VERBAL MORPHOLOGY OF JAPANESE

Michiya Kawai
Huron University College/University of Western Ontario

1. Introduction

In this paper I argue that the regularity underlying Japanese verbal morphology is properly captured by the analysis of verbal morphology proposed by Lasnik (1995).

The intricacy of the Japanese verbal morphology has long been recognized (cf. Bloch 1946, McCawley 1968, Shibatani 1990, among others). In this paper, we focus on so-called dictionary (or plain) form, as in (1), putting aside the polite form.¹

(1) Present Positive Past Positive Present Negative Past Negative
a. taberu tabeta tabenai tabenakatta 
ed eat
b. da datta de(wa)nai de(wa)nakatta be
c. takai takakatta takakunai takakunakatta high
d. aru atta nai nakatta exist

Lasnik (1995) presents an analysis on the English verbal morphology, a minimalist analysis inspired by the analysis in Syntactic Structures (Chomsky 1957). Specifically, Lasnik’s proposes two kinds of Infl’s: featural I (I₀) and affixal I (Iₐ).

(2) a. I₀ agrees with, and subsequently attracts, a lexicalist predicate, i.e., an element with its features fully specified, such as English tensed auxiliaries.
   b. Iₐ PF-merges with a nonlexicalist element, a bare verb stem.
   c. A stranded Iₐ can be salvaged by Do-Support, a PF-rule.

This paper demonstrates that with the assumptions in (3), the Japanese paradigm in (1) can be accounted for properly.

¹ I would like to thank Howard Lasnik, David Michaels, Mineharu JJ Nakayama, Wako Tawa, Timothy J. Vance, Diane Massam, and Yves Roberge for their support. I am also indebted to the editors of this volume, colleagues at Huron University College and University of Western Ontario, and the organizers and participants of the 2006 CLA meeting at York University – in particular, Kenji Oda, Tohru Uchiumi, and Hitay Yukseker – for questions and helpful comments. I owe Jami Bujnowski Kawai for editorial help. As always, remaining mistakes are my own.

¹ The polite form is amalgamation of various dialects, formed (or “formalized”) around the late 19th and early 20th century (Sugimoto 1988). It is quite prescriptive, and, arguably, not fully consistent within itself. Prior to examining it, an extensive corpus-based study is necessary, in order to grasp the actual use of this form.
(3)  a. Japanese main verbs are nonlexicalist and predicate verb ar is lexicalist,
    b. Japanese has an operation of Ar-Support parallel to the English Do-Support, and
    c. Japanese has both I$_F$ and I$_A$.

This paper is organized as follows: A brief outline of the Japanese verbal morphology in Section 2; a brief outline of Lasnik’s (1995) analysis of the English verbal morphology in Section 3; and Lasnik’s analysis applied to the Japanese verbal morphology in Section 4. Brief concluding comments are given in the final section.

2. Japanese Verbal Morphology and Distribution of -aru

In the subsequent discussion we assume that $^+$ru and $^+$ta are temporal suffixes. This assumption is not accurate, since $^+$ru and $^+$ta assume other roles (e.g., aspectual markers (Sakuma 1936, Shibatani 1990, Kawai 1987, 1998, for example)), as well. However, the assumption should suffice for the present purposes. Consider the baseline data again.

<table>
<thead>
<tr>
<th>(1)</th>
<th>Present Positive</th>
<th>Past Positive</th>
<th>Present Negative</th>
<th>Past Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>taberu</td>
<td>tabeta</td>
<td>tabenai</td>
<td>tabenakatta</td>
</tr>
<tr>
<td>b.</td>
<td>da</td>
<td>datta</td>
<td>de(wa)nai</td>
<td>de(wa)nakatta</td>
</tr>
<tr>
<td>c.</td>
<td>takai</td>
<td>takakatta</td>
<td>takakunai</td>
<td>takakunakatta</td>
</tr>
<tr>
<td>d.</td>
<td>aru</td>
<td>atta</td>
<td>nai</td>
<td>nakatta</td>
</tr>
</tbody>
</table>

Observe the presence of atta (marked in bold) in the copular and adjectival inflection, as well as in all of the past negative forms. The surface paradigm (1) arguably involves suppletion, assimilation, and other morphophonemic alterations (McCawley 1968, Kawai 1985, for example). For example, da is likely a suppleted form of de$^+$aru, a formal counterpart (cf. Sugimoto 1988). Likewise, the present positive form of adjectives is likely a suppleted form of takak$^+$aru, which existed as early as 8th century C.E.$^2$ Extending this to the negative forms, which appear parallel to adjectives, we analyze the present negative form as nak$^+$aru.$^3$ Finally, (wa) in (1b) is a contrast-marker, not essential to the verb morphology; it can be omitted; thus, we will exclude it from our discussion. Bearing these in mind, we arrive at the paradigm in (4) (Kawai 1985, 2003, 2006).$^4$

---

$^2$ One finds, in the 793th poem of Man-yoo-shuu (759 C.E.), kanashikarikeri (I am sad), which can be analyzed as kanashik + ari+ keri. This form became very common in Middle Japanese (Okimori 1989).

$^3$ The Japanese negation behaves as if they are adjectival not only morphologically, but also semantically (cf. Jacobson 1995).

$^4$ It is difficult to decide whether or not ar+ru is deleted in the present positive forms in (1b)/(1c); it could be absent, rather than having been deleted. Given that the present indicative interpretation is distinct from that of a tenseless small clause (Ohta 1997, Kawai 2006), and that the past predicates have overt “past” tense ar+ta, we assume the
The distribution of aru, highlighted with bold letters in (4), will be discussed in Section 4.


(5) French verbs raise past the negation to I in a tensed negative clause, as in (7a/b), and to [Spec, C] in an interrogative clause, as in (7c/d).
(6) Unlike English auxiliary verbs (8), English verbs do not move past the negation to I in a tensed negative clause (9a/b), nor to [Spec, C] in an interrogative clause (9c/d).

(7) a. Jean (n’)aime pas Marie.
   b. Jean [i (n’)aime [Neg pas] [VP t1 Marie ] ]
   c. Aime-jean Marie?
   d. [CP Aime [IP Jean [VP t1 Marie ] ] ] ?
(8) a. John is not a student.
   b. John [i isi [Neg pas] [VP t1 a student ] ]
   c. Is John here?
   d. [CP Is [IP John [VP t1 Marie ] ] ] ?
(9) a. * John likes not Mary.
   b. John [i likesi [Neg not] [VP t1 Mary ] ]
   c. * Likes John Mary?
   d. [CP Likes [IP John [VP t1 Mary ] ] ] ?

Chomsky (1991) presents a minimalist analysis of the English verbal morphology, as in (10), inspired by Pollock’s (1989) analysis of the French counterpart.

(10) a. S is the maximal projection of the inflectional morpheme Infl (=C of Chomsky 1957).
   b. Infl takes VP as its complement.
   c. When the head of VP is have or be it raises to Infl, the next head up.
   d. Otherwise Infl lowers to V: Affix Hopping.
   e. Otherwise, do adjoins to Infl. (Adapted from Lasnik 1995 (1999: 98))

underlying presence of ar in ar+ru. The ultimate correctness of this position is still open.
Chomsky (1993) eliminates lowering operation from syntax, rendering (10d) impossible. Instead, a strict lexicalist position is advocated in his 1993 analysis: i.e., English verbs are fully inflected – i.e., with the relevant features present – in the Lexicon. Thus, there is no merger between V and Infl. The difference between French and English is captured, instead, as follows:

(11) a. V-features of Agr in French are strong.
    b. V-features of Agr in English are weak.
(12) a. V-features are not legitimate PF objects.
    b. Strong features are visible at PF, but weak features are not.

Procrastinate
Delay an operation until LF whenever possible, that is, whenever delaying would not cause the derivation to crash.

(13) English-specific stipulation:
Have and be are semantically vacuous, hence not visible to LF operations.

With (11a), French must eliminate the V-features of Agr before Spell-Out, by raising V into Agr. Thus, French verbs must overtly raise to I. In English, on the other hand, V-features need not be eliminated before Spell-Out, given (11b). Further, Procrastinate (13) demands V-features be eliminated after Spell-Out in English; hence, no overt V-raising. (14) is designed to derive the contrast in (15).

(15) a. * John likes not Mary.
    b. John is not happy.
    c. John has not been happy.

Overt raising of the auxiliaries in (15b/c) appear to violate Procrastinate. Yet, it does not prevent be/have from overtly raised, since it compares only among the convergent derivations. With (14), LF operation fails to see the presence of be/have; if they are not overtly raised, then they cannot be raised at all, and their unchecked features will cause the representations to crash at LF. Thus, overt raising of have/be is the only option available, according to Chomsky (1993). This account suffers from many problems. See Lasnik 1995 for problems with Chomsky’s (1993) analysis of English verbal morphology. In order to recognize the problems in the analysis, it suffices us to look at (16), as pointed out by Lasnik (1995 (1999: 104)).

(16) a. * John not likes Mary.
    b. * John likes not Mary.

Notice that (16) satisfies Procrastinate; the V-features does not offend the PF representation. (16a) “must be ruled out, but its derivation must not crash. If it crashed, it could not block [the ungrammatical (16b)], since Procrastinate only chooses among convergent derivations” (Lasnik 1995 (1999: 104)).

As an alternative, Lasnik (1995) proposes (17).
(17) a. There are two types of Infls: I_F and I_A.
    b. The features in I_F agree with the corresponding features in V, and I_F attracts V subsequently.
    c. A verb can be lexicalist, with its features fully present, or nonlexicalist – a bare stem.
    d. Do-Support, a last resort PF-operation, licenses a stranded I_A.

Lexicalist verbs need I_F because their features must be licensed via agreement with the corresponding features in I_F. Nonlexicalist verbs, on the other hand, need I_A because they lack the relevant features for licensing I_F. For the English verbal paradigm, the additional three assumptions in (18) suffice.

(18) a. English main verbs are nonlexicalist.
    b. English auxiliary verbs are lexicalist.
    c. French verbs (i.e., both main and auxiliary verbs) are lexicalist.

(19) a. John [I_A ] [VP like Mary]. //PF-Merge(I_A, like) \( \rightarrow \) likes
    b. John [I_A ] not [VP like Mary]. //PF-Merge(I_A, like)
    c. John [I_A-do] not [VP like Mary]. //Do-Support \( \rightarrow \) does
    d. John [I-isI] [VP t_1 happy]. //is-Raising
    e. John [I-isI] not [VP t_1 happy]. //is-Raising

For example, likes, being nonlexicalist, PF-merges with I_A (19a); it is incompatible with I_F, not being able to agree with it. With a negation, it cannot PF-merge with I_A (19b); thus, Do-Support licenses the stranded I_A (19c). Be-verbs, on the other hand, are lexicalist, and, thus, raises to I_F after agreement (19d). Crucially, in this analysis, V-to-I raising is not blocked by the intervening negation (19e). We assume that this kind of V-to-I raising falls outside of the Head Movement Constraint (Travis 1984). See Lasnik 1995 for several possibilities as to how this may be so. Here, for concreteness, we follow Lasnik’s suggestion that movement obeys a kind of Relativized Minimality proposed by Roberts (1993, 1994), which differentiates A-heads from Ā-heads. Assuming that V and I are A-heads and negation, an Ā-head, negation does not block V-to-I (Lasnik 1995 (1999: 108)). Let us assume the correctness of Lasnik’s analysis of the English verbal morphology for the remaining discussion. See also Matushansky 2006 for a related discussion.

4. **Japanese Verbal paradigms**

The baseline paradigm of Japanese verbal morphology is repeated here again:

<table>
<thead>
<tr>
<th>(4) Present Positive</th>
<th>Past Positive</th>
<th>Present Negative</th>
<th>Past Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tabe+ru</td>
<td>tabe+ta</td>
<td>tabe+nak+ar+ru</td>
<td>tabe+nak+ar+ta</td>
</tr>
<tr>
<td>b. de+ar+ru</td>
<td>de+ar+ta</td>
<td>de+nak+ar+ru</td>
<td>de+nak+ar+ta</td>
</tr>
<tr>
<td>c. takak+ar+ru</td>
<td>takak+ar+ta</td>
<td>takaku+nak+ar+ru</td>
<td>takaku+nak+ar+ta</td>
</tr>
<tr>
<td>d. ar+ru</td>
<td>ar+ta</td>
<td>nak+ar+ru</td>
<td>nak+ar+ta</td>
</tr>
</tbody>
</table>

Let us first consider the verbal paradigm in (4a). Observe in (20a/b) that tabe+ does not raise past the negation to I (20a/b). In other words, Japanese Vs do not
agree with – and, subsequently are not attracted by – a $I_A$, as schematically shown in (20c).

(20) a. $^*$ nak $\text{taberu}_{\{F_1,\ldots,F_n\}}^+ [I_{\{F_1,\ldots,F_n\}}]$ neg eat

b. $^*$ nak $\text{tabeta}_{\{F_1,\ldots,F_n\}}^+ [I_{\{F_1,\ldots,F_n\}}]$ neg ate

c. $^*$ $[\text{VP} \ldots t_1 \ldots] [\text{Neg} \text{nak}] V_1^+ [I \text{ru/ta}]$

Thus, according to Lasnik's (1995) analysis,

(21) a. Japanese verbs are nonlexicalist, and
b. The $I$ in (4a) is an $I_A$.

In the present and past positive forms in (4a), the $I_A$ is linearly adjacent to the verb root $\text{tabe}$. As the theory predicts, the $I_A$ is licensed by the PF-merger with $\text{tabe}$ (22a):

(22) a. $[\text{VP} \ldots \text{tabe}^+ \ldots] [I \text{ru/ta}_A]$ /*PF-merger($\text{tabe}, I_A$)

b. $^* [\text{VP} \ldots \text{tabe}^+ \ldots] [\text{Neg} \text{nak}] [I \text{ru/ta}_A]$ /*PF-merger($\text{tabe}, I_A$)

c. $[\text{VP} \ldots \text{tabe}^+ \ldots] [\text{Neg} \text{nak}] [I ar + \text{ru/ta}_A]$ //Ar-Support

In (22b), PF-merger of $\text{tabe}$ and the $I_A$ is impossible, due to the intervening negation; without any further action, the $I_A$, being stranded, would be an illegal LF object. This is precisely where $ar$ occurs in the verbal paradigm.

(23) $Ar$-Support inserts $ar$, a supporting element for the stranded $I_A$.

For the ease of presentation, we will henceforth use $AR$ for those inserted by $Ar$-Support, being distinct from the non-supportive $ar$.

The present analysis explains the verb paradigm correctly with one exception: the negative forms of $aru$.

(24) Present Positive  Past Positive  Present Negative  Past Negative
a. $ar+ru$, $ar+ta$, $\text{nak}+\text{AR}+ru$, $\text{nak}+\text{AR}+ta$
b. $ar+ru$, $ar+ta$, $^*\text{ar}+\text{nak}+\text{AR}+ru$, $^*\text{ar}+\text{nak}+\text{AR}+ta$

While the positive forms confirm the predicted forms, the negative forms do not. Instead of the correct (24a), the analysis predicts the incorrect (24b); $Ar$-Support is expected by the theory to apply in the negative paradigm, as indicated with capital letters. Why is there such a gap in the paradigm? I believe that we answer this question by examining the adjectival and copula paradigm, as the relevant part of the base-line paradigm in (4b/c).

(4) Present Positive  Past Positive  Present Negative  Past Negative
b. $\text{de}+ar+ru$ $\text{de}+ar+ta$ $\text{de(wa)}+\text{nak}+\text{AR}+ru$ $\text{de(wa)}+\text{nak}+\text{AR}+ta$
c. $\text{takak}+ar+ru$ $\text{takak}+ar+ta$ $\text{takak}+\text{nak}+\text{AR}+ru$ $\text{takak}+\text{nak}+\text{AR}+ta$
It is reasonable to suppose that ar in de+ar and takak+ar is an “auxiliary” verb, as in (25a/b) parallel to the English be (25c).

(25) a. \[ VP \text{[PredP } \ldots \text{ de } \text{ar}\text{]} \text{[i ru]} \]
    b. \[ VP \text{[AP } \ldots \text{ takak}\text{ar}\text{]} \text{[i ru]} \]
    c. \[ [i I] VP \text{is} \text{[AP/PredP]} \]

The positive paradigm is compatible with both lexicalist and nonlexicalist ar, because both Ar-raising to I and PF-merger(ar, I) yield the same surface string. However, the negative paradigm shows the non-involvement of I_A. With I_A Ar-Support would apply to the stranded I_A; thus, for the negative forms of (4b) and (4c), we incorrectly predict da-nai (from de-ar-a-nai) and da-nakatta (from de-ar-a-nakatta), and takakaranai (from takak+ar+a+nai) and takakaranakatta (from takak+a+ar+a+nakatta), respectively.

(26) a. * \[ [ VP \ldots \text{de+ar} \ldots ] \text{[Neg hak]} \text{[i AR+ru/ta\text{is}]} \] //**Ar-Support
    b. * \[ [ VP \ldots \text{takak+ar} \ldots ] \text{[Neg hak]} \text{[i AR+ ru/ta\text{is}]} \] //**Ar-Support

Therefore, I_F must be involved in the adjective/copula paradigms, and ar must be lexicalist; that is, it is merged fully inflected. Reasonably assuming that aru/atta is initially merged inside VP, it must be raised to I_F, as shown in (27).

(27) a. \[ [ VP \ldots \text{de+ t\text{is}} \ldots ] \text{[aru/atta]} \]
    b. \[ [ VP \ldots \text{takak+ t\text{is}} \ldots ] \text{[aru/atta} + \text{I [-past/+past]} \]
    c. \[ [ VP \ldots \text{de+ t\text{is}} \ldots ] \text{[Neg hak]} \text{[i aru/atta} + \text{I [-past/+past]} \]
    d. \[ [ VP \ldots \text{takak+ t\text{is}} \ldots ] \text{[Neg hak]} \text{[i aru/atta} + \text{I [-past/+past]} \]
    e. \[ [IP \text{[I+is}\text{nt}]} \text{[Neg not]} \text{[VP t\text{is} \ldots]} \]

In the negative paradigm, aru/atta is raised past Neg, parallel to English is in (27e). This also accounts for the apparent irregularity with ar that we saw earlier.

Recall the question raised earlier: that the negative forms of aru/atta are nai and nakatta, respectively. Seemingly, the verb root ar disappears in the negative forms. However, the present analysis accounts for this apparent irregularity. Being a lexicalist verb, aru/atta agrees with I_F, and subsequently raises to I_F, vacating its merged position, as in (26b). In other words, the disappearance of aru is only apparent; it only vacates its VP-internal position, raising to I_F.

---

5 Japanese has nonlexicalist ar, as well. It surfaces with the non-standard – and somewhat archaic – negations: –zu, as in arazu, as pointed out by Tohru Uchiumi (personal communication). Likewise, nonlexicalist ar is used with –m(u) or –hen, forms of negation mainly used in western dialects of Japanese: aran and arahen. I believe that those negations are proclitics. That is, they cliticize onto a string-adjacent item with phonetic content. Those proclitics are, naturally, incompatible with the lexicalist ar, since they cannot cliticize onto it once it raises past the negation. The unacceptability of *aranai and *aranakatta in standard Japanese must be due to blocking effects: the inflection with lexicalist ar blocks the use of nonlexicalist ar in this environment. If this is correct, Japanese has three kinds of ar’s: lexicalist-, nonlexicalist-, and supportive-ar.
To sum up, the paradigm in (4) is accounted for by Lasnik’s (1995) analysis of verb morphology with (28).

(28) a. Japanese has both I_F and I_A.
    b. Japanese main verbs are nonlexicalist.
    c. Japanese predicate ar is lexicalist.
    d. Japanese has Ar-Support, a last resort PF-operation.

5 Implications

Supposing the correctness of the present analysis, two questions immediately arise:

(29) a. Why are Japanese main verbs nonlexicalist?
    b. Why is the predicate ar lexicalist?

While I do not have a well-worked out account for each question, the following can be speculated. I suspect that this has something to do with the agglutinative nature of Japanese morphology. Being an agglutinative language, it probably makes more sense to inflect regular lexical materials in the morphological component, rather than storing all the forms fully infected. Storing all the inflected forms would make more sense for a “small number” of irregular predicates. This may also explain why the predicate ar is lexicalist. Cross-linguistically, copula verbs are more likely irregular. Related, why are English main verbs lexicalist, whereas those in French are nonlexicalist? Clearly, the (non-)agglutinative nature plays little in this distinction.

Secondly, the present analysis is incompatible with a claim made by Ishihara (2001) and Ishii (2001) that Japanese verbs obligatorily raise to I. Ishihara’s (2001) argument primarily uses phonological evidence for this claim. Tentatively, I suggest that even though Japanese main verbs do not raise themselves to I, the PF merger of a predicate and an I_A gives rise to the PF effects associated with predicate raising.

The result of this study independently supports Lasnik’s (1995) proposal of verbal morphology, which was not concerned with the range of the data discussed here. A natural extension of the present research may be to examine the Japanese polite-form paradigm, as well as the verbal morphology of the typologically similar languages: for example, Okinawan, Ainu, Korean, Mongolian, Turkish, among others. See Han, Lidz, and Musolino, To appear, on the Korean verbal morphology. They argue that Korean has (at least) two versions of verbal morphology: one with verb raising and the other without. In principle, Lasnik’s (1995) analysis is compatible with such a language since two kinds of I’s are available. Naturally, a detailed analysis is in order before we can make any claim about the Korean verbal morphology. I will leave this and other questions for future research.
References


Kawai, Michiya. 1985. Syllable structure analysis of assimilation phenomena: A case study from Japanese verbal morphology (ms.). The University of Connecticut, Storrs, CT.

Kawai, Michiya. 1987. Barriers to government simplified (ms.). The University of Connecticut, Storrs, CT.


Sakuma, Kanae. 1936. *Gendai-nihongo no hyōgen to gohō.* Kyoto: PHP.

