Teaching verbal morphology in a Japanese classroom
--Learning from the corpus data analysis of native speakers--

Priya Ananth
Middle Tennessee State University
pananth@mtsu.edu

Abstract
There have been several studies in the past literature on tense-aspect acquisition in Japanese that have looked into the form-meaning associations that learners of Japanese make in their choice of verbal morphology. One popular claim is that learners of Japanese associate the perfective/past ending (-ta) with Achievement (e.g. recognize, arrive) and Accomplishment (e.g. run a mile, write a report) verbs. These are verbs that are dynamic and have an inherent endpoint. This study sets out to investigate if the above claim holds true for native Japanese speech from a corpus data of spoken Japanese. This study comprehensively examines the associations of the past (perfective) and non-past (imperfective) verbal morphology with the four classes of verbs, namely Achievement, Accomplishment, Activity and Stative according to Vendler’s classification, in case of toki ‘when’ clauses in Japanese. The study analyzes corpus data collected from spoken discourses such as free conversations, lectures and interview sessions of native Japanese people. The results show that native speakers make form-meaning associations of the past/perfective and non-past/imperfective endings with the semantic characteristics of the verbs. This study has important pedagogical implications in the systematic teaching of verbal morphology in the tense-aspect system of Japanese as a foreign language.

1. Introduction
An extensively investigated issue of tense and aspect development in Japanese second language acquisition studies is the acquisition of tense-aspect morphology. Several researchers have claimed that the development of tense-aspect morphology in first language (L1) acquisition of various languages is strongly influenced by the inherent semantic characteristics of the verbs to which the inflections are attached (Shirai, 1993). The same tendency has been observed in second language (L2) acquisition, especially in English, Spanish and French (Bardovi-Harlig, 1995; Andersen, 1991). A significant number of studies done in Japanese have examined the effect of semantics of the verbs on the verbal endings. The bulk of these studies have looked at the association of the past/ perfective ending ‘ta’ and progressive/ imperfective -tei forms and verb semantics (Shirai, 1995; Sugaya & Shirai, 2007; Sheu, 1997; Nishikawa, 1998). The present
study focuses on the form meaning associations of the past/perfective ‘ta’ and non-past/imperfective ‘ru’ endings with the verb semantics. In order to validate the results obtained from L2 acquisition studies, it is essential to extend the study to native speakers’ data so that generalizations can be drawn about using Japanese universally and not just by foreign or second language learners of Japanese. Hence the need for the present study that specifically aims at examining the association of the perfective and imperfective verbal morphology with the semantic meanings of the verbs that appear in toki ‘when’ clauses as observed in the native speakers’ corpus data of spoken discourses. The study will utilize a corpus linguistic methodology which is defined as “a linguistic methodology that is founded in the use of electronic collection of naturally occurring texts” (Granger, 2002). Starting from the 1980s, the use of corpora in L2 acquisition studies, coupled with the development of software tools and analytical techniques of using computer corpora has increased considerably. Corpus-based approaches have enabled language researchers to describe and analyze the patterns of the target language such as co-occurrence or association patterns and the contexts in which a specific word or phrase is commonly found based on the empirical evidence of large sized real world texts.

The association of the perfective ‘ta’ and imperfective ‘ru’ verbal endings with the semantic characteristics of verbs such as telicity, punctuality and dynamicity is readily observable in toki ‘when’ clauses because toki structures allow the use of both the verbal endings to be used in the subordinate position resulting in grammatical structures. Therefore it becomes easier to observe if the users’ of the language have a preference for a particular verbal ending for a specific class of verbs. For example, taberu toki and tabeta toki can both be used to interpret ‘at the time eating’ and both are grammatical. In the current study by investigating what particular ending (ru or ta) is used more frequently with a verb of a particular semantic category (based on Vendler’s categorization), results will be interpreted quantitatively and their impact on Japanese pedagogy will be discussed.

The organization of this paper is as follows. Section 2 presents previous literature on the tense and aspect studies of Japanese. Section 3 describes the present study including data set, method, analysis and results. Section 4 will present the conclusions and future study including some pedagogical implications of this study.
2. Previous Literature

2.1 Inherent Lexical Aspect

The notion of temporality in language is subdivided into the grammatical categories of tense and aspect. Both tense and aspect are concerned with time, but in ways quite different from each other. While tense looks at the “relative sequential ordering between two time points”, aspect is defined as “different ways of viewing the internal temporal constituency of a situation” (Comrie, 1976). Aspect is typically distinguished into two types: grammatical and inherent lexical aspect. Grammatical aspect, also known as “viewpoint aspect” (Smith, 1991), refers to aspektual distinctions marked explicitly by linguistic devices, usually auxiliaries or inflections, as in (1) and (2).

(1) I slept. (simple past)
(2) I was sleeping. (past progressive)

Inherent lexical aspect (Andersen, 1986), or also known as “situation aspect” (Smith, 1991) or semantic aspect (Comrie, 1976), refers to what is inherent in the lexical items that describe the situation. In (1) and (2) *sleep* has intrinsic duration whether in simple past or past progressive. These inherent distinctions are noted by Vendler’s (1967) framework of lexical aspects, which consists of States, Activities, Accomplishments and Achievements. These four lexical classes can be distinguished by three features: Dynamicity, Telicity, and Punctuality (Andersen, 1991; Comrie, 1976). “Dynamicity” distinguishes dynamic events from states. States (e.g. sleep) exhibit little or no change over time, whereas a dynamic event (e.g. eat) describes a process that changes. “Telicity” refers to whether an action has a naturally defined endpoint such that the action is completed and cannot continue once that endpoint is reached. Telic predicates exhibit processes with an inherent limit (e.g. run a mile), whereas atelic predicates reveal processes without limit (e.g. run in the park). “Punctuality” distinguishes durative from punctual predicates. Durative events take place over a certain period of time (e.g. make a chair), whereas punctual events take place instantly (e.g. recognize).

2.2 Vendler’s Classification

The four-way distinction, based on temporal properties of verbs (or verb phrases) that was introduced by Vendler (1967), is described below and summarized in Table 1.
Achievement or punctual predicates are telic and punctual, and capture the beginning or the end of an action (Lee, 1999). These take place instantaneously, and can be reduced to a single point in time (Shirai, 1995). Examples: arrive, leave, notice, recognize.

Accomplishment or telic predicates are telic and non-punctual, and thus have some duration, but with a single clear inherent endpoint. In English, these are often phrasal. Examples: build a house, make a chair, write a letter, paint a painting.

Activity are atelic and thus do not assume a natural or inherent endpoint. They have inherent duration in that they involve a span of time. Examples: rain, write, read, swim, play.

State are non-dynamic and persist over time without change. Examples: have, believe, think, know, love.

<table>
<thead>
<tr>
<th></th>
<th>Dynamicity</th>
<th>Telicity</th>
<th>Punctuality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Accomplishment</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Activity</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>State</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 1: Vendler’s Classification

2.3 Aspect Hypothesis

A universal tendency observed in the acquisition of tense-aspect morphology, summarized in what is known as the Aspect Hypothesis (Andersen & Shirai, 1994; Bardovi-Harlig, 2000) claims that learners create restricted form-meaning relationships at the early stages of acquiring L1 tense-aspect morphology. Specifically, learners strongly associate past tense and perfective aspect forms with punctual verbs (verbs that express an event that occurs in a moment, e.g. recognize) and telic verbs (verbs that have natural endpoints, e.g. wake up), imperfective aspect forms with atelic verbs (verbs that do not have an inherent endpoint, e.g. swim), and progressive aspect forms with activity verbs (e.g. write). In the current study, it will be tested quantitatively if native speakers make any associations between the past/ perfective ending ‘ta’ and telic verbs and non-past/ imperfective ending ‘ru’ and atelic verbs in toki ‘when’ clauses.
2.4 Toki ‘when’ clauses

In Japanese, *toki* denotes the time at which. In terms of -ru/-ta combination at S1 (subordinate clause) and S2 (main clause) there are four patterns schematized below and illustrated by (3) – (7) from Lee, 1999.

Sentence Patterns
S1 *ru toki* S2 *ru*
S1 *ru toki* S2 *ta*
S1 *ta toki* S2 *ru*
S1 *ta toki* S2 *ta*

(3) Nihon-e iku toki kamera-o kau
Japan-Loc go when camera-Acc buy
‘I’ll buy a camera when (before) I go to Japan.’

(4) Nihon-e itta toki kamera-o kau
Japan-Loc went when camera-Acc buy
‘I’ll buy a camera when (after) I go to Japan.’

(5) Nihon-e itta toki kamera-o katta
Japan-Loc went when camera-Acc bought
‘I’ll buy a camera when (after) I go to Japan.’

(6) Nihon-e iku toki kamera-o katta.
Japan-Loc go when camera-Acc bought
‘I bought a camera when (before) I went to Japan.’

(7) Nihon-e iku/itta toki, kamera-o motte itta.
Japan-Loc go/went when camera-Acc carried
‘I carried a camera, when I went to Japan.’

The four patterns given above present three possible temporal interpretations: the event in S1 is prior to event in S2 as in (4) and (5), or S2 is prior to S1 as in (3) and (6) or both are contemporaneous as in (7). As can be seen in these examples, verbs in *toki* clauses allow both the *ru/ ta* endings leading to perfectly grammatical interpretations. Hence it is useful to see if the users’ selection of the *ru/ ta* endings is governed by the type of verb present in the *toki* clauses.

3. Present Study

The main objective of the current study is to examine the association of the perfective and imperfective verbal endings with the semantic characteristics of the verbs that contain them, namely telicity and punctuality that appear in *toki* ‘when’ clauses in the native speakers’ corpus data of spoken discourses.
3.1 Data Set

Data was collected from recorded spoken discourses such as free conversations, lectures and interview sessions of native Japanese speakers. The following sources were used: KWIC Corpus, Uemura Corpus, Ono Corpus, Monbukagakusho Corpus, Kokken Corpus.

3.2 Method and Analysis

After selecting the target sentences with *toki* clauses from the corpus data, they were classified into four sentence patterns (S1: subordinate clause, S2: main clause)

(a) S1 *ru toki* S2 *ru*
(b) S1 *ru toki* S2 *ta*
(c) S1 *ta toki* S2 *ta*
(d) S1 *ta toki* S2 *ru*

Following are examples of target sentences for each of the four sentence patterns.

S1 *ru toki* S2 *ta*
8. Ie o deru toki denwa ga kakatte kita.
   ‘The phone rang before he stepped out.’

S1 *ta toki* S2 *ta*
9. Gohan o tabeta toki ohasi o zyoozu ni tukatta.
   ‘He used chopsticks skilfully when he ate food.’

S1 *ru toki* S2 *ru*
10. Basu tee de matu toki itumo CD o kiku.
    ‘He always listened to CDs when he waited at the bus stop.’

S1 *ta toki* S2 *ru*
11. Yooso ga atta toki itte kudasai.
    ‘Please let me know if there is an issue.’

After listing the target sentences under each sentence pattern, verbs at S1 position were classified into the four Vendler’s categories. Shirai (1993)’s tests were used to classify the verbs. Percentage occurrence for the Vendler’s category of verbs in S1 position in each of the four patterns was then calculated. Both the raw figures and percentages are shown in Tables 2-5 in the results section below.

3.3 Predictions

Based on the claim of the Aspect Hypothesis, that learners strongly associate past tense/perfective aspect forms with punctual verbs (verbs that express an event that occurs in a moment,
e.g. recognize) and telic verbs (verbs that have natural endpoints, e.g. wake up), and non-past tense/ imperfective aspect forms with atelic verbs (verbs that do not have an inherent endpoint, e.g. swim), two predictions can be made for this study:

**Prediction 1**: Native speakers will associate past tense/ perfective aspect form ‘ta’ with telic verbs.

**Prediction 2**: Native speakers will associate non-past tense/ imperfective aspect form ‘ru’ with atelic verbs.

### 3.4 Result

In this section, the results of the data analysis will be tabulated. Telic verbs refer to Achievements and Accomplishments (with endpoints) and atelic verbs refer to the Activities and States (that lack endpoints).

Table 2 below shows the overall occurrence of verbs at S1 position in native speakers’ data. Table 3 below summarizes the overall percentage occurrence of telic and atelic verbs for the same data set.

<table>
<thead>
<tr>
<th>Verbal Category (Token Count)</th>
<th>Overall Occurrence (521)</th>
<th>Percentage Occurrence (100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement</td>
<td>220</td>
<td>42</td>
</tr>
<tr>
<td>Accomplishment</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>Activity</td>
<td>137</td>
<td>26</td>
</tr>
<tr>
<td>Stative</td>
<td>46</td>
<td>9</td>
</tr>
<tr>
<td>Stative te iru</td>
<td>68</td>
<td>13</td>
</tr>
<tr>
<td>Stative Combined*1</td>
<td>114</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 2: Overall and Percentage Occurrence of verbs at S1 position

<table>
<thead>
<tr>
<th></th>
<th>Highest</th>
<th>Telic</th>
<th>Atelic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>Achievement 42%</td>
<td>52%</td>
<td>48%</td>
</tr>
</tbody>
</table>

Table 3: Summary of Overall Percentages

---

1 Stative Combined is the sum total of regular Statives, such as aru ‘to be’ and derived Statives that have the -te iru ending, such as site iru ‘doing’.
Overall there was an even distribution in the occurrence of telic and atelic verbs in the S1 position. Out of 521 tokens, telic verbs (Achievement and Accomplishment) occupied 52% of the total, while atelic verbs (Activity and Stative Combined) occupied 48%. When seen individually, Achievements were the highest with 42%, Activities with 26%, Statives with 22% and Accomplishment with 10%.

Table 4 below shows the occurrence of verbs grouped by sentence patterns at S1 position. Table 5 below summarizes the percentage occurrence of verbs grouped by sentence patterns.

<table>
<thead>
<tr>
<th>Verbal Category</th>
<th>S1 ru toki S2 ru</th>
<th>S1 ru toki S2 ta</th>
<th>S1 ta toki S2 ta</th>
<th>S1 ta toki S2 ru</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Token / %)</td>
<td>(219) (100)</td>
<td>(40) (100)</td>
<td>(118) (100)</td>
</tr>
<tr>
<td>Achievement</td>
<td>71 32 12 30</td>
<td>49 42 88 61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accomplishment</td>
<td>11 5 4 10</td>
<td>30 25 5 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity</td>
<td>74 34 7 18</td>
<td>22 19 34 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stative</td>
<td>20 9 3 8</td>
<td>9 8 14 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stative te iru</td>
<td>43 20 14 35</td>
<td>8 7 3 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stative Combined</td>
<td>63 29 17 42</td>
<td>17 14 17 12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Raw and Percentage Occurrence of verbs grouped by sentence patterns

<table>
<thead>
<tr>
<th>Verbal Category</th>
<th>Highest</th>
<th>Telic</th>
<th>Atelic</th>
</tr>
</thead>
<tbody>
<tr>
<td>ru-ru</td>
<td>Activity</td>
<td>34%</td>
<td>37%</td>
</tr>
<tr>
<td>ru-ta</td>
<td>Stative</td>
<td>42%</td>
<td>40%</td>
</tr>
<tr>
<td>ta-ta</td>
<td>Achievement</td>
<td>42%</td>
<td>67%</td>
</tr>
<tr>
<td>ta-ru</td>
<td>Achievement</td>
<td>61%</td>
<td>64%</td>
</tr>
</tbody>
</table>

Table 5: Summary of Percentage Occurrence of verbs grouped by sentence patterns

A finer classification of the verbs based on sentence patterns showed a predicted distribution between atelic and telic verbs. In the S1 ru toki S2 ru and S1 ru toki S2 ta sentence patterns, it was predicted that atelic would be higher than the telic verbs (because there will be preference for associating the imperfective ‘ru’ ending with S1 verbs, that do not have end points). As can be observed from the summary in Table 5, in the S1 ru toki S2 ru pattern, the
atelic verbs (63%) are higher than telic verbs (37%), with the highest appearance is of Activity verbs (34%). Similarly in the S1 ru toki S2 ta pattern, the atelic verbs (60%) are higher than the telic verbs (40%) with the highest appearance is of Statives (42%).

In the S1 ta toki S2 ta and S1 ta toki S2 ru sentence patterns, it was predicted that telic would be higher than the atelic verbs because there will be a preference for associating the perfective ‘ta’ ending with telic verbs, that have endpoints. As can be observed from Table 5, in the S1 ta toki S2 ta pattern, telic verbs (67%) are higher than the atelic verbs (33%) with the highest appearance of Achievement (42%). In the S2 ta toki S2 ru pattern also the telic verbs (64%) are higher than atelic verbs (36%) with the highest appearance of Achievement verbs at 61%.

In sum, both the predictions regarding the native speakers’ choice of the ru/ ta morphemes for verbs with specific semantic characteristics have been proved correct quantitatively.

4. Conclusion and Future Study

The overall percentages show an almost even distribution of telic and atelic verbs for native speakers of Japanese. However, finer distinctions in sentence pattern types show that there is a tendency to use ‘ta’ with telic verbs and ‘ru’ with atelic verbs. From the above results it can be quantitatively claimed that native users of Japanese are sensitive to the semantic characteristics of the verbs such as telicity and punctuality and make associations with them in their selection of the ru/ ta endings in case of toki clause structures. They have a clear preference to use ‘ta’ with verbs that have natural endpoints and ‘ru’ with verbs that do not.

The results of this study have some significant pedagogical implications. Since native speakers’ choice of the ru/ta morpheme in toki clauses seems to be governed by the semantic characteristics of the verbs, it would be useful to introduce a verb classification based on the inherent semantic aspect of the verbs in the Japanese language classrooms. Currently most textbooks introduce verb classifications that are based on verb endings. Since the native users (from this study) and non-native users (from previous studies) are already associating the verb endings with their semantic characteristics, it would be easier for the instructors to introduce verbal classifications based on inherent semantic characteristics that would also assist the learners in their understanding, processing and production of verbs and verbal structures.

As a follow up of this study, a comparative analysis of the native and non-native spoken productions of toki clauses using the corpus linguistic methodology will be conducted. This is
needed to validate the status of Aspect Hypothesis as a universal phenomenon that pertains to both native as well as non-native speakers of Japanese. Most previous studies that have been conducted with non-native speakers were based on experimental questionnaires, where the sentences or discourses were pre-planned and prepared by the experimenter. The corpus data used in this study as well as in the follow up study uses spoken discourses that were naturally produced by the speakers in various settings such as conversations, lectures and such. Additionally, another follow up study to confirm the results of the present study in written productions of native and non-native speakers of Japanese can be suggested. That would further corroborate the validity of the form-meaning associations between verb endings and inherent meanings of the verbs across spoken and written discourses among all users of Japanese language.

**Corpus Sources**

KWIC Corpus (http://languagecraft.jp/kwic/)

Uemura Corpus (http://www.env.kitakyu-u.ac.jp/corpus/docs/index.html)

Ono Corpus (personal communication)

Monbukagakusho Corpus (Simulated Spoken Dialog Corpus, Grant in aid for scientific research on priority areas by Ministry of Education, Science and Culture, Research on understanding and generating dialog by integrated processing of speech, language and concept, 4 Vols, 1994, CD).

References


