Japanese honorifics as emotive definite descriptions

Christopher Potts and Shigeto Kawahara
University of Massachusetts, Amherst

1. Introduction

Harada (1976) identifies honorifics as one of "the salient features of the Japanese language" (p. 500) and provides an extensive overview of pressing syntactic issues surrounding this system. But he remains noncommittal about the semantic contribution of these morphemes; the question of what honorifics mean, what they contribute to a discourse, is left unanswered. The present paper offers an answer to this question, in the form of a set of composition rules and interpretation principles that, taken together, can properly distinguish honorific-free forms like (1a) from their honorific and antihonorific counterparts ((1b) and (1c)).

(1)   a. Sam-ga   warat-ta.
             Sam-NOM laugh-PAST

          i. 'Sam laughed.'

     b. Sam-ga   **o**-warai-**ninat**-ta.
             Sam-NOM **subj.hon**-laugh-**subj.hon**-PAST

          i. 'Sam laughed.'
          ii. 'The speaker honors Sam.'   [subject honorific]

     c. Sam-ga   warai-**yagat**-ta.
             Sam-NOM laugh-**antihon**-PAST

          i. 'Sam laughed.'
          ii. 'The speaker views Sam negatively.'   [antihonorific]

We've glossed each of these examples using two separate sentences, a practice we often employ. It represents a specific theoretical claim, namely, that honorifics contribute a meaning that is independent of the content of the sentence containing them. Thus, honorific-containing sentences manifest a special kind of multidimensional semantic content in the sense of Bach 1999 and Potts 2004. It should be noted, though, that such glosses imply too strong a connection between honorifics and the realm of propositional content. We contend that honorifics do not contribute propositional meanings, but rather that they function as a special kind of definite description, in a sense that we adapt from the work of Karttunen (1976) and Heim (1982).

      We construe the term 'honorific' broadly, and we do not attempt to delimit the class. The best known kind of honorifics are the argument-oriented honorifics (called 'propositional honorifics' in Harada 1976). Example (1a) is a subject-honorific; (2) provides a typical instance of object honorification (Boeckx and Niinuma 2004; Bobaljik and Yatsushiro 2004).

(2)   Kathyrn-wa Sam-o   **hai**-ken-shi-nakat-ta.
        Kathryn-TOP Sam-ACC **obj.hon**-see-do-not-PAST

     i. 'Kathryn did not see Sam.'

    ii. 'The speaker respects Sam.'   [object honorific]

We divide our attention between these cases and the honorifics represented in (3) and (4), which associate semantically with some aspect of the propositional content rather than with the denotation of an argument nominal.

(3)   Mary-ga   ringo-o   tabe-**mashi**-ta.
        Mary-NOM apple-ACC eat-**perf.hon**-PAST

     i. 'Mary ate the apple.'

    ii. 'I am speaking nicely to you.'   [performative honorification]

(4)   John-wa [Mary-ga   nesugoshi-**chimat**-ta]   -koto-o shitteiru.
        John-TOP Mary-NOM oversleep-**antihon**-PAST -fact   know

     i. 'John knows that Mary overslept.'

    ii. 'It sucks that Mary overslept.'   [antihonorification]

The first is called *performative honorification* in Harada (1976) (certainly a suggestive name). It goes by the name 'polite speech' as well. We call the second kind *antihonorification*. Here, the morpheme *chimau*, which appears as *chimat-*, signals that the speaker has contempt for the proposition expressed by the clause in which it appears.

There is considerable precedent for treating antihonorifics as part of the same set of expressions as that represented in (1a), (2), and (3). For instance, Tokieda (1940) and Sakuma (1940) discuss them at length, and Tsujimura (1978) stresses their importance for the theory of honorifics. We elaborate on these connections with traditional and descriptive grammar below. It was a pleasant surprise to us to find echos of our formal proposal in the work of traditional Japanese grammarians.

2. Expressives

We aim to situate the study of honorifics in the theory of expressive content (Zimmermann 1991; Kaplan 1999; Kratzer 1999; Potts 2003a, 2004). Other expressive content operators are expressive attributive adjectives like *damn* (*the damn Republicans*), the content of epithets, certain interjections (*ouch*), and probably a large subset of the class of discourse particles; see Potts 2004:§5 for an overview and additional references.

We identify the following central properties of expressives:

(5)   i. Nondisplaceability

             ii. Independence

             iii. Immediacy

             iv. Descriptive ineffability

The idea is that all and only the words and constructions with the conjunction of these properties are expressives and therefore should be analyzable in terms like those of the present work.

The connection with expressive content in some sense predates our work. For instance, Tsujimura (1978) notes that honorifics and discourse particles share important features. The properties in (5) can be regarded as an attempt to substantiate claims of this form.

Over the course of the next few subsections, we review each of the properties in (5), attempting to show how they are relevant to understanding honorifics.

2.1. *Nondisplaceability*

Honorifics tell us about the *speaker's* beliefs in the *utterance situation.* Thus, they never end up in the semantic scope of any operators. Example (4) illustrates: it is the speaker who expresses displeasure at Mary's oversleeping. This displeasure cannot be attributed to John (the matrix subject), and no reading of the sentence entails that John knows that the speaker is displeased about the oversleeping.

We regard this example as particularly important because it shows that honorific content is not 'plugged' by attitude predicates like *know.* Karttunen (1973) famously identified plugs as different from the presupposition holes (e.g., negation) because presuppositions do not project up through them, at least as far as the semantics is concerned. (Conversational implicatures can give rise to the sense that they have projected, but such inferences are significantly more malleable than presuppositions.) This is by no means the only contrast we identify between expressives and presuppositions, but it is an important one, since it impacts how these meanings are defined (lexicalized).

It is worth noting, though, that honorific content does project up out of all the standard presupposition holes. We illustrate with negation:

(6)   Jim-wa   **o**-warai-**ninar**-anakat-ta.
         Jim-TOP **subj.hon**-laugh-**subj.hon**-not-PAST

             i. 'Jim did not laugh.'

             ii. 'The speaker respects Jim.'   [subject honorific]

The facts are parallel for antihonorific plural markers like *domo*, which apply to nouns to dishonor their denotations (in a sense):

(7)　　ano sensei-wa 　[gaki-**domo**-ga　　urusai] -koto-o shir-anai.
　　　　 that teacher-TOP　kids-**antihon**-NOM noisy　-fact　　know-not

> i. 'That teacher does not know that the kids are (annoyingly) noisy.'
>
> ii. 'I don't like the kids.'

The contempt expressed by *domo* is in no way negated by the negation *nai* in the matrix clause.

2.2. *Independence*

Sentences containing honorifics are multidimensional in the sense that each honorific contributes a meaning that is independent of the meaning of the main clause. As noted in section 1, our glosses reflect this, in the sense that they are small discourses. Japanese descriptive grammarians have often noted this property of honorifics; we find this idea in the work of Tokieda (1940:44) and Ooishi (1975:74), and it receives particularly clear expression in Kikuchi 1994, where honorifics are identified as a special case of a broader class of 'Taiguu Hyoogen' (Attitudinal Expressions), which have the property specified in the quotation in (8).

(8)　　"For a set of expressions with honorifics and antihonorifics derived from one neutral sentence $E_1, E_2, \ldots, E_n$, there is a basic and core meaning M, but different attitudinal expressions $A_1, A_2, \ldots, A_n$.

$$
\begin{aligned}
E_1 &= M + A_1 \\
E_2 &= M + A_2 \\
&\vdots \\
E_n &= M + A_n
\end{aligned}
$$

"M is the same in the sense that as long as truth value is concerned, M is invariant."　　　　　　　　　　　　　　　　　　　(Kikuchi 1994:22-23)

We see this independence clearly when we look to examples involving quantifiers, as in (9).

(9)　 ?? hotondo no kyoojyu-ga　　[sono kurasu-o
　　　　 most-of　　professor-NOM　that　class-ACC
　　　　 o-oshie-**ninat**-ta]　　　　　　　-to　omotteiru.
　　　　 subj.hon-teach-**subj.hon**-PAST that believes

* 'Most professors believe that they, who I honor, taught that class.'

Our English gloss suggests why the example fails. The honorific is a kind of nonrestrictive modifier on the subject pronoun (which is silent). It is well known that nonrestrictive modifiers cannot associate with elements that are quantification or bound by quantified elements. It is for this reason that the honorification causes the sentence to be ungrammatical.

In our theory, the embedded clause's subject translates as a variable that necessarily remains free in the honorific meaning, even though it is bound by the matrix quantifier. Thus, to obtain a coherent reading of (9), speakers must restrict attention to assignments that take the variable to a denotation that is equivalent to that of *the professors who believe they taught that class.* The need for this extra-grammatical repair accounts for the fact that speakers judge the very best examples of this form to be highly questionable. Section 6 fleshes out these informal comments about the theory.

2.3. *Immediacy*

Honorifics achieve their intended act simply by being uttered; they do not offer content for inclusion into the common ground so much as inflict content upon it. In this sense, they are performative. This is the link with the theory of performatives, which have a self-verifying semantics (Kamp 1978; Krifka 2004; Potts 2003b).

This property too makes an appearance in more traditional descriptions. For instance, Tsujimura (1978) explicates honorific content by drawing parallels with widely accepted speech-act categories:

(10)   "expressions such as commands, prohibitions, or wishes clearly establish a relationship with the interlocutor, and hence should be treated from the attitudinal viewpoint [just like honorifics]"   (Tsujimura 1978:223)

In the theory we are working towards, a sentence containing an honorific has identical static content as that of the comparable sentence with the honorific removed. The two sentences also have identical primary speech-act force (or potential). But the sentence containing the honorific defines a secondary speech act, one that derives from the presence of the honorific. We illustrate with the following variations on a basic imperative.

(11)   a. iikagen   ni shite!   [command]
             good-extent   do.IMP
             'Knock it off!'

         b. iikagen   ni shite-**kudasai**!   [command; positive honorification]
             good-extent   do.IMP-**hon**
             'Knock it off (I am asking you nicely)!'

         c. iikagen   ni shi-**yagare**!   [command; antihonorification]
             good-extent   do.IMP-**antihon**
             'Knock it off (I am already annoyed)!'

Example (11a) typically has the force of an imperative. Example (11b) typically has imperative force plus positive speaker-to-addressee honorification, and (11c) typically has imperative force plus negative speaker-to-addressee honorification. (We hedge with 'typically' because the relation between clause-type and speech-act-type is extremely complex; Zanuttini and Portner 2003; Truckenbrodt, to appear.)

2.4. *Descriptive ineffability*

Speakers are never fully satisfied when they paraphrase honorifics. This suggests that their content is not propositional. In (12), we illustrate this indeterminacy with some antihonorific uses:

(12)   *chimau*:   i.   attaches to a verbal stem;
                       ii.   emphasizes the completion of the action, or
                       iii. expresses the speaker's displeasure with the action.

    a.   nesugoshi-ta.   ⇒ nesugoshi-**chimat**-ta.
         oversleep-PAST   overslept-**antihon**-PAST
         'I overslept.'   'I overslept, which sucks.'

    b.   mi-rare-ta.   ⇒ mi-rare-**chimat**-ta.
         see-pass-PAST   see-pass-**antihon**-PAST
         'I was seen.'   'I was seen.' ('I wasn't supposed to be seen.')
                             ('This might have bad consequences.')
                             ('I'm unhappy that I was seen.')

Consider what this means for second-language learners and their teachers. The teachers cannot provide paraphrases, as these are bound to fall far short of, or at best approximate, the actual contribution of these morphemes. They cannot provide translations, because many language seem to lack anything like the complex honorific system of Japanese, and those that do (Korean, Thai) divide this lexical class up in significantly different ways. Thus, teachers must describe typical situations in which one uses honorifics, hoping that such examples will allow students to generalize to the full class of appropriate environments. As a result, honorifics present one of the greatest obstacles to learning Japanese as a second language.

Harada (1976) glosses most argument-oriented honorifics as saying 'x is socially superior to the speaker', where x is the argument to the honorific's meaning. This is reasonable for many cases. But one must be careful not to assume that 'socially superior to' is actually the relation involved. Honorific contributions are manifestly different from this. For instance, the 'socially superior to' relation is transitive: if x is socially superior to y and y is socially superior to z, then x is socially superior to z. But this does not hold for honorifics, where we can easily have situations like the following:

6
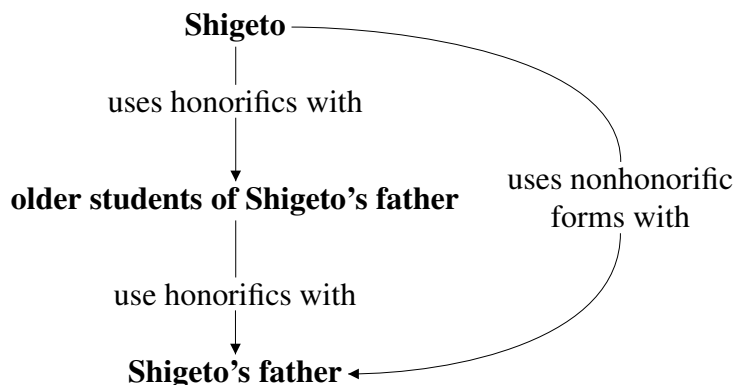
Kikuchi (1994) lists *fifteen* factors that potentially affect the use of honorifics. Some are social and some are psychological.

3. Semantic composition

Potts (2004) establishes a range of logical and conceptual connections between the above definition of expressive (including honorific) content and *conventional implicatures* as they were originally specified by Grice (1975). We continue to think that the connection is robust, and the logical theory that we develop here can be seen as an extension of Potts's.

3.1. *An expressive type*

We want to control where expressives appear in our semantics and what kind of expressions they interact with. The descriptive properties discussed above indicate that we basically want to remove their meanings from the semantic composition. In order to do this, we propose a partially special type theory:

(13)   i. e and t are regular types.

ii. ε is an expressive type.

iii. If σ and τ are regular types, then $\langle \sigma, \tau \rangle$ is a regular type.

iv. If σ is a regular type, then $\langle \sigma, \varepsilon \rangle$ is an expressive type.

v. Nothing else is a type.

If we remove clauses (ii) and (iv), we obtain the usual type theory employed in semantics, with basic types for entities and propositions and a type-constructor, here $\langle \, , \, \rangle$, for forming functional types. So the only additions are the two provisions for expressive types. We first define a special basic type ε (clause (ii); the type ε can be identified with the type e^c of Potts 2004, which goes unused there.) In clause (iv), we allow ε to enter into a limited class of functional types: it can be the output of any type with a regular type as its input. Thus, we have types like $\langle e, \varepsilon \rangle$ and $\langle \langle e, t \rangle, \varepsilon \rangle$. But we do not have types $\langle \varepsilon, \sigma \rangle$, for any σ, nor do we have anything like $\langle \langle e, \varepsilon \rangle, t \rangle$, where a functional expressive type is not an output type in the expression as a whole. (Note that $\langle e, \varepsilon \rangle$ is an expressive type. By clause (iv), it inherits its expressivity from the inclusion of ε as a subtype.)

The guiding intuition here is that the types help us to delimit the space of possible meanings. Thus, simply by translating honorifics as functions denoting somewhere in the type space $\langle \sigma, \varepsilon \rangle$, we are in a position to keep a tight grip on how their content is distributed. For instance, an immediate result of this definition is that nothing ever scopes over an honorific's content. Such an operator would have to have a type that took an expressive type as its input. But we have no such types, by definition (13). This is the heart of the explanation for the *nondisplaceability* property discussed above in section 2.1.

3.2. *A composition rule*

We employ a single new composition rule to model the multidimensional nature of expressive content:

(14)
$$\alpha : \sigma$$
$$,$$
$$\beta(\alpha) : \varepsilon$$
$$\beta : \langle \sigma, \varepsilon \rangle \quad \alpha : \sigma$$

Here, α and β are expressions of a meaning language, with their logical types given after the colons.

Suppose the expression β is the honorific, with α its regularly typed argument. The first thing to note is that this rule passes α on to the root node without doing anything to it. An identity function. But it uses α in a second way as well: the functor β applies to α to return an expression $\beta(\alpha)$ that has an expressive type. So the value on the root of this subtree is multidimensional, in the sense that it denotes both α and $\beta(\alpha)$. These are *independent* expressions of our logic, and thus no other expression can apply to the pair of them. But there is also no indeterminacy about how composition would proceed from here. There are no expressions in $\langle \varepsilon, \tau \rangle$, for any τ, so the only fair target for functional application is α itself. Similarly, it cannot be the case that α is a functional expressive type, because that would mean that the type σ is something of the form $\langle \tau, \varepsilon \rangle$. But the type $\langle \langle \tau, \varepsilon \rangle, \varepsilon \rangle$ is not a member of the set specified in (13). Thus, there is a unique continuation of any subtree of the form in (14).

This rule alone cannot constitute the logic of composition; we refer to Potts 2004:§3 for a more complete set of rules. But we claim that this rule does suffice for describing the particulars of honorific composition.

4. Denotations

The semantic work of section 3 was essentially syntactic. We defined the space of possible semantic translations (the set of types in (13)) and an inference rule (the admissibility condition in (14)). These are merely tools for highlighting honorifics' special semantic contribution, to which we now turn our attention.

4.1. *An interpretation rule*

The parsetree admissibility condition defined in (14) isolates the ε-type expressions from the rest of the composition. This has the virtue of capturing the independence property discussed in (2.2), but it requires us to say something special to ensure that these expressive meanings form part of the interpretation of the whole. The rule in (15) provides a simple and direct method for doing this:

(15)   The interpretation of a parsetree $\mathcal{T}$ is the tuple $\langle A, \mathbf{B} \rangle$, where

    i. A is the semantic value of $\mathcal{T}$'s root node; and

    ii. $\mathbf{B}$ is the set consisting of all and only the interpretations of the type ε expressions in $\mathcal{T}$.

Now that we have included the ε-type expressions in the interpretation scheme, we need to determine what they mean. For this, we venture a substantive connection with the theory of definite descriptions.

4.2. *Honorific definite descriptions*

In the dynamic semantics anticipated by Karttunen (1976) and developed by Heim (1982, 1983) (see also Dekker 1994), definite nominals register something about the current context, namely, that it contains (supports) a discourse referent with the requisite property. Definite descriptions must be *familiar*.

    Honorifics are also subject to a familiarity condition. Just as a definite description is felicitous only in situations in which an entity with its descriptive content has already been introduced, so too an honorific is felicitous only in discourse situations in which it is already established that the speaker bears the appropriate relation to his addressee (in the case of performative honorifics) or to the denotation of the relevant argument (in the case of argument-oriented honorifics). For instance, a speaker of (16) indicates that the input context is one in which he should honor (speak politely with) his addressee.

(16)   Mary-ga   ringo-o   tabe-**mashi**-ta.
        Mary-NOM apple-ACC eat-**perf.hon**-PAST

    i. 'Mary ate the apple.'

    ii. 'I respect you.'

So we propose a theory of honorific denotations that is largely adapted from the theory of discourse anaphora. The heart of the proposal is a special set $EMOTION$ of emotions:

(17) $EMOTION$ is the set of real numbers in the interval $[-1, 1]$.

Bill Ladusaw first suggested to us that the reals might provide the appropriate space in which to define expressive meanings. Throughout this paper, we employ the following very rough metalogical interpretation of $EMOTION$:

(18)   a.   Antihonorific forms (impolite speech): $[-1, -.5]$

    b.   No honorific marking: $-.5 < \mathbf{r} < .5$

    c.   Positive honorific forms: $[.5, 1]$.

We have all the gradations we could possibly want. This freedom enables us to make good on an important insight from Cruse (1986):

(19) "presented [regular] meaning is for the most part coded digitally — that is to say, it can vary only in discrete jumps; expressive meaning, on the other hand, at least in respect of intensity, can be varied continuously, and is therefore analogically coded."   (Cruse 1986:272)

The set $EMOTION$ is essential to our theory of honorifics, but it is not, in itself, enough to yield the interpretations we are after. The targets of honorification can vary widely: argument-oriented honorifics permit speakers to honor individuals they are talking about, and performative honorifics permit honorification of the addressee. The same range of targets is available to antihonorifics. So honorific denotations needs to provide roles for the speaker, a member of $EMOTION$, and the honored individual. We achieve this with clause (iv) of the definition in (20); the other clauses define (entirely standard) denotation domains for our other types.

(20)   i. The domain of type e is the domain of entities, D_e.

   ii. The domain of type t is the set of all propositions, D_t, the power set of the set of worlds.

   iii. The domain of a functional type $\langle \sigma, \tau \rangle$ (regular or expressive) is the set of all functions from D_σ to D_τ.

   iv. The domain of type ε is D_ε, the set of all triples arb, where $a, b \in D_e$ and $r \in EMOTION$.

Thus, ε-type expressions like SUBJ-HON(**john**) denote triples: $a.\mathbf{756}[\![\mathbf{john}]\!]^c$ and the like. These triples do not have force of their own. In the next section, we show how they enter into contexts in such a way as to have potentially significant effects.

4.3. *Contexts*

Broadly speaking, we need to ensure that contexts satisfy two criteria: they must contain enough honorific triples to define the speaker's attitudes towards the people he might talk with and about, and they must be consistent in their honorific content. It is straightforward to achieve these goals by defining a restriction on the theory of contexts in (21), which extends that of Kaplan (1989).

(21)   A context is a tuple $c = \langle c_A, c_P, c_T, c_W, c_{HON} \rangle$, where

   i. c_A is the agent (speaker) of c;

   ii. c_P is the place of c;

   iii. c_T is the time of c;

   iv. c_W is the world of c; and

   v. c_{HON} is a subset of D_ε.

One must, of course, delimit the range of possible contexts in various ways, so that they behave as one wants them to (Kaplan 1989:§18). For honorifics, we identify just one important axiom:

(22)   A context c is admissible only if c_{HON} contains exactly one triple $c_A \mathbf{r} b$ for every contextually salient $b \in D_e$.

This definition ensures that the context's agent (speaker) bears an honorific attitude to each contextually salient entity in the domain. It also ensures honorific consistency, by forbidding contexts in which c_{HON} contains, for instance the triples $a0b$ and $a1b$. This is just to say that no one can 0-honor and 1-honor the same individual in the same context.

    The final step is to connect the above notion of context with our ε-type terms. We do this via the definedness condition in (23).

(23) $[\![\alpha : \varepsilon]\!]^c$ is defined only if $[\![\alpha]\!]^c \in c_{HON}$

    Where defined, $[\![\alpha : \varepsilon]\!]^c \in D_\varepsilon$

4.4. *Some examples*

Example (24) is our semantic parse for an example involving a subject-oriented honorific.

(24)   a. Sam-ga   o-warai-**ninat**-ta.
        Sam-NOM **subj.hon**-laugh-**subj.hon**-PAST

        i. 'Sam laughed.'
        ii. 'The speaker honors Sam.'   [subject honorific]

    b.   **laugh**(**sam**) $: t$

            sam $: e$,   **laugh** $: \langle e, t \rangle$
            SUBJ-HON(**sam**) $: \varepsilon$

        SUBJ-HON $: \langle e, \varepsilon \rangle$ **sam** $: e$

Parsetree interpretation rule (15) instructs us to interpret two nodes in this tree: the root node and the ε-type expression SUBJ-HON(**sam**). Assume that $[\![\text{SUBJ-HON}]\!]^c$ is a function that takes any entity a into the triple $c_A 1 a$. This is maximal honorification in the system suggested by (18). Thus, the above tree is defined for a context c only if the set c_{HON} contains $c_A 1 a$. The exact nature of $[\![\text{SUBJ-HON}]\!]^c$ is not of direct concern to us here. What is important is that we have captured the defining properties of expressives as reviewed in section 2.

Independence   In (24b), the regular semantic value of the root node is logically and compositionally independent from the honorific marking. It is true that the honorific partializes the interpretation in the sense that it renders some contexts inadmissible as parameters on the interpretation function. But the honorific does not modify or restrict the proposition expressed.

    One can see independence at the level of the models as well: the honorific carves out a set of admissible contexts, and the root node's value carves our a set of admissible possible worlds.

Nondisplaceability   We ensure nondisplaceability in essentially the same way that Kaplan (1989) ensures that English indexicals always receive primary scope: the denotations for our honorifics make direct appeal to the context parameter on the interpretation function, which cannot, in our system, be manipulated by other expressions. This limitation is reflected in our type theory.

Since Kaplan's article, linguists have uncovered problems with Kaplan's rigid view of the logic of indexicals; we refer in particular to Schlenker 2003 and Anand and Nevins 2004. For now, we assume that expressives are rigid in the way that Kaplan suggested for indexicals, but we should be on the watch for complications. We return to this issue in section 7.

Immediacy   Honorifics, as defined here, do not directly impact the propositional semantics at all. Rather, they act on the context. This links them directly with the utterance situation, which is presumably also the locus of speech-acts like assertions and commands (Krifka 2001, 2004). Thus, although we cannot offer a theory of speech-acts in this paper, it is easy to see how the account could be closely allied with such a theory.

Descriptive ineffability   At its heart, descriptive ineffability says that speakers will stammer and hedge if pressed for a propositional paraphrase of their honorifics (with all expressives, in fact). Our system captures this by assigning honorifics non-propositional denotations. A demand for a paraphrase is, in the present theory, a demand for a paraphrase *of a triple*, which makes essentially no sense.

It is worth looking at one more example-type, this one involving Korean. Korean is useful at this point because its argument-oriented honorifics can be realize morphologically on both the targeted nominal and on the verb. Thus, we often get two chances to see the same honorific relation expressed in the same clause. The examples in (25) come from Pollard and Sag 1994:92; the honorific triples below certain phrases indicate likely ε-type denotations for them.

(25)   a.   Kim sacang-**nim**-i   o-**si**-ess-ta.
            Kim president-**hom**-DECL come-**hon**-PAST-DECL
                $c_A 0 [\![\textbf{kim}]\!]^c$ $c_A 0 [\![\textbf{kim}]\!]^c$
         'President Kim has come.'

       b. # Kim sacang-i   o-**si**-ess-ta.
             Kim president   come-**hon**-PAST-DECL
                $c_A 0 [\![\textbf{kim}]\!]^c$ $c_A 1 [\![\textbf{kim}]\!]^c$
          'President Kim has come.'

       c. # Kim sacang-**nim**-i   o-ess-ta.
             Kim president-**hon**-NOM come-PAST- DECL
                $c_A 1 [\![\textbf{kim}]\!]^c$ $c_A 0 [\![\textbf{kim}]\!]^c$
          'President Kim has come.'

The first example is consistent in its honorific marking. It is unproblematic. The next two examples have inconsistent honorific marking. Thus, there is no single context that we can use to interpret these sentences; any context we pick will violate the honorific consistency condition that follows from (22).

5. Sketch of a dynamic treatment

Perhaps more than any other area of semantics, the theory of anaphora and presupposition has benefited from the move to a dynamic setting (Heim 1982; Kamp and Reyle 1993; Groenendijk and Stokhof 1991; van der Sandt 1992; Dekker 1994; Chierchia 1995; among many others). Though there is not space in this paper to develop a dynamic theory of honorification in full, we feel that such a move is worthwhile, so we pause now to sketch one way in which it could be done.

At present, our theory makes do with a single context, situated as a parameter on the interpretation function. This is an *extensional* view of the role of contexts. But there can be indeterminacy about the context in the same way that there can be indeterminacy about the world, so we have reason to intensionalize this aspect of the theory, to imbue it with some of the richness of possible worlds semantics.

Our definition of (Kaplanian) contexts, (21), restricted with (22), admits a huge number of contexts, and thus already provides the basis for intensionality in this area. We can say that the initial information state (a state of complete ignorance about the utterance situation) is the full set of admissible contexts. And then we can define expressions that restrict this set, by eliminating potential but nonactual contexts. From this perspective, honorifics play much the same role as definite descriptions: they check the input information state to make sure that it has certain properties. In (26b), we provide a concrete example, alongside the static denotation of (23) above.

(26)   a. $[\![\alpha : \varepsilon]\!]^c$ is defined only if $[\![\alpha]\!]^c \in c_{HON}$
            Where defined, $[\![\alpha : \varepsilon]\!]^c \in D_\varepsilon$

    b. $C + [\![\alpha : \varepsilon]\!]$ is defined only if every $c \in C$ is such that $[\![\alpha]\!]^c \in c_{HON}$
            Where defined, $C + [\![\alpha : \varepsilon]\!] = C$

In (26b), C is a set of contexts and $+$ is an update function, which could be a kind of set intersection.

This is just a sketch of how a dynamic semantics of honorification could work. The immediacy property suggests that it should ultimately connect with the theory of speech-acts (as discussed in Potts 2003b), and the Korean facts in (25) indicate that subsentential dynamics could play a crucial role (Bittner 2001).
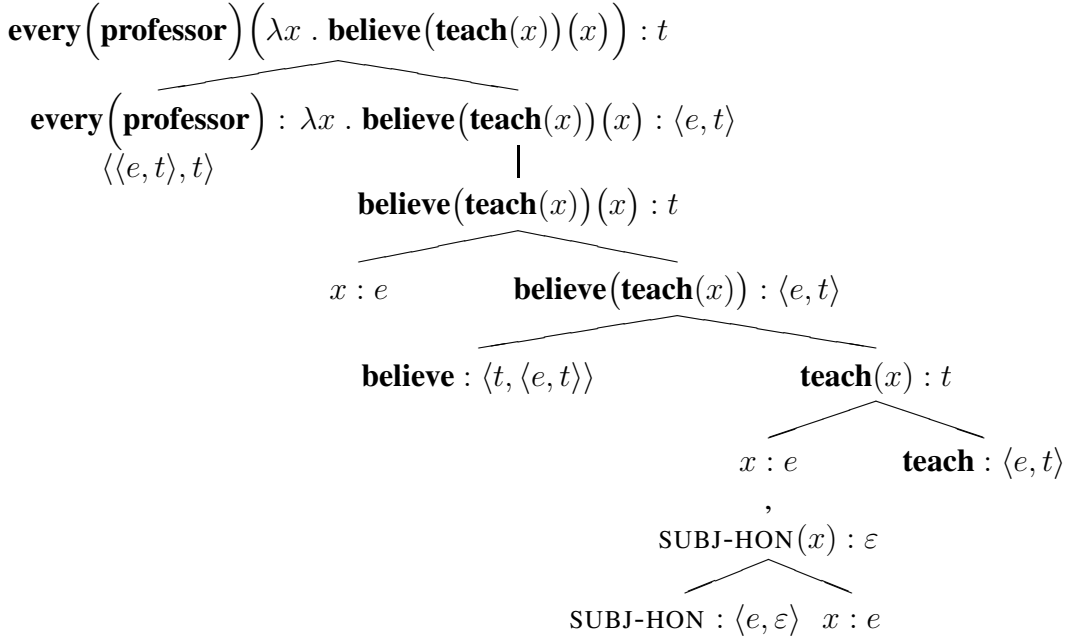
6. Honorifics and quantification

Argument-oriented honorific cannot associate with a pronoun that is bound by a higher quantifier. Here, we provide examples involving universal and proportional

quantification:

(27) dono kyoojyu-mo [*pro* sono kurasu-o oshie-ta] -to omotteiru.
      every faculty-MO   that class-ACC teach-PAST that believes

      'Every professor believes that he taught that class.'

(28) [??] dono kyoojyu-mo [*pro* sono kurasu-o   **o**-oshie-**ninat**-ta]
      every faculty-MO   that class-ACC **subj.hon**-taught-**subj.hon**-PAST
      -to   omotteiru.
      that   believes

      [*] 'Every professor believes that he, who I honor, taught that class.'

(29) hotondo no kyoojyu-ga   [*pro* sono kurasu-o oshie-ta] -to
      most-of   professor-NOM   that class-ACC teach-PAST that
      omotteiru.
      believes

      'Most professors believe that they taught that class.'

(30) [??] hotondo no kyoojyu-ga   [*pro* sono kurasu-o
      most-of   professor-NOM   that class-ACC
      o-oshie-**ninat**-ta]   -to omotteiru.
      subj.hon-taught-**subj.hon**-PAST that believes

      [*] 'Most professors believe that they, who I honor, taught that class.'

Our treatment provides an easy explanation for the markedness of these examples. The essential logical fact is that we have no provision for binding the variable that forms the argument to the honorific. Thus, it remains free in the honorific realm even as it is bound by the higher quantifier in the regular denotation. The following tree illustrates the predicament:

14

(31)   *pro* sono kurasu-wo **o**-oshie-**ninat**-ta.

$$\textbf{every}\Big(\textbf{professor}\Big)\Big(\lambda x\,.\,\textbf{believe}\big(\textbf{teach}(x)\big)(x)\Big) : t$$

$$\textbf{every}\Big(\textbf{professor}\Big) : \qquad \lambda x\,.\,\textbf{believe}\big(\textbf{teach}(x)\big)(x) : \langle e,t\rangle$$
$$\langle\langle e,t\rangle,t\rangle$$

$$\textbf{believe}\big(\textbf{teach}(x)\big)(x) : t$$

$$x : e \qquad\qquad \textbf{believe}\big(\textbf{teach}(x)\big) : \langle e,t\rangle$$

$$\textbf{believe} : \langle t,\langle e,t\rangle\rangle \qquad\qquad \textbf{teach}(x) : t$$

$$x : e \qquad\qquad \textbf{teach} : \langle e,t\rangle$$
,
$$\text{SUBJ-HON}(x) : \varepsilon$$

$$\text{SUBJ-HON} : \langle e,\varepsilon\rangle \quad x : e$$

The variable name x is important to the honorific meaning; $\text{SUBJ-HON}(x)$ denotes the triple $c_A 1 g(x)$, where g is the current variable assignment. But x is irrelevant to the regular denotation, which has no free variables in it. A major lesson of the work on bound and deictic readings of pronouns in elided phrases is that variables cannot lead this kind of dual life.

7. Prospects

Our semantics ensures that no honorific ever embeds. We achieve this effect using roughly the method that Potts (2003a, 2004) uses to ensure that conventional-implicature content does not embed. Here, however, this can been seen as part of the broader generalization that prevents speech-acts from embedding. Krifka (2001) has challenged the view that speech-act operators operate only at the root level. His evidence comes mainly from universal quantification into question acts. We have yet to find evidence that comparable facts obtain in the realm of honorifics, but this could be just a failure to look in the right places or from the proper perspective. We close with a suggestive example. First, Huang and Ochi (2004) discuss embedded *the hell*, as in (32).

(32)   a. Sue asked who **the hell** would favor the Mets.
            (the negativity of *the hell* seems to be attributed to Sue)

        b. You might be wondering who **the hell** would favor the Mets.
            (here, the negative content seems to be the speaker's)

It might be that the antihonorific *yagaru* shows similar variation:

(33)   Kankan-ni-natte John-wa [Mary-ga   nani-o   kui-**yagat**-ta]-ka
       became-angry   John-TOP Mary-NOM what-ACC eat-ANTI-HON-PAST-Q
       shiritagatteiru.
       wondering

       'John is angrily wondering what Mary ate'

(34)   Kankan-ni-natte John-wa [Mary-ga   nani-o   kui-**yagat**-ta]-ka
       became-angry   John-TOP Mary-NOM what-ACC eat-ANTI-HON-PAST-Q
       kiita.
       asked

       'John asked angrily what Mary ate'

Kazuko Yatsushiro suggests the following scenario, though she is unsure whether it works: John is a doctor and Mary is his patient. Mary is suffering from diabetes, so she has restrictions on what she can eat. After seeing the result of Mary's blood test, John knows that Mary ate something sweet. John gets mad and . . . [fill in (33) or (34) here].

Once we move to a view of grammar in which many expressions operate on contexts, and moreover lack regular proposition-oriented denotations, we should expect complex interactions between them and other elements. The semantics presented in this paper reduces the possibility of such interactions. This provides a simple calculus, but it might prove too restrictive. We hope that the simplicity focusses attention on any environments in which richer interactions actually occur.

References

Anand, Pranav and Andrew Nevins. 2004. Shifty operators in changing contexts. In Kazuha Watanabe and Robert B. Young, eds., *Proceedings of SALT 14*. Ithaca, NY: CLC Publications.

Bach, Kent. 1999. The myth of conventional implicature. *Linguistics and Philosophy* 22(4):367–421.

Bittner, Maria. 2001. Surface composition as bridging. *Journal of Semantics* 18(2):127–177.

Bobaljik, Jonathan and Kazuko Yatsushiro. 2004. Problems with honorification-as-agreement in Japanese: A reply to Boeckx and Niinuma, Ms., University of Connecticut, Storrs.

Boeckx, Cedric and Fumikazu Niinuma. 2004. Conditions on agreement in Japanese. *Natural Language and Linguistic Theory* 22(3):453–480.

Chierchia, Gennaro. 1995. *Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar*. Chicago: University of Chicago Press.

Cruse, D. A. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.

Dekker, Paul. 1994. Predicate logic with anaphora. In Lynn Santelmann and Mandy Harvey, eds., *Proceedings from SALT IX*, 79–95. Ithaca, NY: DMLL Publications, Cornell University.

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole and Jerry Morgan, eds., *Syntax and Semantics*, Volume 3: Speech Acts, 43–58. New York: Academic Press.

Groenendijk, Jeroen and Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14(1):39–100.

Harada, S. I. 1976. Honorifics. In Masayoshi Shibatani, ed., *Syntax and Semantics*, Volume 5: Japanese Generative Grammar, 499–561. New York: Academic Press.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachussets, Amherst. [Distributed by GLSA].

Heim, Irene. 1983. File change semantics and the familiarity theory of definiteness. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, eds., *Meaning, Use, and Interpretation of Language*, 164–189. Berlin: Walter de Gruyter.

Huang, C.-T. James and Masao Ochi. 2004. Syntax of *the hell*: Two types of dependencies. In Keir Moulton and Matthew Wolf, eds., *Proceedings of NELS 34*. Amherst, MA: GLSA.

Kamp, Hans. 1978. Semantics versus pragmatics. In Franz Guenthner and Siegfried J. Schmidt, eds., *Formal Semantics and Pragmatics for Natural Languages*, 255–287. Dordrecht: D. Reidel.

Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.

Kaplan, David. 1989. Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In Joseph Almog, John Perry, and Howard Wettstein, eds., *Themes from Kaplan*, 481–614. New York: Oxford University Press. [Versions of this paper began circulating in 1971].

Kaplan, David. 1999. What is meaning? Explorations in the theory of *Meaning as Use*. Brief version — draft 1, Ms., UCLA.

Karttunen, Lauri. 1973. Presuppositions and compound sentences. *Linguistic Inquiry* 4(2):169–193.

Karttunen, Lauri. 1976. Discourse referents. In James D. McCawley, ed., *Syntax and Semantics*, Volume 7: Notes from the Linguistic Underground, 363–385. New York: Academic Press.

Kikuchi, Yasuto. 1994. *Keigo* [*Honorifics*]. Tokyo: Kadokawa Shoten.

Kratzer, Angelika. 1999. Beyond *ouch* and *oops*: How descriptive and expressive meaning interact. A comment on David Kaplan's paper Cornell Conference on Theories of Context Dependency. URL `http://semanticsarchive.net/Archive/WEwNGUyO/`.

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9(1):1–40.

Krifka, Manfred. 2004. Semantics below and above speech acts, Handout from a talk at Stanford University, April 9.

Ooishi, Shootaroo. 1975. *Keigo* [*Honorifics*]. Tokyo: Chikuma Shoboo.

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Potts, Christopher. 2003a. Expressive content as conventional implicature. In Makoto Kadowaki and Shigeto Kawahara, eds., *Proceedings of the North East Linguistics Society 33*, 303–322. Amherst, MA: GLSA.

Potts, Christopher. 2003b. The performative nature of expressive content, Talk given at the University of Connecticut (November 17) and the University of Rochester (November 18).

Potts, Christopher. 2004. *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press. [Revised 2003 UC Santa Cruz PhD thesis].

Sakuma, Kanae. 1940. *Gendai Nihongohoo no Kenkyuu.* [*An Investigation into the Use of Modern Japanese*]. Tokyo: Reprinted by Kuroshio Publishers, 1983.

van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9(4):333–377.

Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26(1):29–120.

Tokieda, Makoto. 1940. *Kokugogaku Genron* [*Principles of Japanese Linguistics*]. Tokyo: Iwanami.

Truckenbrodt, Hubert. To appear. Zur Strukturbedeutung von Interrogativsätzen, *Linguistische Berichte*.

Tsujimura, Toshiki. 1978. Keigo to hikeigo [Honorifics and anti-honorifics]. In Yasuo Kitahara, ed., *Keigo: Ronshuu Nihongo Kenkyuu 9* [*Honorifics: Papers on Research in Japanese*], 218–231. Tokyo: Yuuseido.

Zanuttini, Raffaella and Paul Portner. 2003. Exclamative clauses at the syntax–semantics interface. *Language* 79(1):39–81.

Zimmermann, Thomas Ede. 1991. Kontextabhängigkeit. In Arnim von Stechow and Dieter Wunderlich, eds., *Semantik/Semantics: Ein internationales Handbuch der zeitgenössischen Forschung*, 156–229. Berlin/New York: Walter de Gruyter.