. Thus, if ρ is sufficiently high, there is no incentive to use $\tilde{\psi}$.

Case A-3. At $(\tilde{\omega}, s_1'')$ where $s_1'' \neq s_1'$

If agent 1 uses $\varnothing$ at $(\tilde{\omega}, s_1'')$, then it induces $a_{\tilde{\omega}}$ if $s_2 = s_1''$ and $\rho \in (0,1)$ is sufficiently high. Then, agent 1's expected payoff from $\varnothing$ at $(\tilde{\omega}, s_1'')$, is at least $\rho(s_1'')u(a_{\tilde{\omega}}, \tilde{\omega})$ if $\rho \in (0,1)$ is sufficiently high. If agent 1 uses $\tilde{\psi} \neq \varnothing$ by deviating from σ_{OS}, then the expected payoff from $\tilde{\psi}$ is at most $u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\tilde{\psi})$. Then, since $c(\tilde{\psi}) > 0$, there is no incentive to use such $\tilde{\psi}$ if $\rho \in (0,1)$ is sufficiently large.

Case A-4. At $(\omega_{s_1'}, s_1'')$ where $s_1'' \neq s_1'$

If agent 1 uses $\psi' \in supp(\sigma_{OS}(.|\omega_{s_1'}, s_1''))$ at $(\omega_{s_1'}, s_1'')$, then it induces $a_{\omega_{s_1'}}$ if $s_2 = s_1''$ and $\rho \in (0,1)$ is sufficiently high. Thus, her expected payoff is at least $\rho(s_1'')u(a_{\omega_{s_1'}}, \omega_{s_1'}) - c(\psi')$ if $\rho \in (0,1)$ is sufficiently high. First, if agent 1 uses $\varnothing$ by deviating from σ_{OS}, it induces $a_{\tilde{\omega}}$ if $s_2 = s_1''$ and $\rho \in (0,1)$ is sufficiently high. Thus, her expected payoff from $\varnothing$ is at most

$\rho(s_1'')u(a_{\tilde{\omega}}, \omega_{s_1'}) + (1-\rho(s_1''))u(a_{\omega_{s_1'}}, \omega_{s_1'})$. Then, as in Case A-2, if $\rho \in (0,1)$ is sufficiently large, agent 1 has no incentive to use $\tilde{\psi}$.

Second, if agent 1 uses some on-path description $\tilde{\psi} \neq \varnothing$ by deviating from σ_{OS}, then it always induces some $a \neq a_{\omega_{s_1'}}$. Hence, from Assumption 2, if $\rho \in (0,1)$ is sufficiently large, there is no incentive to use such $\tilde{\psi}$.

Finally, suppose agent 1 uses an off-path $\tilde{\psi}$. If $\tilde{\psi} \notin \Psi_\pi(\omega_{s_1'}|s_2)\backslash\Psi_{\sigma_{OS}}$ for any s_2, then $\tilde{\psi}$ cannot be profitable. Then, consider $\tilde{\psi} \in \Psi_\pi(\omega_{s_1'}|s_2)\backslash\Psi_{\sigma_{OS}}$ for some s_2. If $\tilde{\psi} \notin \Psi_\pi(\omega_{s_1'}|s_2 = s_1'')\backslash\Psi_{\sigma_{OS}}$, then the probability that $\tilde{\psi}$ induces a suboptimal action goes to 1 as $\rho \to 1$. Then, from Assumption 2, agent 1 prefers ψ' to $\tilde{\psi}$ when $\rho \in (0,1)$ is sufficiently large. In contrast, if $\tilde{\psi} \in \Psi_\pi(\omega_{s_1'}|s_2 = s_1'')\backslash\Psi_{\sigma_{OS}}$, the probability that $\tilde{\psi}$ induces $a_{\omega_{s_1'}}$ goes to 1 as $\rho \to 1$. Note that by construction, if $\tilde{\psi}$ is an off-path description in σ_{OS}, it is also an off-path description in σ_{OF}. Thus, if σ_{OF} is an equilibrium strategy under any large $\rho \in (0,1)$, we must have $c(\tilde{\psi}) \geq c(\psi')$. By construction, whenever $\tilde{\psi} \in \Psi_\pi(\omega_{s_1'}|s_2)\backslash\Psi_{\sigma_{OS}}$ for some s_2, $c(\tilde{\psi}) \neq c(\psi')$. Thus, we must have $c(\tilde{\psi}) > c(\psi')$. Then, agent 1 has no incentive to use $\tilde{\psi}$ if $\rho \in (0,1)$ is sufficiently high.

Case A-5. $\omega'' \neq \omega_{s_1'}, \tilde{\omega}$

If agent 1 uses $\psi' \in supp(\sigma_{OS}(.|\omega'', s_1))$ at (ω'', s_1), her payoff is at least $u(a_{\omega''}, \omega'') - c(\psi')$. First, if agent 1 uses some on-path description by deviating from σ_{OS}, then it always induces some $a \neq a_{\omega_{s_1'}}$, then, from Assumption 2, if ρ is sufficiently large, there is no incentive to use such a description. Second, note that agent 1's payoff from ψ' is the same as that in the occasion-free equilibrium with σ_{OF}. Then, since $\Psi_{\sigma_{OF}} \subset \Psi_{\sigma_{OS}}$, there is no incentive to use any off-path description.

Case B: $\sigma_{OF}(\varnothing|\omega, s_1) = 0$ for all ω

Case B-1. At (ω_{s_1}, s_1)

If agent 1 uses $\varnothing$ at (ω_{s_1}, s_1), then it induces $a_{\omega_{s_1}}$ if $s_2 = s_1$ and ρ is sufficiently high. Thus, her expected payoff is at least $\rho(s_1)u(\omega_{s_1}, s_1)$ if ρ is sufficiently high. If agent 1 uses other description $\tilde{\psi}$, agent 1's expected payoff is at most $u(a_{\omega_{s_1}}, \omega_{s_1}) - c(\tilde{\psi})$. Since $c(\tilde{\psi}) > 0$, there is no incentive to use $\tilde{\psi}$ if $\rho \in (0,1)$ is sufficiently high.

Case B-2. At (ω, s_1) where $\omega \neq \omega_{s_1}$

If agent 1 uses $\psi' \in supp(\sigma_{OS}(.|\omega, s_1))$ at (ω, s_1), her expected payoff is at least $u(a_\omega, \omega) - c(\psi')$. First, suppose agent 1 uses $\varnothing$ by deviating from σ_{OS}. Then, it induces some $a \neq a_\omega$ if $s_2 = s_1$ and $\rho \in (0,1)$ is sufficiently high. Thus, agent 1's expected payoff from $\tilde{\psi}$ is at most $\rho(s_1)\max_{a \neq a_{\omega'}} u(a, \omega) + (1-\rho(s_1))u(a_\omega, \omega)$. Then, from Assumption 2, agent 1 prefers ψ' to $\tilde{\psi}$ if

$\rho \in (0,1)$ is sufficiently large.

Second, if agent 1 uses some on-path description $\tilde{\psi} \neq \varnothing$ by deviating from σ_{OS}, it induces some $a \neq a_\omega$ for sure. Then, from Assumption 2, there is no incentive to use $\tilde{\psi}$.

Finally, if agent 1 uses an off-path description $\tilde{\psi}$, agent 1's payoff from ψ' is the same as that in the equilibrium with σ_{OF}. Then, since $\Psi_{\sigma_{OF}} \subset \Psi_{\sigma_{OS}}$, agent 1 has no incentive to use $\tilde{\psi}$ at (ω, s_1) with σ_{OS}.

Claim 3. The equilibrium with σ_{OS} Pareto-dominates the equilibrium with σ_{OF}.

Let $V_\sigma(\omega, s_1)$ be agent 1's expected payoff at (ω, s_1) in the equilibrium with σ.

Case A: σ_{OF} has $\tilde{\omega}$ such that $\sigma_{OF}(\varnothing|\tilde{\omega}, s_1) > 0$.

Case A-1: At $(\omega_{s_1'}, s_1')$

Since σ_{OS} uses $\varnothing$ at $(\omega_{s_1'}, s_1')$ and it is ambiguous in σ_{OS}, $V_{\sigma_{OS}}(\omega_{s_1'}, s_1')$ is at least $\rho(s_1')u(a_{\omega_{s_1'}}, \omega_{s_1'})$ if ρ is sufficiently high. In contrast, any $\psi \in supp(\sigma_{OF}(.|\omega_{s_1'}, s_1'))$ is not ambiguous in σ_{OF} and induces the state-optimal action. Thus, $V_{\sigma_{OF}}(\omega_{s_1'}, s_1') = u(a_{\omega_{s_1'}}, \omega_{s_1'}) - c(\psi_1)$ where ψ_1 is some description in $supp(\sigma_{OF}(.|\omega_{s_1'}, s_1'))$. Hence,

$$V_{\sigma_{OS}}(\omega_{s_1'}, s_1') - V_{\sigma_{OF}}(\omega_{s_1'}, s_1') \geq (\rho(s_1') - 1)u(a_{\omega_{s_1'}}, \omega_{s_1'}) + c(\psi_1)$$

Case A-2: At $(\tilde{\omega}, s_1')$

Since $\psi \in supp(\sigma_{OS}(.|\tilde{\omega}, s_1'))$ can be ambiguous in σ_{OS}, $V_{\sigma_{OS}}(\tilde{\omega}, s_1')$ is at least $\rho(s_1')u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\psi)$ if ρ is sufficiently high. On the contrary, since σ_{OF} uses $\varnothing$ at $(\tilde{\omega}, s_1')$ and it is not ambiguous in σ_{OF}, $V_{\sigma_{OF}}(\tilde{\omega}, s_1') = u(a_{\tilde{\omega}}, \tilde{\omega})$. Thus,

$$V_{\sigma_{OS}}(\tilde{\omega}, s_1') - V_{\sigma_{OF}}(\tilde{\omega}, s_1') \geq (\rho(s_1') - 1)u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\psi_2)$$

where ψ_2 is some description in $supp(\sigma_{OS}(.|\tilde{\omega}, s_1'))$.

Case A-3: At $(\tilde{\omega}, s_1)$ where $s_1 \neq s_1'$

Both σ_{OS} and σ_{OF} use $\varnothing$ at $(\tilde{\omega}, s_1)$ but $\varnothing$ is ambiguous only in σ_{OS}. Thus, as in Case A-1, $V_{\sigma_{OS}}(\tilde{\omega}, s_1)$ is at least $\rho(s_1)u(a_{\tilde{\omega}}, \tilde{\omega})$ if ρ is sufficiently high, whereas $V_{\sigma_{OF}}(\tilde{\omega}, s_1) = u(a_{\tilde{\omega}}, \tilde{\omega})$. Hence,

$$V_{\sigma_{OS}}(\tilde{\omega}, s_1) - V_{\sigma_{OF}}(\tilde{\omega}, s_1) \geq (\rho(s_1) - 1)u(a_{\tilde{\omega}}, \tilde{\omega})$$

Case A-4: At $(\omega_{s_1'}, s_1)$ where $s_1 \neq s_1'$

Since $\psi \in supp(\sigma_{OS}(.|\omega_{s_1'}, s_1))$ can be ambiguous in σ_{OS}, $V_{\sigma_{OS}}(\omega_{s_1'}, s_1)$ is at least $\rho(s_1)u(a_{\omega_{s_1'}}, \omega_{s_1'}) -$

$c(\psi)$ if ρ is sufficiently high. In contrast, since $\psi \in supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))$ is not ambiguous in σ_{OF}, $V_{\sigma_{OS}}(\omega_{s_1'}, s_1) = u(a_{\omega_{s_1'}}, \omega_{s_1'}) - c(\psi)$. Note that, by construction, if $\psi' \in supp(\sigma_{OF}(.|\omega_{s_1'}, s_1))$ and $\psi'' \in supp(\sigma_{OS}(.|\omega_{s_1'}, s_1))$, then $c(\psi') = c(\psi'')$. Hence,

$$V_{\sigma_{OS}}(\omega_{s_1'}, s_1) - V_{\sigma_{OF}}(\omega_{s_1'}, s_1) \geq (\rho(s_1) - 1)u(a_{\omega_{s_1'}}, \omega_{s_1'})$$

Case A-5 $\omega'' \neq \omega_{s_1'}, \tilde{\omega}$

Clearly, agent 1's expected payoff at (ω'', s_1) is the same in both equilibria. That is, $V_{\sigma_{OS}}(\omega'', s_1) - V_{\sigma_{OF}}(\omega'', s_1) = 0$.

From A-1,2,3,4, and 5, the difference between the agent 1's ex-ante expected payoff in the equilibrium with σ_{OS} and that with σ_{OF} is

$$\begin{aligned}
&\sum_{\omega}\sum_{s_1}\sum_{s_2}\sum_{\theta}[V_{\sigma_{OS}}(\omega, s_1) - V_{\sigma_{OF}}(\omega, s_1)]g(s_1, s_2, \theta)\pi_\theta(\omega) \\
=&[(\rho(s_1') - 1)u(a_{\omega_{s_1'}}, \omega_{s_1'}) + c(\psi_1)]\sum_{\theta}\sum_{s_2}\pi_\theta(\omega_{s_1'})g(s_2, \theta|s_1')g(s_1') \\
&+\sum_{s_1 \neq s_1'}(\rho(s_1) - 1)u(a_{\omega_{s_1}}, \omega_{s_1})\sum_{\theta}\sum_{s_2}\pi_\theta(\omega_{s_1'})g(s_2, \theta|s_1)g(s_1). \\
&+[(\rho(s_1') - 1)u(a_{\tilde{\omega}}, \tilde{\omega}) - c(\psi_2)]\sum_{\theta}\sum_{s_2}\pi_\theta(\tilde{\omega})g(s_2, \theta|s_1')g(s_1') \\
&+\sum_{s_1 \neq s_1'}[(\rho(s_1) - 1)u(a_{\tilde{\omega}}, \tilde{\omega})]\sum_{\theta}\sum_{s_2}\pi_\theta(\tilde{\omega})g(s_2, \theta|s_1)g(s_1)
\end{aligned}$$

Note that, by choosing large $\rho \in (0,1)$, we can make the above difference arbitrarily close to

$$c(\psi_1)\sum_{\theta}\sum_{s_2}\pi_\theta(\omega_{s_1'})g(s_2, \theta|s_1')g(s_1') - c(\psi_2)\sum_{\theta}\sum_{s_2}\pi_\theta(\tilde{\omega})g(s_2, \theta|s_1')g(s_1').$$

Recall that, by definition,

$$\sum_{\theta}\sum_{s_2}\pi_\theta(\omega_{s_1'})g(s_2, \theta|s_1') > \sum_{\theta}\sum_{s_2}\pi_\theta(\tilde{\omega})g(s_2, \theta|s_1')$$

Moreover, note that, by construction, $c(\psi_1) \geq c(\psi_2)$. Hence, if $\rho \in (0,1)$ is sufficiently high, agent 1's expected payoff in the occasion-sensitive equilibrium is higher than that in the occasion-free equilibrium.

Case B: $\sigma_{OF}(\varnothing|\omega, s_1) = 0$ for all ω

Case B-1. At (ω_{s_1}, s_1)

Since σ_{OS} uses $\varnothing$ at (ω_{s_1}, s_1) and it is ambiguous in σ_{OS}, $V_{\sigma_{OS}}(\omega_{s_1}, s_1)$ is at least $\rho(s_1)u(\omega_{s_1}, s_1)$ if ρ is sufficiently large. On the contrary, any $\psi \in supp(\sigma_{OF}(.|\omega_{s_1}, s_1))$ is not ambiguous and induces the state-optimal action for sure. Thus,

$$V_{\sigma_{OS}}(\omega_{s_1}, s_1) - V_{\sigma_{OF}}(\omega_{s_1}, s_1) \geq (1 - \rho(s_1'))u(a_{\omega_{s_1}}, \omega_{s_1}) + c(\psi').$$

where ψ' is some description in $supp(\sigma_{OF}(.|\omega_{s_1}, s_1))$.

Case B-2. At (ω, s_1) where $\omega \neq \omega_{s_1}$

In this case, by construction, $\sigma_{OS}(\psi|\omega, s_1) = \sigma_{OF}(\psi|\omega, s_1)$. Thus, $V_{\sigma_{OS}}(\omega, s_1) - V_{\sigma_{OF}}(\omega, s_1) = 0$.

From B-1 and 2, the difference between the agent 1's ex-ante expected payoff in the equilibrium with σ_{OS} and that with σ_{OF} is

$$\sum_{s_1}\sum_{\omega_{s_1}}[(\rho(s_1) - 1)u(a_{\omega_{s_1}}, \omega_{s_1}) + c(\psi')]\sum_{\theta}\sum_{s_2}\pi_\theta(\omega_{s_1})g(\theta, s_2|s_1)g(s_1)$$

Note that, by choosing large $\rho \in (0, 1)$, we can make the above arbitrary close to

$$c(\psi')\sum_{s_1}\sum_{\omega_{s_1}}\sum_{\theta}\sum_{s_2}\pi_\theta(\omega_{s_1})g(\theta, s_2|s_1)g(s_1).$$

Note that $c(\psi') > 0$. Hence, if $\rho \in (0, 1)$ is sufficiently high, agent 1's expected payoff in the occasion-sensitive equilibrium is higher than that in the occasion-free equilibrium.

There can also be occasion-sensitive equilibria that do not exhibit rational miscommunication.

Claim 4. If an occasion-sensitive equilibrium does not exhibit rational miscommunication, then there is a payoff-equivalent occasion-free equilibrium.

Consider any occasion-sensitive equilibrium that does not exhibit rational miscommunication. Let σ' be the equilibrium strategy. From Lemma 1, no equilibrium description is ambiguous. Then, from Proposition 1, the cost of any equilibrium description at ω has to be the same across s_1. Let σ'' be an occasion-free strategy such that $supp(\sigma''(.|\omega, s_1)) = \bigcup_{s_1} supp(\sigma'(.|\omega, s_1))$ for all s_1. Note that $\Psi_{\sigma'} = \Psi_{\sigma''}$ by construction. Clearly, σ'' is an occasion-free equilibrium strategy, and the equilibrium payoff at every (ω, s_1) is the same as that with σ'. Moreover, σ'' is language-comparable to σ'.

As shown earlier, given any occasion-sensitive equilibrium, we can construct a Pareto-dominant language-comparable occasion-sensitive equilibrium that exhibits rational miscommunication if the conditions is Proposition 3 is satisfied.

7.5 Proof of Proposition 4

Suppose there is an equilibrium that exhibits rational miscommunication. Then, from Lemma 1, agent 1 uses a pragmatically ambiguous description ψ' in the equilibrium. Suppose that ψ' is used at (ω', s_1') and (ω'', s_1'').

First, if $g(s_1, s_2, \theta) = g(s_1, \theta)g(s_2)$, then

$$\mu^*(\omega|\psi', s_2) = \frac{\sum_\theta \sum_{s_1} \sigma^*(\psi'|s_1, \omega)g(s_1, \theta)\pi_\theta(\omega)}{\sum_{\omega'} \sum_\theta \sum_{s_1} \sigma^*(\psi'|s_1, \omega')g(s_1, \theta)\pi_\theta(\omega')}$$

That is, $\mu(\omega|\psi', s_2)$ is constant in s_2. From Proposition 1, $f_\sigma(\psi', s_2) = a_{\omega'}$ for some s_2. Then, we must have $f_\sigma(\psi', s_2) = a_{\omega'}$ for all s_2, and ψ' never induces $a_{\omega''}$ at (ω'', s_1''). But then, from Assumption 2, agent 1 has an incentive to use a fully precise description of ω'' at (ω'', s_1''), which induces $a_{\omega''}$ for sure, a contradiction.

Second, if $g(s_1, s_2, \theta) = g(s_2, \theta)g(s_1)$, then $\mu^*(\omega|\psi', s_2)$ depends on s_2. If agent 1 uses ψ', there exists a non-empty set

$$S_\omega^{\psi'} = \{s_2 \in S_2 : f_{\sigma^*}(\psi', s_2) = a_\omega\}.$$

Then, the probability of inducing the optimal action $a_{\omega'}$ with ψ' is $\sum_\theta \sum_{s_2 \in S_{\omega'}^{\psi'}} g(s_2, \theta)$. From Assumption 2, if $\sum_\theta \sum_{s_2 \in S_{\omega'}^{\psi'}} g(s_2, \theta) \leq 0.5$, the agent 1 prefers to use a fully precise description of ω', which induces $a_{\omega'}$ for sure. Thus, we must have $\sum_\theta \sum_{s_2 \in S_{\omega'}^{\psi'}} g(s_2, \theta) > 0.5$. But then, since $S_2 \backslash S_{\omega'}^{\psi'} \supset S_{\omega''}^{\psi'}$, the probability of inducing the optimal action $a_{\omega''}$ with ψ', i.e., $\sum_\theta \sum_{s_2 \in S_{\omega''}^{\psi'}} g(s_2, \theta)$, has to be at most 0.5. Then, from Assumption 2, agent 1 has an incentive to use a fully precise description of ω'' at (ω'', s_1''), a contradiction.

References

Arrow, Kenneth J (1974). *The limits of organization*. WW Norton & Company.

Austin, John Langshaw (1975). *How to do things with words*. Vol. 88. Oxford university press.

Battigalli, Pierpaolo and Giovanni Maggi (2002). "Rigidity, discretion, and the costs of writing contracts". In: *American Economic Review* 92.4, pp. 798–817.

Benz, Anton, Gerhard Jäger, and Robert Van Rooij (2005). *Game theory and pragmatics*. Springer.

Blume, Andreas (1996). "Neighborhood stability in sender–receiver games". In: *Games and Economic Behavior* 13.1, pp. 2–25.

Blume, Andreas and Oliver Board (2013). "Language barriers". In: *Econometrica* 81.2, pp. 781–812.

Blume, Andreas, Yong-Gwan Kim, and Joel Sobel (1993). "Evolutionary stability in games of communication". In: *Games and Economic Behavior* 5.4, pp. 547–575.

Crawford, Vincent P and Joel Sobel (1982). "Strategic information transmission". In: *Econometrica: Journal of the Econometric Society*, pp. 1431–1451.

Cremer, Jacques, Luis Garicano, and Andrea Prat (2007). "Language and the Theory of the Firm". In: *The Quarterly Journal of Economics* 122.1, pp. 373–407.

Dewatripont, Mathias and Jean Tirole (2005). "Modes of communication". In: *Journal of political economy* 113.6, pp. 1217–1238.

Estival, Dominique, Candace Farris, and Brett Molesworth (2016). *Aviation English: A lingua franca for pilots and air traffic controllers*. Routledge.

Glazer, Jacob and Ariel Rubinstein (2001). "Debates and decisions: On a rationale of argumentation rules". In: *Games and Economic Behavior* 36.2, pp. 158–173.

— (2006). "A study in the pragmatics of persuasion: a game theoretical approach". In: *Theoretical Economics* 1, pp. 395–410.

Grice, H.P. (1975). *"Logic and Conversation," Syntax and Semantics,* ed. by P. Cole and J. Morgan. Vol. vol.3. Academic Press.

Grossman, Sanford J (1981). "The informational role of warranties and private disclosure about product quality". In: *The Journal of Law and Economics* 24.3, pp. 461–483.

Huang, Yan (2007). *Pragmatics*. Oxford University Press.

Jäger, Gerhard, Lars P Metzger, and Frank Riedel (2011). "Voronoi languages: Equilibria in cheap-talk games with high-dimensional types and few signals". In: *Games and economic behavior* 73.2, pp. 517–537.

Marschak, Jacob and Roy Radner (1972). *Economic Theory of Teams*. Yale University Press.

Milgrom, Paul R (1981). “Good news and bad news: Representation theorems and applications”. In: *The Bell Journal of Economics*, pp. 380–391.

Rubinstein, Ariel (1996). “Why are certain properties of binary relations relatively more common in natural language?” In: *Econometrica*, pp. 343–355.

Shannon, Claude E (1948). “A mathematical theory of communication”. In: *The Bell system technical journal* 27.3, pp. 379–423.

Sobel, Joel (2012). “Complexity versus conflict in communication”. In: *2012 46th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, pp. 1–6.

— (2020). “Lying and deception in games”. In: *Journal of Political Economy* 128.3, pp. 907–947.

Suzuki, Toru (2017). “Directives, expressives, and motivation”. In: *Theoretical Economics* 12.1, pp. 175–210.

— (2020). “Efficient communication and indexicality”. In: *Mathematical Social Sciences* 108, pp. 156–165.