Syntax and semantics of aspectual constructions in Japanese: Defective T and habituality

Akitaka Yamada, Georgetown University

• **Introduction:** In Japanese, aspectual markers such as *hazime-* 'start,' *tuduke-* 'continue,' and *oe-* 'finish,' can take two different types of complements, as in (1)a (= **Type 1** use) and (1)b (= **Type 2** use). In order to highlight the difference from another construction (aka., the lexical VV-compound), Type 1 and Type 2 are considered synonyomus (Shibatani 1973; Kuno 1983; Matsumoto 1996; Yumoto 2005; Fukuda 2012; Kageyama 1993, 2016) and their difference has not been seriously examined. Exploring the syntax and the semantics of these constructions, this paper argues that Type 1 is different from Type 2 in that Type 1 has an embedded CP-TP_{def} layer, which (i) yields the control construction and (ii) precludes non-habitual readings.

• **Data:** There are four important differences between Type 1 and Type 2. **[FACT 1]** Type 2 use has a reading not available for Type 1. In addition to the **habitual reading** (the habit of his teaching English had begun) — this reading is shared by both (1)a and (1)b —, the **single-event reading** is available for (1)b, but not for (1)a; there is a single seamless event of teaching and the reference time is set to the beginning of this single event. **[FACT 2]** The subject of Type 1 has to be an AGENT, while Type 2 does not have this restriction. (2)a and (3) are illicit because the subject is not the AGENT. **[FACT 3]** In Japanese, %LH...LH% pitch accent is assigned to the constituent in the sisternode of T (Yamada 2018). *Hazime-* and *osie-* form such a single pitch contour in Type 2, but not in Type 1. **[FACT 4]** (4) shows that only Type 1 allows an adverbial intervention between *koto* and *hazime*.

(1) a. *Kare-wa* [*eigo-o* *osier-u* *koto*]-*o* ***hezime***-*ta*. **Type 1**
he-TOP English-ACC teach-PRS C-ACC **begin**-PST

b. *Kare-wa* [*eigo-o* *osie*]-***hazime***-*ta*. **Type 2**
he-TOP English-ACC teach-**begin**-PST
'He began teaching English.'

(2) a. **Kane-wa* [*nar-u* *koto-o*] ***hazime***-*ta*. b. *Kane-wa* [*nari*]-***hazime***-*ta*.
bell-TOP ring-PRS C-ACC **begin**-PST bell-TOP ring-**begin**-PST
'The bell began ringing (intended).' **Type 1** 'The bell began ringing.' **Type 2**

(3) a. **Eigo-ga* [*osie-rarer-u* *koto*]-*o* ***hazime***-*ta*. **Type 1**
English-NOM teach-PASS-PRS C-ACC **begin**-PST

b. *Eigo-ga* [*osie-rare*] ***hazime***-*ta*. **Type 2**
English-NOM teach-PASS **begin**-PST
'English started being taught.'

(4) a. *Kare-wa* [*eigo-o* *osier-u* *koto*]-*o* *kinoo* ***hezime***-*ta*. **Type 1**
he-TOP English-ACC teach-PRS C-ACC yesterday **begin**-PST
'Yesterday, he began teaching English.'

b. **Kare-wa* [*eigo-o* *osie*] *kinoo* ***hazime***-*ta*. **Type 2**
he-TOP English-ACC teach yesterday **begin**-PST
'Yesterday, he began teaching English (intended).'

• **Syntax:** **[FACT 3]** suggests that there is a T-head between the embedded verb and the embedding predicate in Type 1 but not in Type 2. To capture this, I propose the tree in (5) (cf. Fukuda 2012; Kishimoto 2014). Unlike (5)b, the embedded verb in (5)a projects a CP-TP layer. **[FACT 4]** also supports this hypothesis. The fact that *koto*-constituent can precede an adverb that modifies the tense in the main clause suggests that this *koto*-constituent is a possible target of the scrambling. Under the traditional assumption that CP and NP(DP) are the unit that can get preposed by scrambling, we predict that *koto*-complement can be scrambled in front of the adverb while we cannot scramble the VoiceP to the front, yielding the contrast in **[FACT 4]**.

(5) a. [TP [VoiceP he_i [Voice' [vP [CP [TP_{def} [VoiceP PRO_i [vP *teach Eng* Voice]] T_{def}] [C *koto*]] [v v $\sqrt{hazime}$]] Voice]] [T *-ta*]]

b. [TP [AspP [VoiceP he [[vP *teach Eng.*] Voice]] [AspP Asp $\sqrt{hazime}$]] [T *-ta*]]

To account for the other observations, I hypothesize that this embedded T in Type 1 is defective (= tenseless) (Chomsky 2000, 2001), because this clause lacks the tense distinction; the past tense is not allowed in this complement clause. This conclusion gives us two reasonable results. First, we can explain why the subject in Type 1 must be an AGENT (= **[FACT 2]**). In (5)a, TP_{def} has no ability to assign the nominative to the Spec, VoiceP. Thus, PRO needs to appear in this Spec, VoiceP. In addition, the main clause subject in the upper Spec, VoiceP receives the θ-role (AGENT/the person who starts) from Voice (cf. Krazter 1996; or, from the combination of Voice and v+$\sqrt{hazime}$- 'start'). Thus, the referent of he_i is always the AGENT wrt the *starting* event (e.g., the individual who causes/initiates the event of *teaching*). In contrast, (5)b does not have T_{def}. Since there is no split between *he* and PRO, there is only one θ-role that we need to consider; i.e., the theta-role from Voice (or Voice and v+$\sqrt{teach}$). This θ-role depends only on the lexical semantics of the verb $\sqrt{teach}$, which has nothing to do with the verb *start*. Thus, non-Agent θ-role is allowed. Second, we can relate the defective (= tenseless) T with the habituality. Being tenseless, it cannot relate the event(s) with any *particular* reference time. So, the defective T serves as the source of *genericity* in the verbal domain, as convincingly argued by previous studies (Krifka's 1987 I-genericity; Carlson 2011).

• **Semantics:** Based on this, I propose the following semantics. First, along with Krazter (2007) (also, Krifka 1992; Landman 1996; Ferreira 2016), I assume that verbs are born as plurals, where singularities are special cases of pluralities; (6)a is the domain for the eventualities and (7)a is the denotation for VoiceP both in Type 1&2. Second, in Type 1, T_{def} subtracts singular events (= (6)b) from this set (= (6)a), resulting in the set of the non-singular events (= $D_s \setminus D_s^{\text{SG}}$). Thus, we only consider plural events in Type 1. Third, whether it is merged with v or ASP, the chief function of *hazime* is to specify the relation between two intervals. For all the best worlds, there is a relevant period in which an event/state holds (the ***event interval***, EI), such that (i) the ***reference interval*** (RI), i.e., INT(e) in (7)c, is situated within this EI and (ii) the end point of the RI needs to precede the contextually-given threshold, which determines the end point of the 'beginning' part of this interval. In the semantics in (7)c, $\text{MIN}(P(e'))$ and $\text{MAX}(P(e'))$ refer to the initiating/terminating point of the EI and the threshold is the point that divides this EI into $1 : (n-1)$ (where n is the contextually given parameter). This is what Fig 1 shows; the RI is inside this beginning part of the EI (between the blue and the red lines). In Type 1, T_{def} requires us to have multiple events, so the EI consists in several teaching events as illustrated in Fig 2. But Type 2, which lacks T_{def}, does not have such a restriction, allowing a case where the EI only consists in a single event as in Fig 3. This is why only Type 2 is ambiguous (= **[FACT 1]**).

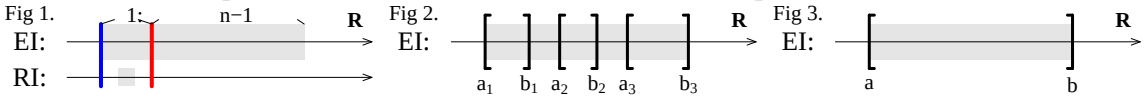
(6) a. $D_s = \{e_1, ..., e_n, e_1 \oplus e_2, e_1 \oplus e_3, e_1 \oplus e_2 \oplus e_3, ...\}$ b. $D_s^{\text{SG}} = \{e_1, e_2, e_3, ..., e_n\}$

(7) a. $[\![$**VoiceP 'he/PRO teach English'**$]\!] = \lambda e \in D_s.{}^*\text{teach}(e) \wedge^* \text{AG}(e, \text{he}) \wedge^* \text{PAT}(e, \text{Eng})$

b. $[\![\text{T}_{\text{def}}]\!] = \lambda P_{<s,t>}.\lambda e.P(e) \wedge e \in D_s \setminus D_s^{\text{SG}}.$

c. $[\![$**hazime- 'begin'**$]\!] = \lambda P.\lambda e.\lambda w.\forall w' \in \text{BEST}(w).\exists e'.$

$\text{INT}(e) \subseteq \left[\text{MIN}(P(e')), \frac{\text{MAX}(P(e'))+(n-1)\times\text{MIN}(P(e'))}{n}\right]$ in w'.



• **Future directions:** The defective T analysis has thus provided reasonable accounts both to the syntax and the semantics; (i) T_{def} yields a control construction (syntax; FACT 2-4), and (ii) it also provides the I-genericity (semantics; FACT 1). The semantics in (7) also contributes to a body of literature with similar conclusions that aspect is a source of modal meaning (imperfective paradox; Dowty 1977; Landman 1992; Portner 1998).

Fukuda, S. 2012. Aspectual verbs as functional heads. **Ferreira, M.** 2016. The semantic ingredients of imperfectivity in progressives, habituals, and counterfactuals. **Landman, F.** 1992. The progressive. **Landman, F.** 1996. Plurality. **Kageyama, T.** 2016. Verb-compounding and verb-incorporation. **Kishimoto, H.** 2014. The layered structure of syntactic V-V compuonds in Japanese. **Krifka, M.** 1992. Thematic relations as links between nominal reference and temporal constitution. **Kratzer, A.** 2007. On the plurality of verbs. **Kuno, S.** 1983. *Shin nihon bunpo kenkyu*. **Matsumoto, Y.** 1996. *Complex predicates in Japanese.* **Portner, P.** 1998. The progressive in modal semantics. **Shibatani, M.** 1973. Where morphology and syntax clash: a case in Japanese aspectual verbs. **Yamada, A.** 2018. Phase-based prosody: evidence from pitch-accent distribution in the Japanese verbal domain. NELS 49. **Yumoto, Y.** 2005. Hukugodosi hasei dosi no imi to togo.