

Linguistic Profiles of Heritage Bilingual Learners of Japanese

Kazue Kanno, Tomomi Hasegawa, Keiko Ikeda, and Yasuko Ito
University of Hawai'i at Manoa

1. Introduction

This paper reports the findings of an ongoing study on the linguistic profiling of heritage (simultaneous) bilingual learners of Japanese.

Heritage language (HL) instruction at the university level is on the rise, and some universities have even created a separate track for heritage bilingual learners of various languages in order to accommodate their needs, which differ from those of foreign language learners (see e.g. McGinnis, 1996; Sohn, 1995, 1996). However, very few studies (e.g. Douglas, 2001) have actually examined the linguistic characteristics of heritage bilinguals. The following statement by Campbell and Rosenthal (2000) summarizes this situation well:

"...The following are descriptions of what we believe to be the competencies and knowledge that average/typical HL students bring to university foreign language programs. These generalizations are based on our observations and those of others who work with this population. Because there are no known experiments, analyses, or publications to support these claims, we offer them as "working hypotheses" (Campbell & Rosenthal, 2000, p. 167).

This study seeks to begin to remedy this situation by examining the linguistic profiles of heritage bilingual learners of similar L2 proficiency—i.e., Advanced or Advanced Plus on the ACTFL Proficiency Scale, as assessed by the Oral Proficiency Interview (OPI).

We adopt two of Campbell and Rosenthal's "Working Hypotheses" (Campbell & Rosenthal, 2000, pp 167-169).

Hypothesis 2: Heritage language learners have acquired 80% to 90% of the grammatical rules that govern word, phrase, sentence, and discourse production and recognition.

Hypothesis 3: Heritage language learners have acquired extensive vocabularies; however, the semantic range of their vocabulary is limited to just a few sociocultural domains including hearth and home, neighborhood and, not infrequently, religious institutions.

2. Instruments

Data were collected using three instruments—a written test, a guided narrative, and conversation (i.e., the OPI).

2.1 Written test

The written test consisted of five sections. The first two tested learners' structural knowledge; the remaining three tested their lexical knowledge.

Items in the written test were selected from the lists of vocabulary and grammatical patterns in the Japanese Language Proficiency Test (JLPT). (This was done because the OPI involves holistic assessment, and only general guidelines are provided.) Two specific references used were the *Nihongo*

Noryoku-shiken Shutsudai Kijun (Japanese Language Proficiency Test: Test Content Specifications; the Japan Foundation, 1996) and *Ichiman-go Goi Bunruishuu* (the Classification of 10,000 words; Senmon Kyooiku Shuppan, 1998).

Based on the guidelines and criteria (e.g., the number of instruction hours, size of vocabulary, etc. for each level of the JLPT), JLPT Levels 1 and 2 roughly correspond to ACTFL Superior and Advanced, respectively, on the OPI scale. JLPT Level 1 is the minimum requirement for acceptance into a regular college undergraduate/graduate program in Japan; hence, it is appropriate to assume that Level 1 corresponds to the Superior level.

2.1.1 Structural knowledge:

According to ACTFL's Proficiency Guidelines, three key characteristics of structural knowledge at the Advanced level are (i) control of frequent structural patterns, (ii) partial control of honorific markings, and (iii) paragraph-level discourse. (See Makino, 1991, for detailed descriptions of each OPI level for Japanese, as well as ACTFL's Proficiency Guidelines). The first two characteristics are examined in the first section of the test; the third characteristic is dealt with in the second section.

a. Structural patterns

This section contains twenty structural items which include (i) persistent error patterns, such as anaphoric use of deictic demonstratives and phrasal particles (see Sakoda, 1993, 1996, 1997; Sakamoto, 1995; Yagi, 1996); (ii) items considered difficult to acquire, such as empathy/perspective-related patterns and giving/receiving verbs (see Tanaka, 1997, 1999); (iii) honorifics; and (iv) other advanced structural patterns (modality and conditionals). These are all JLPT Level 3 or Level 4 test items, except those involving modality, which are Level 2 items.

b. Knowledge of connectives/functional expressions

This section contains ten connectives/functional expressions from JLPT Level 2 (e.g., *...ni kanshite* 'regarding ...') that are essential for the paragraph-level discourse required at the Advanced level.

For each item in both sections, participants were asked to select one of four choices.

2.1.2 Lexical and collocational knowledge

Three sections examine learners' lexical knowledge. A particular focus is knowledge of collocation—i.e., selectional properties of lexical items.

c. Collocations involving pairs of elements

There were ten items from JLPT Level 2 in this section, e.g., *me-o tojiru* (for closing one's eyes, *tojiru* is the only verb that can be used; other 'closing' verbs, such as *shimeru*, are not appropriate.)

d. Collocations involving frequently used fixed idiomatic expressions

Although the JLPT considers idioms as Level 1 items (and the OPI guidelines state that the use of idioms is one of the characteristics of the Superior level), there is no specific list. Eleven idioms were selected from among those introduced in teaching materials for intermediate/advanced level Japanese and from among frequent and colloquial idioms involving body-part vocabulary, such as *hana-ga takai* ('the nose is high = proud').

e. Collocations involving mimetics (i.e., ideophonic expressions, not onomatopoeic words).

Japanese makes quite extensive use of mimetics, far more than English does, and such expressions are an integral part of the language. For instance, whereas English uses different words to distinguish among smiling, giggling, and guffawing, Japanese has only one verb *warau* 'laugh', which is accompanied by a mimetic expression, *nikoniko warau*, *kusukusu warau*, and *geragera warau*. These are considered to be Level 1 vocabulary items in the Japanese Language Proficiency Exam (JLPT).

Sections (c) and (e) involve multiple-choice tasks, while section (d) utilizes a fill-in-the-blank format.

2.2 Guided narrative

One of the characteristics of the Advanced level and above is the ability to narrate and describe in major tense/aspect frames. Looking at a set of four photos depicting the September 11 attacks in New York City, the participants were asked to describe the events and express their thoughts. The September 11 topic was chosen for the following two reasons: (i) its high degree of familiarity to all, due to extensive media coverage, and (ii) the need for use of more than everyday, basic vocabulary (i.e., high-level registers) including *kango* (words of Chinese origin). (Note that the use of *kango* is one of the characteristics of the OPI Superior.)

2.3 Conversation

Parts of the data (i.e., excluding the role-play portions) from the OPI were used for our analyses. This interactive task involves daily and basic vocabulary, and does not necessarily require use of high-level registers. No results from this portion will be reported in this paper.

3. Participants

Fifteen English/Japanese simultaneous bilingual learners of Japanese at the University of Hawaii participated in the study. All were enrolled in sections of the first course of the 'bilingual tract' at the 3rd-year level, and the majority were in their first year at the university. They were placed in this course based on the results of a placement test and their performance on an oral interview in Japanese. We classified them into two groups based on their previous language-learning experience.

Group 1: Bilingual A (Bil. A)

Five learners who once a week had attended a *hoshuukoo* (special 'auxiliary' school) that followed the academic curriculum employed in schools in Japan; they had 'comparable' abilities in all four language skills.

Group 2: Bilingual B (Bil. B)

Ten learners who had been exposed to Japanese from birth, but had weaker reading and writing skills. This group was further divided into subgroup 1 (n=5) and subgroup 2 (n=5) based on amount and kind of prior exposure to Japanese. While the latter subgroup had been exposed to the language only in a family context, the former had not only had extensive exposure to Japanese but also more experience in a wider variety of contexts (e.g. a work place).

Note that seven out of the fifteen learners took only the written test and did not participate in the other two tests. Table 1 shows the proficiency levels of the eight remaining learners as determined by the OPI. Although we do not know the OPI level of the first seven participants, we can assume that they are at the same or similar proficiency level as the other bilinguals because, as mentioned previously, they were placed in the same course based on the results of a placement test as well as their oral performance in Japanese.

Table 1: Participants and their proficiency level

Groups	OPI Levels for those who participated	No. of learners who did not participate	Total
Bilingual A (Bil.A)	3 Advanced Plus	2	n=5
Bilingual B (Bil.B)	Subgroup 1	3	n=5
	Subgroup 2	3* Advanced**	2

* One of the three did not take the written test.

** One subject in Bilingual B, subgroup 2 was placed at the Advanced level, but with a note from the OPI tester that s/he is weak in vocabulary and on the borderline between Intermediate High and Advanced.

4. Results

4.1 Written test

This section reports the results of a written test. On the written test, students received one point for each item correctly answered. To confirm the validity of the test, the same written test was administered to ten native speakers of Japanese. Their response was perfect (i.e., 100%) for every test item except one in the grammar section where insufficient context was given. (This item involved an anaphoric use of a deictic demonstrative and was removed from the data before analysis.)

We began by comparing scores on the structural section and on the lexical and collocational section. We then compared the scores on the subsections within those two sections.

First, the overall performance on the structural and the lexical sections were analyzed across the three groups. Table 2 shows the performance percentage of each group on each section, and the same results are graphically illustrated in Figure 1 below.

Table 2: Accuracy rates on structural and lexical sections

	Structural section	Lexical section
Bil. A	94.67%	72.90%
Bil. B sub. 1	82.67%	50.97%
Bil. B sub. 2	52.50%	25.81%

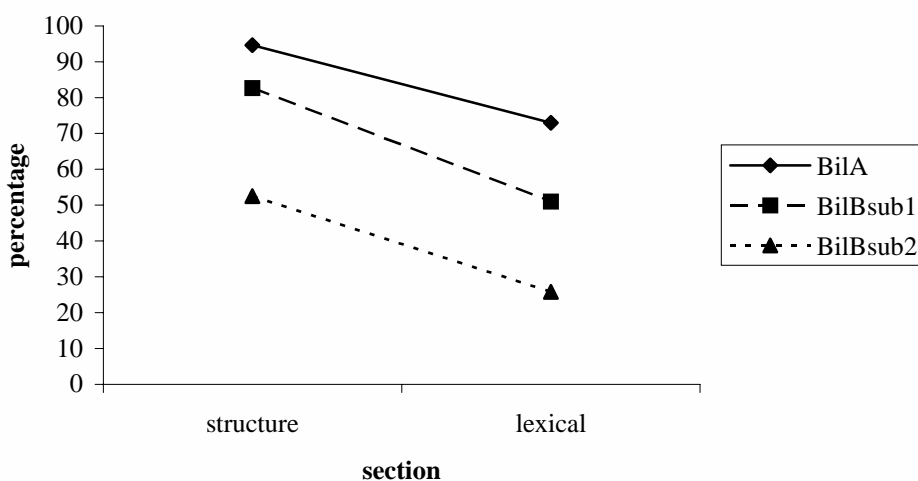


Figure 1: Performance on structural and lexical sections

As shown in Table 2 and Figure 1, Bilingual A outperformed Bilingual B subgroup 1, which also outperformed Bilingual B subgroup 2, on both sections. This may be due to the varying language experience observed across the three groups. Learners in the Bilingual A group were exposed to Japanese from birth and received content-based instruction in Japanese while attending a *hoshuukoo* (special auxiliary school). Not surprisingly, they are competent in the four language skills.

In contrast, learners in both Bilingual B groups were exposed to Japanese from birth, but only within a family context (subgroup 2), or contexts such as the work place or during a short-term summer visit to Japan (subgroup 1). These differing language experiences may have led to different amounts of exposure to the target language, and consequently affected the learners' structural and lexical knowledge. It could also be argued that their overall proficiency rather than different language experience might have directly influenced the test results since overall proficiency in Japanese as measured by OPI was slightly lower in the Bilingual B subgroup 2 (Advanced) than in Bilingual A and Bilingual B subgroup 1 (Advanced Plus). However, the proficiency level of learners in the three

subgroups was more or less comparable. What is interesting here is that despite their comparable proficiency level, performance on the lexical section differed across groups.

Table 3 and Figure 2 below present the results of learners' performance on each subsection.

Table 3: Performance on each subsection

	Structural		Lexical		
	Grammar	Connectives	Pairs of elements	Idiomatic expressions	Mimetics
Bil. A	95.00%	94.00%	96.00%	34.55%	92.00%
Bil. B sub. 1	88.00%	72.00%	68.00%	21.82%	66.00%
Bil. B sub. 2	58.75%	40.00%	50.00%	4.55%	25.00%

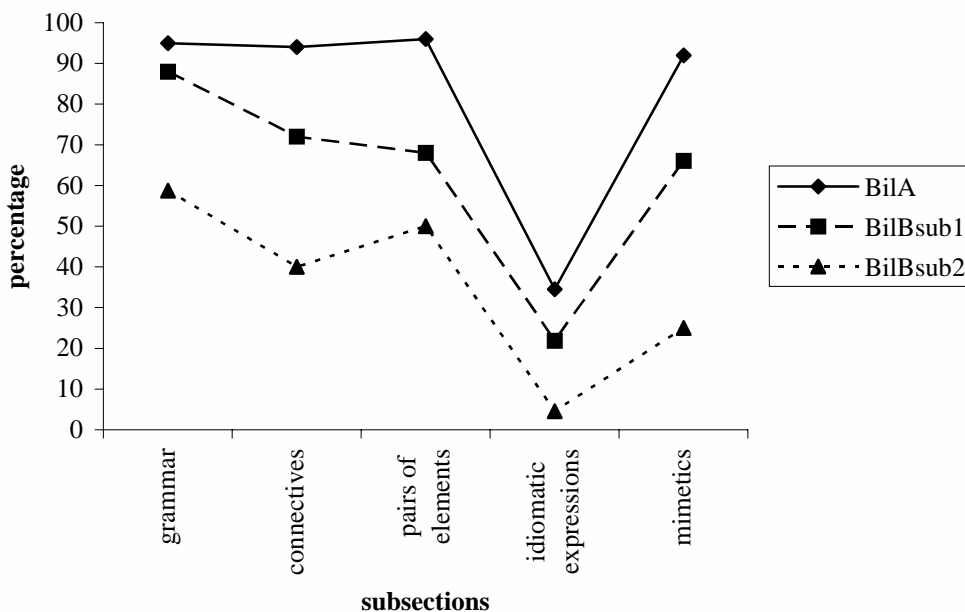


Figure 2: Performance on each subsection

As shown in the Figure 2, the overall pattern of learners' performance is very similar to the results illustrated in Figure 1. In particular, on every subsection Bilingual A outperformed Bilingual B subgroup 1, which outperformed Bilingual B subgroup 2. A key finding is that Bilingual A scored almost perfect on all subsections except the idiomatic expression section. The nature of the task may be a possible explanation for this. Recall that the learners were asked to fill in blanks for this section, whereas a multiple-choice format was used for the other sections. However, this cannot be the sole explanation, since the idiomatic expressions used in this section consist of nothing but basic vocabulary.

Table 4 below provides the mean score and the standard deviation of each group on each subsection. The standard deviation in Table 4 shows that although the overall performance pattern is very clear, as shown in Figures 1 and 2, there is variation in scores within certain groups, especially within Bilingual B subgroup 1. Among the five subsections of the test, the standard deviation for Bilingual B subgroup 1 was larger than for the other two groups on three subsections (connectives, pairs of elements, and mimetics). Learners in the Bilingual A group and the Bilingual B subgroup 2 are homogenous and there is therefore not much variation in terms of their language experience. In contrast, learners who belong to neither Bilingual A nor Bilingual B subgroup 2 are categorized into Bilingual B subgroup 1. Variation in scores was therefore expected within this subgroup, which has learners with varying language learning experience, and this expectation seems to have been borne out based on the standard deviation reported in Table 4.

Table 4: Mean score and standard deviation of each group

		Structural		Lexical		
		Grammar	Connectives	Pairs of elements	Idiomatic expressions	Mimetics
max		20.00	10.00	10.00	11.00	10.00
A	<i>M</i>	19.00	9.40	9.60	3.80	9.20
	<i>SD</i>	1.00	0.89	0.55	2.17	0.45
	<hr/>					
B-1	<i>M</i>	17.60	7.20	6.80	2.40	6.60
	<i>SD</i>	1.52	2.17	1.79	1.95	3.58
<hr/>						
B-2	<i>M</i>	11.75	4.00	5.00	0.50	2.50
	<i>SD</i>	2.22	1.15	0.82	1.00	0.58

Note. A=Bilingual A; B-1=Bilingual B subgroup 1; B-2=Bilingual B subgroup 2.

4.2 Guided narratives

Narrative data were also collected from three native speakers of Japanese (two graduate students and a business person, mixed genders) for comparison. The narratives were transcribed, coded and analyzed¹.

Data were analyzed for complexity and accuracy at both structural and lexical levels. For the purposes of this study, fluency was not included as an analytic criterion. We recognize unresolved language-specific issues in measuring fluency for Japanese due to its linguistic properties (e.g., its agglutinative rather than inflectional morphological features). For instance, the development of a consensus on an appropriate "unit" for analyzing produced utterances is still underway. In addition, particle ellipsis and zero pronouns are prominent features in spoken Japanese, which also complicates the process.

4.2.1 Structural complexity

Structural complexity was measured using three ratios: (i) number of clauses per utterance², (ii) number of dependent clauses divided by the total number of clauses, and (iii) number of passives

¹ For coding and analyses, we followed the method introduced by Tamaru & Yoshioka (1994). For the units of analysis (utterance, T-unit, clause etc.), some of the main criteria are as follows:

- (i) insertion : treat each inserted phrase or clause as a separate unit
- (ii) repetition: count the last repeated one
- (iii) filler : ignore fillers
- (iv) correction: count the last one (i.e., the "repaired" one)

² For the definition of an utterance, we adopted the following criteria from Crookes (1990) and Crookes & Rulon (1985):

An utterance unit can be found:

- (i) under one intonation contour,
- (ii) bounded by pauses, and
- (iii) constituting a single semantic unit

We especially relied on (i) and (ii) above when coding the speech production by non-heritage learners, from whom we collected data in addition to bilingual learners reported in this paper. In their production, pauses sometimes appear even where there is no syntactic or semantic defining unit (e.g., between a noun and a case particle). Hence we relied more on (iii) for such cases. However, all three criteria were appropriate for bilingual (heritage) learners' production, which shows little dysfluency.

divided by the total number of clauses. (See Wolfe-Quintero, Inagaki & Kim, 1998 for other potential measures.) Table 5 below presents the overall mean clause/ utterance (C/ U) ratio for the native speakers and the learners of Japanese.

Table 5: Utterances and clauses in native speaker and learner narratives

	Mean no. of clauses	Mean no. of utterances	C/U ratio
Native speakers	97.7	18.33	5.339
Learners*	40.2	14.47	2.778

*Note: Learners include (i) classroom learners, (ii) naturalistic learners, (iii) bilingual A, (iv) bilingual B (subgroup 1 and 2).

Another measure, the dependent clause ratio, was used to examine the degree of embedding/subordination. This was calculated by dividing the number of dependent clauses (DC) by the total number of clauses (C). Table 6 presents the mean dependent clause ratio (DC/C) for native speakers and learners (sorted by the type of learners). In our data, most of the learner groups frequently produced a particular clause pattern in Japanese with the final conjunctive particle *kedo* ('although'), which is often used at the utterance-final position in spoken Japanese. Taking this into consideration, another DC/C ratio was calculated by dividing the number of dependent clauses without *kedo* clauses by the total number of clauses.

Table 6: Mean dependent clause ratios for native speakers and learners

	Mean no. of all dependent clause/ C	DC/C ratio	Mean no. of dependent clauses without <i>kedo</i> 'although' / C	DC without <i>kedo</i> / C ratio
Natives	61.3 / 97.6	.627	50.0 / 97.6	.511
Bil. A	14.0 / 34.6	.403	12.0 / 34.6	.346
Bil. B sub.1	26.5 / 41.5	.638	16.0 / 41.5	.385
Bil. B sub.2	19.3 / 45.3	.426	17.0 / 45.3	.382

The bilingual groups A and B (both sub-group 1 and 2) uniformly exhibited a high DC/C ratio (with or without *kedo* clauses). Bilingual B subgroup 1, whose members grew up in a Japanese language environment and had exposure to the language in many socio-cultural contexts, approximate the most to the native speakers' production.

The production of passives is another measurement criterion for calculating grammatical complexity (see, e.g., Kameen, 1979, cited in Wolfe-Quintero et al, 1998) with a view to differentiating learners' developmental levels. Passive constructions are one of the essential linguistic devices for maintaining perspective within a sentence or paragraph. In a narrative construction, such devices become an important pivot feature for establishing coherence. Tanaka's studies (1997; 1999) indicate that as proficiency improves, learners of Japanese as a second language improve their use of passives for perspective maintenance. Table 7 below summarizes the use of passive constructions by the participants in our study.

As in the clausal complexity analysis, the two heritage groups (Bilingual A and B subgroup 1) exhibited a high passive ratio, although they produced fewer than the native speaker group. The other group (Bilingual B subgroup 2) had a much lower ratio in contrast to the other subgroup. However, it should be noted here that in terms of frequency and consistency their use of passives is not as developed as that of native speakers as a discourse-oriented device for perspective taking/maintenance. Taking the nature of the guided narrative into account (i.e., describe a "tragedy" from the perspective of those witnessed the incident), use of passives was called for in order to produce a coherent "story."

In summary, the dependent clause and passive ratios illustrated that with respect to structural complexity, the Bilingual A group and Bilingual B subgroup 1 were the most advanced learners, leaving Bilingual subgroup 2 behind.

Table 7: Passive ratios for native speakers and learners

	Mean no. of passives	Mean no. of clauses	Passive ratio
Natives	8.6	97.6	.088
BilA	2.6	34.6	.076
Bil. B. Sub.1	3.0	41.5	.072
Bil. B. Sub. 2	1.0	45.3	.022

4.2.2 Accuracy

In order to assess accuracy, we carried out an error analysis of the narrative data. The coding method used for this analysis is as follows. Two raters identified all errors for each narrative and classified them into four categories, namely (i) particles, (ii) structure, (iii) discourse, and (iv) vocabulary³. Vocabulary was broken down into three subcategories: use of English words, Japanese English words, and others.

Two measures were used to assess learner accuracy. One is the ratio of error-free clauses to total clauses, and the other is the error frequency rate. Table 8 below shows the results.

Table 8: Error free clause (EFC) ratio

	Mean No. of clauses	Mean no. of error free clauses (EFC)	EFC Ratio EFC/C
Bil. A	37	23	.62
Bil. B Subgroup 1	43	26	.60
Bil. B Subgroup 2	46	19	.41

The first two groups exhibited similar error-free clause ratios of around 0.60. Bilingual B subgroup 2 showed the least accuracy (0.41). Given that some clauses contained more than one error, a second measure was developed by dividing the number of errors by the number of clauses. In this measurement, a lower score indicates a greater degree of accuracy. Table 9 shows the results. Bilingual group A had the lowest score (highest accuracy) of 0.30. Bilingual B subgroup 2 had the highest score (lowest accuracy), making it the least accurate group of all.

Summing up so far, the error analysis of the narrative data showed Bilingual A group to be the most accurate, and Bilingual B subgroup 2 to be the least accurate. The results parallel what was found in the written test. In the next section, we discuss further the types of errors attested in the data, and explore the question of whether these types of errors can be used to distinguish learner groups for future profiling.

³ Three pairs of raters participated. Inter-rater reliability coefficients for each pair are:

Rator A- Rator B Error Identification : .936 (kappa = .663805), Error category: .965 (kappa = .95196)

Rator B- Rator C Error Identification : .909 (kappa = .387018), Error category: .931 (kappa = .906452)

Rator C-Rator A Error Identification : .926 (kappa = .55942), Error category: .1 (kappa = 1).

Table 9: Error frequency rate per clause

	Mean no. of clauses	Mean no. of errors	Error Freq. Rate E/C
Bil. A	37	11	0.30
Bil. B Subgroup 1	43	18.5	0.43
Bil. A Subgroup 2	46	31	0.67

4.2.3 Error types

The errors were classified into one of four types as already indicated—particles, structures, discourse (i.e., use of zero pronouns), and vocabulary. Vocabulary errors include substitution of English words, and use of 'Japanized' English words (English words pronounced according to Japanese phonological rules) which are not part of the Japanese language⁴. An error frequency rate (errors divided by total number of clauses) for each type was calculated. Table 10 shows the results.

Table 10: Error types in learner narratives

	Vocabulary	Particles	Structure	Discourse
Bil. A	0.150	0.122	0.027	0
Bil. B Subgroup 1	0.178	0.131	0.043	0.053
Bil. B Subgroup 2	0.326	0.123	0.181	0.035

It is assumed that one of the characteristics of heritage learners is frequent borrowing from the dominant language (e.g. see Campbell & Rosenthal, 2000). For the subjects in our data, English was the dominant language. Table 11 below shows the detailed analysis of vocabulary errors, sorted into three subcategories. The results for vocabulary errors in Table 11 show that borrowing was not prevalent among our subjects. All group showed very few errors of this type⁵.

Table 11 : Number of vocabulary errors in three sub-categories

	Mean no. of English words	Mean no. of Japanized Eng. words	Others	Mean total no. of errors
BilA	0.66	0.33	4.66	11
Bil. B Subgroup 1	1	0.5	8	18.5
Bil. B Subgroup 2	2.3	2.0	9.66	31

4.2.4 Vocabulary in the guided narrative

The vocabulary used in the guided narratives were first coded for syntactic category using JCHANGE ('Japanese Morphology Analyzer') developed by Hidetoshi Shirai, and then checked by a transcriber. Fifteen percent of the data was independently checked by a second rater⁶.

⁴ *Kojien* (a Japanese dictionary; Shinmura, 1998) was used to determine this.

⁵ Following the coding method mentioned above, the use of highly context-specific names, such as 'Twin Towers,' was not included in the error count in our analysis. The occurrence of such cases was very rare throughout the data we collected.

⁶ Inter-rater reliability for each of the same three pairs (as indicated earlier) was kappa .929, .928, and .917, respectively.

The lexical items used in the narratives were categorized based on the syntactic coding, after which types and tokens of the categories were counted. For the purposes of this study only items belonging to lexical categories (i.e., nouns, verbs, adjectives, adjectival nouns, and adverbs) were analyzed. The list of 1,500 lexical items in *Nihongo Noryoku Shiken Shutsudai Kijun* (the Japanese Proficiency Test: Test Content Specification, The Japan Foundation) was used to distinguish Basic and Non-Basic vocabulary. The vocabulary was also classified into (i) Basic, (ii) Non-Basic, and (iii) Other (i.e., English substitutions and Japanese words). Words in Non-Basic category were further classified as (a) *kango* (Chinese origin words), which is also labeled as sophisticated vocabulary, and (b) non-*kango*⁷.

Kango words were identified using *Nihon no Kango* (Sato, 1979) and the *Iwanami Shin-Kango Dictionary* (the Iwanami New-*Kango* Dictionary, 1994). The narrative data were examined with respect to the range of lexical variation and the level of lexical sophistication. Lexical variation was measured by dividing the number of lexical word types (LWT) by the total number of lexical words (LW). Table 12 presents the LWT/LT ratios for all the participants.

Table 12: Lexical type/token ratio for the narrative

	Mean LWT/LW ratio for narratives
Native Speakers	1.817
Bil. A	1.512
Bil. B Subgroup 1	1.640
Bil. B Subgroup 2	1.666

There was little difference in lexical variation, with all groups hovering around 1.5/1.6. These results show that the learners had sufficient lexical variation to avoid repeated 'recycling' of the same vocabulary items, which is an indication of their advanced skill in L2.

A measure of lexical sophistication was obtained by dividing (i) the numbers of Non-Basic word types and (ii) the number of sophisticated *kango*-word types individually by the total number of lexical word types. Table 13 presents the results.

Table 13: Lexical sophistication ratios for the guided narrative

Groups		Lexical sophistication ratio in the narrative
Native Speakers	Basic	53.0%
	Non-basic	47.0%
	[Sophisticated <i>Kango</i>: Other Non-basic]	
Bil. A	Basic	63.7%
	Non-basic	36.3%
	[Sophisticated <i>Kango</i>: Other Non-basic]	
Bil. B Subgroup 1	Basic	62.6%
	Non-basic	37.4%
	[Sophisticated <i>Kango</i>: Other Non-basic]	
Bil. B Subgroup 2	Basic	78.5%
	Non-basic	21.5%
	[Sophisticated <i>Kango</i>: Other Non-basic]	

Overall, learners did not use Non-Basic and sophisticated *kango* as much as native speakers, relying instead on Basic vocabulary. Figure 3 shows the proportions of Basic, other Non-Basic, and Non-Basic sophisticated *kango* vocabulary in the narrative data. Recall that due to the design of this

⁷ There are *kango* (Chinese origin words) vocabulary items which belong to the Basic category. For instance *jiko* ('accident') is a *kango* word, but is considered to be part of the Basic vocabulary set for Japanese according to the reference we used.

particular guided narrative task, the use of *kango* and sophisticated words was contextually required in the narratives. When we examine the distribution of these vocabulary types in the native speakers' narratives, the validity of the design is well supported.

Native speakers employed a roughly equal proportion of Basic and Non-Basic words (53% and 47%, respectively), and half of their Non-Basic words were composed of *kango*. Amongst the bilingual groups, relatively more Non-Basic words were produced by Bilingual B subgroup 1 (37.4%) and by Bilingual A (36.3%), compared to Bilingual B subgroup 2.

Between Bilingual A and Bilingual B subgroup 1, the use of Non-Basic sophisticated *kango* differentiated the two groups. Although the difference is still much smaller in comparison with native speakers, Bilingual A (13.0%) shows a relatively higher portion of sophisticated *kango* in their narratives than Bilingual B subgroup 1 (5.3%).

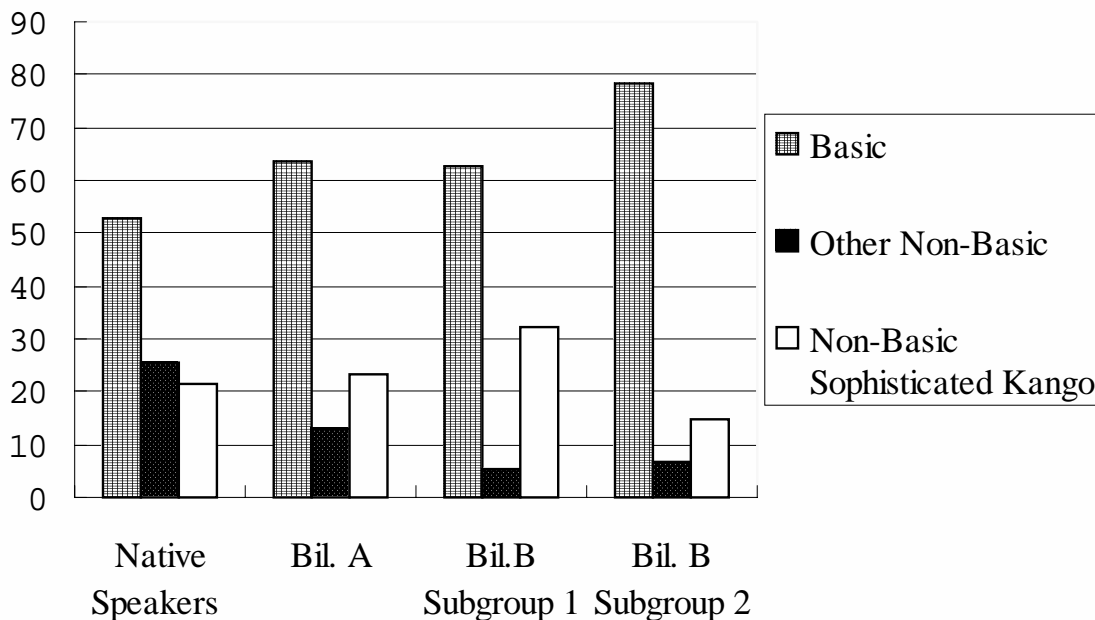


Figure 3: Lexical Sophistication Ratios for the Guided Narrative

5. Discussion and conclusion

This study has examined the linguistic profiles of heritage simultaneous bilingual learners of Japanese who were at a similar proficiency level despite different language learning experiences:

Bilingual A: Bilingual learners who attended a *hoshuukoo* (a special “auxiliary” school) and have comparable abilities in all four language skills.

Bilingual B: Bilingual learners with weaker reading and writing skills. Learners in subgroup 1 have experienced Japanese both inside and outside the family context, while learners in the subgroup 2 have experienced Japanese only within the family context.

The linguistic characteristics of these learners were investigated with reference to the two hypotheses adopted from Campbell and Rosenthal (2000), repeated below.

- Hypothesis 2: Heritage language learners have acquired 80% to 90% of the grammatical rules that govern word, phrase, sentence, and discourse production and recognition.
- Hypothesis 3: Heritage language learners have acquired extensive vocabularies; however, the semantic range of their vocabulary is limited to just a few sociocultural domains including hearth and home, neighborhood and, not infrequently, religious institutions.

Note that Campbell and Rosenthal are referring to 'average/typical heritage bilinguals'.

Hypothesis 2 was partially supported. That is, Bilingual A and Bilingual B subgroup 1 exhibited high scores on the structural sections of the written test (Tables 2 & 3) and relatively low error frequency rates for structures in the guided narrative (Table 9), supporting the hypothesis. On the other hand, the results from Bilingual B subgroup 2 (e.g. the low scores on the written test) do not seem to support it. One might wonder whether the last group of bilinguals are 'average/typical' heritage bilinguals. In terms of the pool of bilingual learners that are enrolled in the Japanese program at the University of Hawaii, they are at least 'average' or above average, given that there are quite a few simultaneous bilinguals who do not meet the criteria for admission to this third-year bilingual course and take a second-year course Japanese. Recall that all these participants were placed in this selective bilingual class at the 3rd-year level based on a placement test and an oral interview in Japanese.

Hypothesis 3 was supported. This is evident first in the low lexical type/token ratios for the narrative (Table 12), which suggest that the learners have sufficient lexical variation to avoid repeated use of the same vocabulary items. Second, as seen in the measure of lexical sophistication (Table 13 and Figure 3), and in the lack of advanced knowledge of items such as idiomatic expressions in the written test (Table 3 and Figure 2), the range of the learners' vocabulary is limited to the low register (i.e. basic vocabulary).

One thing worth noting here is that "borrowing from the majority language (i.e., English in this case)" was not frequent in our data, although it is assumed by Campbell and Rosenthal to be one of the characteristics of heritage bilingual learners.

In summary, our findings are more or less consistent with what Campbell and Rosenthal assume in their working hypotheses. That is, the average/typical heritage bilingual learners have very high grammatical knowledge, but exhibit weakness in their lexical knowledge, specifically lacking vocabulary in a high register. However, our findings differ from those of Campbell and Rosenthal in that the heritage bilingual learners are not homogeneous in their linguistic knowledge, contrary to what they seem to assume. This is evident in the unique linguistic profiles of our three groups of subjects, who were all simultaneous heritage bilingual learners with similar proficiency levels.

Larger scale studies are clearly called for before any conclusions can be drawn, given that this was a small-scale descriptive study. We hope that this study will trigger more research on this topic.

References

- Campbell, R. & Rosenthal, J. 2000. Heritage Languages. In J. W. Rosenthal (ed.), *Handbook of Undergraduate Second Language Education* (pp.165-84). Mahwah, NJ: Lawrence Erlbaum.
- Crookes, G., & Rulon, K. 1985. Incorporation of corrective feedback innative speaker/non-native speaker conversation. *Technical Report No. 3*. Center for Second Language Classroom Research, Social Science Research Institute, University of Hawaii.
- Crookes, G. 1990. The utterance, and other basic units for second language discourse analysis. *Applied Linguistics* 11, 183-99.
- Douglas, M. 2001. Individualized learning: a course for Japanese heritage learners at a college level, a talk given at the ATJ seminar (<http://www.japaneseteaching.org/ATJseminar/2001/douglas.html>).
- Kameen, P. T. 1979. Syntactic skill and ESL writing quality. In C. Yorio, K. Perkins, & J. Schachter (eds.), *On TESOL '79: The learner in focus* (pp. 343-64). Washington, D.C.: TESOL.
- McGinnis, S. 1996. Teaching Chinese to the Chinese: the development of an assessment and instructional modal. In J.E. Liskin-Gasparro (Ed.), *Patterns and Policies: The changing demographics of foreign language instruction* (pp. 107-121). Boston: Heinle & Heinle.
- Makino, S. 1991. ACTFL no gairaigo noryokukijun oyobi soreni motodoku kaiwanoryoku tesuto no rinen to mondai. *Sekai no Nihongo Kyooiku (Japanese-Language Education around the Globe)* 1, 15-32.

- Nihongo Gakuryoku Test Testing Committee. 1998. Ichimango-go Goi Bunruishuu (the Classification of 10,000 words). Tokyo: Senmon Kyooiku Shuppan.
- Sakamoto, T. 1992. On acquisition order: Japanese particles *wa* and *ga*. *Proceedings of the 4th Conference on Second Language Research in Japan*, 105-22.
- Sakoda, K. 1993. Hanashikotoba ni okeru ko-, so-, a no chuukangengo kenkyuu. *Nihongo Kyoiku (Journal of Japanese Language Teaching)* 81, 67-79.
- Sakoda, K. 1997. Nihongo gakushuusha ni okeru shijisi so-, a- no tsukaiwake ni kansuru kenkyuu. *Acquisition of Japanese as a Second Language* 1, 57-70.
- Sato, K. 1979. *Nihon no Kango*. Tokyo: Kadokawa Shoten.
- Shinmura, I. 1998. *Kojien (Dictionary)* 5 Edition. Tokyo: Iwanami Shoten.
- Sohn, S-O. 1995. The design of curriculum for teaching Korean as a heritage language. *Korean Language Education*, 1, 19-35.
- Sohn, S-O. 1996. Issues and concerns in teaching multi-level classes: Syllabus design for heritage and non-heritage learners. *Korean Language Education*, 2, 139-158.
- Tanaka, M. 1997. Shiten Voice, hukubun no shuutoku yoin. *Nihongo Kyoiku (Journal of Japanese Language Teaching)* 92, 107-18.
- Tanaka, M. 1999. Dainigengo shuutoku ni okeru nihongo voice no shuutoku junjo. *Proceedings of the 3rd PacSLRF*, Vol.1, 117-27.
- Tamaru, Y. & Yoshioka, K. 1994. Nihongo hatsuwa-shiryoo bunseki no tani o meguru mondai. *The Language Programs of the International University of Japan. Working Papers* 5, 84-100.
- The Japan Foundation. 1996. Nihongo Noryoku-shiken Shutsudai Kijun (Japanese language Proficiency Test: Test Content Specifications). Tokyo: Bonjinsha.
- Wolfe-Quintero, K., Inagaki, S. & Kim, H-Y. 1998. *Second Language Development in Writing: Measures of fluency, accuracy and complexity*. Honolulu, HI: University of Hawaii Press.
- Yagi, K. 1996. Shokyuu gakushuusha no sakubun ni mirareru nihongo no joshi no seiyoojunjo. *Sekai no Nihongo Kyooiku (Japanese-Language Education around the Globe)* 6, 65-81.
- Yamaguchi, A., & Takeda, A. 1994. *Iwanami Shin-Kango Dictionary*. Tokyo: Iwanami Shoten.

ISB4: Proceedings of the 4th International Symposium on Bilingualism

edited by James Cohen, Kara T. McAlister,
Kellie Rolstad, and Jeff MacSwan

Cascadilla Press Somerville, MA 2005

Copyright information

ISB4: Proceedings of the 4th International Symposium on Bilingualism
© 2005 Cascadilla Press, Somerville, MA. All rights reserved

ISBN 978-1-57473-210-8 CD-ROM
ISBN 978-1-57473-107-1 library binding (5-volume set)

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Press.

Ordering information

To order a copy of the proceedings, contact:

Cascadilla Press
P.O. Box 440355
Somerville, MA 02144, USA

phone: 1-617-776-2370
fax: 1-617-776-2271
sales@cascadilla.com
www.cascadilla.com

Web access and citation information

This paper is available from www.cascadilla.com/isb4.html and is identical
to the version published by Cascadilla Press on CD-ROM and in library binding.