The Economic Journal of Takasaki City University of Economics vol.49 No.3,4 2007 pp.117～129

Grammatical Errors across Proficiency Levels in L2 Spoken and Written English

Abe Mariko

Introduction

Studies of grammatical development in second language acquisition fall into two categories: (a) studies investigating formal characteristics of the acquisition process and (b) studies proposing developmental indices for assessing the overall progress of second language learners (Bardovi-Harlig & Bofman, 1989). This study belongs to the first of these categories; its aim is to identify and characterize features of second language use in Japanese learners of English that explicitly indicate developmental progress. Towards this end, the study also attempts — by making use of learner corpora — to characterize Japanese learners' errors of spoken and written English in terms of noun-, verb-, and other part-of-speech-related errors.

As previous studies of language acquisition have been restricted to relatively small amounts of data, research using larger data sets may lead to significant advances in the understanding of language acquisition (Biber, Conrad, & Reppen, 1998). For this study, a substantial body of spoken and written data were used to investigate differences between spontaneous spoken production and less time-pressured written production to show the acquisition sequence of certain grammatical features in the different production modes.

Purpose

Most learners' error studies to date have not considered differences in the errors made by learners at different stages of development, and consequently have provided a static view of language acquisition (Ellis, 1994) that affords little insight into the developmental sequence of acquisition. This research focuses on error types associated with different second language proficiency levels, and — by using cross-sectional data from different production — how these error types may characterize each developmental stage.

Another problem shared by most learners' error studies is the considerable variation in error frequency among second language learners shown by Chamot (1978, 1979). This 44-month

longitudinal study found significant variation, a finding that has made it difficult to describe learners' language development satisfactorily by quantifying types of errors. To address this problem, the present study increased the amount of data and error categories included as far as possible, and furthermore employed rigorous statistical treatment.

Research questions

The following research questions are pursued in this paper: (1) Are there any differences in rates of part-of-speech accuracy between spoken and written L2 production? (2) What differences can be observed in patterns of noun- and verb-related errors in different production modes?

Methods

Materials

Spoken data were extracted from the ***National Institute of Information and Communications Technology Japanese Learner English Corpus*** (NICT JLE Corpus). The NICT JLE Corpus consists of interview protocol data elicited from 15 minute interviews of approximately 1,200 Japanese EFL learners. Each learner's file is marked with one of nine different grades assessing the subject's spoken-English proficiency (Izumi, Uchimoto, & Isahara, 2004). The present study is based on spoken data from 100 examinees whose proficiency levels were assessed, variously, as SST level 2 (mid-novice), level 3 (high novice), level 4 (low intermediate), level 5 (low-plus intermediate), level 6 (mid-intermediate), level 7 (mid-plus intermediate), level 8 (high intermediate), and level 9 (advanced). Since the numbers of examinees in levels 2 and 9 were low, these levels were combined with levels 3 and 8, respectively.

The examinees all took the ***Standard Speaking Test*** (SST) to determine their proficiency. The SST, which was developed by ACTFL and ALC press, aims to assess the speaking proficiency of Japanese learners of English on the theoretical basis of ACTFL OPI. According to the SST evaluation criteria, performance samples are recorded and evaluated by two or three certificated evaluators. The speaking test consists of five different stages: (1) warm-up questions, (2) a single picture description task, (3) role-playing with an interviewer, (4) a picture sequence task, and (5) wind-down questions; however, data were extracted only from the single picture description task. This task does not target specific grammatical points such as verb tenses, as does the picture sequence task, nor does it target functional expressions, as in the role-playing task; rather, the picture description task presumably elicits the most natural and spontaneous language usage of the various components of the speaking test. In this task, examinees are asked, with no planning time,

to describe a picture; the resulting data therefore consist not of conversation or speech, but rather of story-telling production. The breakdown by proficiency of the 100 files used in this study is shown in Table 1.

Table 1 Corpus Size of Spoken Data

SST Level	SST 2/3	SST 4	SST 5	SST 6	SST 7	SST 8/9	Total
Examinees	22	17	16	19	16	10	100
Tokens	1,222	1,418	1,755	1,891	2,004	1,370	9,660

Written data were extracted from a corpus of compositions produced by Japanese junior and senior high-school EFL learners. The corpus includes compositions on six different topics, which elicited both narrative and argumentative compositions. Only one type of narrative composition, however, was chosen for this research. The compositions used were produced as a 20-minute, in-class, handwritten task, with no additional preparation time and no use of references allowed. Essays by learners from six different academic years (grades 7 to 12) were included. The size of the written data sub-corpus for each academic year was almost the same, since approximately 5,000 token-composition files were randomly extracted from each academic year.

Table 2 Corpus Size of Written Data

School year	Junior1 (J1)	Junior2 (J2)	Junior3 (J3)	Senior1 (S1)	Senior2 (S2)	Senior3 (S3)	Total
Examinees	104	77	87	46	53	55	422
Tokens	4,994	5,004	5,000	4,997	5,000	5,005	30,000

Although this study attempts to address weaknesses of previous error analysis studies by increasing the amount of learner language data analyzed, it nonetheless has the limitation that the tasks employed to elicit spoken and written data were not identical. The examinees who produced the spoken data were asked to describe a picture, whereas those who produced the written data were asked to write about a personal experience. Future studies must determine whether differences in error and accuracy rates for each data type resulted from the different language-production tasks or the different production modes.

Procedure

Error tags have been inserted into the data in the NICT JLE corpus, from which the spoken data were drawn, and for this study these tags were reviewed and modified as necessary (see Appendix). A definition of an error as "a deviation from the norms of the target language" (Ellis, 1994, p. 51) was employed, with the grammar taught in Japanese ELT classrooms being regarded as normative. Written data were error-tagged manually, using the same error tagging guidelines used for the spoken corpus. Figure 1 presents a sample of error-tagged spoken data.

<F>Er</F> yes. <F>Er</F>. All right. It's very sunny. And <F>er</F> there <v_agr crr="are">is</v_agr> a lot of people in this picture. <F>Er</F>. <F>Er</F> <R>two</R> two <n_inf crr="children">childrens</n_inf> are playing <F>er</F> with <g_at crr="a">the</g_at> ball. And <F>er</F> <F>er</F> one girl <F>er</F> <F>uum</F> <JP>nawatobi nante iundarou</JP> <F>er</F> <F>mm</F> is <v_fml crr="playing">play</v_fml> alone.

Figure 1. A sample of error-tagged spoken data

As error analysis studies have been restricted to what learners cannot do, it is essential to examine what learners can produce correctly (Larsen-Freeman & Long, 1991; Ellis, 1994). Part-of-speech (POS) tags were therefore added using the CLAWS tagger using the C7 tagset (Garside, Leech, & McEnery, 1997). To avoid skewing accuracy rates, POS classifications that do not correspond with what is taught in Japanese ELT classrooms were manually changed. Using POS tags, obligatory occasions for the use of each category in an error tag were identified, and occasions on which a feature was employed correctly or incorrectly were identified to enable calculation of error and accuracy rates.

Subsequently, a multivariate statistical method called Correspondence Analysis was performed to clarify the correlation of each proficiency level and error frequency. Correspondence Analysis shows the relationship of two nominal variables by mapping the results such that similarities and dissimilarities of the variables are visually apparent. This statistical method is useful as a first step in reducing the complexity of data, as well as in identifying points for detailed analysis.

Results and Discussion

Accuracy rate for each part of speech

The average accuracy rate for all grammatical categories in each part of speech is shown in

Figure 2 and 3. Developmental characteristics that can be observed in changes in accuracy rates across proficiency levels are summarized in Table 3.

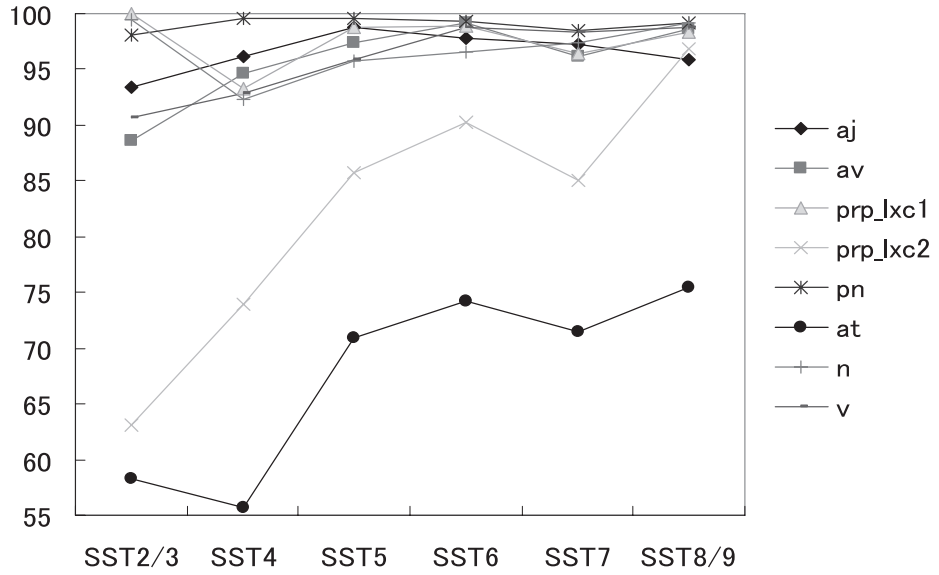

Figure 2. Accuracy rate of spoken L2 data

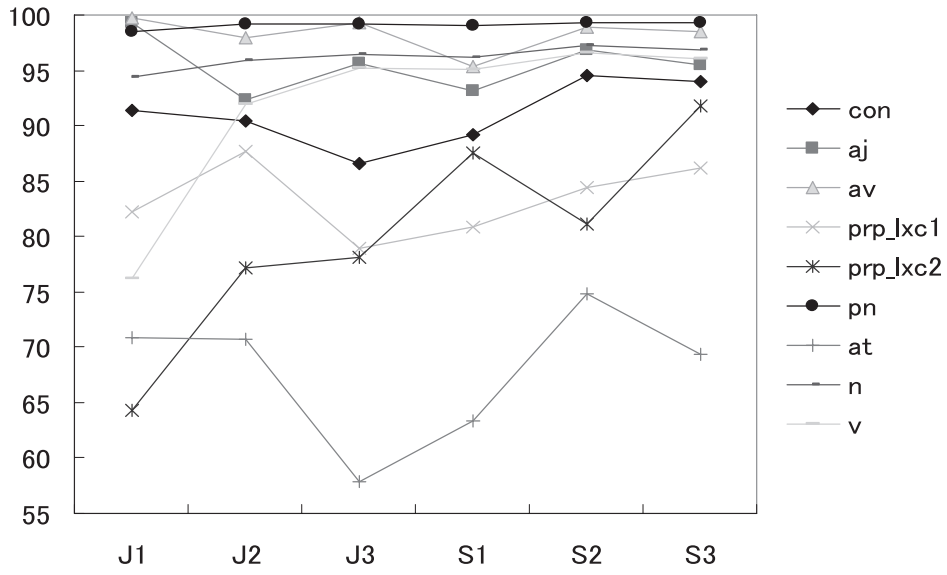

Figure 3. Accuracy rate of written L2 data

Table 3 Tendencies regarding accuracy rate in each production mode

Tendency	SP		WR	
	(%)	POS	(%)	POS
High accuracy rate and not much difference in accuracy rate across proficiency level	90-100%	Pronoun, Adverb, Noun, Adjective, Verb, Preposition 1 (noun and adjective)	90-100%	Pronoun, Adverb, Noun, Adjective, Verb
Accuracy rate increases gradually from 3rd year of junior high school	—	No data	90-80%	Conjunction, Preposition 1 (noun and adjective)
Accuracy rate increases steadily as proficiency level increases	60-100%	Preposition 2 (verb)	65-90%	Preposition 2 (verb)
Low accuracy rate relative to other parts of speech, but increases to around 75%	55-75%	Article	58-73%	Article

First, the accuracy rate for prepositions related to verbs "prp_lxc2" (Tom's teacher accused him *about cheating. / I came *to here.) increased dramatically in both production modes. The prepositions in this category comprise subordinating prepositions and prepositions in phrasal verbs, which language teachers often consider especially challenging for Japanese learners of English. Yet contrary to any expectations that the accuracy rate might remain low in this category, it increased considerably, arriving at 96.7% in spoken mode and 91.2% in written mode respectively at the highest level. Thus, errors involving prepositions associated with verbs may disappear as language acquisition progresses.

Secondly, the accuracy rate for articles increased significantly in both production modes. Notably, the highest accuracy rate in this error category is lower than that for other parts of speech, and accuracy rates in this category are lower than for other parts of speech at every stage except Junior 1 in written mode. The accuracy rate does increase approximately 20% from the lowest to the highest levels, but nonetheless, the article is clearly a problematic item for Japanese learners compared with other parts of speech.

Thirdly, the overall accuracy rates for pronouns, adjectives, adverbs, nouns, and verbs ranged from approximately 90% to 100%. These part-of-speech categories had consistently high accuracy rates at all proficiency levels, but lack of sentence complexity may contribute to the high rates at lower levels. That is, less proficient learners may produce shorter, simpler sentences that are less likely to contain errors than more proficient learners.

While the accuracy rate of spoken production increases, that of written production changes little or shows irregular variations. These irregular results may reveal variability in the development of written English proficiency. It is also possible that learners attempt to use more of the varied grammar points and vocabulary that they have learned at school in written production mode, in which they are under less time pressure than in spoken mode.

Noun- and verb-related errors across proficiency levels

In this section, the error categories for nouns and verbs are examined more closely. Among the part-of-speech errors that maintained high accuracy rates across proficiency levels, noun- and verb-related errors were investigated in detail because they are fundamental in sentence construction.

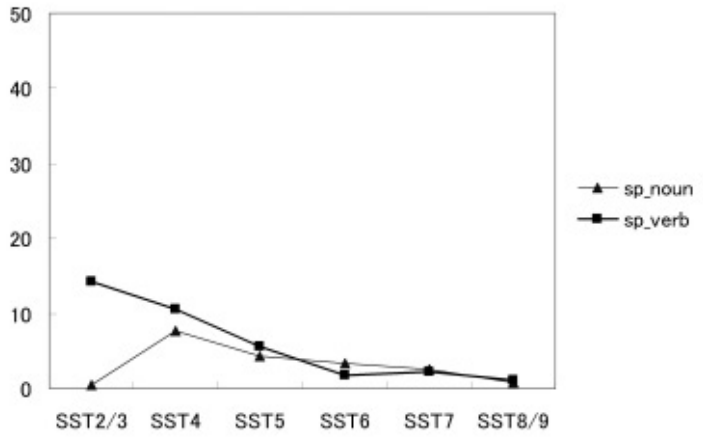

Figure 4. Average error rates for nouns and verbs in spoken data

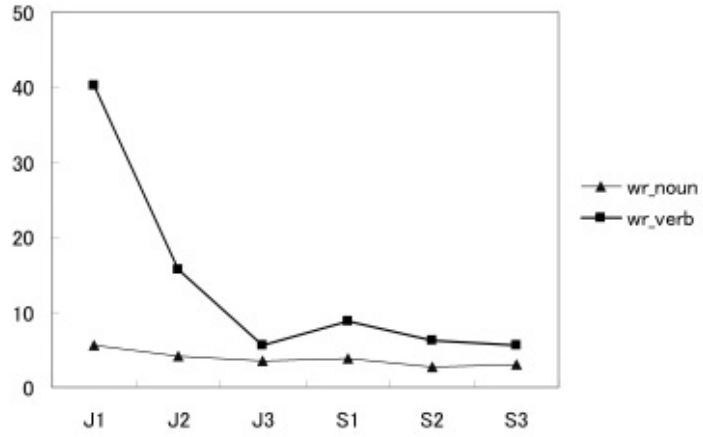

Figure 5. Average error rates for nouns and verbs in written data

As indicated in Figure 4, the verb-related errors in the spoken data for novice learners are relatively high and they gradually decrease as proficiency level increases. Similarly, the error rate for verbs in the written data is pronounced in the first grade (J1), but it too decreases over time. The noun-related error rate is relatively low compared with the rate for verb-related errors, and interestingly, the noun-related error rate remains almost unchanged across the academic years, from J1 to S3, in the written data, which may imply that noun-related errors are not easily overcome in the course of the acquisition of written English.

The next question examined whether or not there were any dissimilarities in patterns of grammatical errors in L2 spoken and written production, and if so, which particular grammatical category of noun- and verb-related error is strongly associated with the two production modes. The frequencies of errors were normalized based on the total size of each developmental stage sub-corpus and the overall frequency of nouns and verbs across the stages. The results of Correspondence Analysis are shown in Figure 6.

In this figure, Dimension 1 shows the difference of mode. Written production is plotted on the left and spoken production on the right. Dimension 2 indicates proficiency level; the lower-level learners are plotted on the upper part and the advanced-level learners on the lower part. Regarding written production, all school academic years except J1 are clustered together, and SST proficiency levels are clearly dispersed.

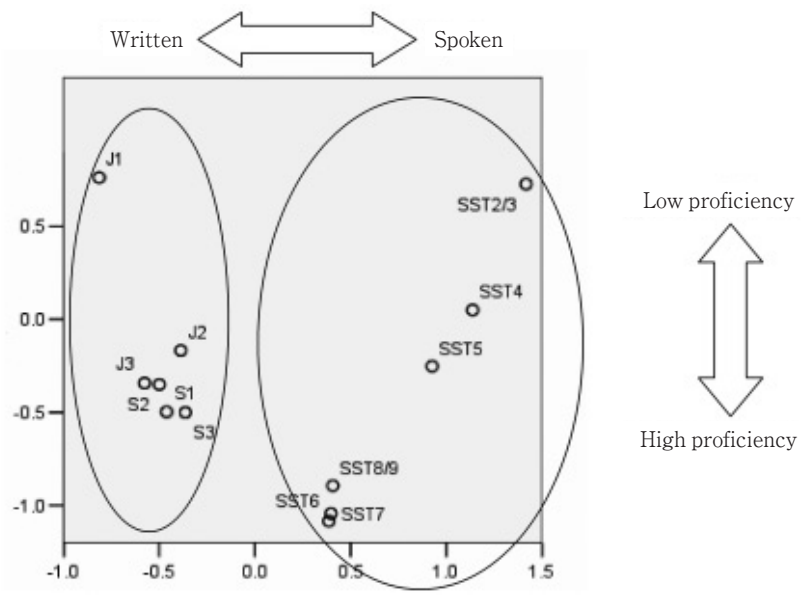

Figure 6. Distribution of noun- and verb-related error categories

The results shown in Figure 7 indicate that verb-related errors tend to be positioned in the upper part of the figure, and noun-related errors in the lower part. Verb-related errors can therefore be assumed to be associated with lower-level learners, and noun-related errors with advanced-level learners in each production mode.

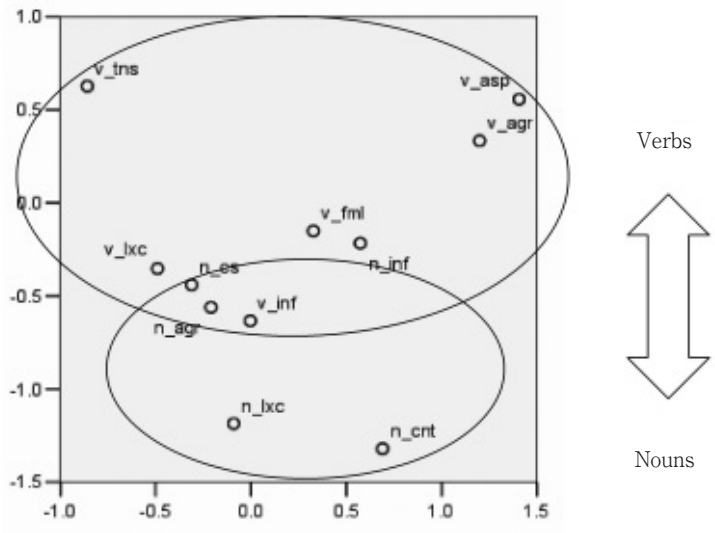

Figure 7. Distribution of noun- and verb-related error categories

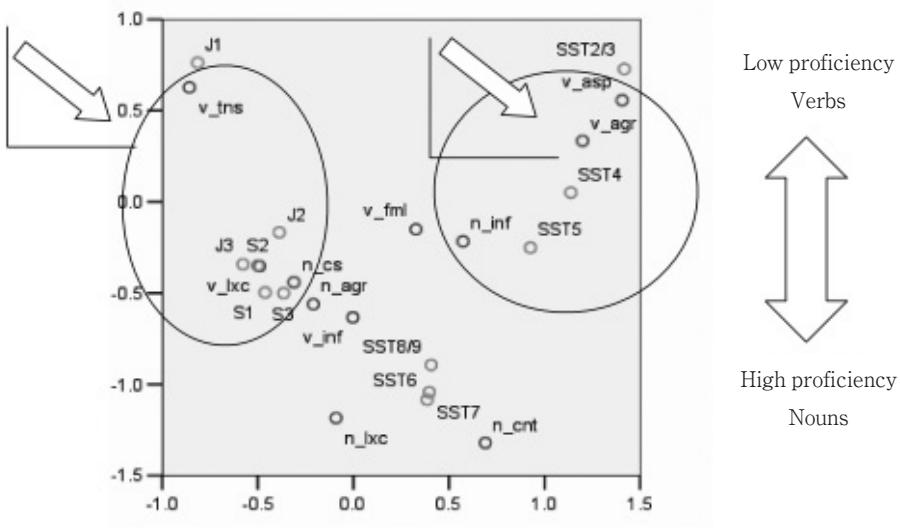

Figure 8. Error categories that are strongly related to error rate patterns

Next, the characteristics of error rate changes in both modes were examined closely (see Figure 8). Since the agreement rule in English is complex, it can be a troublesome grammar point for

learners of English. However, the error rate for agreement decreases as proficiency level increases. Also, it was clear that verb aspect errors and noun inflection errors decrease over the developmental stages of spoken production, and tense and verbal lexical error decrease over the developmental stages of written production. These results represent general patterns in noun- and verb-related errors and their variability over the course of language acquisition. Through detailed examination of error categories, we observed errors that share common developmental patterns and that correlate with particular production modes. The characteristics of error rates in both modes are summarized in Table 4.

Table 4 Characteristics of Error Rates in Both Modes

Characteristics	WR	SP
Error rate decreases as proficiency level increases	v_tns (tense) v_lxc (lexical errors) n_cs (case)	v_asp (aspect) n_agr (agreement) n_inf (inflection)
Error rate gradually increases, then decreases, as proficiency level increases	n_lxc (lexical errors)	n_cnt (countable / uncountable)
Error rate gradually decreases, then increases, as proficiency level increases	n_agr (agreement) v_inf (inflection)	— —
Unpredictable	v_fml (form)	—

Conclusions

This analysis of learners' language development has described how errors vary across stages of language acquisition and production mode. By analyzing errors at different developmental stages and in different production modes, the study has shown the possibilities for using learner corpora for language acquisition research. Furthermore, detailed examination of part-of-speech accuracy rates and use of nouns and verbs in learner data indicates that some patterns in the occurrence of errors are closely associated with learners' production mode.

A few methodological issues remain for future studies to resolve. It would be desirable to investigate other sections of the protocol data in addition to the picture description task. This study mainly investigated grammatical and lexical errors, but much remains to be done with regard to the numerous other types of errors, such as syntactical errors. It may also be important to focus

on what learners are producing correctly, overusing, underusing, or avoiding in the course of development.

The results of this study, nonetheless, reveal patterns in rates of noun- and verb-related errors during the development of proficiency: verbal errors were firmly associated with lower-level learners, and nominal errors were firmly associated with advanced-level learners. Furthermore, noun-related errors in written production do not seem to vanish readily over the course of development. In addition, detailed examinations of error categories suggest that some errors may share common developmental patterns, while others may vary uniquely across proficiency levels. The results also imply that some types of errors do not steadily disappear during the acquisition process.

In conclusion, this study supports the assumption that errors can provide information about the current state of learners' language development, as Corder (1967) argued, and furthermore, that errors can characterize the linguistic competence of learners. This in turn argues for the value of investigating characteristics of learner language that explicitly describe developmental progress.

Since errors provide information on the current state of learners' language development, it remains essential for language teachers to understand learners' errors (Corder, 1967). The better language teachers understand about how much learners have learned — and what problems learners face — the more effective teachers can become in providing targeted instruction and feedback. The ability to recognize what types of errors occur frequently at a given point in language development may enable teachers to create better teaching materials and language tests — that is, materials and tests that target common learner errors. Furthermore, it is not yet clear which errors should be treated lightly and which should be regarded as significant in the classroom; consequently, teachers are currently left to rely on intuition in responding to errors. Hence, further research into learners' errors may afford new insights into second language acquisition that will enable advances in both teaching and learning.

Appendix The NICT JLE Corpus Error Tagging Guideline and Examples of Error Tagged Data

POS	Error Category	Tag	Examples
Adjective	Inflection	<aj_inf>	*more tall
	Comparison	<aj_us>	Jane is taller than Mary, but Mary is the *best basket ball player.
	Quantifier	<aj_qnt>	There was *few traffic on the road.
	Word choice	<aj_lxc>	It is a *genius diamond.
Adverb	Inflection	<av_inf>	*more far
	Comparison	<av_us>	She came back *most quickly than me.
	Position	<av_pst>	I have difficulty *often in understanding her.
	Word choice	<av_lxc>	He worked *hardly today.
Article	Article	<at>	*a apple / *She is in a development of low cost water pumps. / *Her office is on twenty-third floor of Trump Tower. / *The letter was posted on February the second.
Noun	Inflection	<n_inf>	*childerens / *housewifes / *peoples
	Agreement	<n_agr>	many *book / one *books / a *books / each *books
	Countability	<n_cnt>	*a music / *musics
	Case	<n_cs>	my *friend house
	Word choice	<n_lxc>	*type (a typewriter)
Preposition	Complement	<prp_cmp>	I look forward *to see you again.
	Word choice	<prp_lxc1>	It was held *on June. (adjective & noun)
	Omission	<prp_lxc1>	*It was held June.
	Word choice	<prp_lxc2>	Tom's teacher accused him *about cheating. (verb)
	Omission	<prp_lxc2>	*Tom's teacher accused him cheating.
Pronoun	Inflection	<pn_inf>	*themselfes
	Agreement	<pn_agr>	It is a good book. I like *them.
	Case	<pn_cs>	*We school festival is very good.
	Word choice	<pn_lxc>	I often ask *me why I work so hard.
	Omission	<pn_lxc>	*will go back home early today.
Verb	Inflection	<v_inf>	*sleeped
	Subject-verb agreement	<v_agr>	there *are a cat / there *is cats / he *like / I *likes
	Tense	<v_tns>	I *eat breakfast this morning.
	Aspect	<v_asp>	The people *weren't knowing the reality.
	Form	<v_fml>	To *drinks / is *drink
	Word choice	<v_lxc>	She *is black and short hair.
	Omission	<v_lxc>	*She black and short hair.

(Full-Time Lecturer, The Faculty of Economics, Takasaki city University of Economics)

References

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition 11,* 17-34.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Chamot, A. (1978). Grammatical problems in learning English as a third language. In E. Hatch (Ed.), *Second language acquisition* (pp. 175-189). Rowley, MA: Newbury House.

Chamot, A. (1979). Strategies in the acquisition of English structures by a child bilingual in Spanish and French. In R. W. Andersen (Ed.), *The acquisition and use of Spanish and English as first and second languages* (pp. 90-106). Washington, DC: TESOL.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics 5,* 161-169.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Garside, R., Leech, G., & McEnery, T. (Eds.). (1997). *Corpus annotation: Linguistic information from computer text corpora.* Harlow: Longman.

Izumi, E., Uchimoto, K., & Isahara, H. (2004). *A speaking corpus of 1200 Japanese learners of English.* Tokyo: ALC Press.

Larsen-Freeman, D., & Long, M. (1991). *An introduction to second language acquisition research.* New York: Longman.