

# The Dimensions of Verbs of Saying Based on Rasch Modeling and Corpus Analysis\*

Heechul Lee · Jaewoo Shim · Nayu Kim\*\*

(Chonbuk National University)

**Lee, Heechul, Shim, Jaewoo & Kim, Nayu. (2014). The Dimensions of Verbs of Saying Based on Rasch Modeling and Corpus Analysis.** *The Linguistic Association of Korea Journal*, 22(2), 133-153. The purpose of this study is to investigate if there are any hierarchies of difficulty in the learnability of verbs of saying on the basis of Rasch modeling and, if any, to examine what may cause the difficulty hierarchies, using corpus analysis. As noted by Van Valin and LaPolla (1997), such verbs of saying as  *speak, say, talk* and  *tell* can be covered by their shared semantic decomposition which represents  *expressing something to someone in some language*. Although those verbs have parts of their meaning in common, they are different from one another in their syntactic behaviors and semantic selectional restrictions. This research analyzes 121 high school students' acceptability judgements of 24 sentences in which the verbs of saying occur in different syntactic or semantic environments. Then the paper puts forward the reasons of difficult items with respect to corpus analysis and provides educational implications on how to teach those verbs. With respect to item difficulty, this study has found that the verb  *speak* is unmarked and, accordingly, the easiest item, on the one hand, and the verb  *talk* is more difficult than the verb  *speak*, on the other. Regarding educational implications, it has produced a diagram which summarizes the similarities and differences of the verbs of saying.

**Key Words:** verbs of saying, Rasch model, corpus analysis

---

\* This paper was supported by research funds of Chonbuk National University. Heechul Lee received the fund in 2014 and Jaewoo Shim in 2013.

\*\* The first author is Heechul Lee, the corresponding author, Jaewoo Shim, and the third author, Nayu Kim.

## 1. Introduction

Learners of English as a second language use verbs of saying (i.e., *speaking, saying, telling,* and *talk*) in everyday communication (Culicover, 2009). The Korean translations of those verbs do not seem to distinguish them from one another with respect to their meanings. In terms of syntactic and semantic aspects of the verbs of saying, however, there exist some sharp differences among them. Nevertheless, their distinctive behaviors (Van Valin & Lapolla, 1997) are hardly taught and even may be foreign to some teachers of English unless they have studied the syntactic and semantic restrictions of these verbs. With the increased interest in focus-on-form activities, the frequency of learner errors on the uses of verbs of saying may bother some teachers of English and prompt them to wonder if there is any remedy for those learner errors. For example, in the unacceptable sentence, *\*He spoke that his younger sister had lost her doll*, the syntactic restrictions of the verb *speaking* are violated and thus the sentence is grammatically wrong.

In recent years, corpus-based analyses of linguistic phenomena have been made available to find typical patterns of linguistic uses. The corpus tools may also be used for investigating samples of the verbs of saying that are accessed by learners through their textbooks and study materials. The cross-examination of their input learners of English receive from their textbooks and study materials would help researchers trace some sources of errors concerning the verbs of saying.

This study aims to answer the following questions;

1. Which verbs of saying do high school learners of English feel the most difficult or easiest?
2. What corpus data support the hierarchy of item difficulties?

## 2. Literature Review

Verbs of saying in English are classed as activity verbs and their complexity is shown by Wierzbicka (1987), who divides them into thirty-eight subclasses. Following Van Valin and Wilkins (1993), Van Valin and LaPolla (1997) try to

capture similarities of such verbs as *speak*, *say*, *talk*, *tell*, and *discuss* by positing a more complex lexical decomposition, as follows:

$$(1) \text{ do}' (x, [\mathbf{express}(\alpha).\mathbf{to}.\beta).\mathbf{in.language}.\gamma](x,y)]^1)$$

Informally put, (1) shows that the verbs mean *to express something to someone in some language* and have parts of their meaning in common. On the other hand, they differ from one another in their syntactic behaviors and in what kind of nouns they can take as their object. We will restrict our investigation to four verbs of saying, *speak*, *say*, *tell*, and *talk* in this study.

Let us first investigate their syntactic phenomena. In the following examples, the verbs of saying occur together with  $\alpha$ .

- (2) *that*-clause occurring as  $\alpha$
- a. \*Angie spoke that it would snow.
  - b. Angie said that it would snow.
  - c. \*Angie talked that it would snow.
  - d. \*Angie told that it would snow.

Such sentences as (2a), (2c), and (2d) are unacceptable while (2b) is acceptable. Only the verb *say* can take an indirect discourse complement as its immediately following object when  $\alpha$  alone is expressed. In Vorlat's (1982) terms, the verbs *speak* and *talk* do not occur with the message (expressed as *that*-clause here) whereas the verbs *say* and *tell* do. (2d) is unacceptable because the verb *tell* must be followed by an indirect object (Williams, 2004; Huddleston & Pullum, 2002).

Now let us consider the sentences in which the verbs of saying occur together with  $\alpha$  and  $\beta$ , as follows:

- (3)  $\beta$  followed by  $\alpha$
- a. \*Stephen spoke Mary that it would snow.
  - b. \*Stephen said Mary that it would snow.

---

1)  $\alpha$  means the content of an utterance,  $\beta$  the addressee, and  $\gamma$  the language used.

- c. \*Stephen talked Mary that it would snow.
- d. Stephen told Mary that it would snow.

As represented by the asterisks, such sentences as (3a), (3b), and (3c) are inappropriate whereas the sentence (3d) is appropriate. The only verb among the four that can take  $\beta$  without a preposition is the verb *tell* and the others cannot. The verb *tell* can occur when it is followed by an indirect personal object and subsequently by a *that*-clause.

It is time to look into the sentences in which the verbs of saying occur together with  $\alpha$  and  $\beta$  as in (3), but this time  $\beta$  with a preposition, as follows:

- (4)  $\beta$  followed by  $\alpha$ ,  $\beta$  with a preposition
  - a. \*Jessica spoke to Bob that Joan would arrive soon.
  - b. Jessica said to Bob that Joan would arrive soon.
  - c. \*Jessica talked to Bob that Joan would arrive soon.
  - d. \*Jessica told to Bob that Joan would arrive soon.

As seen above, such sentences as (4a), (4c), and (4d) are ungrammatical while (4b) is grammatical. The reason that (4a) and (4c) are incorrect is that the verbs *speak* and *talk* cannot take a *that*-clause as  $\alpha$ , as shown in (2). There is nothing wrong with the verbs taking a personal noun with a preposition, as follows:

- (5)  $\beta$  with a preposition
  - a. Peter spoke to Alissa.
  - b. Peter talked to Alissa.

The contrast between the grammaticality of the examples in (5) and the ungrammaticality of the examples (4a) and (4c) shows that the verbs *speak* and *talk* cannot take a *that*-clause as their complement, as in (4a) and (4c). The reason for the ungrammaticality of (4d) is that the personal noun *Bob* does not need a preposition because it occurs as an indirect object immediately following the verb, which is shown in (3d).

The investigation of (2) to (5) leads to a conclusion that both verbs of *say* and

*tell* foreground the speaker and the message whereas the verbs *speak* and *talk* focus on the speaker and the communicative event itself, as stated in Vorlat (1982). The basic difference between the two pairs of verbs is matched by their syntactic restriction: in their prototypical uses the verbs *say* and *tell* are transitive while the verbs *speak* and *talk* are intransitive (Dirven, 1982). It is evidenced by Vorlat's (1982) corpus-based research showing the high figures of the verbs *talk* and *speak* functioning as one-place predicates occurring with the speaker as subject.

Let us examine the semantic aspects of  $\alpha$  which may determine whether  $\alpha$  can occur in the form of a noun phrase as an object of the verbs of saying, as follows:

(6) NP as  $\alpha$

- a. \*John spoke a story about Sam.
- b. \*John said a story about Sam.
- c. \*John talked a story about Sam.
- d. John told a story about Sam.

As shown above, the sentences (6a), (6b), and (6c) are incorrect whereas (6d) is correct. Only the verb *tell* among the four can take as its object what is so called utterance nouns, e.g. *story*, *joke*, *rumor*, *statement*, as in *tell a story about Sam* (Van Valin & LaPolla, 1997).

Let us take a look at the sentences in which  $\alpha$  is realized in the form of a noun phrase, but this time with a different semantic content from the one in (6) above, as follows:

(7) NP as  $\alpha$ ,  $\alpha$  containing a metalinguistic noun

- a. Raymond spoke only a few words to Jane.
- b. Raymond said only a few words to Jane.
- c. \*Raymond talked only a few words to Jane.
- d. \*Raymond told only a few words to Jane.

Sentences (7a) and (7b) are acceptable, but (7c) and (7d) are unacceptable. The verbs *speak* and *say* can take what is so called a metalinguistic noun, e.g.

*word, syllable*, as their object, as in *speak only a few words* or *say only a few words* (Van Valin & LaPolla, 1997) Examples (6) and (7) show that the (un)acceptability of those is determined by the semantic content of the noun phrase occurring as the object of the verbs of saying.

Let us include the examples in which a language name occurs as the object of the verbs of saying in our consideration, as follows:

- (8) language name as  $\gamma$  without a preposition
- a. Ken spoke French.
  - b. \*Ken said French.
  - c. Ken talked French.
  - d. \*Ken told French.

It is well known that the verb *speak* can take a language name as its object, as (8a) shows. On the other hand, it is interesting, especially to the non-native speakers of English, that the verb *talk* can also take a language name as its object, as in (8c). As seen in (8b) and (8d), the verbs *say* and *tell* cannot take a language name as their object.

With the 24 sentences shown above with regard to their (un)grammaticality in (2), (3), (4), (6), (7), and (8), a survey questionnaire was prepared. In the survey (refer to Appendix), however, the sentences are mixed up across different grammatical or semantic categories displayed here so that the subjects were prevented from guessing what they are tested for. They were instructed to mark O in the parenthesis at the end of each sentence if they think the sentence is acceptable and to mark X if they think that the sentence is unacceptable.

### 3. Research Procedure

#### 3.1 Participants

121 male subjects, all of whom were 12th grade high school students in a city, Korea, participated in this research. The subjects were studying their English, following the 8<sup>th</sup> government-led English education curriculum and

taking four to five credit hours a semester over the course of four consecutive semesters. They were typical EFL learners of English, rarely exposed to any authentic uses of English beyond their classrooms. None of the subjects stayed more than two months in any English speaking country.

### 3.2 Instrument

The proficiency test for verbs of saying consisted of a total of 24 sentences with each having a verb of saying that included *speak*, *say*, *tell* and *talk*. The subjects were given approximately one minute for each sentence. One of the researchers oversaw the pencil and paper test.

### 3.3 Data analysis

The subjects' responses to the set of 24 items were coded and entered as '1's for correct answers and '0's for incorrect answers in Excel version 2010 before they were transported to Winsteps, which is frequently used for Rasch analysis of dichotomous data (Linacre, 2009). The program yields statistics and graphs for analyzing test items including misfit items, item difficulty, item map, and summary statistics (McNamara, 1996).

The Rasch analysis measurement model changes empirical data such as subjects' responses into a logit scale ranging anywhere from -4 to +4 (Bond & Fox, 2007). On the logit scale, the average performance is set as zero. As the number unit on the scale increases, so does the difficulty associated with items (Li & Olejnik, 1997). The data resulted from the Rasch analysis applied to Antconc 3.2.4w, the corpus analysis program. The characteristics of usage of verbs of saying by the subjects were investigated by the functions of concordance, cluster and word list in the program.

### 3.4 Corpus data set

The corpus data set in this research study contained 731,925 words, which were from an array of 42 high school English text books and Education Broadcasting System (EBS) English reference books for 10<sup>th</sup> and 11<sup>th</sup> grade students. English text books involved with this study have been taught general

high school students across the country and EBS reference book have been used as the auxiliary text books in most high school English class. Thus corpus data set from these materials is significantly authentic in that the subjects might have been exposed to them subsequently for two years as well as learned them in priority (Conrad, 2000).

## 4. Results and Discussions

### 4.1 Reliability of test items and persons

The item reliability in Rasch modelling refers to the degree of consistency observed in a similar group of students. The item reliability of test results was .94, indicating high levels of agreement would be expected from another similar proficiency group. The mean of raw score was 69.5 with the standard deviation of 22.4. The maximum score was 99.0 and the minimum one was 33.0. The item difficulty ranged from -1.27 to 1.44, and the misfit index of Infit ZSTD ranged from -3.0 to 3.5.

Figure 1. Summary of Measured Items

SUMMARY OF 24 MEASURED ITEMS								
	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	69.5	121.0	.00	.21	.99	.0	1.01	.1
S. D.	22.4	.0	.89	.02	.11	1.3	.18	1.5
MAX.	99.0	121.0	1.44	.24	1.31	3.5	1.62	4.7
MIN.	33.0	121.0	-1.27	.19	.80	-3.0	.76	-2.9
REAL RMSE	.21	ADJ. SD	.87	SEPARATION	4.06	ITEM	RELIABILITY	.94
MODEL RMSE	.21	ADJ. SD	.87	SEPARATION	4.14	ITEM	RELIABILITY	.94
S.E. OF ITEM MEAN = .19								
UMEAN=.000 USCALE=1.000								
ITEM RAW SCORE-TO-MEASURE CORRELATION = -1.00								
2904 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 3296.02 with 2760 d.f. p=.0000								

### 4.2 Analysis of misfit items

The fit statistics (Figure 2) indicated that items 10, 17 and 22 did not fit the Rasch model as explained above, while most items followed the Rasch model expectations.



- (9) Ken talked French. (item No. 10)
- (10) \*Angie talked that it would snow. (item No. 17)
- (11) \*Stephen spoke Mary that it would snow. (item No. 22)

Figure 2. Item Statistics–Misfit Order

ITEM STATISTICS: MISFIT ORDER													
ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIIT MNSQ	OUTFIT ZSTD MNSQ	IPT-MEASURE ZSTD CORR.	EXACT EXP.	MATCH OBS%	ITEM			
10	38	121	1.23	.20	1.31	3.5	1.62	4.7	A-.25	.28	62.0	69.5	talk(10)
8	33	121	1.44	.21	1.07	.8	1.21	1.5	B.13	.26	72.7	73.0	tell(8)
5	41	121	1.11	.20	1.11	1.4	1.21	1.9	C.10	.28	64.5	67.7	talk(5)
11	37	121	1.27	.21	1.09	1.1	1.17	1.4	D.12	.27	67.8	70.1	say(11)
24	66	121	.17	.19	1.11	1.9	1.13	1.9	E.14	.31	57.9	63.6	speak(24)
2	90	121	-.80	.22	1.11	1.0	1.06	.5	F.15	.28	70.2	75.1	tell(2)
9	52	121	.69	.19	1.03	.5	1.07	1.0	G.25	.30	63.6	63.3	talk(9)
7	88	121	-.70	.21	1.06	.7	1.03	.3	H.21	.29	70.2	73.8	say(7)
3	62	121	.32	.19	1.01	.3	1.03	.6	I.28	.31	59.5	63.0	talk(3)
13	75	121	-.16	.20	1.03	.4	1.02	.3	J.27	.30	66.9	66.3	speak(13)
12	53	121	.65	.19	1.02	.4	1.01	.1	K.28	.30	61.2	63.1	say(12)
14	88	121	-.70	.21	1.01	.1	.99	.0	L.28	.29	75.2	73.8	say(14)
20	93	121	-.94	.22	.99	.0	.99	.0	I.28	.28	80.2	77.3	tell(20)
21	93	121	-.94	.22	.97	-.2	.98	-.1	k.32	.28	80.2	77.3	say(21)
23	39	121	1.19	.20	.97	-.3	.95	-.5	j.32	.28	69.4	68.9	tell(23)
16	53	121	.65	.19	.97	-.6	.94	-.8	i.36	.30	66.1	63.1	tell(16)
15	99	121	-1.27	.24	.93	-.4	.87	-.6	h.36	.26	83.5	81.9	speak(15)
1	85	121	-.57	.21	.91	-1.0	.87	-1.1	g.43	.29	71.1	71.8	speak(1)
18	46	121	.91	.20	.91	-1.4	.90	-1.2	f.42	.29	70.2	65.2	talk(18)
4	98	121	-1.21	.24	.91	-.6	.84	-.9	e.40	.26	82.6	81.1	say(4)
6	89	121	-.75	.22	.88	-1.2	.86	-1.1	d.46	.29	77.7	74.5	speak(6)
19	97	121	-1.15	.24	.86	-1.1	.84	-.9	c.46	.26	81.8	80.4	tell(19)
17	70	121	.03	.19	.83	-.3	.81	-.2	b.56	.31	74.4	64.5	talk(17)
22	82	121	-.44	.20	.80	-2.6	.76	-2.5	a.59	.30	76.9	69.9	speak(22)
MEAN	69.5	121.0	.00	.21	.99	.0	1.01	.1			71.1	70.8	
S.D.	22.4	.0	.89	.02	.11	1.3	.18	1.5			7.4	6.0	

The result shows that the subjects were not consistent or sure whether the forms of  $\{TALK + \gamma(LANGUAGE)\}$ ,  $\{TALK + \alpha(That\text{-}clause)\}$  and  $\{SPEAK + \beta(Addressee) + \alpha(That\text{-}clause)\}$  were correct or not. In an attempt to relate this inconsistency to frequencies of those forms in the textbooks, the verbs *talk* and *speak* were searched on the corpus data set (Table 1).

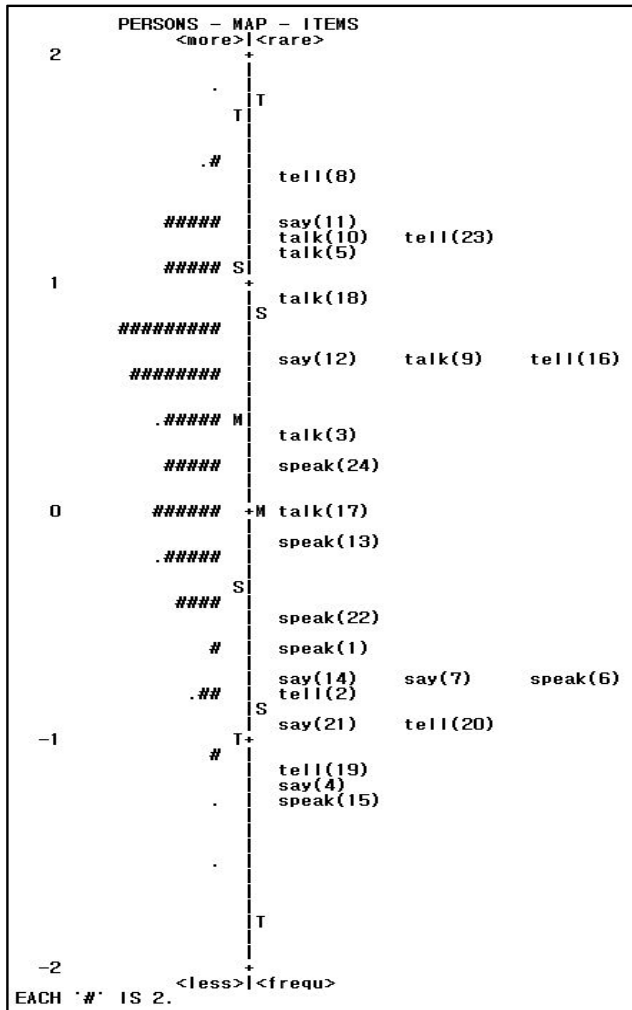
Table 1. Verbs *talk* and *speak* from Corpus

verb	verb (frequency)	verb (number of cluster: two words)	verb phrase (frequency)
TALK	talk(162)	talk(43)	TALK about(82) TALK to(77)
	talked(26)	talked(9)	
	talking(77)	talking(26)	
SPEAK	speak(96)	speak(53)	SPEAK about(18) SPEAK to(18) SPEAK with(16)
	spoke(37)	spoke(22)	
	spoken(46)	spoken(20)	
	speaking(67)	speaking(41)	
the corpus scope: 731,925 words			

The verb *talk* mostly collocates with the prepositions *about* and *to*. There were no cases of [TALK+  $\gamma$  (LANGUAGE)] and [TALK+  $\alpha$  (*that*-clause)]. In addition, in corpus data set the verb *speak* did not contain any examples of the form of [SPEAK+  $\beta$  (Addressee)] at all, whose typical collocations, though, consist of the prepositions *about*, *with* and *to*. In other words, the subjects may have not been exposed to those forms through their texts. Thus regardless of their language ability, they may have experienced some difficulty identifying correct forms.

### 4.3 Analysis of the most difficult and the easiest items

Figure 3. Item Map



In the item map above provided in the Winsteps, item difficulty levels were matched with subjects' ability levels along the logit scale. In this particular analysis of verbs of saying, the map showed most items within 1 and -1 logit scores with the mean score of zero. The following analyses of the most difficult items and the easiest items were based on the item hierarchy reported in the

Winsteps.

In the item map, it was found that items 4, 15 and 19 were the easiest ones. In particular, among the six items containing verb *speak* in the instrument, five items were below the mean. That is, the items 1, 6, 13, 15 and 22, in which the verb *speak* occurs, were marked correctly by most subjects. It could be directly related to the selectional properties of the verb *speak* by Van Valin and LaPolla (1997).

- (12) a. Sandy spoke but a few words.  $y=\alpha$   
 b. Sandy spoke to Kim.  $y=\beta$   
 c. Sandy spoke Telegu.  $y=\gamma$   
 d. Sandy spoke.

(Van Valin & LaPolla, 1997)

In the theory, the verb *speak* could occur with the argument  $x$  alone and none of the internal variables as in (12d). The verb *speak* is the only verb among the four that can occur with each of the three internal semantic variables. Including the verb *speak*, the three easiest items in the test were as follows:

- (13) Angie said that it would snow. (item No. 4)  
 (14) Ken spoke French. (item No. 15)  
 (15) \*Ken told French. (item No. 19)

Most subjects got correct the items containing these forms: [SAY+  $\alpha$  (*that*-clause)] and [SPEAK+  $\gamma$  (LANGUAGE)]. In order to identify the relation between the result and the corpus data, the frequency of those forms was examined.

Table 2. Verbs *say* and *speak* from Corpus

verb	frequency of verb	cluster : frequency
SAY	1,160	SAY+ $\alpha$ ( <i>that</i> -clause) : 209
SPEAK	246	SPEAK+ $\gamma$ (LANGUAGE): 15
the corpus scope: 731,925 words		

The frequency of the form of [SAY+  $\alpha$  (*that*-clause)] was 209 out of the total occurrences of the verb *say*, 1,160. As for [SPEAK+  $\gamma$  (LANGUAGE)], there were 15 hits out of 246. With the applications of Winsteps, the measure order was extracted in order to identify the differences among the items and the result was as follows:

Figure 4. Item Statistics–Measure Oder

ITEM STATISTICS: MEASURE ORDER													
ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	IN FIT MNSQ	OUT FIT ZSTD	PT-MEASURE CORR.	EXACT EXP.	MATCH OBS%	ITEM			
8	33	121	1.44	.21	1.07	.81	.21	1.51	.13	.261	72.7	73.01	tell(8)
11	37	121	1.27	.21	1.09	1.11	.17	1.41	.12	.271	67.8	70.11	say(11)
10	38	121	1.23	.20	1.31	3.51	.62	4.71	-.25	.281	62.0	69.51	talk(10)
23	39	121	1.19	.20	1.37	-.31	.95	-.51	.32	.281	69.4	68.91	tell(23)
5	41	121	1.11	.20	1.11	1.41	.21	1.91	.10	.281	64.5	67.71	talk(5)
18	46	121	.91	.20	1.31	-1.41	.90	-1.21	.42	.291	70.2	65.21	talk(18)
9	52	121	.69	.19	1.03	.51	1.07	1.01	.25	.301	63.6	63.31	talk(9)
12	53	121	.65	.19	1.02	.41	1.01	.11	.28	.301	61.2	63.11	say(12)
16	53	121	.65	.19	1.37	-.61	.94	-.81	.36	.301	66.1	63.11	tell(16)
3	62	121	.32	.19	1.01	.31	1.03	.61	.28	.311	59.5	63.01	talk(3)
24	66	121	.17	.19	1.11	1.91	.13	1.91	.14	.311	57.9	63.61	speak(24)
17	70	121	.03	.19	1.33	-3.01	.81	-2.91	.56	.311	74.4	64.51	talk(17)
13	75	121	-.16	.20	1.03	.41	1.02	.31	.27	.301	66.9	66.31	speak(13)
22	82	121	-.44	.20	1.30	-2.61	.76	-2.51	.59	.301	76.9	69.91	speak(22)
1	85	121	-.57	.21	1.31	-1.01	.87	-1.11	.43	.291	71.1	71.81	speak(1)
7	88	121	-.70	.21	1.06	.71	1.03	.31	.21	.291	70.2	73.81	say(7)
14	88	121	-.70	.21	1.01	.11	.99	.01	.28	.291	75.2	73.81	say(14)
6	89	121	-.75	.22	1.38	-1.21	.86	-1.11	.46	.291	77.7	74.51	speak(6)
2	90	121	-.80	.22	1.11	1.01	1.06	.51	.15	.281	70.2	75.11	tell(2)
20	93	121	-.94	.22	1.39	.01	.99	.01	.28	.281	80.2	77.31	tell(20)
21	93	121	-.94	.22	1.37	-.21	.98	-.11	.32	.281	80.2	77.31	say(21)
19	97	121	-1.15	.24	1.36	-1.11	.84	-.91	.46	.261	81.8	80.41	tell(19)
4	98	121	-1.21	.24	1.31	-.61	.84	-.91	.40	.261	82.6	81.11	say(4)
15	99	121	-1.27	.24	1.33	-.41	.87	-.61	.36	.261	83.5	81.91	speak(15)
MEAN	69.5	121.0	.00	.21	1.39	.01	1.01	.11			71.1	70.81	
S.D.	22.4	.0	.89	.02	1.11	1.31	.18	1.51			7.4	6.01	

Items 8, 11, 10 and 23 were most difficult to answer and items 4, 5 and 19 were so easy that everyone got them correct (Figure 4). The most difficult items were as follows:

- (16) \*Angie told that it would snow.(item No. 8)
- (17) Ken talked French. (item No. 10)
- (18) \*Stephen said Mary that it would snow. (item No. 11)
- (19) \*Jessica told to Bob that Joan would arrive soon. (item No. 23)

The above result shows that the forms of [TELL+  $\alpha$  (*that*-clause)] and [TELL+  $\beta$  (to Addressee)] were considered correct by the subjects, though the forms were actually wrong. The incorrect perceptions of the form of [SAY+  $\beta$  (Addressee)+  $\alpha$  (*that*-clause)] were also observed. In addition, the occurrences of

the verbs *tell* and *say* were examined in the corpus data set as presented in Table 3.

The verb *tell* cannot occur in the structures of [TELL+ $\alpha$ (*that*-clause)] and [TELL+ $\beta$ (To Addressee)]. However, in the analysis of the corpus data set, those forms could be found by 10 hits and by 7 hits, respectively.

- (20)-a. We have been told that a problem cannot be solved at the level where it began.
- b. At the medical center, Fat’s mother, Teresa, was told that her daughter would probably die within a few days.
- c. The customers were told that every time they had their car washed, their loyalty card would be stamped.
- d. The children were also told that if they would wait for 15 minutes, they could have two instead of one.

(from the corpus data)

Table 3. Verbs *tell* and *say* from Corpus

verb	verb (frequency)	verb (number of cluster)	verb phrase (frequency)	
TELL	tell(240)	tell(55)	TELL me(110)	TELL that(10)
	told(243)	told(56)	TELL the(40)	TELL to(7)
	telling(59)	telling(24)		
SAY	say(464)	say(148)	SAY that(209)	SAY the(26)
	said(575)	said(211)	SAY to(69)	SAY a(4)
	saying(121)	saying(67)		SAY an(3)
the corpus scope: 731,925 words				
number of cluster: two words				

However, most subjects may be familiar with the form of [TELL+ $\alpha$ (*that*-clause)] given that the forms of [BE Told+ $\alpha$ (*that*-clause)] occurs frequently in the passive voice.

- (21) a. We are often told to drive defensively.  
 b. A nurse, who had been told to say good-bye to Carlie on Friday at the end of their shift, was shocked to see Russ still there.  
 c. The subjects are told to do their tasks within a certain time.  
 (from the corpus data set)

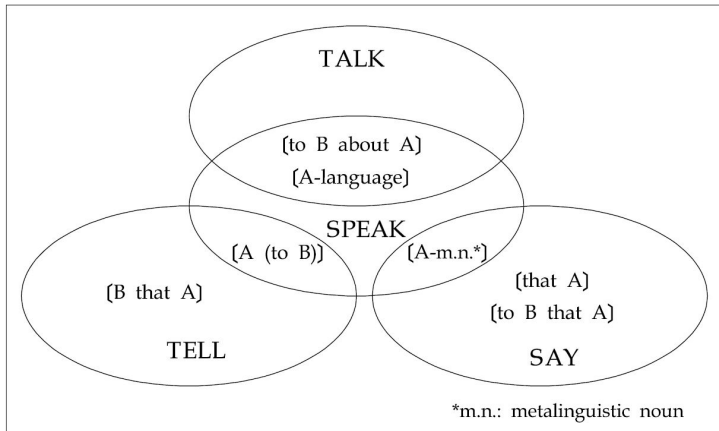
As the result of the corpus analysis above shows, the form of [be TOLD+to-Infinitive] occurs in the passive voice. Since the subjects are familiar with the verb *tell* followed by *to*-Infinitive, they may have been mistaken the form of [be TOLD+to-Infinitive] for the form of [TELL+to  $\beta$ ].

- (22)-a. "There now," said the gods, "all your wishes are granted."  
 b. "I want to repay you," said the nobleman.  
 c. "The need of our passengers is always given top priority," said a spokesman for the British airline.  
 (from the corpus data set)

The form of [SAY+ $\beta$ (Addressee without a preposition)+a(*that*-clause)] cannot occur. But in a direct speech, the verb *say* would induce the inversion of subject-verb as in (22). Accordingly, if the teacher does not explain the actual differences between the seemingly similar forms, the subjects might be confused about the two forms.

The above analysis of verbs of saying reveals their syntactic and semantic aspects. It is necessary to teach syntactic and semantic features of verbs of saying clearly so that Korean students may not be confused about the verbs bearing apparently the same meaning. The following diagram in Figure 5 outlines the similarities and differences among the four verbs and suggests a way of teaching them.

Figure 5. Diagram of Verbs of Saying



The above diagram shows that the verb *speak* shares some similarities with all the other three verbs, *talk*, *tell*, and *say* whereas each of the three verbs have something in common only with the verb *speak*. In that line, it may be possible to say that the verb *speak* is the most unmarked verb and others are marked. In turn, then, it may sound more reasonable to call them verbs of speaking, not verbs of saying.

As shown in Figure 3, the 24 items can be divided into four groups according to their item difficulty, 1) the most difficult items, 2) difficult items, 3) easy items, and 4) the easiest items. In the groups of the easiest items and easy items, there are no sentences in which the verb *talk* occurs. On the other hand, in the group of the most difficult items, there are no sentences in which the verb *speak* occurs. Therefore, it is possible to say that the verb *talk* is neither the easiest nor easy item whereas the verb *speak* is not the most difficult item.

When the groups of the difficult and easy items are examined carefully, it is found that most items in the difficult group are the sentences in which the verb *talk* occurs and that most items in the easy group are the sentences in which the verb *speak* occurs. Therefore, it can be said that there exists a difficulty hierarchy between the verbs *talk* and *speak*, which is the characteristic that can be better found by the Winsteps analysis with respect to the Rasch model. The hierarchy is represented as follows:



(23) talk > speak

The above difficulty hierarchy is consonant with the argument that the verb *speak* is unmarked.

## 5. Conclusion

The findings of this study reveal that the subjects, the high school students, had some difficulty judging the (un)acceptability of the sentences in which the verbs of saying occur. As far as the verb *tell* is concerned, the subjects erroneously took as correct the sentences which take the forms of [TELL+ $\alpha$  (*that*-clause)] and [TELL+to  $\beta$ ]. The reason may be that the corpus data includes some occurrences of seemingly but not exactly the same forms. Since the corpus data contains the forms of [be TOLD+ $\alpha$ (*that*-clause)] and [be TOLD+to-Infinitive], the subjects may have been confused with apparently the same forms.

With regard to the verb *say*, many subjects did not know that the form of [SAY+ $\beta$ (Addressee)+ $\alpha$  (*That*-clause)] ( $\beta$  without a preposition) is ungrammatical. This may be explained by the corpus data which show a high frequency of occurrences of the form of [SAY+Personal noun+ $\alpha$ (part of clause)]. The subjects may have mistaken [SAY+ $\beta$ (Addressee)+ $\alpha$  (*That*-clause)] for [SAY+Personal noun+ $\alpha$ (part of clause)]. The subjects, though, knew very well that the form of [SAY+ $\alpha$  (*that*-clause)] was acceptable and this may be also explained by the corpus data in which that form occurs most frequently among the sentence types containing the verb *say*.

For the verb *speak*, it was found that most of the subjects were aware that the structure of [SPEAK+ $\gamma$ (LANGUAGE)] is correct. The analysis of corpus data containing 731,925 words shows that the structure has a high frequency of occurrences, which, in turn, tells us that the subjects may be familiar with the structure.

With respect to the verb *talk*, most subjects made a wrong judgement about the acceptability of the sentence in which the form of [TALK+ $\gamma$ (LANGUAGE)] occurs. This can be predicted by no occurrence at all of the form in the

corpus data, even though it is claimed to be linguistically possible by Van Valin and LaPolla (1997).

The present research provides an implication that English teachers need to find ways to teach the verbs of saying effectively and efficiently so that Korean high school students can distinguish the verbs from one another with respect to their syntactic behaviors and semantic selectional restrictions even though they share parts of their meaning with one another (Adger, 2013). One way may be as outlined in the Literature Review section. Even though an array of English learning and acquisition approaches such as Communicative Language Teaching focus on fluency, not accuracy, achieving fluency is not meaningful unless accuracy is guaranteed for these high frequency verbs (Bade, 2008; Radwan, 2004).

## References

- Adger, D. (2013). *A syntax of substance*. Cambridge, Mass: MIT Press.
- Bade, M. (2008). Grammar and good language learners. In C. Griffiths (Ed.), *Lesson from good language learners* (pp. 174-181). Cambridge: Cambridge University Press.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human science*. N.J: Lawrence Erlbaum.
- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21<sup>st</sup> century? *TESOL Quarterly* 34, 548-552.
- Culicover, P. W. (2009). *Natural language syntax*. NY: Oxford University Press.
- Dirven, R. (1982). Talk: Linguistic action perspectivized as discourse. In R. Dirven, L. Goossens, Y. Putseys, & E. Vorlat (Eds.), *The scene of linguistic action and its perspectivization by SPEAK, TALK, SAY, and TELL* (pp. 37-84). Amsterdam: John Benjamins.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Li, M., & Olejnik, S. (1997). The power of Rasch person-fit statistics in detecting unusual response patterns. *Applied Psychological Measurement*. 21, 215-231.
- Linacre, M. (2009). A user's guide to WINSTEPS Rasch model computer

- programs. Chicago: Winsteps.com.
- McNamara, T. (1996). *Measuring second language performance*. London & New York city: Longman.
- Radwan, A. A. (2004). The effectiveness of explicit attention to form in language learning. *System* 33, 69-74.
- Van Valin, R. D., & Wilkins, D. P. (1993). Predicting syntactic structure from semantic representations: Remember in English and Mparntwe Arrernte. In R. D. Van Valin (Ed.), *Advances in role and reference grammar* (pp. 499-534). Amsterdam and Philadelphia: John Benjamins.
- Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning and function*. London: Cambridge University Press.
- Vorlat, E. (1982). Framing the scene of linguistic action by means of speak. In R. Dirven, L. Goossens, Y. Putseys, & E. Vorlat (Eds.), *The scene of linguistic action and its perspectivization by SPEAK, TALK, SAY, and TELL* (pp. 9-36). Amsterdam: John Benjamins.
- Wierzbicka, A. (1987). *English speech act verbs: A semantic dictionary*. Sydney: Academic Press.
- Williams, H. (2004). Lexical frames and reported speech. *ELT Journal*, 58(3), 247-257.

## Appendix

@ 아래에 나와 있는 문장은 speak, tell, talk, 또는 say와 같은 동사를 포함하고 있다. 각 문장이 올바른 문장이면 O, 그렇지 않으면 X를 괄호 속에 표시하시오.

- 1) John spoke a story about Sam. ( )
- 2) Stephen told Mary that it would snow. ( )
- 3) John talked a story about Sam. ( )
- 4) Angie said that it would snow. ( )
- 5) Jessica talked to Bob that Joan would arrive soon. ( )
- 6) Angie spoke that it would snow. ( )
- 7) Ken said French. ( )
- 8) Angie told that it would snow. ( )
- 9) Stephen talked Mary that it would snow. ( )
- 10) Ken talked French. ( )
- 11) Stephen said Mary that it would snow. ( )
- 12) John said a story about Sam. ( )
- 13) Jessica spoke to Bob that Joan would arrive soon. ( )
- 14) Raymond said only a few words to Jane. ( )
- 15) Ken spoke French. ( )
- 16) Raymond told only a few words to Jane. ( )
- 17) Angie talked that it would snow. ( )
- 18) Raymond talked only a few words to Jane. ( )
- 19) Ken told French. ( )
- 20) John told a story about Sam. ( )
- 21) Jessica said to Bob that Joan would arrive soon. ( )
- 22) Stephen spoke Mary that it would snow. ( )
- 23) Jessica told to Bob that Joan would arrive soon. ( )
- 24) Raymond spoke only a few words to Jane. ( )

Heechul Lee

Department of English Education, Chonbuk National University  
664-1 Deokjingu, Jeonju, Chonbuk Province  
Phone: 063-270-2719 E-mail: hclee@cbnu.ac.kr

Jaewoo Shim

Department of English Education, Chonbuk National University  
664-1 Deokjingu, Jeonju, Chonbuk Province  
Phone: 063-270-2729 E-mail: shimjw@cbnu.ac.kr

Nayu Kim

Department of English, Wonkwang high school  
9-4, Busongdong, Iksan, Chonbuk Province  
Phone: 063-720-3654 E-mail: kimnayu98@naver.com

Received on March 31, 2014

Revised version received on May 19, 2014

Accepted on May 30, 2014