Linguistic Research 34(2), 163-190
DOI: 10.17250/khisli.34.2.201706.001

Verb class and instrument PPs: A mixed model analysis*

Hye-Won Choi
(Ewha Womans University)

Choi, Hye-Won. 2017. Verb class and instrument PPs: A mixed model analysis. *Linguistic Research* 34(2), 163-190. This paper investigates whether a lexical variance caused by verbs influences the choice and frequency of instrument *with*-PPs. The corpus data shows that observations grouped by verb class demonstrate systematically different behaviors and this individual variance of verb classes can be captured by means of a random effect of a mixed-effects model. Building up on Choi's (2012) research that identifies the syntactic, semantic, and morphological factors that influence the presence of instrument *with*-PPs as fixed effects, the current study classifies the instrument-taking verbs into verb classes, based on Levin's (1993) study, and builds a mixed-effects model taking verb class as a random variable. The new statistical technique of hierarchical, multi-level, mixed-effects modeling (Baayen 2008; Bresnan *et al*. 2007; Gelman and Hill 2007; Johnson 2008; Kuperman 2009; Pinheiro and Bates 2000) can process across-word fixed effects and by-word random effects together. By taking into consideration the subtle syntactic and semantic characteristics of verbs, this new modeling analysis provides a way to incorporate native speakers' lexical knowledge into grammar. **(Ewha Womans University)**

Keywords instrument PP, *with*, verb class, corpus, BYU-BNC, linear regression, mixed-effects model, mixed model, fixed variable, random variable

1. Introduction

Exploring the question of whether an instrument *with*-PP in English[1] is an

* I would like to thank the two anonymous reviewers for their instructive and insightful comments and recommendations, which greatly helped improve the paper in logical clarity and readability. The remaining errors are all mine.

1 An instrument is defined to be an "inanimate force or object causally involved in the action or state identified by the verb" (Fillmore 1968) or "an intermediary between actor and patient in the decomposition of an action" (Jackendoff 1977). In English, instruments are usually marked by the preposition *with*, although not all *with*-PPs are instruments, as shown in (i): the *with*-PP in (b) is comitative and the one in (c) is a manner phrase (Schütze 1995: 124).

(i) a. John stirred the soup with a spoon.

argument or adjunct, Choi (2011) suggests that the issue of argument or adjunct may not be of categorical distinction but rather of gradience in the sense that grammar is a matter of probability (Bresnan and Hay 2008; Manning 2003). Differing from typical adjuncts such as temporal or locational adjuncts, instrument phrases, as shown in (1), behave like arguments, similar to the dative argument PP: instrument PPs pass some of the syntactic tests for argumenthood such as iteration, fronting, clefting, and extraction. On the other hand, differing from typical arguments such as dative PPs, instrument *with*-PPs act like adjuncts: instrument PPs do not participate in valance alternations such as passive (Carlson and Tanenhaus 1988) and fail such syntactic tests for argumenthood as *pro*-form replacement and ordering (Schütze 1995; Schütze and Gibson 1999; see Choi 2010b for a review).

(1) a. Kim wiped the table with a napkin.
b. Kim saw the man with a binocular.

As an alternative to the analyses based on the categorical distinction, Choi (2012) proposes a probabilistic approach along the line of Manning (2003:302), where the subcategorization information of a verb is to be represented as a probability distribution over argument frames, with different verbal dependents occurring with the verb with a certain probability, conditioned on various features. Choi identifies the four features that seem to favorably influence the presence of instrument PPs—Semantic obligatoriness of *with*-PP, Verb's morphological relatedness to instrument noun, Subject's agentivity, and Instrument/subject alternatability—and builds a statistical model that can process these four features as predictor variables. This multiple linear regression model is designed to explain the presence of *with*-PPs (represented by the proportional frequency of *with*-PPs) as a function of all the predictor variables processed simultaneously.

While the linear regression model captures and processes the major variables effecting the presence of *with*-PPs, it still leaves a wide range of variance. Thus, this paper explores the role of lexical item verb as a random variable to improve the model. The purpose of this paper is to investigate whether the random variable "verb

b. John visited his parents with Mary.
c. John opened the box with care.

class" is valid in accounting for the presence of *with*-PPs, and if so, to propose a new mixed-effects model that can process fixed and random effects together (Baayen 2008; Bresnan *et al*. 2007; Bresnan and Hay 2008; Bresnan and Ford 2010; Choi 2010a; Johnson 2008).

2. Linear regression model in Choi (2012)

Before exploring the role of lexical variance, we will first review how Choi's (2012) simple linear regression model works as it is the basic model upon which we will test improvability by adding a further variable.

2.1 Data

The data were collected from the BYU-BNC corpus, a 100 million-word online British National Corpus (1970s-1993) equipped with online search function. The total number of sentences containing instrumental *with*-PPs is 1,286 and the total 428 verbs are identified as used in these *with*-PP token sentences.[2] Then, proportional frequency of *with*-PP per verb was calculated by conducting a search for all transitive sentences that are used with each of these 428 verbs. For example, the proportional frequency for verb *cover* (which takes a *with*-PP most frequently, i.e. 95 times) is 5.34% (95 out of 1778) as the total frequency of *cover* in its transitive usage is 1,778. The same data set coded for the Choi (2012) model is used for the current study.

2.2 Response variable

As our major concern is how to predict the occurrence of *with*-PPs given the factors, the presence of *with*-PPs, represented as the proportional frequency of *with*-PPs (i.e., the number of *with*-PP-containing tokens divided by the total number of sentence tokens per verb), would be the response or dependent variable. The verbs, however, vary a great deal in their frequencies (from 1 to

---

[2] Actually there were 430 verbs identified as containing an instrument *with*-PP but verbs *do* and *have* were discarded because it was impossible to calculate the total number of sentences used with these two extremely frequent verbs.

over 20,000) and 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). To avoid distortion by a few extreme outliers, a logarithmic transformation was applied to the proportional *with*-frequency as a technical solution to this 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.[3] The log-transformed proportional frequency of instrument *with*-PPs (labeled lWithPercent) is used as the response variable in the model.

2.3 A simple model with fixed variables only

In Choi (2012), a multiple linear regression model where the proportional frequency of *with*-PPs (lWithPercent) is modeled as a function of four predictor variables, namely, WithRequired (semantic obligatoriness of *with*-PP), InstV (morphological relatedness to instrument noun), AgentSbj (agentivity of subject), and InstSbj (instrument/subject alternatability). This 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" denotes that the response 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.)

(2) Multiple linear regression model (Choi 2012)
lWithPercent ~ WithRequired + InstV + AgentSbj + InstSbj

The output of the linear model fit in (3) evaluates whether the coefficients are significantly different from zero in a model containing all other predictors, in

---

[3] 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).

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 (see Choi 2012 for details): as all the variables are processed simultaneously, the effects of the variables are weighted and adjusted.

(3) *With*-frequency as a function of four factors

	Estimate	Std Error	t value	Pr(>|t|)
(Intercept)	-0.0419	0.1558	-0.269	0.7883
WithRequired:Required	0.7695	0.2000	3.848	0.0001
InstV:Yes	0.9004	0.3192	2.821	0.0050
AgentSbj:Experiencer	-1.3221	0.3545	-3.730	0.0002
AgentSbj:PsychAgent	1.1870	0.3449	3.442	0.0006
InstSbj:Yes	0.3440	0.1739	1.978	0.0486

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 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 0.7695 (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 0.0001. The third coefficient 0.9004 is for the contrast between the two levels of the factor InstV. The group mean for the Yes group is 0.9004 (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 0.0050. The next two coefficients are about the contrasts among the three levels of the factor AgentSbj. As the reference level is ActionAgent, being mapped onto the Intercept, the first contrast is between ActionAgent verbs and Experiencer verbs. The group mean of the Experiencer verbs is 1.3221 (3.75%) lower (indicated by the minus sign) 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 1.1870 (3.28%) higher than that of the ActionAgent verbs. These two contrasts are significant as indicated by the

p-values, 0.0002 and 0.0006 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 0.3440 (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 0.0486, which is right below 0.05. Although this factor is not so strong, it is not insignificant, so was left in the model.

Then the ANOVA test was run on the model. What the summary in (4) tells us is whether, by means of F-tests, each predictor contributes significantly to explaining the variance in the response variable.

(4) ANOVA on the model

	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	0.0038
AgentSbj	2	93.08	46.540	16.0806	1.860e-07
InstSbj	1	11.32	11.320	3.9114	0.0486

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 (4) is different from that in (3), 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 (4) 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 0.0486 for InstSbj in (4) is the same as the p-value for InstSbj in (3). 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.

While better than single models constructed with each variable, the R^2 (R-squared) of the multiple model is not so great, which means that a considerable portion of variances are still not explained by the model. R^2 is an important measure to show

the explanatory power of the model, which indicates the proportion of variance accounted for by the model. It can range between 0 and 1 and a bigger R^2 value shows that the model is substantially more significant (Szmrecsanyi 2005: 119).

2.4 Problem of lexical variance

While Choi's (2012) linear regression model was a valid and meaningful attempt to capture the direction and degree to which each of the four factors, when combined together, influences the frequency of *with*-PP, there remains the question why we still see a wide range of variance among the tokens that share the same set of features. For example, 9 of the top 20 most frequent verbs, the shaded verbs in (5) below, share the same feature values [Required; No; ActionAgent; Yes]: namely, all these verbs semantically require a *with*-phrase; they are not morphologically related to the corresponding instrument noun; their subject is agentive; and the instrument can be used as the subject in the alternating construction. Yet, the range of the log-transformed frequency of *with*-phrase with these verbs is immense, from 0.029 to 2.759.

(5) Feature values of the top 20 frequent verbs

	Verb	log (WithPercent)	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

The wide range of variance seems to raise the question whether the apparent semantic, syntactic, and morphological effects really hold when they are tested on specific (groups of) verbs. The encoded effects introduced in the liner regression model are called “fixed” variables because their effects can generalize to other words and sentences regardless of speaker or lexical variation (Baayen 2008; Johnson 2008). In a fixed-effects model, the coefficients of the regression do not vary by group (Gelman and Hill 2007: 245). That is, these effects are assumed to hold no matter who utters the sentences or what words are used in those sentences. It is assumed, for instance, that the verbs taking the “agentive subject” would occur with a *with*-PP more often regardless of individual verbs, may it be *hit*, *cut*, *wipe*, or *spray*. However, the wide range of variance among the verbs that share the same value set of the fixed variables seems to suggest that in addition to the shared features, each individual verb or group of verbs may have its own specific characteristics working for or against the co-occurrence with the *with*-PP. If such lexical variation does exist, a model needs to take these variables into account as well as the fixed variables, which constantly apply regardless of lexical items. One way is to make adjustments to the coefficient for each group of lexical items, which are referred to as random intercepts (Baayen 2008: 247). Different from the fixed variables, variables for lexical variation are called “random” as they are item-specific and thus not repeatable to other words (Baayan 2008; Gelman and Hill 2007; Kuperman 2009).

3. Verb class

Verbs, as argument-taking elements, show especially complex sets of properties so that native speakers can make subtle judgments concerning the occurrence of various syntactic expressions (Levin 1993: 2). This should be true of instrument *with*-phrases also, particularly when we take the view that argumenthood is a probabilistic notion (Manning 2003). Verbs can be grouped into classes according to their syntactic behavior and semantic meaning. Levin (1993: 14) observes that a class of verbs whose members pattern together with respect to diathesis alternations, i.e. alternations in the argument expressions, form a semantically coherent class. In other words, diathesis alternations can be used as a probe into the semantically coherent verb classes.

Verbs that take instrument *with*-PPs also participate in various diathesis alternations often accompanied by changes of meaning. One such alternation is Instrument Subject Alternation, examples of which are presented below (Levin 1993: 80). Marantz (1984) notes that a *with*-PP may be interpreted as an "intermediary agent" or a "facilitating/enabling instrument" and that this distinction correlates with the syntactic possibility of using the instrument as a subject.

(6) a. David broke the window with a hammer.
b. The hammer broke the window. (intermediary instrument)

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

Some verbs, like *break* in (6), can take an instrument subject, as demonstrated in (6b), because they take 'intermediary agent' instruments as in (6a). Other verbs, like *eat*, can never take an instrument subject, as shown in (7b) because they only take 'enabling/facilitating' instruments, as shown in (7a).

Similarly, Abstract Cause Subject Alternation and Locatum Subject Alternation show comparable behaviors in that the 'cause' or 'locatum' *with*-phrase can turn into the subject (Levin 1993: 81). Examples are illustrated below respectively.

(8) a. He established his innocence with the letter.
b. The letter established his innocence.

(9) a. I filled the pail with water.
b. Water filled the pail.

As we included in our corpus data these 'cause' and 'locatum' *with*-phrases as "instrument" in a broader sense (see Choi 2011),[4] we called all three Subject

[4] To determine what counts as an instrument *with*-PP, I used the *use*-paraphrase test, as a comitative or manner *with*-PP does not pass this test. A *with*-PP is considered to be an instrument if the prepositional object NP of the *with*-PP (i.e., *X* in *with X*) can be used in the paraphrase . . . *used X to* . . . . According to this criterion, not only the tool-type but also the material-type counts as instrument, and thus the *with*-PPs in the Spray or Fill verbs are included. Also, abstract cause *with*-PPs (Levin 1993: 81) and the *with*-PPs contained in Begin/Conclude type of verbs are

Alternation cases as one single kind and classify the "instrument"-taking verbs into two groups depending on this alternability. In fact, this feature was encoded as a fixed variable InstSbj in the simple linear model developed in Choi (2012).

Another diathesis alternation is the With/Against Alternation, which involves such verbs as *hit*, *swat*, *spank, poke*, *throw*, and *break* (Levin 1993: 67). These verbs can be grouped into 3 categories depending on the alternatability. The Hit verbs can take both the *against*- and *with*-variants, as in (10); the Swat, Spank, Poke verbs can take only the *with*-variant, as in (11); the Throw and Break verbs can take only the *against*-variant, as in (12).

(10) a. Brian hit the stick against the fence.
b. Brian hit the fence with the stick.

(11) a. *Don swatted the newspaper against the mosquito.
b. Don swatted the mosquito with the newspaper.

(12) a. Brian threw the stick against the fence.
b. *Brian threw the fence with the stick.

Other verbs that share the same alternating pattern with each of these verbs above are listed below (Levin 1993: 67). Underlined are the verbs that are found in our corpus dataset. (The same underlining is also used in the discussions below to indicate the items that are found in the dataset.)

(13) Alternating Verbs:
Hit Verbs: bang, bash, batter, beat, bump, butt, dash, drum, hammer, hit, kick, knock, lash, pound, rap, slap, smack, smash (where no effect implicated), strike, tamp, tap, thump thwack, whack

(14) Non-Alternating *with* Only:
a. *Swat Verbs: bite, claw, paw, peck, punch (person), scratch, shoot (gun), slug, stab, swat, swipe

included. Additionally, those with Adorn/Embellish verbs are counted in under the same logic. See Choi (2011) for details.

b. *Spank Verbs: belt, birch, bludgeon, bonk, brain, cane, clobber, club, conk, cosh, cudgel, cuff, flog, knife, paddle, paddywhack, pummel, sock, spank, strap, thrash, truncheon, wallop, whip, whisk
c. *Poke Verbs: dig, jab, pierce, poke, prick, stick

(15) Non-Alternating *against* Only:
a. *Throw Verbs: bat, bunt, ?cast, chuck, ?fire, flick, fling, flip, hit(ball), hurl, kick(ball), knock, lob, ?loft, itch, nudge, pass, pitch, punt, shoot(projectile), shove, slap, sling, smash, tap, throw, tip, toss
b. *Break Verbs: break, crack, crash, crush, fracture, rip, shatter, smash, snap, splinter, split, tear

One assumption that we can make out of this alternating pattern concerning the presence of *with*-PP is that the non-alternating verbs taking the *with*-variant only such as Swat, Spank, Poke verbs will be more likely to cooccur with *with*-PPs than those verbs taking the *against*-variant only such as Throw and Break verbs. And the Hit verbs, which can alternate, may be in the middle in terms of *with*-frequency. The prediction is not that simple, though. The similar kind of verbs participate in another alternation, for instance, the Through/With Alternation (Levin 1993: 68), and in this case, the alternation pattern of the verbs does not exactly match the With/Against Alternation. See examples below.

(16) a. Alison pierced the needle through the cloth
b. Alison pierced the cloth with a needle.

(17) a. *Paula hit the stick through/into the fence.
b. Paula hit the fence with the stick.

Unlike the With/Against Alternation, the Poke verbs such as *pierce* in (16) are alternating, while the Hit verbs are not, as shown in (17); the Swat and Spank verbs again take the *with*-variant only, along with Touch verbs. The alternatability of the participating verbs are illustrated below (Levin 1993: 68).

(18) Alternating Verbs:
Poke Verbs: dig, jab, pierce, poke, prick, stick

(19) Non-alternating *with* Only:

a. *Hit Verbs: bang, bash, batter, beat, bump, butt, dash, drum, hammer, hit, kick, knock, lash, pound, rap, slap, smack, smash(where no effect implicate), strike, tamp, tap, thump, thwack, whack

b. *Swat Verbs: bite, claw, paw, peck, punch(person), scratch, shoot (gun), slub, stab, swat, swipe

c. *Spank Verbs: belt, birch, bludgeon, bonk, brain, cane, clobber, club, conk, cosh, cudgel, cuff, flog, knife, paddle, paddywhack, pummel, sock, spank, scrap, thrash, truncheon, wallop, whip, whisk

d. *Touch Verbs: caress, graze, kiss, lick, nudge, pat, peck (=kiss), pinch, prod, sting, stroke, tickle, touch

Comparing the two alternation patterns above, we can see that a certain subgroups of verbs among the *with*-taking verbs show similar syntactic behaviors. As these behaviors obviously influence the presence of *with*-PPs, we need to encode these features into the model. What we could do is to posit the With/Against Alternatability and the Through/With Alternatability as respective fixed variables in the model, just as we did with the Instrument Subject Alternation case. Then variable With/Against Alternation would have three levels: Yes, *with*-only, *against*-only; whereas variable Through/With Alternation would have two levels: Yes, *with*-only. The values that prefer the presence of *with*-PPs are shaded in the table below.

(20) Verbs classes' alternating patterns

	Instrument Subject Alternation	With/Against Alternation	Through/With Alternation
Hit	Yes	Yes	*with*-only
Swat, Spank	No	*with*-only	*with*-only
Poke	Yes	*with*-only	Yes
Throw, Break	Yes	*against*-only	-
Touch	No	-	*with*-only

As for these three features, the strongest combination for the presence of *with*-PP would be [Yes; *with*-only; *with*-only]. None of the verb classes in (20) have these values; nor does any have the weakest combination [No; *against*-only; Yes] either.

Interestingly, however, if we look at the means of the log-transformed proportional *with*-frequency of these verb classes in our dataset, we can see that those verb classes that have two strong features such as Hit, Swat, Spank, and Poke show higher mean frequency than those that have only one strong feature such as Throw, Break, and Touch. See (21) below.

(21) Means of response variable lWithPercent per verb class

Hit	Swat	Spank	Poke	Throw	Break	Touch
1.3593	1.3107	1.8844	1.6331	0.4674	-0.5141	0.8118

While it looks obvious that different verb classes show different *with*-frequency patterns, there are problems with this approach of adding more fixed-variables. One is that not all *with*-taking verbs can be classified according to the alternatability because of the particular semantics these alternations require. For example, Touch verbs do not participate in the With/Against Alternation and Throw/Break verbs do not participate in the Through/With Alternation. Only those verbs that share the closely relevant meaning can participate in With/Against Alternation or Through/With Alternation. Then quite a few other verbs would remain valueless for these variables. Also, there are more alternations that involve *with*-PPs, and thus it may not be a very efficient way to posit a new fixed variable whenever we find a new alternation pattern.

Another well-known alternation is the Locative Alternation (1993:2,51). The Spray/Load verbs and Fill/Cover verbs are semantically very closely related. Yet, while Spray/Load verbs may express their arguments in two different ways, as shown in (22), Fill verbs are used only in the *with*-variant (not allowing alternation), as in (23). Such verbs as *pour* can never occur in the *with*-variant, as in (24).

(22) a. Jack sprayed paint on the wall.
b. Jack sprayed the wall with paint.

(23) a. *Gina filled lemonade into the pitcher.
b. Gina filled the pitcher with lemonade.

(24) a. Carla poured lemonade into the pitcher.
b. *Carla poured the pitcher with lemonade.

(25) Alternating verbs:
Spray/Load Verbs: brush, cram, crowd, cultivate, dab, daub, drape, drizzle, dust, hang, heap, inject, jam, load, mound, pack, pile, plant, plaster, ?prick, pump, rub, scatter, seed, settle, sew, shower, slather, smear, smudge, sow, spatter, splash, splatter, spray, spread, sprinkle, spritz, squirt, stack, stick, stock, strew, string, stuff, swab, ?vest, ?wash, wrap

(26) Non-alternating *with* only:
*Fill Verbs: adorn, anoint, bandage, bathe, bestrew, bind, blanket, block, blot, bombard, carpet, choke, cloak, clog clutter, coat, contaminate, cover, dam, dapple, deck, decorate, deluge, dirty, douse, dot, drench, edge, embellish, emblazon, encircle, encrust, endow, enrich, entangle, face, festoon, fill, fleck, flood, frame, garland, garnish, imbue, impregnate, infect, inlay, interlace, interlard, interleave, intersperse, interweave, inundate, lard, lash, line, litter, mask, mottle, ornament, pad, pave, plate, plug, pollute, replenish repopulate, riddle, ring, ripple, robe, saturate, season, shroud, smother, soak, soil, speckle, splotch, splot staff, stain, stipple, stop up, stud, suffuse, surround, swaddle, swathe, taint, tile, trim, veil, vein, wreathe

(27) Non-alternating locative preposition only:
a. *Put Verbs: arrange, immerse, install, lodge, mount, place, position, put, set, situate, sling, stash, stow
b. *Verbs of Putting in a Spatial Configuration (except *hang*): dangle, lay, lean perch, rest, sit, stand, suspend
c. *Funnel Verbs: bang, channel, dip, dump, funnel, hammer, ladle, pound, push, rake, ram, scoop, scrape, shake, shovel, siphon, spoon, squeeze, squish, squash, sweep, tuck, wad, wedge, wipe, wring
d. *Verbs of Putting with a Specified Direction: drop, hoist, lift, lower, raise
e. *Pour Verbs: dribble, drip, pour, slop, slosh, spew, spill, spurt
f. *Coil Verbs: coil, curl, loop, roll, spin, twirl, twist, whirl, wind

From this alternation pattern, we could predict to see the Fill verbs, which occur in the *with*-variant only, more often with a *with*-PP than Put, Funnel, Pour, Coil groups of verbs, which occur in the locative variant only. The corpus data do show that the *with*-variant verbs occur more often with a *with*-PP than the locative variant verbs, as illustrated by the group means of these verbs in (28). Again, this alternation is specific to these groups of verbs that involve locative arguments, and thus if we are to posit this feature as a fixed variable, it would be hard to assign values to many other verbs.

(28) Means of response variable lWithPercent per verb class

Spray/Load	Fill	Put	Funnel	Lift
1.9710	2.0156	-1.0777	0.9002	-0.0408

There are two more alternations that involve *with*-PPs that Levin (1993) introduces. One is Fullfilling Alternation, examples of which are illustrated in (29) to (31), and the other is Image Impression Alternation, examples of which are demonstrated in (36) through (38) below (Levin 1993: 65-66). Depending on the syntactic behaviors regarding these alternations, verbs can be grouped into three, the verb lists being shown in (32) to (34) and (39) to (41) respectively.

(29) a. The judge presented a prize to the winner.
b. The judge presented the winner with a prize.

(30) a. The judge offered a prize to the winner.
b. *The judge offered the winner with a prize.

(31) a. *The judge saddled a prize to the winner.
b. The judge saddled the winner with a prize.

(32) Alternating verbs:
Verbs of Fulfilling: credit, entrust, furnish, issue, leave, present, provide, serve, supply, trust

(33) Non-alternating *to* only
*Verbs of Future Having: advance, allocate, allot assign, award, bequeath, cede, concede, extend, grant, guarantee, issue, leave, offer, owe, promise, vote, will, yield

(34) Non-alternating *with* only
*Equip Verbs: arm, burden, charge (with a task), compensate, equip, invest, ply, regale, reward, saddle

Interestingly again, the Equip verbs, which take the *with*-variant only, are most likely to occur with *with*-PPs, whereas the Offer verbs, which take the *to*-variant only, are least likely to occur with *with*-PPs, as shown below. Expectedly, the alternating Present verbs are in the middle, as shown below.

(35) Means of response variable lWithPercent per verb class

Present	Offer	Equip
-0.0296	-2.6825	1.5616

(36) a. The jeweller inscribed the name on the ring.
b. The jeweller inscribed the ring with the name.

(37) a. The jeweller copied the name on the ring.
b. *The jeweller copied the ring with the name.

(38) a. *The jeweller decorated the name on the ring.
b. The jeweller decorated the ring with the name.

(39) Alternating verbs:
Verbs of Image Impression: applique, emboss, embroider, engrave, etch, imprint, incise, inscribe, mark, paint, set, sign, stamp, tattoo

(40) Non-alternating locative preposition only
a. *Scribble Verbs: carve, chalk, charcoal, copy, crayon, doodle,

draw, forge, ink, paint, pencil, plot, pint, scratch, scrawl, scribble, sketch, spraypaint, stencil, trace, type, write

b. *Transcribe Verbs: copy, film, forge (signature), microfilm, photocopy, photograph, record, tape, televise, transcribe, type

(41) Non-alternating *with* only:
*Illustrate Verbs: address, adorn, autograph, brand, date, decorate, embellish, endorse, illuminate, illustrate, initial, label, letter, monogram, ornament, tag

Here again, Illustrate verbs, which take the *with*-variant only, are most likely to occur with *with*-PPs; Scribble or Transcribe verbs, which take the locative preposition only, are least likely to occur with *with*-PPs; finally the alternating Inscribe verbs are in the middle.

(42) Means of response variable lWithPercent per verb class

Inscribe	Scribble	Transcribe	Illustrate
0.7882	0.4054	-0.4959	1.8377

Documenting the extensive list of alternations of verbal arguments, Levin (1993) argues that the differences in verb behavior can be explained if the diathesis alternations are sensitive to particular components of verb meaning. She defines a group of verbs that share the same syntactic behavior in diathesis alternations as belonging to the same "verb class," and proposes 191 distinct verb classes. As seen above, the *with*-taking verbs participate in a variety of alternations, yet these are not the only alternations they participate in. For instance, among the verbs of contact that are listed in (20), the Swat and Spank verbs show the identical behaviors regarding the three alternations, but as shown below in (43) (* indicates 'not participating in the alternation.'), their behaviors differ in other alternations (that may not involve *with*-PPs) (Levin 1993: 148-150). This means that the two classes of verbs differ further in finer-grained meaning and thus cannot be grouped together into one verb class.

(43) a. Swat verbs: *With/Against Alternation, *Through/With Alternation, Conative Alternation, Body-Part Possessor Ascension Alternation, *Causative Alternation, *Middle Alternation, *Instrument Subject, Resultative Phrase, Zero-related Nominal
b. Spank verbs: *With/Against Alternation, *Through/With Alternation, *Conative Alternation, Body-Part Possessor Ascension Alternation, *Causative Alternation, *Middle Alternation, *Instrument Subject Alternation, Resultative Phrase, *-ing* Nominal

As such, the members of each verb class have in common a range of properties, including the possible expression and interpretation of their arguments, as well as the existence of certain morphologically related forms. The existence of regular relationships between verb meaning and verb behavior suggests that even though not all aspects of a verb's behavior need to be listed in its lexical entry, a speaker of English possesses the lexical knowledge of the meaning that determine the syntactic behavior of verbs and probably the general principles that determine behavior from verb meaning.

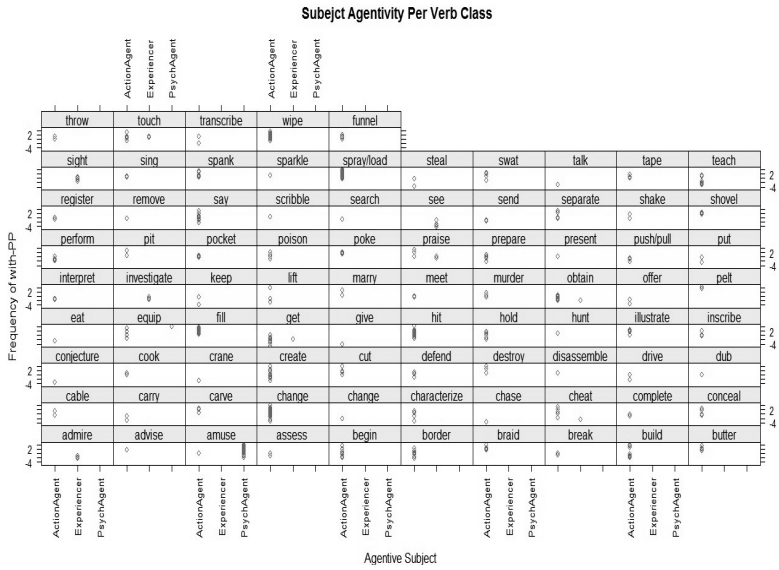
4. A mixed-effects model analysis

4.1 Fixed effects and random effects

In section 3, we have noted that there are alternations that will influence the frequency of *with*-PPs, and that unlike the Instrument Subject Alternation, they are specific to certain verb classes and thus may not be applicable to all *with*-taking verbs. Now, how can we encode these new features into the model? The effects exerted by the variables encoded in the simple regression model (Choi 2012) are "fixed" effects in the sense that the verbs in the data have the same relationship with the variables and that each observation is independent of the other observations. For example, as for variable AgentSbj (whether or not the subject is agentive), the possible levels for this variable (ActionAgent, PsychAgent, and Experiencer) are fixed and all verbs can equally be given a value for this variable. Even though this variable is used for another set of data (which may have nothing to do with instrument *with*-PP), the levels won't change and thus be repeated (Baayen 2008: 241).

As demonstrated above, we could see that observations from different verbs systematically differ from each other (Johnson 2008: 232). Some verbs may be biased more toward the occurrence of *with*-PP and others may be biased more toward the non-occurrence of *with*-PP. This means that we cannot posit the alternation effects reviewed in this section as fixed variables because the verbs would not have the same relationship with each variable and each observation is not independent of other observations. See, for example, how each verb class has a different relationship with variable AgentSbj, demonstrated in (44).

(44) Subject agentivity per verb class



As is graphically represented, each verb class demonstrates a distinct predisposition toward the presence of *with*-PP. It is not surprising to see for instance, on the bottom row, that PsychAgent verbs such as Amuse class show higher distributions of lWithPercent than Experiencer verbs such as Admire class. What is surprising though is that the same ActionAgent verbs show quite different patterns of lWithPercent distribution: see, for example, the third row from the bottom, from Conjecture to Dub.

As such, the effect of verb classes to the presence of *with*-PPs is not fixed but

rather "random," because it is item-specific and not repeatable to other words. To take care of this verb-specific lexical variation evidenced by differing alternation patterns, I categorized the 428 verbs in the data set into verb classes. I manually sorted and classified all 428 verbs into 85 verb classes, mostly based upon Levin's (1993) observations and classifications. For the verbs that are listed in her 1993 book, I followed Levin's verb classifications. For those that are listed under more than one verb class, the most relevant verb class was assigned based upon the usage in the corpus data. Other verbs that are not listed in Levin (1993) were grouped into the closest possible verb classes based on my personal judgment. According to the classifications, the top 20 verbs seen earlier grouped into the following verb classes in (45). How to incorporate this new factor VerbClass into the model will be discussed in the next section.

(45) Verb classes of top 20 verbs in the data set

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

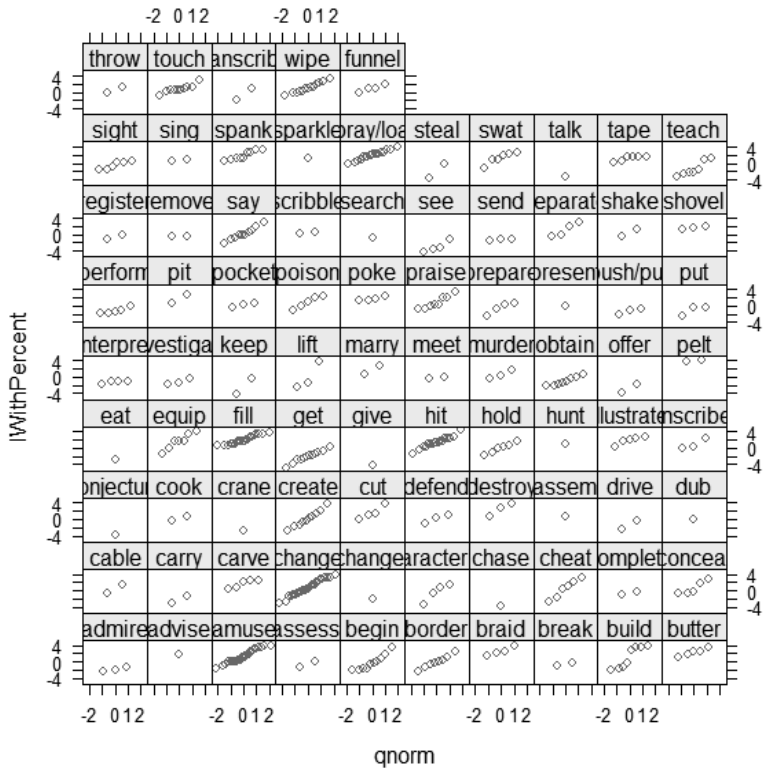
4.2 A mixed model with verb class as a random variable

The most efficient way to capture the individual lexical variance would be to use a hierarchical, multi-level regression, or mixed-effects modeling technique, which builds Verb Class into the model as an additional layer (Baayen 2008; Bresnan *et al*. 2007; Gelman and Hill 2007; Johnson 2008; Kuperman 2009; Pinheiro and Bates 2000).[5] In a mixed-effects model, variance is simultaneously estimated at the level of groups, as well as between individual verb classes. Mixed-effects models assume that a random effect factor (word-specific mean intercepts) follows close-to-normal distribution around the grand average, and has a mean of zero with unknown standard deviation to be estimated from the data (Baayen 2008: 241; Kuperman 2009). Unlike a fixed effect variable, whose coefficient is estimated by means of contrast to a specific reference level (e.g., the coefficient of WithRequired=Required being estimated against WithRequired =Optional), a random variable is adjusted by means of contrasts with respect to the population mean (i.e. with each other) (Baayen 2008: 242; Kuperman 2009). As seen in the previous section, verbs that belong to a certain verb class tend to take *with*-PPs more frequently than other verbs belonging to another verb class. Observations seem to behave differently depending on the verb class, which is visualized in (46). These differences will be captured by adjustments for individual verb classes, which will vary around zero (i.e., the adjustment mean) with some unknown standard deviation.

---

5 There could be several different measures to incorporate individual variance by verb class into the model. For one, we could include verb class as another fixed-effect factor with each verb class as a level of the factor, namely with 85 levels. Although this model is conceptually possible, it suffers from huge waste of degrees of freedom, which will make it an inefficient model (Kuperman 2009). Alternatively, we could run a separate analysis for each individual verb (Lorch and Myers 1990), ending up with 85 models. This will of course be inefficient; moreover, if we needed to consider speakers (subject) as another random variable, we would not be able to analyze both speaker (subject) and verb (item) at the same time.

(46) Variance of lWithPercent per verb class



To capture this variance, we can add the random variable VerbClass to the simple fixed-effect model, as shown in (47). The random-effects term (1|VerbClass) specifies that the model will make such by-VerbClass adjustments for the average frequency (lWithPercent) by means of small changes to the intercept (Baayen 2008: 244-245). Recall that in linear models the intercept provides a kind of baseline mean. Therefore, the random component produces a separate intercept value for each verb: namely, this model does not assume that average probability of *with*-PP is the same for each verb class (Johnson 2008: 237). Lowering the intercept for a verb class implies that the mean probability of *with*-PP for that verb class is lower; by contrast, raising the intercept for a verb class implies that the probability of *with*-PP for that verb class is higher.

(47) Linear mixed-effects model

lWithPercent ~ WithRequired + InstV + AgentSbj + InstSbj + (1|VerbClass)

(48) Fixed effects

	Estimate	Std Error	df	t value	Pr(>|t|)
(Intercept)	-0.2488	0.1880	175.6000	-1.324	0.1873
WithRequired:Required	0.5827	0.2122	414.3000	2.746	0.0063
InstV:Yes	0.7869	0.3123	421.8000	2.519	0.0121
AgentSbj:Experiencer	-1.1293	0.4123	283.5000	-2.739	0.0065
AgentSbj:PsychAgent	2.0153	0.7586	144.8000	2.657	0.0088
InstSbj:Yes	0.2956	0.1715	418.4000	1.724	0.0855

(49) Random effect

Random effects:

Groups	Name	Variance	Std.Dev.
VerbClass	(Intercept)	1.010	1.005
Residual		2.001	1.415

Number of obs: 428, groups: VerbClass, 85

The first to note in this mixed model is the status of fixed variable InstSbj. Unlike that of the simple model in (3), the p-value of InstSbj is bigger than 0.05, so is not valid any more, as illustrated in (48). As a matter of fact, it is not unusual that the effect of a fixed variable decreases or even disappears when a random variable is added, especially when the random variable is related to the fixed variable (Baayen 2008: 281; Kuperman 2009). As we discussed in section 3, Instrument Subject Alternation is also a particular feature relevant only to certain groups of verbs. Now that we factor in the verb classes that share the same syntactic and semantic characteristics as the random variable VerbClass, the role of fixed variable InstSbj naturally weakens. Therefore, the insignificant fixed variable InstSbj will be discarded out of the model and a new model is tried as in (50).

(50) Revised mixed-effect model

lWithPercent ~ WithRequired + InstV + AgentSbj + (1 | VerbClass)

The revised mixed-effects model has all its fixed effects valid as illustrated by the p-values in (51).

(51) Fixed effects

	Estimate	Std Error	df	t value	Pr(>|t|)
(Intercept)	-0.0737	0.1589	102.1000	-0.464	0.6437
WithRequiredRequired	0.6044	0.2123	416.2000	2.846	0.0046
InstVYes	0.7806	0.3130	422.7000	2.494	0.0130
AgentSbjExperiencer	-1.2132	0.4111	282.7000	-2.951	0.0034
AgentSbjPsychAgent	1.9904	0.7631	149.0000	2.608	0.0100

(52) Random effect

Groups	Name	Variance	Std.Dev.
VerbClass	(Intercept)	1.037	1.018
Residual		2.004	1.416

Number of obs: 428, groups: VerbClass, 85

Now, all the fixed variables are valid, as demonstrated in (51). The intercepts of the fixed effects can be read in the same manner as they were in the simple model in (3) although specific numbers are a little different. Let us turn to the random effect VerbClass, in (52). The model shows that the observed 428 individual verbs are grouped into 85 different verb classes. The variance among these 85 verb classes is 1.037 and standard variation is 1.018. It shows that verb classes do vary in their intercepts.[6] Residual stands for the residual error, the unexplained variance. Note that the means of this random variable is not listed, as it is always zero.

Then, to test whether the model's fixed terms are significant, ANOVA (Analysis of Variance) was conducted and as shown in (53), all three fixed variables are significant, as confirmed by small p-values. See (4) again for interpretation of the ANOVA result.

---

6 The variance and standard deviation of a mixed model are used as an indicator to evaluate the model: a better model has smaller variance and standard variation of the random effect. For instance, the random-effect-only model (lWithPercent ~ (1|VerbClass)) has the variance of 1.495 and standard deviation of 1.223), which are bigger than those of the current mixed model. Therefore, we can say that the current mixed-effects model works better than the random-effect-only model.

(53) ANOVA on the model

	Sum Sq	Mean Sq	Num DF	DenDF	F value	Pr(>F)
WithRequired	16.240	16.240	1	416.20	8.1025	0.0046
InstV	12.465	12.465	1	422.67	6.2189	0.0130
AgentSbj	32.277	16.139	2	195.35	8.0519	0.0004

Also, by performing a backward elimination of all effects, we can reconfirm that both the fixed effects and random effect of this linear mixed-effects model are valid. As tested by the likelihood ratio test, the random effect is significant, demonstrated in (54).

(54) Backward elimination on random effect (Likelihood ratio test)

	Chi.sq	Chi.DF	elim.num	p.value
VerbClass	69.41	1		kept < 1e-07

Also, the Least Squares Means method on fixed effects proves their significance with the Confidence Intervals and p-values, shown in (55).

(55) Backward elimination on fixed effects (Least squares means)

	DF	t-value	Lower CI	Upper CI	p-value
WithRequired:Required	180	3.33	0.4811	1.108	0.0010
InstV:Yes	261	3.13	0.4701	2.066	0.0020
AgentSbj:ActionAgent	151	3.35	0.2536	0.984	0.0010
AgentSbj:PsychAgent	144	3.38	1.0821	4.136	0.0009

Finally, the R^2 (R-Squared) calculated for the new mixed model.[7] R^2 indicates the proportion of the variance in the response variable that is predictable from the independent variable, namely, the proportion of the variance of lWithPercent explained by the model.

7 The current mixed-effects model does not provide the R^2 value of the model. So, it was calculated manually as squared correlation coefficient by comparing the expected frequencies with observed frequencies.

(56)

Model	Model with Fixed-Effects	Model with Random Effect	Model with Mixed Effects
R^2	0.1518	0.4886	0.4972

As illustrated in (55), The R^2 of the mixed-effects model has improved a great deal compared with that of the fixed-effects-only model and it is also better than that of the random-effect-only model.[8] Therefore, the new mixed model that considers the by-verb lexical variation as a random variable is justified. In other words, the presence of instrument *with*-PPs can be better explained when we take into account the syntactic and semantic characteristics of individual verbs or verb classes.

5. Conclusion

This paper has investigated whether a lexical variance caused by verb class influences the choice and frequency of instrument *with*-PPs. The corpus data shows that observations grouped by verb class demonstrate systematically different behaviors and this individual variance of verb classes can be captured by means of a random effect of a mixed-effects model. The new statistical technique of hierarchical, multi-level, mixed-effects modeling (Baayen 2008; Bresnan *et al*. 2007; Gelman and Hill 2007; Johnson 2008; Kuperman 2009; Pinheiro and Bates 2000) can process across-word fixed effects and by-word random effects together. The newly proposed mixed model demonstrates that verb class is indeed a significant random variable, and this model proves to be more powerful and accurate than the simple regression model proposed by Choi (2012), which takes only the fixed effects into account.

The corpus modeling analysis proposed in this paper raises a question regarding the traditional view of subcategorization or argument structure, because it assumes the presence and the nature (argumenthood) of such syntactic expression as an instrument *with*-PP is not a matter of categorical distinction. As is visually presented in (46), the degree to which each verb class takes an instrument *with*-PP varies

8 Although it is the best among the three models, the R^2 of the mixed model is still below 0.5. However, this does not mean that the model is not good because in such fields that attempt to predict human behavior (language is a characteristic human behavior), it is expected that the R-squared values will be low, typically lower than 50%.

systematically and this complex lexico-syntactic knowledge of native speakers cannot be possibly explained within the rigid framework of categorical grammar. The current statistical modeling analysis can account for this basic but subtle grammatical knowledge of argument-taking by assuming that grammar can be gradient, grammatical judgment may be on continuum, and grammatical rules or constraints may be applied in various strengths or magnitudes, just like the variables in different effect sizes of the model.

References

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Boume, Irene Kraemer, and Joost Zwarts (eds.), *Cognitive Foundations of Interpretation*, 69-94. Amsterdam: Royal Netherlands Academy of Science.

Bresnan, Joan and Jennifer Hay. 2008. Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118: 245-259.

Bresnan, Joan and Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86: 168-213.

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. *Syntax and Semantics 21: 263-291.* New York: Academic Press.

Choi, Hye-Won. 2010a. Capturing lexical variance with a mixed model: Verb in ordering variation. *Discourse and Cognition* 17(1): 113-137.

Choi, Hye-Won. 2010b. 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. *Studies in Generative Grammar* 21: 75-96.

Choi, Hye-Won. 2012. A linear regressional analysis of *with*-PPs in English. *Linguistic Research* 29(1): 21-44.

Fillmore, Charles J. 1968. The case for case. In Emmon Bach and Robert Harms (eds.), *Universals in Linguistic Theory,* 1-88. New York: Holt, Rinehart, and Winston.

Gelman, Andrew and Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

Jackendoff, Ray. 1977. *X-bar syntax: A study of phrase structure*. Cambridge, MA: The MIT Press.

Johnson, Keith. 2008. *Quantitative methods in linguistics*. Oxford: Blackwell.

Koenig, Jean-Pierre, Gail Mauner, and Breton Bienvenue. 2003. Arguments for adjuncts. *Cognition* 89: 67-103.

Kuperman, Victor. 2009. Mixed-effects modeling. Presented at Stanford University, March 2009.

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Lorch, Robert F. and Jerome L. Myers. 1990. Regression analysis of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16: 149-157.

Manning, Christopher. 2003. Probabilistic syntax. In Rens Bod, Jennifer Hay, and Stefanie Jannedy, *Probabilistic Linguistics*, 289-314. Cambridge, MA: The MIT Press.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: The MIT Press.

Pinheiro, Jose C. and Douglas M. Bates. 2000. Mixed-effects models in S and S-PLUS. *Statistics and Computing*. New York: Springer.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [https://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.

Szmrecsanyi, Benedikt. 2005. Language users as creature of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1: 113-150.

Hye-Won Choi
Department of English Language and Literature
Ewha Womans University
52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea
E-mail: hwchoi@ewha.ac.kr

Received: 2017. 04. 17.
Revised: 2017. 06. 14.
Accepted: 2017. 06. 14.