

Combined Constructions Must be Compatible: Evidence from the Coordinate Structure Constraint

Abigail B. Fergus¹ and Adele E. Goldberg¹

¹Department of Psychology, Princeton University

Author Note

Abigail Fergus <https://orcid.org/0009-0000-0188-9102>

Adele E. Goldberg <https://orcid.org/0000-0001-8636-8890>

We have no conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Adele Goldberg:
adele@princeton.edu.

Abstract

Speakers generally prefer semantically related content to be expressed nearby in the string, yet certain constructions allow semantic long-distance dependencies (LDDs). Why do LDDs exist and how and why are they constrained? Here, new evidence is brought to bear in support of the following claim: LDDs, which make a constituent prominent, may not target backgrounded content, nor can they asymmetrically target an aspect of content construed as symmetric. Studies demonstrate that LDDs from single conjuncts are rated more acceptable to the extent that (i) the targeted conjunct expresses more prominent content or (ii) the entire conjunction is construed as describing a single complex event rather than two events. Processing factors, namely familiarity and intransitivity of the first conjunct, are found to influence the acceptability of LDDs in addition to the discourse-semantic factors.

Keywords: islands; conjunction; long-distance dependencies; information structure

**Combined Constructions Must be Compatible:
Evidence from the Coordinate Structure Constraint**

Languages tend to position semantically related units close together in the string (Futrell, 2017; Gibson, 2000; Hahn & Xu, 2022; Hawkins, 2014). The tendency to express semantic relationships locally makes sense from a processing perspective, since language users need to build hierarchical semantic structure incrementally, and holding content in working memory is costly. Nonetheless, under certain circumstances, many languages allow semantically related terms to be expressed at some distance from one another. For example, the information question (*wh*-Q) in (1a) positions *which section* a relatively long distance away from the verb required to interpret it: *rewrite*. A similar long-distance dependency occurs in English relative clauses (RCs) as shown in (1b):

- 1a. Which section did he think the editor asked the authors to rewrite__? (*wh*-Q)
 b. She liked the section that he had thought the editor asked the authors to rewrite__. (RC)

Perhaps unsurprisingly, given the general preference for locality, long-distance dependences (LDDs) are not uniformly acceptable. Complements of manner of speaking verbs such as *mumble in* (2a-b) are generally less amenable to LDDs than semantically simple verbs (such as *think in* [1a-b]) (Ross, 1967). That is, the examples in (2a) and (2b) are less acceptable than those in (1a) and (1b), respectively, despite their apparently identical syntax.

- 2a. ? Which section did he mumble the editor asked the authors to rewrite __?
 b. ? She liked the section that he had mumbled the editor asked the authors to rewrite__.

In the linguistics literature, constructions that resist being combined with LDDs have been dubbed *islands*, a term intended to metaphorically evoke the idea that content within an island may not “move off of it” (Ross, 1967). Thus, complements of manner-of-speaking verbs are considered islands, while complements of verbs such as *think* are not. Constraints on which constructions LDDs can combine with have traditionally been viewed as resulting from one or more categorical syntactic rules, unrelated to any discourse function that LDDs or island constructions might serve (Bresnan, 1976; Chomsky, 1973, 2014; Fox & Lasnik, 2003; Sag, 1976; Sabel, 2002; Sprouse & Hornstein, 2013).

Others have argued that constraints on LDDs stem from the cognitive demands required to process them in comparison with sentences in which semantic relations are expressed locally. Such accounts have observed that increased interference during processing results in reduced acceptability (Gibson, 2000; Kluender, 2004; Kluender & Kutas, 1993; Ness & Meltzer-Asscher, 2017). A related proposal has argued that differences in acceptability arise from differences in exposure: for instance, the fact that “*think* + clausal complement” is far more frequent than

“*mumble* + clausal complement” is argued to facilitate speakers’ processing of (1a) compared to (2a) (Kothari, 2008). The role of frequencies remains debated; certain studies have found evidence consistent with frequencies playing a role (Dąbrowska, 2013; Liu et al., 2022), while others have not found an effect (Huang et al., 2025; Lu et al., 2025).¹

Notably, a fundamental question has been all but ignored by both syntactic and processing explanations: *Why do languages ever allow long-distance dependencies in the first place?* That is, insofar as it is possible to express semantic relationships locally and local relationships are easier to process, it seems perverse for LDDs to exist at all.

A third approach begins by addressing the question of why LDDs exist, explicitly or implicitly, by recognizing the communicative functions of LDD constructions (Abeillé et al. 2020; Chaves, 2022; Chaves & Putnam, 2020; Deane, 1992; Goldberg, 2006; 2013; Erteschik-Shir, 1973; Kuno, 1973): The point of *wh*-Qs is to request information about the targeted phrase; the point of relative clauses is to identify or modify the targeted content. Formally, each LDD construction positions the meaningful target phrase in a particularly prominent position: at the “edge” of the domain of the LDD. Each LDD construction involves elevating the prominence of the targeted content within the domain of the LDD. The domain of information questions (*wh*-Qs) is the entire sentence, and the domain of relative clauses (RCs) is the noun phrase. Correspondingly, in the *wh*-Qs (1a and 2a), *which section* is sentence-initial; in the relative clauses in 1b and 2b, *section* is NP-initial. That is, LDDs align semantic or discourse prominence with formal prominence.

This suggests an explanation for why LDDs exist *and* why certain LDDs are less acceptable than others. Constraints on LDDs exist because LDDs require the targeted content to be eligible for the special prominence required by the LDD construction. Goldberg (2006; 2013) stated the generalization as follows:

Backgrounded Constructions are Islands (BCI): Constructions are islands to long-distance dependency constructions to the extent that their content is backgrounded within the domain of the long-distance dependency construction.

The BCI predicts that island effects arise because LDD constructions make a constituent prominent and combining them with content that is taken for granted or backgrounded results in a clash of functions. That is, it is infelicitous for a speaker to simultaneously make certain content particularly prominent, by using an LDD, while simultaneously backgrounding it (e.g., by including it in the complement of a verb like *mumble*). Growing experimental evidence is consistent with the BCI (Ambridge & Goldberg, 2008; Cuneo & Goldberg, 2023; Dąbrowska, 2013; Goldberg, 2006, 2013; Goldberg et al., 2024; Lu et al., 2025; Namboodiripad et al., 2022).

The majority of empirical work lending support to the BCI or related discourse accounts has focused on clausal complements of main verbs, as in the examples provided in (1) and (2)

(Ambridge & Goldberg, 2008; Cuneo & Goldberg, 2023; Dąbrowska, 2013; Liu et al., 2022; Lu et al., 2025; Richter & Chaves, 2020). Lu, Pan, and Degen (2025) demonstrate a causal influence of prominence, by providing a context sentence that emphasized either the target content or the main verb, before collecting ratings on *wh*-Qs that followed.

An experiment that included a variety of island constructions (including clausal complements, adjuncts and others) found that the degree of backgroundedness of the targeted content in corresponding base sentences predicted acceptability judgments on LDDs (*wh*-Qs and RCs), while showing little to no correlation with the acceptability of the base sentences themselves, which were more uniformly judged acceptable (Cuneo & Goldberg, 2023). Base sentences are corresponding simple sentences without the LDD (e.g. 3a), and the degree of backgroundedness is then determined by negating the main clause of the base sentence (3b) and querying the degree to which the potential island construction is affected (Ambridge & Goldberg, 2008):

- 3a. He mumbled that the editor asked the authors to rewrite a section. (base sentence; potential island construction underlined)
- 3b. Negation: He didn't mumble that the editor asked the authors to rewrite the section.
 Query: Did the editor ask the authors to rewrite the section?
 (probably not → more backgrounded)

Empirical work on untensed adjunct phrases is also consistent with the claim that Backgrounded constructions are islands (Namboodiripad et al., 2022).

Constraints on Coordination

Here we consider a case that has long been believed to require its own distinct constraint: conjunctions. Ross (1967) posited a special “Coordinate Structure Constraint” to account for the fact that LDDs generally cannot target elements within conjunctions. This constraint is quite stark when an LDD targets an entire conjunct (for discussion, see Altshuler & Truswell, 2022):

- 4a. ?? What was she happy and ____?
- b. ?? Who did he meet __ and Sally?

Note that the notion of backgroundedness cannot explain the unacceptability of (4a) or (4b) since the targeted content is not taken for granted or backgrounded in the corresponding base sentence (e.g., *He met Harry and Sally*). To see this, note that content that is backgrounded is less affected by main clause negation. As is evident in (5), neither conjunct is backgrounded, since main clause negation unambiguously negates each conjunct:

5. She met Harry and Sally.

Negation: She didn't meet Harry or Sally.

Query 1: Did she meet Harry? No → not backgrounded

Query 2: Did she meet Sally: No → not backgrounded

We therefore propose a new constraint, more general than the BCI, namely, Combined Constructions must be Compatible (CCC). We spell this out as it relates to LDD constructions in two parts, corresponding to the BCI constraint in (CCC part 1) and an explanation of constraints on coordinations in (CCC part 2):

CCC: Combined Constructions must be Compatible: The function of an LDD construction, to make an entity especially prominent within its domain, must be compatible with the function of the targeted “base” construction.

- a) It is infelicitous for a speaker to choose to background content while simultaneously choosing to make the same content particularly prominent in an LDD (paraphrase of the BCI).
- b) It is infelicitous for a speaker to choose a symmetric construction (logical conjunction) while simultaneously choosing to make only (part of) *one* conjunct especially prominent with an LDD.

Thus, we hypothesize that symmetrical conjunctions combined with LDDs (e.g., 4a-b) are unacceptable because the LDD construction is incompatible with the conjunction: it is infelicitous for a speaker to target a symmetric conjunction while simultaneously choosing to make only one conjunct especially prominent in a *wh*-question. If the explanation for the Coordinate Structure Constraint is actually this straightforward, it should apply to a wide range of conjunction types, including apparent exceptional cases. This brings us to our current quarry: LDDs that target conjoined verb phrases.

The traditional Coordinate Structure Constraint accurately predicts the unacceptability of the example in Figure 1 (left). Yet intriguingly, as Ross (1967) had already noted, LDDs targeting a single verb phrase conjunct are sometimes acceptable, as illustrated by the example on the right side of Figure 1 (i.e., *What did he run to the store and buy?*).



Figure 1: Example *wh*-Q LDDs from a single conjunct. The example on the left is more degraded than the example on the right.

Apparently acceptable LDDs that target a single verb phrase conjunct have been subjected to a good deal of theoretical discussion (Altshuler & Truswell, 2022; Bošković, 2019; Goldberg, 1998; Goldsmith, 1985; Lakoff, 1986; Na & Huck, 1992; Oda, 2017; Schmerling, 1975; Stjepanović, 2014), but rarely have they been examined experimentally. An exception is Harris (2009), who collected acceptability judgments on examples designed to vary in their degree of symmetry, but this work did not correlate acceptability with an independent measure of symmetry. The current work addresses this lacuna.

Specifically, we test the hypothesis that the degree of asymmetry in verb phrase conjunctions predicts the degree to which LDDs from a single conjunct are acceptable.

The Function of Verb Phrase Conjunction

What is the function of a verb phrase conjunction? Logical conjunction canonically presents multiple entities, events, or states as equally prominent: “A & B” is the same as “B & A.” We will see evidence that the conjunction in examples like (6) represent logical conjunction (underlined):

6. She drinks gin and eats steak.

At the same time, natural language is not logic (e.g., Ariel, 2012; Ariel & Mauri, 2019; Ariel, 2023; Jennings, 1994). Some verb phrase conjunctions are better described as depicting a single complex event, rather than two symmetric events of equal prominence (Ariel, 2012; Harris, 2009; Truswell, 2006; Altshuler & Truswell, 2022; Boguraev & Mahowald, 2026). For example, we will see that the conjunction in (7) is construed as representing a single complex event:

7. He ran to the store and bought apples.

Empirically, we test whether acceptability ratings on LDDs are predicted by independently measured factors determined by the discourse or semantic properties of base sentences with verb phrase conjunctions. We also test whether processing factors—familiarity and

interference—additionally influence acceptability ratings on LDDs. More specifically, we test whether each factor predicts acceptability of LDDs more strongly than it predicts the acceptability of the corresponding base sentences.

Experiment 1 measures the extent to which the to-be-targeted conjunct is more prominent by asking participants which conjunct expresses the main idea. Experiment 2 *manipulates* the prominence of the to-be-targeted phrase, by adapting the methodology of Lu, Pan, and Degen (2025) for use with verb phrase conjunctions.

Several theoretical proposals suggested that exceptions to the coordinate structure constraint involved verb phrase conjunctions that express a single complex event rather than multiple distinct events (Altshuler & Truswell, 2022; Chaves 2022; Goldberg, 1998; Goldsmith, 1985; Harris, 2009; Lakoff, 1986; Ross, 1967; Truswell, 2006). When a conjunction describes a single complex event, the event's internal semantics (e.g., path–goal, means–end, cause–effect) renders one subevent more prominent than the other, satisfying the CCC constraint, namely that combined constructions must be compatible, by allowing the more prominent phrase to be targeted by an LDD, without a functional clash. Conversely, when the conjunction expresses two logically independent events of equal status, neither conjunct is inherently more prominent than the other, so that if an LDD targets only one, the base and LDD constructions are incompatible (CCC part 2).

We therefore also test the claim that LDDs that target a single verb phrase (VP) conjunct will be more acceptable to the extent that the conjoined VPs are construed as a single complex event or state rather than two distinct events. New groups of participants are asked for explicit judgments in experiment 3 or are asked about an implicit measure of single event status in experiment 4.

In exploratory analyses we additionally test two factors to determine whether ease of processing predicts the acceptability of LDDs and whether these factors are sufficient to explain the variability in ratings. First, more familiar conjuncts should be easier to process, so we hypothesize that the log frequency of the first conjunct appearing in verb phrase conjunctions should influence ratings on LDDs, using searches of the billion-word COCA corpus (Davies, 2009). Secondly, when an LDD targets the second of two conjuncts, an initial transitive conjunct should interfere more, because the transitive verb suggests a possible interpretation for the phrase targeted by *who* or *what* (Hofmeister & Sag, 2010; Kluender, 2004), as comprehenders interpret language incrementally (Rayner & Clifton, 2009). Finally, we test whether the discourse-semantic properties confirmed in experiments 1-4 influence acceptability ratings, over and above the combination of processing factors.

Results are consistent with the claim that Combined Constructions must be Compatible (CCC). In particular, it is infelicitous for a speaker to choose a symmetric construction (logical conjunction) while simultaneously choosing to make only (part of) *one* conjunct especially prominent in an LDD. We also find evidence that ease of processing influences acceptability

judgments on LDDs more than base sentences. Finally, we confirm that the processing constraints hold in addition to, not instead of, the discourse-semantic factors found relevant in Experiments 1-4.

Methods

Preregistration

Data sets, R code, and stimuli can be found here. Preregistration is available [here](#). Model outputs are in the SI. Linear mixed models were created in which acceptability ratings were predicted by an interaction of sentence type (base sentences, *wh*-Qs, or RCs) and the factor measured in each Experiment, with base sentences treated as the reference level. Random slopes and intercepts are included for sentence type on items, and random intercepts are included for participants. Slopes on participants are not warranted, except in Exp 2, because separate groups of participants rated the acceptability of base sentences or a single type of LDD. The tasks in each experiment were performed by a unique group of participants, with resulting item means included as the relevant fixed factor for each experiment. The between-participant design ensures that each data set is independent of acceptability judgments and judgments on any other task. We calculate Cohen’s d for effect sizes in the most conservative way available: $d = \beta / \text{total SD}$ ($\sqrt{\text{sum of random-intercept variances} + \text{residual variance}}$) (Borenstein et al., 2009). We diverge from the preregistration in the following ways: first, in accord with the preregistration, we assign 20 participants per item, but due to a miscalculation, this requires 120 participants to judge Base sentences, *wh*-Qs, and Relative clauses (total: 360), rather than the 60 mentioned in the preregistration. Secondly, we report linear rather than ordinal models in the text because cumulative link mixed models with the preregistered random effect structure showed poor convergence properties in all cases, by exceeding clmm’s default gradient tolerance by 34x to 196x. Simplified ordinal models were not uniformly improved, but are included in the SI, and show the same significant effects except in Exp. 3¹. Finally, the incongruent emphasis condition of Exp. 2, and the frequency and intransitivity factors were analyzed after the preregistration was finalized.

Participants

A total of 680 English-dominant participants were recruited from Prolific and paid for their time. After exclusions for either not completing the task or failing more than half of the practice questions intended to be straightforward, 615 participants were left for analyses (337 F, 264 M, 10 nonbinary, 4 prefer not to say; $M_{\text{age}} = 40.19$ [18, 83] $SD = 14.13$).

¹ The one exception is that for Experiment 3, where explicit judgments were analyzed, the ordinal model reveals only a marginal interaction in the predicted direction for *wh*-questions ($p = .0645$) and a non-significant interaction for RCs ($p = .12$).

Stimuli

A set of 96 base sentences was created, each containing two conjoined verb phrases without any long-distance dependencies. The base sentences were used to create two types of corresponding LDDs: information questions (*wh*-Qs) and sentences with relative clauses (RCs). *Wh*-Qs are straightforwardly related to base sentences by querying the argument in the second verb phrase. Sentences with relative clauses contained a main clause agentive subject and transitive verb in the past tense; RCs were attached to direct objects. Because the function of restrictive relative clauses is to identify the noun referent, we used past perfect tense within relative clauses so that the event(s) described in the relative clauses were construed to precede the event denoted by the main verb. Two example stimuli of each type (base, *wh*-Qs, and RCs) are provided in Table 1.

Table 1: Two sample stimuli of each type: base sentences (no LDDs), *wh*-Q, and RCs.

Total stimuli of each type: 96 (= 288 sentences).

Base Sentences

- a. She picked up some apples and ran back for potatoes
- b. She woke up and thought of a great solution.

Wh-Qs

- a. What did she pick up some apples and run back for?
- b. What did she wake up and think of?

RCs

- a. She ate the potatoes that she had picked up some apples and had to run back for.
- b. She wrote down the solution that she had woken up and thought of.

Acceptability Judgments

We collected acceptability ratings on sentences containing verb phrase conjunctions in the three types of sentences from unique groups of participants: i) simple declarative non-LDD base sentences, ii) information questions (*wh*-Qs) that request information from the second of two conjuncts based on each base sentence; iii) relative clauses that modify a noun from only the second of two conjunctions (RCs). Each participant rated 16 of the 96 items on a single type of sentence. Six lists of items for each sentence type were counterbalanced across participants and presented to individual participants in randomized order. Participants were asked to rate how natural the sentences sounded on a 7-point Likert scale labeled [1: completely unnatural] to [7: perfectly natural]. Each participant also saw the same 20 filler items, half intended to be acceptable and half, unacceptable, to avoid saturation effects.

Results

As intended, mean acceptability ratings on base sentences were reasonably high [4.24, 6.95], $M = 6.45$, $SD = .42$, while ratings on both types of LDD showed a good deal of variation across items: *wh*-questions: [1.67, 6.33], $M_{wh-Qs} = 3.66$, $SD = 1.18$; RCs: [1.5, 5.94], $M_{rc} = 3.38$, $SD = 1.02$. Our goal is to explain why some LDDs involving only the second of two conjuncts are more acceptable than others. (Mean scores for each item as base, *wh*-Q, and in sentence with RC are included in the SI.)

Experiment 1: LDDs are More Acceptable when the to-be-targeted Conjunct Expresses the Main Idea

Experiment 1 is motivated by the Backgrounded Constructions are Islands (BCI) hypothesis outlined in the introduction. Specifically, the BCI predicts that *wh*-Qs and relative clauses should be more acceptable when they target content within a more prominent, less backgrounded conjunct. A subgroup of 40 Prolific participants responded to a 3-alternative forced choice task, with the 96 items randomly divided across two lists, so that each participant responded to a total of 48 trials.

The task presented the base sentences, one at a time, and asked participants to select whether the main idea was expressed by the first conjunct or the second conjunct; a third option allowed participants to say they were unsure. We predicted that the greater the proportion of participants who decided the second conjunct expressed the main idea in the base sentences, the more acceptable the corresponding *wh*-Qs and RCs should be (as LDDs always targeted the second conjunct), with the same factor being a weaker influence on the acceptability of the base sentences themselves.

Results

Unsurprisingly, we find a general negative effect of both *wh*-Qs and RCs on ratings compared to base sentences, which served as the reference level (*wh*-Qs: $\beta = -2.69$, $t = -17.15$, $p < .0001$; RCs: $\beta = -3.05$, $t = -20.67$, $p < .0001$). This makes sense since both types of LDDs are more complex and harder to process than the base sentences. Because the same acceptability judgments are used as a factor in each of the models described here, we omit restating the overall negative effects of each LDD in what follows. We also find a general effect of the second conjunct expressing the main idea. Acceptability judgments are higher when the proportion of participants deciding the 2nd conjunct expresses the main idea is higher, perhaps because of a preference for sentence final focus ($\beta = .46$, $t = 3.60$, $p = .0005$).

Critically, as shown in Figure 2, the predicted interactions were also significant: as the proportion of participants judging the second conjunct as conveying the main idea increased, the acceptability of *wh*-Qs and RCs increased more strongly than base sentences (main idea x *wh*-Q interaction: $\beta = 1.32$, $t = 3.75$, $p = .0003$, $d = .66$; main idea x RCs $\beta = 1.28$, $t = 4.57$, $p < .0001$, $d = 0.64$).

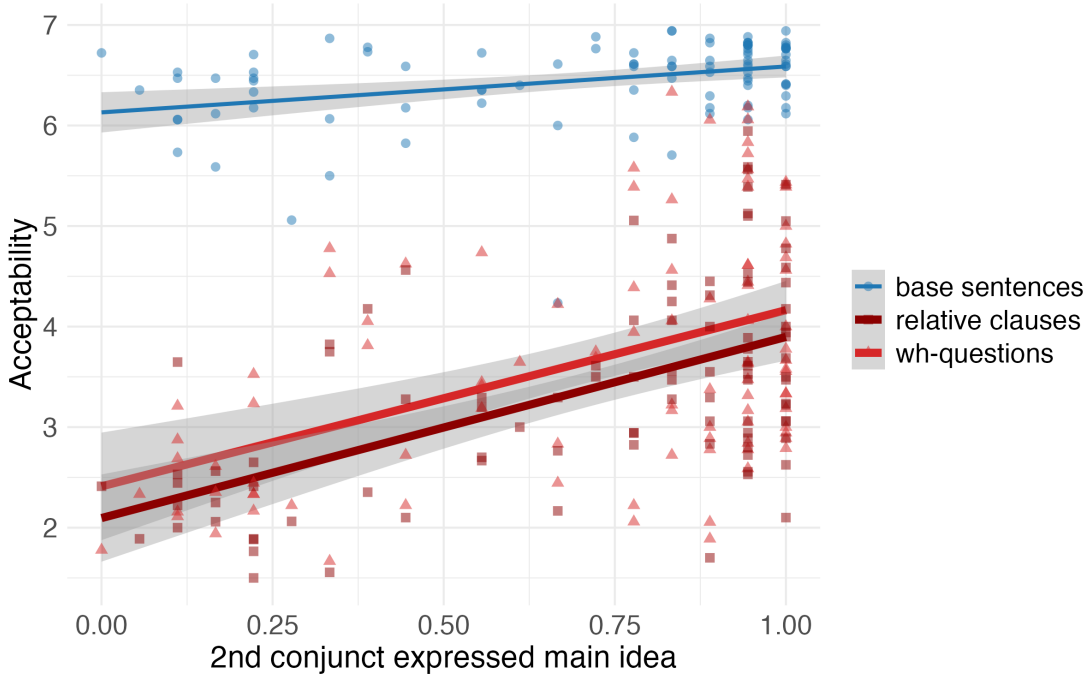


Figure 2: The acceptability of *wh*-Qs and RCs increases, more than base sentences, as the proportion of participants who judge the 2nd conjunct as expressing the main event is higher. Points: Mean acceptability judgments on base sentences (blue circles) and two types of LDDs: *wh*-Qs (brighter red triangles) and RCs (darker red squares).

Experiment 2: Manipulating Discourse Prominence

Experiment 2 *manipulates* the prominence of the to-be-targeted constituent in base sentences by adapting the methodology of Lu, Pan, and Degen (2025), who had developed it to investigate complements of manner-of-speaking verbs. We adapt it to test the influence of manipulating the prominence of a to-be-targeted constituent in a conjunction. Since the method was designed to test the effects of prominence on *wh*-Qs, not RCs, we only test *wh*-Qs in Experiment 2. A subgroup of 160 Prolific participants responded to an acceptability judgment task on a 7-point Likert scale, with the $96 \times 3 = 288$ items divided across 8 lists such that each participant responded to 36 trials.

Participants were presented with a series of mini conversations, in which two people discuss the content of a proposition expressed by a conjoined verb phrase. Participants were always asked to rate the acceptability of a *wh*-Q. Conditions varied in the type of “context sentence” that was read just before the *wh*-Q was rated. Example stimuli in all three conditions are provided in Fig 3. In a congruent prominence condition, the argument that was targeted in the following *wh*-Q was emphasized in the preceding context sentence, with capital letters (e.g., *The team didn’t win a prize and eat PIZZA to celebrate*) (Fig 3: left); a neutral condition included an identical context sentence but with no added prominence (no

capitalization) (Fig 3: center); finally, an incongruent prominence condition emphasized an argument in the first conjunct, while the *wh*-Q queried an argument in the second conjunct (Fig 3: right). Each participant saw items from each condition, while none saw the same *wh*-Q more than once, as item-condition was varied across participants.

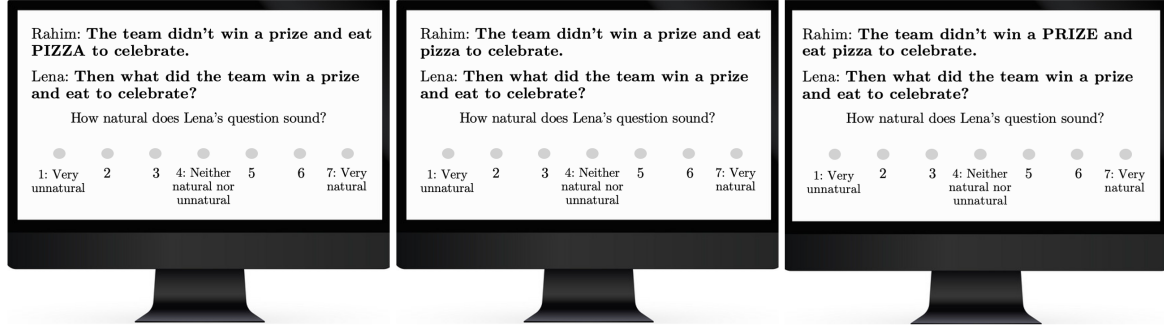


Figure 3: Context sentences with emphasis on congruent (left), incongruent (right) or without emphasis (center) immediately before a *wh*-question that was judged for acceptability.

If prominence in discourse plays a causal role in the acceptability of single-conjunct *wh*-Qs, we predict ratings will increase, in comparison to a neutral context sentence when the to-be-targeted constituent is emphasized in a context sentence presented just before the *wh*-Q is judged. We also predict that ratings will decrease if a different, contextually incongruent constituent is emphasized immediately before the *same wh*-Q (incongruent emphasis here is on an argument of the non-targeted conjunct).

Both predictions were confirmed by a linear mixed model predicting acceptability ratings from a three-level Emphasis factor, with neutral stress as the reference level (Fig. 4 center). Because data for Exp. 2 was collected within-participants, we include random intercepts and slopes on both participants and items. Results confirm that discourse prominence (emphasis) causally influences ratings on *wh*-Qs. The acceptability of *wh*-Qs increases when the to-be-queried constituent is emphasized in the preceding sentence ($\beta = 0.25$, $t = 3.34$, $p = .001$, $d = .11$) (Fig 4: left), and decreases when emphasis is placed within the conjunct that is then *not* targeted in the following question ($\beta = -.31$, $t = -4.15$, $p < .001$, $d = -.14$) (Fig 4: right). Points in Fig 4 represent mean ratings on the 96 *wh*-Qs, with segments connecting the identical *wh*-Q across conditions. See SI for full model.

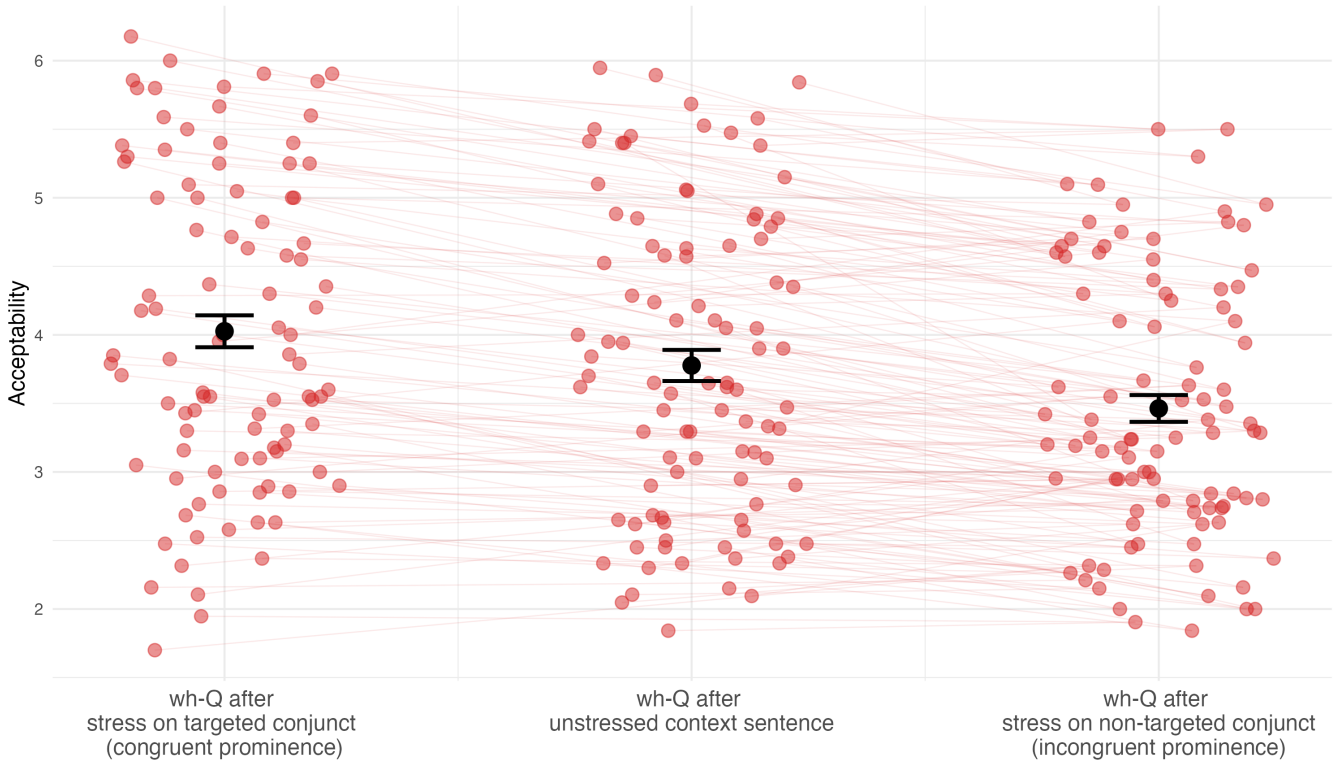


Figure 4: Acceptability of *wh*-Qs from a single conjunct after a context sentence that emphasizes—increases prominence—of the to-be-targeted content (left), a neutral context sentence (center), and after a context sentence with emphasis incongruent with the *wh*-Q (right).

Single Complex Events are More Amenable to *Wh*-Qs and RC LDDs

Experiment 3: Explicit Query

Experiment 3 simply asked participants explicitly whether each base sentence expressed a single complex event (or state) or two separate events (or states), in order to correlate results with acceptability judgments on the LDDs and on the base sentences themselves. As noted, item order was randomized for each participant. A subgroup of 40 Prolific participants responded to a two alternative forced-choice task, with the 96 items randomly divided across two lists, so that each participant responded to 48 total trials.

Results. The factor of interest is the extent to which conjoined VPs are construed as a single complex event. This factor was centered and scaled. VP conjunctions are generally more acceptable when the verb phrases are construed as a single event ($\beta = .22$, $t = 5.89$, $p < .0001$). Critically, as predicted and shown in Figure 5, construal as single event predicts greater acceptability of *wh*-Qs and relative clauses (in red) more than the factor influences base sentences (in blue). That is, the predicted interactions are significant (single event x *wh*-Q

interaction: $\beta = .41$, $t = 3.78$, $p = .0002$, $d = .21$; single event x RC interaction: $\beta = .37$, $t = 4.24$, $p < .00001$, $d = .19$).

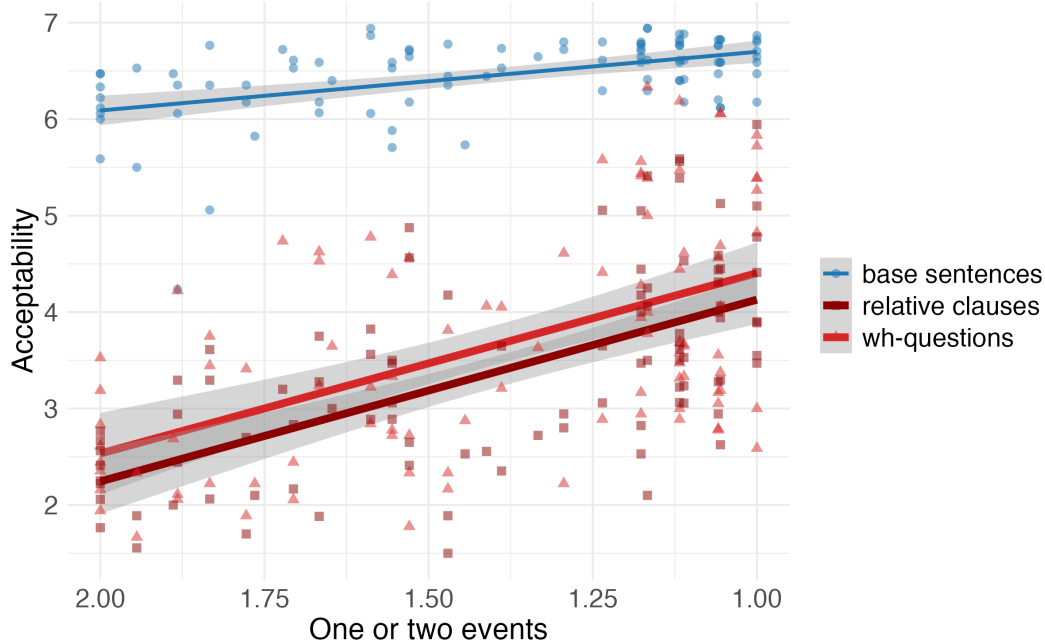


Figure 5: x-axis: Mean construal of each conjunction as depicting one single event (or state) or two distinct events (or states). Points: Mean acceptability judgments on base sentences (blue circles) and two types of LDDs: *wh*-Qs (brighter red triangles) and RCs (darker red squares).

Experiment 4: Implicit Measure: Negating Disjunction vs. Conjunction

While experiment 3 asked participants explicitly whether conjunctions described one complex event (or state) or two distinct events (or states), Experiment 4 offers an implicit measure of the same construct. To see this, first observe that negated logical conjunctions are expressed with disjunction (*or*) rather than *and* (see 6 below). This follows from DeMorgan’s law: $\neg(A \ \& \ B) \iff \neg A \text{ or } \neg B$. This implies that negated disjunctions as in (9a) should be somewhat more acceptable than negated conjunctions (9b), as indicated by the “>” symbol.

- 9a. She doesn’t drink **or** smoke. >
- b. She doesn’t drink **and** smoke.

At the same time, by hypothesis, English conjunction can be used to describe a single complex event. In fact, if (9b) is interpreted as negating a single complex event (“she doesn’t drink and smoke at the same time”), negation with *and* is made acceptable.

Experiment 4 asked participants to rate the acceptability of negated conjunctions (10a, 11a) and negated disjunctions (10b, 11b). Each of the 96 base sentences was negated with *and* in one version and *or* in the other (192 sentence stimuli in total). No participant saw both versions of the same base sentence.

- 10a. The team didn't win a prize and eat pizza to celebrate.
 b. The team didn't win a prize or eat pizza to celebrate.
 11a. She didn't wake up and think of a great solution.
 b. She didn't wake up or think of a great solution.

We predict that the *difference* in acceptability between negated disjunction and negated conjunction should correlate with acceptability ratings on *wh*-Qs and RCs more strongly than it correlates with the acceptability of the base sentences themselves. Specifically, the larger the difference in acceptability between negated disjunction and negated conjunction, the less acceptable the corresponding LDDs should be. A subgroup of 80 Prolific participants rated the acceptability of a random subset of the 192 items, divided across four lists, so that each participant responded to 48 total trials.

Results

The prediction was confirmed, as depicted in Figure 6. The greater the relative acceptability of the *and*-negation compared to the *or*-negation (i.e., the greater the likelihood of construal as a single complex event), the more acceptable corresponding *wh*-Qs and RCs were judged, with this factor influencing judgments of base sentences significantly less. That is, the predicted interactions are significant (*and_friendly* $\times$ *wh*-Qs: $\beta = .36$, $t = 3.26$, $p = .0016$, $d = .18$; *and_friendly* $\times$ RCs: $\beta = .32$, $t = 3.57$, $p = .0006$, $d = .16$). We also find an overall effect of the same factor ($\beta = .17$, $t = 4.49$, $p < .0001$).

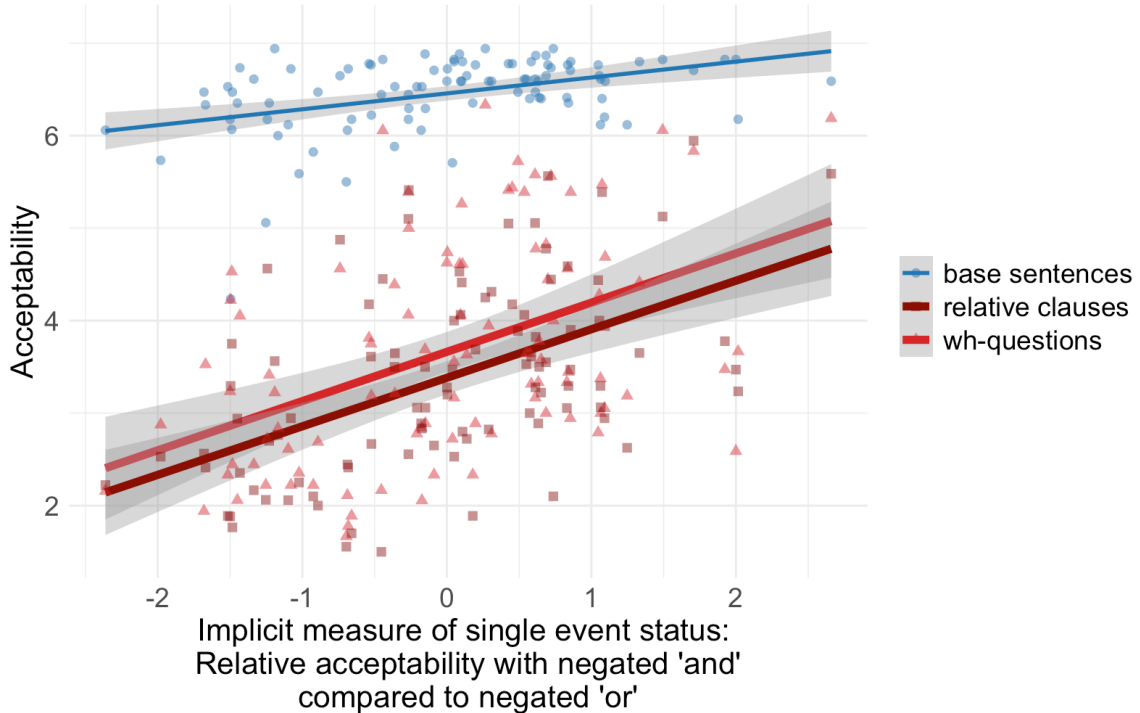


Figure 6: x-axis: Implicit measure of single event status as relative acceptability of negated sentence with ‘and’ compared to negated sentence with ‘or’. Points: Mean acceptability on base

sentences (blue circles) and two types of LDDs *wh*-Qs (brighter red triangles) and RCs (darker red squares).

In exploratory analyses, we next analyze two influences that relate to the ease of processing of LDDs to determine whether the discourse semantic factors demonstrated in Exp 1-4 are reducible to ease of processing.

Probing an Effect of Frequency

As noted in the introduction, phrases that occur with greater log frequency facilitate processing, and have been argued to predict the acceptability of LDDs in particular. To test a potential effect of familiarity, we collected frequency counts of each initial conjunct as part of a verb phrase conjunction, within COCA (Davies, 2009), searching for strings of: <the first conjunct> and VERB. A full thirty conjuncts never occurred in COCA, so we calculated $\log_{1p}$ ($= \log(x + 1)$) in order to include all items in the analysis. As in the other models, linear mixed models were built with acceptability ratings predicted by sentence type (base, *wh*-Q, or RC), now interacting with the centered log frequencies of the first conjunct, with random intercepts for participants and items and random slopes for sentence-type on items. Once again, the reference level was base sentences. Results confirm significant interactions: log frequencies predict acceptability ratings of LDDs more than base sentences ($\log_freq_c \times wh\text{-}Q$: $\beta = .46$, $t = 4.34$, $p < .001$; $\log_freq_c \times RC$: $\beta = .35$, $t = 3.93$, $p = .0002$).

Intransitivity of the First Conjunct

Many examples in the literature used to illustrate acceptable LDDs from verb phrase conjunctions involve an intransitive first conjunct (e.g., *What did Alex go to the store and buy?*). It is possible that intransitive conjuncts facilitate LDDs that target a different conjunct because intransitive phrases are more likely to be backgrounded, serving as scene setters and less likely to push the narrative forward (Hopper & Thompson, 1980; Thompson & Hopper, 2008). Alternatively, intransitive conjuncts may facilitate processing of LDDs in comparison to transitive conjuncts for the following reason. Comprehenders try to resolve LDDs as soon as possible (Crain & Fodor, 1985), and intervening transitive verbs are better candidates for resolving the current LDDs, so they can be expected to interfere more with incremental processing (Stowe 1986; Traxler & Pickering, 1996).

In the current 96 stimuli sets, the first conjunct was intransitive in 36 (37.5%) and transitive in the remaining 60 items. We probed whether intransitivity of the first conjunct influenced the acceptability of LDDs more than base sentences, by treating intransitivity as a centered binary factor. We once again used the same models, now with sentence type interacting with the intransitivity of the first conjunct, and with random slopes of sentence type on items removed for convergence. Results show significant interactions (intransitive $\times wh\text{-}Q$: $\beta = 1.33$, $t = 12.71$, $p < .00001$, $d = .76$; intransitive $\times RC$: $\beta = 1.05$, $t = 10.24$, $p < .00001$, $d = .60$).

We note that intransitivity was not confounded by length in our stimuli. The number of words in transitive and intransitive first conjuncts were not significantly different ($M_{\text{transitive}} = 3.05$ (.81) and $M_{\text{intransitive}} = 3.22$ (1.05); Mann-Whitney U: $W = 928.5$, $p = .217$). Nor did the number of characters approach statistical significance: $M_{\text{transitive}} = 17.72$ (3.86) and $M_{\text{intransitive}} = 17.94$ (5.94); Mann-Whitney U: $W = 1037.5$, $p = .75$).

Robustness Check: Controlling for the Effects of Frequency and Intransitivity

The significant interactions of sentence type with both log frequency and intransitivity of the first conjunct raise an important question: are these processing factors sufficient to explain the gradient differences in ratings observed? That is, are the discourse-semantic factors identified in the first four experiments reducible to ease of processing? We can rule out this possibility in the case of Experiment 2, since identical *wh*-Qs were compared, while the preceding context sentence varied in whether it included increased prominence. That is, intransitivity and frequency vary across items, but the effect reported in Experiment 2 considers changes in ratings on the same items in different contexts.

To test whether the effects found in the other experiments (discourse prominence in Exp. 1, and single event status found in Exps. 3 and 4) can be reduced to a combination of processing factors, we built models that include three interactions: the measure collected in each experiment interacting with sentence type (*wh*-Q, RC, and base), intransitivity interacting with sentence type, and log frequencies interacting with sentence type. Results reveal that all interactions of the theoretically-motivated factors measured here remain significant after controlling for interactions with frequency and intransitivity² (main idea $\times$ *wh*-Q: $\beta = .94$, $t = 3.20$, $p = .002$, $d = .49$; main idea $\times$ RC: $\beta = 1.01$, $t = 4.2$, $p < .0001$, $d = .52$; single event, explicit $\times$ *wh*-Q: $\beta = .32$, $t = 3.51$, $p < .001$, $d = .17$; single event, explicit $\times$ RC: $\beta = .31$, $t = 4.06$, $p = .0001$, $d = .16$; single event, implicit $\times$ *wh*-Q: $\beta = .23$, $t = 2.41$, $p = .018$, $d = .11$; single event, implicit $\times$ RC: $\beta = .23$, $t = 2.85$, $p = .005$, $d = .11$).

In these models, log frequency's interaction with sentence type reached significance in the following cases: main event $\times$ *wh*-Q ($\beta = .20$, $t = 2.15$, $p = .034$), single event, implicit $\times$ *wh*-Q: ($\beta = .21$, $t = 2.07$, $p = .041$); Interactions showed marginal effects in: main event $\times$ RC ($\beta = .13$, $t = 1.68$, $p = .097$) and single event, explicit $\times$ *wh*-Q ($\beta = .19$, $t = 1.96$, $p = .053$) as well. The intransitivity of the first conjunct showed a significant interaction in all models and comparisons (β s = .86 - 1.15, p s < .001), with $d = .45$ - .61. See SI. Thus, results confirm that processing factors—familiarity and reduced interference—play a role in addition to the discourse-prominence and event-structure effects identified in Experiments 1-4.

Correlations Among all Measures

Table 2 reports Pearson's correlations between all measures collected. Mean acceptability ratings on *wh*-Qs and RCs correlated highly with each other: $r = .75$, $t(94) =$

² The models were: $\text{response} \sim \langle \text{factor_of_interest} \rangle * \text{BorLDD} + \log_freq_c * \text{BorLDD} + \text{first_intrans_c} * \text{BorLDD} + (1|\text{participant}) + (1+\text{BorLDD}|\text{item})$.

11.1, $p < .0001$, 95% CI:[.65, .83], despite being collected from separate groups of participants and despite unavoidable differences between the two sets of sentences. This is consistent with the CCC (and BCI) constraint, with its claim that all LDDs make the targeted constituent prominent. It is not consistent with a different discourse proposal, the Focused Backgrounded Constraint (Abeillé et al. 2020: 10; Winckel et al. 2025), which we return to in the general discussion. As shown in Table 1, the acceptability of *wh*-Qs and RC sentences also correlates with the acceptability of base sentences ($r = .32 - .35$), but far less strongly than the two types of LDD constructions correlate with each other ($r = .75$).

Table 2: Pearson correlations among measures. All correlations $> .30$ are significant at $p < .001$; Relevant correlations in bold.

	<i>wh</i> -Q	RC	Base	#1: 2 nd conj. more prominent	#3: single event construal (explicit)	#4: single event construal (implicit)	<u>log_freq</u> <first conjunct> and VERB	<u>Intrans</u> 1 st conjunct
Acceptability								
<i>wh</i> -Q acceptability	1							
RC acceptability	.75	1						
Base acceptability	.32	.35	1					
#1: 2 nd conjunct is more prominent	.46	.55	.34	1				
#3: More likely construed as single event (explicit)	.53	.62	.49	.72	1			
#4: More likely construed as single event (implicit)	.45	.51	.41	.67	.74	1		
<u>Log_freq</u> <first conjunct> and VERB	.45	.41	.18	.22	.24	.27	1	
Intransitive 1 st conjunct	.54	.47	.04	.13	.08	.13	.34	1

Experiments 3 and 4 collected explicit and implicit measures of whether conjunctions are construed as a single event, respectively, which were intended to capture the same construct. We therefore expect them to correlate strongly with each other, and they do ($r = .74$, $t(94) = 10.63$, $p < .0001$, CI= [0.63, 0.82]).

Also important is the strong correlation between the proportion of participants who considered the second conjunct more important (Experiment 1) and the tendency to construe the conjunction as expressing two distinct events, whether judged explicitly in Experiment 3 ($r = .72$) or implicitly, in Experiment 4 ($r = .67$). These correlations suggest that the discourse effect and the semantic effect of construal as distinct events are closely related constructs. We speculate that when verb phrase conjunction is construed as a single complex event, its semantic structure is typically asymmetric: one subevent (often the first) serves an enabling or path-setting role, while the other expresses the culminating outcome. This asymmetry may

itself render the second conjunct more prominent, such that single complex event construal partially mediates the prominence effect.

Finally, the two factors related to ease of processing, frequency and intransitivity of the first conjunct correlate with one another ($r = .34$), more highly than either correlates with the discourse-semantic factors measured experimentally.

General Discussion

The current results demonstrate a role for both discourse semantic and processing factors. The four experiments provide converging evidence that a major aspect of the Coordinate Structure Constraint, namely long distance dependencies (LDDs) resist targeting a single conjunct, is not a primitive syntactic rule but a consequence of a functional clash between the symmetric prominence structure of logical conjunction and the prominence-marking function of LDD constructions.

Logical conjunction links two phrases of equal prominence. The current work establishes and clarifies that it is generally infelicitous for a speaker to combine a construction construed to treat the conjoined phrases as equally prominent, with an LDD construction that treats only one of them as prominent. But combined constructions must be compatible (i.e., the CCC constraint). LDD constructions are acceptable to the extent that the prominence they assign to the targeted constituent is compatible with the prominence structure imposed by the base construction.

Each experiment demonstrated that the factor mentioned influenced acceptability of LDDs more than the base sentences. Experiments 1 and 2 tested whether LDDs were rated more acceptable to the extent that the targeted conjunct was more prominent in the discourse. Experiment 1 collected independent judgments of whether the targeted conjunct expressed the main idea of the conjunction. Results confirmed that both information questions (*wh*-Qs) and relative clauses (RCs) were judged more acceptable when the targeted conjunct was interpreted as expressing the main event in the corresponding base sentence, judged by an independent group of participants. Experiment 2 *manipulated* prominence by varying whether a preceding context sentence, which was presented before the information question (*wh*-Q) was rated, a) emphasized the queried phrase b) emphasized a non-queried phrase c) did not add emphasis. Results demonstrate that, compared to the neutral condition, increasing prominence of a phrase just before a *wh*-Q increased the acceptability of the *wh*-Q while increasing prominence elsewhere reduced the acceptability of the same *wh*-Q.

Experiments 3 and 4 tested the related hypothesis that event-structural asymmetry, the means by which conjunction can escape logical symmetry, predicts LDD acceptability. Experiment 3 asked participants explicitly whether the conjunction in a base sentence described a single complex event or two distinct events, while Experiment 4 probed the same construct implicitly. Both measures of single complex event status predicted ratings on *wh*-Qs

and relative clauses, roughly twice as strongly as they predicted judgments on the base sentences themselves.

There was a strong correlation between discourse prominence of the second conjunct (probed in Experiments 1 and 2) and construal of a single complex event (probed in Experiments 3 and 4). This is consistent with the idea that the two factors may converge on a single construct: LDDs from a single conjunct are acceptable to the extent that the conjunction’s semantics is already asymmetric in prominence, whether that asymmetry arises from event semantics or from discourse context.

The current work also demonstrates an effect of ease of processing in two ways. First, questions and relative clauses that target the second of two conjuncts are more acceptable when the first conjunct is more familiar. Specifically, the acceptability of LDDs increases as the log frequency of the first conjunct phrase in large corpora increased. Secondly, the acceptability of questions and relative clauses that target the second of two conjuncts increased when interference was reduced: i.e., when the first conjunct was intransitive. These findings are consistent with prior work that has found an effect of frequency on the acceptability of LDDs (Kothari, 2008; Dąbrowska, 2013; Liu et al., 2022; Cuneo & Goldberg, 2023), and work that has found LDDs are easier to process when there is less interference (Gibson, 2000; Kluender, 2004; Kluender & Kutas, 1993; Ness & Meltzer-Asscher, 2017).

The current results allow us to compare the CCC/BCI with a different discourse proposal, the Focus Background Constraint (FBC) which states: A focused element may not occur within a backgrounded constituent (Abeillé et al., 2020; Winckel et al., 2025). The FBC shares much with the BCI hypothesis, as both rely on a shared understanding of backgroundedness. When the LDD makes a constituent a focus, as *wh*-questions do, both accounts make the same predictions. Because the head noun of a relative clause need not be a focus, the FBC does not predict any constraints on relative clauses, so current evidence that *wh*-Qs and relative clauses pattern alike poses a challenge (see Cartner, et al. 2026; Cuneo & Goldberg, 2023). The CCC/BCI proposal acknowledges that the type of prominence varies construction by construction but claims that relative clauses and *wh*-questions both position a constituent *prominently* within their domain. To summarize, the current results add to evidence that *wh*-questions and relative clauses pattern alike, which is expected if the proposal is stated in terms of prominence, but is unexpected under an account formulated more narrowly in terms of focus.

Limitations and Future work

The current four experiments tested two types of LDDs when the second of two conjuncts was queried (*wh*-Qs) or modified by a predication (relative clauses). Future work needs to probe LDDs that contain more than two conjuncts and collect judgments from LDDs that target initial or intermediate conjuncts (Lakoff 1986; for discussion see Altshuler & Truswell, 2022). Lakoff (1986) had already observed that LDDs from final conjuncts involve

construal as a single event but suggested that LDDs from initial conjuncts involve violations of expectations. We observe that a recurring type of violation of a complex event may itself be a type of complex event. In just this context, Goldberg (1998) observed that certain individual verbs, which presumably encode single (possibly complex) events, express familiar types of violations of expectations (e.g., *stiff* [a waiter], *renege*, *escape*). Thus, we speculate that events that violate expectations may at least in some circumstances, be just another type of complex single event.

Future work needs to test LDDs from single conjuncts in languages other than English and in LDD constructions other than information questions and relative clauses in order to investigate both the nature of conjunctions and the nature of LDDs more broadly.

Conclusion

To express new messages, language users need to combine constructions. Results from four studies demonstrate that the event structure and information structure of verb phrase conjunctions influence the gradient (un)acceptability of long-distance dependency constructions that target a single conjunct. Moreover, results confirm that ease of processing impacts the acceptability of LDDs in addition to, not instead of, the discourse-semantic factors.

It is infelicitous for a speaker to choose a symmetric base construction such as logical conjunction if they intend to make only an element from a single conjunct prominent. On the other hand, when a verb phrase conjunction is construed to express a single complex event, particularly with one subevent being more prominent, querying an aspect of the prominent subevent correspondingly improves. The larger lesson drawn from the current work is that each construction in a language requires careful analysis of its function(s) to understand its interactions with other constructions. At the same time, two quite distinct LDD constructions, *wh*-Qs and relative clauses are found to pattern remarkably alike, indicating that despite their distinct functions, each makes certain targeted content prominent.

That is, long-distance dependencies exist, despite a general preference for semantically related content to be nearby in the string, because they mark a phrase as particularly prominent within their domain both formally and functionally. The Backgrounded Constructions are Islands (BCI) constraint entails that content targeted by LDDs may not be backgrounded in the corresponding base construction. This remains valid, but it is recognized to be an instance of a far more general constraint, namely that, combined constructions must be compatible (CCC).

The CCC claim covers constraints on LDDs from conjunctions as well as other cases. LDDs, which make a targeted constituent especially prominent, cannot combine with a base construction that has conflicting prominence relations. Backgrounded content in a base construction is incompatible with being made especially prominent by an LDD (the Backgrounded Constructions are Islands, BCI constraint). In addition, base constructions that make conjuncts symmetrically prominent are also incompatible with LDD constructions that make a single conjunct especially prominent; and conversely, LDDs are possible to the extent

that the prominence-marking function of the LDD aligns with asymmetric prominence in the base construction. Coordination, long thought to require its own primitive constraint, turns out to simply confirm the need for combined constructions to have compatible prominence relations. This follows from a very simple and broad generalization: *Combined constructions must be compatible*.

Acknowledgements

We thank Arielle Belluck and Nicole Cuneo for contributions to pilot work on this topic published as Fergus, Belluck, Cuneo, Goldberg (2025). We thank Rui Chaves for helpful comments on an earlier draft. Finally, we thank Princeton University for funding.

Notes

¹ Cuneo & Goldberg (2023) find evidence for the BCI discourse effect to be introduced below *and* an effect of frequencies on clausal verb complements. Noting that the two factors are correlated with one another, they suggest that higher frequency verbs are more likely to background the content of their complement clauses.

References

- Abeillé, A., Hemforth, B., Winckel, E., & Gibson, E. (2020). Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition*, 204, 104293.
- Altshuler, D., & Truswell, R. (2022). *Coordination and the syntax–discourse interface*. Oxford University Press.
- Ariel, M. (2012). Relational and independent and conjunctions. *Lingua* 122:1682-1715.
- Ariel, M. (2023). The long experimental road from semantic meaning to pragmatic truth-compatible inferences. AmPrA Workshop.
- Ariel, M., & Mauri, C. (2019). An “alternative” core for *or*. *Journal of Pragmatics*, 149, 40–59.
- Ambridge, B., & Goldberg, A. E. (2008). The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics*, 19(3), 349–381.
- Boguraev, S., & Mahowald, K. (2026). Causal Drawbridges: Characterizing Gradient Blocking of Syntactic Islands in Transformer LMs. arXiv preprint arXiv:2604.13950.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester, UK: Wiley and Sons.
- Bošković, Ž. (2019). On the coordinate structure constraint and labeling. In *Proceedings of the 36th West Coast Conference on Formal Linguistics*.
- Bresnan, J. W. (1976). On the form and functioning of transformations. *Ling. Inq.* 7(1), 3-40.
- Cartner, M., Kogan, M., Webster, N., Wagers, M., & Sichel, I. (2026). Subject islands do not reduce to construction-specific discourse function. *Cognition*, 271, 106467.
<https://doi.org/10.1016/j.cognition.2026.106467>

- Chaves, R. P., & Putnam, M. T. (2020). *Unbounded dependency constructions: Theoretical and experimental perspectives*. Oxford University Press.
- Chaves, R. P. (2022). Sources of discreteness and gradience in island effects. *Languages*, 7(4), 245.
- Chomsky, Noam (1973/2014). Conditions on Transformations. Reprinted in *An Annotated Syntax Reader: Lasting Insights and Questions*. R. Kayne, R. Zanuttini, T. Leu (Eds). Blackwell Publishing.
- Crain, S., & Fodor, J. D. (1985). How can grammars help parsers. *Natural language parsing: Psychological, computational, and theoretical perspectives*, 94-128.
- Cuneo, N., & Goldberg, A. E. (2023). The discourse functions of grammatical constructions explain an enduring syntactic puzzle. *Cognition*, 240, 105563.
- Dąbrowska, E. (2013). Functional constraints, usage, and mental grammars: An explanation of speakers' intuitions about questions with long-distance dependencies. *Cognitive Linguistics*, 24(4), 633–665.
- 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. D. (1992). *Grammar in mind and brain: Explorations in cognitive syntax* (Vol. 2). Walter de Gruyter.
- Erteschik-Shir, N. (1973). *On the nature of island constraints* (Doctoral dissertation, Massachusetts Institute of Technology).
- Fergus, A., Belluck, A., Cuneo, N., & Goldberg, A. (2025). Islands result from clash of functions: Single-conjunct wh-questions. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 47).
- Fox, D., & Lasnik, H. (2003). Successive-cyclic movement and island repair: The difference between sluicing and VP-ellipsis. *Linguistic Inquiry*, 34(1), 143–154.
- Futrell, R. L. J. (2017). *Memory and locality in natural language* (Doctoral dissertation, Massachusetts Institute of Technology).
- Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). MIT Press.
- Goldberg, A. E. (1998). Semantic principles of predication. In J.-P. Koenig (Ed.), *Discourse and cognition: Bridging the gap* (pp. 41–55). CSLI Publications.
- 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 J. Sprouse (Ed.), *Syntactic islands* (pp. 221–238). Oxford University Press.
- Goldberg, A., Cuneo, N., & Fergus, A. (2024). Addressing a challenge to the backgroundedness account of islands. <https://doi.org/10.31234/osf.io/hmc9n>

- Goldsmith, J. A. (1985). A principled exception to the coordinate structure constraint. In *CLS 21: Papers from the 21st Regional Meeting of the Chicago Linguistic Society* (pp. 133–143).
- Hahn, M., & Xu, Y. (2022). Crosslinguistic word order variation reflects evolutionary pressures of dependency and information locality. *Proceedings of the National Academy of Sciences*, 119(24), e2122604119.
- Harris, J. A. (2009). Extraction from coordinate structures: Evidence from language processing. In *Proceedings of the Chicago Linguistic Society* (Vol. 45, No. 2, pp. 73–88).
- Hawkins, J. (2014). *Cross-linguistic variation and efficiency*. Oxford University Press.
- Hofmeister, P., & Sag, I. A. (2010). Cognitive constraints and island effects. *Language*, 86(2), 366–415.
- Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 56, 251–299.
- Huang, N., Almeida, D., & Sprouse, J. (2025). A nearly exhaustive experimental investigation of bridge effects in English. *Language*, 101(1), 72–108.
- Jennings, R.E. (1994). *The Genealogy of Disjunction*. New York: Oxford University Press.
- Kothari, A. (2008). Frequency-based expectations and context influence bridge quality. In M. Grosvald & D. Soares (Eds.), *Proceedings of WECOL 2008* (pp. 136–149). California State University.
- Kluender, R. (2004). Are subject islands subject to a processing account? In *Proceedings of WCCFL 23* (pp. 475–499). Cascadilla Press.
- Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, 5(2), 196–214.
- Kuno, S. (1987). *Functional syntax: Anaphora, discourse, and empathy*. University of Chicago Press.
- Lakoff, G. (1986). Frame semantic control of the coordinate structure constraint. In *Chicago Linguistic Society Parasession on Pragmatics and Grammatical Theory*.
- Liu, Y., Ryskin, R., Futrell, R., & Gibson, E. (2022). A verb-frame frequency account of constraints on long-distance dependencies in English. *Cognition*, 222, 104902.
- Lu, J., Pan, D., & Degen, J. (2025). Evidence for a discourse account of manner-of-speaking islands. *Language*, 101(4), 627–659.
- Na, Y., & Huck, G. (1992). On extracting from asymmetrical structures. In *The joy of grammar: A festschrift in honor of James D. McCawley* (pp. 251–274).
- Namboodiripad, S., Cuneo, N., Kramer, M. A., Sedarous, Y., Bisnath, F., Sugimoto, Y., & Goldberg, A. E. (2022). Backgroundedness predicts island status of non-finite adjuncts in English. In *Proceedings of the Cognitive Science Society*.
- Ness, T., & Meltzer-Asscher, A. (2017). Working memory in the processing of long-distance dependencies: Interference and filler maintenance. *Journal of Psycholinguistic Research*, 46(6), 1353–1365.

- Oda, H. (2017). Two types of the coordinate structure constraint and rescue by PF deletion. In *Proceedings of NELS 47* (pp. 343–356).
- Rayner, K & Clifton, C. Jr. (2009). Language processing in reading and speech perception is fast and incremental: implications for event-related potential research. *Biological psychology* 80(1), 4–9.
- Ross, J. R. (1967). *Constraints on variables in syntax* (Doctoral dissertation, MIT).
- Richter, S. N., & Chaves, R. P. (2020). Investigating the role of verb frequency in factive and manner-of-speaking islands. In *Proceedings of the Cognitive Science Society* (Vol. 42).
- Sabel, J. (2002). A minimalist analysis of syntactic islands. *Linguistic Review*, 19(3).
- Sag, I. A. (1976). Deletion and logical form. PhD thesis, Massachusetts Institute of Technology
- Schmerling, S. F. (1975). Asymmetric conjunction and rules of conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 211–232). Academic Press.
- Sprouse, J., & Hornstein, N. (2013). *Experimental syntax and island effects: Toward a comprehensive theory of islands*. Cambridge University Press.
- Stjepanović, S. (2014). Left branch extraction and the coordinate structure constraint. In *Proceedings of NELS 44* (pp. 157–170).
- Stowe, L. (1986). Evidence for on-line gap location. *Language and Cognitive Processes*.
- Thompson, S. A., & Hopper, P. J. (2008). Transitivity, clause structure, and argument structure. In *Frequency and the emergence of linguistic structure* (pp. 27–60). John Benjamins.
- Traxler, M. J. & Pickering M. J. (1996). Plausibility and the processing of unbounded dependencies: an eye-tracking study. *Journal of Memory and Language*.
- Truswell, R. (2006). *Locality of wh-movement and the individuation of events* (Doctoral dissertation, University College London).
- Winckel, E., Abeillé, A., Hemforth, B., & Gibson, E. (2025). Discourse-based constraints on long-distance dependencies generalize across constructions in English and French. *Cognition*, 254, 105950.