

Form-Meaning Associations in Japanese --Analysis of Native and Non-Native Corpus Data--

Priya Ananth¹

1.0. Introduction

Form-oriented studies or “form-to-function studies” (Sato 1990) within the domain of tense-aspect acquisition studies are mainly concerned with the emergence of temporally relevant morphology. Several general studies in both first and second language acquisition claim that the development of tense-aspect morphology is strongly influenced by the inherent semantic aspect of the verb before inflection (Shirai 1993, Anderson 1991, Bardovi-Harlig & Reynolds 1995). These studies investigate the possibility of “differential distribution across aspectual categories” in verbal morphology (Bardovi-Harlig 1999). In other words, they seek to establish whether the choice of verbal morphology is distributed across aspectual categories of verbs in any predictable pattern. These studies address two different kinds of questions (Bardovi-Harlig 1999):

- (a) Where do various tense-aspect morphemes occur?
- (b) How is each of the lexical aspectual categories marked by learners?

In what is known as the Aspect Hypothesis, Anderson and Shirai (1996) summarize a universal tendency of L1 and L2 learners to create restricted form-meaning relationships at early stages in the acquisition of tense-aspect morphology. Much of the experimental research in Japanese devoted to this hypothesis has focused on the claims

¹ I am grateful to Masaaki Kamiya, Lauren Gates, Judith Rusciolelli, Patrick Conley, Jun Da and Richard Morris for reading earlier versions of the article and giving their valuable feedback. I would like to express my gratitude to the two anonymous KULR reviewers for their helpful suggestions that greatly improved the article. I would also like to thank Toto Sutarso for guidance on statistical analysis. All remaining errors are my own. The research was supported by a course release grant from Middle Tennessee State University, College of Liberal Arts, TBR Access and Diversity Grants, for which I am very thankful.

that learners apply perfective and imperfective aspect to a restricted set of verb semantics before extending their use to a broader range of verbs.

The purpose of this study is to test the validity of these restricted relationships in the case of *toki* ('when') bi-clausal sentences produced by native and non-native speakers of Japanese. More specifically, this study examines the interaction of the lexical aspect of the verb phrases and the grammatical aspect (based on the inflections of verbal morphology *-ru*, *-ta*, *-te iru* and *-te ita*) in the case of *toki* bi-clausal sentences in the spoken discourse of native and non-native speakers of Japanese. The majority of previous L2 Japanese experimental research, as summarized in Table 2, has tested the associations between the perfective/past marker *-ta* with Achievements (punctual verbs) and imperfective/progressive marker *-te i* with Activities (durative verbs). This study will examine these associations and, in a departure from the primary focus of previous L2 Japanese studies, present the distribution of the non-past ending *-ru* over the four lexical classes. Whereas most previous studies used research methods involving controlled tasks such as grammaticality judgment, picture description and narratives, this study uses natural oral production data from lectures, conversations and speeches conducted in uncontrolled or natural environments.

Significantly, this study seeks to test the presence or absence of any restricted form-meaning relationships in the natural productions of native speakers of Japanese. This data will shed light on a possible lexical aspect-verbal morphology distribution pattern, if any, on the adult native grammar, hence validating the results obtained in Shirai (1995) for native speaker data.

Finally, studies so far have primarily investigated the interaction of verbal morphemes with lexical classes in mono-clausal independent sentences. The present study investigates the interaction at the subordinate verb position of a bi-clausal *toki* sentence, thus expanding the scope of investigation of the claims related to the Aspect Hypothesis.

The paper is organized as follows. Section 2 presents background literature and an overview of the experimental research done in Japanese addressing the claims of the Aspect Hypothesis. Section 3 delineates the research questions and method of study.

Section 4 presents preliminary observations and data analysis. Discussion and concluding remarks follow in sections 5 and 6 respectively.

2.0. Background

This section presents background literature on the concepts of tense and aspect and their interaction in *toki* sentences and a summary of previous experimental research investigating the claims of the Aspect Hypothesis.

2.1. Tense and Aspect

Tense and aspect are concerned with the notion of temporality. Tense places an event on a timeline, relevant to the time of speech, namely, past, present and future (Reichenbach 1947). Aspect refers to how an event unfolds in time, focusing on the internal properties of an event, such as whether an event has already been completed (Comrie 1976, Chung & Timberlake 1985, Smith 1991, 1997). Aspectual distinctions can be explicitly marked using linguistic devices, usually auxiliaries or inflections (e.g., *-ing* for progressive). This is known as grammatical or “viewpoint” aspect (Smith 1991). The grammatical inflections available in Japanese relevant to aspect are *-ru*, *-ta*, *-te iru* and *-te ita*. The following examples illustrate some typical occurrences of these four inflections.

The inflection *-te i* takes both the progressive and resultative readings with certain kinds of verbs. These readings will be illustrated in the examples later in this section.

(1) Hanako-ga asita hasi-**ru**.

Hanako-Nom tomorrow run-NON PAST

‘Hanako **will run** tomorrow.’

(2) Hanako-ga kinoo hasit-**ta**.

Hanako-Nom yesterday run-PAST

‘Hanako **ran** yesterday.’

(3) Hanako-ga ima hasit-**te i-ru**.

Hanako-Nom now run-PROG-NON PAST

‘Hanako **is running** now.’

(4) Hanako-ga kinoo hasit-**te i-ta**.

Hanako-Nom yesterday run-PROG-PAST

‘Hanako **was running** yesterday.’

Regarding the **Japanese grammatical inflections**, researchers hold several opposing views about the status of *-ru* and *-ta* as aspect or tense markers. Some claim that *-ta* and *-ru* are aspectual markers and that the tense interpretation emerges as a derivative from the aspectual meanings (Soga 1983, Hasegawa 1999). Soga (1983) has given two reasons for this view. First, in Japanese conversations and narratives, there appears to be no consistency concerning the occurrences of *-ta* and *-ru*, which may even occur alternatively. Second, so-called tense agreement does not exist, that is, *-ta* and *-ru* are not necessarily tied to the time points of past or non-past but to meanings of completion or incompleteness. Commenting on the Japanese past marker, Horie (1997) states that the Japanese past tense marker is often considered to have a sense of “perfect” or “perfective” because it is still in the process of grammaticizing from a perfect into a perfective aspect marker and then into a simple past tense marker. Due to this transitional nature of the Japanese past tense marker, there have been disagreements as to whether *-ta* is an aspect marker or tense marker (Takahashi 1976, Hasegawa 1999).

Besides using grammatical inflections to explicitly mark aspectual distinctions, these distinctions can also be determined by the lexical semantics of the verb phrase. **Inherent lexical aspect** (also known as situation aspect, semantic aspect) “encodes the aspectual contribution of the verb phrase” (Gabriele 2005) and usually refers to the Vendler/Dowty four-way classification (Vendler 1967, Dowty 1979). This classification contains the four lexical classes of Statives, Accomplishments, Activities and Achievements which are distinguished by the three features of dynamicity, telicity and

punctuality (Anderson 1991, Chung & Timberlake 1985, Comrie 1976). Dynamic events describe a process that changes (e.g., to eat) distinct from states that exhibit little or no change over time (e.g., to sleep). Telic predicates exhibit processes with an inherent limit (e.g., to run a mile), while atelic predicates reveal processes without such a limit (e.g., to run in a park). Punctual events take place instantaneously (e.g., to recognize), whereas durative events take place over a period of time (e.g., to bake a cake). The Vendler/Dowty four-way distinction is described below.

Achievements are telic and punctual predicates, 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 are *toku* ‘to arrive,’ *deru* ‘to leave,’ *hairu* ‘to enter.’

Accomplishments are telic and non-punctual predicates, and, thus, have some duration but with a single, clear, inherent end point. These are often verb phrases. Examples are *gohan o taberu* ‘to eat food,’ *tegami o kaku* ‘to write a letter,’ *keeki o tukuru* ‘to bake a cake.’

Activities are atelic predicates and thus do not assume a natural or inherent end point. They have inherent duration in that they involve a span of time. Examples are *taberu* ‘to eat,’ *kaku* ‘to write,’ *tukuru* ‘to make.’

Statives are non-dynamic predicates that persist over time without a change. Examples are *aru* ‘to be,’ *iru* ‘need,’ *dekiru* ‘be able.’

Table 1 below summarizes Vendler/Dowty’s classification.

Table 1: Vendler/Dowty Verb Classification

	Dynamicity	Telicity	Punctuality
Achievements	+	+	+
Accomplishments	+	+	–
Activities	+	–	–
Statives	–	–	–

Several modified versions of the Vendler/Dowty classification have been proposed (e.g., Smith 1991, Kindaichi 1976) that have additional (semelfactives in Smith) or different (Type IV in Kindaichi) categories of verb phrases. Due to the limited applicability of the new verb phrase categories, these classifications have not been as robustly developed or applied in previous research on tense-aspect acquisition as the Vendler/Dowty classification. Hence, the Vendler/Dowty classification will be adopted for the present study.

As stated previously, the focus of this study is on examining the interaction of the lexical aspect of verb phrases (based on the Vendler/Dowty classification) and verbal inflections in the case of *toki* bi-clausal sentences. *Toki* bi-clausal sentences have the basic form S1 *toki* S2, where S1 is the subordinate clause and S2 is the main clause. *Toki* is compatible with all the four inflections of verbal morphology and, hence, an ideal construction for investigating the interactions of the morphemes with the lexical classes. Examples of *toki* sentences illustrating these interactions are given below. The order of presentation of the verbs or verb phrases in these examples is based on increasing length of their temporal span, that is, Achievements < Accomplishments < Activities < Statives. Note that the inflection *-te i* combines with Achievements to give a resultative reading (example 7), with Activities it prototypically gives a progressive reading (example 13), and with Accomplishments it gives both resultative and progressive readings (examples 10a and 10b).

-ru + Achievement

(5) Ie-o **de-ru** toki, tomodati-ga kita.

House-Acc leave-NON PAST toki friend-Nom came

‘My friend came **before I left** the house.’

-ta + Achievement

(6) Ie-o **de-ta** toki, tomodati-ga kita.

House-Acc leave-PAST toki friend-Nom came

‘My friend came **after I left** the house.’

-te iru + Achievement (Resultative)

- (7) Ohuro-ni **hait-te iru** toki, tomodati-ga kita.
Bath-Loc enter-RESULT NON PAST toki, friend-Nom came
'My friend came **while I was in** the bath.'

-ru + Accomplishment

- (8) Sarada-o **tuku-ru** toki, te-o aratta.
Salad-Acc make-NON PAST toki, hand-Acc washed
'I washed by hands **before making** salad.'

ta + Accomplishment

- (9) Gohan-o **tabe-ta** toki, te-o aratta.
Meal-Acc eat-PAST toki hand-Acc washed
'I washed my hands **after I ate**.'

-te iru + Accomplishment (Progressive/Resultative)

- (10a) Gohan-o **tabe-te iru** toki, haha-ni hanasu tumori da.
Meal-Acc eat-PROG NON PAST toki, mother-Dat speak intent COP
'I plan to talk to my mother, **while eating** dinner.'

- (10b) Daigaku-ni **it-te iru** toki, uti-ni dengon-ga kita.
University-Loc go-RESULT NON PAST toki, home-Loc message-Nom came.
'A message came when **I was gone** (to school).'

-ru + Activity

- (11) Denwa **su-ru** toki, itumo bangoo-o kakunin suru.
Phone do-NON PAST toki, always number-Acc confirm
'I always check the number **before calling**.'

-ta + Activity

- (12) **Yon-da** toki, itumo memo-o totte oita.
Read-PAST toki always note-Acc take down
'I always took notes **after reading.**'

-te iru + Activity (Progressive)

- (13) Benkyoo **si-te iru** toki, itumo jisyo-o tukau.
Study do-PROG NON PAST toki, always dictionary-Acc use
'I always use a dictionary **while studying.**'

-ru + Stative

- (14) Nihon-ni **i-ru** toki, kimono-o kita.
Japan-Loc be-NON PAST toki, kimono-Acc wore
'I wore a kimono **while (living) in Japan.**'

-ta + Stative

- (15) Nihon-ni **i-ta** toki, kimono-o kita.
Japan-Loc be-PAST toki, kimono-Acc wore
'I wore a kimono when **I was (living) in Japan.**'

The next subsection will present an overview of the experimental research done on Japanese addressing the claims of the Aspect Hypothesis.

2.2. Approaches to the Acquisition of Aspect

In the literature covering the acquisition of aspect, a significant number of studies have investigated the Aspect Hypothesis (e.g., reviews in Bardovi-Harlig 1999, Li & Shirai 2000, Slabakova 2001), which is based on the theory of lexical aspect. The current postulation of the Aspect Hypothesis (Anderson & Shirai 1996) claims that lexical aspectual classes influence beginning language learners (both L1 and L2) in

their production of verbal morphology. The main claims of the most current version of the Aspect Hypothesis are as follows (Anderson & Shirai 1996):

- (i) Learners first use (perfective) past marking on Achievements and Accomplishments, eventually extending use to Activities and Statives.
- (ii) In languages that encode the perfective/imperfective distinction, imperfective past appears later than perfective past, and imperfect past marking begins with Statives, extending next to Activities, then to Accomplishments, and finally to Achievements.
- (iii) In languages that have a progressive aspect, progressive marking begins with Activities, and then extends to Accomplishments and Achievements.
- (iv) Progressive markings are not incorrectly overextended to Statives.

The hypothesis predicts that in early acquisition, tense-aspect morphemes will show a differential distribution among the lexical aspect classes of verbs. That is, learners will attach the perfective and imperfective markers to a restricted set of semantic verbs first, and then extend their usage to a broader range of verbs. In the initial stages, past/perfective marking will appear on Achievements/Accomplishments, progressive will appear on Activities, and imperfective marking will appear on Statives. In the later stages, the distribution of aspectual morphemes will be extended to other classes with the exception of progressive on Statives. In other words, initially they use perfective aspect with verbs that denote telic, punctual and resultative events, and imperfective aspect with verbs that denote durative and continuous events. Gradually, they expand their use to less prototypical cases, e.g. perfective aspect markers with Activity verbs and imperfective markers with Achievement verbs (Shirai & Andersen 1996). In order to explain the rationale for these patterns, a number of theories have emerged that attribute the differential distribution to prototypes, distributional bias of input data, and natural affinity of combinations.

In language acquisition, the prototype theory claims that L2 learners acquire a linguistic category starting with a prototype of the category and later expand its

application to the less prototypical cases (Slabakova 2002). Proponents of the Aspect Hypothesis argue that the prototype of past tense is restricted to verbal predicates that describe punctual, completed events (characteristics of Achievements), and the prototypical progressive to an event in progress (characteristic of Activities) [Anderson 1991, Anderson & Shirai 1994, Shirai 1991].

The Distributional Bias Hypothesis states that standard associations of inherent aspect and morphology in general input are marked by high frequency of Activities with progressive inflection and very low frequency of Stative progressives (Anderson 1988). Shirai (1993) opposes this hypothesis, stating that native speakers use the morphemes in less than prototypical ways and that, therefore, learners may not be getting the prototypes from input only.

The earliest explanation for the skewed distribution of tense-aspect morphemes and certain lexical classes on semantic and pragmatic grounds is suggested by Comrie (1985) in the “Naturalness of Combinations” principle. According to this principle, some aspectual morphemes combine more naturally with some verb types than with others. For example, perfective but not imperfective aspect markers combine naturally with punctual verbs because the former present a situation as a single whole without reference to its internal structure, and punctual verbs denote a single point lacking structure.

In previous studies, the above theories have helped to account for the restrictions in form-meaning relationships in Japanese tense-aspect acquisition. This study explores the possibility that these theories can explain the form-meaning distributional patterns of Japanese non-native as well as native oral production data containing *toki* ‘when’ constructions.

2.3. Experimental Research on the Aspect Hypothesis: L2 Japanese

Past experimental research in tense-aspect acquisition of Japanese has focused primarily on the first and third claims of the Aspect Hypothesis. The following table presents an overview of the research summarized from Shirai (2002) and Ishida (2004).

Table 2: Summary of Research on the Aspect Hypothesis: L2 Japanese

Study	Research Method	Result	Claims of AH
1. Kurono 1994	Grammaticality Judgment Task; Mixed L1 learners	ACT + <i>-te i</i> (PROG) easier than ACH + <i>-te i</i> (RESULT)	Claim (iii): Support
1. Shirai 1995	Conversational Interviews; L1 Chinese	Strong Association for: ACH+ <i>-ta</i> ACT+ <i>-te i</i>	Claim (i): Support Claim (iii): Support
1. Sheu 1997	Picture Description; Paper-pencil; L1 Chinese	ACT + <i>-te i</i> (PROG) higher than ACH + <i>-te i</i> (RESULT)	Claim (iii): Support
3a. Nishikawa 1998	Written and Oral Elicitation (L1 Chinese, Korean, English, Spanish)	---same as above---	Claim (iii): Support
3b. Koyama 1998	Grammaticality Judgment; L1 Korean, Chinese	---same as above---	Claim (iii): Support
1. Kurono 1998	Speech Samples	Strong Association for: ACH+ <i>-ta</i> Inconclusive for <i>-te i</i>	Claim (i): Support
1. Shibata 1999	Frog story narrative analysis	Strong Association for: ACH+ <i>-ta</i> ACT + <i>-te i</i> (PROG) lower than ACH + <i>-te i</i> (RESULT)	Claim (i): Support Claim (iii): Oppose
1. Ishida 2004	Conversation Data; 4 learners	ACT + <i>-te i</i> (PROG) lower than ACH + <i>-te i</i> (RESULT)	Claim (iii): Oppose
1. Sugaya & Shirai 2007	Oral Picture Description, Acceptability Judgment	Strong Association for: ACT+ <i>-te i</i>	Claim (iii): Support

ACH: Achievements, ACT: Activities, PROG: Progressive, RESULT: Resultative

As can be seen from the above summary, several studies support the association of *-ta* with Achievements and *-te i* with Activities, while others oppose the latter. Additionally, Shirai (1995) shows that native speakers associate *-ta* with Achievements, but not so strongly. Reviews of these works such as Salaberry and Shirai (2002) acknowledge that both learner internal (i.e. native language, innate constructs), and learner external (i.e., instruction, input, environment) factors need to be considered to explain the differential distribution of the verbal morphology and lexical classes.

In another recent study outside the realm of tense-aspect acquisition, Nara (2011) examines the skewed distribution of past and non-past tense markers over the four lexical classes in a written Japanese narrative. In this study, Natsume Soseki's 1906 novel *Botchan* was chosen as the corpus of analysis, and each sentence was analyzed

for a variety of information including aspectual classification of the sentence (based on Vendler/Dowty's model). Although the data showed a large proportion of the ending *-ru* marking in the entire narrative, "there seemed to be a tendency that the shorter the temporal span described by the sentence, the more likely it is that the sentence would be set in the past and vice-versa" (Nara 2011). In other words, this study supports claim (i) of the Aspect Hypothesis in the case study of a written narrative.

In light of previous experimental research on tense-aspect acquisition in L2 Japanese, the goal of the present study is two-fold. First, to investigate the previously examined claims (i) and (iii) of the Aspect Hypothesis in non-native data to see if the restricted interactions between lexical aspect and grammatical inflections hold true in the case of *toki* 'when' bi-clausal constructions as well. Secondly, to examine the use and distributional patterns of the non-past ending *-ru* over the four lexical classes in non-native data. In addition, native speaker data in this study will shed light on any distributional bias that they may be feeding into their input that could possibly explain the output patterns of the learners.

3.0. Research Questions and Method of Study

3.1. Research Questions

The research questions addressed in this study are:

- (1) Do the verbal morphemes show differential distribution with the lexical classes in the case of *toki* subordinate verbs?
 - (1a) Does the ending *-ta* associate first with Achievements and Accomplishments followed by Activities and Statives (i.e. claim (i) of Aspect Hypothesis)?
 - (1b) Does the ending *-ru* associate first with Activities and Statives followed by Achievements and Accomplishments?
 - (1c) Does the progressive *-te i* associate first with Activities and then extend to Achievements and Accomplishments (i.e. claim (iii) of Aspect Hypothesis)?
 - (1d) Is there a difference in the distribution of *-te iru* and *-te ita* with Activities?

- (1e) Does the sentence structure (with varying combinations of tense endings in the verbs, see section 3.2) show a skewed distribution with the lexical aspect classes?
- (2) Do the native speaker data and non-native speaker data show comparable patterns of distribution?

3.2. Method of Study

For this study, a collection of native and non-native corpora was chosen for analysis from the following sources:

Native corpora: (a) Uemura (b) BTS 2007 (c) Spoken Dialog

Non-Native corpora: (a) Uemura (b) BTSJ 2007 (c) BTSJ 2009

The collection of corpora contained spoken discourses such as free conversations, lectures and interviews. The discourses were produced in an uncontrolled environment and, therefore, were as close to natural speech as possible. Background information about the speakers and content of the production data is included in Appendix A.

The data was carefully scanned for instances of *toki*² used in the subordinate position of bi-clausal sentences and were listed separately. Then each sentence in the list was marked for three types of information: (1) lexical aspect classification of the verb in the *toki* clause, (2) tense-aspect ending of the verb and (3) sentence structure.

The lexical aspect type of the *toki* clause was coded using an integer from 1 to 4: 1 for Achievements, 2 for Accomplishments, 3 for Activities, and 4 for Statives. The tests of lexical aspect as outlined in Shirai (1993) were used to determine the lexical aspect type. See Appendix B for a summary of the tests.

For tense-aspect endings, 1 for *-ru*, 2 for *-ta*, 3 for *-te iru* and 4 for *-te ita* were used.

For sentence structures, *S1 toki S2* was divided into four types. *S1 ru toki S2 ru* (*Ru—Ru*) was coded 1, *S1 ru toki S2 ta* (*Ru—Ta*) was coded 2, *S1 ta toki S2 ru* (*Ta—Ru*) was

² This study includes all instances of subordinate *toki*, whether or not there is a particle following (e.g., *toki-ni*, *toki-wa*, *toki ni wa*) as long as it appears in a bi-clausal sentence.

A tally of the raw data regarding the distribution of the four tense-aspect endings (*-ru*, *-ta*, *-te iru*, and *-te ita*) yielded the following observations. In the non-native data, among 207 tokens, close to half were marked by *-ru* ending, followed by *-ta*. See Table 3 below.

Table 3: Distribution of Tense-Aspect Endings at S1 Position – Non Native

-ru	-ta	-te iru	-te ita	Total
102	79	15	11	207
49.3%	38.2%	7.2%	5.3%	

In the native data, among 541 tokens, close to half were marked by *-ta* ending, followed by *-ru*. See Table 4 below.

Table 4: Distribution of Tense-Aspect Endings at S1 Position – Native

-ru	-ta	-te iru	-te ita	Total
209	263	58	11	541
38.6%	48.6%	10.7%	2.0%	

In the overall distribution of tense and aspect endings, native speakers tend to mark the verbs with *-ta* more often (48.6%) than with *-ru* (38.6%). Non-native speakers, on the other hand show an opposite trend by marking the verbs with *-ru* more often (49.3%) than with *-ta* (38.2%). The progressive *-te i* endings are consistently low in frequency, with *-te iru* slightly higher than *-te ita* in both data sets.

4.2. Distribution of Lexical Aspect Classes

In the non-native data, among 207 tokens, Activities had the highest appearance, followed by Accomplishments, Achievements and Statives, in that order. See Table 5 below. Based on telicity, the distribution was even for telic verbs and atelic verbs. See Table 5a below.

Table 5: Overall Occurrence of Verbs at S1 Position – Non Native

Achievement	Accomplishment	Activity	Stative	Total
44	61	77	25	207
21.3%	29.5%	37.2%	12.1%	

Table 5a: Telicity Distribution – Non Native

	Overall Occurrence	%
Telic	105	50.7
Atelic	102	49.3

In the native data, among 541 tokens, Achievements had the highest appearance followed by Activities, Accomplishments and Statives, in that order. See Table 6 below. Based on telicity, the distribution was higher for telic verbs (Achievements and Accomplishments) than for atelic verbs (Activities and Statives). See Table 6a below.

Table 6: Overall Occurrence of Verbs at S1 Position – Native

Achievement	Accomplishment	Activity	Stative	Total
236	63	195	47	541
43.6%	11.6%	36.0%	8.7%	

Table 6a: Telicity Distribution – Native

	Overall Occurrence	%
Telic	299	55.3
Atelic	242	44.8

In the case of native data, Achievements top the chart with an overall appearance of 43.6%, whereas in the non-native data, Activities show the highest appearance with 37.2%. The telic-atelic distribution is even in the case of non-native speakers, but skewed towards telic verbs (Achievements and Accomplishments) in the case of native speakers.

4.3. Lexical Aspect Correlated with Endings

The preliminary observations of data correlating lexical aspect and ending choice are reflected below. Since the aspect and tense categories are non-contiguous nominal categories, a chi-test⁶ was performed on these tables. For non-native data, the chi-square value (χ^2) was 41.05, d.f.=9. At a confidence interval of $p < 0.0001$, it was found that the lexical aspect/tense ending correlation is statistically significant for native data. Statistical significance means that chance is an unlikely explanation for the observed relationship (Kirk 1990). This result was corroborated by Cramer’s v value (Cramer 1999), which is a post-test to determine the strength of association between lexical aspect and endings that yielded a value of 0.257, indicating a moderate association.⁷ The Lambda value indicated the following: ending without the knowledge of lexical aspect 37.20%; ending from lexical aspect 43.96%; lexical aspect without the knowledge of ending 49.28%; lexical aspect from ending 53.14%. The Lambda value is also a post-test that gives a comparison of strengths of association with or without the knowledge of one of the correlates. An increase in the value from ‘without the knowledge’ to ‘with the knowledge’ (37.20% \rightarrow 43.96% and 49.28% \rightarrow 53.14%) is an indicator of association strength, moderate in this case. In other words, the increase is interpreted as the amount of improvement in predicting the dependent variable (lexical aspect) that can be attributed to the independent variable (tense ending) (Kirk 1990). The values for the telicity distribution were ($\chi^2=18.31$, d.f.=3, $p < 0.001$, Cramer’s v value=0.297, moderate association). See Tables 7 and 7a.

Table 7: Lexical Aspect Correlated with Ending – Non Native

	-ta	-ru	-te ita	-te iru	Total
--	-----	-----	---------	---------	-------

⁶ Summary of Statistical Terms (Nara 2011)

A chi-square is a statistical test that is based on the calculation of the discrepancy between the observed and theoretical frequencies. In other words, it is a measure of uneven distribution. Degrees of freedom (*d.f.*) refer to the number of values in the contingency table that are free to vary while preserving the mean. The confidence interval, shown by the p value, is the probability for these frequencies to appear by chance. Usually a p value below 0.05 is taken as a level suitable for rejecting the null hypothesis and accepting the research hypothesis. P values smaller than 0.05 are generally regarded as statistically significant.

⁷ Measures of association strength range from 0 to 1; the larger the number, the stronger the association (Kirk 1990). A general rule of thumb for interpreting the strength of association (1=strongest; 0=weakest) is: < 0.10 Weak, $0.11 - 0.30$ Moderate, > 0.31 Strong.

Achievement	22	19	1	2	44
Accomplishment	28	32	0	1	61
Activity	14	41	10	12	77
Stative	15	10	0	0	25
Total	79	102	11	15	207

Table 7a: Telicity Distribution – Non Native

	-ta	-ru	-te ita	-te iru	Total
Telic	50	51	1	3	105
Atelic	29	51	10	12	102
Total	79	102	11	15	207

For native data, the chi-square value (χ^2) was 75.59, d.f.=9. At the confidence interval of $p < 0.0001$, it was found that the correlation was statistically significant in the native data. Cramer's v value yielded 0.216, suggesting a moderate association. The Lambda value indicated the following: ending without the knowledge of lexical aspect 43.62%; ending from lexical aspect 50.46%; lexical aspect without the knowledge of ending 48.61%; lexical aspect from ending 53.60%. A marginal increase in the Lambda values in this correlation suggests moderate association strength. The values for the telicity distribution were ($\chi^2=48.29$, d.f.=3, $p < 0.0001$, Cramer's v value=0.299, moderate association). See Tables 8 and 8a.

Table 8: Lexical Aspect Correlated with Endings – Native

	-ta	-ru	-te ita	-te iru	Total
Achievement	137	83	1	15	236
Accomplishment	44	17	1	1	63
Activity	59	85	9	42	195
Stative	23	24	0	0	47
Total	263	209	11	58	541

Table 8a: Telicity Distribution – Native

	-ta	-ru	-te ita	-te iru	Total
Telic	181	100	2	16	299

Atelic	82	109	9	42	242
Total	263	209	11	58	541

In the case of native data, the association between *-ta* and Achievements and that between *-ru* and Activities are the highest. For the progressive marker *-te i*, Activities were the highest for both *-te iru* and *-te ita* (with *-te ita* showing a higher appearance than *-te iru*). In the overall usage of *-te i* to give progressive readings with Activities and resultative reading with Achievements, the result shows that progressives (frequency=51) clearly outnumber the resultative (frequency=16) readings. In the telicity distribution, telic was higher than atelic for *-ta*, but atelic is higher than telic for the rest of the endings (*-ru*, *-te iru* and *-te ita*). In the case of non-native data, Accomplishments associate the most with *-ta*, while Activities associate the most with the rest of the endings. Within Activities, *-te iru* is marginally higher than *-te ita*. As in the case of native data, the progressive *-te i* readings (frequency=22) outnumber the resultative *-te i* readings (frequency=3). In the telicity distribution, telic is higher than atelic for *-ta*, while atelic is higher than telic for *-te ita* and *-te iru*. Telic and atelic are even for *-ru*.

From the above results, it is clear that both native and non-native speakers associate telicity with *-ta* and atelicity with *-te i* and *-ru*. This is an interesting result, especially since native speakers contextually tend to use more *-ta* as well as more telic verbs overall in their utterances (c.f. Tables 3 and 5). The non-native data results point to supporting the claims of the Aspect Hypothesis. More on this will be presented in the discussion section.

4.4. Lexical Aspect Correlated with Sentence Structure

The preliminary observation of data correlating lexical aspect and sentence structure is reflected below. A chi-square test performed on the non-native data showed that there exists a statistically significant difference in the way the cell values are distributed ($\chi^2=18.52$, d.f.=9, $p<0.05$). Cramer's v value was calculated as a post-test and yielded a value of 0.173, suggesting a weak to moderate association. As for predictability, the Lambda value for predictability was obtained, which showed the

following predictability numbers: structure without the knowledge of lexical aspect 37.20%, structure from lexical aspect 42.51%, lexical aspect without the knowledge of structure 43.96% and lexical aspect from structure 44.93%. A very marginal increase in the Lambda values in this lexical aspect-sentence structure correlation suggests weak to moderate association strength. The values for the telicity distribution were ($\chi^2=2.85$, d.f.=3, $p<0.05$, Cramer's v value=0.117, weak to moderate association). See Tables 9 and 9a.

Table 9: Lexical Aspect Correlated with Structure – Non Native

	Ru—Ru	Ru—Ta	Ta—Ru	Ta—Ta	Total
Achievement	16	5	13	10	44
Accomplishment	25	8	6	22	61
Activity	41	12	8	16	77
Stative	9	1	4	11	25
Total	91	26	31	59	207

Table 9a: Telicity Distribution – Non Native

	Ru—Ru	Ru—Ta	Ta—Ru	Ta—Ta	Total
Telic	41	13	19	32	105
Atelic	50	13	12	27	102
	91	26	31	59	207

A chi-square test performed on the native data showed that there exists a statistically significant difference in the way the cell values are distributed ($\chi^2=80.48$, d.f.=9, $p<0.0001$). Cramer's v value was calculated as a post-test and yielded a value of 0.223, suggesting a moderate association. As for predictability, a Lambda value for predictability was obtained, which showed the following predictability numbers: structure without the knowledge of lexical aspect 43.62%, structure from lexical aspect 48.98%, lexical aspect without the knowledge of structure 41.59%, and lexical aspect

from structure 46.77%. A marginal increase in the Lambda values in this correlation suggests moderate association strength. The values for the telicity distribution were ($\chi^2=30.18$, d.f.=3, $p<0.0001$, Cramer's v value=0.236, moderate association). See Tables 10 and 10a.

Table 10: Lexical Aspect Correlated with Structure – Native

	Ru—Ru	Ru—Ta	Ta—Ru	Ta—Ta	Total
Achievement	84	14	88	50	236
Accomplishment	14	4	7	38	63
Activity	107	20	37	31	195
Stative	20	4	14	9	47
Total	225	42	146	128	541

Table 10a: Telicity Distribution – Native

	Ru—Ru	Ru—Ta	Ta—Ru	Ta—Ta	Total
Telic	98	18	95	88	299
Atelic	127	24	51	40	242
Total	225	42	146	128	541

For the native data set, Activities have the highest appearance in *Ru—Ru* and *Ru—Ta*, while Achievements are highest in *Ta—Ta* and *Ta—Ru* patterns. In the telicity distribution, atelic is higher than telic for *Ru—Ru* and *Ru—Ta* patterns, while telic is higher than atelic for *Ta—Ta* and *Ta—Ru* patterns. In the case of the non-native data set, Activities have the highest appearance in *Ru—Ru* and *Ru—Ta* patterns, but Achievements in *Ta—Ru* and Accomplishment in *Ta—Ta* have the highest representations. Looking at the telicity distribution, atelic is higher than telic for *Ru—Ru*, but are exactly even for *Ru—Ta*. For *Ta—Ru* and *Ta—Ta*, telic was higher than atelic by a marginal difference.

Overall, both the native and non-native data sets give comparable results in this category. Telic verbs uniformly trend to a higher appearance on the *Ta—Ta* and *Ta—Ru* structures, while atelic verbs have a higher appearance on *Ru—Ru* and *Ru—Ta*

structures. In order to explain this common pattern of distribution for both sets of data, the field needs a theory that applies universally to all speakers, native and non-native.

5.0. Discussion

First, let us look at the **non-native data set**. Claim (i) of the Aspect Hypothesis which states that *-ta* associates first with Achievements and Accomplishments (i.e. telic verbs), followed by other classes was reinforced in the non-native data. Additionally, it was also observed that *-ru* associates with Activities predominantly over other classes. Claim (iii) of the Aspect Hypothesis that progressive *-te i* attaches first with Activities was also true, with *-te iru* marginally higher than *-te ita* within Activities. Also, in the overall usage of *-te i* to give progressive readings with Activities and resultative reading with Achievements, the result showed that progressives clearly outnumber the resultative readings in frequency. This frequency pattern is in alignment with the results of previous studies (Kurono 1994, Sheu 1997, Shibata 1999, Ishida 2004).

In sum, the distributional pattern in the non-native data set supports the two claims of the Aspect Hypothesis. In order to account for these patterns of interaction between grammatical aspect (tense-aspect endings) and inherent lexical aspect of verbs (semantic verb classes), let us once again consider the “Naturalness of Combinations Principle” as postulated by Comrie (1976). According to this principle natural combinations exist between grammatical and lexical aspects. For example, perfective aspect markers are naturally associated with telic verbs (e.g. *hait+ta* ‘enter’) as both grammatical and lexical aspects denote a situation as a simple whole with no internal structure. On the contrary, imperfective aspect markers are naturally associated with atelic verbs (e.g. *tabe+te iru* ‘eating’) as both denote the internal structure of an event with no fixed temporal boundary. In other words, a learner will associate an inflection with a verb according to how congruent the form of the inflection is with the meaning of the verbal predicate. Progressive inflection aligns with duration and dynamicity and, therefore, matches with Activities. Past inflection implying completion aligns closely with Achievements that achieve completion as soon as they occur or Accomplishments that reach completion in a finite amount of time.

Moreover, the interaction between lexical classes and sentence structure for non-native speakers showed a similar trend with mostly telic verbs appearing in *Ta—Ta* and *Ta—Ru* structures, and atelic (mostly Activities) appearing in *Ru—Ru* and *Ru—Ta*. This result implies that the ending choice in the subordinate position predominantly follows the prototypical associations proposed by the Aspect Hypothesis and explained by the Naturalness of Combinations principle irrespective of the type of the ending (*-ru* or *-ta*) in the main verb position.

Another theory explaining the results of the non-native data is to explore the possibility of a distributional bias in the input provided by the native speakers. In this study it is interesting to note that the **native data set** showed similar patterns of distribution (compared to the non-native data set) in the correlations between lexical aspect and tense endings as well as sentence structure. Like the learners, the native speakers mainly associated the perfective/past ending *-ta* with telic verbs (Achievement and Accomplishment), followed by other verb categories. Additionally, non-past ending *-ru* as well as imperfective ending *-te i* were predominantly attached to Activities, more than any other class. Within Activities, *-te iru* was higher than *-te ita* substantially for native speakers. As in the case of non-native data, the progressive *-te i* readings (with Activities) outnumbered the resultative *-te i* readings (with Achievements). The interaction between lexical classes and sentence structure showed a trend similar to that of the non-native data with mostly telic verbs appearing in *Ta—Ta* and *Ta—Ru* structures, and atelic (mostly Activities) appearing in *Ru—Ru* and *Ru—Ta*. The results from the native data set, in fact, reinforce Shirai's (1995) conclusion that a skewed distribution of the verbal inflections over lexical classes is present in the general input of native speakers. Li and Shirai (2000) claimed that the association as postulated by Comrie's Naturalness of Combinations Principle in fact continues to persist in adults. Since there is a natural relationship in the expression of telicity of an event with the use of a specific aspect marker in reality, adults are still likely to associate telicity of an event with perfective aspect markers and atelic event with imperfective aspectual forms. This makes it possible to explain the similarity in trends in both native and non-native data using the same principle of Naturalness of Combinations. The distribution

patterns in this study indicate that speakers of Japanese (both native and non-native) are mindful of the semantic features such as telicity and punctuality, and make their verbal ending choices based on the presence or absence of these features.

In the present study, although native and non-native data gave comparable results for the most part, there were some differences as well. First, while native data had a predominance of *-ta* forms, non-native data had a predominance of *-ru* forms. This is because native speakers chose more Achievements (hence, more *-ta*), while non-native speakers chose more Activities (hence, more *-ru*) in the spoken discourse. It may be conjectured that the tendency to use one lexical class more frequently than another could be a result of the influence of type and context (scenario) of the spoken discourse. In order to confirm this hypothesis, there is a need to look into the specific context in which each discourse was carried out, the types of topics within those discourses, and participant relationships. This hypothesis can be tested in a future study. The predictions may also be different depending on the genre of discourse under study. For example, the results have been found to be quite different in written narratives, as in the case of Nara (2011) where the native text (a literary novel) had a predominance of *-ru* forms.

Secondly, in the correlation between lexical aspect and sentence structure, although the data was statistically significant ($p < 0.05$), the strength of association was moderate for native speakers and weak for non-native speakers. Small data size can be a possible contributing factor towards the weak association results in the non-native data set. Although six online corpus banks (three each for both sets) were used to extract the data, the number of tokens available for the two categories, especially the non-native data, was small.⁸ It is possible that a larger sample of the non-native data might lead to a slightly better agreement in the results of this study.

6.0. Concluding Remarks

⁸ In native spoken discourses and more so in non-native discourses, the instances of *toki* with verbs in subordinate position are limited in frequency as compared to other combinations with *toki* such as with nouns (*gakusei no toki* 'student days') or with adjectives (*wakai toki* 'younger days'). These latter instances of *toki* were not within the scope of this study and therefore were excluded from the data.

In summary it may be concluded that the claims of the Aspect Hypothesis that learners tend to associate past/perfective marking with telic verbs and progressive marking with Activities turned out to be largely accurate for non-native data in subordinate verb positions in *toki* bi-clausal sentences. Similar distributional patterns were also seen in the native data set. Additionally, there seems to be a strong tendency for the *Ta—Ta* and *Ta—Ru* sentence structures to take telic verbs, and *Ru—Ru* and *Ru—Ta* structures to take atelic verbs in the *toki* constructions in both data sets.

It has become clear from this study that the empirically supported differential distributions among tense-aspect verb endings, sentence structures and lexical aspect classes apply across native and non-native discourses. This proves that language users (L1 and L2) are universally sensitive to the lexical aspectual classes and the semantic features that govern them, such as telicity, punctuality and duration. This conclusion may have useful pedagogical implications for the field of second language education.

Since both L1 and L2 users of Japanese are naturally governed by the semantic features, it would be useful to introduce the lexical aspect based verbal classification (such as Vendler/Dowty's) in Japanese language classrooms. The traditional verb classifications, based on morphological endings (classified into *-u* verbs, *-ru* verbs, and irregulars) and widely used in current Japanese language textbooks, primarily rely on form distinctions. It would be useful to add a function component to the existing classification (by introducing the semantic features of telicity, punctuality and duration) that can explain a wider range of phenomena in Japanese. In addition to explaining the nuances drawn in *toki* clauses based on different tense and aspectual endings, the Vendler/Dowty classification can also be used to explain other key aspectual usages of verbs such as *-te iru* 'ongoing/resultative,' stem + *hazimeru/dasu* 'start of a verbal action,' the contrast between *made/made ni* 'until/by,' etc., that use the concept of durative/non-durative aspect of the verbs to bring out their accurate interpretation, and cannot be explained by the form-based classification alone. A testable hypothesis in this direction would be to see if explicit teaching of the relationships between lexical aspect and verbal inflections causes an improvement in the performance of those L2 learners whose native languages do not have overt tense markers (such as Chinese), and who

may potentially have difficulty in making the form and function associations in the beginning stages of their language learning.

A future investigation, in addition to replicating this study with a larger sample of non-native data, might include analysis of written production of native and non-native use of *toki* and of another temporal structure similar to *toki*, such as *tokoro* ‘place’, which allows all tense and aspect endings in the preceding verb. These suggestions are likely to yield useful results complementing the present study.

References

- Andersen, R.W. 1988. Models, processes, principles and strategies: Second language acquisition in and out of the classroom. *IDEAL*. Vol. 3, pp. 111-138.
- Andersen, R.W. 1991. Developmental Sequences: The emergence of aspect marking in second language acquisition. T. Huebner (Ed.) *Crosscurrents in second language acquisition and linguistic theories*, pp. 305-324. Amsterdam: Benjamins.
- Andersen, R.W. 1993. Four operating principles and input distribution as explanation for underdeveloped and mature morphological systems. K. Hyltenstam and A. Viborg (Eds). *Progression and Regression in Language*. Cambridge: Cambridge University Press.
- Andersen, R.W., Y. Shirai. 1994. Discourse motivations for some cognitive acquisition principles. *Studies in second language acquisition*, Vol. 16, pp. 133-156.
- Andersen, R.W., Y. Shirai. 1996. The primacy of aspect in first and second language acquisition: The pidgin-creole connection. W.C. Ritchie, T.K. Bhatia (Eds.) *Handbook of second language acquisition*, pp. 527-570. San Diego: Academic Press.
- Bardovi-Harlig, K. 1999. From Morpheme studies to temporal semantics: Tense aspect research in second language acquisition. *Studies in Second Language Acquisition*. Vol. 21:3, pp 341-382.

- Bardovi-Harlig, K., D.W. Reynolds. 1995. The role of lexical aspect in the acquisition of tense and aspect: *TESOL Quarterly*. Vol. 29, pp. 107-131.
- Chung, S. and A. Timberlake. 1985. Tense, aspect and mood. In T. Schopen (Ed.), *Language Typology and Syntactic Description, Volume 3. Grammatical Categories and the Lexicon*. Cambridge, UK: Cambridge University Press.
- Comrie, B. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Comrie, B. 1985. *Tense*. Cambridge University Press.
- Cramér, H. 1999. *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press.
- Dowty, D. 1979. *Word meaning and Montague grammar*. Dordrecht and Boston: D. Reidel Publishing Company.
- Gabriele, A. 2005. *The Acquisition of Aspect in a Second Language: a Bidirectional Study Learners of Japanese and English*. Doctoral Dissertation. City University of New York.
- Hasegawa, Y. 1999. Tense-aspect Controversy Revisited: The -ta and -ru forms in Japanese. J. Verschueren (Ed.) *Pragmatics in 1998: Selected papers from the 6th International Pragmatics Conference*, pp. 240-255. Antwerp: International Pragmatics association.
- Horie, K. 1997. Form meaning interaction in diachrony: A case study from Japanese. *English Linguistics*. Vol. 14, pp. 428-449.
- Ishida, M. 2004. The effect of recasts on acquisition of aspect in Japanese as a foreign language. *Language Learning*, Vol. 54, pp. 311-394.
- Kindaichi, H. 1976. *A classification of Japanese verbs*.
- Kirk, R. 1990. *Statistics: An Introduction (Third edition)*. Holt, Rinehart and Winston, Inc.
- Koyama, S. 1998. Nihongo gakusyuuusya ni yoru asupekuto no syuutoku (The acquisition of tense-aspect by learners of Japanese). *9th National meeting of the Association of Second Language Acquisition*. Nagoya University.
- Kurono, A. 1994. *Nihongo gakushusha ni okeru tensu asupekuto no syuutoku ni tsuite*

- (The acquisition of tense aspect by learners of Japanese). MA Thesis. Nagoya University.
- Kurono, A. 1998. Syuuroo o mokuteki tosite taizai site iru gaikokuzin no tensu-asupekuto no syuutoki ni tuite [On the acquisition of tense-aspect by foreign workers]. Unpublished manuscript, Japan International Cooperation center, Tokyo.
- Lee, K. 1999. Relative Tense and Absolute Tense in Adverbial Adjuncts in Japanese. *Function and Structure: In honor of Susumu Kuno*, pp. 215-250. Amsterdam: Benjamins.
- Li, P., Y. Shirai. 2000. Acquisition of Aspect in Japanese. *The Acquisition of lexical and grammatical aspect*. Berlin and New York: Mouton De Gruyter.
- Nara, H. 2011. Aspect and Discourse in Tense Switching: A Case Study of Natsume Soseki's Botchan. *Japanese Language and Literature*, Vol. 45, pp. 273-305.
- Nishikawa, Y. 1998. *Teiru no syuutoku yooiin ni kansuru kenkyuu* (A study of factors that influence the acquisition of te iru). BA Thesis. Department of Japanese Language Teaching. Hiroshima University.
- Reichenbach, H. 1947. *Elements of Symbolic Logic*. New York: Free Press.
- Salaberry, R. and Y. Shirai (Eds.). 2002. *Tense-aspect morphology in L2 acquisition*. Amsterdam: John Benjamins.
- Sato, C. J. 1990. *The syntax of conversation in interlanguage development*. Tubingen: Gunter Narr.
- Sheu, S. 1997. Tyuu-zyookyuu taiwanzin nihongo gakusyusya ni yoru -teiru no syuutoku ni kansuru oodan kenkyuu [A cross-sectional study of the acquisition of –te iru by intermediate and advanced Taiwanese learners of Japanese] *Nihongo Kyooiku*. Vol. 95, pp. 37-48.
- Shibata, M. 1999. The use of Japanese tense-aspect morphology in L2 discourse narratives. *Acquisition of Japanese as a second language*. Vol 2, pp. 68-102.
- Shirai, Y. 1991. Primacy of aspect in language acquisition: Simplified input and prototype. Unpublished doctoral dissertation, University of California, Los Angeles.

- Shirai, Y. 1993. Inherent aspect and acquisition of tense-aspect morphology in Japanese. H. Nakajima, Y. Otsu (Eds.) *Argument Structure: Its Syntax and Acquisition*. Tokyo: Kaitakusha.
- Shirai, Y. 1995. Tense-aspect marking by L2 learners of Japanese. D. MacLaughlin, S. McEwen (Eds.) *Proceedings of the 19th Annual Boston University Conference on language Development*. Vol. 2. Somerville, MA: Cascadilla Press.
- Shirai, Y. 2002. The aspect hypothesis in SLA and the acquisition of Japanese. *Acquisition of Japanese as a Second Language* 5: 42-61.
- Slabakova, R. 2001. *Telicity in the second language*. Amsterdam: John Benjamins.
- Slabakova, R. 2002. Recent research on the acquisition of aspect: An embarrassment of Riches? *Second Language Research*. Vol. 18: 2, pp. 172-188.
- Smith, C.S. 1991. *The Parameter of Aspect*. Dordrecht: Kluwer Academic.
- Smith, C. S. 1997. *The Parameter of Aspect* (2nd Edition). Dordrecht: Kluwer Academic Publishers.
- Soga, M. 1983. *Tense and Aspect in Modern Colloquial Japanese*. University of British Columbia.
- Sugaya, N., Y. Shirai. 2007. The Acquisition of Progressive and Resultative meanings of the imperfective aspect marker by L2 learners of Japanese. *SSLA*. Vol 29, pp. 1-38.
- Takahashi, T. 1976. Commentary: A short history of the study of aspect in Japanese verbs. H. Kindaichi (Ed.) *Aspect in Japanese verbs*. Tokyo: Mugi Shobo.
- Vendler, 1967. Verbs and times. Z. Vendler (Ed.), *Linguistics and philosophy*, pp. 97-121. Cornell University Press.

Appendix A: Background Information about the Speakers and Production Data

Native Corpus

Corpus	Dialog sets	Participants
Uemura	50 sets	Japanese students, teachers, business clerks, housewives
BTS 2007	116 sets (1436 minutes=24 hours)	Local people
Spoken Dialog	93 sets (450 minutes)	Japanese students, teachers, office workers

Uemura: <http://www.env.kitakyu-u.ac.jp/corpus/texts/index.html>

BTS 2007: <http://www.tufs.ac.jp/ts/personal/usamiken/corpora2007.htm>

Spoken Dialogue Corpus: <http://research.nii.ac.jp/src/eng/list/files/PASD.pdf>

(BTS: Basic Transcription System)

Non-Native Corpus

Corpus	Dialog sets	Level of Japanese	Native Language
Uemura	49 sets	beginning to advanced	English, Thai, Korean, Russian, German, Chinese
BTSJ 2007	37 sets (691 minutes=11 hours)	intermediate and advanced	Taiwanese, French, Korean
BTSJ 2009	20 sets	intermediate and advanced	English, Chinese, Korean, Thai, Indonesian

Uemura: <http://www.env.kitakyu-u.ac.jp/corpus/texts/index.html>

BTSJ 2007: <http://www.tufs.ac.jp/ts/personal/usamiken/corpora2007.htm>

BTSJ 2009: http://japanese.human.metro-u.ac.jp/mic-j/mic-J_corpus_fl/btsjdata.html

(BTSJ: Basic Transcription System Japanese)

Appendix B: Summary of Tests to Determine Lexical Aspect Type (Shirai 1993)

Step 1: State or Non-State?

1a) Can it refer to present state in simple present tense without having a habitual or vivid-present interpretation?

If yes -- State (e.g., *Tukue no ue ni hon ga aru.* 'There is a book on the table.')

If no -- Non-State (e.g., *Boku wa gohan o taberu.* 'I will eat food.')

Step 2: Activity or Non-Activity? (Telic or Atelic)

2a) If you stop in the middle of an action, does that entail that you did it?

If yes -- Activity (e.g., *aruku.* 'walk')

If no -- Non-Activity (e.g., *Eki made aruku.* 'walk to the station.')

Step 3: Accomplishment or Achievement? (Punctual or Non-Punctual)

3a) 'If X wa Y de V-ta in' (Y=time; e.g., 10 minutes), does that entail X was involved in Ving (i.e., V-te ita) during that time?

If yes -- Accomplishment (e.g., *Kare wa itimai no e o kaita.* 'He painted a picture.')

If no -- Achievement (e.g., *Kare wa itimai no e ni kizuita.* 'He noticed a picture.')

3b) Can 'V-te iru' have the sense of "action-in-progress"?

If yes -- Accomplishment (e.g., *Kare wa oyu o wakasiteiru.* 'He is boiling water till it is hot.')

If no -- Achievement (e.g., *Kare wa sono e ni kizuiteiru.* 'He has noticed the picture.')

**Form-Meaning Associations in Japanese
--Analysis of Native and Non-Native Corpus Data--**

Priya Ananth

Abstract

Several studies on tense-aspect acquisition in Japanese have investigated the form-meaning associations that learners of Japanese make in their choice of verbal morphology. These studies make two popular claims, namely, that learners associate the perfective/past ending (*-ta*) with telic verbs (Achievements and Accomplishments) and progressive endings (*-te i*) with Activities. The goal of the present study is to test the validity of these claims in the case of *toki* ‘when’ subordinate clauses in spoken production data of native and non-native speakers of Japanese. A quantitative analysis of the data showed that speakers of Japanese in fact do make these form-meaning associations and are sensitive to the inherent semantic features of verbs such as telicity, punctuality and duration.

要旨

日本語テンス・アスペクト習得の先行研究に形態素「タ」形と限界性のある動詞（AchievementsとAccomplishments）または「テイル」形とActivitiesに密着性があるという仮説がよく論じられる。本研究はこの仮説の妥当性を「とき」従

属節に現れるテンス・アスペクトを対象に検討する。本論文の目的は、日本語母語話者および非日本語母語話者の発言における「とき」従属節に現れる「ル」、「タ」、「テイル」、「テイタ」という四つの形態素の選択とそれらの動詞を構成している、継続性、限界性、瞬間性といった意味論的な特徴の関わりを考察することである。オンラインコーパスから得た日本語話者の発言データをもとに数量的な分析を行い、上記の仮説を検討する。

受領日 2011年8月17日
受理日 2011年11月13日