Is syntactic knowledge probabilistic? Experiments with the English dative alternation

Joan Bresnan

Theoretical linguistics traditionally relies on linguistic intuitions such as grammaticality judgments for data. But the massive growth of language technologies has made the spontaneous use of language in natural settings a rich and easily accessible alternative source of data. Moreover, studies of usage as well as intuitive judgments have shown that linguistic intuitions of grammaticality are deeply flawed, because (1) they seriously underestimate the space of grammatical possibility by ignoring the effects of multiple conflicting formal, semantic, and contextual constraints, and (2) they may reflect probability instead of grammaticality. Both of these points are richly exemplified by studies of the English dative alternation (Green 1971; Gries 2003, 2005; Fellbaum 2005; Bresnan & Nikitina 2003; Bresnan, Cueni, Nikitina & Baayen in press; Lapata 1999; Bresnan & Hay 2006; Hay & Bresnan 2006), which is the linguistic domain of the present study.

The present study discusses two experiments following up Bresnan et al. (in press). The first indicates that the "soft" generalizations found in corpus studies of the dative alternation reappear in subjects' intuitions of grammaticality in context, and that language users have substantial knowledge on the basis of these generalizations of what others are going to say (meaning here the choice of syntactic structure to convey the message). The second experiment shows that rare constructions that have been considered ungrammatical by many linguistic theorists are judged natural by speakers when the appropriate soft conditions are met. Intuitive contrasts in grammaticality that many linguists have reported seem to reflect probabilities rather than categorical constraints.

Background

The English dative alternation is illustrated in (1):

(1) a. *Who gave you that wonderful watch?* ← double object construction
b. *Who gave that wonderful watch to you?* ← prepositional dative

Although alternative forms often have different meanings (Pinker 1989; Levin 1993; Rappaport-Hovav & Levin 2005), frequently explained in terms of "the principle of contrast" (E. Clark 1987), the alternatives in (1a,b) are very close paraphrases, and the flexibility afforded by their violation of the principle of contrast appears to have functional advantages in sentence production (V. Ferreira 1996). Moreover, subtle intuitions of fine-grained semantic differences between syntactic constructions have turned out in many cases to be inconsistent and unreliable (Fellbaum 2005; Bresnan & Nikitina 2003; Bresnan, Cueni, Nikitina & Baayen in press; Bresnan 2006). We therefore view the prepositional dative and double object constructions as having overlapping meanings which permit them to be used as alternative expressions or paraphrases.

Previous studies have shown that the probability of using one or the other of these two alternatives – the double object construction or the prepositional dative – is associated with the verb and its semantic class (Lapata 1999; Gries 2005) and is respectively in-creased/decreased when the first phrase following the verb is a pronoun/lexical noun phrase, is definite/indefinite, refers to a highly accessible referent/a referent not previously mentioned, refers to a human/non-human, or is shorter/longer (Bock & Irwin 1980; Thompson 1990; Bock, Loebell & Morey 1992; Hawkins 1994; Collins 1995; Prat-Sala & Branigan 2000; Arnold et al. 2000; Snyder 2003; Wasow 2002; Gries 2003). From these and other variables such as the previous occurrence of a parallel structure (Bock 1986; Pickering, Branigan & McLean 2002; Gries 2005; Szmrecsányi 2005), it is possible to predict the dative alternation (that is, predict which alternative is used: (1a) or (1b)) in spoken English with 94% accuracy (Bresnan, Cueni, Nikitina & Baayen in press).

Bresnan et al. show that their model generalizes beyond the contingencies of the particular collection of telephone conversations that constitutes their spoken dative database and predicts statistical differences in a very different written corpus of edited reportage. The generalizability of the model raises the question of whether it represents some aspects of the implicit knowledge of English language users.

Experiment 1

If the probabilistic model of Bresnan et al. captures the implicit knowledge of English language users, then theoretically *language users could predict the dative syntax choices that speakers make,* as a function of the same

kinds of variables – just as the model does. Where the corpus model predicts high or low probabilities, subjects should also do so, and where the model predicts middle-range probabilities (underdetermining dative syntax choices), subjects should do so as well.

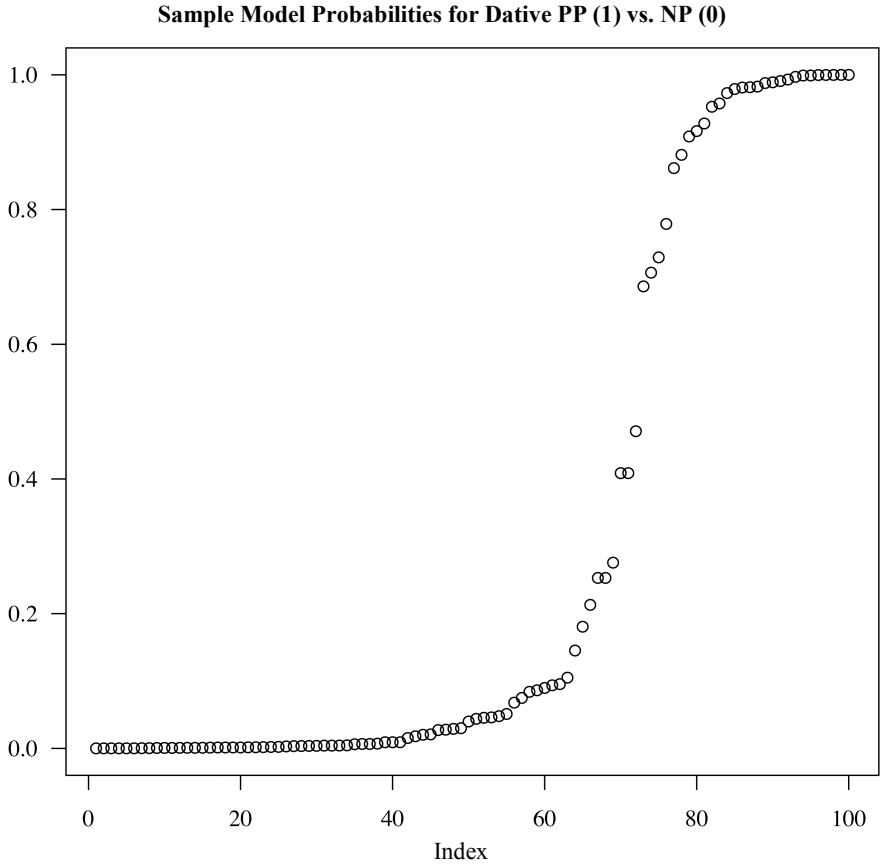

Figure 1. Sample probabilities from the corpus model of Bresnan et al.

Figure 1 shows the model probabilities of a prepositional dative construction for a random sample of one hundred observations of the alternating verbs from the Bresnan et al. spoken corpus dataset of 2360 observations. The data points at the top of the vertical *y* axis scale have probabilities near 1 of being a prepositional construction, those at the bottom have probabilities near 0. In this model of the binary choice between the two alternative

dative paraphrases illustrated in (1a,b), low probability of being a prepositional dative construction is equivalent to high probability of being a double object construction, so the points at the bottom are almost always realized in the double object construction. The prevalence of data points near the zero end of the scale (the bottom of the *y* axis) reflects the overall skewedness of the data toward double object constructions, which constitute 79% of the total observations. The data points in the middle of the *y* axis scale are cases where both of the alternative constructions have substantial probability – 50/50, 60/40, and the like.

Hypothesis

The specific hypothesis investigated in Experiment 1 is this: given the same multivariable information as the corpus model, including contextual information from the original dialogues, subjects will make ratings of alternative dative constructions like (1a,b) that correspond to the corpus model probabilities.

Method

The task was inspired by Rosenbach's (2003) experiment on the genitive alternation, which required subjects to choose between alternative possessive constructions as continuations of edited passages excerpted from a novel. The present experiment introduces several differences in method. First, the items are built from randomly sampled transcriptions of spoken dialogue passages, rather than selected literary passages in accordance with a factorial design. Second, subjects are given a scalar instead of a binary rating task. And third, subjects' responses are analyzed as a function of the original corpus model predictor variables by using mixed effects regression (Pinheiro & Bates 2000; Bates & Sarkar 2006). This type of regression can model the responses as a function of the linguistic predictors while simultaneously taking into ac-count the clusters of data caused by multiple observations from both of the randomly sampled elements – the experimental subjects and the dative verbs.

The experimental items were chosen by randomly sampling observations in the dative corpus data from the centers of five equal probability bins defined by the corpus model, ranging from very low probability of being a prepositional dative to very high. Potentially ambiguous items were excluded. The item probabilities are shown in Figure 2.

For each sampled observation an alternative paraphrase was constructed, and both were presented as choices in the original dialogue context, which

was edited for readability by shortening and by removing disfluencies. Items were pseudo-randomized and construction choices were alternated to make up a questionnaire. The subjects were nineteen paid Stanford undergraduates of both genders who reported that they were monolingual and had not taken a syntax course. Each subject received the same questionnaire, with the same order of items and construction choices. Figure 3 displays a sample item.

Subjects were asked to rate the naturalness of the alternatives in the given context by distributing 100 rating points over the two alternatives in accordance with their own intuitions. Any pair of scores summing to 100 was permitted, including 0–100, 63–37, 50–50, etc.

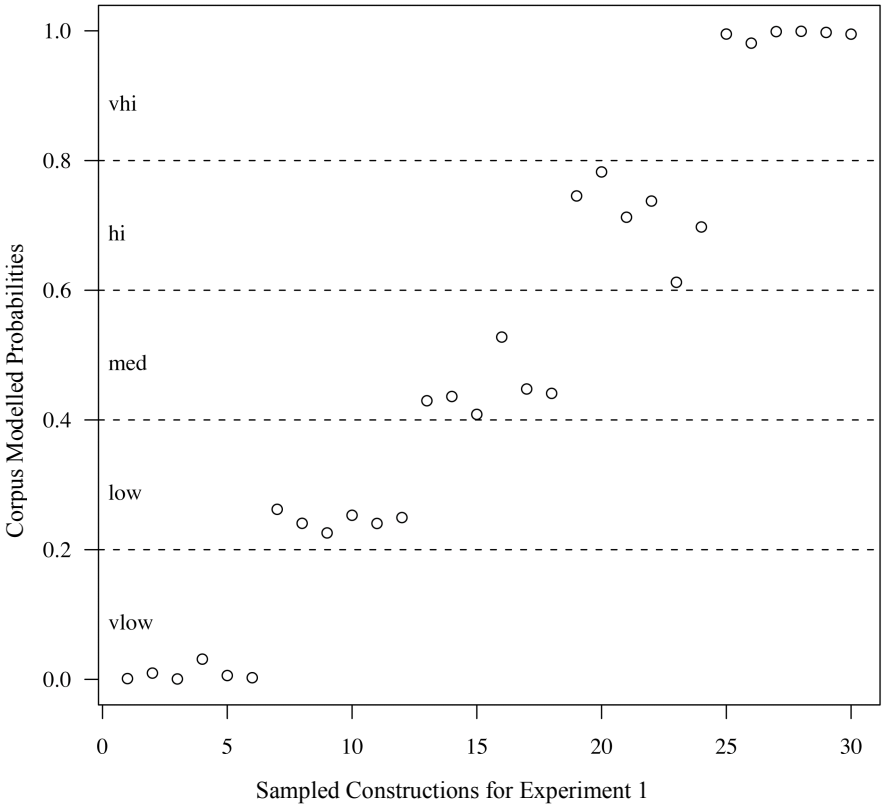

Figure 2. Probability bins of items for Experiment 1

Speaker:

About twenty-five, twenty-six years ago, my brother-in-law showed up in my front yard pulling a trailer. And in this trailer he had a pony, which I didn't know he was bringing. And so over the weekend I had to go out and find some wood and put up some kind of a structure to house that pony,

(1) because he brought the pony to my children.
(2) because he brought my children the pony.

Figure 3. Sample item for Experiment 1

Results

Plots of the data suggest that subjects' scores of the naturalness of the alternative syntactic paraphrases correlate with the corpus model probabilities. Figure 4 shows the mean subject scores for each item plotted against the corpus model probability of the item. The line is a nonparametric smoother which indicates the trend of the data by averaging local values; it shows a roughly linear correspondence between the corpus model probabilities and the mean item scores. Note that the items in the middle probability bins overlap far more in average ratings than those in the extreme bins, indicating that average subjects' scores are most indecisive where the corpus model is least accurate.

In Figure 5 each panel shows a single subject's mean scores for the items in each corpus probability bin. (The subject numbers are not contiguous because data from seven of twenty-six who completed the questionnaire were excluded because they reported they were either bilingual or had taken a syntax class.) All of the subjects' mean ratings of items from the lowest probability bin are below their mean ratings of items from the highest probability bin. The ratings of items from the middle bins tend to fall in the middle of each subject's rating range, though their relative rankings vary quite a bit across subjects, as expected from the original corpus model probabilities (Figure 1).

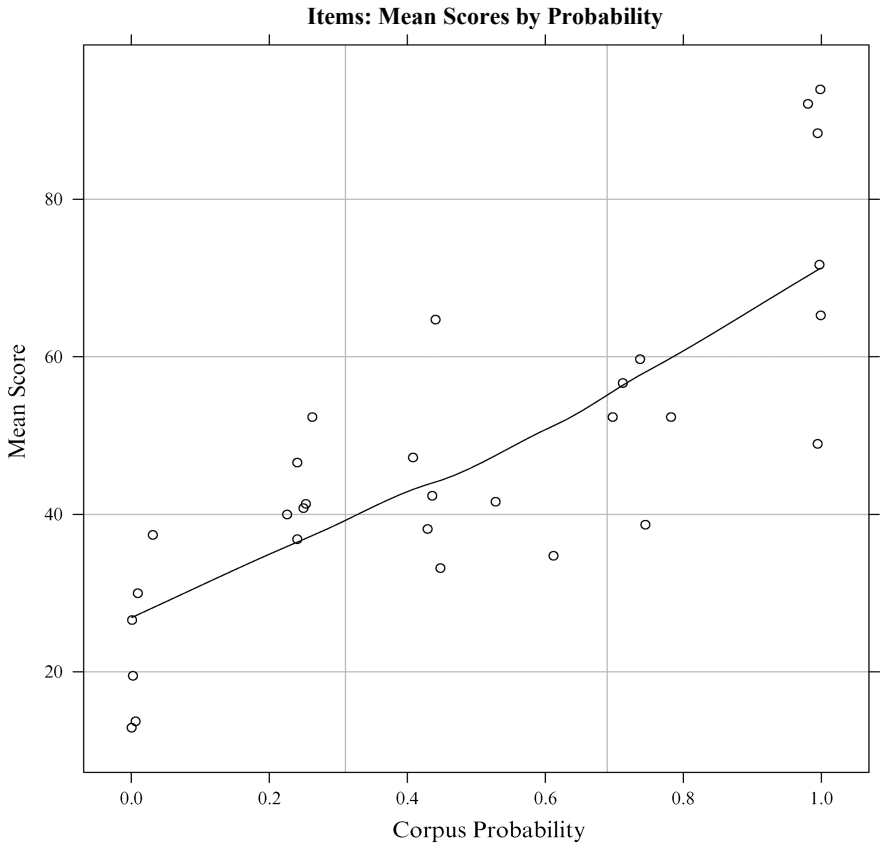

Figure 4. Mean Experiment 1 Scores of Items by Probability Bin

The results were analyzed using a linear mixed effects regression model (Pinheiro & Bates 2000; Baayen 2004; Bates & Sarkar 2006), which fit the scores using adjustments for both subject and verb sense as random effects and adjustments for fixed effects conditioned on the random effects.[1]

The fixed effects were taken from the original corpus mode of Bresnan et al. together with the order of items, the order of construction choice and the lemma frequency of the verbs according to the CELEX database (Baayen, Piepenbrock & Gulikers 1995). The last three effects were eliminated from the model because their coefficients were less than their standard errors (Chatterjee, Hadi & Price 2000: 286–288). Table 1 shows that the 95% confidence intervals of all remaining factors except for givenness of recipient exclude 0, indicating a significant effect on the response.

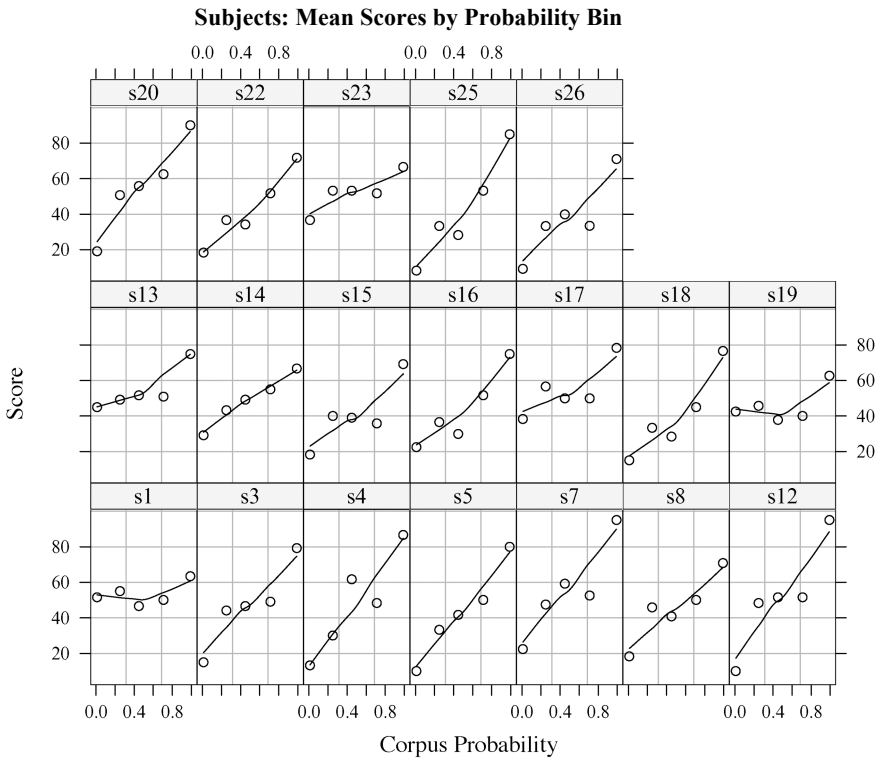

Figure 5. Mean Experiment 1 Scores by Probability Bin for Each Subject

Table 1. Model Coefficients for Experiment 1

Fixed effects:	95% Conf. Limits		
	Estimate	Lower	Upper
(Intercept)	73.19	45.70	102.22
pronominality of theme = pronoun	16.91	10.48	23.28
definiteness of theme = indefinite	–12.48	–17.57	–7.39
givenness of theme = non-given	–14.77	–19.62	–9.92
pronominality of recipient = pronoun	–22.47	–33.25	–11.85
definiteness of recipient = indefinite	14.13	5.58	22.98
givenness recipient = non-given	-9.00	–19.43	1.42
animacy of recipient = inanimate	–29.48	–43.75	–15.66
parallelism of PP	16.70	8.73	24.67
argument length difference (log scale)	–4.77	–9.37	–0.12

Number of observations: 570, groups: subject 19; verb sense 11.

By examining the model coefficients in Table 1 we can interpret the results. The coefficients show the magnitudes and directions of the effects: these are consistent with the *harmonic alignment* effects in the original corpus model (Bresnan et al), which has been observed in many previous corpus studies (Thompson 1990; Collins 1995): nongiven or indefinite themes and pronominal recipients favor V NP NP, pronominal themes and indefinite recipients favor V NP PP. Contrary to Bresnan et al's model, inanimate recipients favor V NP NP, but there are only two such items in the sample used in Experiment 1 and both occur with abstract senses of verbs, which strongly favor the double object construction (Bresnan & Nikitina 2003).

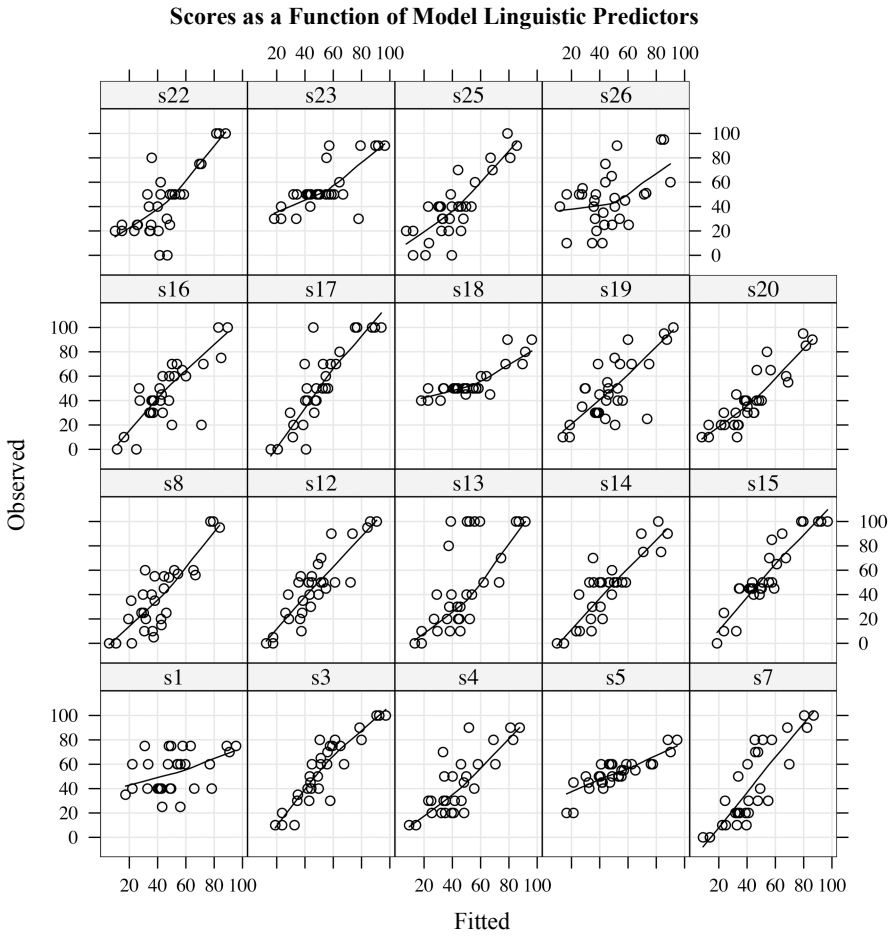

Figure 6. Fit of linear mixed effects model to Experiment 1 scores

Finally, the fit of the experimental model (R^2 =0.61) is displayed in Figure 6, a trellis graph with nonparametric smoothing lines to facilitate visualization of the data (Cleveland 1979). Each panel of the trellis plot is a scatterplot of the data from a single subject, showing all thirty scores (represented on the *y* axis) plotted against the fitted model values (represented on the *x* axis). A roughly linear relation appears in each panel, indicating a good fit of the model variables to the score data.

These results show that subjects' scores of the naturalness of the alternative syntactic structures correlate very well with the corpus model probabilities and can be substantially explained as a function of the same predictors as the original corpus model. In fact, as shown in Table 2, the subjects' preferred choices, which were made according to their own intuitions, reliably tended to pick out the same choices made by the original dialogue participants in the corpus transcriptions. If they had invariably preferred the double object construction in every item, their responses would have matched 57% of the original speakers' choices; this is the baseline in Table 2. In actuality, their responses matched the original choices well over the baseline. Their ratings are thus good predictors of what the speakers would say.

Table 2. Proportions of subjects' ratings favoring actual corpus choices

0.63	0.83	0.80	0.70	
0.80	0.80	0.67	0.77	
0.73	0.83	0.80	0.77	
0.80	0.77	0.77	0.73	
0.73	0.87	0.67		Baseline =0.57

Experiment 2

Experiment 1 suggests that language users' implicit knowledge of the dative alternation in context reflects the usage probabilities of the construction. In Experiment 2 we ask whether linguistic manipulations that raise or lower probabilities influence grammaticality judgments.

Mismatches between grammaticality judgments reported by linguists and the actual language use of speakers and writers are surprisingly common, particularly in areas of theoretical syntax and semantics where subtle contrasts are invoked. A variety of cases are discussed in Bresnan (2006). The English dative alternation provides one such case, illustrated in (2) and (3),

where the double object constructions reported by linguists to be ungrammatical with verbs like *drag* and *whisper* are found in actual usage (Bresnan & Nikitina 2003; Bresnan et al. in press). (In example (2a), *Sumomo* is the name of a small robot servant.)

(2) a. …while Sumomo dragged him a can of beer. ← attested example
b. *I dragged John the box. ← reported grammaticality judgment

(3) a. *She came back and whispered me the price.* ← attested example
b. **Susan whispered Rachel the news* ← reported grammaticality judgment

Although we lack specific probability estimates for all of the relevant verbs, we know that differing alternation classes of dative verbs correspond to differing frequencies of use in internet samples (Lapata 1999), and that different argument types are more likely to occur in different syntactic positions following dative verbs (Thompson 1990; Collins 1995; Bresnan et al. in press). In particular, double object constructions in which a pronoun precedes a lexical NP are far more frequent than those in which two lexical NPs occur, as shown in Table 3, and it is in the more frequent contexts that reportedly non-alternating dative verbs can most readily be found in actual use.

Table 3. Frequency of Dative Double Object Constructions in SWITCHBOARD

V […Pronoun…] NP	V […Noun…] NP
1530	178

Thus *drag* and *whisper* are reported to be ungrammatical in the double object construction, but Google queries yield examples in the more frequent construction types (2a) and (3a), along with *dragged the body to the king* and *whisper the password to the fat lady*. The reportedly ungrammatical examples constructed by linguists as in (2b) and (3b), tend to utilize the far less frequent positionings of argument types, like *drag the king the body* and *whisper the fat lady the answer*.

Hypothesis

Subjects' ratings of the reportedly ungrammatical dative constructions will indicate grammaticality when the probability of the syntactic context is higher.

Method

Experiment 2 used the same task as Experiment 1. Six alternating and eight reportedly non-alternating verbs were sampled from the internet. There were three alternating verbs of communication by instrument verbs ('a cm') – *phone, text, IM* – and three alternating verbs of instantaneous transfer ('a tr') – *flip, throw, toss*. There were four reportedly non-alternating verbs of manner of communication ('n cm') – *whisper, mutter, mumble, yell* – and four reportedly non-alternating verbs of continuous transfer ('n tr') – *carry, push, drag, lower* (Pinker 1989). The verb types are summarized in Table 4.

All of the verbs were sampled in the construction types found to be most frequent in corpus studies—the double object construction with pronoun recipient preceding lexical NP theme or the dative construction with a lexical NP as prepositional object.

Table 4. Verbs used in Experiment 2

Communication		Transfer	
Alternating	Non-Alternating	Alternating	Non-Alternating
'a cm'	'n cm'	'a tr'	'n tr'
phone	*whisper*	*flip*	*carry*
text	*mutter*	*throw*	*push*
IM	*mumble*	*toss*	*drag*
	yell		*lower*

Table 5. Contexts for each verb

V […Pronoun…] NP	(sampled)
V NP to […Pronoun…]	(constructed)
V NP to […Noun…]	(sampled)
V […Noun…] NP	(constructed)

```
-----------------------------------------------------------------
Money in the pot is dead money. It does not belong
to anyone until the hand is over

(1) and the dealer pushes someone the pot.
(2) and the dealer pushes the pot to someone.
-----------------------------------------------------------------
```

Figure 7. Sample item for Experiment 2

With each sample the preceding context was obtained for discourse cohesion and the presence of any parallel structures, which are known to influence syntactic choices (see Szmrecsányi 2005 for references).

An alternative to each sampled sentence was created in the opposite construction type. For example, for the sample sentence containing *whisper me the price* the alternative *whisper the price to me* was created; and for a sample containing *whisper the password to the fat lady*, an alternative *whisper the fat lady the password* was created. Similarly, for the sample sentence containing *toss the ball to Worthy*, the alternative *toss Worthy the ball* was constructed; and for *toss me the socks*, *toss the socks to me* was constructed. Thus, each verb in each semantic class occurred in the four conditions shown in Table 5. (The data also included two instances of *someone* sampled in the prepositional dative construction and one instance of *someone* sampled in the double object construction.)

The same method of creating a questionnaire was used as in Experiment 1. Figure 7 displays a sample item for the reportedly non-alternating verb *push*.

The subjects were twenty paid Stanford undergraduates of both genders who reported that they were monolingual and had not taken a syntax course. They were given the same forced-choice scalar rating task as in Experiment 1.

Results

A plot of the data is given in Figure 8. In this and subsequent plots the vertical axis high score limit now shows the top rating for the *double object* construction, because this is precisely the construction which is at issue—found in actual usage but judged ungrammatical by linguists. Figure 8 shows that the ranges of subjects' mean scores of the double object constructions appear to differ by both semantic class of the verb and pronominality of the recipient.

The columns represent the verb classes shown in Table 4: in each panel, the first and third classes are alternating ('acm', 'atr'), while the second and fourth are reportedly non-alternating ('ncm' and 'ntr'). The black dots designate themiddles of the ranges of mean scores in each verb class, the boxes are the interquartile ranges, and circled points falling outside of the dashed lines are potential outliers. The panel labeled 'V […Noun…] NP' on the left represents the less frequent type in which both objects are lexical NPs; the panel labeled 'V […Pron…] NP' on the right represents the very frequent type in which a pronoun object precedes a lexical NP object.

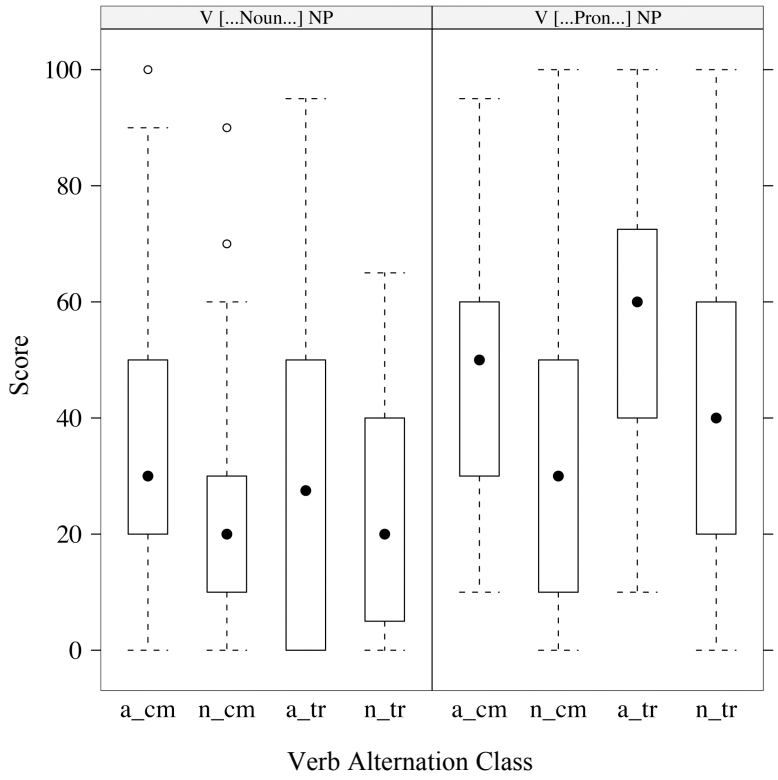

Figure 8. Ranges of subjects' mean scores for double object constructions by semantic class of verb and pronominality of recipient

Looking within each panel of Figure 8, we see that given a particular structure type, V […Pronoun…] NP or V […Noun…] NP, the median scores (black dots) for the alternating verbs appear higher than those for the non-alternating verbs. Looking across the two panels, we see that all the median scores appear higher for double object constructions of the more frequent argument type (V […Pronoun…] NP) than for the less frequent type (V […Noun…] NP), regardless of verb class.

Strikingly, the median scores (black dots) for the reportedly non-alternating verb classes in the V […Pronoun…] NP structure appear as high as or higher than those for the alternating verb classes in the V […Noun…] NP structure. This means that the reportedly ungrammatical verb classes appear to be rated as highly in the frequent context as the grammatical verb classes in the infrequent context. (The latter are supposed to be fully grammatical

by definition as alternating verbs.) In other words, relative frequency of argument types seems to override and reverse linguists' reported classifications of relative grammaticality.

To analyze the significance of the results, a linear mixed effects regression model was fit with both verb and subject as random effects and with the fixed effects of pronominality of recipient, semantic class, and item order. An interaction between the random effect of verb and pronominality of recipient was also included to take account of possible individual differences between verbs in their selectivity for the recipient type (pronoun or lexical noun head) – whether by prosodic, stylistic, or other differences. Such a term allows for variable adjustments to the verb estimates for both recipient types and it significantly improved the overall loglikelihood of the model, $Pr(>Chisq)3.358e-06$. Construction order and verb lemma frequency were not significant and were dropped from the final model because their coefficients were less than their standard errors.

As seen in Tables 3 and 5, the least frequent syntactic contexts for dative verbs – prepositional dative pronouns and lexical noun objects – were constructed, because they were non-occurring in the usage samples for the non-alternating double object verbs. This introduces a possible confound between the syntactic context types and the naturalness of the discourse passage. To measure the influence of the specific context on the choice of syntactic construction, all of the stimuli were annotated for discourse givenness of recipient and theme and the presence of a parallel construction—double object or prepositional dative – in the preceding context. Then these four factors were tested in the model: givenness of recipient and theme in the discourse context and the existence of a prior parallel double object or dative prepositional construction. All four were insignificant, with coefficients less than the standard errors, and were dropped from the final model.[2] All of the recipients were animate and all of the themes inanimate, so these factors were not included in any of the models.

Table 6 shows the 95% confidence intervals of the remaining variables of semantic class, pronominality of recipient, and item order. The intercept is the estimate for the nonalternating communication semantic class ('n cm') of verbs (*whisper, mutter, mumble, yell*) with lexical noun recipients; these constitute the reference set against which the other predictor values are contrasted. These verbs are also by intuitive judgments the lowest-rated class of verbs in the double object construction, as we see from Figure 8: the top of the interquartile range of mean scores (represented by the vertical rectangle) is lower than all the others.

Table 6. Model Coefficients for Experiment 2

Fixed effects:		95% Conf. Limits	
	Estimate	Lower	Upper
(Intercept)	14.50	4.65	24.45
semantic class=n_tr	6.93	–5.64	14.62
semantic class=a_tr	16.86	1.76	25.47
semantic class=a_cm	11.84	0.46	22.24
pronominality of recipient=pronoun	13.89	4.65	22.58
item order	0.42	0.13	0.85

Number of observations: 600, groups: subject 20; verb 14.

We can interpret Table 6 as follows. Both the pronominality of recipient and item order coefficients are positive and both of their 95% confidence intervals exclude 0, indicating that they significantly improve ratings above the intercept reference values. All three semantic class coefficients shown are also positive, increasing the rating level from that of the intercept (which is the nonalternating semantic class in the least frequent context of the double object construction, that of the noun recipients). Because its confidence interval includes 0, the coefficient of the other supposedly nonalternating transfer class (the verbs *carry, push, drag, lower*) does not differ significantly from the intercept – which is not surprising, since both nonalternating classes are rated lowest in double object constructions. The other two semantic classes contrast significantly with the reference class: their coefficients (11.84 and 16.86) indicate a positive increase in rating. However, the coefficient of pronominality of recipient (13.89) is even greater than one of these semantic class coefficients and near the center of the confidence interval of the other, meaning that in the pronoun recipient condition, the scores of the nonalternating classes do not differ significantly from those of alternating classes in the noun recipient condition. This provides confirmation of our observation in Figure 8 that the reportedly ungrammatical verb classes appear to be rated as highly in the more frequent context (the pronoun recipient condition) as the theoretically grammatical verb classes in the infrequent context (the noun recipient condition).

Thus generalizations observed in Figure 8 are significant after adjusting for the experimental subject, verb, item order, and the interaction of individual verbs with pronominality of recipient. This model explains a substantial amount of the variance in the ratings ($R^2 = 0.43$).

Discussion

As observed in Bresnan (2006), experimental work on grammaticality judgments has been advanced by improved techniques for eliciting judgments (Schütze 1996; Cowart 1997; Bard, Robertson & Sorace 1996), but the constructed sentences used in many controlled psycholinguistic experiments are often highly artificial, isolated from connected discourse and subject to assumptions about default referents (Roland & Jurafsky 2002). Contextual information about referents should not be ignored because it influences syntactic preferences in production and comprehension (Bock 1977, 1996; Bock, Loebell & Morey 1992; Bock & Warren 1985; Kelly, Bock & Keil 1986; Prat-Sala & Branigan 2000; Thompson 1990; Collins 1995; Ferreira 1996; Rosenbach 2003, 2005; Bresnan et al. in press).

Accordingly, the experimental items of the present study are built from samples of actual usage of syntactic structures in their natural contexts. Modern statistical models provide controls (Baayen 2004). This approach has two benefits in addition to the provision of essential contextual information. In the first experiment a statistical model of the usage data from the corpus study (Bresnan et al. in press) is used to measure subjects' predictive capacities. In the second experiment subjects' judgments are used to test and validate usage data drawn from the internet. In this way convergent corpus and experimental methods are brought to bear on ecologically natural linguistic materials.

What can be learned from studying this natural usage data? From Experiment 2 we see that linguistic manipulations that raise or lower probabilities influence grammaticality judgments, which have traditionally been the primary and privileged data for categorical grammatical models. The experiment points to ways of establishing sounder empirical foundations for syntactic and semantic theory and suggests why the older ways of doing syntax – by generalizing from linguistic intuitions about decontextualized constructions and ignoring research on actual usage, especially quantitative corpus work – produce unreliable and inconsistent findings.

From Experiment 1 we see that language users' implicit knowledge of their language is more powerful than has been recognized under the idealizations of categorical models of grammaticality: language users can in effect make accurate probabilistic predictions of the syntactic choices of others.[3] The present study is the first to our knowledge to measure the predictive capacities of language users in a syntactic domain by means of a sophisticated statistical model of usage data. This approach opens up a variety of

questions for further research, with potential applications in many areas of linguistics and the cognitive sciences more generally.

Acknowledgements

Thanks to Daniel Casasanto, Jeff Elman, Marilyn Ford, Florian Jaeger, Steve Pinker, Anette Rosenbach, and Shravan Vasishth for helpful discussion and comments. None of them can be blamed for the use I have made of their advice. Thanks also to Nick Romero for creative internet sampling and research assistance. Graphics and models were made using R (R Development Core Team 2006, Bates & Sarkar 2006).

Notes

1. The use of verb senses follows Bresnan et al. (in press). Up to five possible senses of any verb were distinguished based on broad semantic classes of their uses in context. For example, the 'transfer' sense of *give* in *give an armband* is distinguished from the 'communicative' sense of *give* in *give your name*.
2. In a more extensive study, both the discourse and the syntactic type could be separately manipulated.
3. A subsequent still unpublished study by the author shows that similar results are obtained when subjects are simply asked to guess which alternative the original dialogue participant used and to give a numerical estimate of the likelihood of their guess being correct.

References

Arnold, Jennifer, Thomas Wasow, Anthony Losongco and Ryan Ginstrom
2000 Heaviness vs. newness: The effects of complexity and information structure on constituent ordering. *Language* 76: 28–55.

Baayen, R. Harald
2004 Statistics in psycholinguistics: A critique of some current gold standards. *Mental Lexicon Working Papers, Edmonton* 1: 1–45.

Baayen, R. Harald, Richard Piepenbrock & Leon Gulikers
1995 The CELEX lexical database (Release 2) CD-ROM. Linguistic Data Consortium, University of Pennsylvania (distributor).

Bard, Ellen G., Dan Robertson & Antonella Sorace
1996 Magnitude estimation of linguistic acceptability. *Language* 72: 32–68.
Bates, Douglas & Deepayan Sarkar
2006 *lme4: Linear mixed-effects models using S4 classes*. R package version 0.995-2. http://www.R-project.org.
Bock, J. Kathryn
1977 Accessibility theory: An overview. *Journal of Verbal Learning and Verbal Behavior* 16: 723–734.
1986 Syntactic persistence in language production. *Cognitive Psychology* 18: 355–387.
1996 Language production: Methods and methodologies. *Psychonomic Bulletin & Review* 3: 395–421.
Bock, J. Kathryn & David Irwin
1977 Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior* 19: 467–484.
Bock, J. Kathryn, Helga Loebell & Randal Morey
1992 From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review* 99: 150–171.
Bock, J. Kathryn & Richard K. Warren
1985 Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21: 47–67.
Bresnan, Joan
2006 A few lessons from typology. Prepared for the LSA Workshop Typology in American Linguistics: An Appraisal of the Field. LSA 79th Annual Meeting, Oakland, January 9, 2005. Stanford University: http://www.stanford.edu/□bresnan.
Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen
in press Predicting the dative alternation. In *Cognitive Foundations of Interpretation*, G. Boume, I. Kraemer & J. Zwarts (eds.). Amsterdam: Royal Netherlands Academy of Science.
Bresnan, Joan & Jennifer Hay
forthc. Gradient grammar: An effect of animacy on the syntax of *give* in varieties of English. *Lingua*. Special issue on animacy invited submission).
Bresnan, Joan & Tatiana Nikitina
2003 On the gradience of the dative alternation. Stanford University: http://www-lfg.stanford.edu/bresnan/download.html.
Chatterjee, Samprit, Ali S. Hadi & Bertram Price
2000 *Regression Analysis by Example*. 3rd edition. New York: Wiley.
Clark, Eve V.
1987 The principle of contrast: A constraint on language acquisition. In *Mechanisms of language acquisition*, B. MacWhinney (ed.), 1–33. Mahwah, NJ: Erlbaum.

Cleveland, William S.
1979 Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74: 829–836.
Collins, Peter
1995 The indirect object construction in English: An informational approach. *Linguistics* 33: 35–49.
Cowart, Wayne
1997 *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousand Oaks, CA: Sage.
Fellbaum, Christiane
2005 Examining the constraints on the benefactive alternation by using the World Wide Web as a corpus. In *Evidence in Linguistics: Empirical, Theoretical, and Computational Perspectives*, M. Reis & S. Kepser, (eds.), 207–236. Berlin/New York: Mouton de Gruyter.
Ferreira, Victor S.
1996 Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language* 35: 724–755.
Godfrey, John J., Edward C. Holliman & Jane McDaniel
1992 Telephone speech corpus for research and development. In *Proceedings of ICASSP-92*, pp. 517–520.
Green, Georgia
1971 Some implications of an interaction among constraints. *CLS* 7: 85–100.
Gries, Stefan Th.
2003 Towards a corpus-based identification of prototypical instances of constructions. *Annual Review of Cognitive Linguistics* 1: 1–27.
2005 Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34: 365–399.
Hawkins, John A.
1994 *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
Hay, Jennifer & Joan Bresnan
to appear Spoken syntax: The phonetics of *giving a hand* in New Zealand English. *The Linguistic Review: Special Issue on Exemplar-Based Models in Linguistics* 23.
Kelly, Michael H., J. Kathryn Bock & Frank C. Keil
1986 Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language* 25: 59–74.
Lapata, Maria
1999 Acquiring lexical generalizations from corpora: A case study for diathesis alternations. In *Proceedings of the 37th Meeting of the North American Chapter of the Association for Computational Linguistics*, pp. 397–404. College Park, Maryland.

Levin, Beth
1993 *English verb classes and alternations: A preliminary investigation*. Chicago/London University of Chicago Press.
Pickering, Martin J., Holly P. Branigan & Janet F. McLean
2002 Constituent structure is formulated in one stage. *Journal of Memory and Language* 46(3): 586–605.
Pinheiro, José C. & Douglas M. Bates
2000 *Mixed-Effects Models in S and S-PLUS*. New York: Springer.
Pinker, Steven
1989 *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.
Prat-Sala, Mercè & Holly P. Branigan
2000 Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language* 42: 168–182.
R Development Core Team
2006 *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.
Rappaport Hovav, Malka & Beth Levin
2005 All dative verbs are not created equal. Unpublished ms., Hebrew University of Jerusalem and Stanford University, Jerusalem, Israel and Stanford, California.
Rosenbach, Anette
2003 Aspects of iconicity and economy in the choice between the *s*-genitive and the *of*-genitive in English. In *Determinants of Grammatical Variation in English (Topics in English Linguistics/[TiEL]*, G. Rohdenburg & B. Mondorf (eds.), 379–411. Berlin/New York: Mouton de Gruyter.
2005 Animacy versus weight as determinants of grammatical variation in English. *Language* 81: 613–644.
Schütze, Carson T.
1996 *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistics Methodology*. Chicago: University of Chicago Press.
Snyder, Kieran M.
2003 *The relationship between form and function in ditransitive constructions*. Ph.D. thesis, University of Pennsylvania.
Szmrecsányi, Benedikt
2005 Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistics Theory* 1: 113– 149.

Thompson, S. A.
1990 Information flow and dative shift in English discourse. In *Development and Diversity: Language Variation across Time and Space*, J. A. Edmondson, C. Feagin & P. Mühlhausler (eds.), 239–253. Summer Institue of Linguistics and University of Texas at Arlington, Dallas.
Wasow, Thomas
2002 *Postverbal Behavior*. CSLI, Stanford.