

Unit 7 Japanese English at the university Level: tentative analysis based on speech samples

**Michiko Nakano, Kazuharu Owada, Masanori Oya, Norifumi Ueda,
Tae Yamazaki, Eiichiro Tsutsui, and Minako Sunaga
(Waseda University)**

Unit 7-1 Grammatical features

Unit 7-2 Dysfluency features

1. Introduction

The purpose of this unit is (1) to propose a tentative version of easy-to-use tagset for the analysis of Japanese EFL learners' spoken data and (2) to show what types of errors are prone to occur along with the frequency data.

2. Data and method

2.1. Data obtained from the learners who had attended Tutorial English

Waseda University has been offering a program called *Tutorial English*. This program assigns one qualified tutor to four students. In 1997 – 2002 the students had been trained to improve their oral communication skills in English for the period of eight weeks, but more recently the students study for the period of ten weeks. The class met twice a week. During the 1997-2000 course of the *Tutorial English*, the students were required to take a fifteen-minute oral interview test called 'Standard Speaking Test' (SST) three times. After each interview test, they are graded Level 1 to Level 9 based on the criteria of SST. This oral interview test was developed jointly by ACTFL (American Council on the Teaching of Foreign Languages) and ALC Press Inc. This test includes such tasks as Picture Description, Role Play, and Story Telling. The four characteristics of SST are: 1) SST is a descriptive test, 2) SST is a proficiency test, 3) SST is a criterion-reference test, and 4) SST is an efficient test.

We obtained some tentatively transcribed data and tapes from TAO who

developed the system of tagging and created the spoken corpus of Japanese English in 2004 and now available in public.

2.3. Subjects

Thirty students were semi-randomly chosen for this study. They range in grade from Level 4 to Level 9 on the SST. Table 1 below shows the level of each student.

Table 1. Learner profiles

subject/file name	SST level	TOEFL	TOEIC	STEP grade
11	5	NA	725	pre-1st
12	5	420	630	2nd
13	4	NA	NA	NA
14	5	NA	NA	NA
15	9	NA	855	pre-1st
16	6	510	NA	NA
17	6	NA	NA	pre-2nd
18	8	580	NA	NA
19	6	NA	875	NA
20	4	NA	670	NA
21	6	NA	660	NA
22	7	550	850	NA
23	4	NA	NA	NA
24	4	NA	690	NA
25	6	NA	NA	NA
26	7	NA	835	2nd
27	6	540	745	pre-1st
28	4	NA	665	NA
29	4	NA	NA	NA
30	6	NA	750	NA
31	4	NA	NA	2nd
32	4	NA	NA	NA
33	4	NA	NA	2nd
34	7	513	NA	pre-1st
35	6	557	NA	NA
36	4	NA	NA	NA
37	7	NA	NA	NA
38	4	NA	NA	NA
39	4	NA	NA	NA
40	4	NA	NA	NA

2.4. Error tagset used in this study

Our tagset, which was not based on pre-existing codification, was compiled inductively through our experiences as EFL teachers in Japan. Therefore, our coding system may be applicable only to some particular groups of Japanese EFL learners. However, we strongly believe that, with some modifications, this tagset can be used for other groups of EFL learners.

We categorized our tagset into nine error types: 1) noun number, 2) agreement, 3) tense, 4) aspect, 5) lexical choice, 6) omission, 7) addition, 8) word order, and 9) unclear (see Appendix 1).

Table 2. The tagset used in this corpus data analysis

Error Types	error tags	meaning/example	frequency
noun number	<+pl></+pl>	a <+pl>men</+pl>	25
	<-pl></-pl>	two <-pl>hour</hour>	132
			157
agreement	<agr></agr>	agreement	93
tense	<prs></prs>	present	21
	<pst></pst>	past	150
	<fut></fut>	future	8
			179
aspect	<ing></ing>	progressive	16
	<perf></perf>	perfect	6
			22
lexical choice	<eart></eart>	article errors	71
	<epos></epos>	possessive pronoun errors	0
	<edem></edem>	demonstrative pronoun errors	9
	<econj></econj>	conjunct errors	16
	<eadv></eadv>	adverbs errors	8
	<eadj></eadj>	adjective errors	27
	<emorph></emorph>	morphological errors	36
	<eprep></eprep>	preposition errors	86
	<einter></einter>	interjection errors	1
	<epron></epron>	pronoun errors	1
	<elex></elex>	lexical errors	136
	<emda></emda>	modal auxiliary errors	1
	<eindpro></eindpro>	indefinite pronoun errors	1
	<everb></everb>	verb errors	4
	<enoun></enoun>	noun errors	1
			398

addition	<redart></redart>	redandant articles	34
	<redconj></redconj>	redandant conjuncts	1
	<redinter></redinter>	redandant interjections	0
	<redprep></redprep>	redandant prepositions	33
	<redpron></redpron>	redandant pronouns	0
	<redadv></redadv>	redandant adverbs	2
	<redadj></redadj>	redandant adjectives	3
	<redpos></redpos>	redandant possessive pronouns	1
	<redsub></redsub>	redandant subjects	0
	<redobj></redobj>	redandant objects	2
	<redverb></redverb>	redandant verbs	0
	<redcop></redcop>	redandant copulas	9
	<redmda></redmda>	redandant modal auxiliaries	3
	<redapos></redapos>	redandant apostrophes	0
	<rednoun></rednoun>	redandant nouns	0
			88
word order	<WO></WO>	word order errors	19
			19
unclear	<@></@>	uncategorizable errors	89
			89

2.5. Research Questions

We will address the following two research questions.

- 1) What types of errors in general do the learners tend to make?
- 2) What kind of preposition errors do the learners tend to make?

3. Results and data analysis

3.1. The results of the nine error types

The data contained 1,511 errors, according to the criteria of our tagset. Of the nine error types, ‘omission’ errors occurred most frequently (166; 32%), followed by ‘lexical choice errors’ (398; 26%). The breakdown of the errors is shown in Figure 1.

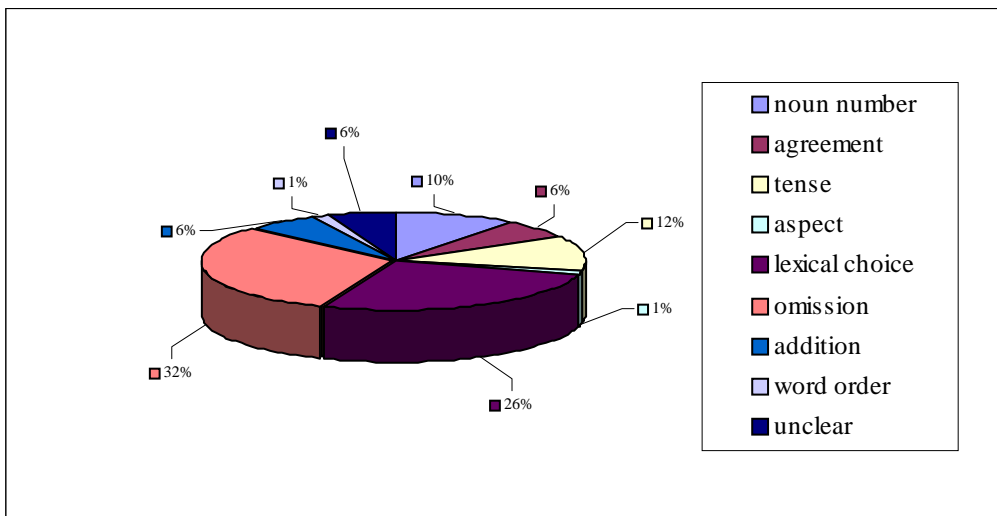


Figure 1. The breakdown of the nine error types

We then calculated the mean number of each error type for ‘Intermediate Group’ (Level 4-6, n=24) and ‘Advanced Group’ (Level 7-9, n=6) in order to see if there is any difference between the two groups. As is seen in Figure 1, ‘Intermediate Group’ made two times more errors than ‘Advanced Group’ in omission error types. Figure 2 indicates that if students can reduce ‘omission’ error type, they can enhance grammatical accuracy.

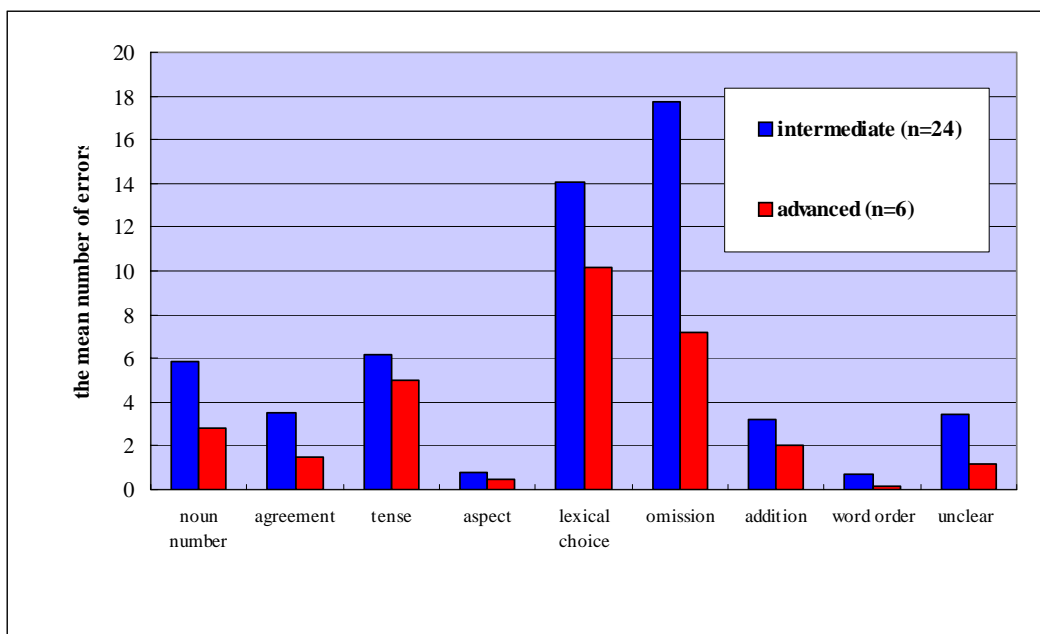


Figure 2. The mean number of each error type for Intermediate Group

(Level 4-6, n=24) and Advanced Group (Level 7-9, n=6)

3.2. The results of errors

In this section we will focus on the error of proposition in ‘omission’ and ‘addition’ error types as a whole, because these two types, being a mirror image of each other, will represent some interesting features worth looking into.

3.2.1. Preposition omission errors

As in Table 4, 466 instances of ‘omission’ errors occurred. The most frequently-occurred errors were ‘article omission’ errors (265; 57%), followed by ‘preposition omission’ errors tagged as <delprep></delprep> (83; 18%).

Table 4. Omission (error type) and preposition omission

Error Type	tag	<i>f</i>
omission	<delart></delart>	265
	<delconj></delconj>	3
	<delinter></delinter>	1
	<delprep></delprep>	83
	<delpron></delpron>	6
	<deladv></deladv>	4
	<deladj></deladj>	5
	<delpos></delpos>	5
	<delsub></delsub>	2
	<delobj></delobj>	47
	<delverb></delverb>	12
	<delcop></delcop>	22
	<delmda></delmda>	3
	<delapos></delapos>	2
	<delnoun></delnoun>	6
	466	

Preposition omission	<i>f</i>
TO	36
IN	13
AT	11
OF	8
ON	5
WITH	3
FROM	2
BY	1
ABOUT	1
OUT OF	1
DURING	1
FOR	1
	83

3.2.2. Preposition addition errors

As in Table 5, 88 instances of ‘addition’ errors occurred. The most frequently-occurred errors were ‘article addition’ errors (34; 40%), followed by ‘preposition addition’ errors tagged as `<redprep></redprep>` (33; 38%).

Table 5. Addition (error type) and preposition addition

Error Type	tag	<i>f</i>
addition	<redart></redart>	34
	<redconj></redconj>	1
	<redinter></redinter>	0
	<redprep></redprep>	33
	<redpron></redpron>	0
	<redadv></redadv>	2
	<redadj></redadj>	3
	<redpos></redpos>	1
	<redsub></redsub>	0
	<redobj></redobj>	2
	<redverb></redverb>	0
	<redcop></redcop>	9
	<redmda></redmda>	3
	<redapos></redapos>	0
	<rednoun></rednoun>	0
	88	

Preposition addition errors	<i>f</i>
to	14
in	10
for	3
at	3
from	2
with	1
	33

3.2.3. Possible learner chunks

When we look closely at ‘preposition’ errors in terms of ‘omission’ and ‘addition’ errors, which are mirror images of each other, we can extract some learner error chunks as follows:

go back my hometown, return my home, come university (omission)

go to there, go to shopping, I always go to with my husband (addition)

4. Conclusion

In this small-scale corpus study we proposed our tagset for the spoken data analysis in order to extract some of the errors produced by Japanese EFL learners. Then we explored the possibility of applying this tagset into the actual spoken data.

One of the unfortunate things about this data is that we have no way of asking the subjects to clarify what we consider to be unclear utterances because a large

period of time elapsed since the test. For future research, we should try to obtain protocol data as soon as the interview is finished so that no utterances will remain unclear for the researchers.

Furthermore, we need to establish clear-cut criteria which researchers all agree on. We also need to expand the corpus to include a sufficient number of subjects from each level.

Appendix 1: Some examples of each Error Type

Noun number

But, <F>er</F>, <reform>from</reform>, <F>a</F>, it takes one or two <-pl>hour</-pl> to go to your home. So, <./> <F>mm</F>, I have no time.

<delart>A</delart> man was driving his car and <F>mm</F> from forward <F>ah</F> <delart>A</delart> <+pl>men</+pl> who was riding his motorcycle <F>um</F> came and <F>mm</F> they crashed <nvs>laughter</nvs>

Agreement

parents and one sister. <repair>My sister is</repair>, <F>er</F>, <./> <F>er</F>, my sister <agr>go</agr> to <delart>A</delart> university in Osaka. Because, <F>er</F>, I go to Waseda university

family. <F>Er, mm</F>, my mother is a housewife and, <F>mh,mm</F>, she <agr>enjoy</agr> her life, playing tennis or meeting with her friends. And my father

Tense

it. <A> <F>Oh</F>, really? And my father and mother, <F>ah</F>, also <prs> didn't </prs> drink, <F>ah</F>, alcohol. <./> <F>Mm</F>, <./> so, <F>ah</F>, <./> my or nine o'clock. So after <repeat>he</repeat>, he <agr>come</agr> back, and if I <prs>could</prs> go, I will go. <A> OK. What time do you think it

Aspect

 <ing> I'm belonging </ing> to, <F>urm</F>, Waseda University, of course, and <ing>I'm belong</ing> to, how can I say, literature, <./> I'm <elex> learning </elex>

And the lion in front of them <repair>is</repair> <./> seems to be <ing>smile</ing> to

<repair>hi</repair> them. So, <F>ah</F> <pst>they're</pst> very happy to see it and

Lexical Choice

choose <eart>the</eart> tie <F>mm</F> and <./> so the <F>mm, mm</F> the <F>mm</F>
<elex>buyer</elex> <F>mm</F> recommended <F>mm</F> <repeat>one</repeat>, one tie
<F>mm</F> the tie <F>mm</F> <cut>li</cut> <pst>looks</pst>

 Yes. Can be quick. <A> But not for you. <./> <eadv>Yes</eadv> <A> Never mind. It
took me much longer. <F>Uhh.</F> How long?

Omission

(file00017) <F>er</F>, <F>oh</F>, the title is, <F>urm</F>, it's about Scotland. Do you know
<delobj>IT</delobj>? <A> <F>Err</F>, I think I know <delobj>IT</delobj>. "Brave Heart".
Yes.

says that "OK I will have that". <repair>He</repair> maybe<F>uh</F>the waiter will pour
<delobj>WINE</delobj> to her wife's <repair>glasses</repair> glass, too. And in the restaurant,

Addition

village <repair>if we are</repair> if we meet at <delart>THE</delart> first time, we
<redcop>are</redcop> smile and <laughter> say "Hello", <A> <F>Mm</F> So <F>um</F> I
in

thirty. OK. <F>um</F> <repair>Could you</repair> can you ask him when he
<redcop>is</redcop> <agr>come</agr> back? <A> Yeah. And do you think you can come back

Word order

of</repeat>, <F>er</F>, every kinds of foods, <F>er</F>, Chinese, Japanese, and French, but
<@>I can't go so often French because the cost </@></?>. <A> <F>Huh-huh</F>.

living in the Island and living in Tokyo? <F>ah</F> <F>umm</F>
<WO>I feel in Tokyo lonely</WO> and <delprep>IN</delprep> the biggest city the people
<./><F>um</F>

for friends? <F>Oh</F>, yes. <./> <repeat>One</repeat>, one thing, <repeat>I</repeat>, I
bought <WO> for my friends New Year's calendar </WO>. <A> <F>Mhmm</F>. Very
beautiful.

 Good. <A> I'll be looking forward to it. Yeah. I'm <WO>forward to looking</WO> to
see you. OK. See you next Sunday. <A> OK.

go on a date? <F>Ahh</F>, <./> actually, <F>uhmm</F>, on next Sunday, <WO>I and my girlfriend</WO> will go to Disneyland. <nvs>laughter</nvs> <A> Disneyland? Yeah.

Advanced

(file00034) So, Naoki,<F>ah</F>can you tell me about yourself? Ok. So, I'm <F>um</F> <WO>junior Waseda University <F>um</F> commercial department</WO>, and now <F>umm</F>I'm major in <F>umm, umm</F>

Unclear

wan</repeat>, I want to change this ticket or <./> <repair>can I, can</repair>, <@>can you change me back the money</@>? <A> <F>Err</F>, I'm sorry, madam. We,

<repair>glasses</repair> glass, too. And in the restaurant, there is a piano and <@> the piano <repeat>is</repeat> is <F>um</F>is playing <repeat>by</repeat> by herself </@>. </task>
<followup>

Appendix 2: Some examples of 'TO' omission

<reform>my <./> older daughter <./> is another one</reform>, my older daughter lives <delprep>IN</delprep> another place. Yes. <F>mmm</F> My apartment is narrow. And it <delcop>IS</delcop> difficult

<A> <F>Mhm</F>. OK. So, <./> <F>mhm</F>, <./> the theater is <cut>Shin</cut>, <delprep>IN</delprep> Shinjuku, so how do you <./> get <eprep>to</eprep> there? <A> <F>Urr</F>, I'm

<A> <F>Mhm</F>. OK. So, <./> <F>mhm</F>, <./> the theater is <cut>Shin</cut>, <delprep>IN</delprep> Shinjuku, so how do you <./> get <eprep>to</eprep> there? <A> <F>Urr</F>, I'm

was <F>um</F> <repeat>ten</repeat> ten years old. And then after that, I stayed <delprep>IN</delprep> Japan for a while but then, <F>um</F> when I was

<reform>my <./> older daughter <./> is another one</reform>, my older daughter lives <delprep>IN</delprep> another place. Yes. <F>mmm</F> My apartment is narrow. And it <delcop>IS</delcop> difficult

Appendix 3: Some examples of 'TO' addition

let me see <./> <F>Ah- Oh</F>, I remembered. <repair>I was</repair> I <prs>went</prs>
<redprep>to</redprep> shopping in Shibuya and I <prs>bought</prs> some nice suits to wear
<eprep>in</eprep>

to go to <./> <repeat>lesson</repeat>, English lesson near <delart>THE</delart>zoo. So I said
<redprep>to</redprep> good-bye to her, <F>er</F>, <./> in front of the gate of
<delart>THE</delart>

<R>wh</R>, why did you choose Oregon? Many Waseda people <F>um</F> go
<redprep>to</redprep> there, so I choose there. And, some, <F>er</F> Waseda <./> University
<F>eh</F>

last week <F>ah</F> <repair>my mother goes to shopping,</repair> <F>ah</F> my mother went
<redprep>to</redprep> shopping <F>er</F> to a <JP>deper</JP> department store
<repeat>and</repeat> <F>ah</F> and <repeat>she want</repeat>,&br/>working on your computer? Working, <F>um</F> <repeat>I</repeat> <./> <F>uh</F> I
like <redprep>to</redprep> <F>ah</F> watching soccer game <./> in <F>er</F> Italy. <A>
<F>Umm</F> why not

Appendix 4: Some examples of 'IN' addition

you so hungry? <laughter>Yes</laughter>. <repeat>I</repeat>, <F>mm</F> I got up early
<F>mm</F> <redprep>in</redprep> this morning, so <./> <F>mm</F> <repeat>I, <F>um</F>
I</repeat>, I'm hungry. <A> I

you live in an apartment or a house? <F>Eh</F> I live <redprep>in</redprep> alone in an
apartment. <A> How do you like living alone?

References

Granger, S. (1998). Learner English on Computer. London & New York: Addison
Wesley Longman.

Maclay, H. and Osgood, C.E. (1959). Hesitation phenomena in spontaneous
English speech. *Word* 15: 19-44.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition* 14: 41-104.

Green, P.S. and Hecht, K. (1993) Pupil self-correction in oral communication in English as a foreign language. *System* 21: 151-163.

Olynak, M., D'Anglejan, A. and Sankoff, D. (1990). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. In Scarcella, R. et al.(Eds.). *Developing communicative competence in a second language*. New York: Newbury Publishers.