March 25, 2024

Addressing a challenge to the Backgroundedness account of islands[1]

Adele E. Goldberg, Nicole Cuneo, Abigail Fergus

Cuneo & Goldberg (2023) offer the following explanation of linguistic "islands", or constructions that resist combination with long-distance dependency constructions (LDDs): The functions of LDDs clash with the functions of certain base constructions, which "background" their content to varying degrees. The critical factor, Backgroundedness, was motivated and operationalized using two independent tasks on the base stimuli, not the LDDs. In a large-scale, preregistered study on three sets of LDDs combined with a wide range of base constructions, results showed that the degree of Backgroundedness predicted acceptability judgments on LDDs more than the base sentences themselves. Here we address a critique posted by Momma & Dillon (preprint) who propose two binary diacritics to account for our data. We reanalyze our data as suggested by Momma and Dillon and show that Backgroundedness remains explanatory, while empirical versions of the proposed diacritics show no effect. We additionally cite new work that manipulates backgroundedness and finds it explanatory.

1. Introduction

Recognizing that each grammatical construction serves a function or range of functions, Cuneo & Goldberg (2023) argue that in order for constructions to combine felicitously, their functions must be compatible. So-called "syntactic islands" involve a clash of functions as follows. Long-distance dependency (LDD) constructions make a certain constituent semantically prominent within the domain of the LDD. For instance, information questions such as (1) ask specifically about the *wh*-constituent, thereby making the semantic object of *read* the most prominent constituent of the question.

(1) What did she think [the students read _]?

Constructions that behave as "islands," on the other hand, are unacceptable to varying degrees when combined with an LDD construction. This is because, we argue, these constructions background their content, to varying degrees. Thus Cuneo & Goldberg (2023) provide support for the **BCI** in (2):

(2) **BCI:** Backgrounded constructions are islands (Goldberg, 2006; 2013).

[1] Data and analyses, linked in the original article are here: https://researchbox.org/1014. Data and models for the current analyses are available here https://osf.io/63chy/.

In order to measure the extent to which each base construction backgrounds its content, we operationalized Backgroundedness by means of two tasks, unrelated to one another: a Negation task (see also Ambridge & Goldberg, 2008; Namboodiripad et al. 2022) and a two-alternative forced choice Discourse task. Neither task involved any LDDs.

For the sake of space, and for reasons we return to in section 2.4, we focus here on the results of the Negation Task, as M&D do. The negation task is quite similar to the test for presuppositions, except that we anticipate judgments to be gradient: A construction backgrounds its content to the extent that its content is impervious to main clause negation. In particular, we presented a group of participants a series of negated sentences and asked them to judge, on a 5-point scale, the extent to which the target construction's content remained true, despite the main clause negation. An example is provided in Figure 1:

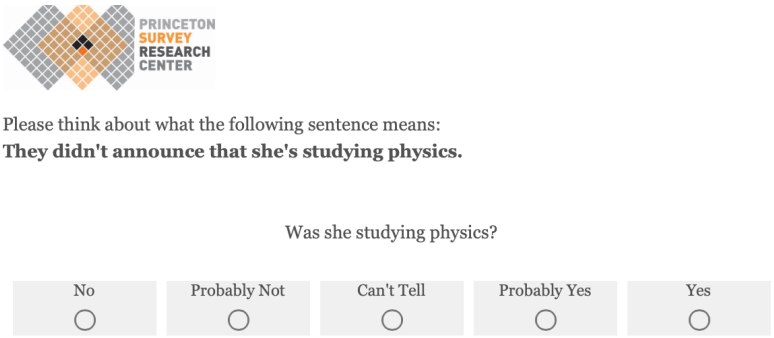

Figure 1: Example negation task trial from Cuneo & Goldberg (2023). The target construction is the clausal complement of the verb *announce*.

Responses that trended closer to "yes" indicated less impact of negation on the content expressed by the base construction, and a higher degree of Backgroundedness.

Whether and to what extent each target construction was judged to be an island was also determined empirically, by collecting acceptability judgments on 144 base sentences and three sets of corresponding LDD constructions: information questions, discourse-linked information questions, and relative clauses. Results confirmed the preregistered prediction (https://researchbox.org/1014): namely that each Backgroundedness measure predicted acceptability judgments on LDDs more than the base sentences ($N = 680$).

2. Addressing the critique by Momma & Dillon (2023)

We welcome Momma & Dillon's (hereafter, M&D) engagement with this work. M&D do not dispute our claim that Backgroundedness predicts acceptability judgments on LDDs more than base sentences across the full set of constructions tested. Rather, they claim that judgments are more directly explained by the assignment of two binary diacritics to semantic subtypes of our base sentences. We respectfully observe that *ceteris paribus*, a single factor that is operationalized and empirically quantified— as Backgroundedness was — is preferable to dividing the data up into four classes according to two binary features. Moreover, the

Backgroundedness measure was gathered independently of the phenomena to be predicted. That is, the Negation task did not include any island violations, nor did it ask for acceptability judgments. Instead, it probed the extent to which negation semantically targeted the content expressed by the base construction.

On the other hand, although the two features were labelled, *Adjunct* (A) and *Multiple events (ME)*, they were assigned in a rather jerry-rigged fashion so as to capture just those sentence types that have traditionally been considered islands. To see this, compare columns labeled *Traditional Islands* and *A or ME* in Table 1.

Table 1: The A and ME features proposed by M&D identify traditional islands

		Traditional Islands	A or ME	A	ME
1.	Main clauses	**0**	**0**	0	0
2.	Parasitic gaps *in adjuncts*	**0**	**0**	0	0
3.	Bridge verbs*[2]	**0**	**0**	0	0
4.	Factive verbs	**1**	**1**	0	1
5.	Causal adjuncts	**1**	**1**	1	0
6.	Manner-of-speaking	**1**	**1**	1	0
7.	Non-parasitic gaps in adjuncts	**1**	**1**	1	1
8.	Temporal adjuncts	**1**	**1**	1	1
9.	Relative clauses	**1**	**1**	1	1
10.	DO	**0**	**0**	0	0
11.	PO	**0**	**0**	0	0

Unsurprisingly, since A & ME are assigned in such a way to predict traditional islands and the Backgroundedness measure was shown to empirically predict island status, it is unsurprising that each feature assignment strongly correlates with Backgroundedness, as also noted by M&D. Specifically, the ME feature correlates with results on the Negation task (r = -.65; t = -10.19, df = 142, p < .00001), and so does the A feature (r = .-55 t = -8.05, df = 142, p < .00001).[3]

However, unlike the Backgroundedness proposal, no explanation is offered regarding the question Cuneo & Goldberg (2023) set out to address, namely *why* certain combinations of LDD and base constructions are less acceptable than other combinations. That is, *why* should adjunct

[2] Factive verbs are said to entail the truth of their complements (Kiparsky & Kiparsky, 1970), while "bridge" verbs are generally defined in a circular manner: they are verbs which allow LDDs that target their complement clause (e.g., verbs like *think, believe*). M&D do not state which of the verbs included in the C&G stimuli they considered to be bridge, manner of speaking, or factive verbs, and we did not stipulate these subcategories. To correlate each feature with Backgroundedness ratings, we used our own intuitions about which verbs M&D likely assigned to which categories.

[3] M&D report a slightly weaker correlation (of r = -.49) between ME and Backgroundedness, perhaps because they judged different verbs to be factive, bridge, or manner-of speaking. The difference in size of the correlation is not particularly relevant, so we do not investigate further.

constructions be islands and *why* might constructions that involve multiple events be islands? The Backgroundedness proposal suggests that the functions of LDD constructions and certain base constructions clash: LDD constructions make a particular constituent prominent while other constructions background the same content to varying quantified degrees. We fully anticipated that constructions would mediate the degree of Backgroundedness and our results indicate that they do (Cuneo & Goldberg 2023: Figure 3). For example, it is rare for main clause arguments to be backgrounded since main clauses generally express the main event; and modification relative clauses are generally backgrounded, while presentational relative clauses are more at-issue.

Do we believe Backgroundedness is the *only* factor that predicts the acceptability of LDDs? Certainly not. There are processing effects related to accessibility demands, confusability and working memory that are known to influence acceptability (Alexopoulou & Keller, 2007, Hofmeister et al. 2007; Kluender 1998; Liu et al. 2022). But is there any reason to believe a diacritic can replace the demonstrated influence of information structure on the acceptability of LDDs? If the proposed features M&D suggest were themselves independently motivated or empirically measured, either or both might lead to a potentially better explanation of island effects. Unfortunately, we highlight certain assignments of A and ME that we feel are *ad hoc*. And note that M&D do not provide any new data. We then offer a theory-neutral interpretation of adjuncts and report new empirical data on the number of events conveyed by the 144 sentences used as items in Cuneo & Goldberg (2023) in order to compare our Backgroundedness measure with theory-neutral versions of the two proposed factors in section 2.4.

2.1. A is for Adjuncts?

Arguments of an event are those entities that are central to the event and therefore semantically obligatory; if an argument is unexpressed, it is assumed to be recoverable in context. Because different types of events involve different types of arguments, arguments apply more narrowly to subsets of event types. In contrast, *adjuncts* are assumed to be less central to the main event. Therefore, adjuncts generally apply to a wide range of event types, do not influence the type of event being conveyed, and may be unexpressed. Although what counts as an argument vs. an adjunct can be debated,[4] the target constructions included in Cuneo & Goldberg (2023) are relatively straightforward. We will treat as adjuncts those cases that are uncontroversial including temporal adjuncts, causal adjuncts and other adjuncts whether or not they happen to co-occur with an additional gap in the main clause.

M&D assert that complements of verbs that express a manner of speaking are adjuncts, labelling them A+ (see also Baltin 1982). However, Ambridge & Goldberg (2008) had addressed this suggestion in earlier work, observing that clausal complements which express communicative content are only combined with a subset of verbs, a hallmark of arguments

[4] Obligatory adjuncts and optional arguments have each been proposed (e.g., Goldberg & Ackerman, 2001) and instrument complements are complicated (Koenig et al. 2003), but these cases were not included in the target stimuli of Cuneo & Goldberg (2023).

rather than adjuncts. And while the content clauses of manner-of-speaking verbs are typically optional, like typical adjuncts are, *unlike* typical adjuncts, the meaning of the verb can be importantly different if no clause is expressed. In particular, *scream* is only a legitimate manner-of-speaking verb in (2); if no content clause is expressed, as in (3), Chris may simply have emoted a sound ("Ahhhh!") (see also *screech, moan, growl).*

(2) Chris screamed that the duck had escaped.
(3) Chris screamed.

In addition, as Kogusuri (2009) cogently observes, manner-of-speaking verbs can alternatively appear with a direct object, as in (4) or with a clausal complement (as in 2), with the same general interpretation.

(4) Chris screamed the news.

Yet few would label a direct object an adjunct. Kogusuri (2009) concludes, in agreement with Ambridge & Goldberg (2008), that the content clauses of manner-of-speaking verbs are arguments rather than adjuncts.

Another unusual assignment M&D make regarding the A feature is that it is asserted that "parasitic" gaps *in adjuncts*, which also involve gaps in the main clause, are *non-adjuncts* (-A). No independent justification for this assignment is offered aside from the fact that it usefully predicts the fact that LDDs with parasitic gaps are generally acceptable (e.g., 5):

(5) What did he review _ without reading _ ?

2.2. ME is for Multiple events?

The second diachritic invoked by M&D is *ME*, which is intended to stand for "multiple events." M&D cite Truswell (2007), who had argued that LDDs from adjuncts are acceptable when the adjunct is construed to form single event with the main clause. Truswell's proposal is interesting and nuanced, but it is not operationalized, so is open to interpretation. M&D simply stipulate that sentences containing so-called factive verbs designate multiple events (+ME), and that manner-of-speaking verbs and so-called "bridge" verbs designate single events (-ME) (Table 1).

In contrast, Cuneo & Goldberg's stimuli were not subdivided into "bridge," "factive," "manner-of-speaking" or any other subcategories. And yet as reviewed in section 3, the gradient measure of Backgroundedness predicted variation in acceptability of the LDD stimuli more than the base sentences across all verb + complements.

Liu et al. (2019) had found that acceptability judgments correlated with the (log) frequency of the main verbs appearing with clausal complements, so Cuneo & Goldberg gathered and included frequencies as an additional factor, confirming that frequency accounted for some variation in judgments while Backgroundedness remained a significant predictor. Intriguingly, as Cuneo and Goldberg further report, Backgroundedness and log frequencies correlated with one another: the more frequently a verb is used with a clausal complement the

more likely was its complement to be considered more at-issue (and less Backgrounded). This latter finding suggests that frequency and backgroundedness are not completely independent, at least in the case of verbal complements: highly frequent verbs tend to become semantically bleached (Givón 1981), allowing the complement clause to be interpreted as relatively more at-issue.

M&D additionally stipulate that temporal verbal adjuncts encode multiple events while causal verbal adjuncts encode a single event. However, separate work which had operationalized and collected judgments on whether a single or multiple event was expressed in a range of adjuncts has found that the negation task better predicted island status among adjuncts than whether multiple events were expressed (Namboodiripad et al. 2022). In section 2.4 we provide empirical evidence of whether each sentence expresses one or two events and test its validity as a predictor of island status. But first we clarify how Cuneo & Goldberg (2023) had provided evidence that Backgroundedness predicts variation *within* grammatical constructions as well as across them.

3. Backgroundedness predicts variation *within* verb + complement clauses and verbal adjuncts

M&D provide a lengthy discussion on mediating variables in the introduction of their critique of Cuneo and Goldberg (2023). This seems to be a claim that Backgroundedness (only) predicts an intermediate level of representation rather than directly predicting acceptability judgments. We had fully expected individual constructions to mediate the degree of backgroundedness, since each construction is used for a particular range of functions. And as mentioned above, there is good reason to believe each construction plays a key role in determining the range of backgroundedness conveyed by its content. In fact, the "backgrounded constructions are islands" hypothesis would seem to make it clear that constructions play a critical role in which parts of a sentence are backgrounded.

The studies conducted by Cuneo and Goldberg (2023) were not designed to test whether Backgroundedness predicted variation *within* each construction tested. If that had been our aim, we could have for instance, included both modificational and presentational relative clauses, but instead all of the RCs were quite similar. Nor did we include a wide variety of adjuncts. Instead, only 12-24 quite similar sentences per construction were included, so we had assumed it would be difficult to detect item variation within individual constructions. Nonetheless, we did provide exploratory analyses comparing items within the pairs of sentence-types used in the 2AFC Discourse task, which allowed us to control for meaning and complexity since each pair of options was designed to be highly similar on these dimensions. M&D are critical of this and suggest that we should have instead included individual constructions as random effects rather than pairs of constructions. Since here we are focused on the Negation task rather than the 2AFC task, we do include constructions as random effects in section 2.4.

It is worth observing, that for two subsets of the stimuli: a) main verbs with clausal complements and b) verbal adjuncts, the only variation among items was semantic. And in both cases, results confirmed an interaction between judgments and Backgroundedness: Backgroundedness predicted acceptability judgments on the LDDs more than the base

constructions (2023: sections 7.3 and 7.6). These cases thus demonstrated that Backgroundedness did indeed account for semantic variation when these syntactic constructions were held constant.

Moreover, we can test the original data set while including random intercepts and slopes on the 8 syntactic constructions tested: main, RC, DO, PO, verbs+clausal_complements, verbal adjuncts (regardless of semantics), parasitic gap adjuncts and non-parasitic gap islands. As shown in Table 2, results confirm the same effect reported in Cuneo & Goldberg (2023), namely a significant interaction between Backgroundedness and LDD status: Backgroundedness predicts decrements in acceptability that are unexpected based on the base sentences ($ß = .31$, $z = 2.85$, $p = 0.004$).

Table 2: Ordinal mixed model confirms significant interaction of independent Backgroundedness measure and whether stimuli are base sentences or LDDs, with random effect structure for subjects, items and constructions. (Each subject judged only Base or only LDD constructions.)

```
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: raw_response ~ Backgrounded * BaseorLDD + (1 | subject) + (1 +      BaseorLDD |
item) + (1 + BaseorLDD | construction)
data:    data

 link  threshold nobs  logLik    AIC      niter       max.grad cond.H
 logit flexible  17279 -24633.88 49299.77 2278(15041) 2.42e-02 9.8e+03

Random effects:
 Groups       Name        Variance Std.Dev. Corr
 subject      (Intercept) 0.9521   0.9758
 item         (Intercept) 0.6516   0.8072
              BaseorLDD   1.0700   1.0344   -0.039
 construction (Intercept) 0.8375   0.9151
              BaseorLDD   2.0351   1.4266   0.964
Number of groups:  subject 480,  item 144,  construction 8

Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
Backgrounded               0.26201    0.07812   3.354 0.000796 ***
BaseorLDD                 -3.46615    0.58611  -5.914 3.34e-09 ***
Backgrounded:BaseorLDD  0.31322    0.10960   2.858 0.004266 **
```

4. Including adjuncts and an empirical measure of multiple events as additional fixed factors

One might ask whether the data we collected on 144 sentences is fully predicted by the features proposed by M&D. Given the reasons to be skeptical of the way A and ME assignments were made, outlined above, we here determine the mean number of events expressed by crowd-sourcing judgments, as described below. We also here treat as adjuncts, temporal adjuncts, causal adjuncts and other adjuncts, whether or not they happen to co-occur with an additional gap in the main clause. Then we test whether either the number of events expressed or adjunct status or a combination of both factors eliminates the influence of Backgroundedness, as M&D assert. On the contrary, results confirm the predicted interaction of Backgroundedness and acceptability in each model: Backgroundedness predicts decrements in acceptability that are unexpected from judgments on the base sentences. Thus, even when we take adjunct status and the number of events expressed into account, Backgroundedness still predicts island status.

Participants

98 participants were recruited from Prolific and paid for their time.

Procedure

The 144 base sentence stimuli from Cuneo & Goldberg (2023) were divided into 4 lists of 36 sentences. Each participant was asked to judge whether each sentence (of one list) expressed one or more than one event. Responses across participants were averaged so that each stimuli sentence was assigned a score [1, 2]. Scores were then centered, so that they ranged [-0.5, 0.5]. Along with new empirical data on how many events are expressed in each stimuli sentence, we use the empirical data on acceptability judgments and the Backgroundness measure (the negation task) from Cuneo & Goldberg (2023).

Results

First, an ordinal mixed model with acceptability judgments as the outcome, Backgroundedness, mean number of events and their interaction with the type of stimuli (Base or LDD) as fixed factors, and random effect structure for subjects, items and constructions was created (Each subject judged only Base or only LDD constructions.) The predicted interaction between Backgroundedness and LDDs is significant ($ß = .33$, $z = 2.70$, $p = 0.007$). The interaction with the number of events shows no effect on judgments ($ß = .14$, $z = 0.33$, $p = 0.74$). See Table 3.

Table 3: Mixed ordinal model predicting acceptability judgments on the basis of Backgroundedness and/or the empirical measure of the number of events, and their interaction with sentence type (base or LDD)

```
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: raw_response ~ Backgrounded * BaseorLDD + mean_events + mean_events:BaseorLDD +
(1 | subject) + (1 + BaseorLDD | item) + (1 + BaseorLDD |
    construction)
data:    data

 link  threshold nobs  logLik    AIC      niter       max.grad cond.H
 logit flexible  17279 -24633.82 49303.64 2659(19458) 3.23e-02 1.3e+04

Random effects:
 Groups       Name        Variance Std.Dev. Corr
 subject      (Intercept) 0.9523   0.9759
 item         (Intercept) 0.6514   0.8071
              BaseorLDD   1.0702   1.0345   -0.039
 construction (Intercept) 0.8295   0.9107
              BaseorLDD   2.0524   1.4326   0.966
Number of groups:  subject 480,  item 144,  construction 8

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
Backgrounded           0.26227    0.08532   3.074  0.00211 **
BaseorLDD             -3.55334    0.67495  -5.265  1.4e-07 ***
mean_events           -0.02463    0.29402  -0.084  0.93323
Backgrounded:BaseorLDD  0.32873    0.12193   2.696  0.00702 **
BaseorLDD:mean_events   0.13464    0.40624   0.331  0.74032
```

We also test an ordinal mixed model with acceptability judgments as the outcome, and *adjunct status*, Backgroundedness, and their interaction with the type of sentence (base or LDD) as fixed factors, with random effect structure for subjects, items and constructions (Each subject judged only Base or only LDD constructions.) As shown in Table 4, the predicted interaction

between Backgroundedness and LDDs is again significant ($ß = .33$, $z = 3.15$, $p = 0.002$). The interaction with adjunct status is not significant ($ß = -1.44$, $z = -1.55$, $p = 0.12$).

Table 4: Mixed ordinal model predicting acceptability judgments on the basis of Backgroundedness and/or adjunct status, and their interaction with sentence type (base or LDD) as fixed effects.

```
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: raw_response ~ Backgrounded * BaseorLDD + adjunct + adjunct:BaseorLDD +
(1 | subject) + (1 + BaseorLDD | item) + (1 + BaseorLDD |      construction)
data:    data

 link  threshold nobs  logLik    AIC      niter       max.grad cond.H
 logit flexible  17279 -24631.85 49299.70 2786(22592) 1.50e-02 1.8e+04

Random effects:
 Groups       Name        Variance Std.Dev. Corr
 subject      (Intercept) 0.9515   0.9754
 item         (Intercept) 0.6562   0.8101
              BaseorLDD   1.0755   1.0371   -0.046
 construction (Intercept) 0.7799   0.8831
              BaseorLDD   1.5311   1.2374   1.000
Number of groups:  subject 480,  item 144,  construction 8

Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
Backgrounded            0.25370    0.07519   3.374  0.00074 ***
BaseorLDD              -3.67026    0.55679  -6.592 4.35e-11 ***
adjunct                -0.50287    0.66589  -0.755  0.45014
Backgrounded:BaseorLDD  0.32730    0.10382   3.153  0.00162 **
BaseorLDD:adjunct      -1.43939    0.93069  -1.547  0.12196
```

5. Why do current results differ from M&D's analyses?

Since current results demonstrate that the degree of Backgroundedness consistently predicts judgments on LDDs more than base sentences, even when adjunct status and/or multiple event status are included as potential additional predictors, it is fair to ask why M&D did not find the same thing? We note the following differences between our analyses and theirs. First, our measure of multiple event status was gathered empirically (section 2.2). We also treat as adjuncts, all adjuncts regardless of semantics or whether there is another gap in the main clause or not (recall section 2.1).

Finally, M&D treat nonfinite verbal and temporal adjuncts as distinct syntactic constructions while we here considered them one grammatical construction with a range of different semantic interpretations. M&D take issue with the idea of treating all non-finite verbal adjuncts as a class because there is a potential confound between the verb form (gerund vs. infinitive) and semantic interpretation. This issue was already addressed by Nambooridiripad et al. (2002) who carefully controlled for verb form and nonetheless found an impact of Backgroundedness. But to be conservative, in a final reanalysis of the original data, we divide verbal adjuncts into two separate constructions: those that receive a temporal interpretation and those that receive a causal interpretation.

This model, which conservatively divides our stimuli into 9 grammatical constructions rather than 8, again demonstrates that Backgroundedness remains a significant predictor of judgments, particularly for the LDD sentences, as evidenced by the significant interaction ($ß = .26$, $z = 2.26$, $p < .03$). In particular, the more negation impacts the base construction, the less

backgrounded the base construction is, and as predicted, the more acceptable the corresponding LDD is.

Thus, Backgroundedness passes the higher bar, suggested by M&D, than what we had claimed: even when one includes adjunct status *and* single-event status and their interactions with BaseorLDD status in the model, and even when we divide our data into 9 distinct constructions, Backgroundedness significantly influences judgments on LDDs more than base sentences, while the empirical measure of multiple events does not and neither does adjunct status (see also Namboodiripad et al., 2022).

Table 5: New model that includes Backgroundedness, adjunct status, and an empirical measure of multiple events as fixed factors.

```
Cumulative Link Mixed Model fitted with the Laplace approximation

formula: raw_response ~ Backgrounded * BaseorLDD + mean_events + mean_events:BaseorLDD +
adjunct + adjunct:BaseorLDD + (1 | subject) + (1 + BaseorLDD |
    item) + (1 + BaseorLDD | Cx)
data:    data

 link  threshold nobs  logLik    AIC      niter       max.grad cond.H
 logit flexible  17279 -24627.48 49294.96 2804(20339) 1.84e-02 1.2e+04

Random effects:
 Groups  Name        Variance Std.Dev. Corr
 subject (Intercept) 0.9520   0.9757
 item    (Intercept) 0.6452   0.8033
         BaseorLDD   0.9951   0.9975   -0.085
 Cx      (Intercept) 0.7628   0.8734
         BaseorLDD   1.6524   1.2855   1.000
Number of groups:  subject 480,  item 144,  Cx 9

Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
Backgrounded            0.20207    0.08066   2.505   0.0122 *
BaseorLDD              -3.67198    0.59205  -6.202 5.57e-10 ***
mean_events             0.09808    0.28576   0.343   0.7314
adjunct                -0.55373    0.60715  -0.912   0.3618
Backgrounded:BaseorLDD  0.25727    0.11379   2.261   0.0238 *
BaseorLDD:mean_events   0.55622    0.38551   1.443   0.1491
BaseorLDD:adjunct      -1.59658    0.88735  -1.799   0.0720 .
```

A direct challenge to M&D's claim that Backgroundedness does not cause island status *directly* comes from new work that manipulates Backgroundedness and finds an impact on acceptability judgments. Lu, Pan, and Degen (2024) introduce this technique in an analysis of manner-of-speaking verbs (section 6.1), which inspired us to conduct a new study on the double object construction (section 6.2).

6. Manipulating Backgroundedness

6.1. Manner of speaking verbs plus complements

Lu et al. (2024) reason that if Backgroundedness influences acceptability judgments on LDDs, then focal stress on a targeted constituent should encourage comprehenders to include that constituent within the focus domain, thereby making it more prominent and less Backgrounded. This should result in higher acceptability of an immediately following LDD question (e.g., 9b), when compared to an example that instead makes the manner-of-speaking verb at-issue, as in (9a). Their results confirm this prediction.

(9a) *Verb-focus:* Sam didn't WHISPER that Chris met with the lawyer.

Then who did Sam whisper that Chris met with?

(9b) O*bject-focus:* Sam didn't whisper that Chris met with the LAWYER**.**
Then who did Sam whisper that Chris met with?

Lu et al. (2024) included frequencies of each verb appearing with a clausal complement in their analysis, since Liu et al. (2021) had found an effect of frequency. No effect of frequency was evident in Lu et al. (2024) data.

In a separate manipulation, Lu et al. (2024) additionally find that so-called "bridge" verbs behave just as manner-of-speaking verbs do when a manner adverb is added, an observation predicted by Deane (1991):

(10) ?? Who did she <u>say softly</u> that Chris met with?

That is, Lu et al. (2024) report that manipulations of backgroundedness involving the same verbs and same constructions *cause* acceptability judgments to *increase* when backgroundedness of the complement clause is reduced (e.g., in 9b), compared to when backgroundedness of the complement clause is increased (e.g., 9a or 10), consistent with the Backgrounded Constructions are Islands proposal.

These results inspired us to revisit the Double Object construction and conduct a new study, described below.

6.1 A New study: Manipulating backgroundedness in the Double Object construction

As reported in Cuneo & Goldberg (2024), information questions of the recipient argument of the Double Object (DO) construction resisted being combined with LDDs, and so appeared to be island-like (Erteschik-Shir 1979; Goldberg 2006, 2013). Yet in the comparison of DO and its prepositional paraphrase, the ways we had operationalized Backgroundedness failed to show any evidence for the BCI. We speculated that this was because it is possible to make the recipient argument of the DO construction prominent, since it is after all a main clause argument, but we had no empirical evidence for this and simply acknowledged it as exceptional.

Here we briefly report a new study that indicates the DO construction is in fact amenable to an explanation in terms of Backgroundedness. Inspired by Lu et al. (2024) we varied whether the recipient argument in a declarative DO is stressed or not just before querying the recipient in an information (*wh-*) question. Acceptability judgments on the wh-questions are collected. We hypothesized that the acceptability of a *wh*-question should be significantly higher when it followed a context-sentence that stressed the recipient argument in comparison to a neutral context-sentence. As described below, this hypothesis was confirmed. Preregistration is available here: https://aspredicted.org/see_one.php.

Participants

120 English-speaking participants were recruited from Prolific and paid for their time. After exclusions for failure on either one of two practice trials, or for lack of confirmation that participants interpreted capitalization to indicate stress, data from 85 participants were included.

Design and Stimuli

Twenty-four sets of stimuli were created. In each, a double-object context-sentence was followed by an information question that queried the recipient argument. Two versions of each stimulus set were created and distributed evenly across four lists. In the stress condition, the recipient argument of the context-sentence was capitalized to indicate lexical stress (see Figure 2 top). The neutral (control) condition was identical, except no particular stress was indicated (no capitalization Fig 2 bottom).

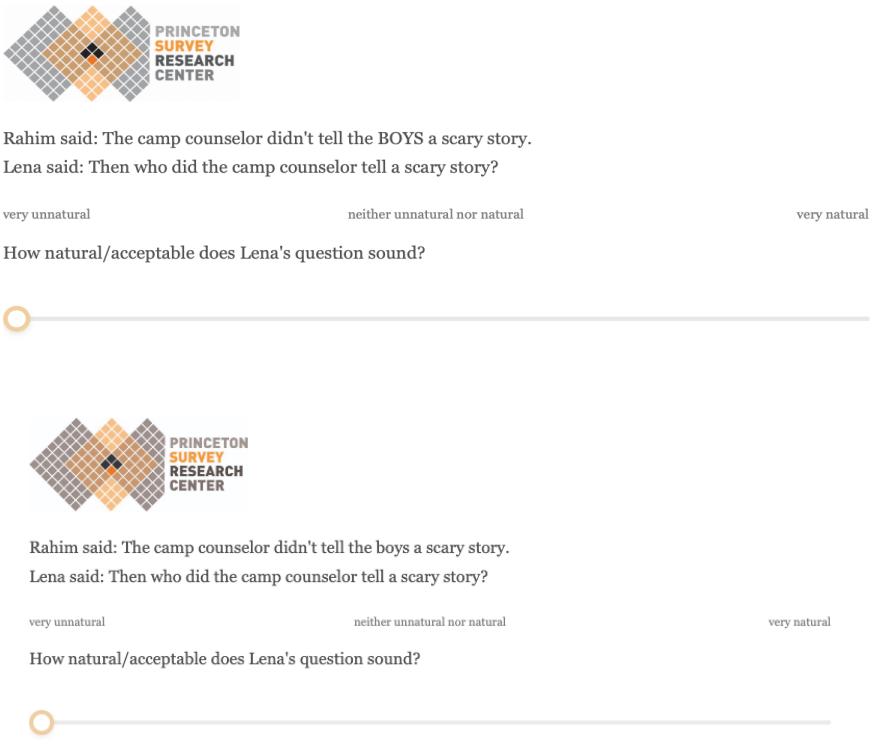

Figure 2: Example stimulus from the stress condition (top): the recipient argument of double-object construction receives lexical stress and the neutral condition (bottom). No participant saw both versions of the same item.

Log frequencies of the main verbs appearing in the double object construction were determined based on COCA (Davies, 2009) in searches performed on Mar 1, 2024 using "VERB PRO a" for each main verb.

The design was within-participants: each participants judged the acceptability of 6 *wh*-questions of the DO construction in the stress condition and 6 in the neutral condition. Items

were counterbalanced across participants so that no one saw the same content twice. Each person also judged the acceptability of 16 filler questions, half of which followed a context-sentence that included stress on the word to be queried, and half of which followed a neutral context-sentence. Participants saw an 16 additional sentences designed for another study (total trials = 44 trials).

Stress was not expected to influence judgments on the fillers, because either the queried content was already prominent in the discourse (acceptable fillers) or because the *wh*-question was unacceptable because of morphological or lexical oddities (unacceptable fillers). An example of each type of filler is provided in Table 6:

Table 6: Example acceptable (A) and unacceptable (B) filler sentences also varied whether the queried argument was stressed in a context sentence (left) or not (right). Acceptability judgments on the questions were collected.

Stressed context sentence	**Unstressed context sentence**
A. Acceptable fillers (example)	
JANET didn't lock the office door	Janet didn't lock the office door.
Then who forgot to lock the office door?	
B. Unacceptable filler (example)	
KEISHA didn't call the company about the power line fallen.	Keisha didn't call the company about the power line fallen.
Then who called the company about the power line fallen?	

Results

A linear mixed model was created with response as output, condition (stress or neutral), and log frequency of the verb in the double-object construction as fixed factors and the most complex random effect structure convergence allowed: random intercepts and slopes for subjects and random intercepts for items. Results confirmed the stress condition showed higher acceptability ratings than the neutral condition (M_{stress} = 53.0, $M_{neutral}$ = 41.1; ß = 9.29, t =5.67, $p <$ .000001). That is, including stress on the recipient argument of a DO construction before querying the recipient argument in a *wh*-question significantly increased the question's acceptability. The relative (log) frequency of the verbs appearing in the double object construction showed no effect (ß= 2.54, t = 1.08, p = .29). Stress did not impact the acceptability judgments on filler items, as confirmed by a mixed model with response predicted on the basis of stress with the maximal random effect structure convergence allowed for participants and items (acceptable fillers: ß = -2.67, t = -0.88, p = 0.41; unacceptable fillers: ß = 0.67, t = 0.31, p = 0.75).

Discussion

It is well-known that the recipient argument of the DO construction resists LDDs to some extent, despite the recipient being a main clause argument. Cuneo & Goldberg (2024) tested

whether this was a result of the recipient being relatively backgrounded in discourse (Bresnan et al. 2007; Erteschik-Shir 1979; Goldberg, 2006; Wasow 2002). As reported, the negation test did not reveal evidence for this when the DO and the prepositional dative were compared, perhaps because wh-questions with the prepositional dative are far more conventional or because the recipient argument of the DO is amenable to being made prominent in the discourse.

Inspired by Lu et al. (forthcoming), here we manipulated Backgroundedness directly by providing a context-sentence in which the recipient was stressed just before a *wh*-question. Results show that manipulating the stress on the recipient argument of the DO construction led to significantly higher acceptability ratings on the following *wh*-questions, consistent with the BCI proposal. Since the identical *wh*-questions were judged in both stressed and neutral conditions, and granted the assumption that lexical stress as indicated by capitalization indicates greater prominence on a particular argument, results demonstrate that manipulating the degree of Backgroundedness of the recipient argument in a DO construction does in fact influence acceptability judgments.

7.Conclusion

The basic idea of the BCI proposal is quite simple: Comprehenders find it odd for a speaker to choose to both background and make prominent the same content. Long-distance dependency (LDD) constructions make a constituent semantically prominent within the domain of the LDD. Constructions are less acceptable when combined with LDD constructions (are islands) to the extent that their content is Backgroundedness. Cuneo & Goldberg (2024) tested this empirically by operationalizing Backgroundedness and crowd-sourcing judgments on 144 base sentences and three types of corresponding LDDs. Here we respond to Momma & Dillon (2023)'s critique of this work. They suggest binary diacritics on semantic subtypes of our stimuli are sufficient to explain the acceptability judgments we had collected. While it is generally possible to stipulate diacritics and call them mediating factors, empirical evidence that is truly independent of the phenomena one is attempting to explain is preferable. Critically neither feature suggests an explanation as to *why* it should result in reduced acceptability of LDDs. Moreover, beyond the fact that one empirically determined gradient measure is preferable to two stipulated binary features, the proposed diacritics were particularly problematic in that they were assigned to stimuli in an ad hoc way. Tellingly, each one as assigned, correlates with our empirical measure of Backgroundedness.

Nonetheless, we took the suggested additional features seriously by collecting an empirical measure of multiple-events and assigning adjunct status in a theory-neutral way. We again find Backgroundedness predicts judgments on LDDs (more than the base stimuli) even with the inclusion of these additional features and find that neither has a significant effect.

Regarding M&D's claim that Backgroundedness predicts (only) an intermediate level of representation, we had been prepared to believe that Backgroundedness applies to the level of construction rather than directly to individual sentences. Constructions after all are constrained to specify aspects of information structure. The Backgrounded Constructions are Islands proposal explicitly appeals to constructions in fact. Nonetheless, in the exploratory analyses reported by Cuneo & Goldberg (2023), Backgroundedness was found to predict variation within

the set of main verbs with clausal complements; Lu, Pan and Degen (forthcoming) have now manipulated Backgroundedness and find that it predicts judgments even within the more narrowly defined class of manner-of-speaking verbs with clausal complements. Here we report new data demonstrating that the double object construction is amenable to manipulating backgroundedness: *wh*-questions that query the recipient are significantly more acceptable when the context makes clear that the recipient is at-issue. Finally, Cuneo & Goldberg (2023) had reported variation within the class of verbal adjuncts, and variation predicted by Backgroundedness has independently been found even when the verb form is tightly controlled for by Namboodiripad et al. (2022).

The goal of Cuneo & Goldberg (2024) was to empirically test our functional proposal of island constraints. Here we address theoretical concerns raised by M&D by applying even more constrained models to our data which include as fixed factors empirical versions of the features they suggest. Results show that Backgroundedness predicts acceptability judgments collected by Cuneo & Goldberg (2024), while the proposed factors do not. Finally, we provide new empirical support for the BCI proposal by manipulating Backgroundedness in the double-object construction. The functions of grammatical constructions must be compatible for them to combine felicitously.

References

Alexopoulou, T., & Keller, F. (2007). Locality, cyclicity, and resumption: At the interface between the grammar and the human sentence processor. *Language*, 110-160.

Ambridge, B., & Goldberg, A. E. (2008). The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics*, *19*(3).

Baltin, M. R. (1982). A landing site theory of movement rules. *Linguistic Inquiry,13*,(1), 1-38.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In *Cognitive foundations of interpretation* (pp. 69-94). KNAW.

Cuneo, N., & Goldberg, A. E. (2023). The discourse functions of grammatical constructions explain an enduring syntactic puzzle. *Cognition*, *240*, 105563.

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International journal of corpus linguistics*, *14*(2), 159-190.

Deane, P. (1991). Limits To Attention: A Cognitive Theory Of Island Phenomena. *Cognitive Linguistics, 2*(1).

Erteschik-Shir, N. (1979). Discourse constraints on dative movement. In *Discourse and syntax* (pp. 441-467). Brill.

Givón, T. (1981). On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica*, *15*(2), 35-54.

Goldberg, A.E. (1995) *Constructions: a Construction Grammar approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, A.E. (2006) *Constructions at Work: the nature of generalization in language*. Oxford University Press.

Goldberg, A.E. (2013). Backgrounded constituents cannot be "extracted". In Sprouse and Hornstein (eds). *Experimental Syntax and Island Effects*, *221*.

Goldberg, A. E., & Ackerman, F. (2001). The pragmatics of obligatory adjuncts. *Language*, 798-814.

Hofmeister, P., Jaeger, T. F., Sag, I. A., Arnon, I., & Snider, N. (2007). Locality and accessibility in wh-questions. *Roots: Linguistics in search of its evidential base*, *185*.

Kiparsky, P and Kiparsky, C. (1970). "Fact" in Berwisch and Heidolph (eds). *Progress in Linguistics*. Mouton.

Kluender, R. (1998). On the distinction between strong and weak islands: A processing perspective. In *The limits of syntax*(pp. 241-279). Brill.

Kogusuri, T. (2009). Manner of speaking verbs and their clausal complements in English. *Tsukuba English Studies*, *27*, 187-203.

Koenig, J. P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, *89*(2), 67-103.

Liu, Y., Ryskin, R., Futrell, R., & Gibson, E. (2019). Verb frequency explains the unacceptability of factive and manner-of-speaking islands in English. In *CogSci* (pp. 685-691).

Liu, Y., Winckel, E., Abeillé, A., Hemforth, B., & Gibson, E. (2022). Structural, functional, and processing perspectives on linguistic island effects. *Annual Review of Linguistics*, *8*, 495-525.

Lu, J., Pan, D., & Degen, J. (2024). *Evidence for a discourse account of manner-of-speaking islands*. Linguistic Society of America Conference. NYC.

Momma, S., & Dillon, B. (2023) Discourse factors do not explain islands. *Available at SSRN 4635713*.

Namboodiripad, S., Cuneo, N., Kramer, M. A., Sedarous, Y., Bisnath, F., Sugimoto, Y., & Goldberg, A. (2022). Backgroundedness predicts island status of non-finite adjuncts in English. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44).

Stoica, I. (2021). *The syntax and the semantics of manner of speaking verbs*. Editura Universității din București-Bucharest University Press.

Truswell, R. (2007). Extraction from adjuncts and the structure of events. *Lingua*, *117*(8), 1355-1377.

Wasow, T. (2002). *Postverbal behavior*. Stanford: CSLI.