Roots of the \textit{wakimae} aspect of linguistic politeness

Modal expressions and Japanese sense of self

Sachiko Ide
Japan Women’s University

Why is Japanese linguistic politeness characterized by \textit{wakimae}? This paper argues that \textit{wakimae}, or linguistic politeness, can be traced back to two roots. First, the Japanese language has abundant modal expressions from the morpheme level to the discourse level that index the context in order to show the speaker’s attitude toward the contextual elements involved. The structure of utterances consists of the proposition and the modality, and it is this modality expression that shows the speaker’s attitude toward the context of speaking. Second, the Japanese ‘sense of self’ consists of two layers of interactional domains: \textit{uchi} (ingroup) and \textit{soto} (outgroup). The use of modal expressions such as addressee honorifics is required to index the interactional domains of \textit{soto} to observe \textit{wakimae}. Finally, in this paper the moral aspiration for the pursuit of \textit{wakimae} is discussed in light of historical development of Japanese philosophy.

1. Introduction

Eloquence and persuasion in verbal discourse are not virtues in Japanese. With this statement, Shibatani (1990) concluded his book on Japanese. As hard as it may be to imagine that eloquence and persuasion are not the goals in Japanese discourse, that is indeed the case. So if speakers are not striving for eloquence and persuasion, what are they striving for? What are the virtues? What are the characteristics of discourse that Japanese speakers agree are worthy of effort? An answer to these questions can shed light on the fundamental differences between successful Japanese discourse and successful Western discourse, and thus can lead to an increased understanding of potential pitfalls in cross-cultural communication.
2. Modality in Japanese

Anytime someone says something, there is an element of subjectivity in the utterance, since the decision to say anything at all is a subjective decision. The way the utterance is phrased is also subjective, and represents one of many possibilities. "It's a bit cool in here," "Would it be ok if I closed the window?" and 'Shut the window!' are all utterances that aim toward the same outcome, but they are not interchangeable in the sense that they are not equally appropriate in different situations. The subjectivity reflected in these utterances results from the speaker's analysis of the relative positions of the participants as well as the context, and this analysis leads to the choice of a statement, a hint, a question, or an imperative.

The term used to describe this whole array of subjective features is modality, and it covers the speaker's belief, attitude, and feeling toward the proposition, as a subjective commentary on the propositional content of an utterance, or as an acknowledgment of the roles, places, and relationships of the participants in a conversation. Modality can be expressed in many different ways in different languages. For example, in English it can be done through intonation, as in 'He s-a-i-d he was sick,' whereby the drawing out of 'said' suggests that the speaker doubts the truth of what he said.

Many languages have what have come to be known as modal particles: French donc or German zwar. It is only recently that these modal particles in Western languages have been elevated to the status of phenomena worthy of academic study. Japanese linguists, however, from Tokieda (1941) through Watanabe (1969), Teramura (1982), Nitta (1989), and Masuoka (1992), have commented on and described modality in Japanese as an obligatory layer or component. Thus, a Japanese sentence can be seen as consisting of two parts - a proposition and a modality - both of which must be expressed. In contrast, English could be seen as consisting of an obligatorily expressed proposition, and a modal component, which may or may not be explicit. It is almost as if, in addition to expressed modality, there is an unmarked modality in English that requires no explicit expression. Only if the modality diverges from the norm, it is verbalized, and then through intonation or additions (such as tag questions, adverbs, and the like), not through grammatical means.

In Japanese, the modal component is obligatory. This may run counter to an often-held belief in the universality of the propositional statement as the core of a sentence. Indeed, upon further reflection, it is not really surprising that the most stripped-down utterance consists of a proposition and a perspective, since language is used among participants who have widely divergent relationships to one another and in situations of great diversity.

Thus, modality is the form the speaker gives to the subjectivity. In Japanese, it reflects the speaker's evaluation of the context in which the utterance is made, and acknowledges the relative positions of the participants, their psychological distance, and the degree of formality appropriate to the situation - and it does so explicitly. A whole set of words or morphemes are available to mark the choices made. A linguistic form is said to be 'marked' when a word or a morpheme is added to an unmarked form to distinguish it with regard to a certain property. Thus, the names in 'Schulz, White, and Matsumoto should report to the office as soon as possible' are unmarked; in 'Dr. White, the person is marked for academic level, whereas the expression 'Mrs. White' is marked only for gender. Modality in Japanese marks decisions as to whether the participants are close or distant, whether the situation is casual or formal, etc., and these markers are, for the most part, placed at the end of the utterance after the predicate. Figure 1 illustrates this by showing the proposition embedded in expressed modality.

Figure 1. Sentence structure of Japanese

Since modality covers a wide variety of subjective factors, it is not surprising that the many choices that must be made are ordered in different layers. The inner layer close to the proposition consists of the speaker's judgment as to the truth or falsity of the proposition (as in 'She s-a-i-d she was ill'). The outer layer contains the features that mark, or more precisely, index, the speaker's relation to the other participant(s); the required expression of politeness toward the addressee, determined by the person, the setting, and the occasion; and the modality of the speaker's mood.

Among the various functions of modal expressions, the most relevant are those that index the context of the conversation. Silverstein (1976) provides a tool for interpreting the function of honorifics and other modal morphemes. In Japanese, there are a great many modal morphemes available for marking a number of choices to index the context of speaking.

The following sentence seems simple enough. It consists of a simple proposition. Nothing more is said or implied.

(1) Taro is ill.
To convey the propositional statement (1), a speaker of Japanese must first choose between (2) and (3). (Modal elements discussed here are marked by boldface letters.)

(2) Taro ga *byooki* *da.*
    Taro *sub* ill *cop*
    'Taro is ill.'
(3) Taro *ga* *byooki* *dat* *te.*
    Taro *sub* ill *cop I hear *
    'I hear Taro is ill.'

The choice is determined by what is called the territory of the source of information (Kamio 1990). The choice reflects whether the information concerning Taro’s illness belongs to the speaker’s territory of information or not, and the choice is made in terms of the speaker’s source of the information. If she is not sure it is sufficiently within her territory for her to be absolutely sure of her facts, she must use (3). These expressions make it clear to the other conversational participants that the information does not belong to the speaker’s territory.

In the English example, the propositional portion of the sentences, namely, ‘Taro is ill’ can stand alone. In Japanese, there is no way to express the propositional content alone. Thus, *Taro ga byooki* cannot be considered a well-formed utterance, and cannot be considered the equivalent of ‘Taro is ill’. Therein lies the difference between Japanese and English, because the Japanese propositional content cannot be expressed without, at the same time, indexing both the speaker’s position with regard to the territory of the propositional information and the speaker’s attitude toward the addressee and/or the referent. This is obligatory in Japanese.

If there are many possibilities in Japanese for modal modification of the propositional content, and choices must be made, how does the speaker decide which possibility to use in each case? The choice is not a matter of volition or the feeling of the moment, but rather one of determining which would be most appropriate in the given situation.

(2) Taro *ga* *byooki* *da.*
    Taro *sub* ill *cop*
    'Taro is ill.'

---

1. The glosses used in this and following examples are: *cop* for 'copula'; *add* for 'addressee'; *hon* for 'honorary'; *aux* for 'auxiliary'; *top* for 'topic'; *sub* for 'subject.'
Having ascertained that there are a number of obligatory choices that the speaker of Japanese must make, the question becomes why this is the case. Why are there these grammatical frames that one is compelled to recognize when one speaks this language? Since there is no neutral form for the propositional content of the statement ‘Taro is ill’ in Japanese, the incorporation of one of the ending forms is obligatory. In a sense, it is almost like the need to choose between singular and plural in English, even when it makes no difference grammatically, as in ‘A book can be useful’ versus ‘Books can be useful’. It is obligatory to do so, and English speakers do so automatically. For speakers of Japanese, the outer layer of the structure of speaking Japanese is equally obligatory and automatic. It could be called a compulsory modal category, parallel to the compulsory grammatical categories of singular and plural familiar in Indo-European languages.

This is not to say that modality is expressed only at the sentence level – on the contrary, parallel phenomena abound at the discourse level as well. The ubiquitous expressions of apology and explanation are characteristic of Japanese discourse. Sugito (1983) termed these phenomena *iwire* (‘explanatory excuses’) *hyogen* (‘expressions’), i.e. ‘expressions of explanatory excuses’, and has collected and analyzed this discourse level modality found in official documents.

Expressions of explanatory excuses can be found in the setting of formal meetings. A professor, for example, at a committee meeting prefaced his second public remark with ‘I apologize for the fact that it is only I who speaks up’. With this introduction to his remark, the professor was expressing his sense of his place in the context of the committee meeting, where most of the participants had not yet spoken once, and here he was taking the floor for a second time. With his preface, he was indicating that it was not his intention to be aggressive, nor to monopolize the proceedings without concern for the feelings of the silent participants. By using the meta-pragmatic expression of explaining the excuses that indicates his awareness of the problem inherent in the situation he had caused, and by apologizing for it, potentially ruffled feathers were smoothed, and he created a situation in which he could not be blamed, and thus saved face.

The treasure trove for the examination of the question of effective communication in Japanese can be found by investigating the concepts of ingroup and outgroup (*uchi* and *soto*), and thus further attention must focus on these.

3. Japanese sense of self

The hypothesis offered here to explain the pragmatically obligatory choices expressed in modal morphemes discussed earlier may seem daring, but evidence points to the source of the layers of modality in the nature of the Japanese sense of self.

Universality has been attributed to the Western concept of an individualistic self. However, psychologists have begun to question whether this individualistic self represents a universal, or perhaps is only a feature of Western individualistic societies. Figure 4 is taken from Markus and Kitayama (1991).

![Diagram](https://example.com/diagram.png)

**Figure 4. Conceptual representations of self**

As Figure 4 shows, the East Asian self can be characterized as the interdependent self, as opposed to the independent self of Westerners. What is meant by the interdependent self is that the self changes in many ways in different contexts, depending on the closeness or distance of the participants. This is particularly noticeable in settings in which people have a close interactional relationship. In this view, the rich varieties of terms of address and self reference arose in response to the need to differentiate the multiple selves in dealing with various addressees.

Even within the context of East Asian cultures, the Japanese interdependent self has a unique structure. The interdependent self is flexible, changes easily, and adjusts to those close to it with whom it interacts. These ingroup members, *uchi*, are surrounded by a hard boundary that separates them from the outgroup, *soto*. Thus, the fragile Japanese self is protected from outsiders by this boundary between ingroup and outgroup. On the inside of this boundary, radically different behavior is expected than outside the boundary, and so, when crossing this boundary, the Japanese shift behavior patterns. Among the many signs of this shift are the linguistic markers of modal expressions, the two sets of choices that mark the inside and the outside of this boundary as we have seen.

Ever since Benedict (1946), American anthropologists have been interested in the seeming contradictions of ideas and behaviors of the Japanese. Against this background, the papers delivered in a panel at the American Anthropological
Association in 1987 entitled *Japanese Selves: Creating a Receiving Culture*, were compiled into a volume (Rosenberger ed. 1994).

Several attempts have been made to capture the intuitive knowledge of the Japanese sense of self in illustrations. One such illustration, seen in Figure 5, contrasts the construal of self of Japanese with that of Americans (Ide 1995).

![Domains of interactional self](image)

Figure 5. Structural construal of self

Following a brief explanation of how this construal of self works, evidence will be provided to show that the choice of modal morphemes indexes the speaker's interactional domains of ingroup (*uchi*) and outgroup (*soto*) toward the addressee.

In order to understand the interactional domain of the ingroup (*uchi*) and the outgroup (*soto*), it is necessary to understand its source, which becomes evident in a contrast between the Japanese and the American construal of self.

The figures should help to clarify the differences between the Japanese construal of self and the American construal of self. At the center of both of the figures is a circle, which represents in both figures the individual 'I' in 'my' world. The obvious difference is that the circle separating the Japanese 'I' from the first ring around it is porous, a dotted line, whereas the circle around the individual in the figure representing the American construal of self is a dark and continuous ring, not a porous, dotted line. In both figures, the ring, regardless of its nature, separates the individual from the ingroup, *uchi*, the group that includes family members, close friends, co-workers, teammates, and other people with whom one is close; in short, all those with whom the individual interacts frequently, and with whom relaxed and informal interaction is allowed and expected. As a rough approximation, those familiar with German might think of oneself speaking with this group as those to whom one says *du*, not *Sie*.

Outside this first circle is a second ring, indicating what is referred to in Japanese as *soto*, the outgroup. These are the people you encounter and interact with in everyday life, but with whom you have no strong bond, and they include a post office clerk, the man wearing a suit standing behind you at the post office, a physician, a secretary, a supervisor at your part-time job, a landlady, a policeman who stopped you for a traffic violation, a clerk at a department store, or a waiter at a coffee shop. (The people in this outgroup are dealt with in more detail in the people categories in Figure 6 later.) In fact, these people do not form a cohesive group, but rather they are people who are not members of your ingroup but with whom you interact in daily life. It is the interaction with members of these categories that feature all the markers of formality, including honorifics and indirect expressions, that indicate the psychological distance between the participants.

A further circle separates this outgroup from a third group, made up of those outside the outgroup. These are the people who are around you every day, but with whom you do not interact. The fact that this third group is lacking in the figure illustrating the American construal of self shows another major difference in the relationship of the self to others in the Japanese context, and it is this difference that has often made visitors to Japan feel the Japanese are impolite. Since the person who bumps into you in the crowded Tokyo subway regards you as outside her outgroup, and since etiquette requires that one be indifferent towards people in this group, she does not say 'excuse me'.

In addition to the difference in the number of rings, representing differences in basic relations between the individual and others, the boundaries of these groups are represented differently. The circle around the individual in the figure representing the American construal of the self is solid, which is intended to indicate that the self is seen as an individual, separated from all others, and constant in the sense that this self remains the same, regardless of the relationship to other interactants and the formality of the situation. It could be seen as a shell surrounding the self. The dotted circle separating the American's ingroup from the outgroup indicates that that boundary is more flowing, that any differences in behavior toward members of the one or the other group are more like points on a continuum.

The circle around the self in the Japanese figure, on the other hand, is a dotted line, indicating that the Japanese sense of self is not seen as rigidly distinct from the self of the members of the ingroup. The self is not seen as defined independently of the situation and the other participant or participants in an interaction. It is this Japanese sense of self that makes it possible to introduce the logic of *ba* (field), the characteristic of which is dual-mode thinking (Ide and Ueno 2011).
4. The choice of modal forms: The case of polite forms

The Japanese sense of self has been the subject for discussion by anthropologists and psychologists within their disciplines, but there is also a linguistic dimension to the question. Empirical data can provide a key to understanding the linguistic code switching resulting from the *uchi/soto* dichotomy, a part of the Japanese sense of self. These data concern the choice between what have come to be called polite linguistic forms and non-polite linguistic forms. The latter are by no means ‘impolite’ forms, but rather forms which do not include the markings of formal politeness.

Figure 6 presents the quantitative data. Students in Japan, the United States, and Sweden were asked to indicate which expressions they would use to request the use of a pen. The questionnaire was based on the assumption that speakers choose different expressions when addressing different people, and that these choices are not random, but rather reflect perceived differences in settings and the people approached. Thus, the students were asked how they would ask to borrow a pen from a professor, a landlady, their sister, etc., and to choose one of about 20 different suggested expressions. A little more than 500 students in both Japan and the United States and about 200 students in Sweden responded to the questionnaires, and Figure 6 shows the ‘correlation of request forms and people/situation’ taken from Hill et al. (1986).

Among the expressions investigated, only the Japanese expressions for asking to borrow a pen are clearly divided into two parts: indicated by densely dotted areas and white areas in the figure. Note that the other two countries do not have this sharp dichotomy.

What, if anything, does this mean? Since the individual expressions from 1 to 10 were chosen almost equally frequently when addressing people in the categories 1 to 13, they seem to be interchangeable. Thus, it appears that the same expressions can be used to address a professor, a well-dressed stranger, a physician, a secretary, a stranger wearing jeans, a post office clerk, the boss at work, the landlady, a police officer, a department store clerk, a younger professor, a clerk in a small store, or a waiter. It turns out that there is a common feature in the expressions from 1 to 10. These expressions are all marked by a formally polite marker: *desu* (*cop* polite), *masu* (*aux add horn*), or *kudasaru* (*receive* polite). None of the expressions 11 to 20 in the lower half of the figure, however, have such a marker. In other words, while the expressions in the upper half are marked by
Explicit, formal modal linguistic markers of politeness, those in the lower half are marked by their absence.

Looking at the horizontal line where the addressees are indicated as categories of certain people in certain settings, there is a distinct difference in the people and settings listed in the categories 1 to 13 and 14 to 20. As mentioned, the people in categories 1 to 13 include professors that belong to soto members. The people in categories 14 to 20 include a co-worker, a boyfriend or girlfriend, a close friend, an elder sibling, your mother, and a younger sibling. These are the people categorized as uchi members.

A Japanese feels intuitively that these are two distinct groups of people, divided clearly from one another by a boundary. The boundary is that between the ingroup and the outgroup, between uchi and soto. This intuitive distinction is clearly
marked linguistically by the presence or absence of formal politeness markers. It is to be noted that discerning the group to which an addressee belongs depends mostly on socially determined norms of behavior, not on the speaker’s volitional choice. This discernment is one of the basic behavior patterns of the Japanese people that is called in terms of a folk terminology wakimae (Ide 2006).

The distinction was illustrated in Figure 5, which showed the Japanese interactional self divided into the domains of *uchi* and *soto*. This could be understood as two layers of the self, and the one or the other comes to the fore depending on the interactants in a conversation. These two layers of the self are indexed linguistically by the use of formal forms or non-formal forms according to the group to which the addressee belongs. When interacting with outgroup members, the speaker presents herself as an outgroup member, using the polite linguistic forms that reflect the distance between the participants. It is this kind of pragmatic concord between the *uchi* and *soto* categories of the interactants and non-formal and formal linguistic forms that Japanese speakers have a moral feeling of wakimae, a shared cultural norm of behavior that is similar to that felt by speakers of a European language when they observe the correct grammatical concord of subject and verb.

On the other hand, when the speaker interacts with ingroup people, she interacts with the inner layer of self where she appreciates the closeness with the addressee. It is in part her switch to non-polite linguistic forms that creates the close relationship. This linguistic choice is made according to the pragmatic concord of (i) the *uchi/soto* group of the interactants, intuitively selected as a reflection of the socially determined rules of categories of people; (ii) the *uchi* and *soto* layers of the sense of self; and (iii) the choice of linguistic forms with and without polite modal morphemes to index the *uchi* and *soto* domains.

Looking again at Figure 6, does this shed light on why the Japanese expressions for requesting the use of a pen have a unique two-way pattern of language choice? It seems that Japanese have certain rules of conduct that are related to the Japanese construal of self and that are expressed by the indexing with linguistic forms of various modal morphemes. A socially well-behaved Japanese would intuitively know the dichotomous grouping of *uchi* and *soto* interactants. For example, an ordinary student categorizes her professor as an outgroup member, no matter how much the student might like her professor. This is an intuitive decision according to the accepted social norms of behavior. The student recognizes that her domain for interacting with her professor is her outer layer of her sense of self. In interacting with the professor, she uses her *soto* domain of self, and thus would control her entire behavior accordingly, dressing decently, behaving politely. One such behavior is her verbal behavior: the topics of conversation chosen, types of questions asked, and the choice of modal morphemes to index her sense of place in relation to the professor.

This discussion has centered around the formal, polite modal expressions called addressee honorifics, but that does not mean that only these linguistic expressions are governed by the interactional choice between *uchi* and *soto*. The ingroup/outgroup choice determines a great many further distinctions, such as that between direct and indirect expressions related to the territory of information, the choice between auxiliary verbs indicating the direction of benefit, the giving and receiving, the choice of speech formula, and much more. All of these choices are controlled by the pattern of behavior of *uchi* and *soto* arising from the Japanese construal of self.

Japanese honorifics have long been seen as markers of politeness. The question remains, however, as to why the use of honorifics functions as a marker of politeness. Based on the empirical data cited here, I would like to argue that the motivation for this use of addressee honorifics is to index the appropriate interactional domains of *uchi* and *soto*. This acknowledgement of the domain of interaction signals that the speaker knows her place in relation to the other participant(s) and in relation to the setting. This is not the kind of politeness a speaker chooses to employ intentionally, but an automatic or passive acknowledgement of the speaker's sense of place according to the social norm. This sense of place is what is meant by the term wakimae, a key concept in the understanding of linguistic politeness prevalent in Japanese (Ide 2002. 2005).

5. Why the Japanese speak the way they do

Now that some major differences in the way English speakers and Japanese speakers use language have been examined, the question is to why the Japanese use language this way arises. The key to appreciate the extent to which social harmony is valued in the society. One of the ways to achieve this harmony is by indexing the acknowledged interactional relationship between the participants and the formality of the setting, either by behavior as in the degree of bowing or the position of the seat or by the choice of modal linguistic forms. Regardless of the size of the group, from the family to the society as a whole, each group is intent on maintaining harmony. In pursuit of this goal, the ideas and intentions of individuals may need to be subordinated for the good of the group.

The priority given to harmony can be traced back to the rise of an indigenous Japanese philosophy between the 17th and the 19th centuries. It was at this time that, for all intents and purposes, Japan closed itself to foreign influences. This was a time when the Japanese focused on themselves and developed a philosophy suitable for an isolated people in an isolated location. Stability and maintenance of the status quo were the goals. Several Japanese philosophers developed a philosophy
of moral doctrine intended to make it possible for the Japanese people to live in peace. This philosophy has been handed down to the modern period. It sets the moral standards for how people should behave according to designated roles in society. One of the key terms used by one of the founders of the indigenous philosophy, Ogyu Sorai, is wakimae (Sorai 1717). In contrast to the Chinese orthodox Confucianism of the time, in which peace was believed to be obtained by the compliance of individuals with the principles of Chinese ethical logic of yin and yang, Sorai thought that peace for an individual is obtained by placing the peace of the society before that of the individual. Thus, for a Japanese, people's interactional relationships in terms of the social norms are the prime behavioral concern. In order for the society to stay at peace, it is imperative for people to stay in their appropriate places, exercising their expected roles. This philosophy, which may be termed the philosophy of wakimae, has been one of the major pillars of contemporary Japanese philosophy.

Now, in the 21st century, feudalism has completely given way to a democratic society, but this radical change has not affected the continuity of the cultural legacy of orientation to the group rather than to the individual. In this society in which ethical behavior is reflected in concern for others, individualistic behavior sometimes carries negative connotations.

It is against this background that patterns of language use that emphasize realizing the expectations of others according to the norms of the situation become understandable. This pattern of language use is part of the larger concept of wakimae, which literally means social norms according to which people are expected to behave in order to be appropriate in the society in which they live. The term wakimae is a noun, and Ato hito ni wakimae ga aru (that person to wakimae suru exist) means 'There is the observation of social norms in that person.' Hill et al. (1986: 348) define it as follows: "Wakimae refers to the almost automatic observation of socially agreed-upon rules and applies to both verbal and non-verbal behavior. A capsule definition would be 'conforming to the expected norms.'"

Thus, the primary goal of the speaker of Japanese is to observe and interpret the context correctly, which makes it possible to choose the appropriate linguistic forms from several sets of possibilities and express it according to wakimae, and thus to observe the social norms.

6. Concluding remarks

Speakers of Japanese are just as concerned as speakers of English that their communication be effective, but the understanding of what constitutes effective communication is different. Just as English speaking people do, Japanese speakers aim to use language so as to have a directly useful effect, and to achieve certain aims. Perhaps it should be seen as being effective indirectly. By observing wakimae through such features as the proper use of honorifics, forms of address, speech formula, the order of framing discourse, and appropriateness in terms of the occasion and the participants the effect of one's discourse on one's audience is that one is considered a sensitive and polite person and an effective communicator.

Understanding that sensitivity and politeness are the foremost features of effective communication in Japanese makes it possible to see the behavior of a student who does not speak up in a class as effective communication. By conforming to the norm of not bothering a busy and respected professor, and other students around who might have their own ideas and hesitate to speak up, is showing sensitivity and politeness. This is the presentation of linguistic politeness according to wakimae. It is in this sense that the Japanese concept of effective communication appears broader, since it encompasses silence.

It has been discussed how and why the Japanese people speak the way they do, while eloquence and persuasion are not believed to be virtues. It has been argued that the obligatory use of modal expressions to index the sense of self, wakimae, in Japanese speaking has relevance to the Japanese sense of self that consists of the uchi domain and the soto domain.

References


