

Investigation into Language Learners' Acquisition Order Based on an Error Analysis of a Learner Corpus

Emi IZUMI and Hitoshi ISAHARA

*National Institute of Information and Communications Technology, 3-5 Hikaridai, Seika-cho,
Soraku-gun, Kyoto 619-0289, Japan*
Graduate School of Kobe University, 1-1, Rokkodai-cho, Nada-ku, Kobe-city, Hyogo, 657-8501, Japan
{emi, isahara}@nict.go.jp

Abstract

In foreign language education, it is important for teachers to know their students' acquisition order of major linguistic items in the target language. This enables them to teach these items more effectively in language classrooms. A hypothesis established in the 1970s based on studies aimed at revealing the natural sequence in second language acquisition is that major grammatical morphemes are acquired in a common order by learners across different backgrounds, such as their L1, ages, or learning environments (hypothesis 1). However, in the 1980s, studies on the acquisition order of Japanese learners of English led to a contradictory hypothesis that differences in learners' backgrounds can cause differences in their acquisition orders (hypothesis 2). These studies revealed that the acquisition order of Japanese learners differs from the sequence supporting hypothesis 1.

In this paper, we tried to see which of these two contradictory hypotheses could be supported by the acquisition order extracted from our NICT JLE (Japanese Learner English) Corpus. In this corpus, learners' grammatical and lexical errors have been annotated manually with 47 types of error tags useful for investigating the acquisition order. The results of the analysis showed no significant correlation between the sequence supporting hypothesis 1 and that extracted from our corpus. On the other hand, there was a significant correlation between our sequence and that supporting hypothesis 2. The most significant difference between our sequence and that supporting hypothesis 1 is that ours indicates Japanese learners acquire articles and plural -s in a later stage. This might arise from L1 transfer because Japanese language does not have any relevant markers for articles and plural -s.

Keywords: *acquisition order, learner corpus, error analysis, error tag*

1 Introduction

In foreign language teaching, for the optimum educative effect grammatical morphemes in the target language should be taught in a specific order. For example, if morpheme *A* cannot be learnt effectively without knowing morpheme *B*, morpheme *B* should be taught first. This is not always possible, however, because all morphemes cannot be sorted in a one-dimensional relationship. On the other hand, knowing which morphemes learners at a certain proficiency level have mastered and which ones they have not could lead to more efficient language teaching.

In the 1970s, many studies were performed in attempts to reveal the second language learners' acquisition order of major grammatical morphemes in the target language, mainly English. As the second languages spoken by the learners of different mother tongues have different features, it is clear that L1 interference has a large effect on second language acquisition (SLA). So, is the acquisition order different among learners of different L1? One popular hypothesis, formulated by comparing the acquisition orders of various groups of learners, holds that grammatical morphemes are acquired in a common sequence across different L1, ages, or learning environments (hypothesis 1). However, in languages spoken by non-native speakers of various mother tongues, we often observe differences not only in pronunciation but also in writing. In the 1980s, several studies, most of which addressed the acquisition order of Japanese learners of English, showed that Japanese learners' acquisition order differs from the one that led to hypothesis 1. Some significant differences were explained based on the

difference of L1. Based on these studies, it was hypothesized that differences in L1 or other aspects of learners' backgrounds cause differences in the acquisition order (hypothesis 2).

In this work, we first extracted the acquisition order from the NICT JLE (Japanese Learner English) Corpus, which is an error-coded spoken corpus of Japanese learners' English. A further analysis was then carried out to determine which of the two contradictory hypotheses our results support. The analysis was performed using tags that specify the learners' grammatical and lexical errors in the corpus. In the following section, we review the studies on the acquisition order in SLA research since the 1970s. Section 3 gives an overview of the NICT JLE Corpus including the error-tagging scheme. Section 4 describes the experiment we have carried out to reveal the correctness ranking of eight grammatical morphemes from the NICT JLE Corpus. Section 5 concludes the paper and mentions future work.

2 Acquisition Order in SLA

2.1 Acquisition order of major grammatical morphemes in Dulay and Burt (1973)

The study by Dulay and Burt (1973), a pioneering one in acquisition order research, revealed how six- to eight-year-old children (151 subjects, L1 Spanish) had acquired eight major English grammatical morphemes. The acquisition order obtained in that study is shown in Figure 1. They found that plural *-s* was acquired first, and the possessive *- 's* mastered last.

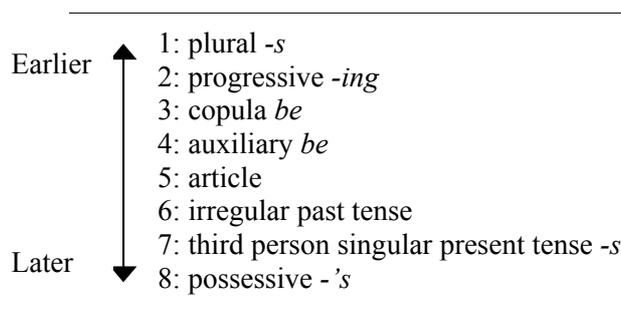


Figure 1 Acquisition order of children whose L1 was Spanish (Dulay and Burt, 1973).

The data were collected using an interview technique called the Bilingual Syntax Measure (BSM). The (Dulay, et al., 1982) BSM was designed to effectively extract the speech samples an examiner wants. For example, the examiner shows the examinee the pictures of a fat man and a thin man and asks, "*Why do you think he's fat?*" The examinee might answer, "*He eats too much.*" or "*He drinks too much beer.*" In this manner, the examiner can determine to what extent the examinee can use linguistic items properly in "natural but controlled" utterances that, though not totally spontaneous, are not nearly as restricted as ones in fill-in-the-blank questions or translation practices.

After this experiment, Dulay and Burt did several experiments in which they changed the number of examinees, used different learner groups, or used other grammatical morphemes (1974, 1975). They obtained similar results in all experiments. Furthermore, in their largest experiment (500 examinees, 13 types of grammatical morphemes) (Dulay, et al., 1984), they found that the acquisition order could not be drawn in a linear way. A hierarchy of several groups of morphemes existed, and the acquisition order of morphemes in each group could change among individuals. Their acquisition hierarchy is shown in Figure 2.

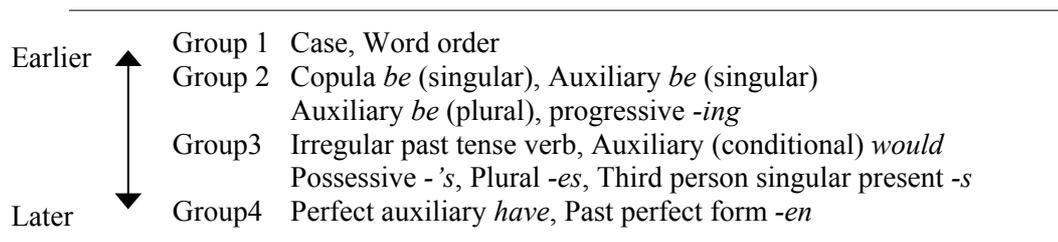


Figure 2 The acquisition hierarchy by Dulay, et al. (1984).

2.2 Studies supporting Dulay and Burt (1973, 1974, 1975)

Based on the acquisition order revealed by Dulay and Burt (1973, 1974, 1975), much follow-up research was carried out by others. In an experiment with adult learners, Bailey, Madden and Krashen (1974) found the same acquisition order as Dulay and Burt (1973, 1974, 1975) did for children. Krashen, et al. (1978), in experiments using 70 university students of four different mother tongues, and Makino (1980), in an investigation of the written English of 800 Japanese university students, found similar acquisition orders. Based on these studies, the hypothesis that a common acquisition order exists across learners of different backgrounds was formulated.

2.3 Studies on Japanese learners of English

However, in the 1980s, several experiments using mainly Japanese learners of English found no significant correlation with those mentioned above. It was obvious that Japanese learners acquire articles much later compared with other learner groups (Terauchi, 1994).

More recently, Tono (2002) revealed the Y8 to Y12 (Junior high school to university level) learners' acquisition order of eight major grammatical morphemes in the JEFFL (Japanese EFL Learner) Corpus. Compared with the results of Dulay and Burt (1973), articles and plural *-s* were mastered later, and possessive *- 's* was acquired earlier. If these three morphemes are excluded, there is little correlation between the works of Tono (2002) and Dulay and Burt (1973). Based on these follow-up studies, it was hypothesized that learners' background can affect the acquisition order.

3 The NICT JLE Corpus

Firstly, we overview the NICT JLE Corpus, focusing on error tagging.

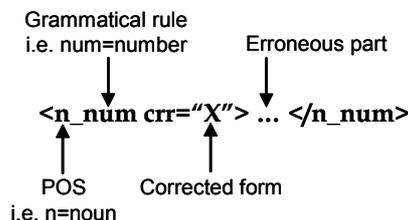
3.1 Overview

The NICT JLE Corpus is a collection of transcripts of English spoken by 1,281 Japanese learners, totalling 2 million words. In the transcribed texts, basic discourse phenomena, such as fillers, repetitions, or self-corrections, have been marked with pre-defined XML (Extensible Mark-up Language) tags. The utterances were obtained from an oral proficiency test system called the Standard Speaking Test (SST). The SST is an interview between one examiner and one test-taker. In a 15-minute interview, the test-taker is asked to perform three tasks: picture description, role-playing, and story-telling. Each test-taker is judged into one of 9 proficiency levels by two or three raters. This means that the learner's proficiency level is indicated in each corpus file. This information can be useful for analysing the learners' developmental stages.

3.2 Annotation on learners' errors (error tagging)

One of the unique features of the NICT JLE Corpus is that a part of the corpus data (167 samples) has been annotated with learners' grammatical and lexical errors based on an originally designed error-tag set. The current tag set consists of 47 types of tags. Since learners' errors vary across different linguistic levels and include grammatical, lexical, pragmatic, and discourse errors (James, 1998), it is quite difficult to categorize all of them systematically. Therefore, only grammatical and lexical are dealt with

in the current tag set. As shown in Figure 3, an error tag contains three pieces of information: a part of speech, a grammatical/lexical rule, and a corrected form. By replacing an erroneous word string with the corrected form indicated in a tag, grammatically and lexically correct sentences can be obtained automatically.



ex) *I belong to two baseball <n_num crr="teams">team</n_num>.

Figure 3 Structure of an error tag and an example of an error-tagged sentence.

Figure 4 shows how the actual error-tagged text looks.

And, for me, a car is just a car, just <F>ah</F> drive and carry us <prp_lxc1 crr="to"></prp_lxc1> somewhere else, but, for my husband, or I assume that <prp_lxc1 crr="for"></prp_lxc1> most <prp_lxc1 crr="of"></prp_lxc1> men, a car is something like a precious <n_lxc crr="thing"></n_lxc>, <at crr="a"></at> kind of like <at crr="an">a</at> accessory, or <./></./> <F>um</F> I don't know.
 <SC>But</SC> <F>eh</F> so the attitude <SC>for</SC> towards <at crr="a"></at> car is very different from the women's attitude.

<F>: Filler
 <SC>: Self-Correction
 <prp_lxc1>: Preposition error
 <at>: Article error

Figure 4 Sample of an error-tagged text.

3.3 Appropriateness of the corpus for the acquisition order analysis

We believe that analysing the data in the NICT JLE Corpus is appropriate for revealing the acquisition order of grammatical morphemes. This is because of the naturalness of the utterances in the corpus and because it contains samples of learners' using the basic linguistic items we wanted to examine. As mentioned above, the corpus data is entirely based on SST interviews involving three tasks. This approach is well designed to keep the test-takers' utterances natural, but controlled so that the examiner can ascertain how the basic language functions are used by the test-takers. For example, in the story-telling task, the examiner gives an instruction such as, "Please retell the story described in this picture as what happened last week." The examiner's aim here is to see how well the test-taker can use past tense forms. This fairly natural but controlled style of interviewing is quite similar to the BSM used by Dulay and Burt. Another reason for the corpus's appropriateness is that error frequency can be easily calculated by using error tags. To know the error frequency, we have to know not only the number of errors, but also the number of appearances of each morpheme in the obligatory context. "Obligatory context" is based on when a morpheme is presumed to be used in a correct usage. These factors can be obtained by converting learners' erroneous sentences into correct sentences by using the corrected forms in the error tags.

4 Correctness ranking of eight grammatical morphemes in the NICT JLE Corpus

4.1 Method

We used the Group Score Method (GSM) introduced by Dulay and Burt (1973). In the GSM, only the frequency of the correct use of each morpheme in the obligatory context is calculated. Firstly, an “expectation score” for each morpheme is defined based on how many times it must appear in the obligatory context. One appearance is counted as 2 points. So, for example, if articles must appear 45 times, the expectation score of articles is 90 points. Next, a “learners’ score” is counted for each morpheme depending on how many times they are used correctly or incorrectly. Each usage pattern is counted as shown in Table 1. If a morpheme is used correctly, two points are counted. If a morpheme is replaced with an erroneous one (replacement error), one point is counted. If a morpheme is omitted (omission error which is a rudimentary type of errors), no points are counted.

Table 1 Distribution of points for conditions of learners’ usage

Condition of learners’ usage	Point(s)
Correct	2
Incorrect: omission of a morpheme	0
Incorrect: an erroneous morpheme	1

Finally, the “morpheme score” is calculated as

$$\text{Morpheme_Score} = \frac{\text{Learner's_Score}}{\text{Expectation_Score}} \times 100$$

$$\text{Expectation_Score} = (\text{frequency of the morpheme in obligatory context}) \times 2$$

$$\text{Learner's_Score} = (\text{number of correct use}) \times 2 + (\text{number of omission errors}) \times 0 + (\text{number of replacement errors})$$

Then, the morphemes are ranked.

In our experiment, the expectation score was obtained by converting the erroneous sentences into correct sentences [crr=“X”] based on the information in the error tags, POS (Part of Speech) tagging them using “TreeTagger (Schmid, 1994)”, and counting the number of POS tags relevant to each morpheme. Secondly, the learners’ score was obtained by counting the number of error tags relevant to each morpheme. The types of values used for counting these scores are shown in Table 2.

In the GSM, only the frequency in the obligatory context is dealt with; “overuse” errors are not counted. Although overuse errors are not as frequent as other types (omission-type and replacement-type) in general, in some morphemes, such as articles, overuse errors sometimes occur frequently. Therefore, overuse errors should not be omitted if one wants to know the “real” situation of learner language. However, their omission is inevitable in GSM. In the future, we would like to account for overuse errors using a different method.

4.2 Data

In the analysis, 167 error-tagged samples from the NICT JLE Corpus were used. The detailed distribution of the data is shown in Table 3.

Table 2 Types of values used for calculating each morpheme score

Morpheme	POS tags and words for counting the expectation score	Error tag for counting the learner's score
plural -s	number of NNS and NPS (plural forms of common and proper nouns)	number of <n_num> (error in noun number) and <n_inf> (error in noun inflection)
progressive -ing	number of VVG, VBG and VHG (-ing form of common, be- and have-verb) in context of progressive tense	number of <v_fm> (error in verb form) and <v_tns> (tense error) whose corrected words are in -ing form
copula be	number of VB* (be-verb) used as copula.	number of <v_lxc> (verb lexical error) whose corrected words are be-verbs
auxiliary be	number of VB* (be-verb) used as auxiliary	number of <mo_lxc> (error in auxiliary verb) whose corrected words are auxiliary be
Article	number of a, an and the	number of <at> (article error) whose corrected forms are a, an, or the
irregular past tense	number of VVD (past tense form of common noun) which is irregular verb, and the number of VBD and VHD (be-verb and have-verb)	number of <v_inf> (error in verb inflection), <v_tns> and <v_asp> (error on verb aspect) whose corrected words are irregular verbs
third person singular present tense	number of VVZ, VBZ and VHZ (common verb, be-verb, have-verb)	number of <v_inf> and <v_agr> (subject-verb agreement error) whose corrected words are in their third singular present tense form
possessive -'s	number of ' (apostrophe)	number of <n_cs> (error in noun case) whose corrected expressions are in "A's B" form.

Table 3 Number of samples, words, sentences, and their distributions

Proficiency level (1 to 3 are novice, 4 to 8 intermediate, and 9 advanced)	Number of samples	Total number of words	Total number of sentences	Distribution
1	1	200	68	0.59%
2	7	1,480	475	4.19%
3	28	12,506	2,646	16.76%
4	43	30,244	5,161	25.74%
5	30	26,547	3,981	17.96%
6	28	28,061	3,864	16.76%
7	16	18,562	2,350	9.58%
8	9	10,507	1,301	5.38%
9	5	6,682	780	2.99%
Total	167	134,699	20,607	-

Although the numbers of samples for the 9 proficiency levels vary widely, their distribution is nearly the same as that for the entire corpus. Therefore, we assume this data set reflects the nature of the entire NICT JLE Corpus. In this experiment, we did not use the proficiency level information. We use 167 samples as one group of learner data.

4.3 Results and discussion

The resulting ranking of the morpheme scores is shown in Figure 5.

1: possessive - 's
2: progressive -ing
3: copula <i>be</i>
4: third person singular present tense -s
5: plural -s
6: auxiliary <i>be</i>
7: irregular past tense
8: article

Figure 5 Ranking of morpheme scores in the NICT JLE Corpus.

It is obvious that our ranking is quite different from Dulay and Burt's (1973) (Figure 1). The Spearman rank correlation between them is -0.48 ($p=0.911$), which is quite low. The most significant differences are listed below.

- In Dulay and Burt (1973), articles ranked fifth, but in our result they ranked eighth. The morpheme score itself was much lower than that of the irregular past tense, which ranked seventh.
- In Dulay and Burt (1973), possessive - 's ranked at the bottom, but in our result it ranked at the top.
- Plural -s ranked at the top in Dulay and Burt (1973), but in our result it ranked fifth.

The reasons why the error rate of articles was much higher than the others might be that the English article system is quite complex and, more importantly, that the Japanese language does not have an article system. Furthermore, as mentioned above, overuse errors were not counted, even though such errors occur frequently, especially in intermediate to advanced learners' data. Therefore, the actual error rate would be even higher.

As for possessive - 's, a lot of overuse of this morpheme can be found in the corpus data. Therefore, the number of omission type errors counted as "zero" decreased, and the overall accuracy rate went up because of the exclusion of overuse errors. The overuse of the "A 's B" form might be caused by the direct translation of the Japanese expression "A *no* B", where "*no*" is a particle for the possessive.

The reason why plural -s ranked fifth in our result might also be L1 interference, as in the Japanese language no distinction is made between singular and plural for most nouns.

Next, we compared our result with Tono's (2002) (Figure 6), who found the same tendency as our results for article, possessive - 's and plural -s in comparison with the results of Dulay and Burt (1973).

Earlier	↑	1: copula <i>be</i>
		2: auxiliary <i>be</i>
		3: possessive - 's
		4: progressive -ing
		5: plural -s
		6: third person singular present tense -s
Later	↓	7: irregular past tense
		8: article

Figure 6 Acquisition order according to Tono (2002)

Firstly, the Spearman rank correlation between our result and Tono (2002) was 0.738 ($p<0.05$), which is statistically significant. However, there are still some differences. For example, compared with Tono (2002), copula *be* and auxiliary *be* ranked low in our result. A possible reason for this might be that Tono (2002) mainly targeted the written language, while our corpus contains spoken data. For instance,

in the NICT JLE Corpus, the total number of errors in auxiliary verbs was 216, of which 94 were for auxiliary *be* (43 %). Further, the total number of lexical errors in verbs was 858, of which 170 were for copula *be* (17%), and most of them were omission errors, which are more elementary than other types of errors. This means that even though *be* is used more frequently than other types of auxiliary verbs or common verbs, learners tend to omit *be*, which does not have an explicit meaning, especially in spoken language where conveying messages is the most crucial task.

Taking all these points into consideration, we conclude that the results of our experiment support the hypothesis that differences in learners' background can cause differences in the acquisition order. Not only the difference in L1, but also the difference in the medium of production (spoken or written) can affect the order. Understanding the L1 interference more deeply requires that we compare the data of different learner groups carefully. It would be useful, for example, to arrange the morphemes that are acquired in the common order across different learner groups and the morphemes that are particularly vulnerable to L1 interference. This kind of information would be helpful for improving foreign language education. For example, it might become possible to design an original way of teaching each learner group to help them learn more effectively.

5. Conclusion

Based on acquisition order hypotheses in SLA research, we determined the Japanese learners' correctness ranking of eight major grammatical morphemes using error tags in the NICT JLE Corpus. The correlation between our results and the acquisition order by Dulay and Burt (1973) was quite low. The most significant differences were that the ranking of articles and plural *-s* went down and the ranking of possessive *'s* went up. We assume that these differences are caused mainly by L1 transfer. On the other hand, our result has a statistically significant correlation with the acquisition order for Japanese learners shown by Tono (2002). Therefore, we conclude that our result supports the hypothesis that differences in learners' backgrounds can cause the differences in acquisition orders.

However, we found several differences between our result and Tono's (2002), such as the ranking of copula and auxiliary *be*, and we attribute them to the difference in the medium of the language production. It is important to consider the extent of factors other than L1 on the learners' acquisition order.

In the future, we plan to continue this analysis for more large-scale data sets and other morphemes. To do this, we first need to increase the number of error-tagged samples. We deal with 47 types of morphemes in our error tag set, and we would like to reveal the correctness ranking of all of them. Furthermore, it is crucial that overuse errors be taken into account. Lastly, we need to analyze learners' errors more thoroughly by subcategorizing the error patterns in each morpheme and investigate the distribution of those patterns.

References

- Bailey, N., Madden, C., and Krashen, S. 1974. Is there natural sequence in adult second language learning? *Language Learning*, 24, 235-243.
- Dulay, H. and Burt, M. 1973. Should we teach children syntax? *Language Learning*, 23, 95-123.
- Dulay, H. and Burt, M. 1974. Natural Sequences in child second language acquisition. *Language Learning*, 24, 37-53.
- Dulay, H. and Burt, M. 1975. A new approach to discovering universals of child second language acquisition. In Dulay, H., Burt, M., and Krashen, S. (Translated by Makino, T.) 1984. *Daini gengo no syutoku*, Tokyo: Taka syobou yumi press, 160-188, (in Japanese).
- James, C. 1998. *Errors in Language Learning and Use*, London: Longman.
- Krashen, S., Butler, J., Birnbaum, R., and Robertson, J. 1978. Two studies in language acquisition and language learning. *ITL, Review of Applied Linguistics*, 73-92.
- Makino, T. 1981. *Acquisition Order of English Morphemes by Japanese Adolescents*, Tokyo: Shinozaki Shorin Press.

- Schmid, H. 1994. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of International Conference on New Methods in Language Processing*, 44-49.
- Terauchi, M. 1994. "Keitaiso no shutoku", In Koike, I. (eds) 1994. *Daini gengo syutoku kenkyu ni motodoku saishin no eigokyouiku*, Tokyo: Taisyukan, 24-39, (in Japanese).
- Tono, Y. 2002. *The Role of Learner Corpora in SLA Research and Foreign Language Teaching: The Multiple Comparison Approach*. Unpublished PhD dissertation, Lancaster University, UK.