Is syntactic knowledge probabilistic?
Experiments described in:


Rosenbach (2003) reports a forced choice study which controls for the overlapping factors (animacy, topicality, prototypicality of the possession relation) that affect genitive choice:

Items and conditions:

[+animate +topical +proto]: the boy’s eyes ∼ the eyes of the boy
[+animate, +topical, −proto]: the mother’s future ∼ the future of the mother
[+animate, −topical, +proto]: a girl’s face ∼ the face of a girl
[+animate, −topical, −proto]: a woman’s shadow ∼ the shadow of a woman
[−animate, +topical, +proto]: the chair’s frame ∼ the frame of the chair
[−animate, +topical, −proto]: the bag’s contents ∼ the contents of the bag
[−animate, −topical, +proto]: a lorry’s wheels ∼ the wheels of a lorry
[−animate, −topical, −proto]: a car’s fumes ∼ the fumes of a car
- Operationalizes animacy as personal, common nouns vs. concrete common nouns (excluding geographical and temporal)

- Operationalizes topicality as second-mention, definite expression vs. first-mention, indefinite expression

- Operationalizes possessive relations as

  for humans: body parts, kin terms, and permanent legal ownership vs. states and abstract ‘possessions’

  for inanimates: part/whole relations vs. non-part/whole relations
A sample question from her questionnaire:

A helicopter waited on the nearby grass like a sleeping insect, its pilot standing outside with Marino. Whit, a perfect specimen of male fitness in a black flight suit, opened [the helicopter’s doors/ the doors of the helicopter] to help us board.

(based on Patricia Cornwell, The Body Farm, 52)
’s and *of* genitives in English (Rosenbach 2002)
Other findings:

the ’s-genitive is spreading across time (older to younger speakers) and space (younger American to younger British speakers)
Note on design and analysis:

– univariable analysis (= ‘basic statistical tests’, such as Chisquare)
– controls (e.g. holds length of possessor and possessum constant; excludes proper nouns)
– stratificational analysis (e.g. age, pp. 396–7)
Compare a subsequent corpus study:

Bresnan (2007):

Hypothesis: If the dative corpus model sufficiently characterizes language users’ implicit linguistic knowledge of usage probabilities, then where the model predicts higher- or lower-probability outcomes, we would expect experiment participants to do so as well in behavioral tasks.
• An indirect task: Rate the naturalness of the alternatives according to your own judgments on a numerical scale of 1 to 100.

• A direct task: Guess the choices made by the original dialogue speakers and rate the likelihood of your guess being correct on a numerical scale of 1 to 100.
Bresnan 2007

Experiment:

Thirty instances of dative constructions were randomly drawn from the centers of five probability bins of the dative corpus model distribution. (Potentially ambiguous items were replaced.)
Sampled Constructions for Experiment 1

Corpus Model Probabilities

- vlow
- low
- med
- hi
- vhi

Sampled Constructions for Experiment 1
The contexts of the sampled instances were retrieved from the full Switchboard corpus transcriptions and edited for readability by removing disfluencies and backchannelings.

- The probability model was not conditioned on speech features (disfluency, prosody, etc)
- The experimental task required reading and not audition.
• An alternative to each target construction was constructed,

• the order of passages was randomized,

• and the order of target constructions alternated.

• A questionnaire was created containing the thirty passages.
Speaker A:
I moved to Arkansas and Texas after living in Ohio and the schools down here rate, you know, bottom ten percent across the country and having been through grade school up there and coming down here to high school I can understand why. Because they’re so far behind and so poorly staffed, half the time the teachers don’t know what’s going on.

Speaker B:
Well, that’s really too bad because

(1) it’s giving some people unfair advantage.
(2) it’s giving unfair advantage to some people.
19 participants from Stanford summer term undergraduates were recruited and paid.\(^\text{a}\)

The participants were instructed to rate the relative naturalness of the alternatives in the given context passage, according to their own intuitions, on a scale of 0 to 100; the ratings of the alternatives must sum to 100.

\(^{a}\)The results from participants who had taken a syntax course were excluded, as well as bilinguals and non-native speakers of English.
Finding: Both as a group and individually, participants’ numerical ratings of the alternative dative continuations showed a direct linear relation to the corpus log odds of those constructions.
Analysis using multilevel multivariable regression showed:

the corpus model probabilities are significant predictors of the ratings, after controlling for random effects of subject and verb as well as item order, order of constructions, and lemma frequency.
Bresnan (2007) also compared each subject’s ratings with the actual choices by the speakers in the original conversations. Baseline = 0.57.

Proportions of Participants’ 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
Participants naturalness ratings are reliably associated with the syntactic alternatives used by the original speakers:

(Wilcoxon signed rank test with continuity correction, $n = 19, V = 190, p < 0.001$)
In a follow-up experiment different participants were asked to guess which choice the original speaker made, and to rate the likelihood that their guess was correct. These likelihood ratings were highly significant—

*they could make reliable guesses about which alternative the original dialogue participant chose*

(Wilcoxon signed rank test with continuity correction, $n = 20, V = 210, p < 0.0001$)
Related work:


