Open Library of Humanities

Looking for default vocabulary insertion rules: Diachronic morphosyntax of the Japanese addressee-honorification system

Akitaka Yamada, Osaka University, Japan, a.yamada.hmt@osaka-u.ac.jp

As opposed to syntactic operations (such as Merge), morphology has been seen as the source of language variation. But few attempts have been made in the field of Distributed Morphology to explicitly and formally identify the role of postsyntactic rules in language change. By analyzing the change in the Japanese addressee-honorification system (Yamada 2019c), this paper proposes that the emergence of the new grammar — which seems to be an instance of a microparametric change in the Agree operation (Kayne 2000; 2005; Biberauer & Roberts 2012; Barbiers 2013) — is, in fact, explained as a morphological reanalysis: the reanalysis is caused by overgeneralization of vocabulary insertion rules resulting from abductive reasoning in pursuit of default rules. It is also suggested from a quantitative corpus survey using Time Series Analysis (the Dynamic Generalized Linear Mixed Effects Model) that the reanalysis is facilitated particularly in the configuration where the addressee-honorific marking is immediately followed by a sentence-final particle.

1 Introduction

In the transition from the P&P model to the Minimalist Program (MP), the view of language variation underwent important changes. In the P&P model, variations are seen as a parametric property attributed to a module of mental grammar (e.g., word order variation is to the X-bar module). These parameters can globally determine a cluster of syntactic variations (macroparameters; e.g., the null Subject Parameter, Rizzi 1982). In the MP, however, only universal operations, such as Merge, are considered in the domain of syntax, and what is learned — and, therefore, is subject to variation — is placed outside the syntax (lexicon, PF, third factors, etc.). Furthermore, the existence of macroparameters has become controversial, with several researchers alternatively proposing the importance of microparameters (Kayne 2000; Biberauer & Roberts 2012; Barbiers 2013). As Kayne (2005: 10) writes, "macroparameter differences might all turn out to dissolve into arrays of microparametric ones."

In addition to researchers of diachronic syntax, those working within the field of Distributed Morphology have also adopted the same assumption for grammatical architecture, attempting to elucidate the role of the postsyntactic module of grammar in (synchronic) language variation (Parrott 2007; Embick 2008; Nevins & Parrott 2010; Havenhill 2016). By examining the recent grammatical change in the Japanese addressee-honorification system from the perspective of Distributed Morphology, the current paper develops this direction, aiming to flesh out the role of postsyntactic rules in language change in a more explicit way than has been discussed in the literature. It is proposed that a microparametric change in Agree is caused by an abductive reanalysis in the postsyntactic module, in pursuit of general default vocabulary insertion rules (i.e., a change in learned lexicons).

This paper is organized as follows. Section 2 provides detailed descriptions of Japanese addressee-honorific (AH) markers, and their historical change. Section 3 presents the main analysis, discussing how the grammars before and after the change are to be analyzed, and why the language changed. Examining data from a historical corpus, Section 4 empirically and quantitatively supports the proposed analysis. The present paper concludes in Section 5 with a summary and future remarks.

2 Japanese addressee-honorific markers

AH markers — or honorific allocutive markers — are morphemes encoding the speaker's respect for the addressee.[1] They have been playing a central role in the recent surge of Performative Hypothesis-based models of the structure of clause periphery, as well as other discourse-oriented expressions (pragmatic markers/particles, Haegeman & Hill 2013; vocatives, Hill 2014; tense

---

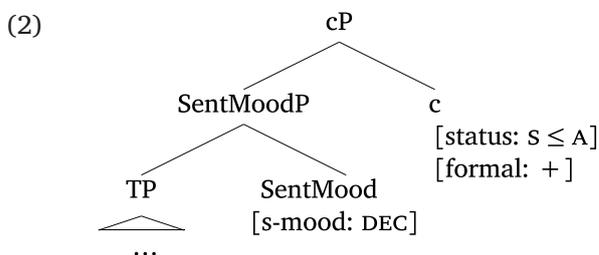[1] AH markings are found in genealogically unrelated languages; see e.g., (Souletin) Basque (*Isolate*, Oyharçabal 1993; Haddican 2018), Thai (*Tai-Kadai*, McCready 2014), Tamil (*Dravidian*, McFadden 2017), Punjabi (*Indo-European*, Kaur 2017), Magahi (*Indo-European*, Alok 2021), Korean (*Koreanic*, Pak 2008; Portner et al. 2019) and Japanese (*Japonic*, Harada 1976; Miyagawa 2017; Yamada 2019c).

ordering, Giorgi 2010; sentence-final particles exhibiting phi-agreement, Zu 2018). For example, observe an example from Korean.

(1)   Inho-ka   choysen-ul ta ha-ess-**supnita.**
        Inho-NOM best-ACC   all do-PST-AH.DECL.FORM
        'Inho did his best' (Portner et al. 2019: 6)

Here the AH marker *supnita* appears at the end of the sentence to encode the speaker's respect for the addressee. Given its orientation to the addressee and its distribution in clause periphery, Portner et al. (2019), for example, propose the structure in (2), where c(ontext)P exits to encode information about discourse relations, and analyze *supnita* as the direct reflex of this head.[2]

(2)   cP
                          ╱────────────────╲
                   SentMoodP   c
                  ╱──────╲   [status: S ≤ A]
               TP   SentMood   [formal: +]
              ╱╲   [s-mood: DEC]
             …

Contemporary Japanese also has an AH system. However, unlike in Korean, where the choice between different AH markers reflects semantic/pragmatic differences, the difference between Japanese AH markers — *-mas* and *des-* — is subject to purely morphosyntactic rules, as listed in (3) (Yamada 2019c).

(3)   The use of Japanese AH depends on:
        a.   whether the sentence is used in a copular sentence
        b.   whether it is used in a negative sentence, and
        c.   whether it is used in a past-tense sentence

But this system underwent an important change in the twentieth century. The older system (Grammar 1) still survives as the prescriptive grammar. But additionally, a new grammatical system (Grammar 2) started being used in a colloquial register. In what follows, we examine how AH markers are distributed in each system.

2.1 Grammar 1

By examining 16 monographs written by renowned linguists published from 1916 to 2009 and government-designed textbooks, Kawaguchi (2014) has revealed that people in the early twentieth century agreed on the summary table in (4). Let us closely examine each case with examples.

---

[2] More accurately, they analyze c and SentMood as being fused and *supnita* as the spellout of the complex head.

(4) Grammar 1

	affirmative		negative	
	PRS	PST	PRS	PST
verb	*-mas* (5)		*-mas* (18)b	*-mas* + *des-* (concord) (19)b
noun	*des-* (low) (15)a		*-mas* (22)a	*-mas* + *des-* (concord) (21)a
nominal adjective	*des-* (low) (15)b		*-mas* (22)b	*-mas* + *des-* (concord) (21)b
canonical adjective	*		*-mas* (22)c	*-mas* + *des-* (concord) (21)c

2.1.1 Affirmative sentences

Verbs. Observe an AH marker used with a verb. Of the two AH markers, the only permitted form is *-mas*, as shown in (5).

(5) a. Alex-wa hasiri-{**mas**/*****des**}-u ka nee.
         Alex-TOP run-AH-PRS   Q SFP
         'Does Alex run?'

     b. Alex-wa hasiri-{**masi**/*****desi**}-ta ka nee.
         Alex-TOP run-AH-PST   Q SFP
         'Did Alex run?'

As mentioned above, the Korean AH marker is seen as the direct spell-out of the c-head. Maintaining this commonly-adopted assumption of Performative Hypothesis that there exists a super-structure representing discourse participants, the literature has also analyzed *-mas* as being involved with such a highest projection (Miyagawa 2012; 2017; Yamada 2019c). However, *-mas* does not appear in the clause periphery, as it is sandwiched between the verb and the tense morpheme, distinct from sentence-final elements, such as an interrogative mood marker *ka* and a sentence-final particle *nee*.[3]

To explain this, Yamada (2019c) proposes that the head of ΣP (or NegP) and the head of cP enter into agreement via the AH feature, analyzing *-mas* as the reflex of the features present in ΣP, as in (6)b.[4]

---

[3] To be more precise, *-mas* is pronounced in a position between an aspectual marker and negation marker:

(i) Mitome-rare-te   i-**mas**-en   desi-ta   ka nee?
     recognize-PASS-CV PRG-AH-NEG COP.AH-PST Q SFP
     'Was (it) not being recognized?'

[4] Two comments are in order. First, some may wish to propose a new functional projection in place of Σ, but the subsequent discussion does not hinge on the choice of a functional projection, as long as it is located somewhere between V and T. Second, in (6)b, an interpretable feature is in a higher c-commanding probe, and the uninterpretable feature

(6)   a.   Korean: [$_{cP}$ ... [$_{TP}$ ... [$_{ΣP}$ ... [$_{vP}$ ...] ... Σ] T] ... c]

$$\Downarrow$$

supnita

AGREE

b.   Japanese: [$_{cP}$ ... [$_{TP}$ ... [$_{ΣP}$ ... [$_{vP}$ ...] ... Σ] T] ... c]

$$\Downarrow$$

mas   [Neg: –] [AH: +]
mas-en [Neg: +] [AH: +]

Nouns and nominal adjectives. When an AH marker is used with nominal and adjectival predicates, another AH marker *des-* is utilized instead of *-mas*:

(7)   Alex-ga   {gakusya/siawase} **des**-u.
        Alex-NOM scholar/happy   AH-PRS
        'Alex is a scholar/happy.' (noun/nominal adjective)

To better understand this, however, we need to understand the system of (i) the copula construction and (ii) classes of adjectives.

Japanese has two types of copulas (Nishiyama 1997; 1999): the predicative copula and the dummy copula. Let us see the PREDICATIVE COPULA, first. Although no overt copula is manifested in English small clauses (= (8)), Japanese obligatorily pronounces a copula even in a non-tense environment (= (9)), and this is called the predicative copula (I use PRED for the gloss).

(8)   Bernie consider [Alex smart].

(9)   a.   Bernie-ga   Alex-o   {gakusya/siawase} **ni**]   si-ta.
            Bernie-NOM Alex-ACC scholar/happy   PRED do-PST
            'Bernie made Alex a scholar/happy.'

        b.   Bernie-ga   [Alex-o   utukusi **ku**]   si-ta.
            Bernie-NOM Alex-ACC beautiful PRED do-PST
            'Bernie made Alex beautiful.'

The predicative copula is pronounced either *ni* or *ku*. This choice is solely dependent on the category of the preceding element; a noun always takes *ni*, but adjectives are split into two types: (i) the one that takes *ni* is called the NOMINAL ADJECTIVE, (ii) while the one with *ku* is the CANONICAL ADJECTIVE.

---

is located in a lower projection in ΣP, which is an assumption different from that in Chomsky (2000; 2001). However, in later studies, the valuation and interpretability are considered two independent properties (Pesetsky & Torrego 2007). Under this view, the honorific feature in Σ is unvalued when externally-merged. Alternatively, one can assume that an Agr-node with a feature is post-syntactically introduced when c-commanded by a triggering feature (Kramer 2009; 2010; Norris 2014; Baier 2015). The analysis of Yamada (2019c) in (6)b adopts the latter strategy.

In the affirmative, tensed-environment, *ni* and *ku* are pronounced *de* and *ku* respectively:

(10) a. Alex-ga   {gakusya/siawase} **de**   at-ta.
          Alex-NOM scholar/happy   PRED COP-PST
          'Alex was a scholar/happy.' (noun/nominal adjective)

  b. ?Alex-ga   utukusi **ku**   at-ta.
          Alex-NOM beautiful PRED COP-PST
          'Alex was beautiful.' (canonical adjective)

Furthermore, they can be optionally contracted with the following element *ar-* (*at-* is its allomorph) to be pronounced *dat-* and *kat-*, respectively:[5]

(11) a. Alex-ga   {gakusya/siawase} **dat**-ta.
          Alex-NOM scholar/happy   PRED.COP-PST
          'Alex was a scholar.'

  b. Alex-ga   utukusi **kat**-ta.
          Alex-NOM beautiful PRED.COP-PST
          'Alex was beautiful.'

Note that this contraction is only permitted when the two elements are adjacent. When a focus particle is attached to the PredP, no contraction is triggered:

(12) a. Alex-ga   [{gakusya/siawase} **de**]-wa   **at**-ta.
          Alex-NOM scholar/happy   PRED-FOC COP-PST
          'It is true that Alex was a scholar/happy.'

---

[5] Although non-contracted examples of *ku* are rarer (when compared with *de*) — this is why I put a question mark in (10)b — there are few cases where the contraction is unaccepted. First, with *beki* 'should,' the non-contracted form is obligatory:

(i)   Hana-ga   utukusi {**ku ar**/*****kar**}-u   beki   da.
          flower-NOM beautiful PRED COP/PRED.COP-u should COP
          'Flowers should be beautiful.'

As in (ii), the *-u* preceded by a non-contracted form does not alternate with *-ta*. Thus, it is likely that the contraction is sensitive to the position where the *ar*-support takes place (i.e., whether it is in T or not).

(ii) *Hana-ga   utukusi {**ku at**/**kat**}-ta   beki   da.
          flower-NOM beautiful PRED COP/PRED.COP-u should COP
          'Flowers should have been beautiful (intended).'

Second, the contraction is prohibited with *-mas-en*:

(iii) Hana-ga   utukusi {**ku   ari**/*****kari**}-mas-en.
          flower-NOM beautiful PRED COP/PRED.COP.AH-NEG
          'Flowers are not beautiful.'

b. Alex-ga   [utukusi **ku**]-wa   **at**-ta.
Alex-NOM beautiful PRED-FOC COP-PST
'It is true that Alex was beautiful.'

The second type of copula is the DUMMY COPULA. It is the element with which the predicative copula is fused. As the name suggests, it is used when a suffix needs morphological support, just as the English *do*-support. Without *ar-*, the sentences are ungrammatical:

(13) a. *Alex-ga   {gakusya/siawase} **de**   -ta.
Alex-NOM scholar/happy   PRED -PST
Alex was a scholar/happy (intended).'

b. *Alex-ga   utukusi **k**   -ta.
Alex-NOM beautiful PRED -PST
'Alex was beautiful (intended).'

The negation marker also triggers a dummy element, again just like in English (e.g., *I *(do)-ed not know*):

(14) a. Alex-wa   hasit-ta.
Alex-TOP run-PST
'Alex ran.'

b. Alex-wa   hasir-anak *(**at**)-ta.
Alex-TOP run-NEG   COP-PST
'Alex did not run.'

Having understood the basics of Japanese copular sentences, let us now turn to the AH system in the affirmative copular sentence. In the polite sentence, the head amalgam *dat*-in (11) is replaced by *des-* (*desi* is its phonologically-conditioned allomorph):

(15) Alex-ga   {gakusya/siawase} {**desi**/*masi}-ta.
Alex-NOM scholar/happy   PRED.COP.AH-PST
Alex was a scholar/happy.' (noun/nominal adjective)

One cannot use *-mas* in this copula position. Conversely, *des-* cannot be placed where *-mas* is supposed to appear. They are, thus, in complementary distribution.

Canonical adjectives. What becomes crucial in the subsequent discussion is that neither *-mas* nor *des-* is permitted with a canonical adjective:

(16) a. *Alex-ga   [utukusi ku]-**masi**-ta.
Alex-NOM beautiful PRED-AH-PST

b. *Alex-ga   utukusi **desi**-ta.
Alex-NOM beautiful PRED.COP.AH-PST
'Alex was beautiful (intended).'

First, *-mas* cannot be used, because it is not a copula. Second, *des-* is illicit, because it is the replacement of *dat-* but not of *kat-*. Therefore, the AH meaning fails to be expressed with a canonical adjective: the AH system in Grammar 1 is 'incomplete.'[6]

A natural question to be addressed is how people convey the intended meaning. The aforementioned survey by Kawaguchi (2014) reveals that up until the 1970s, grammarians recommended the following construction to meet the demand:

(17) Alex-ga   utukusy-uu **gozai-masi**-ta.
      Alex-NOM beautiful-ly exist.AHU-AH-PST
      'Alex was beautiful (lit., Alex existed in a beautiful manner).'

Here, the canonical adjective *utukusi* 'beautiful' is converted into the adverb *utukusy-uu* 'beautiful-ly,' accompanied by a verb of existence: because there is a verb, *-mas* can be attached. However, as its idiosyncratic property, this verb *gozar-* (*gozai* is its allomorph) enhances the politeness level already encoded by *-mas* (hence it is called an AH UPGRADER, AHU; Yamada 2019c; Ikawa & Yamada 2022); the politeness level of (17) is higher than the intended level (i.e., that of (15)). In the strictest sense, this construction is, thus, not a genuine alternative; for this reason, an asterisk is placed in the bottom left corner of (4).[7]

2.1.2 Negative sentences

Verbs. When the sentence is negated, the AH marking exhibits different patterns depending on the tense. When a verb phrase is used with the present tense, *-mas* is followed by the negation marker *-en*:[8]

(18) a. Alex-wa   hasira-nai ka nee?
          Alex-TOP run-NEG   Q   SFP
          'Does Alex not run?' (plain)

      b. Alex-wa   hasiri-{**mas**/*des}-en ka nee?
          Alex-TOP run-AH-NEG   Q   SFP
          'Does Alex not run?' (polite)

In the past tense, the sentence is slightly more complex because of the dummy copula. Earlier, we saw that the negation marker triggers the *ar*-support. In a polite sentence, the dummy copula must be replaced by *des-*, although the AH meaning is already encoded by *-mas*:[9]

---

[6] For *gozar-*, see Section 2.3.1.

[7] As for nouns and nominal adjectives, *gozai-mas* can also be used to encode the highest level of politeness; hence, three honorific levels are clearly identified. Based on this observation, Mikami (1959 [2002]: 157) convincingly argues that the canonical adjective lacks the intermediate polite meaning.

[8] When the negation marker is adjacent to *-mas*, *-nak* (*-nai* is its allomorph) changes its from to *-en* (Yamada 2019c).

[9] In contrast, such an *ar*-support is not obtained with a present tense marker. The so-called present tense marker *-u*

(19) a. Alex-ga   hasiri-**masi**-ta.
          Alex-NOM   run-AH-PST
          'Alex ran.'

    b. Alex-ga   [hasiri-**mas**-en] {*at/**desi**}-ta.
          Alex-NOM   run-AH-NEG   COP.AH-PST
          'Alex did not run.'

Despite the multiplicity of AH markers, the politeness level of (19)b is no higher than that of (19) a. Hence, unlike the previous *des-* 'PRED.COP.AH,' this *des-* 'COP.AH' is considered a realization of an uninterpretable honorific feature introduced only for morphosyntactic reasons, not for semantic purposes, just as the gender feature spreads out in the nominal in well-studied languages (e.g., *la̲ casa̲ blanca̲*). For presentational purposes, I call it *des-* (concord) 'COP.AH,' and the one in an affirmative sentence *des-* (low) 'PRED.COP.AH.'[10]

Nouns and nominal/canonical adjectives. Below are examples of a plain form with nonverbal-predicates in a past-tense, negative environment:[11]

(20) a. Alex-ga   {gakusya/siawase} **de**-wa   nak-**at**-ta.
          Alex-NOM   scholar/happy   PRED-FOC NEG-COP-PST
          'Alex was not a scholar/happy.'

    b. Alex-ga   utukusi **ku**(-wa)   nak-**at**-ta.
          Alex-NOM   beautiful PRED-FOC NEG-COP-PST
          'Alex was not beautiful.'

Notice that *de/ku* and *ar-* are no longer adjacent, so neither *dat-* nor *des-* is available. If the intended meanings are conveyed in a polite form, *-mas* is used with the verb *ar-*, as in (21). Here, too, *des-* (concord) is obligatory:

(21) a. Alex-wa   {gakusya/siawase} de-wa   *(*ari*)-**mas**-en **desi**-ta.
          Alex-TOP   scholar/happy   PRED-FOC be-AH-NEG   COP.AH-PST
          'Alex was not a scholar/happy.'

    b. Alex-wa   utukusi ku(-wa)   *(*ari*)-**mas**-en **desi**-ta.
          Alex-TOP   beautiful PRED-FOC be-AH-NEG   COP.AH-PST
          'Alex was not beautiful.'

---

does not seem to be a genuine tense suffix, but it is more reasonable to see it as a postsyntactically inserted vowel (Yamada 2019a; 2020). By seeing that the stranded affix filter is only applicable to the suffix present in the narrow syntax, we can explain the absence of the *ar*-support in (18).

[10] Technical analysis of how it concords with *-mas* does not affect the subsequent discussions (for an attempt at a detailed analysis, see Yamada 2019c; 2020); for our purposes, it suffices to state that there is a concord between the two positions.

[11] In most cases, a focus particle *-wa* in inserted to mark the scope of the negation.

Unlike the past tense suffix, the present tense does not require an *ar*-support. Thus, the *des(i)-* is also absent:

(22) a. Alex-wa {gakusya/siawase} **de**-wa *(*ari*)-**mas**-en.
 Alex-TOP scholar/happy PRED-FOC be-AH-NEG
 'Alex was not a scholar/happy.'

 b. Alex-wa utukusi **ku**(-wa) *(*ari*)-**mas**-en.
 Alex-TOP beautiful PRED-FOC be-AH-NEG
 'Alex was not beautiful.'

2.2 Grammar 2

Grammar 1 has still been used even by the latest/youngest Japanese native speakers. But in the 20th century, a new grammar system also emerged originally starting from a colloquial register, and started competing with the old system. In this new system, a sentence like (23) can be felicitously generated, enabling a canonical adjective to be used with *des-* (-*u* is considered an epenthetic vowel):[12]

(23) [[$_{TP}$ Alex-ga utukusi kat-ta] {**des**/*mas}-u].
 Alex-NOM beautiful PRED.COP-PST AH-u
 'Alex was beautiful.'

There are several reasons for treating the new variant differently from *des-* (low/concord). First, the *des-* in (23) does not appear in the expected dummy copula position; it is located after the past tense marker (clause periphery). Notice also that, unlike in (16)b, the predicative copula in (23) remains as *kat-*, not *des-*.

Second, the new *des-* does not alternate with *de ar-/dat-*. Compare (23) with (24).[13]

(24) [[$_{TP}$ Alex-ga utukusi k-at-ta] (*{**de ar-u**/**da**})].
 Alex-NOM beautiful PRED-COP-PST PRED COP-PRS/COP
 'Alex was beautiful.'

Third, it is well-known that clauses differ in size (Minami 1974), and the new *des-* cannot be distributed in what is considered to be a reduced clause. For example, while *des-* (low) can be present in *te-* and *tara*-clauses (= (25)a and (26)a), the *des-* in (23) is completely ungrammatical when embedded, as demonstrated in (25)b and (26)b, suggesting the need to treat it differently from *des-* (low).

---

[12] Nowadays, some (young) people use Grammar 2 even in a formal context, and the boundary is quite complex.

[13] This also applies to *des-* (concord):

(i) *Alex-wa utukusi ku-wa *ar*-mas-en {**de at/dat**}-ta.
 Alex-TOP beautiful PRED-FOC COP-AH-NEG PRED COP/COP-PST
 'Alex was not beautiful (intended).'

(25) a. Alex-wa [gakusya **desi**-te],   amerika-ni ryuugaku   si-tei-masi-ta.
          Alex-TOP scholar   PRED.COP.AH-CV America-to study abroad do-PROG-AH-PST
          'Alex, being a scholar, studied abroad in the US.'

      b. *Alex-wa [utukusi k-at-ta   **desi**-te], moderu-o   si-tei-masi-ta.
          Alex-TOP beautiful PRED-COP-PST AH-CV   everything know-PROG-AH-PST
          'Alex, having been beautiful, was a model (intended).'

(26) a. [Gakusya **desi**-tara]   suzuki sensei-o   syookai   si-mas-yoo ka?
          scholar   PRED.COP.AH-if Suzuki prof.-ACC introduce do-AH-shall Q
          'If (you are looking for) a scholar, shall I introduce Prof. Suzuki?

      b. *[Utukusi k-at-ta   **desi**-tara] doraihurawaa-ni si-ta   kamosire-mas-en.
          beautiful PRED-COP-PST AH-if   dried flower-into make-PST may-AH-NEG
          'If (it) was beautiful, I would have made it into a dried flower (intended).'

Given these observations, I call this new variant *des-* (high) 'AH' to distinguish it from the aforementioned uses of *des-* for presentational purposes.

2.2.1 Negative sentences

Now that a canonical adjective can be used with an AH marker (*des-* high), there seems to be no necessity for more change. Nonetheless, the expansion of the use of *des-* is seen in other environments in Grammar 2.

As an alternative to the -*mas* + *des-* concord construction, people have started using *des-* (high), as shown below (Tanomura 1994; Noda 2004; Ozaki 2004; Kawaguchi 2014; Yamada 2019b; c; Ogawa et al. 2020):[14]

(27) a. [[$_{TP}$ Alex-ga   hasir-anak at-ta]   **des**-u].
              Alex-NOM run-NEG   COP-PST AH-u
          'Alex did not run.'

      b. [[$_{TP}$ Alex-ga   {gakusya/siawase} de-wa   nak-at-ta]   **des**-u].
              Alex-NOM scholar/happy   PRED-FOC NEG-COP-PST AH-u
          'Alex was not a scholar/happy.'

      c. [[$_{TP}$ Alex-ga   utukusi ku(-wa)   nak-at-ta]   **des**-u].
              Alex-NOM beautiful PRED-FOC NEG-COP-PST AH-u
          'Alex was not beautiful.'

---

[14] The sentences sound more natural when the nominative marker -*ga* is replaced by -*wa*, the focus particle. However, for expository purposes, the current paper uses a sentence with a nominative subject.

In all of these sentences, *-mas* is no longer present. Consequently, *des-* (concord) is suppressed. Instead, just like (23), the AH marker is found in the clause periphery. Clearly, they are distributed after the tense marker, and as shown below, they do not alternate with *de ar-/dat-*, and cannot be embedded in a *te*-clause or a *tara*-clause. For these reasons, I will also use *des-* (high) to refer to these instances.

(28) a. *[[$_{TP}$ Alex-ga gakusya de-wa   nak-at-ta] **de ar/dat**-u].
          Alex-NOM scholar PRED-FOC NEG-COP-PST PRED COP/PRED.COP-PST-u
       'Alex was not a scholar (intended).'

   b. *Alex-wa [gakusya de-wa   nak-at-ta   **desi**-te], amerika-ni ryuugaku
       Alex-TOP scholar PRED-FOC NEG-COP-PST AH-te   America-to study abroad
       si-tei-masi-ta.
       do-PRG-AH-PST
       'Alex, not being a scholar, studied abroad in the US.'

The same holds for present-tense sentences:

(29) a. [[$_{TP}$ Alex-ga   hasir-anai] **des**-u].
          Alex-NOM run-NEG   AH-u
       'Alex does not run.'

   b. [[$_{TP}$ Alex-ga   {gakusya/siawase} de-wa   nai] **des**-u.]
          Alex-NOM scholar/happy   PRED-FOC NEG AH-u
       'Alex is not a scholar/happy.'

   c. [[$_{TP}$ Alex-ga   utukusi ku(-wa)   nai] **des**-u].
          Alex-NOM beautiful PRED-FOC NEG AH-u
       'Alex is not beautiful.'

2.2.2 Affirmative sentences

In an affirmative sentence, the use of *des-* (high) is less common.[15] Yet we can find the following real-use instances:

---

[15] Some native speakers (especially people in older generations) may not like the examples in (27) through (31), because they sound too casual. But in younger generations, some speakers who do use (27), (28), (29), and (31) find (30) too colloquial to use, with a nuance of the speaker's being jokey or deliberately cutesy/childish. A possible reason for this may have to do with the fact that a verb is the only part-of-speech that requires *-mas*, but not *des-*, in Grammar 1. The strong connection with *-mas* may have a negative 'carry-over' effect in Grammar 2, to the extent that people are less comfortable using a verb with *des-* particularly in formal situations, so they are limited in very casual registers. Marked as they are, it remains true that they are produced in a daily conversation, so following the literature (Ogawa et al. 2020), this paper takes the widest scope treating examples found in casual registers as also reflecting the mental grammar of contemporary Japanese native speakers.

(30) Verb

 a. Moo tukare-ta **des**-u-ne.
   already get tired-PST AH-u-SFP
   '(We) already got tired.'[16]

 b. Minasan-to onazi-yooni, […] watasi-mo ganbar-u   **des**-u.
   you-with same-way   I-also   try to do my best-PRS AH-u
   'Like you, […] I will do my best.'   (Ogawa et al. 2020: 259)

(31) Noun/nominal adjective

 a. [[Ii mise   dat-ta]   **des**-u].
   good restaurant PRED.COP-PST AH-u
   '(It) was a good restaurant.' (noun)[17]

 b. [Pan-wa ["kyarottopan"-toka suki dat-ta]   **des**-u-ne].
   bread-TOP carrot bread-such as fond PRED.COP-PST AH-u-SFP
   'As for bread, I was fond of (bread) such as "carrot bread".' (nominal adjective)[18]

Note that as for the affirmative, present-tense sentence, the difference between the prescriptive form and new form is neutralized when it is preceded by a noun or nominal adjective; the following sentence is analytically ambiguous:[19]

(32) Noun/nominal adjective
 Alex-wa {gakusya/siawase} **des**-u.
 Alex-TOP scholar/happy   (PRED.COP.)AH-PRS/u
 'He is a scholar/happy.'

---

[16] Source: https://twitter.com/mmm_drama/status/872447904169943041?lang=fr.

[17] Source: https://www.hotpepper.jp/strJ001194986/report/detail_om0001452178_rn20180611A00066/.

[18] Source: https://www.city.yonago.lg.jp/item/45473.htm.

[19] This is because of the ambiguous status of -*u*. In Grammar 1 (e.g., the sentence in (i)), the -*u* is analyzed as a non-past (present tense) marker, because it triggers an *ar*-support. In contrast, in Grammar 2 sentence like (ii), -*u* is considered a vowel inserted only for phonological purposes.

 (i) Alex-ga gakusya de ar-**u**.
   Alex-NOM scholar PRED COP-PRS
   'Alex is a scholar.'

 (ii) [[Ii mise   dat-ta]   des-**u**].
   good restaurant PRED.COP-PST AH-u
   '(It) was a good restaurant.' (noun)

On the one hand, if -*u* is analyzed as a tense marker, the sentence in (32) is seen as an instance of *des*- (low). On the other hand, if it is seen as an epenthetic vowel, the same sentence can be seen as an instance of *des*- (high).

The table in (33) summarizes the observations so far. As previously mentioned, Grammar 1 is still permitted. So the case in point is not only an instance of diachronic change in that a new grammar system has evolved, but at the same time it is a case of synchronic competition of multiple grammars (aka Grammar Competition, Kroch 1995; Pintzuk 1999; 2002; Embick 2008). Gradualness of language change has been attributed to an individual's having multiple grammars, the alternation of which is regulated by certain conditions. Roughly speaking, the alternation in question can be characterized as genre-conditioned multiple grammars: Grammar 1 is favored in a formal/prescriptive genre, whereas Grammar 2 is preferred in a colloquial/casual context (though genre is not the only factor, with some other intra- and extra-linguistic factors jointly affecting the choice between the old and new variant in a probabilistic manner, as we will see in Section 4.3). Our goal is to provide a theory explaining how Grammar 2 emerged — though limited to certain genres — and started competing the older system.

(33) Grammar 2

	affirmative		negative	
	PRS	PST	PRS	PST
verb	*des-* (high)		*des-* (high)	
	(30)b	(30)a	(29)a	(27)a
noun	*des-* (low/high)	*des-* (high)	*des-* (high)	
	(7)a/(32)	(31)a	(29)b	(27)b
nominal adjective	*des-* (low/high)	*des-* (high)	*des-* (high)	
	(7)b/(32)	(31)b	(29)c	(27)c
canonical adjective	*des-* (high)		*des-* (high)	
	(23)		(29)d	(27)d

2.3 Comments on possibly related constructions

A reviewer asks some questions about constructions that seem to be related. We will now turn to these before building our analysis in the subsequent sections.

2.3.1 The *gozar*-construction

When we discuss a gap in Grammar 1, it is mentioned that the *gozai-mas*-construction was recommended by renowned linguists, although the politeness level is higher than the intended level. However, some readers may wonder if *gozar-* could be used alone without *-mas*, as shown below:

(34) Alex-ga   utukusyuu **gozat**-ta.
      Alex-NOM   beautifully exist.AHU-PST
      'Alex was beautiful (lit., Alex existed in a beautiful manner).'

This kind of construction was indeed used in the past around in the Muromachi and Edo periods (Yuzawa 1982: 163–166, Kinsui 2011), but during the Edo period, the use of *gozar-* gradually disappeared, save for the *gozai-mas*-construction (Yuzawa 1954: 205, Sato 1990: 306), and by the time Grammar 1 was used (the early twentieth century), the construction in (34) did not compete with other AH constructions. Hence, the present paper does not examine this archaic expression any further.[20]

2.3.2 The *no des*-construction

One might also wonder if the new variant in (35)a is derived as a phonological reduction from the *no des*-construction, as shown in (35)b.

(35) a. Alex-ga utukusi i   **des**-u.
    Alex-NOM beautiful PRED.PRS AH-PRS
    'Alex is beautiful.'

   b. Alex-ga utukusi i   **no** **des**-u.
    Alex-NOM beautiful PRED.PRS COMP PRED.COP.AH-PRS
    'It is that Alex is beautiful.'

However, this analysis is difficult to maintain. First, if an (optional) erosion of *no* is possible, the sentence in (36)a should also be as accepted, contrary to the fact.

(36) a. *Alex-ga utukusi i   **dat**-ta.
    Alex-NOM beautiful PRED.PRS COP-PST
    'Alex was beautiful (intended).'

   b. Alex-ga utukusi i   **no** **dat**-ta.
    Alex-NOM beautiful PRED.PRS COMP AH-PST
    'It was that Alex was beautiful.'

---

[20] Looking at the CHJ, we can find some examples of the *gozar-*, even in the government-designed textbook published in 1947. This paper, however, maintains that the 'gap' did exist in Grammar 1, as has been proposed in the literature. First, the apparent counter-examples in CHJ turn out to be examples from sections designed to familiarize students with old Japanese cultures (e.g., to learn *Kyoogen* texts, which were written in early Edo period Japanese); hence, such examples are not representative of twentieth-century Japanese. Second, if *gozar-* had been used in the twentieth century, linguists should have mentioned as such, which is convincingly denied by the survey of Kawaguchi (2014). Finally, previous studies examining the history of *gozar-* have revealed that after it had obtained the use of an AH marking, it underwent several morphosyntactic/semantic changes in the Edo period, before eventually being only observed in the fossilized grammatical context (namely, the *gozai-mas-* construction). Of course, why the *gozar*-based system became obsolete is another important question. However, I leave this issue to future research, because the change took place during a different historical period.

Second, if it is a phonological erosion (at PF), the semantics of the two sentences must be truth-conditionally equivalent. However, the two sentences are not interchangeable in meaning; only the *no des*-construction has a certain nuance that the prejacent is an explanation to the context (Kuno 1973), and/or that is has a focus reading (Hiraiwa & Ishihara 2002). So this construction will not be investigated any further in the subsequent discussion.[21]

2.3.3 Uses with an epistemic marker

Unlike the two constructions above, there is one important construction more relevant to the discussion of the development of Grammar 2 — a sentence with an epistemic marker. Independently of, and prior to, the transition of the AH system, the system of epistemic markers had undergone a change (Nakamura 1948: 156–170, Ogawa et al. 2020: 252).

In Early Middle Japanese, *-amu* (the old form of *-oo*) was used as an epistemic suffix, and it appeared in a position directly adjacent to a verb phrase, as shown in (37)a.

(37)   Verb
   a.   [[$_{vP}$ Hasir] -amu].
            run   EPI

   b.   [[[$_{vP}$ Hasir-u] **de**] **ar**-oo].
             run-u   de   COP-EPI
        '(He) will run'

However, it could not be directly preceded by a noun or nominal adjective (= (38)a); instead, a predicative and dummy copula were needed (= (38)b).

---

[21] A reviewer asks how the sentence in (35)b is generated in Grammar 1. One possible analysis is that the *no* is a nominalizer/complementizer making a nominal clause:

(i)   (Zyookyoo-wa) [Alex-ga   utukusi   i   **no**]   **des**-u.
      situation-TOP   Alex-NOM beautiful PRED.PRS COMP PRED.COP.AH-PRS
      'It (= the situation) is that Alex is beautiful.'

This is consistent with several observations. First, *no* clauses can receive a case marker, just like a noun:

(ii) Watasi-wa [Alex-ga   utukusi   i   **no**]-**o**   sit-ter-u.
      I-TOP   Alex-NOM beautiful PRED.PRS COMP-ACC know-PRF-PRS
      'I know that Alex is beautiful.'

Second, the preceding element takes the adnominal form (the form for a relative clause):

(iii) (Zyookyoo-wa) [Alex-ga   gakusya {**na**/*****da**}   no] **des**-u.
      situation-TOP   Alex-NOM beautiful PRED.COP.ADN/PRED.COP COMP PRED.COP.AH-PRS
      'It (= the situation) is that Alex is beautiful.'

Third, notice that the English translation is given 'it is [that …].' Since the subordinate marker *that* introduces a clause appearing in an argument position, it is no wonder why the corresponding *no* introduces as a noun clause.

(38)   Noun

    a. *[[$_{NP}$ Gakusya] {-amu/oo}].
            scholar   EPI

    b. [[[$_{NP}$ Gakusya] **de**] **ar**{-amu/oo}].
             scholar   PRED COP-EPI
    '(He) will be a scholar'

Presumably as an extension from (38)b, the verb came to be able to take an epistemic suffix with *de* and *ar-*, as in (37)b, making (37)a obsolete in contemporary Japanese.[22]

    The spread of this spurious *de* started from this verbal construction, and after a while, the same shift happened to the canonical adjective:

(39)   Canonical adjective

    a. [[$_{PredP}$ Utukusi k]   ar-{amu/oo}].
             beautiful PRED COP-EPI

    b. [[[$_{PredP}$ Utukusi i]   **de** **ar**-oo].
               beautiful PRED de COP-EPI
    '(He) will be beautiful'

Although (39)a used to be the right form for a canonical adjective, it became obsolete, and alternatively, (39)b has become the standard form. As before, *de* and *ar-* can contract to *dat-* in the plain sentence (= (40)a); in the polite sentence, this *dat-* is replaced by *des-* (= (40)b).[23]

(40)   a. [[$_{PredP}$ Utukusi i]   **dar**-oo].
              beautiful PRED de.COP-EPI

    b. [[$_{PredP}$ Utukusi i]   **des**-yoo].
              beautiful PRED de.COP.AH-EPI
    '(He) will be beautiful'

    If so, one might wonder if this *des-* is *des-* (low). Certainly, it alternates with *de ar-/dar-*. However, it coexists with a predicative copula. Hence, the *de* in (39)b cannot be treated as a genuine predicative copula.

    One might also wonder if this *-oo* is pronounced at T just like the English auxiliary *will*. However, unlike in English, *-oo* is pronounced separately from — and, thus, can coexist with — the tense marker:

---

[22] In contemporary Japanese, the form in (37)a can no longer encode an epistemic meaning, but nonetheless, the form *hasir-oo* has continued to be used to encode the volitionality of the subject; to put it another way, in the old days, the form in (37)a was ambiguous between the volitional reading and epistemic reading, just like the English work *will*.

[23] The *y* in *yoo* is not an inserted consonant. I use the Cabinet-ordered romanization system, in which [ʃ] is expressed as a combination of *s* and *y* as in *des-yoo*. Under the IPA system, however, it is expressed as [des-o:] > [deʃo:]. So what looks like an allomorph of *oo* is in fact a phonological change in *des*.

(41) [[$_{TP}$ Hasit-**ta**] des-**oo**].
      run-PST   de.COP-EPI
      'it is probable that (he) ran.'

Therefore, *-oo* must not be treated as an element in T. Rather, it should be seen as an instance of another sentence-final element.[24] We will come back to this example in Section 3.2 when we examine the influence of (41) on the development of *des-* (high).

3 Analysis

The emergence of *des-* (high) seems to have been caused by the need to fill a gap in the AH-system. However, the true question is why and how. The fact that there is a gap does not fully explain why *des-* (high) was recruited instead of other possible constructions, and how this change started. After presenting my morphosyntactic analysis of Grammars 1 and 2, this section proposes a hypothesis answering the questions by proposing that the change is driven by abductive reasoning in pursuit of obtaining default vocabulary insertion rules.

3.1 Grammar 1

First, consider *des-* (low) in Grammar 1. As previously mentioned, Yamada (2019c) proposes that *-mas* is obtained as a result of an AH-agreement between a speech act projection and a lower head Σ (= (6)b). Advancing this view, we propose that *des-* (low) is obtained as a result of the agreement between a speech act projection and another lower head Pred.

3.1.1 With a noun or a nominal adjective

To instantiate the main idea, consider the sentences below.

(42) a. Alex-ga   gakusya **de**   at-ta.
          Alex-NOM scholar   PRED COP.AH-PST

      b. Alex-ga   gakusya **dat**-ta.
          Alex-NOM scholar   PRED.COP.AH-PST
          'Alex was a scholar.'

(43) Alex-ga   gakusya **desi**-ta.
      Alex-NOM scholar   PRED.COP.AH-PST
      'Alex was a scholar.'

---

[24] For representational purposes, I use Epi(stemic)P à la Cinque (1999) for presentational purposes, but the choice of this label does not affect the analysis of this paper.

Derivation of (42)a. Following the seminal work of Nishiyama (1997; 1999), who gives an elaborate theory of Japanese copular sentences, we assume that *de* is the head of PredP, and *ar-* is a dummy element. When this PredP analysis is combined with Performative Hypothesis, the structure in (44) is obtained (Portner et al. 2019).[25]

(44)

```
                    cP
                   /  \
                 TP    c
                /  \
            PredP    T
            /   \
          NP    Pred
          |
          N
```
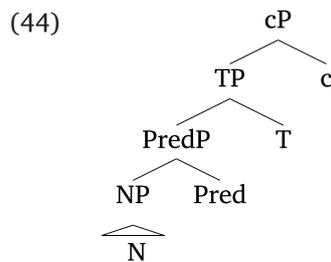
Based on this structure, several postsyntactic modifications are applied. First, the stranded suffix T needs a dummy element. This problem is solved by the *ar*-support:

(45) $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots Pred] \ldots [_T \; dummy \; T]] \ldots c]$
$\uparrow$

With the vocabulary insertion rule in (46), we obtain the terminal nodes, as represented in (47) (= (42)a).

(46)   a.   Pred   ↔   *de*   (Version 1 out of 5; Grammar 1)
        b. $T_{[PST:+]}$ ↔   *ta*
        c.   *dummy*   ↔   *ar*

(47) $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots Pred] \ldots [_T \; dummy \; T]] \ldots c]$
                                                              | | |
                                                             de   *ar* *-ta*

Derivation of (42)b. The probabilistic alternation between (42)a and (42)b is attributed to the optionality of morphological merger (Marantz 1984; Embick & Noyer 2001). When Pred and T are adjacent, Pred can optionally adjoin the T-node:

(48) $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots t_i] [_T [Pred_i a \; dummy] T]] \ldots a \; c]$

The rules in (46), however, wrongly predict *de at-ta* for (48). To avoid this, an allomorphy rule is added, as in (49)b, providing the terminal nodes in (50) (= (42)b).

---

[25] To be more precise, Nishiyama (1997; 1999) assumes that the dummy element is the head of VP (cf., Miyama 2011; Watanabe 2013), whereas in (45), the *ar-* is analyzed as a post-syntactically inserted supporting element. However, this difference does not make any significant differences to the subsequent discussions.

(49) a. Pred ↔ *de* (Version 2 out of 5; Grammar 1)

     b. **Pred** ↔ **d**/[_ **dummy**]

     c. $T_{[PST:+]}$ ↔ *-ta*

     d. *dummy* ↔ *ar*

(50) $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots t_i \quad]$ $[_T [\text{Pred}_i \; dummy] T]] \ldots c]$

                                                                        | | |

                                                                       d   *ar* *ta*

Derivation of (43). Pred enters into an agreement relation with the AH feature in c. The current paper leaves open the question of the timing regarding the morphological merger; the subsequent discussion will not be affected, regardless of whether the agreement takes place before the morphological merger (= (51)a) or not (= (51)b).

(51) a. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots \text{Pred} \quad]$ $[_T \, ar\text{-} T] \ldots c]$

AGREE (arc over Pred ... c)

     b. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots t_i \quad]$ $[_T [\text{Pred}_i \, ar\text{-}] T]] \ldots c]$

AGREE (arc over ar ... c)

Finally, when the Pred has a positive honorific feature, Pred and the dummy element are fused into a single terminal node before vocabulary insertion (= (53)), and with the rules in (52), *des* is inserted, as shown in (54).[26] With an epenthetic vowel *i*, (43) is derived.[27]

---

[26] Given that the pattern in the plain form is well-organized, we do not have motivation for analyzing *dar* and *kar* as a single fused form, and I assume that the fusion is applied when there is an honorific feature in Pred.

[27] A reviewer asks what would happen if we dispensed with the morphological merger, proposing (i)a or (i)b.

(i) a. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots \text{Pred}^{[AH:+]} \quad]$ $[_T \, dummy \, T] \ldots c]$

                                                     | | |

                                                    d   *es* *ta*

     b. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NP} \ldots N \ldots] \ldots \text{Pred}^{[AH:+]} \quad]$ $[_T \, dummy \, T] \ldots c]$

                                                                     | |

                                                                des *ta*

Unfortunately, these analyses run into an empirical problem: a focus particle can be introduced between *de* and *ar* for the *de ar*-sentence, as in (ii)a, indicating that a focus particle can select a PredP. The structures in (i) predict (ii)b and (ii)c, contrary to the fact.

(ii) a. Alex-ga gakusya-de-**wa** ari-masi-ta.
            Alex-NOM scholar-PRED-FOC be-AH-PST

     b. *Alex-ga gakusya-d-**wa** esi-ta.
            Alex-NOM scholar-PRED-FOC COP-PST

     c. *Alex-ga gakusya-∅-**wa** desi-ta.
            Alex-NOM scholar-PRED-FOC COP-PST
            'It is true that Alex was a scholar.'

(52) a. Pred ↔ *de* (Version 3 out of 5; Grammar 1)
   b. Pred ↔ *d*/[_ *dummy*]
   c. **X**$_{[AH:+]}$ ↔ ***des*/[_ T]**
   d. T$_{[PST:+]}$ ↔ *ta*
   e. *dummy* ↔ *ar*

(53) [Pred$_{[AH:+]}$ *dummy*] → # Pred$_{[AH:+]}$ *dummy* # (Version 1 out of 2)

(54) [$_{cP}$ ... [$_{TP}$ ... [$_{PredP}$ ... [$_{NP}$ ... N ...] ... t$_i$] [$_T$ #$\underline{\text{Pred}_i^{[AH:+]}\ \textit{dummy}}$# T]] ... c]
           | |
          des   *ta*

3.1.2 With a canonical adjective

Now consider the derivation of a canonical adjective:

(55) a. Alex-ga   utukusi   **ku**   at-ta.
          Alex-NOM beautiful PRED COP-PST

   b. Alex-ga   utukusi   **kat**-ta.
       Alex-NOM beautiful PRED.COP-PST

   c. *Alex-ga   utukusi   **dat**-ta.
       Alex-NOM beautiful PRED.COP-PST
       'Alex was beautiful.

(56). *Alex-ga   utukusi   **desi**-ta.
       Alex-NOM beautiful PRED.COP.AH-PST
       'Alex was beautiful (intended).

Assuming the structure in (57), one might propose the rules in (58).

(57) [$_{cP}$ ... [$_{TP}$ ... [$_{PredP}$ ... [$_{CAP}$ ... CA ...] ... t$_i$] [$_T$ [Pred$_i$ *dum.*] T]] ... c]

(58) a. Pred ↔ *de*
   b. **Pred** ↔ ***ku*/[CA _]**
   c. Pred ↔ *d*/[_ dummy]
   d. X$_{[AH:+]}$ ↔ *des*/[_ T]
   e. T$_{[PST:+]}$ ↔ *ta*
   f. *dummy* ↔ *ar*

However, this analysis — while certainly explaining the sentence in (55)a — runs into a problem, because, as shown below, it wrongly predicts that the sentences in (55)c and (56) are both acceptable, contrary to the fact.

(59) a. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{CAP} \ldots CA \ldots] \ldots t_i] \quad [_T [Pred_i \quad dum.] T]] \ldots c]$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\; d \qquad ar \quad ta$

    b. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{CAP} \ldots CA \ldots] \ldots t_i] [_T [\underline{Pred_i^{[AH:+]} \quad dum.}] T]] \ldots c]$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\; des \qquad\quad ta$

To obtain the correct form, we need to somehow remember what projection Pred is initially merged with, even after it morphologically merges with T. In Minimalist Program, c-selection has been treated as a special case of agreement (Chomsky 2000: 132, 138; Collins 2002; Adger 2010).[28] Following this tradition, particularly the work of Adger (2010), we treat a selectional feature as a tuple of attribute and value, with the value being unsaturated, thus probing for the value of the goal (n.b. the adoption of Adger's (2010) system is for purposes of illustration, and different agreement notations also do the job). At the time when Pred is externally-merged, as in (60), the unvalued selectional feature obtains its value from the category feature of the head of the projection with which it merges.[29]

(60) $[_{PredP} \ldots [_{CAP} \ldots CA \ldots] \ldots Pred]$

$\qquad\qquad\quad [\text{Cat: } \langle CA, \; 1 \rangle] \quad\; [\text{Cat: } \langle Pred, \; 2 \rangle]$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad [\text{Sel: } \langle CA, \; \_ \; \rangle]$

Being saturated, the selectional feature is no longer active in future 'syntactic' operations. But it remains, and is taken into consideration at the time of 'postsyntactic' operations. Since Pred undergoes a morphological-merger, the feature is pied-piped to the sister position of the dummy element:[30]

---

[28] In Adger (2003), the c-selection is done via checking. However, the choice between checking and agreement is rather conceptual, and does not result in a significant change in the subsequent discussion.

[29] For the ease of illustration, I adopt the bracket notation in place of the set notation. The number used in the value slot is intended to capture the scopal hierarchies (i.e., the Hierarchies of Projection, HoPs), and constrains the distribution of the features. For instance, $\langle CA, 1\rangle$ and $\langle Pred, 2\rangle$ ensure that Pred appears in a position higher than the CA in the structure (Adger 2010: 198).

[30] For limited space, the current paper only presents examples of past-tense sentences. However, our analysis can explain the data in the present tense (e.g., *de ar-u* and *des-u*), if the following rule is added to the list in (62): $T_{[PST:-]} \leftrightarrow u$. A potential problem that has been acknowledged by the previous literature is the treatment of *i* that appears in the non-polite, present-tense canonical adjective, as shown below (Nishiyama 1997; 1999). Note that this is not a vowel harmony: even if the root does not end with *i* (e.g., *omosiro* 'funny'), it is pronounced *i*.

(i) Alex-ga   utukusi **i**.
    Alex-NOM beautiful i
    'Alex is beautiful.'

(61) a. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NAP} \ldots NA \ldots] \ldots t_i] [_T [Pred_i^{[AH:-]}^{[Sel:\langle NA,1\rangle]} \ dum.] T]] \ldots c]$

                                                                   d   *ar*   *ta*

     b. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{CAP} \ldots CA \ldots] \ldots t_i] [_T [Pred_i^{[AH:-]}^{[Sel:\langle CA,1\rangle]} \ dum.] T]] \ldots c]$

                                                                   k   *ar*   *ta*

Then the vocabulary insertion rules in (62) and the fusion rule in (63) are utilized, which are sensitive to the difference between [Sel:⟨CA, 1⟩] and [Sel:⟨NA, 1⟩], hence resulting in different phonological exponents.[31]

The status of this *i* can be analyzed in different ways. One can see it as an allomorph of *u* 'PRS.' Under this view, it remains unclear why it does not trigger an *ar*-support, as in *ku ar-u*. Nishiyama (1999) tries to answer this question by proposing that that "present tense does not require an overt form of the (dummy) copula (ibid.: 191–192)." This explanation raises a deeper question of why a dummy copula tends to disappear cross-linguistically, and why there is a split in Japanese between *u* and *i*.

An alternative is to see it as an allomorph of *k(u)* 'PRED.' By seeing *k* as the real morpheme and *u* as an epenthetic vowel, this line of thinking seems plausible with an additional rule such that $k \rightarrow i$. Under this view, we assume a silent tense morpheme at T, and the absence of a dummy copula is easily derived; since ∅ is silent, it does not need substance to which it attaches. This analysis, too, is not without problems. For example, the choice of T (*u* and ∅) must be sensitive to the phrase to which Pred is merged with. However, T and CA/NA are not in a sister node relation, so we need to devise a mechanism as to how this works.

The current paper does not have a good solution to the problem of this morpheme, either, leaving this issue to future research, but it is not likely that a treatment of this morpheme will substantially change our main analysis.

[31] The sequence *de ar-u* can be contracted to *da*, as shown below. So, one may wish to propose an additional fusion rule ($[[Pred_i^{[AH:-]} \ dummy] T_{[PST:-]}] \rightarrow \# Pred_i^{[AH:+]} \ dummy] T_{[PST:-]} \#$) and a vocabulary insertion rule ($Pred_i^{[AH:+]} + dummy + T_{[PST:-]} \leftrightarrow da$).

(i) Alex-ga   gakusya {**de**   **ar**-u/**da**}.
    Alex-NOM scholar PRED COP-PRS/PRED.COP.PRS
    'Alex is a scholar.'

But this analysis is empirically challenged; this morpheme cannot be embedded in a relative clause, as in (ii) although it can be embedded in a quotative and an interrogative clause (= (iii) and (iv)).

(ii) [Alex-ga   gakusya {**de**   **ar-u**/*****da**}]   zizitu-o sitteir-u.
     Alex-NOM scholar PRED COP-PRS/PRED.COP.PRS fact-ACC know-PRS
     'I know the fact that Alex is a scholar.'

(iii) [Alex-ga   gakusya {**de**   **ar-u**/**da**}]-to   sitteir-u.
     Alex-NOM scholar PRED COP-PRS/PRED.COP.PRS-COMP know-PRS
     'I know that Alex is a scholar.'

(iv) [Alex-ga   gakusya {**de**   **ar-u**/**da**}   ka]-o sitteir-u.
     Alex-NOM scholar PRED COP-PRS/PRED.COP.PRS Q-ACC know-PRS
     'I know whether Alex is a scholar.'

(62) a. **Pred**$_{[AH:-]}$ $\quad\quad$ $\leftrightarrow$ $\quad$ *de* $\quad\quad\quad\quad\quad\quad\quad\quad\quad$ (Version 4 out of 5; Grammar 1)

$\quad\quad$ b. **Pred**$_{[AH:-]}$ $\quad\quad$ $\leftrightarrow$ $\quad$ *d*/[_ *dummy*]

$\quad\quad$ c. **Pred**$_{[Sel:\langle CA, 1\rangle], [AH:-]}$ $\quad$ $\leftrightarrow$ $\quad$ *ku*

$\quad\quad$ d. **Pred**$_{[Sel:\langle CA, 1\rangle], [AH:-]}$ $\quad$ $\leftrightarrow$ $\quad$ *k*/[_ *dummy*]

$\quad\quad$ e. X$_{[AH: +]}$ $\quad\quad\quad$ $\leftrightarrow$ $\quad$ *des*/[_ T]

$\quad\quad$ f. T$_{[PST: +]}$ $\quad\quad\quad$ $\leftrightarrow$ $\quad$ *ta*

$\quad\quad$ g. *dummy* $\quad\quad\quad$ $\leftrightarrow$ $\quad$ *ar*

(63) [Pred$_{[AH: +]}$ *dummy*] → # Pred$_{[AH: +]}$ *dummy* #

$\quad\quad$ iff Pred bears [Sel:$_{\langle N, 1\rangle}$] or [Sel:$_{\langle NA, 1\rangle}$] $\quad\quad\quad\quad\quad\quad\quad$ (Version 2 out of 2)

With these, we can derive (42)a and (55)a, as follows:

(64) a. $[_{cP}$ ... $[_{TP}$ ... $[_{PredP}$ ... $[_{NAP}$ ... NA ...] ... Pred$^{[AH:-]}_{[Sel:\langle NA,1\rangle]}$] $[_T$ *dum.* T]] ... c]

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ | $\quad\quad\quad$ | $\quad$ |

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ *de* $\quad\quad\quad$ *ar* $\quad$ *ta*

$\quad\quad$ b. $[_{cP}$ ... $[_{TP}$ ... $[_{PredP}$ ... $[_{CAP}$ ... CA ...] ... Pred$^{[AH:-]}_{[Sel:\langle CA,1\rangle]}$] $[_T$ *dum.* T]] ... c]

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ | $\quad\quad\quad$ | $\quad$ |

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ *ku* $\quad\quad\quad$ *ar* $\quad$ *ta*

As for a nominal adjective (as well as a noun) with [AH : +], neither (62)a nor (62)b can be applied, so the sentence is ill-formed without a fusion (= (65)a). In contrast, it is grammatical once the terminal nodes are fused (= (65)b), because (62)e is available. As for a canonical adjective, however, the fusion cannot be applied, due to the categorical specification in (63); hence, a canonical adjective cannot be used with *des* (= (65)c).

(65) a. $[_{cP}$ ... $[_{TP}$ ... $[_{PredP}$ ... $[_{NAP}$ ... NA ...] ... t_i] $[_T$ [Pred$^{[AH: +]}_{i}$ $_{[Sel:\langle NA,1\rangle]}$ *dum.*] T]] ... c]

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ | $\quad\quad$ | $\quad$ |

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ * $\quad\quad$ *ar* $\quad$ *ta*

Therefore, we must make the vocabulary insertion rule sensitive to the status of it being in a relative clause. Two analyses seem plausible. The first is to assume an agreement between C$_{[Rel: +]}$ and T (or Pred), and use the rule in (vi). Another is to assume a further morphological merger, making the c and Pred, dummy, and T within the same terminal node when there is C$_{[Rel: +]}$: with the rule in (vi), this analysis works, as desired.

(v) a. Pred$^{[AH: +]}_{i}$ + *dummy* + T$_{[PST:-][Rel:-]}$ $\leftrightarrow$ *da*

$\quad\quad$ b. Pred$^{[AH: +]}_{i}$ + *dummy* + T$_{[PST:-][Rel: +]}$ $\leftrightarrow$ *na*

(vi) a. Pred$^{[AH: +]}_{i}$ + *dummy* + T$_{[PST:-]}$ + C$_{[Rel:-]}$ $\leftrightarrow$ *da*

$\quad\quad$ b. Pred$^{[AH: +]}_{i}$ + *dummy* + T$_{[PST:-]}$ + C$_{[Rel: +]}$ $\leftrightarrow$ *na*

Admittedly, the current paper does not have any further data to show which is better, leaving this problem to future research. However, in either way, the rules in (62) remain intact and the analysis of *da* does not affect analysis of the historical development of *des-*.

b. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NAP} \ldots NA \ldots] \ldots t_i] [_T \#\underline{Pred_i^{[Sel:\langle NA,1\rangle]\,[AH:+]}dum.}\# T]] \ldots c]$

$$\underset{des}{\underline{\qquad}} \qquad \underset{ta}{|}$$

c. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{CAP} \ldots CA \ldots] \ldots t_i] [_T [Pred_i^{[Sel:\langle CA,1\rangle]\,[AH:+]}dum.] T]] \ldots c]$

$$\underset{*}{|} \qquad \underset{ar}{|}\ \underset{ta}{|}$$

Finally, when the morphological-merger does not take place, no well-formed sentence is obtained, because (62)a and (62)c are only applicable to a non-AH Pred:

(66) a. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{NAP} \ldots NA \ldots] \ldots Pred^{[Sel:\langle NA,1\rangle]\,[AH:+]}] [_T dum. T]] \ldots c]$

$$\underset{*}{|} \qquad \underset{ar}{|}\ \underset{ta}{|}$$

b. $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{CAP} \ldots CA \ldots] \ldots Pred^{[Sel:\langle CA,1\rangle]\,[AH:+]}] [_T dum. T]] \ldots c]$

$$\underset{*}{|} \qquad \underset{ar}{|}\ \underset{ta}{|}$$

3.1.3 With a verb and an epistemic marker

Let us complete the list of the rules, by adding a few more for verbs, and epistemic markers.

Verbs. As mentioned earlier, the derivation of *-mas* is analyzed as a node-sprouting at Σ (Yamada 2019c) (i.e., creating an adjunct to Σ):

(67) $[_{cP} \ldots [_{\Sigma P} \ldots [_{VP} \ldots V] \ldots [_{\Sigma} X_{[AH:+]} \Sigma]] \ldots c_{[AH:+]}]$

Having this in mind, one might wish to propose the rules in (68). However, the rule in (68)a runs into the problem of overgeneralization: not only does it generate (69)a, but it also allows (69)b, where PredP and $[_{\Sigma} X_{[AH:+]} \Sigma]$ are in the sister node relation.

(68) a. $X_{[AH:+]}$ $\qquad\qquad \leftrightarrow$ ***mas*** $/[\_ \Sigma]$
    b. $\Sigma_{[NEG:+]}$ $\qquad\qquad \leftrightarrow$ ***en*** $/[X_{[AH:+]} \_]$

(69) a. Alex-wa   hasiri-**mas**-en.
      Alex-TOP run-AH-u
      'Alex does not run.'

    b. *Alex-wa   {gakusya/siawase} (de)   **mas**-en.
       Alex-TOP scholar/happy   PRED AH-u
       'Alex is not a scholar/happy (intended)'

In order to prevent the rule from being applied to sentences like (69)b, we need to impose a condition on Σ, as below:

(70) $X_{[AH:+]}$ ↔ **mas** $/[\_ \ \Sigma_{[Sel:\langle v,1\rangle]}]$

This rule states that a node with an AH-feature is realized as *mas* if and only if it appears in the sister node of Σ which selects a verb.[32]

 Note that the rule in (62)e cannot be applied, because the relevant terminal node is not the sister node of T. Hence, the following sentence is ungrammatical.

(71)   *Alex-wa   hasiri-**desi**-ta.
        Alex-TOP   run-AH-PST
        'Alex ran (intended).'

Lastly, the following rules are needed for non-polite sentences, as in (73).

(72)   a. $\Sigma_{[NEG:-]}$ ↔ ∅
         b. $\Sigma_{[NEG:+]}$ ↔ **nai**

---

[32] In addition to a verb, some auxiliaries are followed by *-mas*, but not by *des-*. For example, consider the example in (i)b, where the passive auxiliary is attached to the verb.

(i)   a.   Bernie-ga   Alex-o   home-{masi/*desi}-ta.
            Bernie-NOM   Alex-ACC   praise-AH-PST
            'Bernie praised Alex.'

       b.   Alex-ga   home-**rare**-{masi/*desi}-ta.
            Alex-NOM   praise-PASS-AH-PST
            'Alex was praised.'

However, Japanese auxiliaries are, in fact, constituted by miscellaneous elements, and there are 'auxiliaries' that are used with *des-*, such as *hazu* 'should' and *yoo* 'seems':

(ii)   Alex-ga   siawase   na   {hazu/yoo}   {**desi**/*masi}-ta.
       Alex-NOM   happy   COP.ADN   should/seem   PRED.COP.AH/AH-PST
       'Alex {should/seems to be} be happy.'

These expressions differ from *(r)are*-type auxiliaries in three important ways. First, unlike *-(r)are* in (i)b, they can be preceded not only by verbs, but also by adjectives and nouns; n.b., **siawase-rare* 'happy-PASS.' Second, etymologically, these expressions used to be nouns referring to a specific part of an arrow, and to a manner, respectively. Third, the copula preceding them takes the adnominal form *na* (the form for a relative clause), not *da*, (see also Footnote 21). With all these considerations into account, *hazu* and *yoo* are morphosyntactically 'nouns' (or 'nounish auxiliaries') being modified by a relative clause, despite their semantics which would be expressed by auxiliaries in English. In other words, only 'verbish' auxiliaries are used with *-mas*.

 The rule in (70) has room for capturing this selectional restriction. Notice that the idea that categorical information can be further decomposed into more primitive features has been advocated since old days (e.g., Chomsky 1974; Stowell 1981: 21). If a primitive feature is shared by a verb and a verbal auxiliary, we can replace the "V" in (70) with that feature, and correctly capturing the distributional property of *-mas*. Although a full resolution awaits further work, it should be noted that at least as far as the rules for *-mas* restricted to verbs and auxiliaries (i.e., something 'verbish'), the subsequent discussion is unaffected.

(73) a. Alex-wa   hasir-u.
        Alex-TOP run-PRS
        'Alex runs.'

     b. Alex-wa   hasira-**nai**.
        Alex-TOP run-NEG
        'Alex does not run.'

     c. Alex-wa   {gakusya/siawase} de   ar-u.
        Alex-TOP scholar/happy   PRED COP-PRS
        'Alex is a scholar/happy.'

     d. Alex-wa   {gakusya/siawase} de   **nai**.
        Alex-TOP scholar/happy   PRED NEG
        'Alex is not a scholar/happy.'

Epistemic markers. Recall from Section 2.3.3 that when used with an epistemic marker, *des-* is used independently of the past tense marker:

(74) [[$_{TP}$ Hasit-**ta**] des-**oo**].
        run-PST   de.COP-EPI
     'It is probable that (he) ran.'

Just as Σ and Pred, Epi also receives a node-sprouting operation when c-commanded by $c_{[AH:+]}$, as in (75). By extending the list of rules as in (76), now (74) is correctly derived.

(75) [$_{cP}$ ... [$_{EpiP}$ ... [$_{TP}$ ... [$_{VP}$ ... V] ... T] [$_{Epi}$ X$_{[AH:+]}$ Epi]] ... c$_{[AH:+]}$]
                                     | | |
                                     ta   *des*   oo

(76) a. Pred$_{[AH:-]}$ ↔ *de*   (Version 5 out of 5; Grammar 1)
     b. Pred$_{[AH:-]}$ ↔ *d/[_ dummy]*
     c. Pred$_{[Sel:\langle CA, 1\rangle],[AH:-]}$ ↔ *ku*
     d. Pred$_{[Sel:\langle CA, 1\rangle],[AH:-]}$ ↔ *k/[_ dummy]*
     e. Σ$_{[NEG:-]}$ ↔ ∅
     f. Σ$_{[NEG:+]}$ ↔ ***nai***
     g. Σ$_{[NEG:+]}$ ↔ ***en*** /[X$_{[AH:+]}$ __]
     h. X$_{[AH:+]}$ ↔ ***des*/[__ T]**
     i. X$_{[AH:+]}$ ↔ ***des*/[__ Epi]**
     j. X$_{[AH:+]}$ ↔ ***mas*** /[__ Σ$_{[Sel:\langle v,1\rangle]}$]
     k. T$_{[PST:+]}$ ↔ *ta*
     l. *dummy*   ↔ *ar*

3.2 Grammar 2

Now that our analysis for Grammar 1 has been proposed, let us turn to our central question: how did Grammar 2 emerge in an environment where Grammar 1 is the only available system, resulting in a variation within a single speaker? To consider this question, this paper inherits the following assumptions from the previous literature:

(77) a. The competing grammars model of (morphosyntactic) variation and change (Kroch 1995; Pintzuk 1999; Yang 2000; Fuss & Trips 2002; Embick 2008; Amaral & Roeper 2014; Simonenko & Carlier 2020; Truswell 2021)

  b. Reanalysis in language change as a result of learners' abductive reasoning (Andersen 1973; Lightfoot 1979; 2006; Roberts 2007)

First, the competing grammars (or grammar competition) model refers to the view widely-held in diachronic syntaticians that a single individual speaker possesses (more than) two distinct grammatical options. Specifically, this theory rejects the view that syntactic change progresses by means of an abrupt transition in grammar, that is, the view that an individual during a language change must possess either the old grammar or the new grammar. This is because a sudden change cannot explain the fact that "during a period of syntactic change, variation may exist not only on the level of the language community, but also within the grammars of the individual speakers of that community (Pintzuk 1999: 6)." To account for such intra-speaker variations, the competing grammars model proposes that a single individual can have multiple grammars which alternate in some regulated fashion. Hence, the variation is attributed not to the heterogeneity within one single grammar: each grammar is considered stable, and only allows a single output (Single Output property, Embick 2008: 65). Rather, the gradualness in language change is analyzed as coming from the competition between competing grammars. With "the new option competing with the old one and gradually replacing it (Pintzuk 2002: 510)," language change proceeds in a continuous manner.

Embick (2008) develops this idea by fleshing out technical details within the framework of Distributed Morphology. For example, the variation between *dove* and *dived* is analyzed as a consequence of two competing grammars, as formulated below (Embick 2008: 66–67):

(78) *dove*-grammar
  a. $T_{[PST]} \leftrightarrow$ *-t* $/\_\{\sqrt{\text{LEAVE}}, \sqrt{\text{BEND}}, \cdots\}$

  b. $T_{[PST]} \leftrightarrow$ *-∅*$/\_\{\sqrt{\text{HIT}}, \sqrt{\text{SING}}, \sqrt{\textbf{DIVE}}, \cdots\}$

  c. $T_{[PST]} \leftrightarrow$ *-ed*

(79) *dived*-grammar
  a. $T_{[PST]} \leftrightarrow$ *-t* $/\_\{\sqrt{\text{LEAVE}}, \sqrt{\text{BEND}}, \cdots\}$

b. $T_{[PST]} \leftrightarrow \text{-}\varnothing/\_\{\sqrt{\text{HIT}}, \sqrt{\text{SING}}, \cdots\}$

c. $T_{[PST]} \leftrightarrow \text{-}ed$

Second, this paper also assume a language acquisition-based explanation for language change (Paul 1920; Andersen 1973; Harris & Campbell 1995; Lightfoot 2006; Roberts 2007). Note that the grammar competition theory does not say anything about the way a new generation started having multiple grammars despite their parents' not having the new grammar as their second option. In order to explain the emergence of a new grammar, the literature has hypothesized that abductive reasoning in language acquisition is the source of language change. Children exposed to language data make an inference about the set of rules that are supposed to generate the language output. However, there is no guarantee that the conclusion that they reach is identical to the rules of the grammar of their parents (Andersen 1973). The reasoning of this kind is called an abductive inference, and is considered responsible for language change.

The grammar competition model and the theory of acquisition-based language change are not in contradiction. The latter is concerned about the initial stage of language change when the new grammar emerges, and the former describes the situation after the new grammar has started being used. Adopting the two, this paper argues that the emergence of Grammar 2 (the grammar that produces *des-* (high)) is triggered by learners' abductively de-categorizing or overgeneralizing a set of vocaublary insertion rules, and the traditional grammar and the new grammar started being in competition.

Notice that one important characteristic of the AH rules in (76) is that there is no default rule. Although these vocaublary insertion rules are empirically learnable, the lack of a default rule is not what is assumed in a textbook scenario (Halle & Marantz 1993). With this in mind, we can reasonably assume that the emergence of new grammar is motivated by children's looking for a default rule; that is, in pursuit of a default vocaublary insertion rule, children abductively overgeneralized and acquired the rule in (80) via abductive reasoning (Roberts 2007) by synthesizing (76)h and (76)i.

(80) $X_{[AH:+]} \leftrightarrow des/[\_\ Y]$ (where X and Y are an arbitrary category)

When Y is c, the rule in (80) enables *des* to appear in a sister node of a sentence-final particle, as in (81).

(81)   a. $[_{cP}[_{TP}$ Moo   tukare-ta] $[_c$ **desu**-ne].
               already get tired-PST   AH-SFP
           '(We) already got tired.'

      b. $[_{cP}[_{TP}$ Alex-ga   utukusi kat-ta   ] $[_c$ **desu**-ne]].
               Alex-NOM beautiful PRED.COP-PST   AH-SFP
           'Alex was beautiful.'

This overgeneralization, in fact, kills two undesirable issues with one stone. In addition to providing a default vocabulary insertion rule, it can provide a sequence that better-fits the prediction of the Performative Hypothesis (Miyagawa 2017): since the AH-feature can now be directly pronounced at c, agreement between Pred/Σ and c is no longer needed.[33] Consequently, a plain form is inserted in Pred and Σ (*u* is an epenthetic vowel): hence, (81)b is derived, as shown in (82).[34]

(82) $[_{cP} \ldots [_{TP} \ldots [_{PredP} \ldots [_{CAP} \ldots] \ t_i \ ] \ [_T \ [\ Pred_i^{[Sel:\langle CA,1 \rangle]} \ dum.] \ T \ ] \ ] \ [_c \ c_{[AH:+]} \ c \ ] \ ]$

   $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad | \quad\quad\quad\quad\quad | \quad\ | \quad\quad\quad | \quad\quad\quad |$
   $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\ k \quad\quad\quad\quad\ ar \quad ta \quad\quad des \quad\ ne$

But (80) is not unmarked enough to be called a 'genuine' default rule. Exposed to sentences like (82), children are likely to make another overgeneralization by completely removing the distributional requirement, as in (83), enabling *des(u)* to appear in c without a sentence-final marker. This is how (84) (= (23)) is derived.[35]

(83) $X_{[AH:+]} \leftrightarrow des$

(84) $[_{cP}[_{TP}$ Alex-ga   utukusi   kat-ta   ] $[_c$ **desu**]].
         Alex-NOM beautiful PRED.COP-PST   AH
       'Alex was beautiful.'

As a result of these abductive reasonings (a sequence of overgeneralizations), a new grammar emerged in the colloquial register, where the following rules are at our disposal:

(85)   a. $Pred_{[AH:-]}$ $\leftrightarrow de$ (Grammar 2)
       b. $Pred_{[AH:-]}$ $\leftrightarrow d/[\_ dummy]$
       c. $Pred_{[Sel:\langle CA, 1 \rangle], [AH:-]}$ $\leftrightarrow ku$
       d. $Pred_{[Sel:\langle CA, 1 \rangle], [AH:-]}$ $\leftrightarrow k/[\_ dummy]$
       e. $\Sigma_{[NEG:-]}$ $\leftrightarrow \varnothing$
       f. $\Sigma_{[NEG:+]}$ $\leftrightarrow nai$

---

[33] This chain reduction could be a side-effect of pursuing an economical grammatical architecture (Chomsky 2005), and the diachronic loss of Agree has been discussed in the literature (Roberts 2007: 123, 193). However, the detailed mechanism needs further elaboration in future research: since agreement applies before vocabulary insertion, one cannot tell (in advance) how the AH-feature is pronounced at the time when Agree is ready to apply, which appears to cause an ordering paradox.

[34] (80) predicts that when *des-* can also be distributed in positions other than T/c. This prediction is borne out by a phrase-final particle construction. Japanese can put a particle in every phrase-boundary. When politely used, *des* is used with each phrase-final particle:

   (i)   [Watasi-wa **desu-ne**] [totemo **desu-ne**] tukare-ta   desu-yo.
         Alex-TOP   AH-PFP   very   AH-PFP   get tired-PST AH-SFP
         'I got so tired. '

[35] Considering the phrase-final particle (see Footnote 34), $X_{[AH:+]}$ is more reasonable than $c_{[AH:+]}$.

g. $\Sigma_{[\text{NEG}:+]}$ $\quad\quad\quad\quad$ $\leftrightarrow$ *en* /[X$_{[\text{AH}:+]}$ __]

h. $\mathbf{X}_{[\text{AH}:+]}$ $\quad\quad\quad\quad$ $\leftrightarrow$ ***des***

i. $\mathbf{X}_{[\text{AH}:+]}$ $\quad\quad\quad\quad$ $\leftrightarrow$ *mas* /[__ $\Sigma_{[\text{Sel}:\langle v,\,1\rangle]}$]

j. $\mathbf{T}_{[\text{PST}:+]}$ $\quad\quad\quad\quad$ $\leftrightarrow$ *ta*

k. *dummy* $\quad\quad\quad\quad$ $\leftrightarrow$ *ar*

Lastly, why does the reanalysis cause a genre-sensitive grammar competition, instead of Grammar 2 abruptly replacing Grammar 1? Although this paper does not have a decisive answer, we can imagine a possible scenario: language acquisition starts in a casual register with children surrounded by caretakers, but as they grow up, children become more familiar with more prescriptive situations, for example, by entering school or by being exposed to adult conversations not necessarily directed to them. The social pressure to use the prescriptive grammar of earlier generation forces them to relearn a more restrictive rule, making the vocabulary insertion sensitive to the category, and they may end up using Grammar 1 for formal situations, as well as maintaining Grammar 2 for casual situations; hence, an individual has competing grammars relativized by registers. Thinking this way, we can answer why, in general, a new grammar is found in casual registers, not the other way round. Speakers who do not like casual registers may not accept using Grammar 2, because it sounds cutesy or jokey. Hence, there emerge an inter-speaker variation.

In sum, the emergence of Grammar 2 is boiled down to the problem of acquiring a set of new vocabulary insertion rules: (76)h and (76)i are replaced by (85)h.[36] After Grammar 2 is established, Grammars 1 and 2 are in competition within a single speaker, just in the case of the *dove/dived*-alternation. As previously mentioned, the competition is subject to genres and registers in which an utterance is produced, but it is also affected by other factors. This point will be discussed in more detail in the next section.

4 Verifying the hypothesis

Under the proposed analysis, instances with a sentence-final particle are interpreted as providing a 'catalytic' environment facilitating the overgeneralization — because (80) is assumed as an intermediate step prior to (83). If so, it is predicted that there should be a stage where *des-* (high) is only permitted when there is a sentence-final particle. Even though we cannot access the I-language of speakers in the past, we can alternatively examine E-languages of the given period to see if there is any indication that the new use truly started out from such environments.

In fact, this prediction is partially verified by previous quantitative studies (Tanomura 1994; Noda 2004; Ozaki 2004; Kawaguchi 2014), which all show that the presence of a sentence-final

---

[36] One could ask what other morphological rules are responsible for multiple grammars. Although this paper does not have a decisive answer, proposals for synchronic variations — such as probabilistic impoverishment rules (Nevins 2008; Nevins & Parrott 2010; Havenhill 2016) — are also other candidates that deserves our attention.

particle facilitates people's producing of the new form. Yet the findings of the previous literature have some room for improvement.

First, the data examined in the literature are samples produced in the late twentieth and twenty-first centuries. The change from Grammar 1 to Grammar 2 is, however, estimated to take place in the early twentieth century (Kawaguchi 2014). To assess and verify our analysis in Section 3, we need to observe the data from much earlier days.

Second, the methods used in the literature do not utilize appropriate methods in the inferential statistics, preventing us from making an inference of the population beyond the samples accidentally at our disposal.

Third, without an elaborate statistical model, it is quite hard to tease an effect of one factor from the other. The use of a new form is quite likely to be simultaneously affected by multiple factors. For example, genres may affect the use as well as the use of *-ne*. We, thus, need to separate the influence of such extralinguistic factors from intralinguistic factors, so we can avoid making the wrong inference about the extent to which each factor contributes to the change in language.

For these reasons, we cannot conclude that our prediction is fully supported by the literature. However, by closely examining data from the first half of the twentieth century with elaborate statistical reasoning, this section now shows that our prediction is borne out.

4.1 Data

To look at examples from the first half of the twentieth century, the current paper examines the tokens in the CHJ (Corpus of Historical Japanese), one of the largest corpora covering the early twentieth century (National Institute for Japanese Language and Linguistics, version 2021.3; Last accessed Nov 24, 2021) — and investigates the alternation between the *gozai-mas*-form and the *des*-form, as shown below.

(86)   a.   [$_{cP}$[$_{TP}$ Alex-ga   utukusyuu **gozai-masi**-ta]   (**ne**)].
                Alex-NOM beautifully exist.AHU-AH-PST SFP

        b.   [$_{cP}$[$_{TP}$ Alex-ga   utukusi   k-at-ta]   **des**-u(-**ne**)].
                Alex-NOM beautiful PRED-COP-PST AH-u-SFP
          'Alex was beautiful.'

Accessed through Chunagon 2.5.2, the instances of (86) are obtained with the following query formulae:

(87)   **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

(88)   **New form**:

    a.   POS LIKE "Canonical Adjective%" AND FOLLOWING WORDS: LEXEME *des-* ON 1 WORDS 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

The results, however, contain unwanted examples, so we restrict our results to a subset. First, to avoid unintended homophonous instances from the earlier period, the current study only focuses on the observations from the Meiji era and onward (1868–1947). Second, since the annotation scheme of the CHJ classifies some negation markers as a canonical adjective, we exclude such spurious cases.[37] Finally, to accurately make an inference on the random-effects, adjectives whose raw frequency is less than five are excluded; the remaining 1,031 instances (32 canonical adjectives) are used to estimate the parameters of the statistical models (see also Appendix A).

4.2 Statistical modeling

The diachronic development of the new form against the old prescriptive variant is reasonably modeled as a type of (i) logistic regression, one having (ii) a mixed-effects component combined with (iii) a time series structure.

---

[37] Apparently *nai* and *utukusi i* look similar regarding the conjugation (they both result in a sequence of *k-at-ta*), so one might wish to treat the sentence in (i)a as an instance of a canonical adjective in Grammar 2. But there is good reason to believe that this is not the case.

(i)   a.   [$_{cP}$[$_{TP}$[$_{ΣP}$[$_{PredP}$ Alex-ga   gakusya de]-wa   **nak**] at-ta]   des-u].
                  Alex-NOM scholar PRED-FOC NEG COP-PST AH-u
             'Alex was not a scholar.'

    b.   [$_{cP}$[$_{TP}$[$_{ΣP}$[$_{PredP}$ Alex-ga   gakusya de]-wa   ari-mas-**en**] desi-ta]].
                  Alex-NOM scholar PRED-FOC COP-AH-NEG AH-PST
             'Alex was not a scholar.'

First, the corresponding sentence in Grammar 1 is given in (i)b, which is predicated by a noun. If one treated (i)a as a canonical adjective construction, it would follow that the sentence changes the sentence type depending on whether it is in Grammar 1 or not. Second, a canonical adjective takes the adverbial form when used in Grammar 1, as in (17), but *nak* in (i) does not. Note, however, that this does not mean that *nai* completely lacks the use of a canonical adjective. When used alone, that is, not being preceded by a PredP, it can be used as an adjective that means 'absent.' The *nai* in (i) is, thus, seen as a negation marker, and the one in (ii) as a canonical adjective. In the present study, I manually distinguished all the instances of *nai*, and exclude the former cases only; however, the frequency of such examples is less than five, and as a result, the remaining examples are ruled out based on an independent criterion.

(ii)   a.   [$_{cP}$[$_{TP}$[$_{PredP}$ Sato-ga   **na**   ] k-at-ta]   des-u].
                  sugar-NOM absent   PRED-COP-PST AH-u
             'There was no sugar (lit., sugar was absent).'

    b.   [$_{cP}$[$_{TP}$[$_{PredP}$ Sato-ga   **noo**   ] gozai-masi-ta]   ].
                  sugar-NOM absently AHU.exist-AH-PST
             'There was no sugar (lit., sugar existed absently).'

In the linguistic literature and other fields, logistic regressions are used in modeling the choice of binary variants (Cedergren & Sankoff 1974; Levshina 2018; Yamada 2019b); in the current study, the following variables are used as fixed-effects independent variables:[38]

(89)   a.   **Textbook** x_1: Dummy variable taking the value of 1 iff the sample is taken from a textbook, and 0 otherwise.
   b.   **Literature** x_2: Dummy variable taking the value of 1 iff the sample is taken from the literature, and 0 otherwise.
   c.   **Sentence-final particle** x_3: Dummy variable taking the value of 1 iff the AH is used with a sentence-final particle, and 0 otherwise.
   d.   **Epistemic model suffix** x_4: Dummy variable taking the value of 1 iff the AH is used with an epistemic modal suffix -oo, and 0 otherwise.
   e.   **Tense** x_5: Dummy variable taking the value of 1 iff the AH is used in a past tense, and 0 otherwise.

The CHJ provides information about the genre from which each sentence is sampled: whether it is from a textbook, the literature (e.g., novels and essays), or the non-literature (the remaining genres).[39] We create two dummy variables that indicate whether it is from a textbook and the literature (x_1 and x_2), making the remaining miscellaneous non-literature genre as the reference category. The variables in (89)c through (89)e are designed to capture the linguistic factors. To test the predictions from the previous section, we use dummy variables for a sentence-final particle and epistemic marker. For example, if *ne* is pronounced in (86)a or (86)b, x_3 is coded 1, and 0 otherwise. Since the variants can take both present tense and past tense sentences, the tense variable is created to see the difference (x_5).

Furthermore, the following random-effects variable is taken into consideration to scrutinize the idiosyncrasies among canonical adjectives (Levshina 2018; Yamada 2019b).

(90)   **Idiosyncrasies among canonical adjectives** $u_{01}, u_{02}, \ldots, u_{032}$: The idosyncrasy of the *j*-th adjective, which is assumed to follow $N(0, \tau^2)$.

In addition, to incorporate chronological information, the intercept is assigned a time-series structure (the State-Space Model; Shumway & Stoffer 2017; Durbin & Koopman 2001; Hagiwara

---

[38] Each sentence of this corpus is provided with an annotation about the genre (textbooks, literature, and nonliterature) and the neighboring morphemes; based on this annotation scheme, the variables $x1$, $x2$, $x4$, and $x5$ are created. As for $x3$, the researcher manually annotated the information for all instances.

[39] The textbook in this data set consists of government-designed textbooks on Japanese published from 1904 to 1947. Rather than explicitly teaching grammar rules, they comprise essays and writings for elementary school (the 1st to the sixth year) and the first year middle school students. Such censored writings are reasonably considered representing the prescriptive grammar at that time. The literature includes novels and essays written not for the purpose of promoting the prescriptive grammar, but for general purposes, reasonably seen as being sensitive to the use of day-to-day conversation. The non-literature genre consists of miscellaneous remaining categories (e.g., magazines).

2021, a.m.o). Here each intercept (the state) at time t is designed to depend only on the state at the previous time point (i.e., a random walk).

These structural components are combined to yield to the model provided in (91); $y_{ij}^{(t)}$ and $\pi_{ij}^{(t)}$, respectively, denote the dependent variable and parameter of the Bernoulli distribution of the i-th observation of the j-th adjective at time t; $y_{ij}^{(t)} = 1$ iff it takes the new variant, and 0 otherwise. $x_{1i}^{(t)}, x_{2i}^{(t)}, \ldots, x_{5i}^{(t)}$ are the values of the aforementioned independent variables at time t, and $\beta_1, \beta_2, \ldots, \beta_5$ are the corresponding effect sizes; $\beta_0^{(t)}$ is the intercept at t. u_{0j} ($j \in \{1, 2, \ldots, 32\}$) are the random effects representing the idiosyncrasies attributed to the j-th adjective.

(91)
$$y_{ij}^{(t)} \sim \text{Bern}(\pi_{ij}^{(t)})$$
$$\pi_{ij}^{(t)} \sim \text{logistic}(\eta_{ij}^{(t)})$$
$$\eta_{ij}^{(t)} = \beta_0^{(t)} + \beta_1 x_{1i}^{(t)} + \beta_2 x_{2i}^{(t)} + \beta_3 x_{3i}^{(t)} + \beta_4 x_{4i}^{(t)} + \beta_5 x_{5i}^{(t)} + u_{0j}$$
$$u_{0j} \sim N(0, \tau^2)$$
$$\beta_0^{(t)} \sim N(\beta_0^{(t-1)}, \sigma_\zeta^2)$$

The parameters are estimated using a Hamiltonian Monte Carlo algorithm with Stan (Stan Development Team 2020) on R (Gelman et al. 2013; R Core Team 2020). The model comparison is conducted by the R package `loo`, which provides us with an efficient approximate leave-one-out cross-validation information criterion (LOOIC), approximate standard errors for the estimated predictive errors, and the Watanabe-Akaike/widely applicable information criterion (WAIC). As a result, among the 2^5 (five fixed- and random-effects variables) $= 32$ constructed models, the model in which x_5 is unselected turns out the best model (see Appendix B). The subsequent discussion proceeds based on the results of the posterior distribution of the parameters of this model.

4.3 Results

Figure 1 summarizes the results of the estimated random-effect of each adjective (i.e., u_{0j}'s). Each bar represents the 95% and the 66% credible intervals of the random-effects u_{0j} (the white circles are the posterior medians and the number in the square brackets shows how many times the adjective is used with an AH form in the CHJ). *Arigata* 'rare' has the largest negative value, showing its strong preference for the prescriptive. However, the results are not surprising, because *arigatoo gozai-mas* is entrenched as a fixed expression meaning *thank you,* which should reasonably increase the relative ratio of the old variant. On the other hand, there are no adjectives with an extremely large positive value (a candidate leading to a change). The largest is *atu* 'hot,' but the posterior median is not as high as $+2$; although the upper limit of the 95% credible interval slightly goes beyond the commonly-assumed $+3$ threshold, the sample size for this adjective is as many as twelve instances, and with few more examples, it is expected that
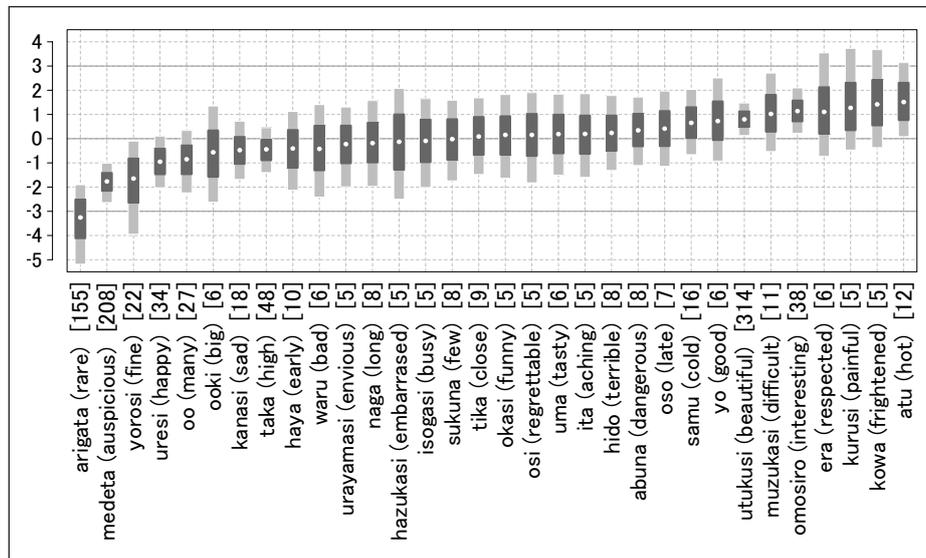
Figure 1: Posterior Distributions for the Random-Effects.

the posterior range becomes smaller. It is, thus, reasonable to conclude that other than *arigata*, the adjectives are almost randomly distributed with no significant idiosyncratic tendency in the change in language.

Figure 2 shows the posterior distribution of $\pi^{(t)}$ from 1873 to 1947. For each year, the posterior median (the solid line) and its 95% credible intervals (the shaded gray area) are drawn. The clear trend of increasing values corroborates the findings of previous studies that the new variant is gradually becoming popular in the first half of the twentieth century (Kawaguchi 2014).[40]

Figure 3 represents the results of the posterior distributions of the fixed-effects. β_1 and β_2 are text genres, which are associated with formality and other sociolinguistic factors. The textbook (β_1) takes a large negative value, suggesting that examples in this genre are resistant to a change; given that textbooks are written maximally respecting the prescriptive grammar, its preference for the old variant seems reasonable. In contrast, the literature (β_2) (novels and essays) has a positive value, suggesting examples in this genre prefer taking the new form. Since novels and essays reasonably reflect everyday conversations to a greater degree, this tendency is also understandable.[41]

---

[40] One unexpected move is the 'spike' between 1897 and 1907, which may be attributed to the idiosyncrasies of the sampled texts. It could, however, be caused by a social movement: in those days, some novelists started writing novels in a colloquial register (aka the Genbun-Icchi movement, that is, the movement to unify the written and spoken forms of a language; Yamamoto 1971, Kamei et al. 1965: 184, Nomura 2013: 269). However, to convincingly prove the influence, finer-grained analysis would be necessary, which I would like to leave open to future research.

[41] Although it does give us a natural interpretation, we need to be aware that the effect size of this variable is not as strong as the others. Furthermore, in terms of LOOIC and WAIC, the second best model in the model selection lacks this variable (see **Figures 4** and **5**). Thus too much emphasis on the difference would not be helpful; it is the least important among the four fixed-effects factors.
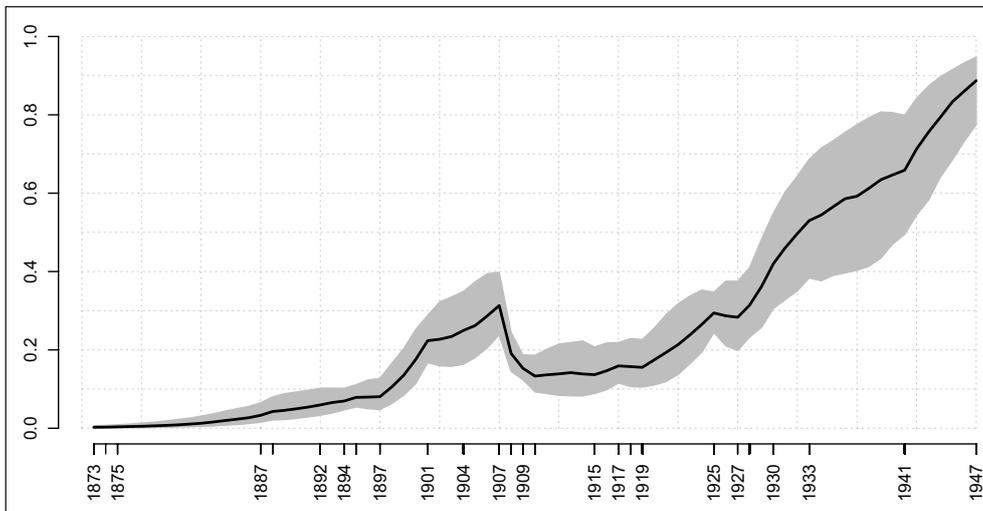
Figure 2: Posterior Inference on $\{\pi_0^{(t)} : t \in \{1873, 1874, \ldots, 1947\}\}$.
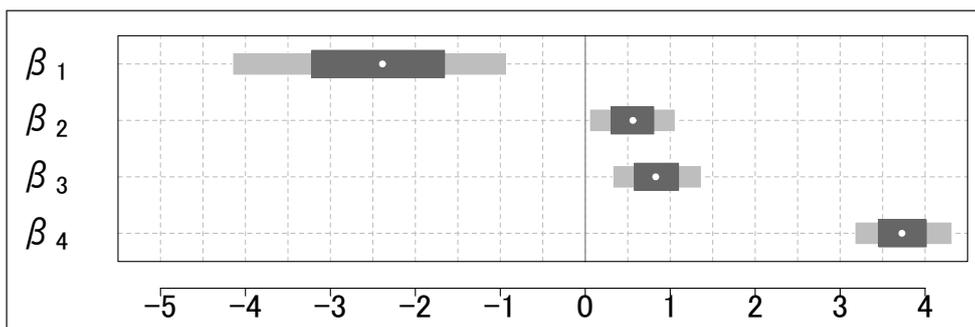
Figure 3: Posterior Distributions for the Fixed-Effects.

Finally, let us turn to intralinguistic factors. First, tense (x_5) is excluded from the best model in (91), suggesting that when other factors are controlled, it is not useful in predicting the use of new variants: the spread of the new form is not sensitive to the present/past tense distinction.[42] Second, the extreme large value for β_4 (an epistemic marker) stems from the fact that *-oo* is used with des- (high), not with -mas, as discussed in Section 2.3.3. Lastly, what is most relevant for our discussion is β_3 (a sentence-final particle). As clearly seen, it has a positive effect size, whose 95% credible intervals are far from 0, strongly supporting our prediction that a new form is facilitated in the structure of [X$_{[AH:+]}$ SFP (c)].[43]

---

[42] Even when we include x_5 and refit the model, the posterior median of the past tense shows a slightly negative value but the 95% credible interval includes 0.

[43] Due to the limitation of the corpus used for this study which does not annotate the speaker of each instance, intraspeaker variations cannot be detected. With a fully annotated-corpus, future research is expected to improve the model by including the variable to such variations.

5 Conclusion

Grammar 1 in Japanese is peculiar in that AH markers are not distributed in clause periphery. The language change, however, enabled people to additionally use Grammar 2 in colloquial register, where *des-* is pronounced in clause periphery, in agreement with the prediction of Performative Hypothesis. From the perspective of diachronic syntax, this change is seen as a microparametric change in the Agree operation (activity condition in the sense of Baker (2008)). This paper further develops this direction from a morphosyntactic perspective, by proposing that what is analyzed as a microparametric change can be given a finer-grained explanation that it was driven by children's abductive reasoning in pursuit of default vocabulary insertion rules, thus elucidating the role of morphological, postsytnactic operations in language change.

Several issues are left for future study. First, while this paper zooms in on the emergence of Grammar 2, it is as important to ask how Grammar 1 was established, attempting to answer why the grammar without a default rule was permitted in the early twentieth century to begin with. Second, as already mentioned, among the new variants, affirmative sentences with verbs and noun/nominal adjectives (= (30) and (31)) are relatively reluctant for the change. The current version of our theory does not explain such finer-grained tendency in use. An improvement is needed in this respect. Lastly, our analysis predicts that when a default vocabulary insertion rule is missing, an abductive reasoning is invited in any languages and beyond the AH-markers. Examination of other languages are to verify the direction developed in this current study.

Appendix A: Raw corpus analysis

	1868	1869	1870	1882	1883	1887	1889	1890	1892	1896
Old form	46	1	2	2	1	11	22	45	2	22
New form	0	0	0	0	1	0	9	5	0	27
	1899	1902	1903	1904	1905	1910	1912	1913	1914	1920
Old form	22	8	11	148	20	2	18	14	7	136
New form	1	28	12	103	9	5	31	1	11	118
	1922	1923	1925	1928	1936	1942				
Old form	2	6	2	27	27	15				
New form	0	13	5	6	8	19				

Table 1: Raw frequencies of the two variants in each year.

	yorosi 'fine'	*arigata* 'rare'	*yo* 'good'	*waru* 'bad'	*uresi* 'happy'	*oo* 'many'	*medeta* 'auspicious'	*haya* 'early'
Old form	175	155	79	32	26	23	22	13
New form	33	0	235	16	8	4	0	5
	omosiro 'interesting'	*kanasi* 'sad'	*samu* 'cold'	*abuna* 'dangerous'	*atu* 'hot'	*sukuna* 'few'	*ooki* 'big'	*taka* 'high'
Old form	12	8	6	5	5	5	5	4
New form	26	2	10	3	7	3	1	2
	uma 'tasty'	*hazukasi* 'embarrassed'	*oso* 'late'	*naga* 'long'	*isogasi* 'busy'	*okasi* 'funny'	*tika* 'close'	*hido* 'terrible'
Old form	4	4	4	4	4	3	3	3
New form	4	1	3	1	4	6	2	3
	osi 'regrettable'	*urayama* 'envious'	*ita* 'aching'	*utukusi* 'beautiful'	*muzukasi* 'difficult'	*kowa* 'frightened'		
Old form	3	3	3	3	2	1		
New form	2	2	2	3	9	4		
	kurusi 'painful'	*era* 'respected'						
Old form	0	0						
New form	5	6						

Table 2: Raw frequencies for each canonical adjective.

Appendix B: Statistical Modeling
Priors

This study uses the flat uniform prior, which is the default setting in Stan; for standard deviations, the lower limit is set to 0 so that they do not take negative values.

Convergence

By taking into account all combination of the variables, the present paper builds $2^5 = 32$ models. Under the condition of THIN = 15, ITER = 25000, and WARMUP = 18000, most models are converged under the commonly-used threshold $\hat{R} < 1.01$ (Vehtari et al. 2021).

Model Comparison

The performance of their models are assessed using Leave-One-Out Cross-Validation (LO-OIC) and the Watanabe–Akaike/Widely Applicable Information Criterion (WAIC), which are methods used in estimating pointwise out-of-sample prediction accuracy from a fitted Bayesian model (Watanabe 2010; Vehtari et al. 2017), which are computed by loo package in R (Vehtari et al. 2016). The results are summarized in **Figures 4** and **5**; the models are ordered in increasing order. The barcode plots in the lower panel represent the variables included in the given model. For example, the leftmost models contain the variables from x_1 to x_4 lacking x_5. The model with the lowest LOOIC/WAIC is the best model, so this leftmost model is identified as the best model.
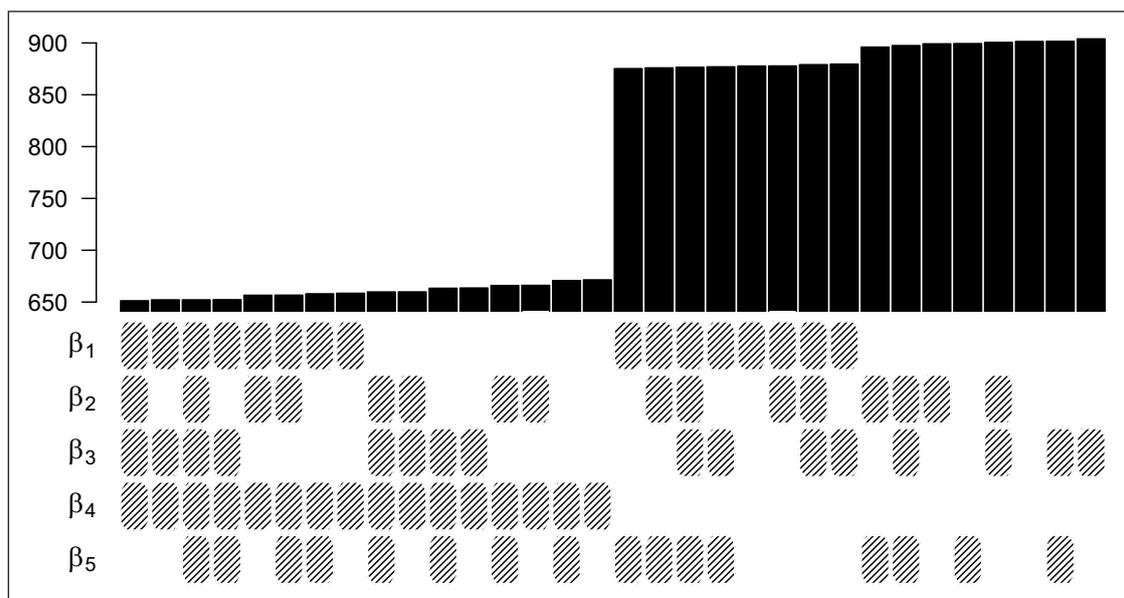
Figure 4: The LOOIC of each model.
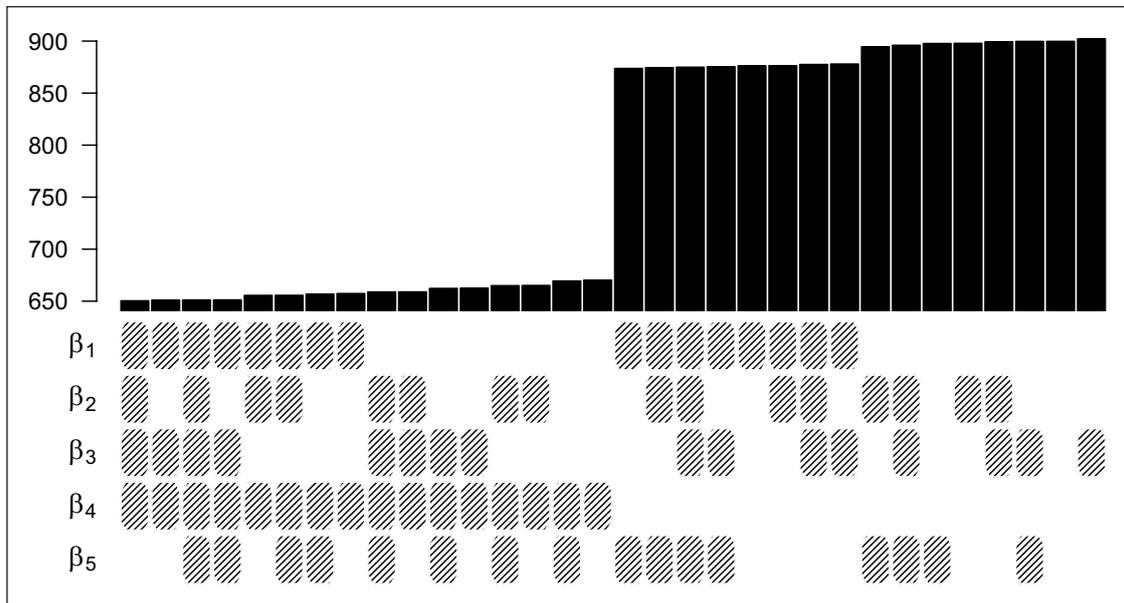
Figure 5: The WAIC of each model.

Abbreviations

This paper uses the Leipzig glossing rules, with the following additions: ADN = adnomial form, AH = addressee-honorific, AHU = addressee-honorific upgrader, COP = dummy copula, CV = converb suffix, EPI = epistemic marker, FORM = formal, PRED = predicative copula, SFP = sentence-final particle

Funding information

Acknowledgements

Competing interests

The author has no competing interests to declare.

References

Adger, David. 2003. *Core syntax: a minimalist approach*. Oxford: Oxford University Press.

Adger, David. 2010. A Minimalist theory of feature structure. In *Features: perspectives on a key notion in linguistics*, 185–218. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199577743.003.0008

Alok, Deepak. 2021. The morphosyntax of Magahi addressee agreement. *Syntax* 24. 263–296. DOI: https://doi.org/10.1111/synt.12213

Amaral, Luiz & Roeper, Tom. 2014. Multiple grammars and second language representation. *Second Language Research* 30(1). 3–36. DOI: https://doi.org/10.1177/0267658313519017

Andersen, Henning. 1973. Abductive and deductive change. *Language* 49. 765–793. DOI: https://doi.org/10.2307/412063

Baier, Nicholas. 2015. Adjective agreement in Noon. In *Proceedings of NELS 45*, vol. 1. 67–80.

Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511619830

Barbiers, Sjef. 2013. Micorsyntactic variation. In *The Cambridge handbook of generative syntax*, 899–926. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511804571.031

Biberauer, Theresa & Roberts, Ian. 2012. Towards a parameter hierarchy for auxiliaries: diachronic considerations. *Cambridge Occasional Papers in Linguisrics* 6. 267–294.

Cedergren, Henrietta J. & Sankoff, David. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50(2). 333–355. DOI: https://doi.org/10.2307/412441

Chomsky, Noam. 1974. The Amherst Lectures, unpublished lecture notes distributed by Documents Linguistiques. University of Paris VII.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Martin, Roger & Michaels, David & Uriagereka, Juan (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT press.

Chomsky, Noam. 2001. Derivation by phase. In Kenstowicz, Michael (ed.), *Ken Hale: a life in language*, 1–52. Cambridge, MA: MIT press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22. DOI: https://doi.org/10.1162/0024389052993655

Cinque, Guglielmo. 1999. *Adverbs and functional heads*. Oxford University Press: Oxford.

Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, 42–64. Oxford: Blackwells. DOI: https://doi.org/10.1002/9780470755662.ch3

Durbin, J. & Koopman, S. J. 2001. *Time series analysis by state space methods*. Oxford: Oxford University Press.

Embick, David. 2008. Variation and morphosyntactic theory: competition fractionated. *Language and Linguistics Compass* 2(1). 59–78. DOI: https://doi.org/10.1111/j.1749-818X.2007.00038.x

Embick, David & Noyer, Rolf. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595. DOI: https://doi.org/10.1162/002438901753373005

Fuss, Eric & Trips, Carola. 2002. Variation and change in Old and Middle English: on the validity of the Double Base Hypothesis. *Journal of Comparative Germanic Linguistics* 4. 171–224. DOI: https://doi.org/10.1023/A:1016588502117

Gelman, Andrew & Carlin, John B. & Stern, Hal S. & Dunson, David B. & Vehtari, Aki & Rubin, Donald B. 2013. *Bayesian data analysis [3rd edition]*. London: CRC Press. DOI: https://doi.org/10.1201/b16018

Giorgi, Alessandra. 2010. *About the speaker: towards a syntax of indexicality*. Oxford: Oxford Universtiy Press.

Haddican, Bill. 2018. The syntax of Basque allocutive clitics. *Glossa: a journal of general linguistics* 3(1). DOI: https://doi.org/10.5334/gjgl.471

Haegeman, Liliane & Hill, Virginia. 2013. The syntacticization of discourse. In Folli, Raffaella R. & Sevdali, Christina & Truswell, Robert (eds.), *Syntax and its limits*, 370–390. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199683239.003.0018

Hagiwara, Junichiro. 2021. *Time series analysis for the state-space model with r/stan*. Singapore: Springer. DOI: https://doi.org/10.1007/978-981-16-0711-0

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and teh pieces of inflection. In Hale, Ken & Keyser, Samuel Jay (eds.), *The view from Building 20: essays in linguistics in honor of Sylavain Bromberger*, 111–176. Cambridge, MA: MIT press.

Harada, S.-I. 1976. Honorifics. In Shibatani, Masayoshi (ed.), *Syntax and semantics*, vol. 5, 499–561. New York: Academic Press. DOI: https://doi.org/10.1163/9789004368835_013

Harris, Alice C. & Campbell, Lyle. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge University Press: Cambridge. DOI: https://doi.org/10.1017/CBO9780511620553

Havenhill, Jonathan. 2016. Relative clauses in Bavarian: A Distributed Morphology approach to morphosyntactic variation. In *The proceedings of the 50th annual meeting of the Chicago Linguistics Society*, 249–264.

Hill, Virginia. 2014. *Vocatives: how syntax meets with pragmatics*. Leiden and Boston: Brill. DOI: https://doi.org/10.1163/9789004261389

Hiraiwa, Ken & Ishihara, Shinichiro. 2002. Missing links: cleft, slucing, and "no da" construction in Japanese. *MIT Working Papers in Linguistics* 43. 35–54.

Ikawa, Shiori & Yamada, Akitaka. 2022. Territory feature and a distributed-morphology approach to clause periphery. In *Proceedings of The 29th Japanese/Korean Linguistics Conference*, 319–328.

Kamei, Takashi & Yamada, Toshio & Ohto, Tokihiko (eds.). 1965. *Nihongo no rekisi 6 [the history of Japanese 6]*. Tokyo: Heibonsha.

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

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

Kayne, Richard. 2000. *Parameters and universals*. New York: Oxford University Press.

Kayne, Richard S. 2005. Some notes on comparative syntax, with special reference to english and french. In *The Oxford handbook of comparative syntax*, 3–69. Oxford: Oxford University Press.

Kinsui, Satoshi. 2011. Teineigo no gogen to hattatsu. In *Rekisigoyooron nyuumon [an introduction to historical pragmatics]*, 164–173. Tokyo: Taishukan Shoten.

Kramer, Ruth. 2009. *Definite markers, phi-features, and agreement: a morphosyntactic investigation of the Amharic DP*: UC Santa Cruz dissertation.

Kramer, Ruth. 2010. The Amharic definite marker and the syntax-morphology interface. *Syntax* 13. 196–240. DOI: https://doi.org/10.1111/j.1467-9612.2010.00139.x

Kroch, Anthony. 1995. Morphosyntactic variation. In *Proceedings of the 30th Annual Meeting of the Chicago Linguistic Society*, vol. 2. 180–201.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, Mass.: The MIT Press.

Levshina, Natalia. 2018. Probabilistic grammar and constructional predictability: Bayesian generalized additive models of help + (to) infinitive in varieties of webbased english. *Glossa: a journal of general linguistics* 55(3(1)). 1–22. DOI: https://doi.org/10.5334/gjgl.294

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 2006. *How new languages emerge*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511616204

Marantz, Alec. 1984. *On the nature of grammatical relations*. MIT Press: Cambridge, Mass.

McCready, Eric. 2014. A semantics for honorifics with reference to Thai. In *Proceedings of the 28th Pacific Asia conference on language, information and computing*, 503–512.

McFadden, Thomas. 2017. The morphosyntax of allocutive agreement in Tamil. Ms., Leibniz Zentrum Allgemeine Sprachwissenschaft, Berlin.

Mikami, Akira. 1959 [2002]. *Kobun no kenkyu [a study of constructions]*. Tokyo: Kurosio Publishers.

Minami, Fujio. 1974. *Gendai nihongo no koozoo [the structure of contemporary Japanese]*. Tokyo: Taishukan Shoten.

Miyagawa, Shigeru. 2012. Agreements that occur mainly in the main clause. In Aelbrecht, Lobke & Haegeman, Liliane & Nye, Rachel (eds.), *Main clause phenomena: New horizons*, 79–112. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.190.04miy

Miyagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/10958.001.0001

Miyama, Mioko. 2011. Da and the zero form as the two contracted forms of the Japanese copula. *Linguistic Research: Working Papers in English Linguistics* 27. 41–61.

Nakamura, Michio. 1948. *Tookyoo-go no seikaku [characteristics of the Tokyo Japanese]*. Tokyo: Kawada Shobo.

Nevins, Andrew. 2008. Cross-modular parallels in the study of phon and phi. In Harbour, Daniel & Adger, David & Bèjar, Susana (eds.), *Phi-theory: Phi-features across modules and interfaces*, 329–368. Oxford: Oxford University Press.

Nevins, Andrew & Parrott, Jeffrey K. 2010. Variable rules meet Impoverishment theory: patterns of agreement leveling in English varieties. *Lingua* 120(5). 1135–1159. DOI: https://doi.org/10.1016/j.lingua.2008.05.008

Nishiyama, Kunio. 1997. what does the copula do? In *University of pennsylvania working papers in linguistics*, vol. 4. 227–243.

Nishiyama, Kunio. 1999. Adjectives and the copulas in japanese. *Journal of East Asian Linguistics* 8. 183–222. DOI: https://doi.org/10.1023/A:1008395915524

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]. *Mathematical Linguistics* 24(5). 228–244.

Nomura, Takeshi. 2013. *nihongo sutandaado no rekisi [the history of standard Japanese]*. Tokyo: Iwanami Publishers.

Norris, Mark. 2014. *A theory of nominal concord*: UC Santa Cruz dissertation.

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

Oyharçabal, Bernard. 1993. Verb agreement with nonarguments: On allocutive agreement. In de Urbina, Jon Ortiz & Hualde, José (eds.), *Generative studies in Basque linguistics*, 89–114. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.105.04oyh

Ozaki, Natsu. 2004. Hitei no teineikei "naidesu" to "masen" ni tuite [on the polite negative forms "naidesu" and "naidesu"]. *Okayama Daigaku Gnegogaku Ronso* 11. 29–42.

Pak, Miok D. 2008. Types of clauses and sentence end particles in korean. *Korean Linguistics* 14(1). 113–156. DOI: https://doi.org/10.1075/kl.14.06mdp

Parrott, Jeffrey K. 2007. *Distributed morphological mechanisms of Labovian variation in morphosyntax*. Washington D.C.: Georgetown University dissertation.

Paul, Hermann. 1920. *Prinzipien der Sprachgeschichte [principles of the history of language]*. Tübingen: Niemeyer.

Pesetsky, David & Torrego, Esther. 2007. The syntax of valuation and the interpretability of features. In Karimi, Samin & Samiian, Vida & Wilkins, Wendy K. (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*, 262–294. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.101.14pes

Pintzuk, Susan. 1999. *Phrase structures in competition: variation and change in Old English word order*. New York: Garland.

Pintzuk, Susan. 2002. Verb–object order in Old English: variation as grammatical competition. In *Syntactic effects of morphological change*, 276–299. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199250691.003.0016

Portner, Paul & Pak, Miok & Zanuttini, Raffaella. 2019. The speaker-addressee relation at the syntax-semantics interface. *Language* 95(1). 1–36. DOI: https://doi.org/10.1353/lan.2019.0008

R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris. DOI: https://doi.org/10.1515/9783110883718

Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.

Sato, Toru. 1990. *Edozidaigo no kenkyu [a study of Edo period Japanese]*. Tokyo: Ohfusha.

Shumway, Robert H. & Stoffer, David S. 2017. *Time series analysis and its applications: with R examples*. New York: Springer. DOI: https://doi.org/10.1007/978-3-319-52452-8

Simonenko, Alexandra & Carlier, Anne. 2020. Between demonstrative and definite: a grammar competition model of the evolution of French l-determiners. *Canadian Journal of Linguistics* 65(3). 393–437. DOI: https://doi.org/10.1017/cnj.2020.14

Stan Development Team. 2020. *Stan modeling language users guide and reference manual version*. 2.21.0, https://mc-stan.org.

Stowell, Timothy Angus. 1981. *Origins of phrase structure*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

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.

Truswell, Robert. 2021. Grammar Competition and word order in a North Early Middle English text. *Languages* 6(59). 1–29. DOI: https://doi.org/10.3390/languages6020059

Vehtari, Aki & Gelman, Andrew & Gabry, Jonah. 2016. *loo: Efficient leaveone-out cross-validation and WAIC for Bayesian models. R package version 0.1.6*. https://github.com/stan-dev/loo.

Vehtari, Aki & Gelman, Andrew & Gabry, Jonah. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27. 1413—1432. DOI: https://doi.org/10.1007/s11222-016-9696-4

Vehtari, Aki & Gelman, Andrew & Simpson, Daniel & Carpenter, Bob & Bürkner, Paul-Christian. 2021. Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis* 16(2). 667–718. DOI: https://doi.org/10.1214/20-BA1221

Watanabe, Akira. 2013. Non-neutral interpretation of adjectives under measure phrase modification. *Journal of East Asian Linguistics* 22(3). 261–301. DOI: https://doi.org/10.1007/s10831-013-9103-5

Watanabe, Sumio. 2010. Asymptotic equivalence of Bayes cross validation and Widely Applicable Information Criterion in singular learning theory. *Journal of Machine Learning Research* 11. 3571–3594.

Yamada, Akitaka. 2019a. Phase-based prosody: evidence for pitch-accent distribution in the Japanese verbal domain. In Baird, M. & Pesetsky, J. (eds.), *Proceedings of NELS 49*, 305–318. GLSA.

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

Yamada, Akitaka. 2019c. *The syntax, semantics and pragmatics of Japanese addresseehonorific markers*. Washington DC: Georgetown University dissertation.

Yamada, Akitaka. 2020. Do-support and be-support as defective movements. In *Proceedings of the 55th annual meeting of the chicago linguistic society*, 435–450.

Yamamoto, Masahide. 1971. *Genbun itti no rekisi ronkoo*. Tokyo: Ohfusha.

Yang, Charles D. 2000. *Knowledge and learning in natural language*. Cambridge, MA: MIT dissertation.

Yuzawa, Koukichirou. 1954. *Edo kotoba no kenkyu [a study of Edo period language]*. Tokyo: Kazama Shobo.

Yuzawa, Koukichirou. 1982. *Tokuga zidai gengo no kenkyu [a study of languages in the Tokugawa period]*. Tokyo: Kazama Shobo.

Zu, Vera. 2018. *Discourse participants and the structural representation of the context*. New York: New York University dissertation.