Linguistic Research 29(1), 21-44

A Linear Regressional Analysis of *With*-PPs in English*

Hye-Won Choi
(Ewha Womans University)

Choi, Hye-Won. 2012. A Linear Regressional Analysis of *With*-PPs in English. *Linguistic Research* 29(1), 21-44. Building up on Choi's (2011) research that identifies the features of the instrument-taking predicates that influence the presence of instrument *with*-PPs, the current study explores to analyze the BYU-BNC corpus data of instrument *with*-PPs with a linear regression model. By building a multiple linear regression model which can deal with identified predictor variables simultaneously, this paper tries to explain the frequencies of instrument *with*-PPs and their nature of argumenthood by combinations of predictor variables. This modeling analysis can be a useful step toward studying gradient argumenthood via a more complex evaluation of a variety of morphological, syntactic, and semantic factors. **(Ewha Womans University)**

Keywords *with*-PP, instrument, argument, adjunct, corpus, BYU-BNC, linear regression model

1. Introduction

Instrument has been known to challenge the categorical distinction of argument and adjunct. Differing from typical adjuncts such as temporal or locational adjuncts in (1c), instrument phrases sometimes behave like arguments, similar to the dative argument PP in (1b): instrument PPs pass some of the syntactic tests for argumenthood such as iteration, fronting, clefting, and extraction (Schütze 1995, Schütze and Gibson 1999; see Choi 2010 for an overview). Yet, also differing from typical arguments such as dative PPs, instrument PPs act like adjuncts such as those in (1c): instrument PPs do not participate in valence alternations such as passive

* I am grateful to the two anonymous reviewers for their careful comments, which helped clarify the discussions in the paper. I would also like to thank Sunghye Cho for assisting me with extracting the instrument data from the BYU-BNC corpus, and Peter Kipp, Karen Muckenhirn and Kelly Song for helping judging the instrument/subject alternatability of the data. Any errors in the coding or analysis of the data are of course mine.

(Carlson and Tanenhaus 1988) and they are syntactically optional (Dowty 1989); in addition, they fail such tests as pro-form replacement and ordering (Schütze 1995, Schütze and Gibson 1999; see Choi 2010 for a review).

(1) a. Chris opened the door with a pin.
 b. Chris gave chocolate to the dog.
 c. Chris danced a shuffle on the street in December.

Choi (2010) examines argumenthood of instrument PPs by surveying the syntactic, semantic, and psychological tests and diagnostics proposed in literature, and concludes that instrument appears to be neither argument nor adjunct. Moreover, some psychological tests even seem to suggest that instrument acts as if it were something in between argument and adjunct (Boland 2005; see Choi 2010 for a review of psychological tests).

Argumenthood of a phrase is in essence determined by the predicate because arguments are phrases that are semantically or syntactically selected by the predicates in the sense that "their presence and the form they take are under the control of individual predicates" (Grimshaw 1990:108). The information about arguments is to be encoded as argument structure in the representation of lexical entries of predicates; adjuncts, by contrast, are not lexically encoded via argument structure, which means that the presence of adjuncts are not dictated or constrained by predicates.

That the instrument is neither a clear argument nor a clear adjunct suggests that the semantic/syntactic selection by the predicate may not be clear cut but a matter of gradience. In fact, noting that there are "things in the middle ground . . . often classified back and forth as arguments or adjuncts," Manning (2003:302) suggests to represent subcategorization information as a probability distribution over argument frames, with different verbal dependents expected to occur with a verb with a certain probability, conditioned on various features, such as in (2), for instance.

(2) a. P(NP[SUBJ]|V=*retire*) = 1.0
 b. P(NP[OBJ]|V=*retire*) = .52
 c. P(PP[from]|V=*retire*)= .05
 d. P(PP[as]|V=*retire*)= .06 (Manning 2003:303)

This kind of model does not tell whether *from*-PP or an *as*-PP is categorically an argument or an adjunct. The model does show, however, that the subject and object NPs are a lot more strongly dependent on or controlled by the verb than the *from*-PP or the *as*-PP, and thus a lot more argument-like. On the other hand, the *from*-PP or the *as*-PP is less argument-like and more adjunct-like, but not necessarily an adjunct either.

The idea that the argumenthood is not an all-or-nothing notion but a matter of probability takes the instrument PP out of its awkward in-between position. We can say that the instrument PP is more dependent on the verb than a pure adjunct such as *on the street* or *in December* in (1) is, although it is less dependent on the verb than the dative PP is on its verb. Furthermore, certain instrument PPs may be said to be more argument-like than other instrument PPs depending on their dependency on the verbs. As a matter of fact, Koenig, Mauner, and Bienvenue (2003) distinguish two types of instrument PPs and argue that only those that are semantically required by the predicate are arguments and those that are not semantically required are adjuncts. Yet, an instrument PP that is categorized as an adjunct according to Koenig et al. (2003) passes syntactic tests for argumenthood, whereas an "argument" instrument PP does not pass the syntactic tests for argumenthood. For instance, *stir*, which semantically requires the presence of an instrument, does not pass such syntactic tests as pro-form replacement or pseudo-clefting, as illustrated in (3). By contrast, *open*, which does not require the presence of an instrument, passes such tests as clefting and *tough*-extraction, as shown in (4).

(3) a. Sue stirred the soup with a spoon, but Fred did so with a fork.
    b. What John did with the knife was cut the bread.
    (Schütze 1995:125)

(4) a. It is this flimsy key that he convinced her to be willing to open such a heavy door with.
    b. This flimsy key is extremely hard to convince yourself to be willing to open such a heavy door with. (Schütze 1995:131)

Therefore, Koenig et al.'s (2003) categorization of instrument PPs can be reinterpreted such that some instrument PPs are more argument-like, and others are less argument-like and more adjunct-like. In other words, the degree to which an instrument PP is dependent on the predicate varies and semantic obligatoriness is one of the deciding factors. Then the question to ask is not whether an instrument PP is an argument or not any more, but rather what kinds of features of a predicate would make its instrument PP more argument-like or less.

Assuming that various factors may contribute to the degree of the predicate's selection of instrument, Choi (2011) investigates potential features or properties of predicates that affect the licensing of instrument PPs, using the BYU-BNC corpus data. Four features have been identified as valid factors: semantic obligatoriness of *with*-PP, verb's morphological relatedness to instrument noun, subject agentivity, and instrument/subject alternatability.

Building up on Choi's (2011) study that identifies the four features that seem to favor the presence of instrument PPs, the current study aims to explore what happens if all four features are put together and processed in a single model. After each feature is tested for statistical significance with a single linear regression model, a multiple linear regression model is fit to the data, where the frequency of *with*-PPs is a function of all the predictor variables processed simultaneously. By building a multiple linear regression model, this paper tries to explain the presence of instrument *with*-PPs by combinations of predictor variables. The modeling analysis facilitates studying gradient argumenthood by enabling complex evaluations of a variety of morphological, syntactic, and semantic factors.

2. Data

The data for the current study have been collected from the BYU-BNC corpus. The BYU-BNC is a 100 million-word online British National Corpus (1970s-1993) equipped with online search function. We first extracted all the transitive sentences that contain a *with*-phrase (i.e. the sentences that have the syntactic structure of S+V+O+*with*) and then manually discarded all the ones where the *with*-PP is not instrumental; namely, the sentences with the comitative or manner *with*-PPs are removed.[1]

---

1 A comitative or manner *with*-PP as shown in (i) below was not included in the data. These *with*-PPs do not pass the *use*-paraphrase test, which was used as the filtering test.

(i) a. Chris went to the movies <u>with Jon</u>. (comitative)

As what counts as an instrument phrase is not so clear, the *use*-paraphrase test is used to determine the instrumentality of a *with*-PP. A *with*-PP is considered to be instrumental if the prepositional object NP of the *with*-PP (*with* NP) can be used in the paraphrase *use* NP *to.* See an example in (5); typical tool-type phrases easily pass this test. (The following examples are from the BYU-BNC corpus and the source documents are provided in square brackets below.)

(5) a. He caught a fish with a piece of string. [ARB]
b. He used a piece of string to catch a fish.

While *a piece of string* in (5a) is a typical tool-type instrument, *honey* in (6) may not sound like an instrument at first. Yet, there seems to be no reason to treat the *with*-PP in (6a) any differently from the one in (5a), so the material-type *with*-PP in (6a) is treated to qualify as an instrument PP in our data.

(6) a. You catch more flies with honey. [A4W]
b. You use honey to catch more flies.

Similarly, other material-type *with*-PPs such as *You clean it with lots of water* [F8D] are included as instruments as they pass the *use*-paraphrase test just as in (5) and (6). *With*-PPs in the *spray*-type or *fill*-type verbs are also included under the same logic. Additionally, abstract cause (Levin 1993:81) is included as instrumental.

(7) a. He proves his worth with a characterisation. [CAD]
b. He uses a characterisation to prove his worth.

Similarly, *with*-PPs contained in *begin/conclude* type of verbs, as in *he concluded his speech with the slogan of the Quebec separatist movement* [HXU], are included. Finally, those with *adorn/embellish*-type verbs, such as in *Modern motor-bike boys adorn their jackets with stud designs* [EDE], are counted in.

The total number of sentences containing instrument *with*-PPs is 1,286. These sentences are then sorted according to the verb; the total of 430 verbs are identified

---

b. Chris did his homework with ease. (manner) as containing an instrumental *with*-PP. Then, we did another search for transitive sentences (whose structure is S+V+O) that are used with each of these verbs, this time without specifying the presence of a *with*-PP. The purpose of the second search is to calculate the proportional frequency of the instrumental PP per verb. The *with*-sentences containing the verbs *do* and *have* are left out additionally because it was impossible to calculate the total number of sentences used with these two extremely frequent verbs. This leaves us the total of 428 verbs.

To give an idea of what the data look like, the 20 most frequent verbs that appear with instrument *with*-PPs from the dataset are illustrated below in (8).

(8) Top 20 *With*-PP-Frequent Verbs

	Verb	Freq of *with*-PPs	Freq of V sentences	Prop Freq of *with*-PP (%)
1	cover	95	1778	5.343
2	hit	59	2231	2.645
3	make	44	29663	0.148
4	wipe	31	277	11.191
5	threaten	27	398	6.784
6	kill	27	1695	1.593
7	beat	21	813	2.583
8	touch	20	1041	1.921
9	attack	16	677	2.363
10	cut	15	1457	1.030
11	strike	13	955	1.361
12	buy	13	3987	0.326
13	prod	12	55	21.818
14	spray	12	76	15.789
15	rub	11	335	3.284
16	open	11	3181	0.346
17	brush	10	250	4.000
18	grab	10	591	1.692
19	catch	10	2370	0.422
20	see	10	26179	0.038

The first column (Freq of *with*-PPs) after the verb column lists the frequencies of instrument *with*-PPs, namely, the number of sentences where the verb appears with an *with*-PP out of the 1286 *with*-sentences collected. The next column (Freq of V sentences) lists the total number of sentences where the verb appears, with or without an instrument *with*-PP. The last column shows the proportional frequencies of *with*-PPs, which are calculated by the following formula: (Freq of *with*-PPs)/(Freq of V sentences)*100. The proportional frequency of *with*-PPs indicates how frequently a verb turns up with an instrument *with*-PP in terms of percentage, and is taken to be an indicator that shows how strongly the verb requires or controls the presence of an *with*-PP.

The frequency of instrumental *with*-PPs per verb varies from 1 to 95. The verb that takes a *with*-PP most frequently (i.e. 95 times) is *cover*, as illustrated in (8). However, this does not necessarily make *cover* the most likely verb to require the presence of an instrument PP (thus taking it more as an argument) because the total frequency of *cover* in its transitive usage is 1,778; therefore, the proportional frequency is only 5.34% (95/1778*100), which is lower than the average frequency 6.26%.

Among the top 20 *with*-frequent verbs, the verb with the biggest proportional *with*-frequency is *prod*. The verb *prod* shows up in 55 sentences total and among those 55 times, it appears with a *with*-PP in 12 sentences, hence the proportional *with*-frequency of 21.818. On the other hand, the verb with the smallest proportional *with*-frequency is *see*: it is as small as 0.038.

Now, I will examine in what follows what features or properties of the verbs make the difference in the *with*-frequencies. I will first review the four factors introduced in Choi (2011) and test their statistical significance with single linear regression models and then analyze the combined effects with a multiple linear regression model.

3. A Statistical Analysis

In this section, I will analyze the relationship between the verb and the presence of an instrument *with*-PP, more specifically, between the various argument-related features of the verb and the actual presence of an *with*-PP. Then, I will build a model where the presence of an *with*-PP is a function of the argument-related verbal features.

3.1 Dependent Variable

As the major concern of the paper is to study how the various factorial features of verbs affect the presence of *with*-PPs, represented as the proportional frequency of *with*-PPs, the dependent variable will be the proportional *with*-frequency. The verbs, however, vary a great deal in their frequencies: there are a few extreme low-probability verbs (which appear only once) and extreme high-probability verbs (which appear over 20,000 times). Also, quite a few verbs appear with a *with*-PP only once, while only a portion of verbs (i.e. top 20 verbs) show noticeable frequencies (over 10 times). This skewness poses a technical problem to statistical analyses because a few extreme outliers could be overly influential and distort the general trend of data points. One technical solution is to apply a logarithmic transformation to the numerical variable to remove a certain amount of skewness by "bringing many straying outliers back into the fold" (Baayen 2008:92).

The range of the proportional frequencies of *with*-PPs is from 8.384e-03 to 100; when they are log-transformed, the range becomes smaller, from -4.7815 to 4.6052.[2] As the log-transformation decreases the skewness by a great deal, the log-transformed proportional frequency of instrument *with*-PPs (labeled lWithPercent) will be used as the dependent variable in the model.

3.2 Predictor Variables

3.2.1 Does the Verb Semantically Require an Instrument? (WithRequired)

One obvious candidate for the verbal factors that influence the presence of *with*-PPs is semantic obligatoriness of instrument PPs, namely, whether the verb semantically requires the presence of an instrument. Semantic obligatoriness seems to be a natural predictor variable: that the verb semantically requires an instrument

2 The logarithm of a number to a given base is the exponent to which the base must be raised to produce that number. For example, the logarithm of 1000 to base 10 is 3, because 1000 is 10^3. Here, the natural logarithm is used where the base is the constant *e* (approximately 2.718).

means that the instrument is conceptually necessary to complete the meaning of the verb. As semantic obligatoriness is a natural basis for syntactic obligatoriness, it will increase the frequency of *with*-PPs.

Jackendoff (1977), Dowty (1982), and Koenig, Mauner, Bienvenue (2003) among others have used the semantic obligatoriness as a criterion for argumenthood. Ironically, however, Dowty (1982) categorizes instruments as adjuncts, whereas Koenig et al. (2003:79) classify at least a subgroup of verbs as taking instruments as arguments. Furthermore, recall the discussion in the Introduction that even the verbs that semantically require an instrument according to the Koenig et al.'s (2003) criterion do not necessarily pass the syntactic tests for argument, while those that do not require an instrument do pass them. Therefore, semantic obligatoriness cannot be an absolute criterion for argumenthood nor for syntactic obligatoriness. Nevertheless, it is no doubt a potential factor that facilitates the presence of instrument *with*-PPs.

Depending on whether an instrument is conceptually necessary to complete the meaning of the verb, each verb is coded as "Required" or "Optional." Additionally, Choi (2011) marked those verbs that conceptually require a body part (although not requiring an instrument) as "BodyRequired": such verbs as *bite, fist, grab, kick, nudge, paw, punch, rub,* and *tap* are coded "BodyRequried"; verbs that require static body parts such as *see* (eyes) and *hear* (ears) are excluded. Out of the total 428 verbs that take *with*-PPs, 79 verbs semantically require an instrument (Required), 49 verbs require a body part (BodyReq), and the remaining 300 verbs require neither an instrument nor a body part (Optional).

We can use one-way analysis of variance (ANOVA) to test whether there is a difference in mean *with*-frequencies among the three groups of verbs: the semantically required (Required), the body required (BodyReq), and the semantically optional (Optional).[3]

(9) ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
WithRequired	2	95.39	47.696	15.077	4.721e-07
Residuals	425	1344.48	3.163		

3 In fact, Choi (2011) used the t-tests for pairwise comparisons of the three groups of verbs for this variable. The multiple t-tests on the same data face the problem of significance inflation (Baayen 2008:105-108), hence discarded in this paper. For more discussion, see section 2.3.2

The ANOVA analysis reports that the F-value of 15.077 for 2 and 425 degrees of freedom is highly significant (p=4.721e-07). This test tells us that there are significant differences in the mean frequencies of *with*-PPs among the three kinds of verbs, Required, BodyReq, and Optional. However, it does not specify which of the possible differences in the means might be involved: BodyReq vs. Optional, Required vs. Optional, BodyReq vs. Required?

In fact, as lWithPercent is a numerical vector and WithRequired is a factor, we can use linear regression to see the inner contrasts of the analysis. Linear regression takes the relationship between the frequency of *with*-PPs and the semantic obligatoriness as a function where the frequency of *with*-PPs is the dependent variable and the semantic obligatoriness is a predictor variable.

(10) *With*-Frequency as a Function of Semantic Obligatoriness
(lWithPercent ~ WithRequired)

	Estimate	Std Error	t value	Pr(>|t|)
(Intercept)	0.9104	0.2541	3.583	0.000379
WithRequiredOptional	-0.7292	0.2741	-2.661	0.008095
WithRequiredRequired	0.4377	0.3234	1.353	0.176642

Residual standard error: 1.779 on 425 degrees of freedom
Multiple R-squared: 0.06625, Adjusted R-squared: 0.06186
F-statistic: 15.08 on 2 and 425 DF, p-value: 4.721e-07

From the table above, we can see that the model takes the BodyReq verbs of the factor WithRequired as the default or reference level.[4] The intercept 0.9104 represents the group mean of *with*-PP frequencies of the BodyReq verbs because BodyReq is the default. The t-value and the corresponding p-value (0.000379) answer the question as to whether the group mean for the BodyReq verbs, 0.9104, is significantly different from zero, which is not surprising at all.

What we are more interested is, of course, the contrasts between the default level

4 The default or reference level is so determined simply because BodyReq precedes other levels (Optional, Required) in the alphabet order; this is why there is no row labeled BodyReq in the table. If the default were Optional or Required, the resulting contrasts of course would remain the same (although the actual coefficients may differ).

and the other levels. The second coefficient -0.7292 represents the contrast (i.e. the difference) between the group mean of the BodyReq verbs and that of the Optional verbs. In other words, the Optional verbs' group mean of *with*-PP frequency is 0.1812, the outcome of subtracting 0.7292 from the intercept 0.9104 (0.9104−0.7292=0.1812). What the t-test in the table tells us is that this adjustment of -0.7292 is statistically significant, shown by the low p-value 0.008095. In other words, we have ample reason to believe that the two group means differ significantly. By contrast, the coefficient 0.4377 for the Required verbs is not statistically significant (p=0.176642 (>0.05)). This indicates that there is no reason to believe that the group mean 1.3481 (0.9104 + 0.4377) of the Required verbs is significantly different from the group mean of the default BodyReq verbs.[5]

In other words, this outcome shows that while Required verbs and BodyReq verbs each are different from the Optional verbs in their mean *with*-frequency, Required and BodyReq are not different from each other in their means of *with*-frequency. This shows that there is no reason to separate BodyReq verbs from Required verbs. In other words, BodyReq and Required verbs together are distinguishable from the Optional verbs in that they show higher *with*-frequencies. For this reason, the data have been recoded such that Required and BodyReq are collapsed into one level, as opposed to Optional. The collapsed level is termed generally as "Required." Now, the variable WithRequired has only two levels, Optional and Required. When linear regression is run again on the newly coded data, the following outcome results.

(11) *With*-Frequency as a Function of Semantic Obligatoriness (revised) (lWithPercent ~ WithRequired)

	Estimate	Std Error	t value	Pr(>|t|)
(Intercept)	0.1812	0.1028	1.763	0.0786
WithRequiredRequired	0.9993	0.1880	5.317	1.71e-07

[5] The table does not show one last comparison, i.e., the contrast between the Optional verbs and the Required verbs because the table lists only those pairwise comparisons that involve the reference level that is mapped onto the intercept. Yet, as multiple comparisons on the same data inflates the statistical significance, the last comparison will be left out. This problem arises whenever a factor has more than two levels. The contrasts among three levels should be understood to be significant if at least one contrast is significant. See Baayen (2008:105-108) to see how to recalculate the inflated significance when multiple comparisons are conducted on the same data.

Residual standard error: 1.78 on 426 degrees of freedom
Multiple R-squared: 0.06223, Adjusted R-squared: 0.06002
F-statistic: 28.27 on 1 and 426 DF, p-value: 1.709e-07

Now, since the default reference level is Optional, the Intercept coefficient 0.1812 is the mean *with*-frequency of the Optional verbs, and the coefficient of the Required, 0.9993, is the difference in the mean between the Optional and Required verbs. That is, the mean *with*-frequency of the Required verbs are 0.9993 higher than that of the Optional verbs; the Required verbs appear with a *with*-PP 2.7% more often (i.e. exp(0.9993)=2.7) than the Optional verbs. The analysis tells us that this contrast is significant as shown by the very small p-value (1.71e-07).

The model provides more summary statistics, as shown at the bottom in (11). The Residual standard error is a measure of how unsuccessful the model is; it gauges the variability in the dependent variable that we can't handle through the predictor variables. The better a model is, the smaller its residual standard error will be. The next line states the R-squared. The R-squared (R^2), the squared correlation coefficient, quantifies, on a scale from 0 to 1, the proportion of the variance that the model explains (Baayen 2008:88).[6] Finally, the F-value (the ratio of the variance estimates) goes with an overall test of whether the linear model as a whole succeeds in explaining a significant portion of the variance. Given the small p-value listed in the summary, there is no question about statistical significance. As a matter of fact, the F-value (Df 1, 426) is exactly the same as the result of ANOVA analysis, as both methods are underlyingly identical.

To summarize, WithRequired, the predictor variable that represents the semantic obligatoriness of instrument PPs to the verbs, is a statistically significant variable that affects the dependent variable lWithPercent, the proportional frequencies of *with*-PPs. The verbs that semantically require an instrument (Required) do appear with instrument *with*-PPs more frequently (by 2.7%) than the verbs that do not semantically require an instrument (Optional); the mean *with*-frequency for the Required verbs is higher than that for the Optional verbs.

---

6 Actually, the R^2 of 0.06 is rather small, which means that there are quite a lot of variances that the model cannot explain. In other words, while WithRequired (semantic obligatoriness) is a significant variable that influences the frequencies of *with*-PPs, the predictability of that variable alone is not so big. This is true for other variables.

3.2.2 Does the Verb Inherently Imply the Use of an Instrument? (InstV)

The next candidate for the predictor variable that influences the *with*-frequency is whether the verb is morphologically zero-related to an instrument noun, in other words, whether the verb originates from an instrument noun. For example, *whip* as a verb originates from the noun *whip*, which is an instrument noun; namely, verb *whip* is zero-related to noun *whip*. The total of 37 verbs in the dataset belong to this Instrument Verb (InstV) category, including *anchor, belt, bolt, brush, fan, hammer, lash, needle, rake,* and *shield*.

Instrument verbs can be regarded as a special group of semantically instrument-requiring verbs because the use of an instrument is conceptually implied or entailed as part of the meaning of the verb; the use of an instrument is an inherent meaning of the verb. For instance, *to brush* means 'to comb something with a *brush*'; *to hammer* means 'to hit something with a *hammer*.'

What makes these verbs interesting is their ambivalent nature in regard to the presence of an *with*-PP. On one hand, the InstV verbs may require the presence of an instrument *with*-PP just as the Required verbs (in section 3.2.1) do, because the verbs conceptually require the presence of an instrument. However, the actual syntactic presence of a *with*-PP may be a different matter. Since the use of an instrument is already implied or entailed in the meaning of the verb, the verb may not need the overt presence of a *with*-PP.

Just as the WithRequired was treated, another linear regression model is fit to the data, where the dependent variable lWithPercent is a function of a predictor variable InstV. InstV is a factor with two levels: Yes (Instrument verbs) and No (Non-instrument verbs).

(12) *With*-Frequency as a Function of Instrument-Verbness
(lWithPercent ~ InstV)

	Estimate	Std Error	t value	Pr(>|t|)
(Intercept)	0.35645	0.09072	3.929	9.94e-05
InstVYes	1.43012	0.30854	4.635	4.75e-06

Residual standard error: 1.794 on 426 degrees of freedom

Multiple R-squared: 0.04801, Adjusted R-squared: 0.04578
F-statistic: 21.48 on 1 and 426 DF, p-value: 4.749e-06

Parallel to the WithRequired case, the coefficient for the Yes level (i.e. Instrument verbs) is 1.43012, which exhibits the difference in the mean of *with*-frequencies (lWithPercent) between the Non-instrument verbs (No, which is mapped to the Intercept) and the Instrument verbs (Yes). The model shows that this difference 1.43012 is statistically significant (p=4.75e-06). In other words, the Instrument verbs appear with instrument *with*-PPs 4.18% (exp(1.43012)=4.18) more often than the Non-instrument verbs.[7]

3.2.3 Is the Subject of the Verb Agentive? (AgentSbj)

The third factor that affects the *with*-frequencies has to do with the verb's argument structure, another syntaco-semantic property of the predicate. That is, it seems that instruments can occur only with a proportion of verbs containing an agent, either expressed or implied (Schütze 1995:126). As discussed in detail in Choi (2010, 2011), instrument *with*-PPs can occur with active verbs that explicitly require an agent, or with passive verbs that implicitly require an agent, but not with middle verbs, which do not require an agent. Also, intransitive verbs that do not require an agent are not compatible with instrument *with*-PPs (Schütze 1995:126-127). Similarly, Spivey-Knowlton and Sedivy (1995) note in their corpus study that action verbs tend to take *with*-PPs as instruments modifying the verbs, whereas psychological and perception verbs do not. These observations can be interpreted such that only action verbs, which take agent subjects, can take instrument *with*-PPs.

Choi (2011) notes, however, that some psychological verbs take agentive subjects

---

7 A potential problem with this variable is that the numbers of observations for the two levels of InstV differ widely. Thus, both methods, linear regression and ANOVA, which are underlyingly identical, may be inappropriate. It is crucial, therefore, to check whether a non-parametric test also provides support for differences in the *with*-frequencies for different types of verb. The test illustrated here is Kruskal-Wallis Rank Sum Test (Baayen 2008:108). The small p-value (1.379e-06) supports the prediction that the Instrument Verbs favor instrument PPs better than Non-instrument PPs.

(i) Kruskal-Wallis Rank Sum Test
Kruskal-Wallis chi-squared = 23.3091, df = 1, p-value = 1.379e-06

and that these should be categorized separately from the psych verbs that take experiencer subjects. Actually, Levin (1993:191) classifies the *amuse*-type psych verbs, "whose subject is the cause of the change in psychological state," differently from the *admire*-type psych verbs with experiencer subjects. Grimshaw (1990) also argues that psych verbs, such as *amuse*, allow the subject/stimulus argument to receive an agentive interpretation.

Thus, the *amuse*-type psych verbs in the dataset, such as *delight, disturb, frighten, impress, surprise, threaten,* and *trouble*, are coded as "PsychAgent"; the *admire*-type psych verbs such as *enjoy* are coded "Experiencer" together with perception verbs such as *feel, hear, see,* and *view*. Other verbs that take experiencer subjects such as *afford, deserve, experience*, etc. are also coded "Experiencer." Action verbs are of course coded as "ActionAgent." The majority of the verbs, 375 out of 428, are ActionAgent verbs, and there are 27 PsychAgent verbs and 26 Experiencer verbs.

A linear model was fit to the data that are classified into three categories, ActionAgent, PsychAgent, and Experiencer. The dependent variable again is lWithPercent (proportional frequencies of *with*-PPs) and the predictor variable is AgentSubj (the agentivity of the subject), which has three levels: ActionAgent, PsychAgent, and Experiencer.

(13) *With*-Frequency as a Function of Subject Agentivity
(lWithPercent ~ AgentSbj)

	Estimate	Std Error	t value	Pr(>|t|)
(Intercept)	0.53323	0.09172	5.814	1.20e-08
AgentSbjExperiencer	-1.76170	0.36020	-4.891	1.43e-06
AgentSbjPsychAgent	0.85401	0.35391	2.413	0.0162

Residual standard error: 1.776 on 425 degrees of freedom
Multiple R-squared: 0.06888, Adjusted R-squared: 0.0645
F-statistic: 15.72 on 2 and 425 DF, p-value: 2.594e-07

The table of coefficients suggests that there are significant contrasts among all three levels, ActionAgent (the reference level), Experiencer and PsychAgent. The coefficient for the Experiencer level is -1.76170, which means that the mean frequency for this group is 1.76170 lower (i.e. lower by 5.82%) than that for the ActionAgent group; therefore, the mean *with*-frequency of the Experiencer verbs is -1.22847 (0.53323-1.76170=-1.22847), which translates to 0.29%. As expected, Experiencer verbs occur with instrument *with*-PPs only very rarely. The difference between the ActionAgent and the Experiencer is significant, as confirmed by the small p-value (1.43e-06). By contrast, PsychAgent verbs occur with *with*-PPs even more often than ActionAgent verbs: the mean *with*-frequency of the PsychAgent verbs is 0.87401 higher than that of the ActionAgent verbs. Although the difference is not big, it is not an ignorable difference (p=0.0162). To summarize, the PsychAgent verbs appear with instrument *with*-PPs most often, next is the ActionAgent verbs, and the Experiencer verbs occur with *with*-PPs least often.[8] Lastly, the F-test result (which is the same as the ANOVA result) reports that there are significant differences in the number of *with*-frequencies as a function of subject agentivity (AgentSbj).[9]

3.2.4 Does the Verb Allow the Subject/Instrument Alternation? (InstSbj)

The last factor considered is whether the verb allows the instrument to alternate to be the subject, as shown in (14) and (15).

---

[8] We need to take it into account that predictor variable AgentSbj has three levels. As mentioned above when we discussed WithRequired, when a factor has more than two levels, there is one comparison that is left out because the linear regression table lists only those pairwise comparisons that involve the default level, the reference level that is mapped onto the Intercept. However, if we carry out three separate t-tests, we run the risk of serious inflation in significance. One remedy to this problem is a Bonferroni Correction (Baayen 2003:106). For n comparisons, simply divide a (=0.05) by n. Any comparison that produces a p-value less than a/n (0.05/2=0.025 in our case) can be regarded as significant at the a significance level. If we were to carry out three pairwise comparisons of two means, any comparison that yields a p-value less than 0.0167 (0.05/3=0.0167) could be accepted as significant. The second remedy to avoid inflation in statistical significance is Tukey's Honestly Significant Difference. However, we cannot use this test because the numbers of observations are not equal among the three groups.

[9] It should be noted again that the numbers of observations for the three levels of AgentSbj differ widely. It is crucial, therefore, to check whether a non-parametric test also provides support for differences in the *with*-frequencies with different subject types. Again, Kruskal-Wallis Rank Sum Test is used. The small p-value (5.716e-07) supports our intuition that the groups means are different.

(i) Kruskal-Wallis Rank Sum Test
Kruskal-Wallis chi-squared = 28.7496, df = 2, p-value = 5.716e-07

(14) a. David broke the window with a hammer. (Levin 1993:80)
b. The hammer broke the window. (intermediary agent)

(15) a. Doug ate the ice cream with a spoon.
b. *The spoon ate the ice cream. (enabling/facilitating instrument)

As seen in the contrast in (14), verb *break* allows the instrument *a/the hammer* in (14a) to alternate to become the subject in (14b), whereas *eat* in (15) does not allow the instrument *a/the spoon* in (15a) to be the subject as in (15b). As noted by Marantz (1984) and Levin (1993) among others, the instrument/subject alternation is correlated with the distinction in interpretation of the instrument phrase. Instruments that can alternate as in (14a) are interpreted more as "intermediary" agents/instruments while those that cannot as in (15a) are interpreted as "facilitating/enabling" instruments. The former, unlike the latter, are able to perform the action independently, as reflected in their occurrence as subjects. The difference in interpretation of an *with*-PPs does not emerge from considering the PPs in isolation but depends on the identity of the verb. Schütze (1995:128) argues that instruments are like arguments in their dependency on the verb for their interpretation.

Taking the instrument/subject alternatability as a potential factor that influences the *with*-frequency comes from the idea that the presence of an "intermediary agent" may be more crucial to the verb than the presence of a "facilitating" instrument. Hence, verbs that allow instrument/subject alternation (i.e taking "intermediary agent" PPs) will be more likely to occur with *with*-PPs than those that do not allow instrument/subject alternation (i.e. taking "facilitating/enabling instrument PPs). The 428 verbs are thus coded with respect to this factor: 273 verbs allow instrument/subject alternation (Yes) and 155 verbs do not (No).[10]

A linear regression with instrument/subject alternatability (InstSbj) as a predictor

[10] The instrument/subject alternatability depends on the choice of instrument as well as the verb (Levin 1993:80). Therefore, judgments as to whether or not each verb allows the alternation varies among the three informants depending on the choice of instruments. Testing all three sets of judgments by the three informants shows that the more liberal the judgments are (i.e. the more verbs are marked Yes), the smaller the p-value is (i.e. the more significant the variable is), although all three sets yield the same result that the instrument/subject alternatability is a significant variable. Choi (2011) used the most conservative informant's judgments which mark 230 verb as Yes, and 198 verbs as No. In this paper, the judgments by the informant in the middle position are used: a little more verbs (273 verbs) are coded Yes, and 155 verbs as No.

variable yields the following result.

(16) *With*-Frequency as a Function of Subject/Instrument Alternatability (lwithPercent ~ InstSbj)

	Estimate	Std Error	t value	Pr(>|t|)
(Intercept)	0.1383	0.1462	0.946	0.34475
InstSbjYes	0.5359	0.1831	2.927	0.00360

Residual standard error: 1.82 on 426 degrees of freedom
Multiple R-squared: 0.01972, Adjusted R-squared: 0.01741
F-statistic: 8.568 on 1 and 426 DF, p-value: 0.003605

The regressional analysis in (16) shows that the verbs that allow instrument/subject alternatability (Yes), i.e., those that take "intermediary agent" instruments, are more likely to occur with instrument *with*-PPs. The mean *with*-frequency for the Yes verbs is higher by 0.5359 (1.71%) than the No verbs (which are mapped onto the Intercept), and this difference is statistically significant as indicated by the p-value. Note that p-value 0.00360, while it is smaller than 0.05, is not remarkably small compared to the p-values for other variables. It seems that InstSjb is not a strong factor.

I have analyzed whether each of the four predictor variables identified in Choi (2011) is a significant variable using linear regression models. Now, I will explore what happens when all four variables are considered simultaneously. Note that each of the 428 verbs that take *with*-PPs is now coded with respect to the four predictor variables: theoretically, verbs can be coded in 24 different combinations of variable features. See the top 20 verbs again in (17). For example, *brush* [Required, Yes, ActionAgent, Yes] has the strongest combination of variable features, while *see* [Optional, No, Experiencer, No] has the weakest combination; other verbs have a variety of mixed (sometimes conflicting) combinations. In other words, instrument-taking verbs are not uniform in their feature combinations, and depending on the combinations, verbs will have varying degrees of dependency relationships with instrument *with*-PPs, which will result in different frequencies.

(17) Top 20 Verbs with Variable Features

	Verb	log(With Percent)	With Required	InstV	AgentSbj	InstSbj
1	cover	1.676	Required	No	ActionAgent	Yes
2	hit	0.973	Required	No	ActionAgent	Yes
3	make	-1.908	Optional	No	ActionAgent	No
4	wipe	2.415	Required	No	ActionAgent	Yes
5	kill	0.466	Optional	No	ActionAgent	Yes
6	threaten	1.915	Optional	No	PsychAgent	Yes
7	beat	0.949	Required	No	ActionAgent	Yes
8	touch	0.653	Required	No	Experiencer	No
9	attack	0.860	Optional	No	ActionAgent	No
10	cut	0.029	Required	No	ActionAgent	Yes
11	buy	-1.121	Optional	No	ActionAgent	No
12	strike	0.308	Required	No	ActionAgent	Yes
13	prod	3.083	Required	No	ActionAgent	No
14	spray	2.759	Required	No	ActionAgent	Yes
15	open	-1.062	Optional	No	ActionAgent	Yes
16	rub	1.189	Required	No	ActionAgent	No
17	brush	1.386	Required	Yes	ActionAgent	Yes
18	catch	-0.863	Optional	No	ActionAgent	Yes
19	grab	0.526	Required	No	ActionAgent	Yes
20	see	-3.265	Optional	No	Experiencer	No

3.3 The Model

We now fit to our instrument data a multiple linear regression model where the presence of *with*-PPs (lWithPercent) is to be modeled as a function of the four predictor variables, namely, WithRequired (semantic obligatoriness of *with*-PP), InstV (instrument-verbness), AgentSbj (agentivity of subject), and InstSbj (instrument/subject alternatability). A multiple linear regression model can simultaneously process more than one predictor variable. That is, when the verbs have conflicting combinations of variable features, the model is capable of resolving clashes by negotiating variable weights. In this case, the term "linear" does not directly show the relation between the dependent variable and the predictor(s). What "linear" denotes is that the dependent variable can be expressed as the sum (or linear combination) of a series of weighted predictor variables. The weights of the predictors are the estimated coefficients (Baayen 2008:96). (The simple main effects are separated by plus symbols in the formula for the linear regression model.)

(18) Multiple Linear Regressional Model
lWithPercent ~ WithRequired + InstV + AgentSbj + InstSbj

(19) *With*-Frequency as a Function of Four Factors

	Estimate	Std Error	t value	Pr(>|t|)
(Intercept)	-0.04185	0.15579	-0.269	0.788346
WithRequiredRequired	0.76948	0.19998	3.848	0.000138
InstVYes	0.90042	0.31921	2.821	0.005017
AgentSbjExperiencer	-1.32205	0.35446	-3.730	0.000218
AgentSbjPsychAgent	1.18704	0.34489	3.442	0.000635
InstSbjYes	0.34397	0.17392	1.978	0.048610

Residual standard error: 1.701 on 422 degrees of freedom
Multiple R-squared: 0.1518, Adjusted R-squared: 0.1417
F-statistic: 15.1 on 5 and 422 DF, p-value: 1.178e-13

The output of the linear model fit in (19) evaluates whether the coefficients are significantly different from zero in a model containing all other predictors, in other words, whether each variable is still significantly contributing when all the variables are considered at the same time. The model has six coefficients. Note that these coefficients are different from those when each variable is modeled alone as seen in the previous sections; as all the variables are processed simultaneously, the effects of the variables are weighted and adjusted. Note also that the multiple model is better than each single linear model in explaining variances, represented by the improved R-squared (R^2).[11]

[11] While better than single models, the R^2 of the multiple model is not great either, which means that a considerable portion of variances are still not explained by the model.

The first is a coefficient for the Intercept, which is the reference level that is contrasted with the other level(s) for each variable. The second coefficient is for the contrast between the two levels of the factor WithRequired. The group mean for the subset of Required is 0.76948 (2.16%) higher than that for Optional, the reference level mapped onto the Intercept. This difference between the Optional verbs and the Required verbs is significant, as indicated by the low p-value 0.000138. The third coefficient 0.90042 is for the contrast between the two levels of the factor InstV. The group mean for the Yes group is 0.90042 (2.46%) higher than the group mean for the No group (mapped onto the Intercept). Again, this difference is significant, as indicated by the p-value 0.005017. The next two coefficients are about the contrasts among the three levels of the factor AgentSbj. As the reference level is ActionAgent mapped onto the Intercept, the first contrast is between ActionAgent verbs and Experiencer verbs. The group mean of the Experiencer verbs is 1.32205 (3.75%) lower than that of the ActionAgent verbs. The next coefficient indicates the difference between the ActionAgent and the PsychAgent verbs. The group mean of PsychAgent verbs is 1.18704 (3.28%) higher than that of the ActionAgent verbs. These two contrasts are significant as indicated by the p-values, 0.000218 and 0.000635 respectively. Finally, the last coefficient is about the contrast between the two levels of the factor InstSbj. The Yes verbs, which allow instrument/subject alternation, have the 0.34397 (1.41%) higher mean than the No verbs, which do not allow such alternation. The difference is not very big and the statistical significance of the contrast between these two groups is on the border line, as is shown by the p-value 0.048610, which is right below 0.05. Although this factor is not so strong, it is not insignificant, so will be left in the model.

Now, we can run the ANOVA test on the model. What the summary in (20) tells us is whether, by means of F-tests, each predictor contributes significantly to explaining the variance in the dependent variable.

(20) ANOVA on the Model
Response: lWithPercent

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
WithRequired	1	89.60	89.597	30.9579	4.692e-08
InstV	1	24.55	24.545	8.4809	0.003779
AgentSbj	2	93.08	46.540	16.0806	1.860e-07
InstSbj	1	11.32	11.320	3.9114	0.048610
Residuals	422	1221.33	2.894		

ANOVA on a linear model is referred to as a Sequential Analysis of Variance because it shows in a sequential way whether a predictor further down the list has anything to contribute in addition to the predictors higher on the list (Baayen 2008:166). The output here in (20) is different from that in (19), which shows whether each variable is significant when all other variables are considered together as well. Each successive row in a sequential ANOVA table in (20) evaluates whether adding a new predictor is justified, given the other predictors in the preceding rows. The small p-value at the end of each row shows that adding the variable on the row is justified. Note that the p-value 0.048610 for InstSbj in (20) is the same as the p-value for InstSbj in (19). This makes sense because on the last row, adding the last variable means having all the other variables as well. Again, the p-value for InstSbj is not so small (while still valid), which shows that this variable's contribution to explaining *with*-frequencies is not so great.

4. Conclusion

Following up on Choi's (2011) study of potential factors that contribute to licensing the presence of instrument *with*-PPs in English, the current research has explored to analyze the BYU-BNC corpus data of instrument *with*-PPs with a linear regression model. After each factor variable is tested for statistical significance with a single linear regression model, a multiple linear regression model is fit to the data, where the frequency of *with*-PPs is a function of all the predictor variables processed simultaneously. The modeling analysis shows that all four variables are still significant when considered all together although their effects are adjusted, and therefore confirms that an instrument *with*-PP is more likely to occur with a verb if (a) the verb semantically requires an instrument; (b) the verb is morphologically zero-related to the corresponding instrument noun; (c) the subject of the verb is agentive; and (d) the instrument is an intermediary agent that can participate in the instrument/subject alternation.

The fact that instrument-taking verbs have a variety of combinations of variable features suggests that verbs may require instrument *with*-PPs to varying degrees. That is, those verbs that have stronger combinations of variable features that favor the presence of *with*-PPs will be more likely to take instrument PPs more frequently and more like arguments, whereas those that have weaker combinations of variable features will be more likely to take instrument PPs less frequently and more like adjuncts. The modeling analysis can combine linguistic theories and quantitative data in a scientific way, and opens a way to studying gradient argumenthood through complex evaluations of various morphological, syntactic, and semantic factors.

References

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Boland, Jule E. 2005. Visual Arguments. *Cognition* 95: 237-274.

BYU-BNC: British National Corpus. Architecture and Interface by Mark Davies. http://corpus.byu.edu/bnc/.

Carlson, Greg and Michael Tanenhaus. 1988. Thematic Roles and Language Comprehension. *Thematic Relations*. *Syntax and Semantics*. New York: Academic Press, 263-291.

Choi, Hye-Won. 2010. The Distinction of Argument and Adjunct as a Gradient Notion. *Language and Information* 14: 25-48.

Choi, Hye-Won. 2011. Licensing Instrument Phrases: A Corpus Study of *With*-PPs in English. *Generative Grammar* 21: 75-96.

Dowty, David. 1982. Grammatical Relations and Montague Grammar. In Pauline Jacobson and Geoffrey K. Pullum (eds.), *The Nature of Syntactic Representation*. Dordrecht: Reidel, 79-130.

Dowty, David. 1989. On the Semantic Content of the Notion of 'Thematic Role.' *Properties, Types, and Meaning,* eds. G. Chierchia, B. Partee, and R. Turner, Dordrecht: Kluwer,

79-130.
Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
Jackendoff, Ray. 1977. *X-bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.
Koenig, Jean-Pierre, Gail Mauner, and Breton Bienvenue. 2003. Arguments for Adjuncts. *Cognition* 89: 67-103.
Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
Manning, Christopher. 2003. Probabilistic Syntax. *Probabilistic Linguistics*, ed. by R. Bod, J. Hay, and S. Jannedy, 289-314. Cambridge: MIT Press.
Marantz, Alec. 1984. *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.
R Development Core Team 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.
Schütze, Carson T. 1995. PP Attachment and Argumenthood. *MIT Working Papers in Linguistics* 26: 95-151.
Schütze, Carson T. and Edward Gibson. 1999. Argumenthood and English Prepositional Phrase Attachment. *Journal of Memory and Language* 40: 409-431.
Spivey-Knowlton, Michael J. and Julie C. Sedivy. 1995. Resolving Attachment Ambiguities with Multiple Constraints. *Cognition* 55: 227-267.

Hye-Won Choi
Department of English Language and Literature
Ewha Womans University
11-1 Daehyun-dong Seodaemun-gu
Seoul 120-750, Korea
Tel: 02-3277-2175 / Fax: 02-3277-2863
E-mail: hwchoi@ewha.ac.kr

Received: 2012. 03. 26
Revised: 2012. 04. 20
Accepted: 2012. 04. 20