

Definite DP island effects across dependency types

Zheng Shen and Meghan Lim

June 2024

Abstract

The source of island constraints in filler-gap dependency is under debate. Syntactic analyses of island effects propose that island constructions are opaque to all types of filler-gap dependency regardless of their discourse functions. On the other hand, a recent discourse-based analysis, the Focus-Background Conflict (FBC) constraint proposed in Abeillé et al. 2020, argues that some island effects result from a conflict between the discourse status of the island construction and the discourse function of certain filler-gap dependencies. Abeillé et al. (2020) explicitly argue that island effects under the FBC constraint are predicted to only show up in filler-gap dependencies involving focus, e.g. *wh*-questions; and should be absent in constructions like relative clauses where a dependency is established but its discourse function does not involve focus.

This paper uses the definite DP island in English as a case study to evaluate these distinct predictions. Abeillé et al. (2020), citing Grosu (1981), claims that the definite DP island effect only holds of *wh*-questions and not relative clauses, thus is also accounted for by the FBC constraint, whereas its effect in a relative clause was observed since Ross (1967). This paper teases apart

the contradictory empirical claims with experimental methods. We show that both movement in wh-questions and movement in relative clauses show the definite island effect, challenging Abeillé et al. (2020)'s approach to locality restrictions.

1 Theories on the sources of island effects

Filler-gap dependencies are constrained by locality restrictions, commonly known as island effects. Sentences in (1) showcase the subject island effect where wh-movement out of the subject in (1a) is unacceptable whereas movement out of the object in (1b) is acceptable.

- (1) Subject island (Chomsky 1973)
- a. *Who did [stories about ___] terrify John?
 - b. Who did you hear [stories about ___]?

Although an abundance of generalizations on island effects across languages have been claimed over the past 60 years, the source of these effects continues to be debated. Analyses under the structural approach to islandhood, for example, propose that island-inducing constructions (e.g. subjects in (1)) are structurally opaque to filler-gap dependencies. This general approach predicts that movement across islands is blocked regardless of the discourse or semantic function of the movement. For specific analyses of island effects that take such a structural approach, see Chomsky (1973); Rizzi (1990); Huang (1982); Müller (2011) among many others.

The recent discourse-based approach proposed by Abeillé et al. (2020, 2022), on the other hand, relates island effects to discourse properties of the construction and movement. According

to this approach, the island effect arises when the discourse status of the island construction is at odds with *the discourse function of the movement*. Specifically, the authors propose the constraint in (2) to account for the subject island effect in (1) as well as adjunct island effects:

- (2) Focus-background conflict (FBC) constraint: A focused element should not be part of a backgrounded constituent.

Abeillé et al. (2020) claim that the subject of a sentence is usually backgrounded while the wh-element in wh-questions is focused. Thus the FBC constraint predicts a penalty in sentences with wh-movement out of the subject as is shown in (3a). On the other hand, movement that does not involve a focused element would not violate the FBC constraint. Relative clauses in English, for example, involve movement, but the extracted element is not focused. The FBC constraint thus predicts that relative clauses with movement from the subject is acceptable shown in (3b). Experimental data from English and French in Abeillé et al. 2020 confirm the contrast in (3).

- (3) a. *Of which sportscar did the color ___delight the baseball player because of its surprising luminance?
b. The dealer sold a sportscar, of which the color ___delighted the baseball player because of its surprising luminance.

This contrast in (3) is unexpected under various approaches to islands, including the structural approach mentioned above, as well as other discourse-based analyses. The dominance analysis by Erteschik-Shir and Lappin (1979) and Erteschik-Shir (1981) and the backgroundedness analysis by Goldberg (2006) and Cuneo and Goldberg (2023), for example, argue that island effects result from

the backgroundedness of the island construction (the subject in (3)). Nevertheless, both analyses predict island effects in relative clauses and *wh*-questions alike, as they assume that the restriction applies to movement in general regardless of its discourse function. As far as we are aware, Abeillé et al. (2020)'s FBC analysis is unique in predicting the contrast between *wh*-questions and relative clauses. In other words, contrasts like (3) pose a challenge to almost all other analyses of island effects, structural or otherwise.¹

Against this background, it is important to investigate how the FBC constraint fares with other island effects. This paper evaluates the FBC constraint with data from the definite DP island effect, another island effect that has been argued to fall under the FBC analysis (Abeillé et al. 2020). Our contribution is two-fold: 1. we provide novel experimental evidence that the definite DP island affects *wh*-questions and relative clauses alike; 2. we argue that the findings pose a challenge to the FBC analysis of the definite DP island.

2 Definite DP island effect

The definite DP island effect, observed since Ross (1967) as well as Chomsky (1973, 1977), refers to the contrast in (4): movement out of definite DP objects is degraded (4b), as opposed to movement out of indefinite objects (4a).

(4) Abeillé et al. 2020, ex. 39a,b

¹The debate over the subject island and the adjunct island is not the focus of this paper. See McInnerney and Sugimoto (2022) for arguments that the contrast in (3) results from a confound and is compatible with non-FBC analyses, as well as evidence against the FBC constraint. See Šimík et al. (2022) for evidence against the FBC analysis for adjunct island effects. See Cuneo and Goldberg (2023) for experimental evidence against the FBC analysis for adjunct island effects among others.

- a. Which actress did you buy [a picture of ___]?
- b. #Which actress did you buy [that picture of ___]?

Abeillé et al. (2020) explicitly claim that the definite DP island ‘follows naturally from the FBC constraint’ (p20), citing the contrast between (4) and (5) from Grosu (1981). According to these judgments, although interrogative wh-movement out of the definite object is unacceptable, relativisation out of the definite object is acceptable. The FBC analysis for this island effect would be similar to the one for the subject island: the definite object introduces backgrounded information while the wh-element conveys focus, so sentences like (4b) would violate the FBC constraint. Indefinite objects, on the other hand, introduce a new entity, and are thus compatible with the wh-element as in (4a). When the movement does not involve focus, e.g. in relativisation, the FBC constraint is not applicable regardless of the definiteness status of the object, so movement out of both types of objects is OK shown in (5).

- (5) Abeillé et al. 2020, ex. 39c
- a. That is the actress who I bought [a picture of ___].
 - b. That is the actress who I bought [that picture of ___].

However, the original observation of the definite DP island effect by Ross (1967) claims that this effect is present in relative clauses as well as wh-questions. Sentences in (6) show the island effect in relative clauses.

- (6) Ross 1967, 224

- a. The man who I gave John a picture of ___ was bald.
- b. ??The man who I gave John this picture of ___ was bald.

Based on the contrast in (4) and (6), structural analyses propose that definite objects are structurally different from indefinite objects. For example, Davies and Dubinsky (2003), following Bowers (1987), propose that indefinite objects are NPs while definite objects are DPs. NP objects are not blocking categories while DP objects are, thus movement out of definite DPs induces a subjacency violation. In more contemporary terms, Jiménez-Fernández (2012) proposes that indefinite objects are not a phase while definite objects are. Movement out of a definite object thus violates the Phase Impenetrability Condition (Chomsky 2000) while movement out of indefinite ones does not. Beyond these syntactic analyses, previous semantic or discourse-based analyses of the definite DP island effect also predict the effect to be present in relative clauses. Fiengo and Higginbotham (1981) propose a semantic constraint, the Specificity Condition, for the definite DP island effect, arguing that operator-variable binding relations cannot cross a definite DP boundary. The discourse-based dominance analysis by Erteschik-Shir and Lappin (1979) argues that elements within a dominant NP cannot be extracted, and definite objects are dominant, thus inducing the island effect.

The debate between the FBC approach and other analyses of the definite DP island effect thus boils down to an empirical question. The former does not predict an island effect in relative clauses while the latter predicts an island effect. Although the definite DP island effect has been experimentally verified in a few studies (Keller 2000; Neal and Dillon 2021; Tollan and Heller 2015; Shen and Lim 2022; Huang 2022; Shen and Huang 2023), all of these works consider only interrogative wh-movement. To our knowledge, there is yet to be any previous experimental studies

on the definite DP island effect in relative clauses or other types of movement.²

In the following sections, we report four experiments testing the definite DP island effect in *wh*-questions and relative clauses, addressing both the empirical and the theoretical debate. To preview the results, the definite DP island effect is observed in both *wh*-questions (Experiment 1) and relative clauses (Experiment 2) alike, with no contrast found between the movement types.

In addition to P-stranding configurations, we also tested PP movement in the context of the definite DP island as shown in (7) (Experiment 3-4). The inclusion of PP movement is motivated by the experiments in Abeillé et al. (2020). They demonstrate that while PP movement out of subjects shows a contrast between *wh*-questions and relative clauses, P-stranding movement out of subjects is unacceptable regardless the construction. Our experiments show the definite DP island effect in both P-stranding and PP movement alike, contrary to the predictions of the FBC analysis (more details in Section 3.3).

(7) a. **About which country** did you see [a/that movie ____].

b. That is the nation **about which** I saw [a/that movie ____].

²Cuneo and Goldberg (2023) tested restrictions including Adjunct Island, Complex NP island, and factive island on both *wh*-questions and relative clauses and found the island effects in both constructions, challenging the FBC approach. The logic of the experiments in Cuneo and Goldberg (2023) is similar to this paper in that restrictions in *wh*-questions and relative clauses are compared, however, except for the adjunct island, the islands tested by Cuneo and Goldberg (2023) are not explicitly claimed by Abeillé et al. (2020) to result from the FBC constraint. The definite DP island, on the other hand, is claimed by Abeillé et al. (2020) to be subsumed under the FBC constraint.

3 Experiments

3.1 Experiment 1: *wh*-movement out of definite DPs

Experiment 1 compares the acceptability of movement out of the indefinite objects and definite objects in *wh*-questions. Given that all previously discussed theories of the definite DP island effect predict an island effect in *wh*-questions, Experiment 1 simply establishes a baseline for what the definite DP island effect would look like in our the experimental design.

3.1.1 Design and materials

Following previous experimental work on island effects (Sprouse et al. 2016 among others), the 2x2 factorial design manipulates two independent factors. First, NP TYPE: the definiteness of the object: (indefinite vs. definite). Second, DEPENDENCY LENGTH: whether the sentence involves movement of the subject or movement out of the object (short vs. long). This gives us four conditions in (8): (i) a sentence with an indefinite object but no movement out of it (8a), (ii) a sentence with a definite object but no movement out of it (8b), (iii) a sentence with an indefinite object and movement out of it (8c), and (iv) a sentence with a definite object and movement out of it (8d). The gap (___), indicating the original position of the *wh*-element, is for illustrative purposes only and was not included in the stimuli.

- (8) a. I wonder which journalist ___ watched a movie about Wonder Woman. (IND.SHORT)
b. I wonder which journalist ___ watched that movie about Wonder Woman. (DEF.SHORT)
c. I wonder which superhero Amy watched a movie about _____. (IND.LONG)
d. I wonder which superhero Amy watched that movie about _____. (DEF.LONG) = island

The 2x2 factorial design allows us to isolate the effects of DEPENDENCY LENGTH, NP TYPE, and any effect specifically due to the extraction out of the definite object. The effect of DEPENDENCY LENGTH will be indicated by the difference between (8a) and (8c). The effect of NP TYPE will be the difference between (8a) and (8b). For there to be an island effect, the difference between (8c) and (8d) should be larger than the combined effects of NP TYPE and DEPENDENCY LENGTH. The island effect can thus be quantified with a Difference in Difference (DD) score: $((8b)-(8d))-((8a)-(8c))$. A positive DD score reflects a super-additive effect which would indicate that a definite DP is an island. A DD score of zero would indicate the absence of the penalty of moving out of a definite DP. Statistically, a significant interaction between the two factors would indicate the super-additive effect.

In the experiments, the definite objects are headed by the demonstrative *that* and the indefinite objects are headed by the indefinite article *a/an*. The demonstrative *that* was chosen to remain faithful to the examples given in Abeillé et al. (2020) cited from Grosu (1981). In addition, the definite article *the* in English has been argued to be ambiguous between a strong and weak determiner (see e.g. Schwarz 2009), which would potentially affect the locality restriction (Simonenko 2016). The use of demonstrative DPs sidesteps this confound. Following the sentences in Abeillé et al. (2020), the *wh*-elements are complex/D-linked, of the shape *which* + noun. Compared to bare *wh*-elements *what/who*, D-linked/complex *wh*-elements are easier to process (Hofmeister and Sag 2010; Sprouse et al. 2016) thus using them in the experiment let us avoid processing difficulties as a potential confound. Moreover, using complex *wh*-elements helps us control for the effects of lexical complexity between experiments on *wh*-questions and relative clauses, as relative clauses necessarily involve a complex filler, e.g. *the superhero who* in *I recognized the superhero who*

*Amy watched a movie about.*³ See Sprouse et al. (2016) Section 6 for a discussion of the effects on complex wh-elements. The head nouns of the objects are picture nouns as most studies on NP extraction focus on this type of noun. In choosing the main verbs, we avoided verbs of creation like *write (a book)* or *take (a picture)*, as verbs of creation have been observed to obviate island effects of definite NPs. Replacing *watch* in (8d) with *film* in (9) makes extraction acceptable despite the definite NP. See Davies and Dubinsky (2003) and Huang (2022) for discussion.

(9) I wonder which superhero Amy **filmed** that movie about.

All the target items are embedded questions under ‘I wonder’ or ‘I know’ with an equal distribution between the two. Using embedded questions keeps the complexity and length of the items in Experiment 1 similar to the target items for relative clauses in Experiment 2, as relative clauses are always embedded inside a clause, shown in (10).

(10) I recognized [the journalist who ___ watched a movie about Wonder Woman.]

As P-stranding is the default method of extraction out of these NPs, the long movement conditions always leave the preposition (*about* in (8)) behind, only extracting the DP argument (e.g. *which superhero* in (8c)-(8d)). We investigate Pied Piping/PP movement (whereby the preposition is extracted alongside the DP) as in (11) in Experiments 3 and 4.

(11) **About which** country did Amy watch that movie ___?

16 lexical variants for each condition were constructed. Each participant saw 16 target items, 4

³We thank a reviewer for bring this point to our attention.

items for each condition, and 32 filler items and 5 attention check items.

Each participant also saw 1 attention check item (12a), 2 acceptability check items (12b-c), and 2 short answer questions. A native speaker vented their answers (12d-e).

- (12)
- a. This sentence is an attention check, so please do not rate its acceptability and instead just click number 4.
 - b. The children are forming a circle on the playground.
 - c. The children is forming a circle on the playground.
 - d. Your boss wants to give you a raise. But it turns out that it is for a project where a colleague did most of the work. What would you do? Please provide a short answer with full sentences (max. 2 lines).
 - e. Your best friend is getting married in one week. You just saw your friend's partner with someone else at a restaurant and you suspect he/she is having an affair. How would you behave? Please provide a short answer in full sentences (max. 2 lines).

Responses from participants who failed the attention check item (12a) or rated the grammatical (12b) equal to or lower than the ungrammatical (12c) were removed from the analysis. Naturalness of the answers to the short answer questions (12d-e) was evaluated by a native speaker of English (a co-author) to decide whether the participants were likely to be a native speaker. Responses from participants were removed from the analysis if their answers had major grammatical errors, or sounded awkward to the native ear. This procedure was used for all four experiments and it was done before the rest of the experimental items were analyzed.

3.1.2 Procedure

The experiment used an acceptability judgement task whereby participants were asked to evaluate how well-formed the sentences are on a 7-point Likert scale. 1 was labeled as “completely unacceptable” which participants were told corresponds to a sentence that would never naturally occur in that form, while 7 was labeled “completely acceptable” where there was absolutely nothing wrong about the form of the sentence. Participants were instructed to follow their intuitions, rather than prescriptivist rules they might have encountered. The experiment took less than 15 minutes to complete. The experiment was hosted on Prolific (Zehr and Schwarz 2018.)

3.1.3 Participants

72 self-reporting native speakers of American English completed the experiment, recruited via Prolific. They were paid 2GBP for their participation. Responses from 6 participants were removed due to failing the attention checks or the short answer questions. Responses from 66 participants were included in the analysis

3.1.4 Predictions

As mentioned above, all existing theories including the structural approach and the FBC constraint predict the definite DP island effect to appear in the LONG+DEFINITE condition in wh-questions (8d).

3.1.5 Results

Mean raw ratings for all the experiments presented in this paper were z-score transformed based on each participants’ ratings of the test and filler items. Z-score transformation standardises otherwise variant scale usage by different participants. The average z-score and mean raw ratings are shown

in Table 1 and plotted in Figure 1. The DD score based on the z-scores $((\text{DEF.SHORT} - \text{DEF.LONG}) - (\text{IND.SHORT} - \text{IND.LONG}))$ is 0.34.

NP TYPE	DEPENDENCY LENGTH	z-score mean	raw judgment mean
indefinite	short	0.85	6.27
indefinite	long	0.45	5.49
definite	short	0.94	6.41
definite	long	0.20	4.99

Table 1: Mean ratings from Experiment 1, n = 66

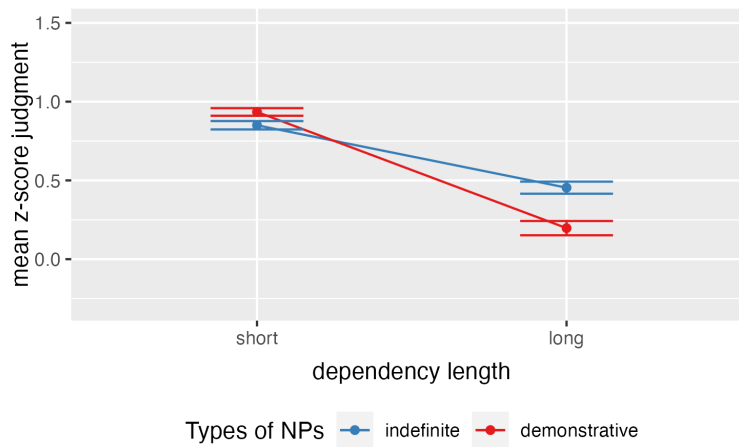


Figure 1: condition means of Experiment 1, n = 66

The distribution of the raw judgments of each condition is shown in Figure 2. Judgments of all four conditions were normally distributed, indicating that the participants formed a homogeneous group regarding their judgments.

Z-scored acceptability ratings were modeled using linear mixed effects models with the R packages lme4 (Bates et al. 2015) and lmerTest (Kuznetsova et al. 2017). NP TYPE and DEPENDENCY LENGTH were specified as fixed effects and lexical items and subjects were specified as random effects.⁴ For all four experiments in this paper, we applied treatment coding for both factors (in-

⁴The model for Exp 1: $z\text{-scores} \sim \text{NP Type} * \text{Dependency Length} + (1 + \text{NP Type} * \text{Dependency Length} | \text{subject}) + (1 + \text{NP Type} * \text{Dependency Length} | \text{Lexical})$. The variable ‘Lexical’ refers to the lexical variant which is coded as the item numbers in the experiment.

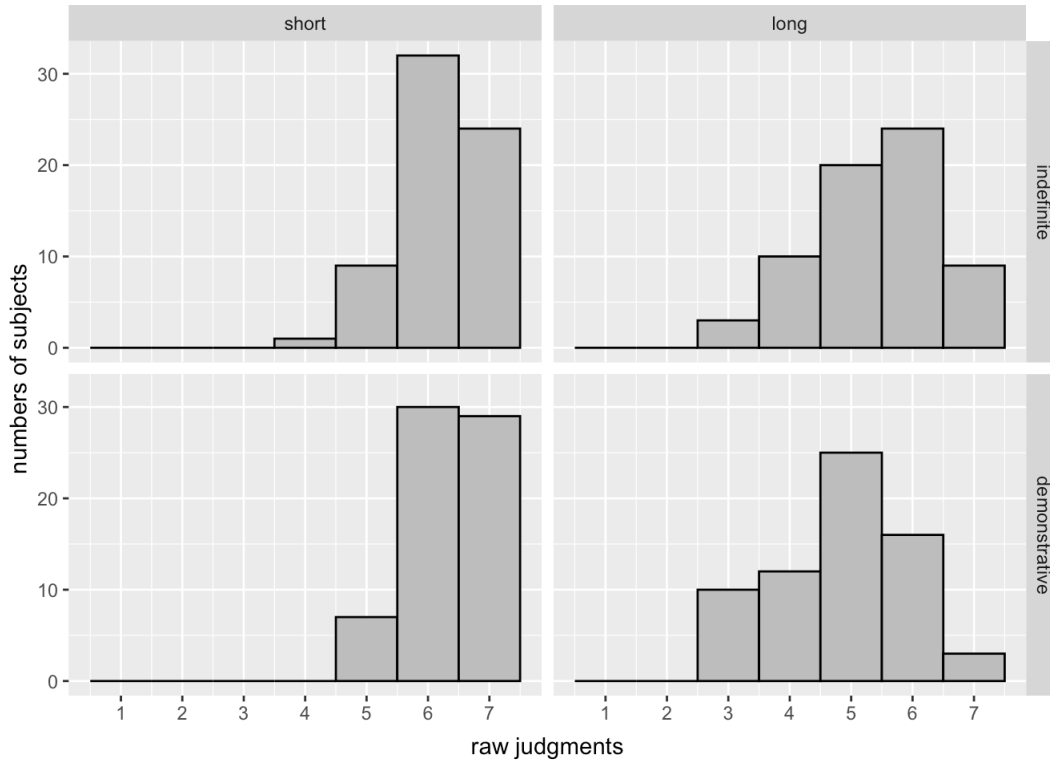


Figure 2: judgment distribution of Experiment 1, n = 66

definite: 0, demonstrative: 1; short: 0, demonstrative: 1).

The model revealed a marginal effect of NP TYPE ($\beta = 0.09$, $t = 1.81$, $p = .08$) and a main effect of DEPENDENCY LENGTH ($\beta = -0.40$, $t = -4.85$, $p < .0001$). Crucially there is a significant interaction effect between the two factors ($\beta = -0.34$, $t = -3.44$, $p = .0024$), suggesting a super-additive effect.

3.1.6 Discussion

The observed interaction between the factors indicates that the DEFINITE.LONG condition with movement out of a definite object is penalized beyond the combination of effects of DEPENDENCY LENGTH and NP TYPE. In other words, the definite DP island effect is observed for wh-questions.

We also constructed Cumulative Link Mixed Effects models to model the raw judgments. The model converged is $\text{judgment} \sim \text{NP Type} * \text{Dependency Length} + (1 + \text{Dependency Length} \mid \text{subject}) + (1 + \text{Dependency Length} \mid \text{lexical})$. The model revealed significant main effects of NP Type ($\beta = 0.52$, $\text{SE} = 0.19$, $z = 2.76$, $p < 0.01$) and Dependency Length ($\beta = -1.71$, $\text{SE} = 0.27$, $z = -6.31$, $p < 0.0001$) as well as a significant interaction between the two ($\beta = -1.24$, $\text{SE} = 0.25$, $z = -4.90$, $p < 0.0001$).

The results of Experiment 1 not only confirm that this experiment design can detect the definite DP island effect, but also provide a DD score (0.34) as a baseline of the *size* of this island effect. The DD score of 0.34 observed in Experiment 1 is similar to previous experiments (0.28 in Shen and Lim (2022), 0.22-0.25 in Neal and Dillon (2021)). Although there is a super-additivity effect found between the two factors, indicating an island effect, the mean rating of the DEFINITE.LONG condition is above the middle of the 7 point scale and the positive z-score indicates that this condition is rated above the mean of the test items and filler items. In other words, while violating the definite DP island did induce a penalty, the island violating sentences were not completely unacceptable. This property is constant in all our experiments and consistent with other studies into definite DP island mentioned above. The reason for the moderate size of island effect is unclear at this point. We are not aware of any analysis that predicts the size of definite DP island violation to be more acceptable than other islands.⁵ We want to point out that in the formulation of Abeillé et al. (2020), the FBC constraint predicts the presence and absence of the penalties involved in moving out of a definite DP, not the size of such penalties. Thus we will leave this issue for future research.

3.2 Experiment 2: Definite DP island in relative clauses

3.2.1 Design and materials

Experiment 2 used the same design as Experiment 1, with NP TYPE and DEPENDENCY LENGTH as the factors, but in relative clauses instead of embedded wh-questions. An example of the four conditions is in (13). The relative clauses modify the matrix object in each item. Half of the items

⁵A reviewer asks if the relatively high ratings of the condition could be due to the inclusion of very clearly unacceptable fillers. The mean rating of ungrammatical fillers was 2.94 for Experiment 1. None of the fillers were rated at the bottom of the scale, with the lowest one rated as 1.98. Thus we do not think the filler played a crucial role in this pattern.

use *I recognized* and the other half use *I met* as the main clause subject and verb.

- (13) a. I recognized the journalist who ___ watched a movie about Wonder Woman. IND.SHORT
b. I recognized the journalist who ___ watched that movie about Wonder Woman. DEF.SHORT
c. I recognized the superhero who Amy watched a movie about _____. IND.LONG
d. I recognized the superhero who Amy watched that movie about _____. DEF.LONG

3.2.2 Predictions

The FBC analysis predicts that there would be no interaction between NP TYPE and DEPENDENCY LENGTH. Relativisation out of a definite DP should not show any super-additive effects. Unlike interrogative wh-movement, relativisation does not involve focus, and non-focus movement out of a backgrounded constituent (the definite NP) should not trigger island effects. Other theories of the definite DP island, including the syntactic approach and the dominance/backgroundedness approach, predict a super-additive effect in relative clauses as specified in Section 2.

3.2.3 Procedure

The procedure of Experiment 2 was identical to Experiment 1.

3.2.4 Participants

72 participants were recruited over Prolific. All participants were self-reported to be native speakers of American English. There was no overlap of participants between Experiments 1 and 2. Participants were all paid 2GBP. Responses from 8 participants were removed due to failing the attention checks or the short answer questions. Responses from 64 participants were included in

the analysis

3.2.5 Results

The average z-score and mean raw ratings of Experiment 2 are shown in Table 2 and plotted in Figure 3. The DD score for the z-scores $((\text{DEF.SHORT} - \text{DEF.LONG}) - (\text{IND.SHORT} - \text{IND.LONG}))$ is 0.24.

NP TYPE	DEPENDENCY LENGTH	z-score mean	raw judgment mean
indefinite	short	0.79	5.86
indefinite	long	0.26	4.84
definite	short	0.87	6.01
definite	long	0.10	4.52

Table 2: Mean ratings from Experiment 2, n = 64

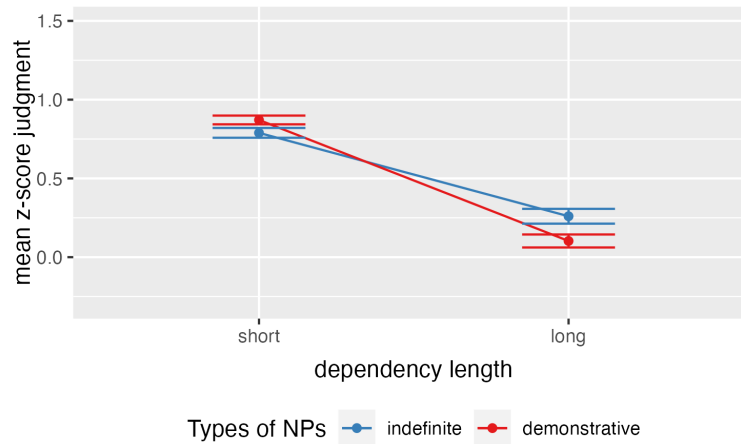


Figure 3: condition means of Experiment 2, n = 64

The distribution of the raw judgments of each condition is shown in Figure 4. Judgments of all four conditions were normally distributed, indicating that the participants formed a homogeneous group regarding their judgments.

We constructed linear mixed effect models to predict z-scores, starting with the maximal models and with random slopes removed until convergence. The linear mixed effect model did not reveal a main effect of NP TYPE ($\beta = 0.07$, $t = 1.25$ $p = .22$) but there is a main effect of DEPENDENCY

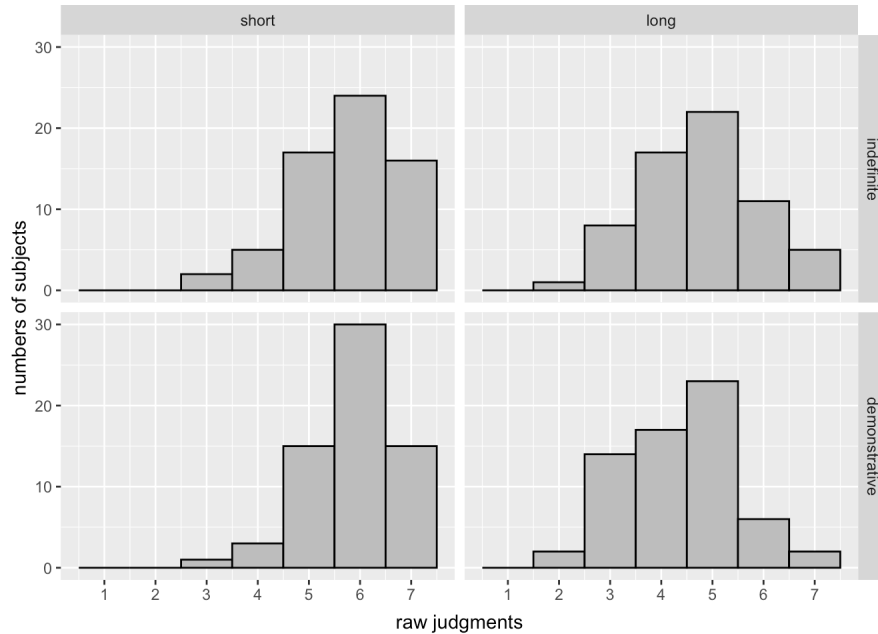


Figure 4: judgment distribution of Experiment 2, $n = 64$

LENGTH ($\beta = -0.53$, $t = -5.91$, $p < .0001$).⁶ Crucially, we still see an interaction effect between the two factors ($\beta = -0.26$, $t = -3.15$, $p < .001$), indicating that the super-additive effect persists for relativisation.

3.2.6 Discussion

The results from Experiment 2 show that the definite DP island effect is active even in relative clauses where the moved element is not a focus. Empirically, the results are at odds with the informal observations made in Grosu (1981), used in Abeillé et al. (2020). The original observation of a definite DP island effect affecting both *wh*-questions and relative clauses by Ross (1967) is verified.

⁶The linear mixed effects model: $z\text{-scores} \sim \text{NP Type} * \text{Dependency Length} + (1 + \text{NP Type} * \text{Dependency Length} \mid \text{subject}) + (1 + \text{NP Type} * \text{Dependency Length} \mid \text{lexical})$.

We also constructed Cumulative Link Mixed Effects models to model the raw judgments. The model converged is $\text{judgment} \sim \text{NP Type} * \text{Dependency Length} + (1 + \text{Dependency Length} \mid \text{subject}) + (1 + \text{Dependency Length} \mid \text{lexical})$. The model revealed a significant main effect of Dependency Length ($\beta = -1.85$, $\text{SE} = 0.33$, $z = -5.72$, $p < 0.0001$), but no main effect of NP Type ($\beta = 0.23$, $\text{SE} = 0.27$, $z = 0.86$, $p = 0.39$). The model also revealed a significant interaction between the two ($\beta = -0.95$, $\text{SE} = 0.29$, $z = -3.28$, $p < 0.01$.)

Regarding theories of the definite DP island, the results are surprising for the FBC analysis, as it predicts the absence of the island effect in relative clauses. On the other hand, the results are predicted by other theories of the definite DP island, which predict no difference in islandhood between movements with different discourse functions. Our results thus argue against the application of the FBC constraint to definite DP island effects.

Given the presence of the island effect in relative clauses, one could potentially argue for a weaker version of the FBC analysis, namely that the FBC constraint does exist and induces the definite DP island effect in wh-questions, but there is an additional mechanism that gives rise to the definite DP island effect in both wh-questions and relative clauses. The candidates for such an additional mechanism include the ones proposed in the syntactic analyses (PIC, Barriers, etc), or the semantics/pragmatic analyses, as long as this analysis applies to both wh-questions and relative clauses. Such a hybrid account (FBC + Source 2) would predict the island effect in both relative clauses and wh-questions. In wh-questions, the island effect is driven by both the FBC constraint and the second mechanism. In relative clauses, on the other hand, the island effect results only from the second mechanism as the FBC constraint does not apply. Under this analysis, the size of the island effect is predicted to be smaller in relative clauses than in wh-questions.

We evaluate this possibility by combining the data from Experiments 1 and 2. We ran a post hoc analysis with CONSTRUCTION TYPE (relative clause vs. wh-question) as the third factor, in addition to NP TYPE and DEPENDENCE LENGTH. If the size of the definite island is different between the two constructions, we would expect a significant three way interaction effect between NP TYPE, DEPENDENCY LENGTH, and CONSTRUCTION TYPE. We applied treatment coding for all three factors (indefinite: 0, demonstrative: 1; short: 0, demonstrative: 1; relative clause: 0, wh-question: 1)

The model with the three fixed factors shows a main effect of DEPENDENCY LENGTH ($\beta = -0.53$, t

= -6.22, $p < 0.0001$), but no main effect of NP TYPE ($\beta = 0.07$, $t = 1.51$, $p = 0.13$), or CONSTRUCTION TYPE ($\beta = 0.05$, $t = 1.12$, $p = 0.26$).⁷ Additionally, we see a two way interaction effect between NP TYPE and DEPENDENCY LENGTH ($\beta = -0.26$, $t = -2.95$, $p < .01$), but no interaction between NP TYPE and CONSTRUCTION TYPE ($\beta = 0.01$, $t = 0.23$, $p = 0.82$), or DEPENDENCY LENGTH and CONSTRUCTION TYPE ($\beta = -0.13$, $t = 1.60$, $p = .11$). Crucially, we do not see a significant three-way interaction ($\beta = -0.08$, $t = -0.76$, $p = 0.45$), which indicates that CONSTRUCTION TYPE does not significantly affect the size of the definite island effect.⁸ The weaker hybrid analysis is not supported.⁹

Based on experimental data, no evidence of the FBC constraint driving the definite DP island can be found. This paper thus joins previous work including McInnerney and Sugimoto (2022), Šimík et al. (2022), and Cuneo and Goldberg (2023) in arguing against the FBC constraint as a locality restriction.

3.3 PP movement and the definite DP island: Experiments 3 & 4

Sentences involving movement out of objects (the LONG conditions) in Experiments 1 and 2 feature P-stranding where only the DP is moved, stranding the preposition, shown in (14).

⁷The model for the post hoc analysis: $z\text{-scores} \sim \text{NP Type} * \text{Dependency Length} * \text{Construction Type} + (1 + \text{NP Type} * \text{Dependency Length} | \text{subject}) + (1 + \text{NP Type} * \text{Dependency Length} | \text{lexical})$

⁸We also constructed a cumulative link mixed model for raw judgments: $\text{judgment} \sim \text{NP Type} * \text{Dependence Length} * \text{Construction} + (1 + \text{NP Type} * \text{Dependence Length} | \text{subject}) + (1 + \text{NP Type} * \text{Dependence Length} | \text{lexical})$. The model revealed a main effect of Dependency Length ($\beta = -1.80$, $SE = 0.30$, $z = -6.00$, $p < 0.0001$), construction ($\beta = 0.93$, $SE = 0.35$, $z = 2.68$, $p < 0.01$) but no main effect of NP Type ($\beta = 0.25$, $SE = 0.21$, $z = 1.176$, $p = 0.24$). Two way interactions were found between NP Type and Dependency Length ($\beta = -0.93$, $SE = 0.31$, $z = -3.03$, $p < 0.05$). No interaction between NP Type and Construction ($\beta = 0.31$, $SE = 0.26$, $z = 1.19$, $p = 0.24$) or between Dependency Length and Construction ($\beta = 0.08$, $SE = 0.32$, $z = 0.23$, $p = 0.82$) were found. No three way interaction was found ($\beta = -0.42$, $SE = 0.37$, $z = -1.14$, $p = 0.25$).

⁹A reviewer suggests conducting an experiment with both types of dependencies so that the island effects can be directly compared within subjects, dividing them into two experiments like we did. While including both dependencies in a single experiment would remove the need to run a post-hoc comparison between experiments, it would double the number of sentences that participants see to 96 sentences. We wanted to avoid the possible fatigue effect and opted for between subjects design. Moreover, we wanted to stay close to the experiments reported in Abeillé et al. (2020), which also allocated each dependency type to its own experiment.

- (14) a. I wonder **which superhero** Amy watched a/that movie **about** ____.
- b. I recognized the superhero **who** Amy watched a/that movie **about** ____.

Experiments 3 and 4 look into PP movement out of objects and test if a contrast in island effect between wh-questions and relative clauses can be observed in PP movement, shown in (15).

- (15) a. **About which country** did Amy watch a/that movie __yesterday?
- b. I remember the country **about which** Amy watched a/that movie ____.

Almost all the empirical claims about definite DP islands in English (or extraction out of DPs in general) are built on sentences with P-stranding, as it is the more natural option for English speakers. PP movement out of DPs, on the other hand, is less common and used in a higher register, and is therefore studied less in works on locality constraints. Note that the empirical claim used as evidence for the FBC analysis for definite DP islands in Abeillé et al. (2020) (cited from Grosu 1981) also involves P-stranding and not PP movement (5).

However, it is PP movement out of the subject that provided evidence for FBC's relevance in subject islands in Abeillé et al. 2020. The claim that subject island effects are only present in wh-questions and not in relative clauses is based on materials with PP movement in (16). The subject island effect is observed as in (16a) and not in (16b).

- (16) PP movement
- a. Of which sportscar did [the color ____] delight the baseball player because of its surprising luminance? wh-question (Abeillé et al., 2020, ex. 24a)
- b. The dealer sold a sportscar, of which [the color ____] delighted the baseball player

because of its surprising luminance. relative clause (Abeillé et al., 2020, ex. 16a)

Sentences with P-stranding as in (17), on the other hand, showed the subject island effect regardless of the construction type. (Abeillé et al., 2020, p10), citing (Chaves, 2013; Chaves and Dery, 2019, p483), claim that P-stranding differs from PP movement because “the grammatical function of the fronted phrase PP is clearer from the onset than if NP were fronted, given the presence of the preposition: there are fewer potential gap sites that are consistent with the extracted constituent, aiding processing and improving acceptability”. In other words, they suggest that sentences with P-stranding might suffer from a processing difficulty that gives the illusion of a subject island effect while sentences with PP movement do not.

(17) P-stranding

- a. Which sportscar did [the color of ___] delight the baseball player because of its surprising luminance? wh-question (Abeillé et al., 2020, ex. 24b)
- b. The dealer sold a sportscar, which [the color of ___] delighted the baseball player because of its surprising luminance. relative clause (Abeillé et al., 2020, ex. 21a)

Evaluating the source of the difference between PP movement and P-stranding is beyond the scope of this paper. To control for the potential effect of P-stranding claimed by Abeillé et al. (2020), we conducted Experiments 3 and 4 to probe the definite DP island effect in wh-questions and relative clauses with PP movement. This is, to our knowledge, the first experimental study of the definite DP island effect with PP movement. To preview the results, Experiments 3 and 4 show that the definite DP island effect is observed in both wh-questions and relative clauses, even with

PP movement. Since Experiments 3 and 4 are identical in design and procedure, we present them together.

3.3.1 Design and materials

Experiments 3 and 4 use the same factorial design as Experiments 1 and 2, with NP TYPE (definite v. indefinite) and DEPENDENCY LENGTH (short v. long) as the two fixed factors. The conditions in both experiments are illustrated in (18) and (19).

(18) Conditions of Experiment 3

- a. Which journalist ___ watched a movie about Finland yesterday? (IND.SHORT)
- b. Which journalist ___ watched that movie about Finland yesterday? (DEF.SHORT)
- c. About which country did Amy watch a movie ___ yesterday? (IND.LONG)
- d. About which country did Amy watch that movie ___ yesterday? (DEF.LONG)

(19) Conditions of Experiment 4

- a. I remember the journalist who ___ watched a movie about Finland. (IND.SHORT)
- b. I remember the journalist who ___ watched that movie about Finland. (DEF.SHORT)
- c. I remember the country about which Amy watched a movie _____. (IND.LONG)
- d. I remember the country about which Amy watched that movie _____. (DEF.LONG)

There are several modifications in the materials from those in Experiments 1 and 2. First, the moved PPs in the LONG conditions have inanimate DP arguments inside them, e.g ‘about **Finland**’ in (18) and (19), as opposed to ‘about **Wonder Woman**’ in Experiment 1 and 2. An animate DP argument would introduce uncertainty as to whether ‘about who’ or ‘about whom’ would be more

appropriate; while both are acceptable, the latter is often dispreferred by some. Using inanimate DP arguments avoids this issue. The moved DPs in the *SHORT* conditions remain animate given the animate preference for subjects in general. Second, as a result, the main verbs in Experiment 4 have been changed to ‘I remember’ and ‘I know’ which are compatible with both animate and inanimate objects, from ‘I met’ and ‘I recognised’ in Experiments 1 and 2. Third, in Experiment 3, the conditions use matrix wh-questions as opposed to the embedded wh-questions in Experiment 1. We used embedded wh-questions in Experiment 1 to balance the complexity of conditions in Experiment 1 and 2 since relative clauses are always embedded. We made the decision to switch to matrix questions in Experiment 3 to be more similar to the items in Abeillé et al. (2020).

Like previous experiments, 16 lexical combinations were created for each condition. Each participant saw 4 items for each condition, thus 16 test items in total. In addition, each list included 32 filler items and 5 attention check items.

3.3.2 Procedure

The procedure of Experiments 3 and 4 was identical to Experiments 1 and 2.

3.3.3 Participants

There were 72 participants in Experiments 3 and 4, respectively. They were recruited via Prolific and paid GBP2 for their participation. All the participants were self-reported native speakers of American English. After filtering with the attention checks or the short answer questions, response from 64 participants were included in the analysis for Experiment 3 and responses from 63 participants were included for Experiment 4.

3.3.4 Predictions

The predictions for Experiments 3 and 4 are identical to the ones for Experiments 1 and 2 for all theories. The FBC analysis of the definite DP island predicts an island effect in wh-questions but not relative clauses, since the former involves focus movement and the latter does not. Other theories, including the syntactic analyses (PIC/barrier) and other discourse-based analyses (dominance/backgroundedness) predict the island effect to be present in both constructions.

3.3.5 Results

The mean ratings of z-scores and raw judgments for each condition are listed in Tables 3 and 4 and are plotted in Figures 5 and 6. The DD scores $((\text{DEF.SHORT} - \text{DEF.LONG}) - (\text{IND.SHORT} - \text{IND.LONG}))$ are 0.21 for Experiment 3 and 0.2 for Experiment 4.

NP TYPE	DEPENDENCY LENGTH	z-score mean	raw judgment mean
indefinite	short	0.99	6.39
indefinite	long	0.22	4.86
definite	short	0.94	6.32
definite	long	-0.03	4.37

Table 3: Mean ratings from Experiment 3, n = 64

NP TYPE	DEPENDENCY LENGTH	z-score mean	raw judgment mean
indefinite	short	0.92	6.26
indefinite	long	-0.12	4.20
definite	short	0.94	6.29
definite	long	-0.30	3.83

Table 4: Mean ratings from Experiment 4, n = 63

The distributions of the raw judgments of each condition is shown in Figure 7 for Experiment 3 and Figure 8 for Experiment 4. Judgments for the demonstrative long condition in Experiment 3 show a bimodal distribution where a portion of participants chose 3 and another portion chose

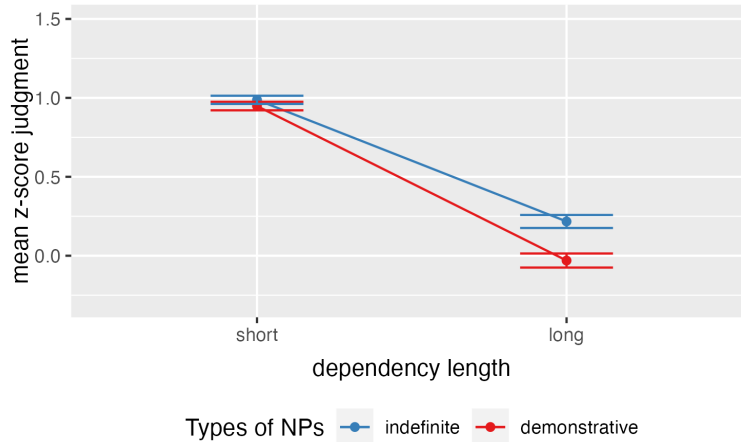


Figure 5: condition means of Experiment 3, n = 64

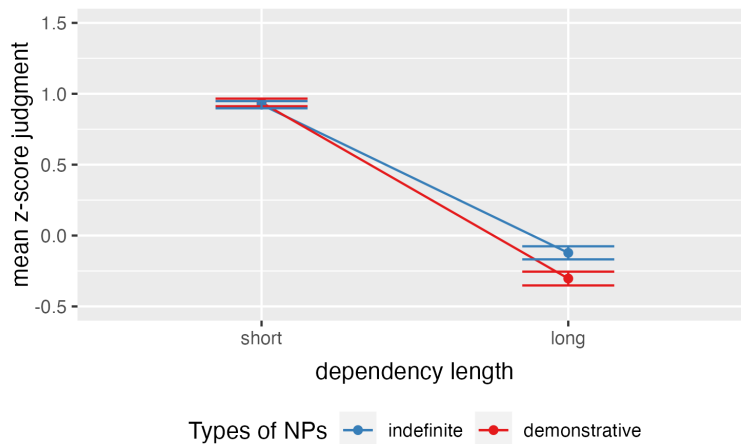


Figure 6: condition means of Experiment 4, n = 63

5. The reason for such distribution is not known to us at this point, though it is worth pointing out that the z-score judgments for this condition (and all other conditions) are normally distributed. Judgments for the other three conditions in Experiment 3 and all four conditions in Experiment 4 were normally distributed, indicating that the participants formed a homogeneous group regarding their judgments.

We constructed linear mixed effect models for the z-scores. In Experiment 3, we do not see a main effect of NP TYPE ($\beta = -0.04$, $t = -0.86$, $p = .39$), but there is a main effect of DEPENDENCY

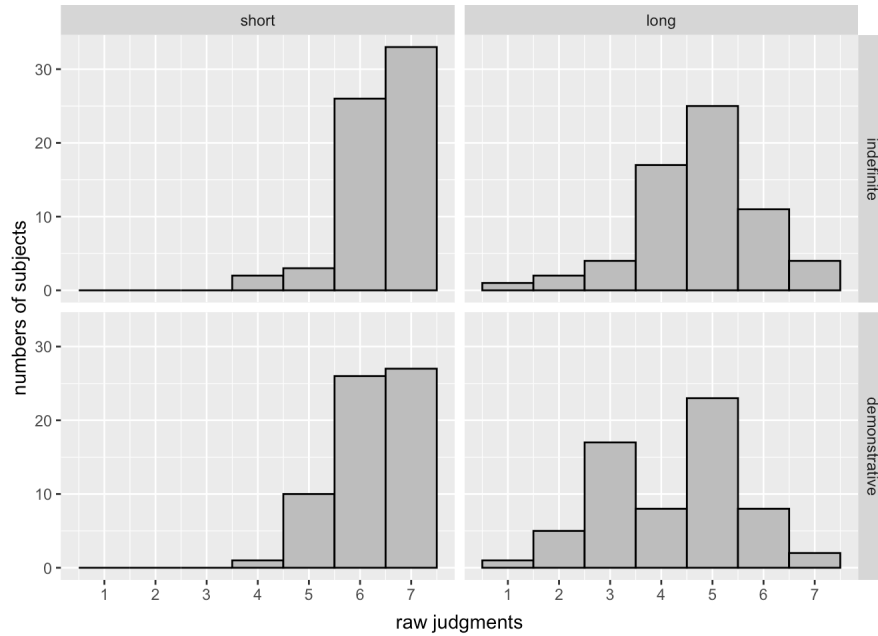


Figure 7: judgment distribution of Experiment 3, n = 64

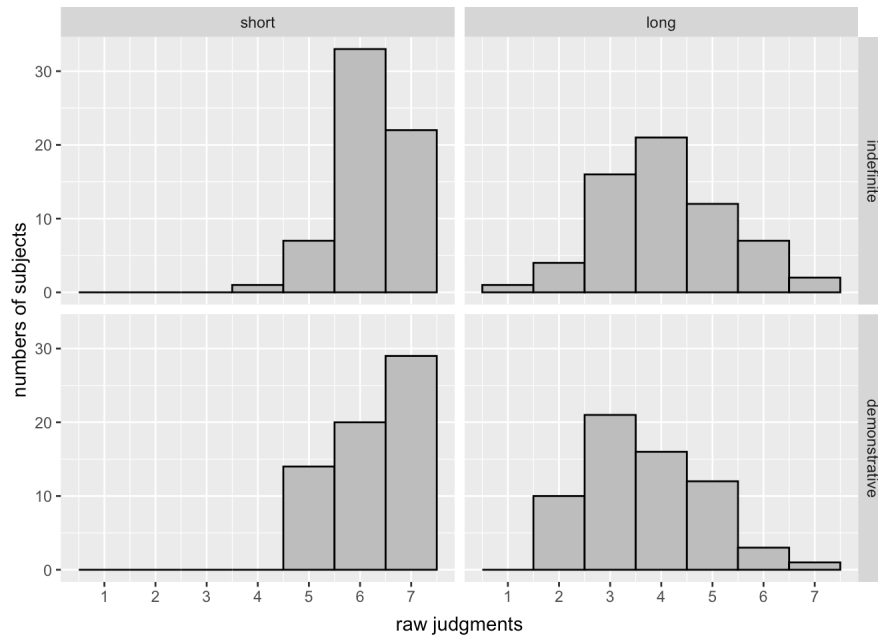


Figure 8: judgment distribution of Experiment 4, n = 63

LENGTH ($\beta = -0.78, t = -10.94, p < .0001$). Crucially, there is a significant interaction between the two factors ($\beta = -0.19, t = -2.57, p = 0.0122$).¹⁰

¹⁰The linear mixed effects model: z-scores \sim NP Type * Dependency Length + (1 + NP Type * Dependency Length | subject) + (1 + Dependency Length | lexical)

We also constructed Cumulative Link Mixed Effects models to model the raw judgments. The model con-

In Experiment 4, there is no main effect of NP TYPE ($\beta = 0.02$, $t = 0.39$, $p = .7$), but there is a main effect of DEPENDENCY LENGTH ($\beta = -1.04$, $t = -11.58$, $p < .0001$). Crucially, there is a significant interaction effect between the two factors ($\beta = -0.2$, $t = -3.04$, $p = 0.002$).¹¹

3.3.6 Discussion

Results from Experiments 3 and 4 show a super-additive effect in both wh-questions and relative clauses, indicating the definite DP island effect in both constructions, even with PP movement. The DD scores were similar (0.21 for wh-questions and 0.2 for relative clauses), indicating that the size of the island effects does not differ either. This goes against the prediction made by the FBC analysis but follows from other theories of definite DP islands.

As with Experiments 1 and 2, we conducted a post hoc analysis using CONSTRUCTION TYPE as the third factor in addition to NP TYPE and DEPENDENCY LENGTH, combining data from Experiments 3 and 4. A weaker version of the FBC analysis would predict a three way interaction, suggesting that the size of the island effect is smaller in relative clauses than in wh-questions as discussed in Section 3.2.6. The model showed no main effects of NP TYPE ($\beta = 0.02$, $t = 0.36$, $p = 0.72$)

verged is judgment \sim NP Type * Dependency Length + (1 + NP Type*Dependency Length | subject) + (1 + NP Type*Dependency Length | lexical). The model revealed a significant main effect of Dependency Length ($\beta = -3.51$, $SE = 0.31$, $z = -11.39$, $p < 0.0001$), but no main effect of NP Type ($\beta = -0.26$, $SE = 0.27$, $z = -0.99$, $p = 0.32$). The model **did not** reveal a significant interaction between the two ($\beta = -0.49$, $SE = 0.33$, $z = -1.48$, $p = 0.14$). This is the only experiment where the cumulative link mixed models on raw judgments show a different pattern from the linear mixed effects models on z-scores. We suspect the absence of the interaction for this CLMM model is partially driven by the bimodal distribution of raw judgments for the demonstrative long condition in Experiment 3 shown in Figure 7. So far, no previous claims have been made that PP movement in wh-questions would not induce an island effect. None of the analyses for the definite DP island effect predict the absence of the definite DP island effect in PP movement in wh-questions, including the FBC constraint analysis, which predicts an island effect as the wh-questions involve focus movement of the PP. Note also that the cumulative link mixed model comparing Experiments 3 and 4 did not find a three way interaction of DP Type, Dependency Length, and Construction, indicating that the island effect does not differ between wh-questions and relative clauses, see Footnote 13. Thus we will put this difference between the cumulative link mixed models and the linear mixed effect models aside for now and take the results from the linear mixed effect models, i.e. PP movement in wh-questions does show the definite DP island effect.

¹¹z-scores \sim NP Type * Dependency Length + (1 + Dependency Length | subject) + (1 + Dependency Length | lexical)

or a marginal effect of CONSTRUCTION TYPE ($\beta = 0.09$, $t = 1.7$, $p = .09$).¹² There is a main effect of DEPENDENCY LENGTH ($\beta = -1.00$, $t = -12.18$, $p < 0.0001$). The two way interaction of NP TYPE and DEPENDENCY LENGTH was significant ($\beta = -0.20$, $t = -3.14$, $p < 0.01$), indicating the definite DP island effect in both constructions. There was also an interaction between DEPENDENCY LENGTH and CONSTRUCTION TYPE ($\beta = 0.22$, $t = 2.32$, $p = 0.02$): movement out of indefinite and definite objects is more degraded in relative clauses than in wh-questions. The interaction between NP TYPE and CONSTRUCTION TYPE was not significant ($\beta = -0.06$, $t = 0.90$, $p = .37$). Crucially, there is no three way interaction of NP TYPE, DEPENDENCY LENGTH, and CONSTRUCTION TYPE ($\beta = 0.015$, $t = 0.17$, $p = 0.89$). Thus the weaker version of the FBC analysis suggested in Section 3.2.6 is not supported as no evidence indicates that the size of the definite DP island effect differs between constructions.¹³

Comparing the same conditions in P-stranding and PP movement, we observe that PP movement out of NP objects is more degraded than sentences with P-stranding, regardless of the definiteness of the objects. As shown in Table 5, in wh-questions (Experiments 1 and 3), both IND.LONG and DEF.LONG conditions with P-stranding are rated 0.23 in z-score higher the same conditions with PP movement; in relative clauses (Experiments 2 and 4), the IND.LONG condition with P-stranding is rated 0.38 higher than the same condition with PP movement and the DEF.LONG condition with P-stranding is rated 0.40 higher than PP movement. These findings confirm that PP movement out of NP object is dispreferred compared to movement with P-stranding. This preference is independent

¹²Z-scores \sim NP Type * Dependency Length * Construction Type + (1 + Dependency Length | subject) + (1 + Dependency Length | lexical)

¹³We also constructed a cumulative link mixed model to model the raw judgments: judgment \sim Np Type * Dependency Length * Construction + (1+Np Type * Dependency Length|subject) + (1+Np Type * Dependency Length|lexical). The model showed a main effect of Dependency Length ($\beta = -3.93$, SE = 0.23, $z = -12.31$, $p < 0.0001$), no main effect of NP Type ($\beta = 0.19$, SE = 0.23, $z = 0.85$, $p = 0.40$), or Construction ($\beta = 0.47$, SE = 0.37, $z = 1.27$, $p = 0.20$). The model also revealed an interaction of NP type and dependency length ($\beta = -0.83$, SE = 0.29, $z = -2.86$, $p < 0.01$). Interaction of dependency length and construction was not significant ($\beta = 0.48$, SE = 0.39, $z = 1.23$, $p = 0.22$), neither was the interaction of NP Type and Construction ($\beta = -0.43$, SE = 0.28, $z = -1.52$, $p = 0.13$). Crucially, like the linear mixed effect models, no three way interaction of NP Type, dependency Length, and Construction were found ($\beta = 0.33$, SE = 0.38, $z = 0.88$, $p = 0.38$).

from the definite DP island effect, as it applies to indefinite objects to the same extent.

Dependency Type	Condition	P-stranding	PP movement	Difference
wh-questions	IND.LONG	0.45	0.22	0.23
	DEF.LONG	0.20	-0.03	0.23
relative clauses	IND.LONG	0.26	-0.12	0.38
	DEF.LONG	0.10	-0.30	0.4

Table 5: Comparing z-scores of LONG conditions between P-stranding and PP movement

4 Concluding remarks

Data from four experiments show that the definite DP island effect affects wh-questions as well as relative clauses for both DP movement and PP movement. Neither the presence nor the size of this island effect is affected by the construction type, disputing the empirical claims reported in Abeillé et al. (2020), following Grosu (1981). Given the distinct discourse functions of interrogative wh-movement and relativisation regarding focus, the findings present a challenge for Abeillé et al. (2020)’s recent FBC constraint that predicts that locality restrictions are sensitive to discourse function of the movement. The patterns observed in our experiments, on the other hand, are compatible with other theories of locality in general and analyses of the definite DP island effect in particular. Syntactic analyses from Bowers (1987), Davies and Dubinsky (2003), and Jiménez-Fernández (2012), the semantic analysis from Fiengo and Higginbotham (1981), the hybrid approach of syntactic and semantic analyses from Shen and Huang (2023), and the dominance analysis from Erteschik-Shir and Lappin (1979) all predict definite DP islands to affect all sorts of constructions, as long as movement out of a definite DP is involved.

While this paper argues against the FBC analysis of the definite DP island in particular, it is possible that the FBC constraint is active in other island constructions, including adjunct island

and subject island as proposed by Abeillé et al. (2020, 2022). However, it is curious why the effect of the FBC constraint is only observed in these islands, but not in definite DP islands. Another possibility is that the FBC constraint is not what causes subject island effects or adjunct island effects in the first place. While pointing out several problems with the FBC analysis, McInnerney and Sugimoto (2022) attribute the results from Abeillé et al. (2020) where *wh*-questions and relative clauses show different sensitivity to the subject island to a confound. They argue that the ‘extracted PPs’ in the relative clauses (20) could be base-generated topics in the left periphery, similar to (21). If this alternative analysis is on the right track, the absence of subject island effect in (20) is not relevant to the discussion of locality constraints as it does not involve movement at all. The differences between P-stranding and PP movement observed by Abeillé et al. (2020) is also accounted for as the former involves movement out of the subject and the latter does not.

(20) The dealer sold a sportscar, of which [the color ___] delighted the baseball player because of its surprising luminance. (repeated from (16b))

(21) Of the stories about Watergate, only yesterday’s was truly surprising. (Akmajian and Lehrer 1976, 401, cited from McInnerney and Sugimoto 2022, ex. 27)

Leaving a full evaluation of the FBC constraint for future research, readers are referred to McInnerney and Sugimoto (2022) and Šimík et al. (2022) for other arguments against the FBC analysis for subject islands and adjunct islands.

Data availability statement

The stimuli, data for the test and filler items, and analyses for the four experiments reported in this paper can be found here: https://osf.io/fwsgn/?view_only=6d141f703b134244a7a4aa3cb40218b1.

References

- Abeillé, Anna, Barbara Hemforth, Elodie Winckel, and Edward Gibson. 2020. Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition* 204.
- Abeillé, Anne, Barbara Hemforth, Ruihua Mao, and Edward Gibson. 2022. Acceptability of extraction out of English adjunct clauses: the role of topicalization. the 35th annual meeting of the Society for Human Sentence Processing, UC Santa Cruz.
- Akmajian, Adrian, and Adrienne Lehrer. 1976. NP-like quantifiers and the problem of determining the head of an NP. *Linguistic Analysis* 2 (4): 395–413.
- Bowers, John. 1987. Extended X-bar theory, the ECP, the Left Branch Condition. In *Proceedings of the sixth West Coast Conference on Formal Linguistics (WCCFL 6)*. Stanford: The Stanford Linguistics Association.
- Chaves, Rui P. 2013. An expectation-based account of subject islands and parasitism. *Journal of Linguistics* 2 (49): 285–327.
- Chaves, Rui P., and Jeruen E. Dery. 2019. Frequency effects in subject islands. *Journal of Linguistics* 55 (3): 475–521.

- Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, eds. Stephen R. Anderson and Paul Kiparsky, 232–286. New York: Holt, Rinehart & Winston.
- Chomsky, Noam. 1977. Conditions on transformations. In *Essays on form and interpretation*, 81–162. New York, New York: Elsevier North-Holland, Inc..
- Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, eds. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. MIT Press.
- Cuneo, Nicole, and Adele E. Goldberg. 2023. The discourse functions of grammatical constructions explain an enduring syntactic puzzle. *Cognition* 240 (105563).
- Davies, William D., and Stanley Dubinsky. 2003. On extraction from NPs. *Natural Language and Linguistic Theory* 21 (1): 1–37.
- Erteschik-Shir, Nomi. 1981. On extraction from noun phrases (picture noun phrases). In *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*, eds. Adriana Belletti, Luciana Brandi, and Luigi Rizzi, 147–169.
- Erteschik-Shir, Nomi, and Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6: 41–86.
- Fiengo, Robert, and James Higginbotham. 1981. Opacity in NP. *Linguistic Analysis* 7 (4): 395–421.
- Goldberg, Adele. 2006. *Constructions at work*. Oxford: Oxford University Press.
- Grosu, Alexander. 1981. *Approaches to island phenomena*. North-Holland.

- Hofmeister, Philip, and Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86: 366–415.
- Huang, C. T. James. 1982. Logical relations in Chinese and the theory of grammar. PhD diss, Massachusetts Institute of Technology.
- Huang, Nick. 2022. How subjects and possessors can obviate phasehood. *Linguistic Inquiry* 53 (3): 427–458.
- Jiménez-Fernández, Ángel L. 2012. A new look at subject islands: The phasehood of definiteness. *Anglica Wratislaviensia* 50: 137–170.
- Keller, Frank. 2000. Experimental and computational aspects of degrees of grammaticality. PhD diss, University of Edinburgh.
- McInnerney, Andrew, and Yushi Sugimoto. 2022. On dissociating adjunct island and subject island effects. In *Proceedings of the Linguistic Society of America*, Vol. 7, 5207. LSA.
- Müller, Gereon. 2011. *Constraints on Displacement. A phase-based approach*. Vol. 7 of *Language faculty and beyond*. Amsterdam: Benjamins.
- Neal, Anissa, and Brian Dillon. 2021. Definitely Islands? An investigation into the offline and online status of definite islands. Poster presented at the annual meeting of Linguistic Society of America 2021.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Massachusetts: MIT Press.
- Ross, John. 1967. Constraints on variables in syntax. PhD diss, Massachusetts Institute of Technology.

- Schwarz, Florian. 2009. Two types of definites in natural language. PhD diss, UMass at Amherst.
- Shen, Zheng, and Nick Huang. 2023. The definiteness effect in wh-fronting and wh-in situ languages. Talk given at the 46th GLOW Colloquium, the University of Vienna and the University of Graz.
- Shen, Zheng, and Meghan Lim. 2022. Extraction from definite, indefinite, and superlative DPs: An experimental approach. In *NELS 52: Proceedings of the 52nd Annual Meeting of the North East Linguistic Society*, eds. Özge Bakay, Breanna Pratley, Eva Neu, and Peyton Deal, Vol. 3, 99–109. GLSA, University of Massachusetts.
- Šimík, Radek, Petr Biskup, Kateřina Bartasová, Markéta Dančová, Eliška Dostálková, Kateřina Hrdinková, Gabriela Kosková, Jaromír Kozák, Klára Lupoměská, Albert Maršík, Edita Schejbalová, and Illia Yekimov. 2022. Extraction from clausal adjuncts in Czech: A rating experiment. In *NELS 52: Proceedings of the 52nd Annual Meeting of the North East Linguistic Society*, eds. Özge Bakay, Breanna Pratley, Eva Neu, and Peyton Deal, Vol. 3. GLSA Publications.
- Simonenko, Alexandra. 2016. Semantics of DP islands: the case of questions. *Journal of Semantics* 33: 661–702.
- Sprouse, Jon, Ivano Caponigro, Ciro Greco, and Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory* 34 (1): 307–344.
- Tollan, Rebecca, and Daphna Heller. 2015. Elvis Presley on an island: wh-dependency formation inside complex NP objects. In *Proceedings of the forty-sixth annual meeting of the North*

East Linguistic Society, eds. Christopher Hammerly and Brandon Prickett. GLSA, University of Massachusetts.

Zehr, Jeremy, and Florian Schwarz. 2018. PennController for Internet Based Experiments (IBEX).
<https://doi.org/10.17605/OSF.IO/MD832>.