A Persona-based Analysis of Politeness in Japanese and Spanish*

Akitaka Yamada[1][0000−0002−6895−7189] and Lucia Donatelli[2][0000−0002−5974−7454]

[1] Osaka University, 1-8 Machikaneyama-cho, Toyonaka, Osaka, Japan
a.yamada@lang.osaka-u.ac.jp
[2] Saarland University, Saarbrücken, Germany
donatelli@coli.uni-saarland.de

Abstract. We present descriptive accounts of 'politeness' in Japanese and Spanish by analyzing Japanese subject- and addressee-honorifics and Spanish pronominal addressee forms. Our accounts focus on inter- and intra-speaker variation in the use of these expressions. Observing this variation, we ask the question of how expressive content interacts with context. We develop a model of Bayesian Dynamic Pragmatics [23], and propose an algorithm for how the use of an politeness-oriented marker contributes to the dynamic creation of the speaker's persona or publicized self-image. Our model captures multiple pragmatic factors that impact politeness-usage and persona simultaneously, which is designed to explain comparable phenomena in other languages that employ politeness-oriented expressions.

Keywords: Bayesian Dynamic Pragmatics · Politeness · Honorifics · Persona · Real Number-based Pragmatics

1 Introduction

The principles governing 'politeness' in social interaction, or the grammatical expressions of social relations between speaker and addressee, are a current research question in formal semantics and pragmatics [15, 17, 23]. Exploring these principles in Japanese and Spanish — two well-documented languages that grammaticalize social relations and accompanying politeness in distinct manners — we propose a general pragmatic model regarding the way politeness information interacts with the context to create a specific persona of discourse participants.

In particular, our central concern lies in the source of the inter-/intra-speaker variation in usage of politeness-oriented expressions. In Japanese and Spanish, the way the speaker uses these expressions is affected by many different sociolinguistic/pragmatic factors. Certainly, as a rule of thumb, there is a general tendency regulating the use of politeness expressions. But at each time the speaker produces an utterance, they can strategically switch the forms, reflecting the multiplicity of the relevant contextual factors. These factors include age-difference, emotional engagement, formality, psychological distance, among others. As a result, the speaker's use of a 'polite' form is subject

* This work was supported by Grant-in-Aid for Research Activity Start-up [Grant Number: 20K21957 to Akitaka Yamada].

to a variation. By observing the speaker's tendency in use of such a polite form, the audience keeps learning what kind of person the speaker is. Our model intends to capture this dynamism in discourse and, we will argue, can be extended to other languages that make use of similar politeness-oriented expressions.

Our paper is structured as follows. In Section 2, we present the relevant data from both Japanese (2.1) and Spanish (2.2). In Section 3, we develop our analysis within the framework of Dynamic Pragmatics and extend it with Bayesian probability modeling to capture the specific dynamics of speaker relations [23]. Section 4 demonstrates how the concept of persona allows us to refine our model in 3 and capture the speaker intent behind politeness usage. We conclude our discussion in Section 5 with remarks on our work's theoretical implications and ideas for future studies.

2 Evidence

2.1 Japanese

Subject-honorifics and addressee-honorifics. Japanese has several honorific markers to encode the speaker's 'politeness.' For example, observe the sentences in (1).

(1) Subject-honorifics

a. *yamada-san-wa asita koogi-o* ***sur****-u.*
Yamada-Ms.-TOP tomorrow lecture-ACC do-PRS
'Ms. Yamada will have a lecture tomorrow.'

b. *yamada-san-wa asita koogi-o* ***nasar****-u.*
Yamada-Ms.-TOP tomorrow lecture-ACC do.SH-PRS
'(i) Ms. Yamada will have a lecture tomorrow;
(ii) the speaker respects the referent of the subject (= Ms. Yamada).'

Truth-conditionally, these sentences are equivalent. The second sentence, however, differs from the first one in that it uses the predicate *nasar-* 'do.SH,' in place of *sur-* 'do.' As a result, this sentence delivers secondary information that the speaker expresses their respect for the referent of the subject (= *Ms. Yamada*). Since *nasar-* obligatorily targets the subject of the sentence, it is called the SUBJECT-HONORIFIC MARKER (hereafter SH), and there are many different verbs and affixes used for subject-honorification.

Now, consider a different type of honorific marker provided in (2). Again, the sentences in (2) are equivalent in the at-issue dimension of the meaning. But unlike the first sentence, the second sentence contains an additional marker *-mas* 'AH,' with which the speaker shows their respect for the addressee. Since the target of the respect is always fixed to the addressee, it is called the ADDRESSEE-HONORIFIC MARKER (hereafter AH).

(2) Addressee-honorifics

a. *yamada-san-wa asita undoo-o sur-u.*
Yamada-Ms.-TOP tomorrow exercise-ACC do-PRS
'Ms. Yamada will do exercise tomorrow'

b. *yamada-san-wa asita undoo-o si-****mas****-u.*
Yamada-Ms.-TOP tomorrow exercise-ACC do-AH-PRS

'(i) Ms. Yamada will do exercise tomorrow;
(ii) the speaker respects the addressee.'

Differences. Traditional Japanese linguistics has treated SH and AH as being instances of the same honorific property with their difference only lying in the target. Yet, the detailed honorific meanings they encode seem to be different.

To see how, consider a case where the referent of the subject noun phrase coincides with the addressee. If these politeness meanings are regulated by the same principle, we predict that either SH and AH are both present (= (3-a)), or both absent. However, this prediction is not borne out, as seen in (3).[3] These sentences are both grammatical, and the acceptability of the sentence in (3-b) suggests that the condition in which *-mas* is used is different from the one for the SH.[4]

(3) a. *asita hapyoo-o* ***nasai-mas****-u-ka?*
tomorrow presentation-ACC do.SH-AH-PRS-Q
'(i) Are you having a presentation tomorrow?;
(ii) the speaker respects the referent of the subject (< *-nasar*)';
(iii) the speaker respects the addressee (< *-mas*).

b. *asita happyoo-o si-****mas****-u-ka?*
tomorrow presentation-ACC do-AH-PRS-Q
'(i) Are you having a presentation tomorrow?;
(ii) the speaker respects the addressee (< *-mas*).'

So, how do they differ? At the most rudimentary level, the following test in (4) serves as a nice criterion for classification. Normally, a teacher does not use the sentence in (3-a) without violating the social expectation, where they can use (3-b) felicitously. This suggests that an SH involves the speaker's assumption about the social hierarchy that the speaker has a social status lower than the target of the honorification.

(4) **Teacher-Student Test**
Can a teacher/president (someone with a higher social status) use the honorific form to a student/employee (someone with a lower social status) without intentionally violating the expectation in the society?

[3] **Avoidance of a pronoun.** In Japanese, using a pronoun to refer to the addressee is considered rude, so the sentence in (3) lacks an overt second person subject pronoun.

[4] **Real use examples.** A reviewer asked us whether there is yet another pattern acceptable in Japanese; that is, a sentence with an SH but not with an AH. Admittedly, such a sentence is not frequently observed. But as shown below, there are indeed some real use examples.

(i) *imasara mata nani-o* ***ossyar****-u.*
this pass oh boy what-ACC say.SH-PRS.
'(i) Now that things have come to this pass, oh boy, what are you saying?;
(ii) the speaker respects the referent of the subject (= the addressee).'
(http://dazai.or.jp/modules/novel/index.php?op=viewarticle&artid=108&page=54)

What makes our language description slightly more complicated is that this social convention/assumption can be strategically violated if the speaker has a good reason to do so. For example, for the aforementioned reason, normally, teachers do not use the sentence in (3-a) when talking to their students, because teachers have a social status higher than the students. But when they do, they can give the audience an impression that the they are so polite so much so that they treat the students as having a higher social status. So, teachers who wish to create a superpolite/humble publicized self-image would prefer using the sentence in (3-a).

One may wonder whether (a) the social hierarchy is the only deterministic factor, or (b) it is a dominant factor, but just one of many factors for an SH. We will put forth the argument that (b) is the correct analysis. As far as we are aware, no literature in the Japanese linguistics has proposed the former, stronger view. In addition, empirically, it is not necessary for a speaker to keep using SH markers to the same addressee, especially when the difference in social status is very small. If the social hierarchy is the only factor, the speaker obligatorily uses a SH, however small the difference may be. For these reasons, we take the latter view in (b) and elaborate our argument in Section 3.

In contrast, the addressee of (3-b) does not have to be someone whose social status is higher than the speaker. A teacher, for example, can utter this sentence to a student without violating the expectation of the social convention. Of course, a social status would be one important factor in deciding whether to use an AH, but there are many more possible motivations. Formality is one such factor. If a teacher is casually talking to a student after the class, they would use the sentence in (5), where no AHs are used.

(5) *asita* *happyoo* *sur-u?*
tomorrow presentation do-PRS
'Are you having a presentation tomorrow?'

If the same teacher is in a conference, and a similar conversation takes place in a Q and A session, the teacher would use the sentence in (3-b). In both contexts, the social relation remains the same. But the formality of the contexts is different, and when it is formal, the 'probability' of the teacher's using (3-b) is strengthened.

Again, this observation is nothing more than a general tendency, obtained at the rudimentary level. Speakers can violate the social expectation when they strategically create their own self-image, resulting in inter- and intera-speaker variations. Even in an informal context, a teacher can use the sentence in (3-b). Thus, a study of politeness-oriented expressions inevitably touches the issue of sociolinguistic/pragmatic factors and variation, and as will be seen, this is also the case in Spanish.

2.2 Spanish

Pronominal address forms. Here we consider the grammaticalization of politeness in Spanish proniminal adress forms; cf., Spanish also expresses politeness in the form of intonation [2]; prosody and gesture [5]; and discourse markers [14]. Spanish address systems differ in whether they are bipartite or tripartite in nature. Bipartite systems make use of the T-V distinction found in many languages with Latin origins, such that the singular pronoun *tú* conveys familiarity and confidence while *usted* connotes for-

mality and respect [4]. Tripartite systems employ a third singular form, *vos*, which allows speakers to convey more solidarity and/or intimacy in its use to signal more equality and/or horizontality with their interlocutors [1]. In all varieties except for Castillian (Spain) Spanish, the plural address form is invariant and adheres to the V form, *ustedes*; in Spain, the informal variant is present as plural *vosotros*.

Though usage of pronominal address forms in Spanish tends to adhere to factors that distinguish familiar (T) from formal (V) uses, usage in practice can vary depending on geographical area and particular community norms. Contemporary Spanish exhibits a preponderance towards the general usage of *tú* or *vos* (depending on geographical location), in large part due to social and political movements for equality [20]. Thus, settings that previously typically called for the use of formal addressee pronouns such as with parents, grandparents, and professors, now frequently permit the use of familiar pronouns. The use of formal pronouns persist in situations with unknown addressees who are senior in age or profession.

Relevant to our study here, the intraspeaker variation in pronoun usage has been a focus of recent discourse-based studies analyzing how speakers flout known, institutionalized norms and instead create personal identities through their use of pronouns. Fine-grained corpus investigations analyzing discourse makers, pronominal and verbal forms, and intonation patterns have demonstrated that address forms are dynamic discourse tools used to construct relevant, time-specific aspects of speaker identities. Helinks [12] in particular adopts two key theoretical notions related to this: (i) the notion of *face*, or a socially attributed and temporary aspect of self evident for the duration of the interaction based on the speaker's conscious use of politeness; and (ii) *facework*, understood as 'efforts made by the participants in verbal interaction to preserve their own face and the face of others' [22]. In other words, a speaker can choose to play with the personal pronouns that he or she uses, given that this variation remains within certain social and conversational constraints.

Variation in tripartite address systems. The ability to manipulate personal identity by varying pronoun usage is particularly evident in tripartite addressee systems such as those found in Chile, Argentina, and Uruguay. The *vos* pronoun used in certain dialects offers speakers an additional dichotomy within T forms to create a more nuanced identity and relationship with the addressee. An example of how a grandmother employs all three forms in the course of talking to her baby granddaughter is below [9]; cf., a comparable performative effect is also reported for Japanese addressee-honorifics [23] and Korean addressee-honorifics [17].

(6)

1.	*cómo está mi niña?*	'how are **you**$_{\text{U}}$ honey?'
2.	*cómo está?*	'how are **you**$_{\text{U}}$?'
	⋮	
3.	*para adónde vai cabrita*	'where are **you**$_{\text{V}}$ going young lady'
	⋮	
4.	*que eres fresca eh?*	'**you**'re$_{\text{T}}$ such a rascal, eh?'
5.	*sí*	'yes'

In the example, *vos* is used simultaneously with *tú* and *usted* to navigate authority, respect, and family relations. Gladys, the grandmother, first addresses her infant granddaughter, Gabriela, in the unexpected *usted* form to indicate familial respect; the pronoun is understood as a more affectionate and tender address than the expected *tú*, as indicated by the accompanying possessive expression *mi niña* ('my girl','honey') and use of rising intonation [7, 19]. Following the use of *ustedeante* verbal forms, the grandmother switches to *voseante* forms to indicate authority and, possibly, annoyance [21]. This is understood with the lack of possessive pronoun following the verbal form (*cabrita* 'young lady') and the fact that the address form is triggered by the granddaughter crawling away from her grandmother. This invoked authoritative figure then changes to a friskier one closer to that of the loving grandmother: an address shift from *voseante* to *tuteante* forms is evident as the grandmother playfully scolds Gabriela by telling her that she is a *fresca* 'rascal' for crawling away. Additional discursive and interactional resources contribute to this interpretation.

The example above illustrates intraspeaker variation in use of Spanish personal pronouns in a single conversation. Though constrained in a very specific and private domain, institutional contexts where both speaker and addressee actively contribute to the discourse exhibit similar variation in how speakers make use of pronouns to construct identities. For example, a policewoman taking an emergency call from a teenage girl can switch between addressing her as *usted* and *tú* to convey both the seriousness of the matter and the reassurance she hopes to provide [9]. Nevertheless, such institutional demonstrate less variation in pronominal usage, such that speakers are more constrained by public social expectations of pronimal use and do not have as much freedom to stray from these expectations.

2.3 Interim summary

In this section, we have seen the basics of Japanese and Spanish honorific and politeness systems. Certainly, these systems differ in several ways. First, Japanese uses verbal suffixes to encode honorificity, whereas politeness is encoded in the nominal domain in Spanish. Second, unlike in Spanish, Japanese has two distinct honorific systems, which are regulated by different pragmatic/social factors.

But the Japanese and Spanish systems do also have some important commonalities. First, at the coarse-grained level, there is a general consensus regarding when to use a politeness-oriented expression. Generally speaking, the use of an SH involves a social-hierarchy. Likewise, the use of *usted* generally relates to the formality/respect. For referential purposes, let us call this anticipation of expected social norms the PRIOR CONDITION. One may wish to describe this meaning as a kind of presupposition, just as we do for the meaning of phi-features; indeed, in Spanish, these same phi-features are utilized to express differing values of politeness [8]. Some survey- and interview-based methods present a static picture of form usage determined by speaker, interlocutor, and setting that is in line with such an analysis [1].

The static view alone fails to capture the expressive aspect of politeness, which is the second important property of politeness-oriented features: SH, AH and the Spanish pronouns can violate strategically the above-mentioned social anticipation, giving a performative and dynamic change in discourse and this is the source of intra-speaker

variation. Since this is an effect triggered by the use of these markers, let us call it the POSTERIOR CONDITION. Returning to the Teacher-Student Test in (4), Spanish allows the same flexibility in that a teacher may use a formal pronoun with a student temporarily to express respect. However, this switch is only a temporary one: it would be infelicitous for a teacher to repeatedly use a formal pronoun with a student if such a relationship were not previously established. On the flip side, if a teacher would like to express a more accessible, even 'hip' version of themselves to their students, the teacher can consciously begin to use a T form in conversation. Informal data collection reveals that this pattern occurs in German, too.

Finally, the use of a politeness-oriented feature is involved with a combination of multiple social and/or pragmatic factors, e.g., formality, psychological distance, social status, age difference, etc. These many factors are, however, not taken into account with the same weights; some are more prioritized than others, and the distribution of the weights is different from speaker to speaker.

(7) a. **Prior condition**: prior to the new utterance, the context expects the speaker to use/not to use a politeness-oriented expression.
b. **Posterior condition**: by (not) being produced by the speaker, a politeness-oriented expression changes the context in a certain way.
c. **Relation with social/pragmatic factors**: more than one factors contribute to the choice of the politeness-oriented form. When a language has more than two honorific systems, each system may have different weights to these factors.

Importantly, the interaction of the presuppositional and performative uses of politeness is quite distinct in Japanese and Spanish. Thus, we also need to address how a language's grammatical resources (i.e. SH/AH markers versus pronominal address forms) constrain possible interpretation of politeness.

3 A Bayesian Dynamic Pragmatics Account of Politeness

The decision-making process in which politeness-oriented form to use is 'probabilistic,' rather than 'deterministic.' For example, formality is not the only factor regulating the use of the Japanese AH. The choice between (3-b) and (5) cannot be categorically determined, because the speaker may place more weights on factors other than the formality. If the addressee is the speaker's best friend, and the speaker may wish to emphasize this psychological proximity, they may avoid using an AH. This decision process is evident in the Spanish dialogue in (6), too.

For this reason, it is appropriate to consider a combination of several pragmatic factors, following the proposal of Politeness Theory [3] and the recent literature of Dynamic Semantics/Pragmatics [15, 23]. We leave the exhaustive identification of such factors to future study. Developing the insights from the aforementioned previous studies, we rather propose a general scheme by assuming p anonymous parameters. Without losing generality, we use $f_i \in \mathbf{R}$ to denote the *i*-th social/pragmatic factor of our discourse model, and let $w_i \in \mathbf{R}$ be the weight of this *i*-th element ($i \in \{1, 2, ..., p\}$). Now,

our decision making process is dependent upon their linear combination, as shown below.

(8) $w_1 f_1 + w_2 f_2 + \ldots + w_p f_p$

As was previously mentioned, the factors involved in the Japanese SH system and the AH system are different. Thus, it is reasonable to assume that the weight vector for AH, $\mathbf{w}^a = (w_1^a\ w_2^a\ \ldots w_p^a)^T$, is different from the weight vector for SH, $\mathbf{w^s} = (w_1^s\ w_2^s\ \ldots w_p^s)^T$, and the weight parameter for the social hierarchy in $\mathbf{w^s}$ is significantly larger than that in $\mathbf{w^a}$; n.b., here, T indicates that the vector/matrix is transposed.

3.1 Dynamic Pragmatics

In what follows, we develop a model within the framework of Dynamic pragmatics that meets the desiderata given in (7). Dynamic Pragmatics is a theoretical framework that aims to capture the discourse dynamicity and the communicative effect of an utterance by articulating pragmatic principles [16]. In the recent literature of this paradigm, a discourse context is modeled to be a tuple of different information component. For example, in (9), the context (c) is modeled to be constituted by the common ground (cg), the question set (qs), and the to-do list (tdl). The effect of a command-expressing sentence is, for example, analyzed as the process of replacing an old tdl in the prior context with a new tdl' for the new, posterior context.

(9) Structured Discourse Context (Version 1 out of 3)
$c =< cg, qs, tdl >$

Model 1 (to be improved). While the model in (9) is useful especially when we examine the sentence mood, our main goal is to develop a model for the meaning associated with honorifics. With the formula in (8) in mind, let us consider the following preliminary model, in which a component p is added to (9), which is a set of tuples of a discourse participant, and the estimated weights:

(10) Structured Discourse Context (Version 2 out of 3)

a. $c =< cg, qs, tdl, p >$

b. $p = \left\{ \begin{array}{c} < alice, \mathbf{w}^a_{\text{alice}}, \mathbf{w}^s_{\text{alice}} >, < bob, \mathbf{w}^a_{\text{bob}}, \mathbf{w}^s_{\text{bob}} >, \\ \ldots < zelda, \mathbf{w}^a_{\text{zelda}}, \mathbf{w}^s_{\text{zelda}} > \end{array} \right\}$

For each individual there are weight vectors summarizing their personal idiosyncrasies in their use of honorific markers. These vectors are represented as $\mathbf{w}^a_{\text{alice}}$ (for addressee-honorifics) and $\mathbf{w}^a_{\text{alice}}$ (subject-honorifics), respectively, and in this regard, the Japanese system is assumed for (10-b). For the Spanish system, we do not need two different vectors; the following would be used in place of (10-b). Since the Spanish system is seen as a simpler version of the Japanese system, our subsequent discussions are based on the model in (10-b).

(11) $p = \{< alice, \mathbf{w}^{usted}_{\text{alice}} >, < bob, \mathbf{w}^{usted}_{\text{bob}} >, \ \ldots, \ < zelda, \mathbf{w}^{usted}_{\text{zelda}} >\}$

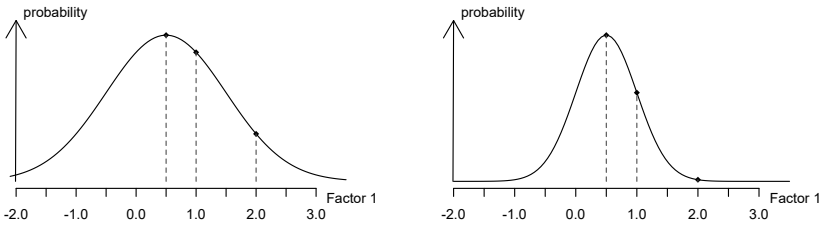


Fig. 1. Modeling uncertainties for a single factor.

To see more clearly what these vectors are, we would benefit from a concrete example. In (12), we give specific values to the weight vectors.

$$(12) \qquad < alice, \begin{pmatrix} 0.5 \\ 8 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} >$$

This represents the situation where the discourse participants all believe that 0.5 and 8 are the right parameter values for $\mathbf{w}^a_{\text{alice}}$ and 1 and 1 for $\mathbf{w}^s_{\text{alice}}$. The first element of the vector corresponds to the weight of the first factor. In this context, Factor 1 does not affect her use of addressee-honorifics that much (0.5), whereas the effect of Factor 2 is quite substantial (8). If Factor 1 represents the formality, and Factor 2 the age difference, then Alice is seen as a person who produces addressee-honorific markers towards elder people, but is not too much sensitive to the formality of the context. In this way, the vector $\mathbf{w}^a_{\text{alice}}$, or the tuple $< alice, \mathbf{w}^a_{\text{alice}}, \mathbf{w}^s_{\text{alice}} >$ represents her character (what kind of person she is) inferred by her use of honorific expressions.

Model 2. The model in (10) has a problem in practice, however: discourse participants (except for Alice) would never know the true values in $\mathbf{w}^a_{\text{alice}}$.[5] What the audience can do is to keep estimating what values are appropriate for the weights in $\mathbf{w}^a_{\text{alice}}$. It is, thus, more appropriate to think that what is stored in the structured discourse context is not the true values, but the estimated values.

The fact that we have to estimate Alice's parameters means that we have some uncertainty about the parameters. In other words, rather than pinning down a single value, we are also open to other many possible candidate values. Following the practice in Bayesian statistics, we use a probability distribution to represent our uncertainty.

To make this idea much clearer, consider how the first element in $\mathbf{w}^a_{\text{alice}}$ (i.e., the weight for Factor 1) is estimated. If we are to estimate the weight of Factor 1, all real numbers are potential candidates, and the audience assigns different probability to these real numbers. By relating a real number with its probability, we can draw curves as in Fig. 1. The x-axis represents candidate values for the weight of Factor 1, and the y-axis is for the corresponding probability assigned by the audience. Some values are more likely than others. For example, consider the curve in the left panel in Fig: 1. This curve

[5] **The speaker's ignorance.** Perhaps, Alice herself does not know the true value, either. The model we are developing here does not matter whether Alice knows the exact value or not.

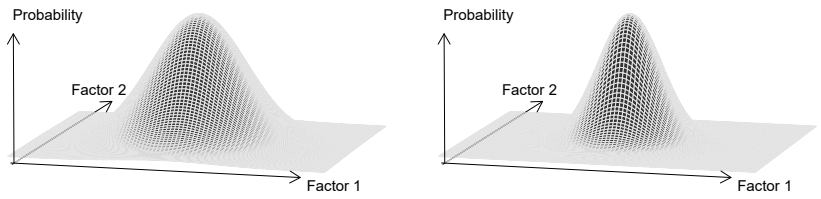


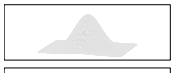
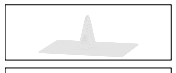
Fig. 2. Modeling uncertainties for two factors.

shows that the value 0.5 is most likely. Other values near 0.5 are also possible and are, thus, given a reasonably high probability. For 1.0, we may assign a smaller, and yet relatively high probability. But an even smaller probability is assigned to 2.0. In this way, by drawing a curve (a probability distribution), we can represent the audience's uncertainty state.

With a large amount of exposure to Alice's utterances, her audience will become more confident about certain values. Imagine, for instance, that after a few more utterances, our uncertainty represented by the left panel of Fig: 1 is updated to the one shown by the right panel in Fig: 1. The probability distribution is more narrowly distributed around 0.5, which means that the audience is more confident in this value. This means that the context update for honorific states is now seen as a change of our uncertainty state represented by a probability distribution.

For simplicity's sake, we have been concerned with the first parameter of $\mathbf{w}^a_{\text{alice}}$. But the same argument can be obtained for the other weights. When we track two factors, the uncertainty shift is modeled as a change of a 3-dimensional probability distribution, as shown in Fig. 2. A context change is an update from a particular uncertainty state (e.g., the one in the left panel) to a new state (e.g., the one in the right panel of Fig. 2).

If we take all these into considerations, it is reasonable to have probability distributions for each individual, as show in (13), unlike the single value approach, as assumed in (12).

(13) a. $< alice,$  ,  $>$

b. $< alice,$  ,  $>$

Since (for some distributions) it is possible to identify its profile by parameters. For example, for a Normal distribution, we can uniquely identify its shape of the distribution by providing the values for μ (the parameter for its center) and σ (the parameter for its variance). So, the tuples in (13) can be re-written as:

(14) a. $< alice, \left(\begin{pmatrix} 0.5 \\ 8 \end{pmatrix}, \begin{pmatrix} 1.5^2 & 0 \\ 0 & 1.3^2 \end{pmatrix} \right), \left(\begin{pmatrix} 0.5 \\ 7.3 \end{pmatrix}, \begin{pmatrix} 0.5^2 & 0 \\ 0 & 0.4^2 \end{pmatrix} \right) >$

b. $< alice, \left(\begin{pmatrix} 0.5 \\ 8 \end{pmatrix}, \begin{pmatrix} 0.3^2 & 0 \\ 0 & 1.3^2 \end{pmatrix} \right), \left(\begin{pmatrix} 0.3 \\ 7.1 \end{pmatrix}, \begin{pmatrix} 0.2^2 & 0 \\ 0 & 0.1^2 \end{pmatrix} \right) >$

From this view point, the nature of context shift is seen as an update of these vectors and matrices. The proposed model is now formally defined as follows:

(15) Structured Discourse Context (Version 3 out of 3)

a. $c =< cg, qs, tdl, p >$

b. $$p = \left\{ \begin{array}{c} < alice, (\boldsymbol{\mu}^a_{\text{alice}}, \boldsymbol{\Sigma}^a_{\text{alice}}), (\boldsymbol{\mu}^s_{\text{alice}}, \boldsymbol{\Sigma}^s_{\text{alice}}) >, \\ < bob, (\boldsymbol{\mu}^a_{\text{bob}}, \boldsymbol{\Sigma}^a_{\text{bob}}), (\boldsymbol{\mu}^s_{\text{bob}}, \boldsymbol{\Sigma}^s_{\text{bob}}) >, \\ \vdots \\ < zelda, (\boldsymbol{\mu}^a_{\text{zelda}}, \boldsymbol{\Sigma}^a_{\text{zelda}}), (\boldsymbol{\mu}^s_{\text{zelda}}, \boldsymbol{\Sigma}^s_{\text{zelda}}) > \end{array} \right\}$$

3.2 Discussion

In Section 2.3, we saw three important desiderata for the pragmatics of honorific elements. Let us now articulate how these properties are explained by our proposal.

Multiple factors. The issue of multiplicity of social/pragmatic factors is simply a matter of the length of the vector. For example, when one wishes to propose three-dimensional weight vector for $\boldsymbol{\mu}^a_{\text{alice}}$, something like the tuple in (16) would be proposed. Notice first that, although it is impossible for us to draw the probability distribution when we have more than two factors, the algebra natural extends. Second, it is also noted that the length of $\boldsymbol{\mu}^a_{\text{alice}}$ does not necessarily match the length of $\boldsymbol{\mu}^s_{\text{alice}}$. Finally, if one wishes to propose a more complicated structure among factors, one can improve the variance-covariance matrix by assuming correlations among factors.

(16) $$< alice, \left(\begin{pmatrix} 0.5 \\ 8 \\ 5.2 \end{pmatrix}, \begin{pmatrix} 0.3^2 & 0 & 0 \\ 0 & 1.3^2 & 0 \\ 0 & 0 & 0.2^2 \end{pmatrix} \right), \left(\begin{pmatrix} 0.3 \\ 7.1 \end{pmatrix}, \begin{pmatrix} 0.2^2 & 0 \\ 0 & 0.1^2 \end{pmatrix} \right) >$$

Prior context. The audience's expectation about the speaker's 'politeness' prior to the their utterance is modeled by the calculation of the formula in (8) together with the information, for example, in (16).

For example, suppose Factor 1 represents the difference in social status, Factor 2 the difference in age, and Factor 3 the formality of the given context. When the new person is her new colleague (= no difference in social status), but he is 2 years older than she is, and the conversation is about to take place in a very formal setting, then we can predict the probability of her producing an addressee-honorific marker, despite her new encounter with this addressee. Notice that we can predict how Alice behaves even before she has started talking to a new addressee. This kind of our expectation is seen as a type of 'presupposition,' because this predicted probability formally states how the discourse participants expect her to behave.

Posterior context. Of course, Alice does not have to behave as she is expected to. If she violates the audience's expectation, it is understood that their estimation is not accurate, which, as a consequence, serves as a hint for the newly-updated/estimated values for

Fig. 3. Distribution of personae

her parameters. This can be seen in (6), in which the speaker consciously violates the expected norms of pronominal use in a grandmother/infant granddaughter setting. This particular example concerns an addressee who cannot perceive this violation; it thus highlights a speaker-oriented effect of this behavior in that the speaker actively chooses to flout the pre-established conversational parameters. We return to this in our analysis in the next section.

Now consider a case where Alice behaves as expected. A context update still takes place. In this case, the audience's uncertainty is reduced: the probability distribution is more narrowly distributed. Suppose that on the first day Alice met Bob, she heard Bob use two SHs out of his three utterances. At this moment, she would vaguely infer that Bob is a relatively polite person, but she is not so certain about her inference. But a semester later, suppose she has heard 1,999 SHs out of his 2,000 utterances. Now she should be more certain about her estimation. The larger exposure to the utterances, the more certain we are about someone's persona.

4 Dynamic update as Persona learning

For each individual, the posterior mean vectors in (15-b), e.g., $\boldsymbol{\mu}^a_{\text{bob}}$ and $\boldsymbol{\mu}^s_{\text{bob}}$, reflect how the give person is analyzed by discourse participants. Let the vector $\boldsymbol{\mu}_{\text{bob}}$ represent the combined vector made of $\boldsymbol{\mu}^a_{\text{bob}}$ and $\boldsymbol{\mu}^s_{\text{bob}}$; e.g., for (14-b), $\boldsymbol{\mu}_{\text{bob}} = (0.5\ \ 8\ \ 0.3\ \ 7.1)^T$.

By comparing these vectors, we can classify individuals. Since the dimension of the vector does not affect our discussion, let us consider the simplest case, that is, the case where $\boldsymbol{\mu}_{\text{bob}}$ consists of two elements; i.e., $\boldsymbol{\mu}_{\text{bob}} = (\mu_1\ \ \mu_2)^T$. When the $\boldsymbol{\mu}$ of each individual is known, we can assign a location to each person in the 2-dimensional space, as illustrated in Fig. 3.

Seen this way, our proposal is understood as a natural extension of Conceptual Spaces framework for lexical semantics [10, 11], or the studies working on Persona-based semantics/pragmatics [6, 15]. Putting aside the theoretical choice of partitioning the space in Fig. 3, we can easily see that Alice and Bob are different from the other individuals in the way they choose honorific forms. For instance, if μ_1 is the age difference, and μ_2 is the formality, Alice and Bob do not put strong priority in age difference, but they are people so sensitive to the difference in formality. In other words, we can identify a behavioral pattern for the group of Alice and Bob, a type of personal charac-

teristics. On the other hand, we can also find a cluster for the other individuals. In terms of the size of the group, this could be seen as a 'mainstream' persona (cf., [6]).

Upon this view, mainstream and anti-mainstream personae are not fixed single locations, but seen as emergent regions that become apparent only after we have estimated many people's locations. We thus predict that a mainstream persona can change, which is borne out by a well-known fact that that an existing politeness-oriented expression gradually becomes analyzed less 'polite' and a new respect-expressing strategy gets utilized to encode a high-level politeness. To create a superpolite publicized self-image people try to deviate from the mainstream location. The synchronic deviation in Fig. 3 is seen as a symptom of an emergent, diachronic language change.

5 Conclusion and theoretical implications

In this paper, we have described how Japanese and Spanish 'politeness'-oriented expressions are associated with sociolinguistic/pragmatic factors, and then, we have developed a pragmatic model explaining the inter-/intra-speaker variation and their prior/posterior relation with the context. In the presented model, the audience is invited to infer the persona of each speaker by observing how they take the contextual factors into consideration when they make a decision regading the use of a 'politeness'-oriented expression.

Although we have chiefly examined the data from Japanese and Spanish, we believe our proposal can be extended to account for data in other languages. The following anecdote from a native German speaker serves as an example (Arne Köhn, personal communication). The speaker, a man in his mid-thirties, went to the barbershop to have his hair cut. His barber was male and appeared to be the same age. The two men spent the duration of the haircut speaking in the third-person (equivalent to English 'one') because neither of them was confident to commit to a formal (*Sie* 'you') or familiar (*du* 'you') address form and corresponding relationship. The speaker relayed this story still with ambivalence about how he should have acted. Anecdotes from other German speakers reveal similar hesitation about committing to certain personal identities, instead opting to be publicly agnostic about their own personalities. Most speakers defer to the formal *Sie*, given that respect is a good default but also, notably, because this form allows the speaker more 'distance' from the addressee in the conversation. In German and other T/V languages (Spanish included), speakers may also use a back-off strategy to avoid committing to a specific singular address form by using a plural address form. Under our analysis, the fear of revealing one's personality is understood as an avoidance of locating oneself in the persona space as shown in Fig. 3.

We predict that some languages place larger/smaller weights on social hierarchy than Japanese and Spanish. With appropriate data sets, future studies can identify the mainstream use in each language, making us easily detect the variation among language communities.

It is also important to examine how each language encodes honorificity. Honorificity can be encoded in the verbal domain (Japanese), and in the nominal domain (Spanish). Is the pragmatic effect we have discussed read off from the same feature in the syntax? If so, where does the feature exist? Is this a feature provided in the NP-periphery, or is it

a clause-level feature?[23] Detailed examinations of the logical forms will surely allow us to promote our understanding of politeness-oriented phenomena across languages.

References

1. Bishop, K., Michnowicz, J.: Forms of address in Chilean Spanish. *Hispania* **93** (3), 413-429. (2010)
2. Borrás-Comes, J., Sichel-Bazin, R., Prieto, P.: Vocative intonation preferences are sensitive to politeness factors. *Language and Speech 58.1*, 68-83. (2015)
3. Brown, P., Levinson, S.: *Politeness: some universals in language usage*. (1987 [1978])
4. Brown, R., Gilman, A.: The pronouns of power and solidarity. In: Sebeok, T. A. (ed.) *Style in Language*, pp. 253–276. MIT Press, Cambridge, Mass. (1960)
5. Brown, L., Prieto, P. (Im)politeness: Prosody and gesture. In *The Palgrave handbook of linguistic (im) politeness*, pp. 357-379. Palgrave Macmillan, London. (2017)
6. Burnett, H.: A persona-based semantics for slurs. *Grazer Philosophische Studien* **97**, 31–62. (2020)
7. Carricaburo, N.: Las fórmulas de tratamiento en el español actual, second ed. Arco Libros, Madrid. (2015)
8. Donatelli, L. *The morphosemantics of Spanish gender: evidence from small nominals.* Ph.D. thesis, Georgetown University. (2019)
9. Fernández-Mallat, V.: Forms of address in interaction: evidence from Chilean Spanish. *Journal of Pragmatics* **161**: 95-106. (2020)
10. Gärdenfors, P.: *Conceptual spaces: the geometry of thought*. MIT press, Cambridge, Mass. (2000).
11. Gärdenfors, P. *The geometry of meaning: semantics based on conceptual spaces*. MIT Press, Cambridge, Mass. (2014).
12. Helincks, K.: Negotiation of terms of address in a Chilean television talk show. *Bulletin of Hispanic Studies 92.7*: 731-753. (2015)
13. Henderson, R., McCready, E.: Dogwhistles, trust and ideology. In: *Proceedings of the Sixteenth International Workshop of Logic and Engineering of Natural Language Semantics* **16** (LENLS16), pp. 1–2. (2019)
14. Landone, E.: Discourse markers and politeness in a digital forum in Spanish. *Journal of Pragmatics* 44.13: 1799-1820. (2012)
15. McCready, E.: *Honorification and social meaning*. Oxford University Press, New York. (2019)
16. Portner, P.: Commitment to priorities. In: Fogel, D., Harris, D., & Moss, M. (eds.) New work on speech acts, pp. 296–316. Oxford University Press, Oxford, (2018)
17. Portner, P., Pak, M., Zanuttini, R.: The speaker-addressee relation at the syntax-semantics interface. *Language* **95**, 1–36 (2019)
18. Potts, C.: *The logic of conventional implicatures*. Oxford University Press, Oxford. (2005)
19. Prieto, P., Borrás-Comes, J., Crespo-Sendra, V., Thorson, J., Vanrell, M.M.: Entonación y pragmática en los enunciados interrogativos absolutos del espa nol en un corpus de habla dirigida a niños. Oralia 14, 227-255. (2011)
20. Real Academia Española (RAE). *Nueva gramática de la lengua espa nola manual.* Espasa (2010).
21. Rivadeneira Valenzuela, M.: Sociolinguistic variation and change in Chilean *voseo*. In: Moyna, M.I., Rivera-Mills, S. (Eds.), *Forms of Address in the Spanish of the Americas.* John Benjamins, Amsterdam, pp. 87-117. (2016)
22. Watts, R. J.: *Politeness.* Cambridge, Cambridge University Press. (2003)
23. Yamada, A.: *The syntax, semantics and pragmatics of Japanese addressee-honorific markers,* PhD. Thesis, Georgetown University, Washington D.C. (2019)