9-1-2012

Effects on the Particle Verb Alternation across English Dialects

Bill Haddican
Queens College-CUNY, bill.haddican@googlemail.com

Daniel Ezra Johnson
Lancaster University, danielezrajohnson@gmail.com

Effects on the Particle Verb Alternation across English Dialects

Bill Haddican and Daniel Ezra Johnson*

1 Introduction

This paper focuses on regional and grammatical effects on the English particle verb alternation. We illustrate this variation in (1), which shows that, with a class of verb+particle combinations, a direct object may appear either to the right or the left of the particle. We refer to these word orders as the *continuous* and *discontinuous* orders respectively.

(1) a. She cut open the melon.   (continuous order)
    b. She cut the melon open.   (discontinuous order)

While a considerable body of literature has focused on different syntactic and processing constraints on the variation in (1) (Dehé 2002, den Dikken 1995, Svenonius 1996, Toivonen 2001), relatively little work has discussed regional effects. In particular, this paper reports on a controlled judgment experiment and a Twitter corpus study designed to address Hughes et al.'s (2005) claim, based on non-controlled evidence, that the continuous order is favored in Scotland while the discontinuous order is favored in Southern England. We also examine the possibility of regional effects on particle placement variation in North American dialects. We test for these possible effects with a controlled judgment experiment with 297 native speakers from the British Isles and North America, and a Twitter corpus from tweets in the UK and US. The results from both the acceptability judgment study and the Twitter corpus reveal no support for a North-South difference across UK dialects, but instead show a trans-Atlantic difference: respondents from the UK and Ireland favored discontinuous orders while US and Canadian participants favored continuous orders. Based on a preliminary analysis of historical corpus data, we speculate that this difference reflects change toward an innovative discontinuous order that has proceeded more quickly in Old World dialects than in North America.

Our discussion is organized as follows. Section 2 of this paper reviews previous literature on social and linguistic effects on particle placement. Section 3 describes a judgment experiment testing regional and focus effects on particle placement. Section 4 of this paper reports and discusses results from a Twitter corpus providing additional support for the regional analysis in the experimental data. Section 5 summarizes the discussion.

2 Social and Linguistic Effects on the Particle Verb Alternation

Much of the formal and sentence processing literature on English particle verbs has focused on two kinds of constraints on particle placement. One set of studies has discussed the length, or prosodic weight of the object as a processing or a phonological phrasing constraint on particle placement. Kroch and Small (1978), Lohse et al (2004), and Gries (2001) all report evidence from corpus studies showing that "heavy" objects such as those in (2) tend to favor the continuous order and disfavor the discontinuous order.

(2) a. She turned off the fan that I brought in.   (continuous order)
    b. ?She turned the fan that I brought in off.   (discontinuous order)

Lighter objects, on the other hand favor the discontinuous order. Speakers generally find the discontinuous order obligatory for unstressed, weak pronouns as in (3).

(3) a. *She turned off it.   (continuous order)
     b. She turned it off.   (discontinuous order)

Lohse et al. (2004) explain the object length effect in terms of a more general processing constraint, namely that processing is facilitated by a short distance between members of a syntactic dependency. Lohse et al. take the relation between the verb and the particle to be a dependency governed by this principle. In the case of discontinuous orders but not continuous orders, heavy objects as in (2) incur a heavy processing cost because they create a large gap between the two elements in the particle verb dependency. Continuous orders are therefore preferred in proportion to increasing object length.

A second set of studies has focused instead on information structural constraints on particle placement. Bolinger (1971), Svenonius (1996), Kayne (1998) and Dehé (2002) note that given objects (topics) favor the discontinuous order while focused objects favor the continuous order. Svenonius (1996) notes that, as an answer to the object wh-question in (4), the continuous order is much more natural than the discontinuous order for many people.

(4) Q: Who will you pick up?
     A: I'll pick (?the girls) up (the girls).   (Svenonius 1996)

In contrast, as an answer to the question in (5), where the object is previously introduced, Svenonius notes that many speakers prefer the discontinuous order. Svenonius reports that this effect is mild for many speakers and that other speakers report no such effect.

(5) Q: How are Turid and Ingrid going to get here?
     A: I'll pick (the girls) up (?the girls).   (Svenonius 1996)

Svenonius explains the contrast between (4) and (5) in terms of syntax-prosody mapping: new information foci are intonationally prominent and prefer to be sentence-final. Focused constituents in the discontinuous order are dispreferred since they are separated from the right edge of the sentence by the particle.

The most detailed treatment of focus effects on particle placement in English is by Dehé (2002), who also explains the facts in cases like (4) and (5) in terms of syntax-prosody mapping constraints. Dehé proposes that the continuous order is the "neutral" order in that it is (i) derivationally prior, and (ii) corresponds to a discourse neutral, "out of the blue" interpretation. Dehé's generalization is that objects must only appear in the discontinuous order when they are defocused and within a syntactic domain bearing focus; the continuous order will be preferred otherwise. Dehé illustrates this with the question-answer pair in (7), in which the discontinuous order is preferred in the answer according to Dehé. In the answer in (7), the object is presupposed by the question and the verb+particle has new information focus. Dehé, proposes that the preference for the discontinuous order in this sentence is a consequence of the fact that the presupposed object sits inside a Focus bearing constituent, namely YP.[1]

(6) Q: What did Durban do to the camera?
     A: Durban [$_{YP[+Foc]}$turned [$_{XP[-Foc]}$the camera] OFF].   (Adapted from Dehé 2002:132)

---

[1]Dehé formalizes this constraint as in (i).
(i)   Condition on Focus Domains   (Dehé 2002:248)
      "Within a focus domain, a [+F] focus feature must be bound by some kind of verbal affix if there is
      a mismatch with regard to focus features."
Dehé proposes that the Verb raises out of VP. In discontinuous orders, where the particle stays inside VP, it may bind the Focus-feature on VP. In continuous orders, the particle will raise out of VP along with the verb and the Focus-feature on VP will be unbound, in violation of (i). See Svenonius (2004) for a critique of this proposal.

Dehé reports two kinds of experimental evidence in support of this generalization. In one experiment, 28 native-speakers of UK English were presented with vertically arranged lists of a verb, particle, and object, as in (7), on a computer screen. Each of the six unique possible orders for these elements were counterbalanced across subjects and items, such that each order and item was seen an equal number of times by the participants. For each such list presented, subjects were asked to assemble a sentence with the subject *she*.

(7) the tray
     in
     carried   (Dehé 2002:95)

Dehé found that continuous orders were generally preferred, which she takes as evidence that the continuous order is "neutral," that is, the order preferred in out-of-the-blue contexts. Dehé does not directly address the extent to which these results are likely to reflect the frequency of continuous vs. discontinuous orders in reading and/or written production. In addition, as Svenonius (2005) notes, because this experiment does not directly bias focus interpretation, it does not illuminate possible sensitivity of word order variation to information structure.

Dehé's second experiment examines the effect of word order and focus on pitch. Ten native speakers of UK English were recorded reading short passages containing a sentence with a particle verb construction. Subjects were asked to familiarize themselves with the passage before each recording. The passages were constructed to bias an interpretation of the object as given or new information. New-information-biased objects appeared in the continuous order and given-biased objects appeared in the discontinuous order, as illustrated in (8) and (9).

(8) Continuous order
     It's late and I want to go to bed. Peter, I would like you to **turn down the radio**. The music
     is too loud, I won't be able to go to sleep.   (Dehé 2002: 179)
(9) Discontinuous order
     "Peter, do you know where that noise is coming from?"
     "Yes, Ann, I do. It's the radio of our next-door neighbor, a student. She likes her music
     loud."
     "Fine, but I can't stand it. I'll go and ask her to **turn the radio down**."   (Dehé 2002: 179)

Dehé reports that for continuous orders, the pitch (F0) range for the particle in these sentences is smaller than that of the object and in the discontinuous orders vice-versa. Dehé takes this result as evidence that word order has an effect on the intonation of particle verb constructions. As Svenonius (2005) notes, however, because word order is confounded with focus bias, these results do not help illuminate the effect of focus on word order in these constructions.

A first goal of the judgment experiments described below is to examine possible focus effects more directly by biasing different kinds of focus interpretation independently of word order. We also include in our design object weight, in an effort to test the possible interaction of object weight and discourse status. We describe these experiments in sections 3 and 4 below.

A second goal of this paper is to test Hughes et al's (2005) claim of a dialectal difference in particle placement preference. Specifically Hughes et al (2005:23) propose that Scottish speakers tend toward continuous orders (1a), while speakers from the south of England tend toward discontinuous forms (1b). The authors report no supporting evidence for this claim, however, and as far as we are aware, no other published literature has reported evidence to this effect in contemporary UK dialects. Based on limited historical corpus evidence, however, Elenbaas (2007: 273–279) speculates that in the early Modern English period, continuous orders were favored in areas most exposed to Scandinavian varieties, that is the Danelaw in Northern and Eastern parts of England, while discontinuous orders were favored elsewhere.

While no literature to date has discussed geographic correlates of this variation in the US, Hughes et al.'s claim of a Scottish-Southern English difference, suggests the possibility of founder effects in North American dialects. That is, if Hughes et al.'s regional difference indeed exists and dates back at least to the time of North American settlement, then we might expect preferences for

continuous vs. discontinuous orders to appear in areas settled by Scots/Scotch-Irish migrants and Southern English migrants respectively. In particular, we might expect much of New England, which was mainly settled by speakers of Southern English dialects to favor discontinuous orders and that Appalachian dialects, which were founded largely by Scotch-Irish settlers, to favor continuous orders (See Krapp 1925, Kurath 1949, and Montgomery 2006 for discussion of early North American migration and settlement patterns and their possible consequences for the emergence of North American regional dialects). We assess evidence in favor of possible regional effects in Sections 3 and 4 below.

3 An Acceptability Judgment Study

3.1 Data and Method

The first data set we report on comes from an online judgment experiment conducted in the spring and summer of 2011.

Subjects. Subjects for the experiment were 297 self-described native speakers of English recruited online through personal contacts of the authors. 145 of these were from the UK or Ireland and 152 were from the US and Canada. Almost all had BA/BS-level degrees or higher. Subjects ranged in age from 18 to 84 (mean = 30). 37% were men.

Materials. The experiment crossed three within-subjects factors, each with two levels: particle-object order, object length, and focus status of the object. The particle-object-order factor had the levels continuous and discontinuous as illustrated in (1) above. Object length was operationalized as an additional binary factor: "short" DPs were all 3-syllable constituents with the definite article and a two-syllable noun, for example *the melon*; "long" DPs were all 7-syllable DPs with a definite article, two two-syllable adjectives and a noun, for example *the heavy juicy melon*.

We followed Dehé (2002) in operationalizing focus as a binary factor by biasing new vs. old information interpretations of the object. We did this using a cataphoric pronoun in a *because*-clause or temporal adverbial clause, bound by either the object of the particle verb in the main clause or the subject of the main clause. In the former case, the object was "given information" in that it was introduced in the preceding clause. In the latter case, the subject was given and the VP was biased as new information.[2] Crossing the above three factors yields (2 x 2 x 2 =) 8 conditions, which we illustrate in (10) and (11)

(10) It$_i$ was about to spoil, so Andrea cut (open) the (heavy juicy) melon$_i$ (open).
       (given object)
(11) Her$_i$ kids wanted a snack, so Andrea$_i$ cut (open) the (heavy juicy) melon (open).
       (new object)

Four lexicalizations were created for each of these (2 x 2 x 2 =) 8 conditions. The particle verbs chosen were all non-aspectual and compositional as described in Lohse et al. (2004). Lexicalizations were blocked by Latin square, such that each block contained a different lexicalization for each of these eight conditions. These blocks were then grouped into 32 lists, with each list containing 4 blocks; each subject therefore saw each condition four times. The 32 experimental sentences in each list were pseudo-randomized within blocks with 32 filler sentences, half grammatical and half ungrammatical. Subjects were semi-randomly assigned to lists by the experimental software.

Procedure. Subjects judged each of these 64 sentences in a self-paced online judgment experiment using Ibex Farm (Drummond 2011). The experiment was anonymous and subjects

---

[2]An alternative way of biasing focus on the object is with wh-questions as the context, e.g. *What did* SUBJECT *do*? (to bias VP-focus readings) and *What happened to* OBJECT? (to bias a given-object reading). A disadvantage of this approach is that it requires repeating the subject and object in the question and answers, which speakers typically find pragmatically odd, particularly for heavy DPs. For this reason we chose the cataphor binding approach explained above.

were neither paid nor did they receive academic credit for participating. Subjects rated each sentence on an 11-point scale by clicking an icon for a value ranging from 0 to 10 in a horizontal array, with endpoints labeled "Bad" and "Good" respectively.

3.2 Results and Discussion

The data for each subject were first normalized by converting to z-scores, subtracting the mean and dividing by the standard deviation of the ratings of the 32 filler sentences. Since some of the fillers were downright ungrammatical, the experimental sentences with particle verbs tended to have positive z-scores, between +0.5 and +1.0 units on average.

Using the *lme4* package in R, we then fit a series of linear mixed effects models, with fixed effects for subject region/country and the above within-subjects factors, and random intercepts and slopes by subject and by item. For example, to test whether subject region significantly affected preference for the continuous or discontinuous order, two models were fit. Both had a random-effect structure consisting of region*order | subject and region*order | item. The more complex model had a fixed-effect term for the region*order interaction while the simpler model had only main effect terms for region and order. A likelihood-ratio test was used to compare the two models and arrive at a *p*-value representing the significance of the region*order interaction.

The results support three main findings. First, the regional analysis revealed no support for any regional distinctions within North America (six regions, *p* = .65) or within the British Isles (twelve regions, *p* = .98). That is to say, there was no significant region*order interaction on either side of the Atlantic. The analysis did, however, reveal a significant trans-Atlantic difference. When subjects were recoded into a factor with three levels corresponding to the subjects' home country–US (N=113) vs. Canada (N=32) vs. UK/Ireland (N=152)–the analysis revealed a significant country*order interaction (*p* = .001), with US subjects preferring continuous orders by .08 units, UK/Ireland subjects preferring discontinuous orders by .03 units and Canadian subjects showing no preference in either direction (Figure 1). We return to these results shortly.
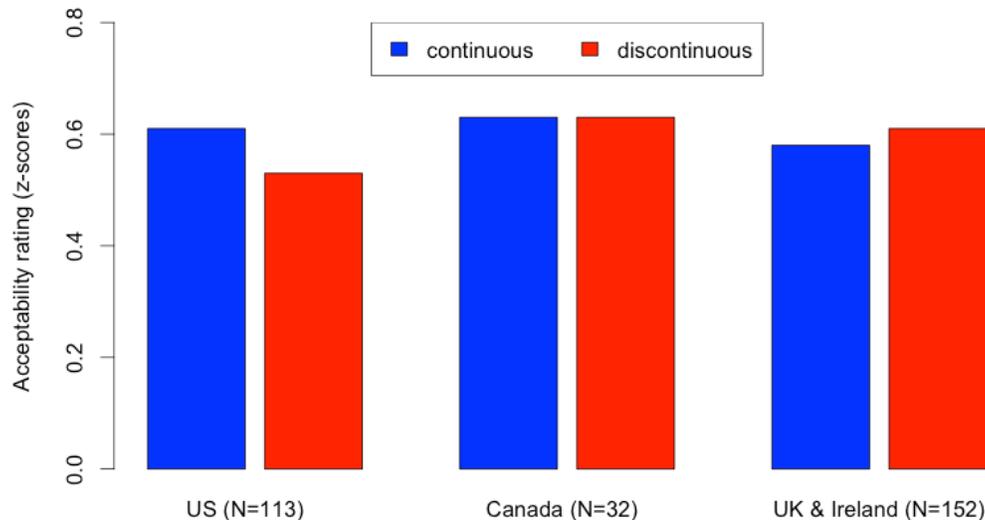

Figure 1: Mean normalized ratings for continuous and discontinuous orders by country.

The second finding is that there was no significant effect for the focus*order interaction. Figure 2 shows that continuous orders were in fact favored somewhat by the new-object condition, in keeping with Dehé's (2002) and Svenonius' (1996) discussion, but the difference of .04 units between conditions was not significant (*p* = .12). We speculate that the cataphoric pronoun technique used for biasing given vs. new information interpretations of the object may not have been successful with this set of subjects.
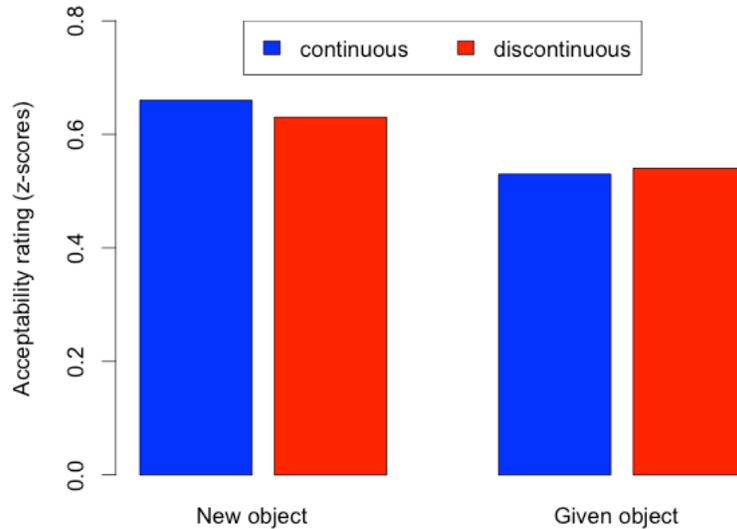
Figure 2: Mean normalized ratings for continuous and discontinuous orders by object type.

Figure 3 illustrates the third main finding: a significant weight*order interaction ($p = .00003$). In sentences with light objects, discontinuous orders are preferred by .05 units, and in sentences with heavy objects, continuous orders are preferred by .07 units. That is, there is a difference of .12 units between the two conditions.
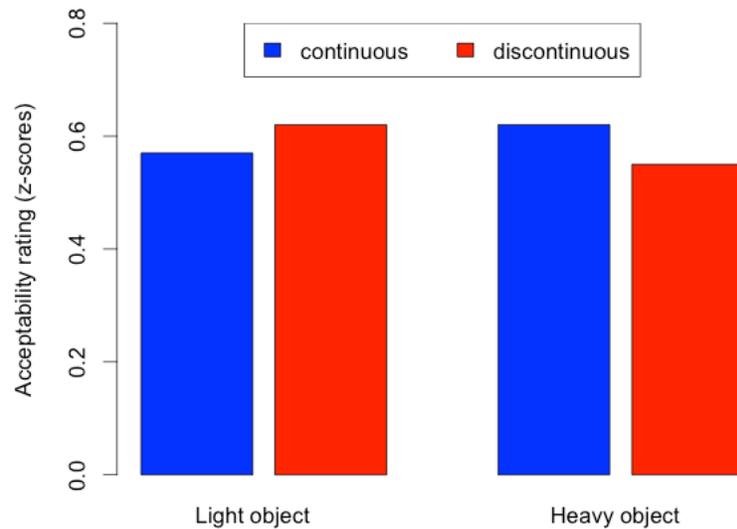

Figure 3: Mean normalized ratings for continuous and discontinuous orders by object weight.

This third result aligns with much previous corpus based work on placement, which has shown that heavy objects tend to be placed after the particle, while lighter objects and pronouns tend to precede the particle (Kroch and Small 1978, Lohse et al. 2004, Gries 2001). These processing and phonological accounts of the "weight effect" correctly predict that a heavy object is judged worse than a light object in the discontinuous order, where the object is interposed between the verb and the particle. In our study, this difference was .07 units ($p = .0009$).

In previous studies of this type subjects have been asked to choose between two syntactic alternants or distribute a fixed number of rating points between them (Bresnan 2007, Melnick et al. 2011). Such designs would make it impossible to independently assess the factors affecting the acceptability of the continuous and discontinuous orders. The present design, in which each order

is evaluated independently, reveals an effect not predicted in the literature, namely that heavy objects are actually judged *better* than light objects in the continuous order, by .05 units (p = .03). This effect is unlikely to derive from processing constraints, but could be explained if subjects implicitly evaluate sentences exhibiting one structure (e.g., continuous order) with respect to the equivalent sentences with the other structure (e.g., discontinuous order). If this is true, the well-motivated weight effect that disfavors heavy objects in the discontinuous order would create a preference for heavy objects in the continuous order.

Finally, we note that the analysis revealed no significant higher-order interaction between country and focus (country*focus*order, p = .75) or country and weight (country*weight*order, p = .43), meaning there is no evidence for trans-Atlantic differences in these effects. Nor was there any significant interaction between focus and weight (focus*weight*order, p = .83)

4 A Twitter Corpus Study

4.1 Data and Method

To test for the possibility of similar regional effects in production, we examined variation between continuous and discontinuous orders in a bespoke Twitter corpus. The corpus consisted of tweets containing a variation on one of two base strings, *turn on the light* (continuous) and *turn the light on* (discontinuous). The volume of tweets was augmented by including examples with *turns* and *turned* as well as *turn, off* as well as *on,* and *lights* as well as *light.* Before analysis, the data were cleaned by hand of song lyrics, quotations, memes, and other examples that did not reflect the production of the user.

The tweets were gathered between February and May of 2011 from Twitter API. The corpus was geocoded to areas within a 150-mile radius of four population centers in the UK and US: Oxford, England and Glasgow, Scotland; Pittsburgh, Pennsylvania and Concord, New Hampshire. The search on the area centered around Glasgow gathered 236 tweets with the relevant strings from Scotland and Northern England. The Oxford-centered search gathered 1472 tweets from an area spanning most of the rest of England (it did not overlap with the Glasgow-centered search). The Concord, New Hampshire-centered search, which yielded 296 tokens, encompassed most of New England, in an effort to target an area founded by Southern English settlers. Finally, the Pittsburgh-centered search gathered 343 tweets and targeted an area of Appalachia and western Pennsylvania, whose founding settlers were predominantly of Scottish or Scotch-Irish origin (Montgomery 2006). We illustrate the different catchment areas for the US and UK dialect areas in Figures 4 and 5, respectively.

Figure 4: Twitter corpus catchment for US dialects.

Figure 5: Twitter corpus catchment for UK dialects.

4.2 Results

The results, again, show no evidence of regional effects within the UK (contra Hughes et al. 2005; Fisher's Exact Test p = .61), nor within the US (p = .87). But the national results do align with the acceptability judgment results reported above, in that the US Twitter users tend slightly toward the continuous order (53% continuous), while UK users tend more decidedly toward the discontinuous order (64% discontinuous). For this trans-Atlantic difference, p = 6×10^{-13}. We illustrate these effects in Figure 6.
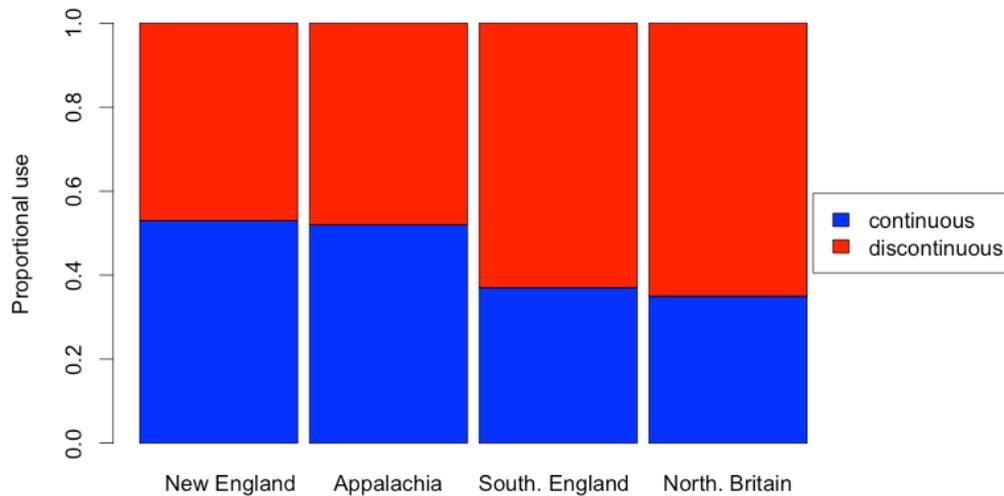

Figure 6: Proportional use of continuous and discontinuous orders by region.

A question that arises in light of the acceptability judgment data and the Twitter corpus data is how to explain the trans-Atlantic difference. We see three main possible explanations: a first possibility is that UK dialects have been innovative in moving toward discontinuous orders after

the period of North American colonization; a second possibility is that US speakers have been innovative in tending toward the continuous order; a third possibility is that both dialects are changing but doing so at different rates; for example, both UK and US dialects could be moving toward the discontinuous order, but UK dialects have moved further and/or faster.

We have not yet fully explored the kind of data most helpful in adjudicating among these possibilities, namely parsed historical corpus data from Middle English and early Modern English. Some insight into these issues, however, comes from a sample of common particle verb sequences from the Corpus of Historical American English (COHA). This data shows historical change toward the discontinuous order in American dialects, supporting the third possibility given above.

The COHA is a 400-million-word corpus of American texts balanced by genre and style from 1810 to the present. We extracted 685 tokens of particle verbs consisting of the strings shown in the upper left hand corner of Figure 7 along with the number of tokens/string. Figure 7 plots the proportion of discontinuous forms by decade. The results show change toward the discontinuous order during this period. Assuming a constant rate of change, a logistic regression (with a term controlling for the individual string) returns a slope of +0.01 log-odds per year in favor of the discontinuous order ($p = .0003$).
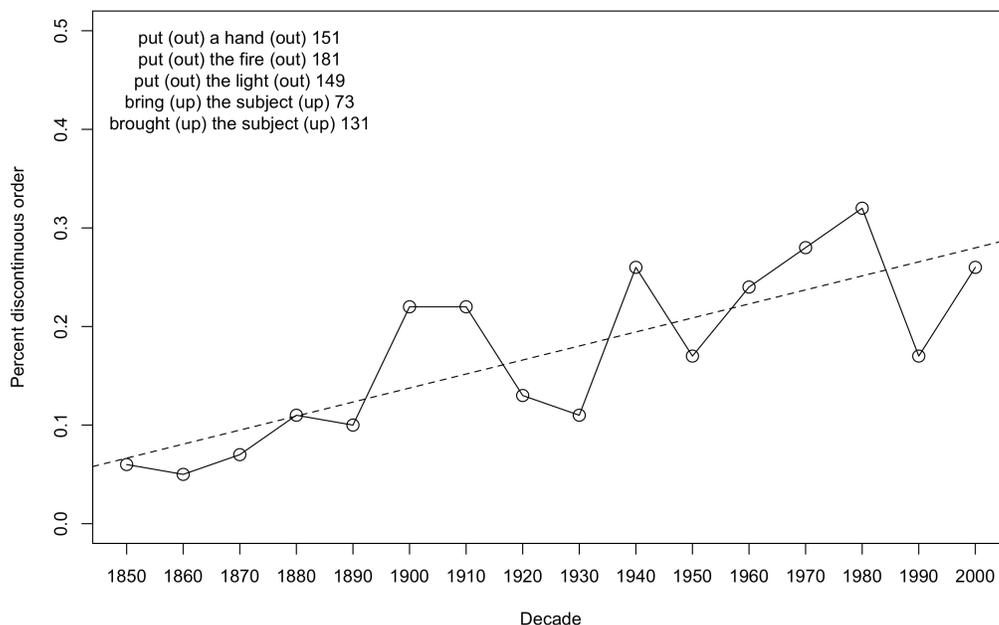

Figure 7: Discontinuous order in the Corpus of Historical American English 1850-2009.

The results therefore suggest evidence of change toward the discontinuous order over time in American English, a result in keeping with the third possibility suggested above.

5 Conclusion

This paper has focused on regional and other effects on the particle verb alternation in English. Our main finding is evidence of a trans-Atlantic difference where UK and Irish English speakers tend toward the discontinuous order in both production (in a Twitter corpus) and perception (in a judgment experiment) to a greater extent than U.S. and Canadian speakers. We also demonstrate the benefits of an experimental methodology whereby syntactic alternants are evaluated independently, rather than as a forced choice. We present results from a historical written corpus of American English suggesting change toward the discontinuous order. We speculate, based on these results and the above Twitter and judgment data, that the discontinuous order is innovative in both US and British Isles dialects but that this change has progressed more quickly in the British Isles. Further historical corpus work, using parsed corpora from UK English varieties, might usefully address this issue.

References

Bolinger, Dwight L. 1971. *The phrasal verb in English*. Cambridge: Harvard University Press.

Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In *Roots: Linguistics in Search of Its Evidential Base*, ed. S. Featherston and W. Sternefeld, 77–96. Berlin: Mouton.

Dehé, Nicole. 2002. *Particle Verbs in English: Syntax, Information Structure, and Intonation.* Amsterdam: John Benjamins.

Dikken, Marcel den. 1995. *Particles: On the syntax of verb–particle, triadic and causative constructions*. Oxford: Oxford University Press.

Drummond, Alex. 2011. Ibex Farm. spellout.net/ibexfarm.

Elenbaas, Marion. 2007. The synchronic and diachronic syntax of the English verb-particle combination. Doctoral dissertation, Radboud University Nijmegen.

Gries, Stefan T. 2001. A multifactorial analysis of syntactic variation: particle movement revisited. *Journal of Quantitative Linguistics* 8:33–50.

Hughes, Arthur, Peter Trudgill, and Dominic Watt. 2005. *English Accents and Dialects*. London: Hodder Arnold.

Kayne, Richard S. 1998. Overt versus Covert Movement. *Syntax* 1, 128–191.

Krapp, George P. 1925. *The English Language in America*. New York: Ungar.

Kroch, Anthony, and Cathy Small. 1978. Grammatical ideology and its effect on speech. In Linguistic variation: models and methods, ed. D. Sankoff, 45–55. New York: Academic Press.

Kurath, Hans. 1949. *A word geography of the Eastern United States*. Ann Arbor: University of Michigan Press.

Lohse, Barbara, John Hawkins, and Thomas Wasow. 2004. Processing domains in English verb-particle constructions. *Language* 80:238–261.

Melnick, Robin, T. Florian Jaeger, and Thomas Wasow. 2011. Speakers employ fine-grained probabilistic knowledge. Paper presented at LSA, Pittsburgh, PA.

Montgomery, Michael. 2006. How Scotch-Irish Is Your English? *Journal of East Tennessee History* 77:65–91.

Svenonius, Peter. 1996. The optionality of particle shift. *Working Papers in Scandinavian Syntax* 57:47–75.

Svenonius, Peter. 2005. Review of Dehé (2001, Benjamins) Particle Verbs in English: Syntax, Information Structure, and Intonation. *Linguistische Berichte.* 202:261–265.

Toivonen, Ida. 2001. The structure of non-projecting words. Doctoral dissertation, Stanford University.

Department of Linguistics and Communication Disorders
Queens College-CUNY
65-30 Kissena Blvd.
Flushing, NY 11367
United States of America
bill.haddican@googlemail.com

Department of Linguistics and English Language
Lancaster University
County South
Lancaster LA1 4YL
United Kingdom
danielezrajohnson@gmail.com