

What is Keigo?

Keigo (敬語) is a critical system of honorific language in Japanese, used to express politeness and respect. It forms an integral part of communication, especially in formal situations. Keigo plays a vital role in maintaining social harmony by adjusting speech levels based on the social status or relationship between the speaker and the listener.

Importance of Keigo in Business and Interviews

In business interactions and job interviews, keigo becomes essential as it reflects professionalism, humility, and understanding of social etiquette. Using the correct keigo shows respect to the interviewer or business counterpart, contributing to a positive impression.

Misuse, on the other hand, can lead to misunderstandings or be perceived as rude, affecting personal or business relationships.

Types of Keigo

We can classify keigo into three types, each designed to adjust the speaker's relationship with the listener.

1. **Teineigo (丁寧語) - Polite Speech:** (The one that is most expected)
 - This is the most neutral form of keigo, used to maintain basic politeness. It's commonly used in everyday formal speech.
 - **Purpose:** Ensures politeness, regardless of social status.
 - **Example:** 「行きます。」(Ikimasu) – "I will go."

2. **Sonkeigo (尊敬語) - Honorific Speech:**
 - This elevates the subject of the conversation, often used when speaking about superiors or clients.
 - **Purpose:** Shows respect to the person being referred to.
 - **Example:** 「社長がいらっしゃいます。」(Shachou ga irasshaimasu) – "The president is here."

3. **Kenjougo (謙譲語) - Humble Speech:**
 - This form humbles the speaker or the in-group, lowering their status while elevating the listener or another person.
 - **Purpose:** The speaker shows humility, making themselves or their actions seem less important.
 - **Example:** 「お持ちいたします。」(Omochi itashimasu) – "I will bring it."

Common Forms of Keigo: Prefixes and Suffixes

We can break down keigo into identifiable linguistic patterns, including both **prefixes** and **suffixes** that mark different types of keigo.

1. Teineigo (Polite Speech)

- **Prefixes:**
 - **"o-"** (お) for nouns (e.g., **"o-cha"** お茶 – tea)
 - **"go-"** (ご) for some nouns and verbs (e.g., **"go-aisatsu"** ご挨拶 – greeting)
- **Suffixes:**
 - **"~masu"** (ます) for verbs (e.g., **"ikimasu"** – to go)
 - **"~desu"** (です) for adjectives/nouns (e.g., **"kirei desu"** – It's beautiful)
 - **"~gozaimasu"** (ございます) for heightened politeness (e.g., **"arigatou gozaimasu"** – thank you)

2. Sonkeigo (Honorific Speech)

- **Prefixes:**
 - **"o-"** (お) for verbs/nouns (e.g., **"o-ikimasu"** – to go)
 - **"go-"** (ご) (e.g., **"go-shoukai"** ご紹介 – introduction)
- **Suffixes:**
 - **"~rareru"** (られる) (e.g., **"mirareru"** – to see [honorific])
 - **"~nasaru"** (なさる) (e.g., **"nasaimasu"** – to do)
 - Specific verbs: **"irasshaimasu"** (いらっしゃいます – to go, come, be), **"ossharu"** (おっしゃる – to say), **"meshiagaru"** (召し上がる – to eat)

3. Kenjougo (Humble Speech)

- **Prefixes:**
 - **"o-"** (お) for nouns/verbs (e.g., **"o-tsutae"** お伝え – message)
 - **"go-"** (ご) for some nouns/verbs (e.g., **"go-teinei"** ご丁寧 – politeness)
- **Suffixes:**
 - **"~itasu"** (いたす) (e.g., **"omochi itashimasu"** – to humbly bring)
 - **"~mousu"** (申す) (e.g., **"moushimasu"** – to humbly say)
 - Specific verbs: **"ukagau"** (伺う – to humbly ask), **"mairu"** (参る – to humbly go)

Designing an algorithm to identify and analyze keigo

To design an algorithm for identifying and analyzing keigo in text, you could consider the following metrics and approaches:

a) Verb form classification:

- Identify and count specific verb endings associated with each type of keigo.

b) Honorific prefixes:

- Count the occurrences of honorifics.

c) Context analysis:

- Implement natural language processing techniques to understand the context and determine if keigo is used appropriately. [**CHALLENGE**]
- Can use OpenAI directly

Metrics to consider:

- 1. Keigo density:** The ratio of keigo expressions to total words or sentences. (Only this was being considered in the latest model)
- 2. Keigo variety:** The number of different keigo forms used.
- 3. Appropriateness score:** Based on context analysis, how well the keigo matches the assumed social situation.

Methods to Detect and Count Keigo

1. Romaji-based Detection with Regex and Python Libraries

We can convert Japanese text into **romaji** using Python libraries and detect keigo patterns with **regex**. Here's the step-by-step process:

1. Convert Japanese to Romaji:

- We can use libraries like **romkan** to convert kana and kanji into romaji.
 - Example: 「行きます」 → "ikimasu"

2. Define Regex Patterns:

- We can create regex patterns to detect common keigo forms.

2. Using MeCab for Keigo Detection

MeCab, a morphological analysis tool for Japanese, can analyze Japanese text and tag words with their parts of speech. We can leverage this tagging to detect keigo forms.

Steps:

1. **Parse the text:** We can use MeCab to split sentences and analyze words with their parts of speech.
2. **Identify keigo forms:** We can define rules to detect keigo by identifying verb stems, conjugations, and honorific markers (e.g., specific verbs like "**irassharu**", "**itasu**", etc.).

Keep in Mind

Complete Change of Words:

For example: *suru* (do) becomes *itasu*, and *morau* (receive) becomes *itadaku*. This becomes a challenge

Density:

Ideally one keigo per sentence. But no concept of sentences because we are directly tokenizing.

Kanji:

Romaji conversion can face challenges in kanji conversion.