Canadian Journal of Linguistics/Revue canadienne de linguistique, 68(2): 329–343, 2023
doi: 10.1017/cnj.2023.13

En bref
Short

Social role effects on English particle verb variation fail to replicate

NAOMI LEE AND LAUREL MACKENZIE
New York University, New York, NY, USA
naomilee@nyu.edu; laurel.mackenzie@nyu.edu

Abstract

The English particle verb alternation has been argued to be sensitive to the social role occupied by speakers on radio broadcasts; Kroch and Small (1978) argue that radio show hosts and in-studio guests' greater sensitivity to prescriptive norms makes them more likely to use the joined variant of the alternation than listeners calling in to the show. This study analyzes 10,521 tokens of variable particle verbs from the RadioTalk Corpus (Beeferman et al. 2019) to try to replicate the effect of speaker role. Our analysis confirms that direct object length, register, a measure of frequency, semantic compositionality of the particle verb, and the particle's prosody all condition the alternation. However, the effect of social role does not replicate.

Keywords: particle verbs, syntactic variation, language ideology, corpus linguistics

Résumé

L'alternance des verbes à particule en anglais a été décrite comme sensible au rôle social occupé par les locuteurs dans les émissions de radio; Kroch et Small (1978) soutiennent que la sensibilité aux normes prescriptives des animateurs d'émissions de radio et des invités en studio les rend plus susceptibles d'utiliser la variante jointe de l'alternance que les auditeurs qui téléphonent de l'extérieur. Notre étude analyse 10 521 exemples de verbes à particules variables du RadioTalk Corpus (Beeferman et al. 2019) pour essayer de reproduire l'effet du rôle du locuteur. Notre analyse confirme que la longueur de l'objet direct, le registre, une mesure de la fréquence, la composition sémantique du verbe à particule et la prosodie de la particule conditionnent tous l'alternance. Cependant, l'effet du rôle social vu dans Kroch et Small n'a pas été confirmé.

Mots-clés: verbes à particule, variation syntaxique, idéologie du langage, linguistique de corpus

Thanks to Richard Kayne, Gary Thoms, Mary Robinson, audiences at NWAV 49 and NYU LingLunch, Heather Newell, Daniel Siddiqi, and two anonymous CJL reviewers. This research was supported by faculty research funds to the second author from New York University.

1. Introduction

English has a class of transitive verb-particle combinations that exhibit word order alternation. In particle verbs like *clean out*, the particle may remain JOINED with the verb; the particle verb's object follows, as *your gutters* follows *clean out* in (**1a**).[1] Alternatively, the particle verb may be SPLIT by its object; in (**1b**), the verb *clean* and the particle *out* are linearly separated from each other by *the gutters*.

(1) a. JOINED: Look, just **clean out** your gutters and increase the grading away from your house (8297d9d3)

b. SPLIT: A lot of people didn't get up to **clean** the gutters **out** before winter (5e572544)

The English particle verb alternation is one of a few well-studied examples of word order variation (Robinson and MacKenzie 2019). Because previous work has established which/how *internal* linguistic factors matter, this alternation is fertile ground for investigating whether *external* factors condition syntactic variation. In an influential early variationist study, Kroch and Small (1978) contend that they do. Specifically, they argue that radio show hosts and in-studio guests show greater sensitivity to prescriptive norms, and that this makes them more likely to use the joined variant of the alternation than listeners calling in to the show. Our study attempts to replicate this finding with a much larger data set and more modern statistical methods, but finds it does not hold up: social roles predicted to be associated with adherence to prescriptive language ideologies do not condition the rate of joined variant use for particle verbs in our data.

2. Background

Kroch and Small (1978) investigate the role of prescriptive norms in conditioning syntactic variation. They hypothesize that speakers who occupy different social roles will differ in their adherence to the grammatical ideology of the standard language, and that this can be observed in how they produce syntactic structures that vary in their "correspondence between propositional form and surface syntax" (Kroch and Small 1978: 48). One such variable structure is the English verb-particle construction, for which "the surface syntactic configuration V Prt NP would correspond to propositional form more directly than would V NP Prt because only the first reflects in its word order the semantic unity of the verb and particle (compare the split infinitive)" (Kroch and Small 1978: 49). In other words, the prediction is that speakers who are more sensitive to prescriptive norms will be more likely to use what we are calling the joined variant (Kroch and Small's "V Prt NP").

Kroch and Small test this with a small corpus study of talk radio broadcasts from Philadelphia. To operationalize the factor of adherence to standard language ideology, they divide speakers in their radio data into two sociolinguistic groups: those

[1]Examples from the RadioTalk Corpus (Beeferman et al. 2019) are labeled with "signatures" (in parentheses), 8-character identifiers of unique transcriptions of speech events.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

who have called in to a talk show ("callers") on one hand, and, on the other, the show hosts and their studio guests ("hosts/guests"). They further assume that "for reasons both of social status and role in the radio talk-show interaction, the host/guest group generally speaks a more standard English than the caller group" (Kroch and Small 1978: 48). They find that this is borne out quantitatively: the host/guest group is found to use the joined variant of the verb-particle alternation at a rate of 63% (N = 167), significantly different from the caller group's 47% rate for the same variant (N = 138) via a chi-squared test ($p < 0.01$).[2] Further analysis subsetting the data by linguistic factors (length of object, degree of semantic transparency of verb-particle construction, stress placement) and comparing caller and host/guest rates within each level of the linguistic factor under study generally confirms the basic pattern quantitatively (though in section 5.2 we cast some doubt on these findings). Finally, Kroch and Small report the results of a judgment task carried out by Temple University students (N = 32): 44% of students report that the sentence *John called up Mary* is "substantially" more correct than the alternant *John called Mary up*. This further supports the notion that language users are sensitive to prescriptive norms against the split variant of this alternation.

Kroch and Small's findings have been influential. They continue to be cited in discussions of the role of prescriptive norms and standard language ideology in conditioning language variation (Romaine 1981, Irvine 1985, Guy and Bayley 1995, Johnstone and Bean 1997, Cameron 2000, Díaz-Peralta and Almeida 2000, Bresnan and Ford 2010, D'Arcy and Tagliamonte 2010, Adli 2013, Wiechmann and Lohmann 2013, Bouchard 2018), in literature reviews of the English verb-particle alternation (Gorlach 2004, Bleaman 2020, Haddican, Johnson, Wallenberg, et al. 2020, Röthlisberger and Tagliamonte 2020), and in considerations of the question of whether syntactic variation more generally can be sensitive to social factors (Meyerhoff 2000, Röthlisberger and Tagliamonte 2020). On this latter point, there is a large literature proposing or asserting that syntactic variation is unlikely to show social conditioning (see Levon and Buchstaller 2015 for a recent review); Kroch and Small's host/guest vs. caller difference is sometimes held up as a counterexample to this (Cheshire 1987, Meyerhoff 1997, Meyerhoff 2000, Röthlisberger and Tagliamonte 2020).

The lasting influence of their findings notwithstanding, Kroch and Small were careful to point out the need to replicate their study with more data and with multivariate analysis (1978: 51). There has been considerable recent large-scale corpus work on the particle verb alternation which has demonstrated that it has robust social correlates, including register, regional variety, speaker age, and social class (Gries 2003; Haddican and Johnson 2012; Röthlisberger and Tagliamonte 2020). Here, we ask: Can Kroch and Small's host/guest vs. caller difference be replicated on a larger data set using more modern statistical methods?

[2]Kroch and Small also find a similar quantitative difference in the rates at which the two groups variably delete the *that* complementizer, with the host/guest group deleting significantly less than the caller group, again consistent with standard language ideology, which should favour using *that* to explicitly link the matrix verb to the complement clause.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

3. Methods

This study extracted, annotated, and analyzed 10,521 tokens of particle verb constructions.

3.1 The corpus

The data used in this study come from the RadioTalk Corpus (Beeferman et al. 2019). RadioTalk consists of 2.8 billion words of automatic speech recognition transcripts. The transcribed snippets were sampled from 284,000 hours of talk radio broadcasts that aired in the continental United States between October 2018 and March 2019 and were subsequently made available online. The majority of sampled stations were either news, talk, or public radio stations; as a result, the corpus includes conversational speech both from radio show hosts and from listeners who called in to their shows. Using unsupervised spectral clustering to judge "whether the underlying audio came from a telephone or studio audio equipment," a process that mostly relied on "the narrow frequency range of telephone audio," Beeferman et al. estimate that 10.0% of the corpus is telephone speech (Beeferman et al. 2021).

3.2 Data extraction and coding

We used Google BigQuery Standard SQL to extract instances of 249 transitive verb-particle combinations in either the joined or split orders. Because RadioTalk was automatically transcribed and not checked by human annotators, our query excluded any snippets for which Beeferman et al.'s speech recognizer had less than 98% confidence in its own transcription accuracy. Other incorrect transcriptions whose actual values could not be recovered, as judged during our own data annotation process, were also excluded.

Rate of use of the joined variant is known to increase with the length of the object, becoming nearly categorical with objects above six words in length (Grafmiller and Szmrecsanyi 2018: 389). Reflecting this, our query also excluded any matches where more than four words intervene between the verb and the particle; objects of greater than four words found with particle verbs in the joined order were excluded by hand. We opted for a maximum object length of four words because exploratory queries that expanded the window between verb and particle any further resulted in an excessively high proportion of false string matches relative to legitimate split particle verb tokens (e.g., *eating* ***clean*** *is eating less junk and working* ***out*** *is being less sedentary*). False string matches of this type were removed by hand.

Pronominal objects are nearly categorical in requiring the split order, in both our data (n = 5339/5350, 99.8%) and previous work (Grafmiller and Szmrecsanyi 2018: 389). When the particle in a particle verb is modified by elements like *right*, *straight*, or *back*, the split order is again highly preferred (n = 30/32, 93.8%, including only non-pronominal objects). Fronted objects, as in passives or relative clauses, are also non-alternating. We excluded all such non-alternating items by hand.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

3.3 Coding for conditioning variables

Following these exclusions, the dataset contained 10,521 tokens of particle verb constructions. Each token was annotated for our critical variable of speaker modality, and for five control variables that previous work or our own explanatory data analysis found to affect word order.

Speaker modality: The core finding of Kroch and Small (1978) that we hope to test is the effect of the social roles occupied by various talkers on radio show broadcasts. Radio show hosts and in-studio guests are hypothesized to adhere more closely to prescriptive language ideologies — use more joined variants — than listeners calling in to the show. This hypothesized difference could stem from differences in the groups' ranges of occupations and their attendant participation in the linguistic market (Sankoff and Laberge 1978), as well as from the asymmetric power dynamic of the host-caller interaction itself.

Audio from listeners calling in to radio shows via telephone is limited to narrowband frequencies (300–3400 Hz), whereas audio recorded in the studio (i.e., radio show hosts and in-person guests) is not subject to that limitation. Unsupervised spectral clustering software made use of this difference to automatically supply "a flag for whether a given utterance was recorded in a studio or came from a telephone call-in" (Beeferman et al. 2019: 2).

Length of the direct object: Increasing the length of the object has been found to increase the likelihood of using the joined verb-particle order, as mentioned in section 3.2. Longer, prosodically "heavier" objects, which may also be more complex (see Wasow 1997, Wasow and Arnold 2011 for discussion of endweight), tend to occur at the ends of clauses. We operationalize object length as number of orthographic words (following Gries 2003, Grafmiller and Szmrecsanyi 2018, see also Röthlisberger and Tagliamonte 2020), ranging from one to four.

NPR show as a proxy for written register: Some previous studies on British English have shown an effect of written vs. spoken language, where written particle verb constructions are more likely to be found in the joined order (Gries 2003, Cappelle 2005). As a proxy for written register, we search for the string "NPR" in the name of the show associated with each token. NPR shows like *Morning Edition* often include hosts reading prepared or scripted news briefs, which we take to be more like written language. We therefore expect particle verbs coded as coming from an NPR show to be more likely to be joined.

Meaning frequency: Particle-verb collocations are highly polysemous. Each token was coded for its specific meaning, as demonstrated in example (2) for meanings of *knock off*.

(2) Different meanings of *knock off*

a. literally strike, causing to fall: *an asteroid knocked a piece of Mars off* (c99995fc)

b. remove: *I was working a plan and seein' progress so we just started knocking off baby steps like no tomorrow* (34cccaf1)

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

c. defeat a competitor: *In the NHL yesterday the New Jersey devils knocked off the Boston Bruins five to one* (28ceee30)

d. stop some behaviour: *Knock that singing off!*

Since measures of particle verb frequency from other corpora are not specified for these meanings, we operationalized meaning frequency using our own data set. We counted all occurrences of each meaning, including in non-alternating contexts (e.g., *the piece of Mars that the asteroid knocked off*), and log10-transformed the resulting meaning counts to arrive at a measure of frequency.

Semantic compositionality: English particle verbs' meanings range from entirely predictable from their component pieces' literal meanings (e.g., *blow away the feather*) to not at all compositional (e.g., *carry out a task*). That degree of compositionality has been argued to robustly affect a given particle verb's rate of alternation, with less transparently predictable, more idiomatic particle verb meanings corresponding to higher rates of the joined order (Fraser 1976, Kroch and Small 1978, Chen 1986, Dehé 2002, Gries 2003).

We hand-coded each specific particle verb meaning as either transparently compositional or non-compositional (Grafmiller and Szmrecsanyi 2018, Röthlisberger and Tagliamonte 2020). Following previous corpus studies, we use Lohse et al. (2004)'s diagnostic that a particle verb is only compositional if both the bare transitive (3a, 4a) and predication of the particle over the object (3b, 4b) are entailed. Thus, *knock off* meaning 'literally strike, causing to fall' is compositional, but *knock off* meaning 'stop some behaviour', is not compositional.

(3) compositional:

I knocked off his hat.

a. ✓ I knocked his hat.

b. ✓ His hat is off.

(4) non-compositional:

Knock off that singing!

a. *Knock that singing!

b. *That singing is (finally) off!

Particle prosody: Finally, exploratory visualization of word order variation by particle revealed that particle verbs containing *away* and *around*, the only two iambic particles, had visibly lower rates of joined usage than the rest of the particles (i.e., *down*, *in*, *off*, *on*, *out*, *over*, and *up*). We therefore investigate the prosody of the particle as a possible factor, along the lines of previous work that has explored whether phrasal prosodic considerations such as satisfying endweight (Ryan 2019) or conforming to alternating stress (Shih 2017, Grafmiller and Szmrecsanyi 2018) could affect word order alternations.

3.4 Statistical modeling

Particle verb order is a binary measure that is analyzed using logistic mixed-effects models with speaker modality as the main predictor, treatment-coded with studio as the baseline. The model includes control predictors of object length, NPR, log10 of meaning frequency, particle verb compositionality, and particle iambicity. The

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

model also includes a random intercept for each particle verb meaning, and was fit in R (R Core Team 2020) using the *glmer* function in the *lmer4* package, which provides p-values based on asymptotic Wald tests (Bates et al. 2015: 101).

4. Results

All control predictors were found to have significant effects on the variation. Complete model output is available in Appendix A, but to summarize: the joined variant is favoured with longer objects, more frequent particle verb meanings, non-compositional particle verbs, and on NPR shows. Object length and compositionality also interact significantly, such that the preference for joined order is stronger for longer objects with non-compositional particle verbs. The split variant is favoured with the iambic particles *around* and *away*. However, the predictor of speaker modality (studio vs. telephone) does not reach significance ($\beta = 0.09$, SE = 0.11, $p = 0.41$).

There are many data points in our "studio" category from radio shows that do not themselves have a call-in component. In order to see whether our null effect of speaker modality is somehow due to non-call-in show studio data skewing the results, we narrowed down the studio data to only those data points from shows that also have a telephone component in our data. But modality does not reach significance in this subset either ($\beta = 0.08$, SE = 0.11, $p = 0.48$). Figure 1 plots the rate of use of the joined variant among studio speakers in the full data set, the subset of studio speakers from shows that also have a telephone component ("call-in shows"), and telephone speakers. The three groups have identical rates of use of the joined variant at 81%.

We also checked for the possibility that speaker modality interacts with one of our other predictors. Speaker modality does not interact significantly with compositionality ($\beta = 0.05$, SE = 0.30, $p = 0.87$), object length ($\beta = 0.04$, SE = 0.15, $p = 0.79$), particle prosody ($\beta = -0.05$, SE = 0.57, $p = 0.93$), or meaning frequency ($\beta = 0.03$,

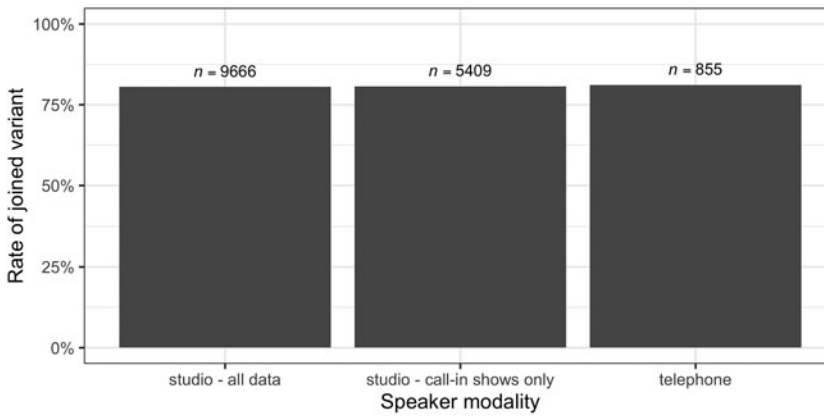

Figure 1: Rates of joined variant by speaker modality.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

SE = 0.15, $p = 0.86$). In all four cases, inclusion of the interaction raises AIC and BIC compared to the model with no interaction, and does not change the non-significant main effect of speaker modality.

5. Discussion

What might explain this failure to replicate Kroch and Small (1978)? One option is that the coding of studio vs. telephone modality is unreliable in our data. We think this unlikely, given the clear acoustic signature of telephone speech, as discussed in section 3.3. In the rest of this section, we discuss two possible sources of the difference between Kroch and Small (1978)'s results and ours.

5.1 Changes since the 1970s

It is possible that our null result represents a change in the prescriptive norms surrounding particle verb order since the 1970s. An acceptability study in Haddican, Johnson, and Hilton (2016) reveals that younger participants have a weaker preference for the joined order than older participants do. Moreover, Haddican, Johnson, Wallenberg, et al. (2020) find that both US and UK written Englishes are changing toward the split variant in real time, possibly suggesting "the loosening of a prescription against the VOP order in written English" (p. 220), which may or may not have a parallel in spoken English. Finally, in a matched guise study, Robinson and MacKenzie (2019) find that participants do not socially evaluate speakers who produce all particle verb tokens in the joined order differently from speakers who use all split.

Conversely, it is possible that prescriptive norms on particle verb order persist, but that radio hosts/in-studio guests do not contrast with callers on their adherence to those norms. This could either be because the two groups are too socially disparate with respect to factors that may matter more for adherence to prescriptivist norms,[3] or because talk radio's orientation towards those prescriptive norms has changed since the 1970s. The radio talk shows of 2018–19 inhabit and generate different social and interactional contexts than the radio talk shows of 1978. In the 1970s and 1980s, talk show audiences – and callers in particular – were socially peripheral (Turow 1974, Bierig and Dimmick 1979, Armstrong and Rubin 1989), but today's talk radio listenership is more integrated into general society (Hofstetter et al. 1994, Turow et al. 1996, Davis and Owen 1998). Moreover, regulatory shifts in the late 1980s and 1990s led to increasingly polarized content appealing to "talk-radio listeners[, who are] fairly united in their strong anti-government sentiments – sentiments that are often conditioned and reinforced by talk-show hosts" (Owen

[3]RadioTalk does not have other reliable social differences to test this hypothesis within the same dataset. Beeferman et al. (2019) did guess speaker gender for each snippet automatically, but those guesses are demonstrably inaccurate (repeated identical snippets of speech are often assigned different genders). RadioTalk also provides station city and state for each snippet, but speakers may not be from the locale of the broadcasting station, and syndication is very frequent, meaning that speakers may be associated with multiple stations in very different regions.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

1996: 134, Davis and Owen 1998). It is possible that today's talk radio hosts' mistrust of government and other powerful institutions extends to established standards in language, too. Perhaps talk radio hosts no longer orient themselves towards prescriptive linguistic norms in general; perhaps their stance towards dominant language ideologies has changed. If so, then future work with corpora with speech from contexts or speakers more susceptible to prescriptivist pressures may yet uncover social conditioning of particle placement.

5.2 Data and statistical problems with Kroch and Small (1978)

Finally, we have to recognize the possibility that the original effect in Kroch and Small (1978) was spurious. Specifically, their finding could be driven by differences in the length of direct objects used by each speaker group. As discussed above in section 3, particle verb constructions become nearly categorical in requiring the joined order with objects of greater than six words (Grafmiller and Szmrecsanyi 2018: 389), and we excluded tokens with objects of greater than four words for this reason. Kroch and Small, on the other hand, split tokens into those with short "L1" objects (<three words in length) and those with longer "L2" objects (three+ words). Longer objects which would categorically require the joined order seem to be included in that "L2" category (see, for instance, the underlined six-word object in their example (10c) *Harry mulled over the idea that his wife suggested*). Although the difference between speaker groups in the "L2" (three+ words) condition is not significant, this is likely due to the relative sparsity of caller data (36 tokens, 58% joined) vs. host/guest data (64 tokens, 73% joined).[4] It is therefore possible that the strongly joined host/guest "L2" sample could contain many very-long objects, which might then drive the overall effect of modality within the full dataset.

Another factor complicating efforts to replicate Kroch and Small's finding is lack of clarity concerning the envelope of variation and the operationalization of predictors in their study. For one, it is possible that the verb-particle constructions that Kroch and Small studied do not fully overlap with our definition of the phenomenon. In exemplifying their semantic dependence (compositionality) predictor, Kroch and Small write that "Cases where the particle is an adverb and functions as such in non-particle as well as particle constructions (e.g., *bring home the bacon*) were **excluded from the analysis of this factor**" (emphasis ours), which seems to imply that they included *bring home* elsewhere in the analysis (1978: 50). In our data, only constructions where *around*, *away*, *down*, *in*, *off*, *on*, *out*, *over*, or *up* function as the particle, and not elements like *home*, are included; it is possible that differences in defining the notion of particle verb explain the divergence between Kroch and Small's findings and our own.

Another such possible difference in defining the phenomenon under study arises from Kroch and Small's "D0" semantic category of particle verbs, which they

[4]The difference between groups in the "L1" (<three words) condition was reported to be significant (their Table 3), but that result may in fact also suffer from insufficient data: our own experience running a chi-square test on the numbers reported in Kroch and Small's Table 3 returned a *p*-value of 0.056.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

describe as cases where "the particle has no semantic content except as part of the verb" (p. 50). This category is both the best-represented in their data (153 (62%) out of the 246 tokens they coded for compositionality), and the only semantic category where the host/guest vs. caller distinction is significant (their Table 4). However, it is not clear which particle verbs exactly map to this category. Kroch and Small represent "D0" using their example (13) *He called the mayor up*, and subsequent work has argued compellingly that *up* (as used in *call up*) and other "aspectual particles" actually do have semi-predictable semantic contributions to a verb's aktionsart (Bolinger 1971, Brinton 1985, Tenny 1994, Jackendoff 2012, i.a. on such particles as telicity markers; see Cappelle 2005, Walkova 2013, Larsen 2014 for alternative approaches to the semantics of "aspectual" *up*). Kroch and Small's definition, however, also seems compatible with non-compositional particle verbs that are more opaque and idiomatic, like *carry out (duties)* or *knock off (behaviour)*. We include and code such tokens as non-compositional in our own dataset, but it is unclear to us whether Kroch and Small did so in theirs. In any case, our results show no interaction of speaker modality and semantic compositionality, regardless of whether we split aspectual particle verbs like Kroch and Small's example (13) from other types ($\beta = -0.31$, SE = 0.27, $p = 0.24$), or consider them together with all other non-transparently-compositional particle verbs (as described in section 3.3 and reported above). This both shows the indeterminacy of which particle verbs Kroch and Small were including for study, and demonstrates that our analysis is not missing a modality effect that is hidden in a particular syntactic/semantic class of particle verbs.

6. Conclusion

This study shows that Kroch and Small's finding – that social roles (in radio talk show contexts) condition the rate of joined variant use for particle verbs – does not replicate. It is possible that either the specific prescriptive linguistic norm under study has shifted, or that the way that talk radio relates to that norm has changed. Another possibility is that Kroch and Small's original study suffered from statistical flaws or issues with defining the phenomenon consistently.

The results of this study make the search for positive cases of social evaluation of syntactic variation even more pressing. Future work should both seek evidence of social influence on particle verb order in other types of data (including large-scale matched guise studies), and cast a wider net for the social correlates of syntactic variation in a variety of languages, variables, and data sources.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/cnj.2023.13

References

Adli, Aria. 2013. Syntactic variation in French *Wh*-questions: A quantitative study from the angle of Bourdieu's sociocultural theory. *Linguistics* 51(3): 473–515.

Armstrong, Cameron B., and Alan M. Rubin. 1989. Talk radio as interpersonal communication. *Journal of Communication* 39(2): 84–94.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48.

Beeferman, Doug, William Brannon, and Deb Roy. 2019. RadioTalk: A large-scale corpus of talk radio transcripts. *Proceedings of the 20th Conference of the International Speech Communication Association (INTERSPEECH 2019)*.

Beeferman, Doug, William Brannon, and Deb Roy. 2021. *RadioTalk GitHub Repository*. Laboratory for Social Machines. <https://github.com/social-machines/RadioTalk>.

Bierig, Jeffrey, and John Dimmick. 1979. The late night radio talk show as interpersonal communication. *Journalism Quarterly* 56(1): 92–96.

Bleaman, Isaac L. 2020. Implicit standardization in a minority language community: Real-time syntactic change among Hasidic Yiddish writers. *Frontiers in Artificial Intelligence* 3(35): 1–20.

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

Bouchard, Marie-Eve. 2018. Subject pronoun expression in Santomean Portuguese. *Journal of Portuguese Linguistics* 17(5): 1–29.

Bresnan, Joan, and Marilyn Ford. 2010. Predicting syntax: processing dative constructions in American and Australian varieties of English. *Language* 86(1): 168–213.

Brinton, Laurel J. 1985. Verb particles in English: Aspect or aktionsart? *Studia Linguistica* 39 (2): 157–168.

Cameron, Richard. 2000. Language change or changing selves? Direct quotation strategies in the Spanish of San Juan, Puerto Rico. *Diachronica* 17(2): 249–292.

Cappelle, Bert. 2005. Particle patterns in English: A comprehensive coverage. Doctoral dissertation, KU Leuven.

Chen, Ping. 1986. Discourse and particle movement in English. *Studies in Language* 10(1): 79–95.

Cheshire, Jenny. 1987. Syntactic variation, the linguistic variable, and sociolinguistic theory. *Linguistics* 25(2): 257–282.

D'Arcy, Alexandra, and Sali A. Tagliamonte. 2010. Prestige, accommodation, and the legacy of relative *who*. *Language in Society* 39(3): 383–410.

Davis, Richard, and Diana Marie Owen. 1998. *New media and American politics*. Oxford: Oxford University Press.

Dehé, Nicole. 2002. *Particle verbs in English: Syntax, information structure and intonation. Linguistics Today*, vol. 59. Amsterdam: John Benjamins Publishing.

Díaz-Peralta, Marina, and Manuel Almeida. 2000. Sociolinguistic factors in grammatical change: the expression of the future in Canarian Spanish. *Studia Neophilologica* 72(2): 217–228.

Fraser, Bruce. 1976. *The verb-particle combination in English*. Cambridge, MA: Academic Press.

Gorlach, Marina. 2004. *Phrasal constructions and resultativeness in English: A sign-oriented analysis*. Amsterdam: John Benjamins.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

Grafmiller, Jason, and Benedikt Szmrecsanyi. 2018. Mapping out particle placement in Englishes around the world: A study in comparative sociolinguistic analysis. *Language Variation and Change* 30(3): 385–412.

Gries, Stefan Thomas. 2003. *Multifactorial analysis in corpus linguistics: A study of particle placement*. New York: Continuum.

Guy, Gregory R., and Robert Bayley. 1995. On the choice of relative pronouns in English. *American Speech* 70(2): 148–162.

Haddican, Bill, and Daniel Ezra Johnson. 2012. Effects on the particle verb alternation across English dialects. In *University of Pennsylvania Working Papers in Linguistics 18.2: Selected papers from NWAV 40*, ed. Hilary Prichard, 31–40.

Haddican, Bill, Daniel Ezra Johnson, and Nanna Haug Hilton. 2016. Constant effects and the independence of variants in controlled judgment data. *Linguistic Variation* 16(2): 247–266.

Haddican, Bill, Daniel Ezra Johnson, Joel Wallenberg, and Anders Holmberg. 2020. Variation and change in the particle verb alternation across English dialects. In *Advancing socio-grammatical variation and change: In honour of Jenny Cheshire,* eds. Karen V. Beaman, Isabelle Buchstaller, Susan Fox, and James A. Walker, 205–228. New York: Routledge.

Hofstetter, C. Richard, Mark C. Donovan, Melville R. Klauber, Alexandra Cole, Carolyn J. Huie, and Toshiyuki Yuasa. 1994. Political talk radio: A stereotype reconsidered. *Political Research Quarterly* 47(2): 467–479.

Irvine, Judith T. 1985. Status and style in language. *Annual Review of Anthropology* 14(1): 557–581.

Jackendoff, Ray. 2012. English particle constructions, the lexicon, and the autonomy of syntax. In *Verb-particle explorations,* eds. Nicole Dehé, Ray Jackendoff, Andrew McIntyre, and Silke Urban, 67–94. Boston: De Gruyter Mouton.

Johnstone, Barbara, and Judith Mattson Bean. 1997. Self-expression and linguistic variation. *Language in Society* 26(2): 221–246.

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

Larsen, Darrell. 2014. Particles and particle-verb constructions in English and other Germanic languages. Doctoral dissertation, University of Delaware.

Levon, Erez, and Isabelle Buchstaller. 2015. Perception, cognition, and linguistic structure: The effect of linguistic modularity and cognitive style on sociolinguistic processing. *Language Variation and Change* 27(3): 319–348.

Lohse, Barbara, John A. Hawkins, and Thomas Wasow. 2004. Domain minimization in English verb-particle constructions. *Language* 80(2): 238–261.

Meyerhoff, Miriam. 1997. *Be I no gat*: constraints on null subjects in Bislama. Doctoral dissertation, University of Pennsylvania.

Meyerhoff, Miriam. 2000. *Constraints on null subjects in Bislama (Vanuatu): Social and linguistic factors*. Canberra: Pacific Linguistics.

Owen, Diana Marie. 1996. Who's talking? Who's listening? The new politics of radio talk shows. In *Broken contract?: Changing relationships between Americans and their government,* ed. Stephen C. Craig, 127–146. Boulder, CO: Westview.

R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org>.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

Robinson, Mary, and Laurel MacKenzie. 2019. *Socially-evaluated syntactic variation? A perception study of the English particle verb alternation*. Poster presented at the 2019 Annual Meeting of the Linguistic Society of America.

Romaine, Suzanne. 1981. The status of variable rules in sociolinguistic theory. *Journal of Linguistics* 17(1): 93–119.

Röthlisberger, Melanie, and Sali A. Tagliamonte. 2020. The social embedding of a syntactic alternation: Variable particle placement in Ontario English. *Language Variation and Change* 32(3): 317–348.

Ryan, Kevin M. 2019. Prosodic end-weight reflects phrasal stress. *Natural Language & Linguistic Theory* 37(1): 315–356.

Sankoff, David, and Suzanne Laberge. 1978. The linguistic market and the statistical explanation of variability. In *Linguistic variation: Models and methods,* ed. David Sankoff, 239–250. New York: Academic Press.

Shih, Stephanie S. 2017. Phonological influences in syntactic alternations. In *The morphosyntax-phonology connection: Locality and directionality at the interface,* eds. Vera Gribanova, and Stephanie S. Shih, 223–252. Oxford: Oxford University Press.

Tenny, Carol. 1994. *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer.

Turow, Joseph. 1974. Talk show radio as interpersonal communication. *Journal of Broadcasting & Electronic Media* 18(2): 171–180.

Turow, Joseph, Joseph N. Cappella, and Kathleen Hall Jamieson. 1996. Call-in political talk radio: Background, content, audiences, portrayal in mainstream media. *The Annenberg Public Policy Center of the University of Pennsylvania Report Series* (No.5).

Walkova, Milada. 2013. The aspectual function of particles in phrasal verbs. Doctoral dissertation, University of Groningen.

Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9 (1): 81–105.

Wasow, Thomas, and Jennifer Arnold. 2011. Post-verbal constituent ordering in English. In *Determinants of grammatical variation in English,* eds. Günter Rohdenburg, and Britta Mondorf, 119–154. Berlin: De Gruyter Mouton.

Wiechmann, Daniel, and Arne Lohmann. 2013. Domain minimization and beyond: Modeling prepositional phrase ordering. *Language Variation and Change* 25(1): 65–88.

CORPUS

RadioTalk Corpus. <https://github.com/mit-ccc/RadioTalk>

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

Appendix A: Model results

	Dependent variable:
	Use of joined variant
Compositionality (vs. Compositional)	
Non-compositional	0.889*** (0.227)
Object length (words, centered around mean)	0.484*** (0.090)
Speaker modality (vs. studio)	
Telephone	0.089 (0.108)
Broadcast service (vs. non-NPR)	
NPR	0.684*** (0.138)
Meaning frequency (log scale)	0.516*** (0.127)
Particle prosody (vs. non-iambic)	
Iambic	-2.376*** (0.425)
Compositionality × object length	
Non-compositional × object length	0.358*** (0.100)
Intercept	0.024 (0.253)
Observations	10,521
Log Likelihood	-4,088.192
Akaike Inf. Crit.	8,194.384
Bayesian Inf. Crit.	8,259.734

Note: *p <0.05; **p <0.01; ***p <0.001

Table 1: Logistic regression model of full data set. Accompanying each predictor are coefficient, standard error (in parentheses), and significance level (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). Coefficients for treatment-coded predictors should be interpreted in relation to the reference level, given in parentheses alongside the predictor name. other predictors are continuous.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press

	Dependent variable:
	Use of joined variant
Compositionality (vs. compositional)	
Non-compositional	0.614* (0.254)
Object length (words, centered around mean)	0.393** (0.121)
Speaker modality (vs. studio)	
Telephone	0.081 (0.113)
Broadcast service (vs. non-NPR)	
NPR	0.733*** (0.154)
Meaning frequency (log scale)	0.530*** (0.139)
Particle prosody (vs. non-iambic)	
Iambic	-2.693*** (0.492)
Compositionality x object length	
Non-compositional x object length	0.419** (0.133)
Intercept	0.169 (0.303)
Observations	6,264
Log Likelihood	-2,493.691
Akaike Inf. Crit.	5,005.382
Bayesian Inf. Crit.	5,066.065

Note: *p < 0.05; **p < 0.01; ***p < 0.001

Table 2: Logistic regression model of call-in show data set. Accompanying each predictor are coefficient, standard error (in parentheses), and significance level (*p < 0.05; **p < 0.01; ***p < 0.001). Coefficients for treatment- coded predictors should be interpreted in relation to the reference level, given in parentheses alongside the predictor name. Other predictors are continuous.

https://doi.org/10.1017/cnj.2023.13 Published online by Cambridge University Press