Familiarity with Catenative Verb Constructions through a JEFLL Corpus Analysis

能登原　祥　之

Yoshiyuki NOTOHARA

Abstract

This study investigates Japanese EFL learners' relative familiarities with thirteen catenative verb constructions (henceforth, CVC) through an analysis of *the JEFLL Corpus* (Tono, 2007). The results of 4-gram analyses show that learners tend to use SV, SVC, and SVO patterns in six timed-narrative essays. Furthermore, correspondence analysis reveals that learners are mainly familiar with *be+-ed, be+-ing, like+to*-infinitive, and *like+-ing*. On the other hand, they are unlikely to use the "accusative with infinitive" constructions (Jespersen, 1933, p.340) such as *like+*(person/thing)+ *to*-infinitive (SVO). These results suggest that on the basis of four familiar constructions, low and intermediate learners at the Japanese university level should gradually enrich their productive catenative verb constructions through pedagogical examples and also use them fluently.

1. Introduction

Given a timed-narrative essay, Japanese EFL learners, especially at the low and intermediate levels, are likely to repeat the same verbs with the same verb patterns. They probably do not know a sufficient amount of verbs in their minds, much less use their verb patterns accurately or fluently. However, they try to enrich and extend their own EFL knowledge at schools, although it might be fossilized easily. Here, based on the *dynamic usage-based model* (Langacker, 1987; 2000)[1], their dynamically developing interlanguage is focused on and treated as *learning-driven data* (Seidelhofer, 2002).

Notohara (2010) investigated Japanese EFL learners' verb patterns in terms of event schemata and sentence patterns through the analysis of a large learner corpus, *the JEFLL corpus* (Tono, 2007). After this study, their immature interlanguage was expected to appear in *links* (e.g. *bare/to-* in the infinitive clause, and *-ing/-en* in the participle clause, etc.) in CVCs.

This problem could lead them to repeatedly use inappropriate English verb patterns. To precisely describe this aspect of Japanese EFL learners' interlanguage, this study addresses how they use the links in CVCs. Based on the previous verb-based study (Notohara, 2010) mentioned above, learners' recurrent use of CVCs with the aforementioned links are described through *the JEFLL Corpus*.

2. Literature Review

2.1. Links in CVCs

Palmer (1988, p.173) classifies links in CVCs into four types in terms of form: (1) *bare* infinitive, (2) *to*-infinitive, (3) *-ing* form, and (4) *-en* form. Furthermore, in terms of diachronic semantic change, Nakao (2004, p.85) explains the shades of the meanings of links referring to the semantic relationships between catenative verbs and links. Table 1. summarises the features of four types of links in CVCs.

Table 1. *Links in CVCs* (*Palmer, 1988, p.173*)

Links	Forms	Meanings
(1) *bare* infinitive	*He helped **wash** up.*	without time lag / directionality [futurity, modality]
(2) *to*-infinitive	*He wants **to go** to London.*	time lag / directionality [futurity, modality]
(3) *-ing* form	*He keeps **talking** about it.*	progressiveness / generality [factivity]
(4) *-en* form	*He got **shot** in the riot.*	passiveness / perfectness

Recently, using a large corpus, *Bank of English,* Bullon, Krishanamurthy, Manning, & Todd (1990, pp.184-193) made a corpus-based list of catenative verbs with special reference to links. Some years later, some corpus-driven research (Francis, Hunston, & Manning, 1996; Hunston & Francis, 2000) refined catenative verb lists from the perspective of *Pattern Grammar*. It seems extremely laborious work for Japanese EFL learners to appropriately acquire all verbs in the lists in their daily lives due to the lack of English exposure. Specifically, low and intermediate Japanese EFL learners do not usually have a fixed linguistic knowledge about verb patterns. To make matters worse, they often neglect the importance of their semantic relationships between catenative verbs and links in learning English. Naturally, they tend to simply combine two verbs and repeat them without thinking about the semantic relationships. As a result, learners' knowledge about links is likely to be easily fossilized.

2.2. Two pedagogical approaches towards links in CVCs

Most pedagogical grammar at the high school level focuses on links and suggests learners should remember verb lists such as *-ing* form verbs, *to*-infinitive verbs, and *-en* form verbs. Thus, learners are gradually getting used to CVCs under a *links-based approach* (see Figure 1.). However, in this way, they have little chance to develop their awareness of the semantic relationships between catenative verbs and links. They might be in trouble then, especially when using links in CVCs if they do not remember the verb lists accurately because they are assumed to have their own vocabulary not as patterns but as individual words.

Here, the importance of semantic relationships between catenative verbs and links are revisited. Altenatively, the *catenative verb-based approach* should more often be considered when developing Japanese EFL learners' productive vocabulary. This is because this approach could develop learners' semantic sensitivities towards links in CVCs. Then their new sensitivities could lead them to acquire appropriate links in CVCs systematically and use CVCs fluently without any errors. These two pedagogical approaches can be illustrated as follows (Figure 1.):
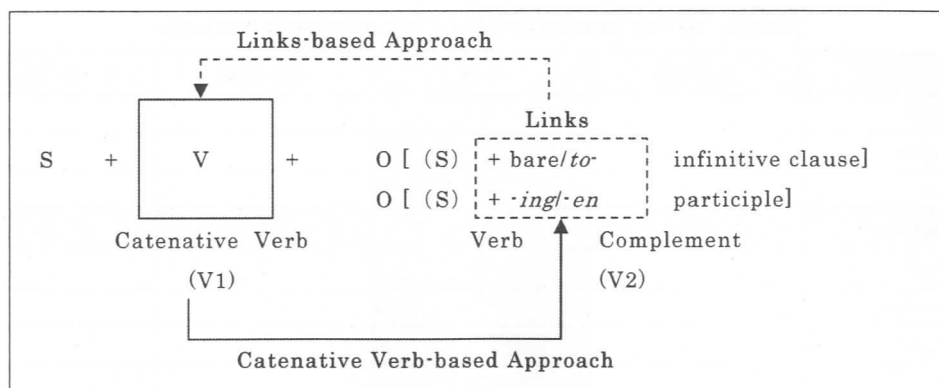
Figure 1. *Two Pedagogical Approaches towards Links in CVCs*

As Figure 1. shows, CVC can be categorized as an SVO pattern (Onions, 1971, p.7; Quirk, Greenbaum, Leech & Svartvik, 1972, p.837). In this pattern, links are described as parts of connections between V 1 and V 2 such as *bare/to-* in the infinitive clause and *-ing/-en* in the participle clause. With CVCs, learners could express two kinds of actions or situations at a time. Ideally speaking, they are expected to use CVCs while feeling the semantic connections between catenative verbs and links. Here is an example: If learners try to use *hope + to*-infinitive, it would be better for them to imagine the semantic connection between the futurity of a catenative verb, such as *hope* and the time lag and modality of a link, such as *to* in the infinitive. Developing such linguistic sensitivities towards semantic features of links in CVCs could encourage and enrich learner's creative language use. Thus, taking such a *catenative verb-based approach* (as Figure 1.) should be essential in teaching English expressions if we are to think much of communicative language teaching in Japan.

2. 3. Japanese EFL learners' familiarity with links in CVCs

Palmer (1988, pp.191-205) classifies catenative verbs (i.e., V 1) into nine types: (1) *Futurity,* (2) *Causation,* (3) *Report,* (4) *Perception,* (5) *Process,* (6) *Achievement,* (7) *Attitude,* (8) *Need,* and (9) *Appearance and Chance.* On the basis of his classification, the relationships between nine catenative verbs and four links can be summarised as seen in Table 2.

In terms of lexical collocation, Table 2. shows us that native speakers often use *to*-infinitive and *-ing* form with nine catenative verbs. As for links, Perception verbs (e.g., see) are assumed to be used with all four links. However, there are few empirical studies with large native English corpora, let alone a corpus-based one with Japanese learners' large English corpora. A small corpus-based study implies that Japanese EFL learners at the university level (esp. low and intermediate level) tend to use *to*-infinitive in timed narrative emails (Notohara, 2004). Learners could be exposed more often to *to*-infinitive in learning English even in input-poor EFL environments and acquire that form with great ease. The next frequently exposed link could be *-ing* form. On the other hand, most Japanese EFL learners are expected to be relatively less familiar with *bare*-infinitive and *-en* form. They might have some cognitive difficulties in grasping the "without time lag" of *bare*-infinitive and the passiveness and perfectness of *-en* form. To make matters worse, in some cases, these types of CVCs could have an accusative object (i.e., person or thing) between catenative verbs and links (e.g., (S) in Figure 1.). Such complicated CVCs must be difficult for Japanese EFL learners to acquire and use.

Table 2. *The Relationships between Catenative Verbs and Links*

	bare-inf	*to*-inf	*-ing* form	*-en* form
(1) Futurity		○		
(2) Causation	○	○		○
(3) Report		○	○	
(4) Perception	○	○	○	○
(5) Process		○	○	
(6) Achievement		○		
(7) Attitude		○	○	
(8) Need		○	○	
(9) Appearance and Chance		○		

A recent corpus-based research paper (e.g., Notohara, 2010) showed that in six narrative essays Japanese EFL learners at the junior and senior high school levels are likely to use five verbs: *be* (Occurrence/SVC), *have* (Possession/SVO), *like* (Emotion/SVO), *go* (Self-motion/SV), *and see* (Perception & Cognition/SVO). In this paper, among twenty variations between five verbs and four links, the following thirteen CVCs are selected and their relative familiarities with them will be surveyed through a JEFLL Corpus analysis: (1) *be+to*-infinitive, (2) *be+-ing,* (3) *be+-ed,* (4) *have+*(person)*+bare*-infinitive, (5) *have+*(thing)*+-ed,* (6) *like+to*-infinitive, (7) *like +* (person/thing) *+to*-infinitive, (8) *like+-ing,* (9) *go+-ing,* (10) *see+*(person)*+ bare* -infinitive, (11) *see +*(person/thing)*+to*-infinitive, (12) *see +* (person/thing) *+-ing,* (13) *see+* (person/thing)*+-ed.*

3. Method

3.1. Materials

The JEFLL Corpus (Tono, 2007) is one of the largest Japanese EFL learner corpora complied by Tono's research group. Six different timed-essays were given to 10,038 students at junior and senior high schools in Japan. In a period of twenty minutes, they were asked to write essays without any dictionaries. All free compositions were collected cross-sectionally. Descriptive statistics of the JEFLL Sub-corpus can be summarised as Table 3.

Table 3 . *Descriptive Statistics of the JEFLL Sub-Corpus*

	J 1 (n=1393)	J 2 (n=2635)	J 3 (n=1589)	S 1 (n=1255)	S 2 (n=1977)	S 3 (n=1189)	Total (n=10038)	*M* (*SD*)
Tokens	51149	159736	117766	91096	170555	78979	669281	111547 (46839.38)
Types	2639	6259	5161	4293	7847	4958	4958	31157 (1764.61)
TTR (%)	5.16	3.92	4.38	4.71	4.60	6.28	29.05	4.84 (0.81)
GRTTR*	11.67	15.66	15.04	14.22	19.00	17.64	93.23	15.54 (2.58)

*GRTTR(Guiraud' s Root TTR)= Types/$\sqrt{}$ Tokens (Malvern, Richards, Chipere, & Duran, 2004, p. 26)
*J=junior high, S=senior high

3. 2. Data Analysis

At the first stage, the 4-gram analyses of main four catenative verbs in thirteen CVCs were conducted with the web query system developed by Shogakukan Corpus Network (2008-present), to investigate the tendencies of sentence patterns produced by Japanese EFL learners. The analyses ranged from Left 1 to Right 2. Next, according to three identification criteria (e.g., thirteen CVCs) and eleven exclusion criteria (e.g., Japanese, misspellings, and ungrammatical patterns), the learners' thirteen target CVCs were identified by the author's naked eyes through concordances of six JEFLL Sub-Corpora. The identification work was conducted twice in the same way in order to be reliable. Their raw frequencies were normalized per 100,000 words and summarised in a 13×6 cross tabulation (see Appendix 1.). Finally, based on this cross tabulation, correspondence analysis between CVCs and grades was conducted using *SPSS 18.0* to investigate learners' relative familiarities with the thirteen CVCs.

4. Results & Discussion

4. 1. 4 -gram Analysis

As the results of 4-gram analyses, Japanese EFL learners' tendencies when using sentence patterns were investigated and are summarised in Table 4. Generally speaking, it could be said that learners are likely to use SVO patterns with *have, like,* and *see,* while they often use copula-based *be* with SVC pattern and *go* with SV pattern.

Table 4. *Tendencies of Sentence Patterns from 4 -gram Analyses*

Sentence Patterns	Verbs		
(1) SV	go		
(2) SVC	be		
(3) SVO	have	like	see

4. 2. Correspondence Analysis

Normalized relative frequencies of catenative verbs in the JEFLL Corpus are illustrated in Figure 2. As Figure 2. shows, *be*-based CVCs (e.g., *be+-ing* and *be+ed*) and *like*-based CVCs (e.g., *like+to*-infinitive and *like+-ing*) were found to be often used. Unexpectedly, the *be+-ed* construction was also frequently used. Most specifically, this CVC could be influenced by junior high English textbooks because learners from the second year of the junior high school level and above in JEFLL Corpus often use this type of CVC.
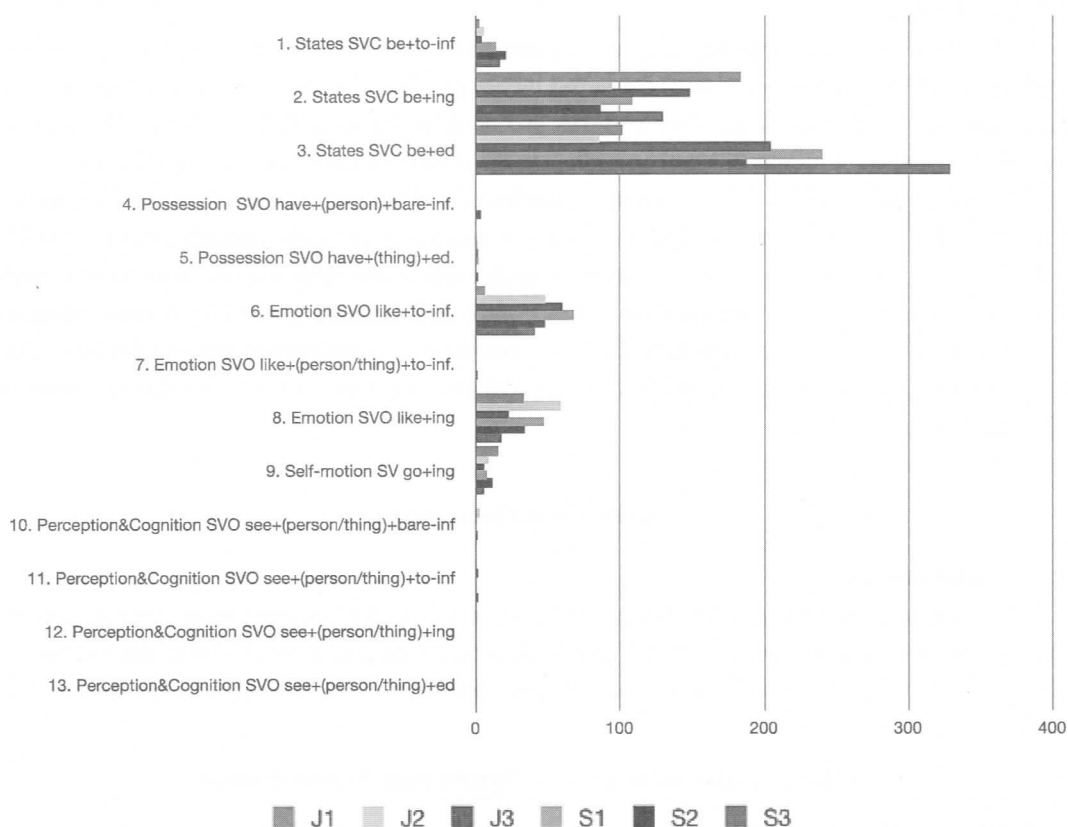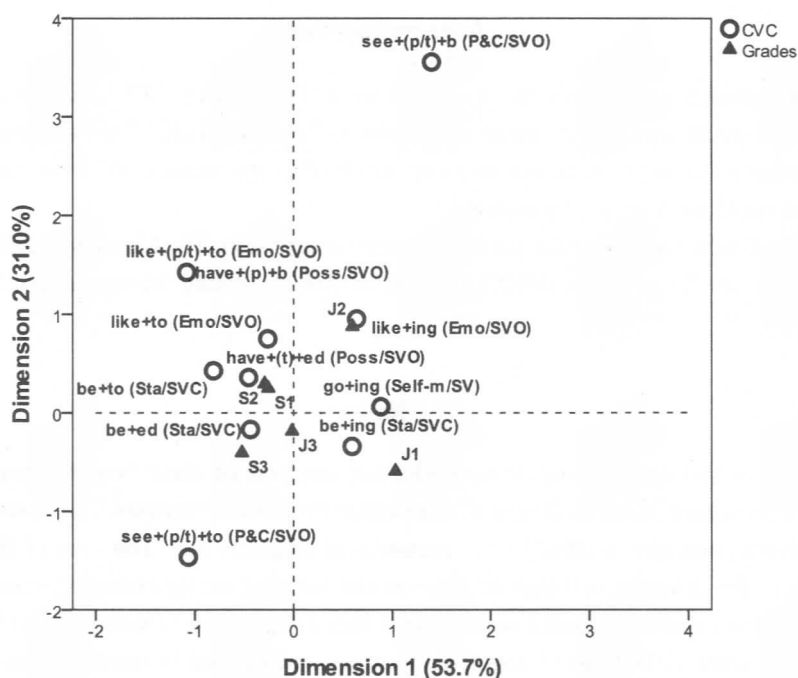
Figure 2. *Relative Frequencies of Catenative Verbs in JEFLL Corpus*
(Normalized per 100,000 words)

Based on this data of normalized frequencies, correspondence analysis was conducted using *SPSS 18.0.* As a result, a chi-square test for total inertia showed there was a statistically significant difference between CVCs and grades (χ^2 (60)=348.319 *p*=.000). The proportion of the inertia accounted for by two dimensions (Dimension 1: 53.7% and Dimension 2: 31.0%) was 84.7%. As Figure 3. shows, *be+-ing* (SVC) construction and *go+-ing* (SV) construction were relatively often used at all levels, especially at junior high school levels. As for *like+ to*-infinitive (SVO) construction, this was used at the senior high school levels. Surprisingly, *like+ing* (SVO) construction was often used by the second year student at junior high schools. From the data in the JEFLL Corpus, learners are familiar with this construction with the meaning of *generality or factivity*.

On the other hand, the SVO construction embedded with an accusative object was not found in the JEFLL Corpus such as *have+* (person) *+bare*-infinitive (SVO), *have +* (thing) *+-ed* (SVO), *like+*(person/thing) *+ to*-infinitive (SVO), *see+*(person/thing) *+bare*-infinitive(SVO), *see+*(person/thing) *+to*-infinitive(SVO), *see+* (person/thing) *+-ing* (SVO), and *see+* (person/thing) *+-ed* (SVO).

In terms of event schemata, learners in the JEFLL Corpus are relatively familiar with Occurrence *(be) and* Self-motion *(go)* schemata. The next frequently used schemata are Possession *(have)* and Emotion *(like).* However, Perception & Cognition *(see)* was relatively rare when used by learners in

*J=junior high, S=senior high, (p)=person, (t)=thing, (p/t)=person/thing

Figure 3. *Correspondence Analysis between Catenative Verbs and Grades*
(Normalized per 100,000 words)

the JEFLL Corpus. The tendencies of learners' relative familiarities in event schemata were strongly associated with not only learners' implicit or explicit event schemata in their minds, but also the cognitive loads due to the complexity of the sentence pattern (see Figure 3.).

4. 3. Limitations

The JEFLL corpus can be considered pseudo-longitudinal data. Therefore, it would be helpful for us to grasp the common tendencies of relative familiarities from this cross-sectional data (e.g., from most familiar to unfamiliar) at all levels. On the other hand, however, we must admit the fact that it is not possible for researchers to follow Japanese EFL learners' long-term interlanguage developments individually from the data here (e.g., from J 1 to S 3).

5. Conclusion

This study showed the Japanese EFL learners' relative familiarities of thirteen canonical CVCs from two perspectives: event schemata and sentence patterns. From the results of this study, the following approach to designing pedagogical examples for learners in teaching free composition is proposed: (1) Basic Level: *be* with Occurrence/SVC, *like* with Emotion/SVO, (2) Syntagmatic Extension Level: *see* with Perception & Cognition/SVO, *like* with Emotion/SVO, and (3) Paradigmatic Extension Level: other verbs with Action/SVO, Transfer /SVO, and Caused-motion/SVO.

Acknowledgements

Notes

1. Our characterization of schematic networks has emphasized their "static" properties, but it is important to regard them as dynamic, continually evolving structures. A schematic network is shaped, maintained, and modified by the pressures of language use. The locus of these pressures is coding, i.e. the interactive relationship−in the form of categorizing−between established conventional units and the specific usage events they are involved to sanction. In the final analysis, a schematic network is a set of cognitive routines, entrenched to varying degrees: despite our inevitable reifications, it is not something a speaker *has,* but rather than what he *does.* (Langacker, 1987, pp. 381-382)

References

Bullon, S., Krishanamurthy, R. Manning, E. & Todd, J. (1990). *Collins COBUILD English Grammar.* London: Harper Collins Publishers.

Francis, G., Hunston, S., & Manning, E. (1996). *Collins COBUILD Grammar Patterns 1: Verbs.* London: Harper Collins.

Hunston, S. & Francis, G. (1998). Verbs observed: A corpus-driven pedagogic grammar. *Applied Linguistics, 19,* 45-72.

Hunston, S. & Francis, G. (2000). *Pattern Grammar. A corpus-driven approach to the lexical grammar of English.* Amsterdam: John Benjamins.

Jespersen, O. (1933). *Essentials of English Grammar.* London: George Allen & Unwin.

Langacker, R.W. (1987). *Foundations of Cognitive Grammar. Volume I Theoretical Prerequisities.* Stanford: Stanford University Press.

Langacker, R.W. (2000). A dynamic usage-based model. In M. Barlow., & S. Kemmer. (Eds.) *Usage Based Models of Language.* (pp.1-63.). Stanford: CSLI Publications.

Malvern, D.D., Richards, B.J., Chipere, N. & Duran, P. (2004). *Lexical Diversity and Language Development. Quantification and Assessment.* New York: Palgrave.

Nakao, Y. (2004). *Tsunagi no Jyuuyousei-Eigo Kyoushi ni Motomerareru Senmonsei (1)*−[The Importance of Links: A required teacher professionality (1)−] In Michikazu Kaneda's Committee for the Publication Project (Eds.). (2004). *Studies in English Language Education: A Collection of Essays: In Commemoration of Professor Michikazu Kaneda.* (pp.85-102). Hiroshima: Keisuisha.

Notohara, Y. (2004). Data-Driven Learning no E-mail Writing Katsudo eno Dounyuuhou. [Integrating Data-Driven Learning into E-mail Writing Activities] *The CASELE Research Bulletin, 34,* 113-122.

Notohara, Y. (2009). E-mail ni Mirareru Nihonjin Eigo Gakushusha no Hinshutsu Doshi no Tokuchou. [Recurrent Verbs of Japanese Adult EFL Learners in Narrative E-mails.] *Bulletin of Hijiyama University, 16,* 29-39.

Notohara, Y. (2010). Nihonjinn Eigo Gakushusha no Ibento Suki-ma to Bunkei eno Shinmitsudo.— JEFLL Corpus no Bunseki o Tooshite—[Japanese EFL Learners' Familiarity with Event Schemata and Sentence Structures through a JEFLL Corpus Analysis.] *English Corpus Studies, 17,* 33-48.

Onions, C.T. (1971). *Modern English Syntax. New Edition of An Advanced English Syntax prepared from the author's materials by B.D.H. Miller.* London: Routledge.

Palmer, F.R. (1988). *The English Verb 2nd edition.* London: Longman.

Quirk, R., Greenbaum, S. & Leech, G. & Svartvik, J. (1972). *A Grammar of Contemporary English.* London: Longman.

Seidelhofer, B. (2002). Pedagogy and local learner corpora. Working with learning-driven data. In Granger, S., Hung, J. & Petch-Tyson, S. (2002). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching.* (pp.213-234). Amsterdam: John Benjamins.

Shogakukan Corpus Network. (2003-present). *The BNC Online.* Net Advance Inc. [Database] Available from Shougakukan Corpus Network Web site, http://www.corpora.jp/

Shogakukan Corpus Network. (2008-present). *The JEFLL Corpus* Net Advance Inc. [Database] Available from Shougakukan Corpus Network Web site, http://scn02.corpora.jp/~jefll03/jefll_top.html

Tono, Y. (2007). *Nihonjin Chuukousei Ichiman nin no Eigo Corpus "JEFLL Corpus"* [10,000 Japanese Junior and Senior High School Students' English Corpus "The JEFLL Corpus"] Tokyo: Shougakukan.

Appendix 1. *Relative Frequencies of Catenative Verbs in the JEFLL Corpus* (*Normalized per 100,000 words*)

Event Schema	SP	CVC	Grades								
			J 1	J 2	J 3	S 1	S 2	S 3	Total	*M*	*SD*
1. States	SVC	*be* + *to*-inf	1.96	5.01	3.40	14.27	21.11	17.00	79.00	10.46	8.03
2. States	SVC	*be* + ing	183.78	94.53	148.60	108.68	86.78	130.00	797.00	125.40	36.58
3. States	SVC	*be* + ed	101.66	85.77	204.64	240.41	187.63	329.00	1298.00	191.52	90.26
4. Possession	SVO	*have* + (p) + bare-inf	0.00	0.00	0.00	0.00	2.93	0.00	5.00	0.49	1.20
5. Possession	SVO	*have* + (p) + ed	0.00	0.63	0.85	1.10	0.00	1.00	4.00	0.60	0.49
6. Emotion	SVO	*like* + *to*-inf	5.87	48.20	60.29	68.06	48.08	41.00	336.00	45.25	21.60
7. Emotion	SVO	*like* + (p/t) + *to*-inf	0.00	0.00	0.00	0.00	0.59	0.00	1.00	0.10	0.24
8. Emotion	SVO	*like* + ing	33.24	58.85	22.93	47.20	34.01	18.00	257.00	35.71	15.19
9. Self-motion	SV	*go* + ing	15.64	8.76	5.94	7.68	11.73	5.00	61.00	9.13	3.97
10. Perception&Cognition	SVO	*see* + (p/t) + bare-inf	0.00	1.88	0.00	0.00	0.59	0.00	4.00	0.41	0.76
11. Perception&Cognition	SVO	*see* + (p/t) + *to*-inf	0.00	0.00	0.85	0.00	0.00	1.00	2.00	0.31	0.48
12. Perception&Cognition	SVO	*see* + (p/t) + ing	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13. Perception&Cognition	SVO	*see* + (p/t) + ed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total			342.15	303.63	447.50	487.40	393.45	542.00	2844.00	419.36	178.79
M			26.32	23.36	34.42	37.49	30.27	41.69	218.77		
SD			55.15	35.46	66.31	69.80	53.92	93.48	396.53		

*SP=Sentence Patterns, CVC=Catenative Verb Constructions, p=person, t=thing, p/t=person/thing

能登原祥之（言語文化学科）