Constructionalization of the Japanese addressee-honorification system

Akitaka Yamada (Osaka University)
a.yamada.hmt@osaka-u.ac.jp

1   Introduction

The addressee-honorific (AH) construction (or, *teinei-go*) grammatically encodes the relation between discourse participants and, therefore, has gained the attention of researchers working on the interplay between grammatical forms and pragmatics/ discourse-oriented meanings. Traditionally, researchers have investigated linguistic and sociolinguistic functions of AH markers as more or less a peculiarity to East Asian languages (i.e., Japanese and Korean). However, in the 2010s, comparable AH markings were "discovered" and subsequently examined extensively in genealogically-unrelated languages beyond Korean and Japanese — such as Punjabi (Kaur 2017; 2020), Tamil (McFadden 2020), Basque (Haddican 2018), Thai (McCready 2014), Burmese (Yamada 2019b), and Ryukyuan languages (Antonov 2015). Since then, researchers have begun to relativize the commonalities and differences in these AH markings in order to better understand how natural languages reflect and encode the speaker's construal of the discourse relations.

With this background in mind, the current study zooms in on the peculiarity in the Japanese language. One important property of the Japanese AH construction that has not been observed in other languages is that the forms are category-dependent. First, as shown in (1), the AH meaning is realized as *-mas* when preceded by a verb, but in other cases, it is pronounced *des-* (n.b., NA and CA refer to a nominal adjective and a canonical adjective, respectively). Second, in early 20th century Japanese, the canonical adjective (or the *i*-adjective) was disallowed when *des-* was present; instead, the variant in (2) was recommended by prescriptive grammarians (Kawaguchi 2014; Yamada 2019b; Ogawa et al. 2020).[1]

(1)   The early 20th-century Japanese

	a. VP	b. NP	c. NAP	d. CAP
	*tobi-**mas**-u.*	*inu **des**-u.*	*kanrei **des**-u.*	**tumetai **des**-u.*
	fly-AH-PRS	dog COP.AH-PRS	cold COP.AH-PRS	cold AH-PRS
	'(I) fly.'	'(It) is a dog.'	'(It) is cold.'	'(It) is cold.'

(2)   *tumetoo **gozai**-**mas**-u.*
coldly   exist.AHU-AH-PRS
'(It) is cold.'

However, the AH system underwent a major change in the 20th century, as seen in the contrast between (1) and (3), and what used to be an unacceptable sentence in (1)d became well-formed, as in (3)d.

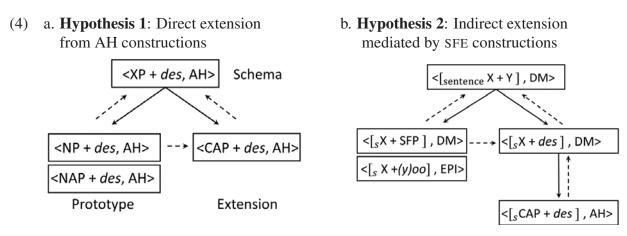[1] The conflict between a canonical adjective and a copula is still observed with the plain form:

(1)   **tumetai **da**.*
cold   COP
'(It) is cold (intended).'

(3) The 21st-century Japanese

	a. VP	b. NP	c. NAP	d. CAP
	*tobi-**mas**-u.*	*inu **des**-u.*	*kanrei **des**-u.*	*tumetai **des**-u.*
	fly-AH-PRS	dog COP.AH-PRS	cold COP.AH-PRS	cold AH-PRS
	'(I) fly.'	'(It) is a dog.'	'(It) is cold.'	'(It) is cold.'

The contrast between (2) and (3)d is, thus, to be seen as an instance of diachronic constructional alternation. The chief goal of this corpus-driven study is to quantitatively examine the tendency in this constructional alternation and propose a model to demonstrate how the new construction (canonical adjective + *des-*) is sanctioned in this language. More specifically, instead of a naive direct extension model in (4)a, the hypothesis that constructions with sentence-final particles mediate the extension is empirically more adequate in modeling the change, as shown in (4)b, where X and Y are used to represent an unspecified category/slot and DM refers to a *discourse-oriented meaning*.

(4) a. **Hypothesis 1**: Direct extension from AH constructions   b. **Hypothesis 2**: Indirect extension mediated by SFE constructions
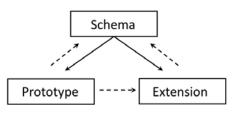

2 Previous literature

2.1 Constructionalist view of language change

In Construction Grammar, as well as in other domains of Cognitive Linguistics, the fundamental linguistic unit is identified as a construction, that is, a pair of form and meaning (i.e., sign; Langacker 2008 a.m.). Under this view, an AH construction is a pair of a certain grammatical pattern (form) and honorific allocutivity (meaning), so it is formally defined as follows (Yamada 2019b).

(5) **The AH construction**: the form-meaning pair of $< X$ *(form)*, AH *(meaning)* $>$, where the form X is associated with the allocutive honorificity.

If a construction is a pairing of form and meaning, then a historical change of a construction is to be seen as a change in one of the poles in a sign (Traugott and Trousdale 2013: 1; Noël and Colleman 2021). Such changes do not take place in a haphazard manner without a (cognitive) motivation: our cognitive capacities, such as metaphor and metonymy as well as our encyclopedic knowledge, allow us to abstract a particular schema from prototypical uses, which then sanctions the establishment of a new form-meaning pair, as shown in (6) (Prototype Theory, Langacker 1987; 2008, a.m.).

(6)

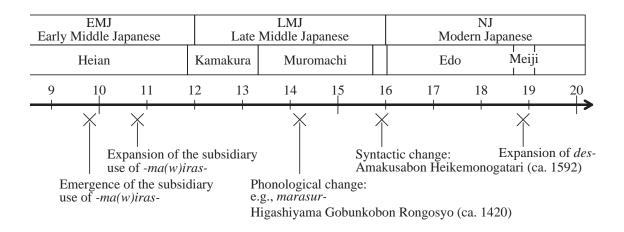| EMJ Early Middle Japanese		LMJ Late Middle Japanese			NJ Modern Japanese	

Figure 1: Historical development of addressee-honorific markers (Yamada 2019b: 163)

Hypothesis 1 (H1) in (4)a is, in fact, a simplistic extension of this model: the prototypes are the instances of *des-* with a noun phrase (NP) and a nominal adjective phrase (NAP), from which a schema is abstracted with the category of the preceding phrase unspecified. Having established such an acategorical phrase in the schema, *des-* is expected to be used with a canonical adjective phrase (CAP). While this may seem reasonable at first, this hypothesis runs into several empirical challenges, as discussed next.

2.2 The historical development of AH markers

Before we delve into a detailed examination of the development of CAP + *des-*, let us briefly look at the previous literature related to the historical change of the Japanese AH system (Figure 1; for a comprehensive overview, see Yamada 2019b: Chapter 2).

Of the two contemporary AH markers (-*mas* and *des-*), -*mas* has the longer history. The historical ancestor of this marker can be dated back to the EMJ *ma(w)iras-*, which is etymologically derived from a content-honorific marker (more specifically, the combination of object-honorific verb *ma(w)ir* 'come.OH' and -*as* 'CAUS'); hence, it is treated as an instance of grammaticalization and intersubjectification (Ohori 2005). Pertinent to *des-*, what is known, at least, is that it became popular in the 19th century; however, its earlier uses are not clear. In the early 20th century, it was preceded only by a noun phrase or a nominal adjective phrase, but a canonical adjective phrase was unacceptable, as we see in (1). In addition to the change in question (the expansion of CAP + *des-*), *mas*-based past tense negation (-*mas-en desi-ta*) is being replaced by *des*-based past tense negation (*nak-at ta des-u*), and several corpus-based quantitative studies have been conducted to reveal the characteristics of this change (Tanomura 1994; Noda 2004; Ochiai 2012; Banno 2012; Yamada 2019a). However, with respect to the development of *CAP+des-*, no elaborate quantitative statistical approaches have been applied.

3 Analysis

Given the standard view of Diachronic Construction Grammar, one might wish to propose a direct extension hypothesis, as shown in (4)a, that assumes that the new construction node (canonical adjective + *des-*) has developed by being sanctioned by a schema established from the existing AH constructions (henceforth, H1). However, this section shows that despite the apparent plausibility, this line of text-book scenario encounters a number of empirical challenges, and thus, an alternative hypothesis (henceforth, H2) is proposed that the change is mediated by sentence-final particle constructions.

3.1 Problems related to Hypothesis 1

If H1 is on the right track, formal properties of the source constructions are predicted to also be shared and inherited by CAP + *des-*. Unfortunately, however, this prediction is not borne out; the *des-* in CAP + *des-* is not distributed in the same position as in the other AH constructions. To see this, let us consider some empirical tests.

First, whether a *te*-clause can embed an AH construction tells us how large the construction is. Since the pioneering work of Minami (1974), Japanese clauses are known to have different sizes. For example, as shown in (7) and (8), the *des* with NP or NAP can be embedded within a *te*-clause.

(7) *musuko-wa* [*mada* [NP *hiyokko*] *desi-**te***] *yononaka-no koto-o   nani-mo   siri-mas-en.*
     son-TOP   still   young child AH-**te**   world-GEN   thing-ACC anything-also know-AH-NEG
     'My son, being a young child, does not know anything about the world.'

(8) *toohoo-wa* [[NP*omonaga*] *desi-**te**],   *meiku-demo kakus-e-nai   baai-ga   yoku*
     I-TOP   long-faced   COP.AH-te make-up-with conceal-can-NEG case-NOM often
     ari-mas-u.
     exist-AH-PRS.
     'As I am long-faced, makeup cannot hide a long face.'

In contrast, the *des-* with CAP cannot be embedded within a *te*-clause, as clearly shown by the following example, thus suggesting that the size of CAP + *des-* is much larger than NAP/NP + *des-*.

(9) **kanozyo-wa* [[CAP*utukusi(i)*] *desi-**te**],   *moderu-o si-tei-mas-u.*
     she-TOP   beautiful   COP.AH-te model-ACC do-PRG-AH-PRS.
     'She, being beautiful, is a model. (intended)'

Second, the order of morphemes also tells us where the *des-* is located. As seen in (10), *des-* is place before the past tense suffix in NP and NAP + *des-*. In stark contrast, CAP cannot be followed by *desi-ta*, but it must be used with *ta des-u*, as shown in (11).

(10) {*orenzi/ondan*} *desi-ta.*
     orange/warm   COP.AH-PST
     '(It) was orange/warm.'

(11) a. **omosiro(i) desi-**ta**.*   b. *omosirok at-**ta**   des-u.*
         interesting COP.AH-PST   interesting COP-PST AH-u
         '(It) was interesting. (intended)'   '(It) was interesting.'

Taken together, we consider that the formal (or distributional/syntactic) property of CAP + *des-* is quite distinct from the other AH markers. As shown in (12)a, *des-* appears in an inner layer of the sentence with NP and NAP, while it is distributed in sentence-periphery, as shown in (12)c. Since H1 predicts the pattern in (12)b, this hypothesis is empirically challenged.

(12) a.   NP/NAP: [sentence [... [... NP/NAP *des-* ...] ... *-te/-ta* ...]   ...]
     b.   CAP: *[sentence [... [...   CAP *des-* ...] ... *-te/-ta* ...]   ...]
     c.   CAP: [sentence [... [...   CAP   ...] ... *-te/-ta* ...] *des-* ...]

Two remarks are in order. First, the configuration in (12)c is, in fact, what is expected for the distribution of discourse-oriented expressions. In fact, AH markers in other languages appears in this sentence-peripheral region, and they are known to be unembeddable, as shown in (13) and (14)

(Yamada 2019b: 190). Second, the change in question cannot be seen as an instance of intersubjectification or grammaticalization because, before and after the change, the meaning of AH remains the same. Rather, this is a creation of a new form-meaning pair and is, thus, analyzed as a case of constructionalization/constructional change.[2]

(13) Korean
    a. *Ecey*   *pi-ka*   *o-ass-**supnita**.*
        yesterday rain-NOM come-PST-DECL.AH
        'It rained yesterday (polite).'
    b. **Inho-ka* [*ecey*   *pi-ka*   *o-ass-**supnita**-ko*]   *malhayss-supnita.*
        Inho-NOM yesterday rain-NOM come-PST-DECL.AH-COMP said-DECL.AH
        'Inho said [that it rained yesterday (polite)] (intended).'

(14) Thai
    a. *lian* *yêɛ*   *ləy lɔ̌* ***khráp**.*
        study problematic PP Q AH.M
        'She studies so badly? (polite)'
    b. **Sakol bɔ̀ɔk* [*waaa kháw maa* ***khráp**] khráp.*
        Sakol say that he come AH.M AH.M
        'Sakol says [that he comes (polite)] (intended).'

3.2 Hypothesis 2: Proposal

Like in other languages, Japanese sentence-final elements are used to encode information regarding some aspect of pragmatic meaning, such as the relation between the speaker and the addressee. A clear example follows.
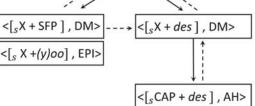
(15) a. *are-wa ookii.*   b. *are-wa ookii **ne**.*
      that-TOP big   that-TOP big SFP
      'That is big.'   'That is big, isn't it?'

A construction like this creates the schema of $<[_{\text{sentence}} ... \text{X} + \text{SFP}]$, Discourse Meaning (DM)$>$. However, sentence-final elements are not necessarily sentence-final particles. For example, the epistemic marker appears in the sentence-final position, as shown below, and encodes a discourse-oriented meaning. These constructions with sentence-final elements create the schema of $<[_{\text{sentence}} ... \text{X} + \text{Y}]$, DM$>$, as shown in (4)b (= (18)).

(16) [$_{\text{sentence}}$[*ame-ga hut-ta*] *dear-**oo**.*]
    rain-NOM   fall-PST COP-EPI
    'It is likely that it rained.'

---

[2]Since CAP + *des-* seems slightly more casual than the other cases, its meaning side can be analyzed as undergoing a change. If this is considered, it is seen as a constructionalization; if, on the other hand, one does not put weights on the change in meaning, it should be analyzed as constructionalization. This paper does not commit to this distinction and uses "constructionalization" as a cover-all term.

(17)

Although neither *-mas* nor *des-* appeared in sentence periphery in the early 20th century, it had a discourse-oriented meaning (i.e., conveying the speaker's respect for the addressee). Thus, via analogical extension, the subschema of <[sentence ... X + *des*], DM> is newly created. Since the slot X in (18) is unspecified, this can be replaced by CAP, yielding the construction node of <[sentence ... X + *des*], DM>. The schema behind the extension is not the AH construction schema, but the schema of the DM construction.

Thinking this way, we can easily explain the aforementioned otherwise mysterious property of CAP + *des-*. First, it cannot be embedded because the properties of <[sentence ... X + Y], DM> are inherited to the new construction: as shown below, the construction of <[sentence ... X + Y], DM> is unembeddable.

(18) a. *are-wa* [*ookiku-te*] *nagai-ne.*   b. **are-wa* [[*ookii ne*]*-te*] *nagai-ne.*
     that-TOP big-te   long-SFP   that-TOP big   SFP-te long-SFP
     'That is big and long, isn't it?'   'That is big and long, isn't it? (intended)'

Second, it is preceded by the past tense marker, simply because a sentence-final element is distributed after the tense marker. In this way, H2 can explain the empirical data that has created a problem for H1, and it is, therefore, superior to the naive direct extension model.

4 Verifying the hypothesis: a corpus survey and the State-Space Model

If H2 is on the right track, we should be able to see clear diachronic evidence of sentence-final elements' historical influence on the use of CAP + *des-*. In what follows, we show that this prediction is borne out; an elaborate statistical analysis of a historical corpus shows that the use of CAP + *des-* is facilitated by use of sentence-final elements.

4.1 Data and statistical model

The previous literature roughly identified the change in the AH system as taking place during the early 20th century. For this reason, the Corpus of Historical Japanese (version 2021.3; last accessed Nov. 24, 2021), which consists of written data produced from the Nara period to around after the WWII.

The following queries are set to examine the diachronic constructional alternation between CAP + *gozai mas* (as in (2)) and CAP + *des-* (as in (3)d).

(19) **Prescriptive form**: POS LIKE "Canonical Adjective%" AND FOLLOWING WORDS: LEXEME *gozar-* ON 1 WORDS FROM KEY AND FOLLOWING WORDS: LEXEME *-mas* ON 2 WORDS FROM KEY

(20) **New form**:
    a. POS LIKE "Canonical Adjective%" AND FOLLOWING WORDS: LEXEME *des-* ON 1 WORDS

Variable	Type	Description
Variant $y_{ij}^{(t)}$	Outcome	Binary variable taking the value of 1 iff the i-th sample at time t with an adjective j takes the CAP + *des*, and 0 iff it takes the CAP + *gozai mas* form.
Textbook $x_{1i}^{(t)}$	Fixed	Dummy variable taking the value of 1 iff the i-th sample at time t is taken from a textbook, and 0 otherwise.
Literature $x_{2i}^{(t)}$	Fixed	Dummy variable taking the value of 1 iff the i-th sample at time t is taken from the literature, and 0 otherwise.
Sentence-final particle $x_{3i}^{(t)}$	Fixed	Dummy variable taking the value of 1 iff the i-th sample at time t has a sentence-final particle, and 0 otherwise.
Epistemic model suffix $x_{4i}^{(t)}$	Fixed	Dummy variable taking the value of 1 iff the i-th sample at time t has an epistemic modal suffix *-oo*, and 0 otherwise.
Tense $x_{5i}^{(t)}$	Fixed	Dummy variable taking the value of 1 iff the i-th sample at time t has a past tense marker, and 0 otherwise.
Idiosyncrasies among canonical adjectives u_{0j}	Random	The idosyncrasy of the *j*-th adjective, which is assumed to follow $N(0, \tau^2)$.

Table 1: Variables in the model.

FROM KEY
b. POS LIKE "Canonical Adjective%" AND FOLLOWING WORDS: LEXEME *-ta* ON 1 WORDS FROM KEY AND FOLLOWING WORDS: LEXEME *des-* ON 2 WORDS FROM KEY

Since the *gozai mas* form is nowadays quite archaic and is no longer productively used, CAP + *des-* is expected to gradually outnumber the old variant over time. However, not only this global tendency, but also intra/extra-linguistic factors are expected to jointly affect the choice between the two in a probabilistic manner, just as the choice between ditransitive and dative constructions is simultaneously and probabilistically affected by multiple (socio)linguistic factors (Röthlisberger et al. 2017). To incorporate the chronological state change and multiple independent variables, this study conducts a Time Series Analysis, in particular, the State Space Model (Hagiwara 2021), by assuming the structure in (21) for the population with the variables in Table 1.

(21) State-Space Model:
$$y_{ij}^{(t)} \sim \text{Bern}(\pi_{ij}^{(t)})$$
$$\pi_{ij}^{(t)} \sim \text{logistic}(\eta_{ij}^{(t)})$$
$$\eta_{ij}^{(t)} = \beta_0^{(t)} + \beta_1 x_{1i}^{(t)} + \beta_2 x_{2i}^{(t)} + \beta_3 x_{3i}^{(t)} + \beta_4 x_{4i}^{(t)} + u_{0j}$$
$$u_{0j} \sim N(0, \tau^2)$$
$$\beta_0^{(t)} \sim N(\beta_0^{(t-1)}, \sigma_\zeta^2)$$

4.2 Results

The posterior distributions of the parameters in the model are estimated using a Hamiltonian Monte Carlo algorithm with Stan (Stan Development Team 2020) on R (Gelman et al. 2013; R Core Team 2020).
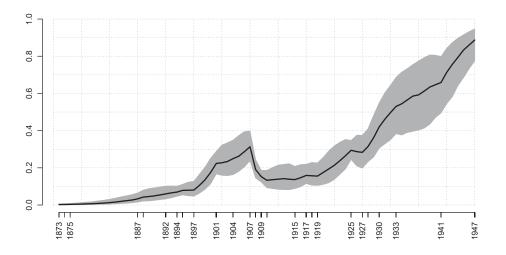
Figure 2: **Posterior inference on** $\{\pi^{(t)} : t \in \{1873, \cdots, 1947\}\}$: The posterior median of $\pi^{(t)}$ (the solid line) and its 95% credible intervals (the shaded gray area) for each year.
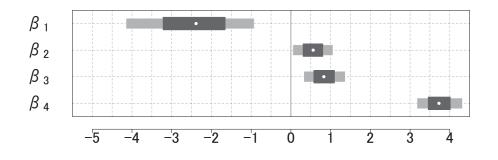

Figure 3: **Posterior inference on the coefficients**: The posterior medians of β's (the white circle) and its 66% and 95% credible intervals (the dark and light gray areas).

First, the posterior distribution of π at time t is shown in Figure 2. Evidently, there is an increasing global trend, suggesting that the probability of taking the new construction (CAP + *des-*) is getting larger as time passes.
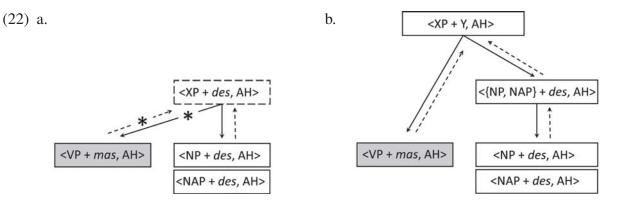
Second, in addition to this time effect, we can also detect the effects of (socio)linguistic factors. Figure 3 summarizes the posterior distributions of each coefficient. The first two are the genre effects (β_1: textbooks, β_2: novels), indicating that textbooks conservatively favor the old, recommended form, while novels seek to have the new construction. The other two variables are our main concerns, which both take a positive value, indicating that CAP+*des-* is favored when used with a sentence-final particle (β_3) and an epistemic marker (β_4). That these sentence-final elements make language users choose CAP+*des-* empirically corroborates our hypothesis (H2) that the establishment of the new construction is mediated and supported by these sentence-final elements.

5 Discussion

If H2 is on the right track, we should ask why $< \mathrm{NP} + des, \mathrm{AH} >$ and $< \mathrm{NAP} + des, \mathrm{AH} >$ did not extend to create $< \mathrm{CAP} + des, \mathrm{AH} >$, or alternatively, what prevented these construction nodes from creating a schema that sanctions the new construction node.

One possible answer worthy of our attention is the heterogeneity of AH constructions. As we saw in (1), the Japanese AH system employs two phonological exponents — *des-* and *-mas*. The presence

of *-mas* is likely to have prevented speakers in the early 20th century from creating a constructional schema $<$ XP $+$ *des-*, AH $>$, because if they had done so, the construction schema should have licensed $<$ VP $+$ *des-*, AH $>$, not $<$ VP $+$ *-mas*, AH $>$, contrary to the fact. This is what is schematically shown in (22)a. If $<$ VP $+$ *-mas*, AH $>$ is a very marginal weight in the constructional network, the schema of $<$ XP $+$ *des-*, AH $>$ could probably be established; however, since many verbs can fill in the VP slot, this construction node has a very large token and type frequencies, so it is considered that this construction node has a very negative effect on creating the schema in (22)a. Hence, in order to avoid such an unwarranted result, they needed to create a less abstract schema at the intermediate level, as shown in (22)b.

(22) a.   b.

```
        ┌ ─ ─ ─ ─ ─ ─ ─ ─ ┐                                    ┌──────────────────┐
        │  <XP + des, AH>  │                                   │   <XP + Y, AH>   │
        └ ─ ─ ─ ─ ─ ─ ─ ─ ┘                                    └──────────────────┘
      *          *                                            ╱                 ╲
   <VP + mas, AH>   <NP + des, AH>                     <{NP, NAP} + des, AH>

                    <NAP + des, AH>            <VP + mas, AH>      <NP + des, AH>

                                                                  <NAP + des, AH>
```

Still, in this constructional network in (22)b, neither $<$ XP $+$ Y, AH $>$ nor $<$ {NP, NAP} $+des-$, AH $>$ is adequate for sanctioning $<$ CAP $+$ des, AH $>$. First, $<$ XP $+$ Y, AH $>$ is too vague; we do not know what to have in the slot of Y. Second, $<$ {NP, NAP} $+$ *des-*, AH $>$ is too restricted; it only allows NP and NAP in the slot preceding *des*. The lack of $<$ CAP $+$ des, AH $>$ in the early 20th century is, therefore, attributed to a natural consequence of our categorization and schema-creation processes, and the effect of type/token frequencies in the constructional network. For this reason, a different schema (i.e., $<$[$_{sentence}$ ... X + Y], DM$>$) was used as a catalyst in developing the construction of $<$ CAP $+$ des, AH $>$.

6 Conclusion and future directions

The current paper has examined the development of the Japanese AH system, in particular, the emergence of $<$ CAP$+$des, AH $>$. While the proposal that the new AH construction node is sanctioned by the extant AH construction schema may seem plausible, this line of direct extension hypothesis (H1) has several empirical problems. As an alternative analysis, this paper proposes a hypothesis in which the new construction node is licensed by the construction schema of discourse-oriented, sentence-final elements. Thus, the sentence-final properties of $<$ CAP $+$ des, AH $>$ are clearly explained. The historical influences of such sentence-final elements are, in fact, quantitatively supported by the corpus analysis with an elaborate statistical analysis (the Generalized Mixed-Effects Time Series Analysis), and the reason why $<$ CAP $+$ des, AH $>$ does not evolve from the AH construction schema is argued to be attributable to the heterogeneity of AH constructions?

The findings and discussions in this paper can be extended in several future directions. First, this case study is seen as an attempt to study the interaction within the constructional network, preventing and allowing constructionalization. Certainly, the prototype and schema theory is useful in explaining why extensions have happened in the past, but the arguments suggesting factors preventing language change that would be likely to happen are lacking. In this paper, however, the view is presented that the heterogeneity within a constructional network has a negative effect for a schema to be abstracted. Future research is expected to carried out to determine the extent to which this prevention is commonly

observed in other constructional networks.

Second, as previously mentioned, the AH markings in Korean and Thai are distributed in sentence periphery, so the change in question has made Japanese akin to these languages. Nonetheless, a few more languages have been known to encode AH meanings in non-sentence-peripheral regions (e.g., *Basque,* Haddican 2018; *Magahi*, Alok 2021; *Burmese*, Yamada 2019b). Among these, AH markers in non-contemporary Japanese — such as *tamap-*, *saburaw*, and *paber* — deserve our attention. For example, *saburaw-* is known to be used, firstly, as a verb (that is, in a mid-sentence region), but it is later distributed in sentence periphery. Whether the direction of historical development is put forth by the same extension in constructional network needs to be extensively discussed.

Third, this paper elucidates how elaborate Time Series analysis reveal a language shift that would otherwise be quite difficult to discern. In most cases, in historical studies, much simpler models have been employed, such as Chi-square analysis. However, the shortcomings of Chi-square analysis in historical language studies have been pointed out in the literature (Yamada 2022). Utilization of Generalized Mixed-Effects Time Series Analysis is also expected to be useful in other historical language research.

Acknowledgements

References

Alok, Deepak. 2021. The morphosyntax of Magahi addressee agreement. *Syntax* 24:263–296.

Antonov, Anton. 2015. Verbal allocutivity in a crosslinguistic perspective. *Linguistic Typology* 19(1):55–85.

Banno, Eri. 2012. Koopasu o tsukatta jutsugo hitei kei "masen" to "naidesu" no shiyoo jittai choosa [a corpus-based study of the negative predicate forms "masen" and "nai desu"]. *Ryuugakusei Kyooiku [Journal of International Students Education]* 17:133–140.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian data analysis [3rd edition]*. London: CRC Press.

Haddican, Bill. 2018. The syntax of Basque allocutive clitics. *Glossa: a journal of general linguistics* 3(1).

Hagiwara, Junichiro. 2021. *Time Series Analysis for the State-Space model with R/Stan*. Singapore: Springer.

Kaur, Gurmeet. 2017. Variation in subject-triggered clitic restrictions: A case of Punjabi. In M. Erlewine, ed., *Proceedings of Glow in Asia XI*, pages 111–126. Cambridge, MA: MIT Working Papers in Linguistics.

Kaur, Gurmeet. 2020. Allocutivity as the locus of imperative syntax. In M. Baird and J. Pesetsky, eds., *Proceedings of the 49th North East Linguistics Society*, vol. 2, pages 165–174. Amherst, MA: GLSA publication.

Kawaguchi, Ryo. 2014. *Teineitai hiteikei no barieeshon ni kansuru kenkyuu [A study of variations among negative polite forms].*. Tokyo: Kuroshio Publishers.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Vol.1 Theoretical Prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.

McCready, Eric. 2014. A semantics for honorifics with reference to Thai. In *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing*, pages 503–512.

McFadden, Thomas. 2020. The morphosyntax of allocutive agreement in Tamil. In P. W. Smith, J. Mursell, and K. Hartmann, eds., *Agree to agree: agreement in the minimalist programme*, pages 391–424. Berlin: Language Science Press.

Minami, Fujio. 1974. *Gendai Nihongo no Koozoo [The Structure of Contemporary Japanese]*. Tokyo: Taishukan Shoten.

Noda, Harumi. 2004. Hitei teineikei "masen" to "naidesu" no siyō ni kakawaru yōin: yōrei tyōsa to zyakunensō ankēto tyōsa ni motozuite [factors relating to the use of "-masen" and "-nai desu" as polite negative form — based on a usage investigation and on a questionnaire investigation of young generation speakers]. *Methematical Linguistics* 24(5):228–244.

Noël, Dirk and Timothy Colleman. 2021. Diachronic construction grammar. In W. Xu and J. R. Taylor, eds., *The Routledge Handbook of Cognitive Linguistics*, pages 661–674. New York: Taylors & Francis Group.

Ochiai, Tomoko. 2012. Kakikotoba ni arawareru "masen" to "naidesu" ["masen" and "nai desu" found in written texts]. *Kokubun Mejiro [Japanese Literature Mejiro]* 51:14–22.

Ogawa, Yoshiki, Yasuaki Ishizaki, and Hirofumi Aoki. 2020. *Bunpoka, goika, koobunka [Grammaticalization, lexicalization, and constructionalization]*. Tokyo: Kaitakusha.

Ohori, Toshio. 2005. Nihongo no bunpooka kenkyu ni atatte: gaikan to rironteki kadai [perspectives on grammaticalization in Japanese: survey and theoretical issues]. *Nihongo no Kenkyuu [Studies in the Japanese Language]* 1(3):1–17.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Röthlisberger, Melanie, Jason Grafmiller, and Benedikt Szmrecsanyi. 2017. Cognitive indigenization effects in the English dative alternation. *Cognitive Linguistics* 28(4):673–710.

Stan Development Team. 2020. *Stan Modeling Language Users Guide and Reference Manual Version*. 2.21.0, https://mc-stan.org.

Tanomura, Tadaharu. 1994. Teineitai no zyutugo hiteikei no sentaku ni kansuru keiryō teki tyōsa: "-masen" to "-nai desu" [a corpus-based study on the polite negative predicate forms in japanese]. *Journal of Osaka University of Foreign Studies* 11:51–66.

Traugott, Elizabeth C. and Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Yamada, Akitaka. 2019a. A quantitative approach to addressee-honorific markers:identification of crucial independent variables and prototypes. *Methematical Linguistics* 32(3):117–132.

Yamada, Akitaka. 2019b. *The syntax, semantics and pragmatics of Japanese addressee-honorific markers*. Ph.D. thesis, Georgetown University, Washington DC.

Yamada, Akitaka. 2022. Historical pragmatics using State-Space Models. Oral Presentation in the 25th Annual Meeting of Pragmatic Society of Japan.Kyoto University/Online. Nov 26-27.

＜要旨＞

日本語丁寧語(聞き手敬語)体系の構文化

山田彬尭（大阪大学）
a.yamada.hmt@osaka-u.ac.jp

本研究は、日本語における丁寧語（聞き手敬語）の体系の通時的変化、とりわけ、イ形容詞とともに用いられる「です」の用法の台頭を、定量的データを用いながら、通時的構文文法の枠組みから論じたものである。「イ形容詞＋です」という構文が次第に容認されるようになったのは、一見すると「名詞＋です」や「ナ形容詞＋です」といった他の「です」を用いた敬語構文からの構文拡張が生じたためだという仮説（仮説1）が成り立ちそうではある。しかし本稿では、この仮説1の下では、多くの経験的なデータが説明できないことを示し、これを退ける。代わりに、この構文拡張は文末要素構文スキーマというものを介して進展したのだという仮説（仮説2）を提案する。すなわち、終助詞や認識的モーダルが文末に生起し談話機能を表す構文スキーマがまず存在し、同じく対他的意味を持つ「です」がこの構文スキーマに認可される形で、文末語としての「です」が生じ、結果として「イ形容詞＋です」という構文が登場するに至ったという歴史的変化を想定する。この仮説2が正しければ、コーパスの中でも、文末語の生起が「イ形容詞＋です」という構文の成立に影響を与えたことを示す明確なパターンが見られるはずであり、実際、コーパスデータに状態空間モデル(State-Space Model)を用いた一般化混合効果時系列分析（Generalized Mixed-Effects Times Series Model）の解析結果からこの予測が経験的に支持されることが示される。仮説1通りの「教科書的」な変化が生じなかった原因には丁寧語に「です」だけでなく「ます」も存在しているという聞き手敬語の不均質性が構文スキーマの立ち上げを阻害した可能性を指摘する。

キーワード：通時的構文文法（ Diachronic Construction Grammar ）・プロトタイプ理論（Prototype Theory）・構文ネットワーク（Constructional Network）・敬意アロキュティビティ（Honorific Allocutivity）・一般化混合効果時系列分析（Generalized Mixed-Effects Time Series Analysis）