

# ANALYSIS OF LEARNER'S PRODUCTION OF ADJECTIVES USING THE JAPANESE LANGUAGE LEARNER'S CORPUS C-JAS: THE CASE OF *TAKAI*

**Irena SRDANOVIĆ**  
University of Ljubljana  
irena.srdanovic@gmail.com

**Kumiko SAKODA**  
National Institute for Japanese Language and Linguistics  
sakodak@ninjal.ac.jp

## Abstract

In this paper, we explore learner production of adjectives using the Japanese language learner's corpus C-JAS (Corpus of Japanese As a Second language). Firstly, we describe the overall usage of adjectives in the corpus and discuss the distribution of the adjectives among learners including their correct and incorrect usages. Then, we take the frequently used adjective *takai* "high/tall/expensive" as an example and show how the learners' production of adjectives develops in terms of form, correct/incorrect usages, and lexico-semantic coverage.

**Keywords:** Japanese language adjectives; C-JAS corpus; second-language acquisition; language production; learner's errors

## Izveček

V tem članku raziskujemo, kako učenci japonskega jezika uporabljajo pridevnike v korpusnem gradivu C-JAS (Corpus of Japanese As a Second language, korpus japonščine kot tujega jezika). Najprej predstavimo celotno rabo pridevnikov v korpusu in opišemo distribucijo pridevnikov pri učencih posamezno vključno s pravilno in nepravilno rabo pridevnikov. Potem se osredotočimo na pogost pridevnik *takai* "visok/drag" in pokažemo, kako se raba pridevnikov razvija glede na obliko, pravilno in nepravilno rabo, ter leksikalno-semantično pokritost.

**Ključne besede:** pridevniki v japonskem jeziku; korpus C-JAS; učenje tujega jezika; jezikovna raba; napake učencev

## 1. Introduction

The development and analysis of second language learner's corpora is of an essential importance since it can show us the way language learners' use and understand the target language, as well as point out their language learning progress and potential obstacles during the learning process. Correct usages and mistakes, as well as the gaps between covered and not covered language usages can also provide valuable information for understanding second language learning acquisition. The objective of this paper is to explore learner's production of adjectives using the Japanese language learner's corpus C-JAS (Corpus of Japanese As a Second language).

Dixon (2004, p.12) states that an adjective class can be recognized for every language, but the criteria for distinguishing adjectives from nouns or verbs are subtle in some cases. There are numerous researches that state that there are no adjectives in some languages. Some approaches still treat adjectives as a sub-type of verbs. Here we take the view that the differences between the function and the structural properties of adjectives in different languages can indicate some possible learning burdens for second language learners. It is therefore important to explore language learners' usage of adjectives in greater detail.

This paper, first, describes the overall usage of adjectives in the corpus and offers discussion on the distribution of the adjectives among learners, with a particular focus on comparing their correct and incorrect usages. Then, we take the frequently used adjective *takai* "high/tall/expensive" as an example and show how learners' production of adjectives develops in terms of form, correct and incorrect usages, and lexico-semantic coverage.

## 2. C-JAS corpora and its characteristics

C-JAS stands for the Corpus of Japanese As a Second language. It is a learner corpus consisting of natural conversations of Japanese learners that has been collected as part of a longitudinal study for three years. The target students are three Korean native speakers, two males and one female (K1 ~ K3), and three Chinese native speakers, females (C1 ~ C3), who were attending the same Japanese school in Japan during their first year of Japanese language studies. The interactions between each of the students with the native Japanese speaker was recorded on tape once every 3-4 months, with each session consisting of 60- or 30-minute conversations recorded at 8 different periods. The total recorded time is 46 hours, which corresponds to script data of about 87 million words. The corpus is tagged for morphological information and for learners' errors and it is possible to search through the corpus systematically using the web page available at <https://ninjal-sakoda.sakura.ne.jp/c-jas/web/>.

In general, the following topics are covered during the different periods:

Period 1: Memories of my primary or secondary school teacher; Period 2: Looking back at the first year of study abroad; Period 3: My Japanese friend; Period 4: My student life; Period 5: About Japanese people; Period 6: How I spend holidays; Period 7: Food, clothing and housing in Japan; Period 8: Looking back at the last 3 years in Japan. Although the topics set in the spoken corpus were the same for all the students, the discussion developed differently with each student and covered various related subtopics.

The C-JAS corpus is an essential contribution to the study of second language acquisition of Japanese language and is especially useful in observing language learning through the viewpoint of its development.

### **3. Analysis of learners' production of adjectives in C-JAS**

This section analyzes learner production of adjectives in C-JAS, shows the developmental sequence of used adjectives, and classifies learner mistakes into different types.

#### **3.1 Overall production of adjectives**

The number of adjectives that appear in the corpus is 8459, out of which on average approximately 85% (7204) are produced correctly and 15% (1255) are produced as an error. Table 1 shows the usage of adjectives by six learners (Korean K1-3 and Chinese C1-3) within eight different periods – a) shows the overall correct and incorrect usage of adjectives, b) covers only incorrect usages, and c) shows only correct usage of adjectives. As can be expected, there are obviously differences among students in their production of adjectives, but above the mere observation of differences, there are some tendencies in the adjective usage that can be noticed. Chinese student C2 produces the most adjectives and he/she is followed by Korean students K3 and K2 and Chinese student C1 who produce adjectives at an average or above average rate (where an average is calculated based on the overall usage of the target students). This group also shows a relatively correct usage of adjectives, where Korean student K2 displays above average performance. On the other hand, the Chinese student C3 is characteristic due to the below average usage of adjectives, but with a high percentage of correct usage, while the Korean student K1 has by far the lowest usage of adjectives with the highest production of mistakes, which might be an indicator of the lowest level of Japanese language proficiency among the target students.

Further on, Figure 1 summarizes wrong usages of adjectives and shows how the error production is at the highest level in the learning phases at the beginning, then lowers down, being especially low in the periods 3 and 4 or 4 and 5, and then goes up

again. Only C3 shows a slightly different tendency with less mistakes in the beginning than later on and with a number of ups and downs during the learning curve.

**Table 1:** Adjectives -i in the C-JAS corpus per speaker and language learning period

| <b>a) Overall usage (correct and incorrect)</b> |           |           |           |           |           |           |           |           |              |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------------|
| <b>S</b>  | <b>P1</b> | <b>P2</b> | <b>P3</b> | <b>P4</b> | <b>P5</b> | <b>P6</b> | <b>P7</b> | <b>P8</b> | <b>TOTAL</b> |
| <b>K1</b>                                       | 10        | 36        | 134       | 121       | 101       | 126       | 112       | 113       | 753          |
| <b>K2</b>                                       | 230       | 181       | 188       | 199       | 191       | 171       | 181       | 168       | 1509         |
| <b>K3</b>                                       | 123       | 202       | 212       | 203       | 262       | 260       | 249       | 281       | 1792         |
| <b>C1</b>                                       | 69        | /         | 158       | 298       | 183       | 226       | 215       | 249       | 1398         |
| <b>C2</b>                                       | 249       | 221       | 210       | 220       | 237       | 235       | 253       | 226       | 1851         |
| <b>C3</b>                                       | 137       | 190       | 131       | 147       | 103       | 133       | 164       | 151       | 1156         |
|   |           |           |           |           |           |           |           |           | <b>8459</b>  |
| <b>b) Only wrong usage</b>                      |           |           |           |           |           |           |           |           |              |
| <b>S</b>  | <b>P1</b> | <b>P2</b> | <b>P3</b> | <b>P4</b> | <b>P5</b> | <b>P6</b> | <b>P7</b> | <b>P8</b> | <b>TOTAL</b> |
| <b>K1</b>                                       | 3         | 11        | 47        | 25        | 15        | 25        | 20        | 25        | 171          |
| <b>K2</b>                                       | 61        | 29        | 15        | 7         | 8         | 14        | 18        | 14        | 166          |
| <b>K3</b>                                       | 29        | 54        | 35        | 30        | 42        | 21        | 31        | 27        | 269          |
| <b>C1</b>                                       | 12        | /         | 23        | 41        | 32        | 40        | 27        | 36        | 211          |
| <b>C2</b>                                       | 87        | 40        | 22        | 26        | 35        | 18        | 19        | 26        | 273          |
| <b>C3</b>                                       | 13        | 35        | 22        | 19        | 14        | 25        | 15        | 22        | 165          |
|   |           |           |           |           |           |           |           |           | <b>1255</b>  |
| <b>c) Only correct usage</b>                    |           |           |           |           |           |           |           |           |              |
| <b>S</b>  | <b>P1</b> | <b>P2</b> | <b>P3</b> | <b>P4</b> | <b>P5</b> | <b>P6</b> | <b>P7</b> | <b>P8</b> | <b>TOTAL</b> |
| <b>K1</b>                                       | 7         | 25        | 87        | 96        | 86        | 101       | 92        | 88        | <b>582</b>   |
| <b>K2</b>                                       | 169       | 152       | 173       | 192       | 183       | 157       | 163       | 154       | <b>1343</b>  |
| <b>K3</b>                                       | 94        | 148       | 177       | 173       | 220       | 239       | 218       | 254       | <b>1523</b>  |
| <b>C1</b>                                       | 57        | /         | 135       | 257       | 151       | 186       | 188       | 213       | <b>1187</b>  |
| <b>C2</b>                                       | 162       | 181       | 188       | 194       | 202       | 217       | 234       | 200       | <b>1578</b>  |
| <b>C3</b>                                       | 124       | 155       | 109       | 128       | 89        | 108       | 149       | 129       | <b>991</b>   |
|   |           |           |           |           |           |           |           |           | <b>7204</b>  |

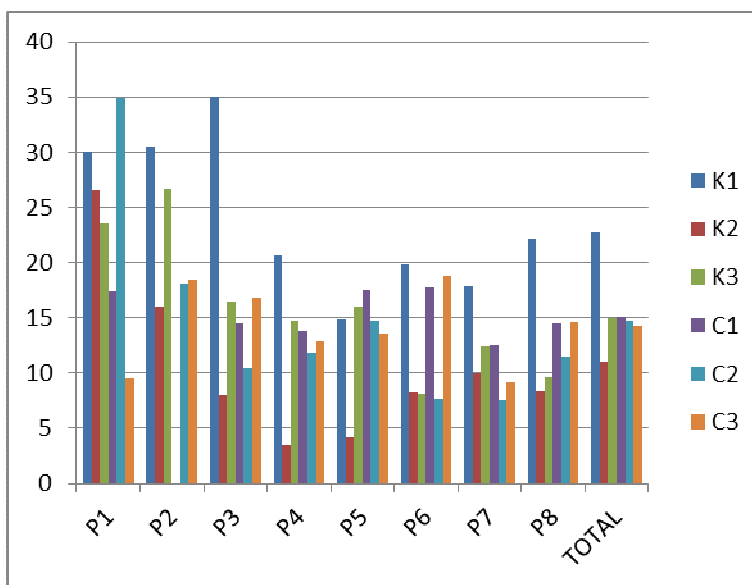


Figure 1: Only wrong usage of adjectives (%)

### 3.2 Production and analysis of mistakes

This section describes the types of mistakes that appear in the production of adjectives and provides a few examples of incorrect production and possible explanations in the case of the adjective *takai*.

#### 3.2.1 Overall production

Table 2 shows a) the overall usage of correct and incorrect forms and b) incorrect usage of the adjective *takai* by six learners within different periods. Similar to the overall usage of adjectives described above in section 3.1, the differences in performance can be observed and summarized as:

- rare usage of the adjective *takai* and no mistakes (K1),
- quite a productive usage of the adjective *takai* and almost no error (actually, one mistake out of 39 usages of the adjective) (K2),
- frequent usage of the adjective *takai* and a high rate of produced mistakes (K3),
- relatively frequent usage of the adjective *takai* and a few errors produced (C1,2,3).

Korean student 1 performs with the lowest usage of the adjective *takai*, which is in line with his overall adjective usage. Korean student K2 displays above average performance with a largely productive usage of the adjective *takai* and almost no errors, which is the case for his overall adjective usage as well.

Looking into the percentage of errors out of the overall *takai* usage per period reveals that the number of errors is higher in the beginning and lowers down later on,

but there is a tendency to produce a high number of errors in the middle periods 4 and 5. The number of errors lowers down with period 6 and stays low until the last periods covered by the data.

**Table 2:** Language learners' usage of *takai*

| a) Overall usage (correct and incorrect) |             |             |             |             |             |            |            |            |             |              |
|--|-------------|-------------|-------------|-------------|-------------|------------|------------|------------|-------------|--------------|
| S  | P1          | P2          | P3          | P4          | P5          | P6         | P7         | P8         | Total       |              |
| K1                                       | 0           | 0           | 1           | 0           | 1           | 2          | 0          | 2          | 6           |              |
| K2                                       | 8           | 17          | 2           | 1           | 5           | 1          | 2          | 3          | 39          |              |
| K3                                       | 4           | 4           | 1           | 2           | 12          | 0          | 5          | 2          | 30          |              |
| C1                                       | 1           | 0           | 3           | 1           | 0           | 3          | 11         | 6          | 25          |              |
| C2                                       | 4           | 3           | 0           | 0           | 3           | 5          | 6          | 4          | 25          |              |
| C3                                       | 4           | 0           | 2           | 4           | 4           | 2          | 8          | 3          | 27          |              |
| <b>Total</b>                             | <b>21</b>   | <b>24</b>   | <b>9</b>    | <b>8</b>    | <b>25</b>   | <b>13</b>  | <b>32</b>  | <b>20</b>  | <b>152</b>  |              |
| b) Only wrong usage                      |             |             |             |             |             |            |            |            |             |              |
| S  | P1          | P2          | P3          | P4          | P5          | P6         | P7         | P8         | Total       | %of mistakes |
| K1                                       | 0           | 0           | 0           | 0           | 0           | 0          | 0          | 0          | 0           | 0            |
| K2                                       | 1           | 0           | 0           | 0           | 0           | 0          | 0          | 0          | 1           | 2.6          |
| K3                                       | 0           | 3           | 0           | 2           | 3           | 0          | 0          | 0          | 8           | 26.7         |
| C1                                       | 0           | 0           | 0           | 0           | 0           | 0          | 2          | 1          | 3           | 12.0         |
| C2                                       | 2           | 1           | 0           | 0           | 1           | 0          | 0          | 0          | 4           | 16.0         |
| C3                                       | 0           | 0           | 1           | 0           | 1           | 1          | 0          | 0          | 3           | 11.1         |
| <b>Total</b>                             | <b>3</b>    | <b>4</b>    | <b>1</b>    | <b>2</b>    | <b>5</b>    | <b>1</b>   | <b>2</b>   | <b>1</b>   | <b>19</b>   | <b>12.5</b>  |
| <b>% of mistakes</b>                     | <b>14.3</b> | <b>16.7</b> | <b>11.1</b> | <b>25.0</b> | <b>20.0</b> | <b>7.7</b> | <b>6.3</b> | <b>5.0</b> | <b>12.5</b> |              |

Table 3 shows the types of mistakes per period and student in the production of the adjective *takai*. Mistakes related to lexical selection, either missing a collocate or choosing a wrong one, are quite often in the case of *takai*. Grammar mistakes are the next type of mistake that often appears. The number of mistakes is still relatively small to be able to draw conclusions about tendencies per separate periods or between different students. With some caution it can be noticed that grammar mistakes are spread all over the periods while lexical mistakes tend to appear more in earlier phases, though they are present later on as well. However, this needs to be clarified using larger data and expanded to cover a greater variety of adjectives. Some students are prone to make some type of mistakes more than others; in the case of adjectives, for

example, the Chinese student C2 does not display grammar mistakes, while making mostly lexical ones with some pronunciation mistakes as well, when producing the adjective *takai*.

**Table 3:** Types of language learners' mistakes per period and per student (the case of *takai*)

| Mistakes per period | Lexical (missing collocate) | Grammar  | Pronunciation | Lexical (wrong collocation) | Discourse | General knowledge | Total     |
|---------------------|-----------------------------|----------|---------------|-----------------------------|-----------|-------------------|-----------|
| 1                   | 2                           |          |               | 1                           |           |                   | 3         |
| 2                   | 1                           | 1        |               | 1                           | 2         |                   | 5         |
| 3                   |                             | 1        |               | 1                           |           |                   | 2         |
| 4                   |                             | 1        | 1             |                             |           |                   | 2         |
| 5                   | 2                           | 1        | 3             |                             |           |                   | 6         |
| 6                   |                             | 1        |               |                             |           |                   | 1         |
| 7                   | 1                           | 1        |               | 1                           |           |                   | 3         |
| 8                   |                             |          |               |                             |           | 1                 | 1         |
| <b>Total</b>        | <b>6</b>                    | <b>6</b> | <b>4</b>      | <b>4</b>                    | <b>2</b>  | <b>1</b>          | <b>23</b> |

| Mistakes per student | Lexical (missing collocate) | Grammar  | Pronunciation | Lexical (wrong collocation) | Discourse | General knowledge | Total     |
|----------------------|-----------------------------|----------|---------------|-----------------------------|-----------|-------------------|-----------|
| K3                   | 2                           | 3        | 1             | 1                           | 2         |                   | 9         |
| C2                   | 3                           |          | 2             | 1                           |           |                   | 6         |
| C1                   | 1                           | 1        |               | 1                           |           | 1                 | 4         |
| C3                   |                             | 2        |               |                             |           |                   | 2         |
| C3                   |                             |          |               |                             |           |                   | 1         |
| K2                   |                             |          | 1             | 1                           |           |                   | 1         |
| <b>Total</b>         | <b>6</b>                    | <b>6</b> | <b>4</b>      | <b>4</b>                    | <b>2</b>  | <b>1</b>          | <b>23</b> |

### 3.2.2 Usage examples

Example 1 is produced by the Chinese learner C2 in period 2: When asked to describe a person, the learner uses adjectives *takai* and *ookii* "big", both in the correct conjunctive form (*renyou-kei*). The grammar is correct but the lexical mistake (missing collocate) can be noticed in line 87L (L stands for learner, and N for native): in Japanese 背 *se* "back" needs to be used to specify the attribute (*se ga takai hito* "a tall person"). The following native speaker's question in the conversation *nani ga* (line 88N) "What is?" also indicates the lack of that part of the collocate. The next line (line 89L) shows a lexical mistake (wrong collocate) where *kao* "face" is produced instead of *se*. The reason for this mistake needs to be further explored, but one possible

interpretation is that the newly learnt words *kao* and *se* were mixed up. Here, the native speaker jumps in and suggests *se ga* “the back is” (in the corpus indicated with brackets), which helps the learner to produce the intended correct expression.

Example 1: Chinese learner C2, period 2, lexical mistake (missing collocate, wrong collocate) for *sei ga takai*

86 N ふーん, どんな人だった?

Fuun, donna hito datta?

“What kind of person he was?”

87 L うーん, 高くて, 大きくて

**Uun, takakute, ookikute**

**“Well, a high, big”**

88 N 何が

Nani ga

“What was high, big?”

89 L 顔が<背が?>せい[背], 背がたか, 高い, あとは, 優しいの顔してるけどね

**Kao ga <sei ga?> sei, sei ga taka, takai, ato wa, yasashii no kao shite ru kedo ne**

**“The face <the back?> the back, the back is high** (meaning “he is tall”), and,

he has a very friendly face” [Note by translator: The learner makes a mistake and

does not specify “what is high”, which is actually required in Japanese for the

adjective *takai* “tall/ high/expensive” where “a tall person” is literary formed as

“the back is high + person”.]

90 N ほんと

Honto

“Really”

91 L うん

Un

“Yea”

While in English and some other languages the expression *a tall person* is formed by combining an adjective and a noun, in the case of Japanese, 背 *se* “back” is used to specify the attribute (*se ga takai hito*). The need to specify the attribute comes from the semantic range of the adjective *takai* (see also Section 3.4) covering not only high and tall things but also those expensive and high in quantity or quality. Since *takai hito* can also be referred to, for example, using the expressions *kyuuryou ga takai hito* “a person with a high salary”, *komyunikeeshon nouryoku ga takai hito* “a person with a high level of communication ability” the attributive role of *takai*. As discussed in detail within Srdanović (2013), the form of this combination is unpredictable<sup>1</sup> by Japanese language learners who are native speakers of English, but not exclusively, and as such directly related to language burden and learner’s possible mistakes. Therefore, such kinds of unpredictable combinations need to be paid special attention to in the teaching/learning

<sup>1</sup> The phenomenon of predictability and unpredictability of collocations is introduced in Nation 2001.



process. This is also confirmed by the error analysis. Such kind of mistakes tends to appear with some learners. In C-JAS data and in the case of *takai*, this kind of mistake reappeared a number of times especially in the case of the Chinese learner C2.

Example 2 is an obvious example of wrong usage of the adjective *takai* in connection with one's age. Although in Japanese it is possible to say 高齢者 *koureisha* "an old person" using the character 高い *takai/kou*, *takai* is not used as a predicate or noun attribute to refer to the noun 歳 *toshi* "year(s)". Similarly the noun is not used in combination with the adjective 多い *ooi* "a lot of". The reason that these two language mistakes happened can be found in the influence of the native language of the speaker, Korean, where adjectives with the meaning *ooi* and *takai* appear as predicates of the noun denoting someone's age.

Example 2: Korean student K2, period 1, lexical mistake (wrong collocate)

226 L はい、私は友達が、多いおお多くてー、ぺんきよ[勉強]がー、よくしー、ん、ぺんきよ[勉強]するのがーよくできま、ませんです、できません、それでも、私は、もう父と、父がー、今歳が、ほんとに、んー、高いです、歳が、多いい{方言}です、ん、まだー、ん、今もう、66歳?

*Hai, watashi wa tomodachi ga, ooi oo ookutee, penkyo[benkyou]gaa,yokushii, n, penkyou [benkyou] suru no gaa yoku dekima, masen desu, dekimasen, soredemo, watashi wa, mou chichi to, chichi gaa, ima sai ga, hontoni, nn, takai desu, sai ga, ooii {hougen} desu, n, madaa, n, ima mou, 66 sai?*

"Yes, I have a lot, a lots of friends, I don't, I cannot study enough, and my father, he is old [Note by translator: the learner makes a mistake and says lit. \*he has high years], is old [Note by translator: the learner tries to correct and says lit. \*he has lots of years {nonstandard, used in some dialects}]. Now he \*still, he is already 66 years."

227 N んーんー、あそう

*nn nn, asou*

"I see"

Example 3 shows the usage of *takai* and the particle *kara* "because", where a grammar mistake appears since *da* is used after the plain form of the adjective. This type of mistake appears due to the transfer of learnt grammar rules for nouns and na-adjectives into the i-adjective. Besides this mistake, the form of the verb and the particle usage is not appropriate as well. The suggested corrections in the corpus are *moraeru kyuryou ga takai kara* "it is because the salary they can get is high" or *moraeru kyuryou ga sara ni fueru kara* "it is because the salary they can get gets even higher".

Example 4 Korean student 3, period 5, grammar mistake

106 L はい、仕事をもったい[もちたい], もちたいです

*Hai, shigoto wo mottai [mochitai], mochitai desu*

"Yes, I want to get a job. [Note by translator: the learner makes a mistake in the verb form *mottai*, which is corrected into *mochitai*]"

- 107 N でも、別に大学に行かなくても仕事  
Demo, betsu ni daigaku ni ikanakutemo shigoto  
“But, a job without going to a university?”
- 108 L うんそうですよでも、専門的な仕事はないと思います  
Un sou desu yo demo, senmontekina shigoto wa nai to omoimasu  
“Yes, that’s right, but I think there is no specialized job”
- 109 N 高校だけでは?  
Koukou dake dewa?  
“Only with a high school?”
- 110 L はい高校だけでは、*〈うん〉* 専門的な、*んー* たつとえば [例えば]、*〈はい〉*  
*〉* 高校そちゅぎよ [卒業] したら、*〈はい〉* もらうきゅうりよ [給料] は少ないん  
だけどー、*〈うん〉* たいがく [大学] そちゅぎよ [卒業したら] したら、もらう  
給料は、**もっと高いだ** からー、*〈うんうんうん〉* 4年間行ったー、結果ーが  
あるんじゃないかなーと思います  
Hai koukou dake de waa, *〈un〉* senmontekina, nn tattoeba [tatoeba], *〈hai〉*  
koukou sochugyo [sotsugyou] shitara, *〈hai〉* morau kyuuryo [kyuuryou] wa  
sukunai n dakedoo, *〈un〉* taigaku [daigaku] sochugyo [sotsugyou shitara]  
shittara, morau kyuuryou wa, **motto takai da karaa**, *〈ununun〉* yon  
nenkan okonattaa, kekkaa ga aru n janai kanaa to omoimasu.  
“Yes, only with a high school, a specialized one, for example when one  
graduates from a high school, the salary he gets is lower, and when one  
graduates from a university, the salary he gets is higher, so I think there is  
a merit in studying four years” [Note by translator: the learner makes some  
mistakes in pronunciation of *tattoeba* instead of *tatoeba*, *sochugyo* instead  
of *sotsugyou*, *shittara* instead of *shitara* etc. The teacher often confirms  
that she follows the conversation by backchannel markers *hai* and *un*]
- 111 N あーなるほどね  
Aa naru hodo ne  
“Oh, I see”

### 3.3 Developmental sequence of adjectives

In this section we explore the developmental sequence of adjectives through the case of the frequent adjective *takai* “high, tall, expensive”. We take the Chinese student C2, who showed an average performance on adjective usage and production of *takai*, as an illustrative example.

Sakoda et al. (2012) explores the developmental sequence of verbs by making a C-JAS survey of the verbs *omou* “think” and *taberu* “eat”. The study revealed that there are phenomena both similar to and different from the patterns of first language acquisition. Deriving a new correct form (e.g. *omou kara* “because (I) think” based on the plain form acquired before (e.g. *omou* “think”) is common to both first and second language acquisition, whereas plain verb forms characterize Japanese children’s verbs in their first appearances which is in contrast to the polite forms seen in learners’ verbs.

Additionally, in the development of verb acquisition, a unique learners' interlanguage form “*plain verb + desu*” (e.g. *omotta desu* “thought COP (POLITE)”) appears, which seems to be a transitional form.

Table 4 shows the development sequence of the adjective *takai* in the case of the Chinese student C2. In the first period, simple plain forms of the adjective in its predicative role can be observed. It can be noticed that the learner is already in the beginner phase and fluent in colloquial forms *un*, *desho* “yes, probably”, and particle *mo* “also”. The following period shows that *takai* is appropriately produced in its continuous form (*renyou-kei*). From the later periods, 5-8, we can notice a slight transition from the usage of *takai* in plain forms and short simple sentences into usage of *takai* inside more complex sentence structure and various combinations of *takai* with other elements in a sentence and with various functions. In period 5, the attributive role (*rentai-kei*) and the adverbial role (*renyou-kei*) appear: *takai gakureki* “a high educational history”, *takaku mottara ikenai* “it shouldn't get that high”. The complex sentence structure *N1 wa N2 ga takai* “N1 has a high N2” and the conditional form with *-tara* “if/when” also appears in this period and is repeated later on. The following period brings conjunctive usages with *-kara* “because/from”, *-shi* “and/as well as”, while the final periods 7 and 8 show clause-final modality forms such as *no do\* ka* (*no ..to ka*) “[marks the preceding utterance as a reason/explanation and hints at additional reasons]”, *naa to omotta* “I thought that”, and *n de* “[marks the preceding utterance as a reason/explanation]”.

**Table 4:** Development sequence of *takai* in the case of the Chinese student C2

| Per. | Explan.         | Examples  |   |
|------|-----------------|---|---|
| 1    | Form            | Un, <i>takai takai</i>                              | N mo <i>takai</i>   |
|      | (Wider) context | (sotsugyou shitara, kyuuryou ga)                    | senmon gakkou mo <i>takai</i><br>(sotsugyou shitara, kyuuryou ga) |
|      | Form            | *N, <i>takai desho</i>                              | * <i>takai</i>  |
|      | (Wider) context | otousan, oniisan, minna, <i>takai desho</i>         | Hontou ni <i>takai</i> (otousan, oniisan, minna)                  |
|      | Correction      | → otousan, oniisan, minna, se ga <i>takai desho</i> | → Hontou ni se ga <i>takai</i> (otousan, oniisan, minna)          |
| 2    | Form            | *Un, <i>takakute</i> ,                              | *kao ga <N:se ga> se ga <i>taka, takai</i>                        |
|      | (Wider) context | *Un, <i>takakute</i> , ookikute (hito)              | (hito)  |
|      | Correction      | → Un, se ga <i>takakute</i> , karada ga ookikute    |   |
|      | Form            | N <i>takai desho</i>                                |   |
|      | (Wider) context | *seikatsudai <i>takai desho</i>                     |   |
|      | Correction      | → seikatsuhi <i>takai desho</i>                     |   |

| Per. | Explan.                      | Examples  |   |
|------|------------------------------|---|---|
| 5    | Form                         | *Ano, N wa, ne, <i>takkai</i> N motte tara, yoku nai            | *N <i>takkaku</i> mottara ikenai                                    |
|      | (Wider) context              | *Ano, josei wa, ne, <i>takkai</i> gakureki motte tara, yoku nai | *Gakureki <i>takkaku</i> mottara ikenai                             |
|      | Correction                   | → Ano, josei wa, ne, <i>takai</i> gakureki motte tara, yoku nai | → Gakureki <i>takaku</i> mottara ikenai/Gakureki takakattara ikenai |
|      | Form                         | * <i>Takkai</i> N wa...   | Se ga <i>takakute</i> ,   |
|      | (Wider) context              | * <i>Takkai</i> hito wa ...                                     |   |
|      | Correction                   | → <i>Takai</i> hito wa ...                                      |   |
| 6    | Form                         | Se ga <i>takai</i> ?  |   |
|      | (Wider) context              | (repeating native speaker's expression)                         | (repeating native speaker's expression)                             |
|      | Form                         | Ato, N wa, N ga <i>takai</i> desho,                             | , N mo <i>takai</i> shi,  |
|      | (Wider) context              | Ato, Nihon wa, seikatsuhi ga <i>takai</i> desho,                | , hikoukidai mo <i>takai</i> shi,                                   |
| Form | , nandemo <i>takai</i> kara, |   |   |
| 7    | Form                         | *N mo ... konna ni <i>takai</i> no do ka                        | <i>Takai</i> (3x)   |
|      | (Wider) context              | *Heya mo konna ni semai de, konna ni <i>takai</i> no do ka      | (ie, an answer to a question by native speaker; nikuman 2x)         |
|      | Correction                   | → Heya mo konna ni semakute, konna ni <i>takai</i> no to ka     |   |
|      | Form                         | *Shikamo <i>takai</i> da shi                                    | Dakara, N nandemo <i>takai</i> naa to omotta                        |
|      | (Wider) context              | (oniku)   | Dakara, Nihon nandemo <i>takai</i> naa to omotta                    |
|      | Correction                   | → Shikamo <i>takai</i> shi                                      |   |
| 8    | Form                         | Ano, <i>takai</i> N wo V <sub>past</sub> n de,                  | ... nanka <i>takai</i> N  |
|      | (Wider) context              | Ano, <i>takai</i> kutsu wo *haita n de,                         | Ano, kutsu no ichiban, ushiro no hou, nanka <i>takai</i> bubun      |
|      | Correction                   | → Ano, <i>takai</i> kutsu wo haite ta n de,                     |   |
|      | Form                         | *Ano, ushiro <i>takai</i> N                                     | ... kara, V no N ga N ga <i>takai</i>                               |
|      | (Wider) context              | *Ano, ushiro <i>takai</i> bubun ga ochite shimatte              | ... kara, deru *no hou ga kyuuryou ga <i>takai</i>                  |
|      | Correction                   | → Ano, ushiro no <i>takai</i> bubun ga totte shimatte           | → ... kara, deru hou ga kyuuryou ga <i>takai</i>                    |

It is interesting to observe that although some grammar items are used correctly and seem to be already acquired, after some time they are reused as a transitional learner's interlanguage form. For example, although *takai shi* "as well as high/tall/expensive" is correctly used in period 6, the transitional interlanguage form is created in period 7 *takai \*da shi* "as well as high/tall/expensive [with an unnecessary copula *da*]", where the usage of *da* seems to be influenced by noun and adjective *-na* forms.

In addition, the lexical mistake mentioned in section 3.2.2 seems to re-appear and it is interesting to bring its progress to attention. The omission of *se* in *se ga takai* already appears in period 1, and then repeats in the following period. During the conversation in period 6, the native speaker uses the expression *se ga takai*, and then the learner repeats it twice, which can also be noted as a good practice for overcoming learner's errors.

Finally, it is interesting to notice the practical role of the adjective *takai*. As a basic adjective it is used by learners to descriptively denote some notions for which lexical representations are unfamiliar to them. For example, in period 8, *takai* is used to denote high heels: *Ano, kutsu no ichiban, ushiro no hou, nanka takai bubun* "the high part that is the most behind on a shoe".

### 3.4 Lexical domains used by language learners

This section describes the lexical domains of the adjective *takai* that are covered by language learners. Nouns that are modified by the adjective are grouped based on their meaning and are observed in comparison to the lexical domains of the same adjective used by native speakers.

As described in Srdanovic (2013), the adjective *takai* in combination with the modified nouns covers three large lexical domains: positional relations (*takai yama* "a high mountain", *takai tokoro* "a high place", *takai kabe* "a high wall/a high barrier"), quantitative relations (*takai kakuritsu* "a high probability", *takai wariai* "a high percentage") and superior/inferior relations (*takai hyouka* "high evaluation", *takai nouryoku* "a high ability"). Each of the domains is further divided into subgroups of meanings and sorted from more concrete to more abstract, from more natural to more artificial, from ordinary to metaphorical meanings. Figure 2 shows a lexical map of the most frequent domains used with the adjective *takai*. For example, the positional relations domain starts with the concrete nouns observable in nature (*yama* "mountain", *ki* "tree"), further covers those created by humans (*biru* "building", *kabe* "wall") and finally introduces metaphorical meanings (*kabe* "barrier", *haadoru* "barrier")

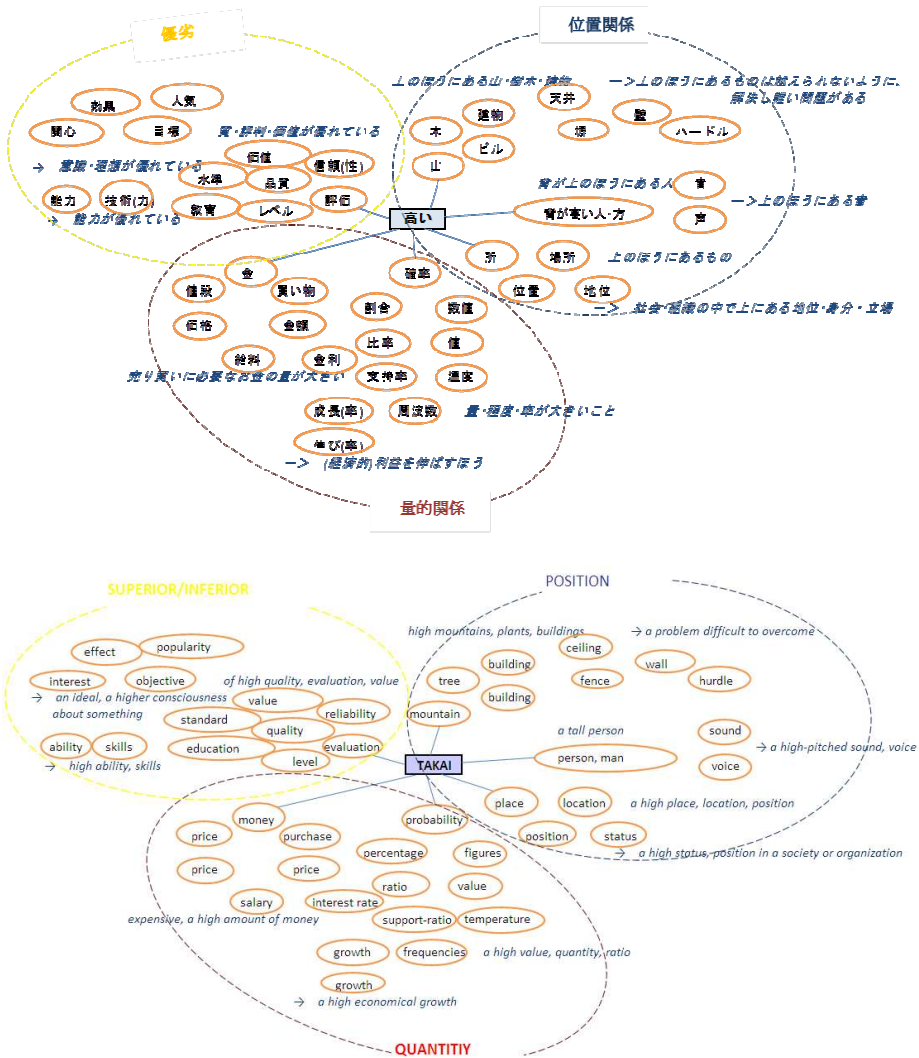


Figure 2: Lexical map of the adjective *takai* and the modified nouns, Japanese and English version (Figure reproduced from Srdanovic,2013)

Analysis of learners' production of the adjective *takai* shows that learners overwhelmingly use the first two semantic domains: positional and quantitative, but do not use the more abstract superior/inferior relation. Only one Korean learner (K2) out of 6 uses the third superior/inferior domain. The same learner is described as "being above average in his performance" based on the quantitative and qualitative analysis of the overall correct and incorrect adjective production (see Section 3.1). Observed wider usage of lexical domains might be related to the learner's richer vocabulary.

## 4. Conclusion

In this paper we present learners' production of Japanese language i-adjectives using the Japanese language learner's corpus C-JAS. First we showed how adjectives are produced by learners and pointed out differences in learners' language performances. Analysis of errors showed that error production is at the highest level in the learning phases in the beginning, then lowers down, is especially low in the middle periods, and then again grows a bit.

Interestingly, the overall usage of adjectives reveals a big similarity with the differences in learners' performances observed for the adjective *takai*. The analysis of the production of *takai* and overall adjectives provide good indications about each learner's language competence. Analysis of error types in the case of *takai* showed that lexical errors (omitting and wrong usage of a collocate) appear often, followed by errors in grammar.

Analysis of the developmental sequence of adjectives showed the development from plain simple forms of *takai* to usage of *takai* in more complex sentence structures with various roles. The transitional learner's interlanguage form is also observed (*takai da\* shi* ← *takai shi* "as well as high [*\*with* and without the copula *da*]").

Analysis of the lexical domains of the adjective *takai* used by the language learners revealed that the majority of target learners use only two domains relating to positional and quantity relations, while the third more abstract lexical domain referring to quality relations (superior) is widely used only by one learner, who seems to be the most proficient and with the richest vocabulary.

This study explained the overall usage of adjectives in C-JAS corpus and analyzed in detail the usage of the adjective *takai*. In the future, there is a need to do the analysis for other frequent adjectives, as well as expand the analysis to encompass various Japanese language learners' corpora. The development of a large-scale Japanese language learner corpus of written and speech data for learners who are native speakers of twelve different languages is currently in progress at the National Institute for Japanese Language and Linguistics (Sakoda 2013) and will surely contribute to the empirical study of Japanese learners' data as introduced in this research.

## References

- Dixon, R.M.W. (1982). *Where Have all the Adjectives Gone? and other Essays in Semantics and Syntax*. The Hague: Mouton.
- Nation, P. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press
- Srdanović, I. (2013, forthcoming) Daikibo koopasu wo mochiita keiyoushi to meishi no korokeishon no kijutsuteki kenkyuu: Nihongo kyouiku no jisho sakusei ni mukete (Description of Adjective and Noun Collocations Based on Large-Scale Corpora: Towards

Dictionary for Japanese Language Learners). *Kokuritsu kokugo kenkyuujo ronshuu (NINJAL Research Papers)* 6.

Sakoda, K., Kinoshita, A., Konishi, M., Lee, J. (2012) Nihongo gakushuusha no juudanteki kaiwa koopasu no kouchiku to shuutoku kenkyuu: 3 nennkan no deeta kara bunpou shuutoku no katei wo saguru (The development of the Japanese learners' conversational longitudinal corpus and the language acquisition research: observing the process of grammar acquisition from the three-year data). *2012 nendo nihongo kyouiku kokusai kenkyuu taikai yokoushuu (The proceeding of the international research conference on Japanese language didactics for year 2012)*

Sakoda, K. (2013) Nihongo gakushuusha no hatsuwa koopasu to doushi no hattatsu. (Learners' Spoken Corpus of Japanese and Developmental Sequence of Verbs) *Kokugoken purojekuto rebyuu (NINJAL Project Review)*, 107-116. Kokuritsu kokugo kenkyuujo (National Institute for the Japanese Language and Linguistics)

Sakoda, K. (2012) Hibogowasha no komyunikeishon kufuu (Ingenuity of communication for non-native speakers). *Nihongo kyouiku no tame no komyunikeishon kenkyuu (Studies in communication for Japanese language didactics)*, 105-124. Tokyo: Kuroshio shuppan