

# Understanding of linguistic politeness

Erica J. Yoon

ejyoon@stanford.edu

Department of Psychology

Stanford University

## Abstract

To be added.

**Keywords:** Pragmatics; language; politeness; probabilistic model; cognition

## Introduction

Communication is an act of cooperation between two or more partners (Grice, 1980), and people normally assume that their communicators would speak in a way that is maximally truthful and informative. Imagine two people in a conversation, Tom and Sally, who are talking about pets they own. Tom says to Sally: “I have two dogs.” Assuming that Tom is a cooperative speaker and Sally also believes that he is, Sally will imagine the true state of the world is such that Tom owns at least two dogs, which makes Tom’s utterance truthful. Furthermore, she will infer that Tom owns two dogs, not three or more, assuming Tom is informative and would have used a term of greater number otherwise. Now imagine the following situation: Tom and Sally are talking about a presentation that Sally just gave, and Sally asked for Tom’s opinion on her talk. Tom says to Sally: “It was okay.” We can apply the same analysis as the previous one in this situation: if we assume Tom has some scale on his mind that ranges from ‘terrible’ to ‘excellent,’ and the term ‘okay’ covers the positive part of the scale starting from the midpoint to the ‘excellent’ end, then we might infer that Tom meant to indicate that Sally’s presentation was “above average but not excellent.” But is that as far as we can infer? For example, if Sally’s performance was exactly average, would Tom necessarily use the word ‘average’ instead of ‘okay’ to be more informative and exact? Is it also possible that Tom used the word ‘okay’ even though Sally’s performance was actually below average? In that case, should we conclude that Tom was not a cooperative speaker? But people do share a strong intuition to maximize praise for their conversational partner, even by telling white lies and compromising truthfulness (REF). Even children show the tendency to tell white lies and think of white lies as better than malicious lies or even truths in situations in which the listener is harmed by truth-telling (REF). Past literature has explained this phenomenon of ‘linguistic politeness’ in terms of people’s want to maintain each other’s ‘face’: every human’s wants to maintain positive public self-image and freedom from impositions. Going back to the example above, speakers maximize praises for the listener in order to save the listener’s positive face, or self-esteem. Leech (1983) furthered this idea and proposed Politeness Principle (PP), which includes Maxim of Approbation: speakers must

maximize praise and minimize dispraise for the listener. Despite the rich literature on related phenomena of linguistic politeness observed across different situations and cultures, there has thus far been few formal models for the common phenomenon of linguistic politeness. The goal of the current paper is to propose a generative model that looks at mechanisms of polite utterance generation and of listeners’ inferences based on the utterances about speakers’ underlying goals.

## Model

The current model<sup>1</sup> is a variant of Kao et al. (2014)’s model and follows its central assumptions: communicators are rational and cooperative agents, and listeners make inferences based on the assumption that speakers choose utterances to maximally achieve their communicative goals. The current model assumes there are two potential goals (and another goal that is combination of the two): the goal to convey the state, and goal to be polite. These goals are built in to the current model as possible QUD’s. The current model assume that there are four possible states, which represent objective, evaluative scores for some item (e.g., Sally’s presentation). The state of 1 denotes the worst possible score, 4 the best. This positive or negative association of states is formalized in the model in the form of valence prior, which is more likely for more desirable states (e.g., 4) and less likely for less desirable states (e.g., 1). Linguistic terms are assigned to particular state(s): “bad” denotes state of 1; “not bad” states of 2, 3, or 4; “great” state of 4, “not great” state of 1, 2, or 3. Given this scale of states and state-associated terms, let’s suppose there was only pressure to be truthful and informative for speakers. Then given a state of 1, speakers would probably use the word “bad,” and given a state of 4, speakers would probably use the word “great,” to be truthful and maximally informative in each case. Given a state of 2, it would be equally likely for speakers to say “not bad” or “not great.” However, with an additive pressure for speakers to be polite, “not bad” would be a better option than “not great” to describe the state of 2, assuming that, given the same amount of ambiguity, speakers would want to maximize praise (i.e., inclusion of maximally desirable state of 4 in ambiguity) and minimize dispraise (i.e., exclusion of minimally desirable state of 1). The ‘literal listener’ in the model hears an input utterance and draws QUD from QUD prior, and infers the state (1, 2, 3, or 4) and politeness level (which is for now binary: 0 for impolite and 1 for polite). The ‘speaker’ reasons about this literal listener

<sup>1</sup>Link: [https://github.com/ejyoon/polite\\_adj/blob/master/polite\\_adj-model.txt](https://github.com/ejyoon/polite_adj/blob/master/polite_adj-model.txt)

and outputs an utterance that aligns best with the given QUD. Then the ‘pragmatic listener’ hears this speaker’s utterance and infers the state and politeness level, taking into account speaker’s reasoning about literal listener. Finally, ‘speaker2’ accesses state and politeness as inferred by the pragmatic listener, and generates an utterance that most accurately conveys the speaker2’s communicative goal to the pragmatic listener. The model makes predictions that closely match the intuitive predictions. For the pragmatic listener, state of 3 is more likely for the speaker’s utterance “not bad” than state of 2, whereas this pattern flips for the utterance “not great”, which reflects the pragmatic listener’s inference about speaker’s potential goal to be polite and use the more favorably ambiguous expression. Also, the pragmatic listener is predicted to attribute politeness more to a speaker who says “not bad” compared to one who says “not great.” Then the speaker2, who speaks with the pragmatic listener on her mind, is more likely to say “not bad” than “not great” given the state of 2 if she is polite, but more likely to say “not great” than “not bad” if she is impolite (and cares more about the goal of conveying the state).

## **Experiment 1**

### **Method**

To be added.