Information structure alone cannot account for subject islandhood: an experimental study

Nikolas Webster[1/*], Mandy Cartner[2], Matthew Kogan[1], Matt Wagers[1], Ivy Sichel[1]
[1]UCSC, [2]TAU; [*]newebste@ucsc.edu

Introduction: Recent experimental and theoretical work calls into question the traditional syntactic view of subjects as strong islands: is there a truly syntactic component to islandhood, which cannot be reduced to pragmatic or semantic factors (Abeillé et al 2020; Winckel et al to appear; Goldberg 2006; Cuneo & Goldberg 2023)? We argue yes: when controlling for the independent costs of DP complexity and dependency length, we observe degraded acceptability of sub-extraction from subjects vs. objects across multiple construction types, each with different information structure (IS) characteristics. Three large scale acceptability judgment studies investigating three constructions (topicalization (TOP), wh-questions (WHQ), relative clauses (RC)) that differ in their IS profile, utilizing a superadditive design (Sprouse 2007; Sprouse et al. 2012), show evidence for a subject island effect, despite differences in the different constructions' IS that are predicted to modulate the presence of a subject island violation under a purely discourse-based account of islandhood. **Background:** The Subject Condition (Ross 1967; Chomsky 1973; Pesetsky 1982; Huang 1982), asserts that constituents within a syntactic subject cannot be targeted for movement. There have been different formulations of the Subject Condition; the strongest version commits to the generalization that all specifiers are opaque for extraction (Privoznov 2021). In both (1-2) below, the same wh-word moves out of the same complex DP, so it is not possible to attribute the contrast in grammaticality of (1-2) to the extractee, nor to anything about the internal makeup of the DP as an extraction domain. Under a syntactic lens, the source must be configurational/structural: in (1), a wh-word is sub-extracted out of a DP in subject position (SpecTP), whereas in (2) it is sub-extracted out of a DP in the direct object position.

(1) *This is the person [who [[a friend of _] invited Rosa to the party]].
(2) This is the person [who [Rosa invited [a friend of _] to the party]].

A genealogy of research (Erteschik-Shir 1973; Kuno 1987; Goldberg & Ambridge 2006; among others) challenges the claim that the source of the (un)acceptability of (1) vs. (2) is syntactic. Based on findings that PP sub-extraction is rated less acceptable out of subjects vs. objects in WHQ constructions, but not in RCs, Abeillé et al (2020) propose that unacceptable sub-extraction out of a subject arises from a "clash" in IS, formalized as the Focus Background Constraint (FBC): "a focused element should not be part of a backgrounded constituent." This analysis hinges critically on the assumption that subjects are typically backgrounded (i.e. discourse familiar, presupposed, unfocused), and objects/post-verbal constituents are part of the focus. WHQ constructions place an element in a focused position, a position associated with the introduction of discourse-new material, whereas RCs specify or attribute a property of the RC head. Subject sub-extraction is therefore unacceptable for WHQs, but not RCs, because the extractee is focused in a WHQ construction, which conflicts with its direct relation to a backgrounded extraction domain (i.e. a subject). **Present Work:** A central prediction of the FBC is that the acceptability of a sub-extraction from a subject should differ across constructions with different IS characteristics, despite the fact that backgroundness, or more concretely, presuppositionality— the 'givenness' of a nominal referent—is known to vary by multiple factors (e.g. definiteness, specificity, and existentiality; for English, see Milsark 1974, et seq; Diesing 1992), but does not directly vary by syntactic position (e.g. subject v. non-subject), as Abeillé et al (2020) assume. Given these doubts about the mapping of subjects to backgrounded and objects to non-backgrounded, we investigate the FBC by probing the acceptability of subject

sub-extraction across three constructions (TOP, WHQ, RC). The FBC predicts no island effect for topicalization out of subjects: in terms of the FBC, topicalization places part of a backgrounded constituent into a non-focused position, and thus does not create the IS 'clash' that underpins subject island effects. TOP constructions are crucially distinct from WHQs in that the extracted element in a TOP construction is typically referentially given (e.g. definite, familiar), whereas the extracted element in WHQs is typically new and focused (Gundel & Fretheim 2006). As discussed above, RCs are not typically associated with either topic or focus, and similarly to TOP constructions, do not engender an IS 'clash,' as the relativized constituent is not focused. Thus, the FBC predicts subject island violations to only arise for WHQs, but not TOP constructions nor RCs. While Abeillé et al (2020) examine PP sub-extraction in WHQs and RCs (i.e., with pied-piping), we chose to examine DP sub-extraction (p-stranding), which is (a) more frequent in a number of syntactic environments (Pullum & Huddleston 2002), and (b) does not introduce a potential attachment ambiguity.

No extraction		
Simple	(a)	Mary realized the news had completely shocked the member.
Complex obj.	(b)	Mary realized the news had completely shocked the member of the council.
Complex subj.	(c)	Mary realized the news about the city had completely shocked the member.
Simple full extraction		
Object	(d)	That member, Mary realized the news had completely shocked _.
Subject	(e)	That news, Mary realized _ had completely shocked the member.
Complex full extraction		
Object	(f)	That member of the council, Mary realized the news had completely shocked _.
Subject	(g)	That news about the city, Mary realized _ had completely shocked the member.
Sub-extraction		
Object	(h)	That council, Mary realized the news had completely shocked the member of _.
Subject	(i)	That city, Mary realized the news about _ had completely shocked the member.

Figure 1. Example itemset from TOP experiment.

Experiments: We conducted three separate experiments (TOP, WHQ, and RC) with 72 participants each. Each utilized a factorial design manipulating POSITION (extraction domain, *subject* or *object*), DP COMPLEXITY (*simple* or *complex*), EXTRACTION TYPE (*no*, *full*, or *sub-extraction*) to estimate the *superadditive cost* of sub-extraction (Sprouse 2007; Sprouse et al 2012; Vincent et al 2022). This factorial design allows for the isolation of island effects by factoring out independent variables that may influence the acceptability of island violations: DP complexity, extraction, sub-extraction from a DP. In terms of the present design, an island effect is defined as the additional penalty observed for sub-extraction conditions that exceeds the predicted cost of complexity (comparison of the acceptability of sentences with simple vs complex DPs) and the predicted cost of extraction (comparison of the acceptability of sentences with/without full DP extraction). In each experiment, participants rated the acceptability of 36 target items and 72 fillers on a 6pt scale. **Results:** Across all three experiments, we found evidence for subject island effects with TOP constructions (cf. Kush et al 2019), WHQs, and RCs: the ratings of sentences with sub-extraction from subjects were significantly lower than the combined cost of DP complexity and extraction for subjects. This indicates the presence of an additional penalty associated with sub-extraction from subjects that is not predicted by DP complexity or extraction. Sub-extraction from objects also somewhat exceeds the predicted cost, but to a significantly smaller extent than subjects. **Conclusions:** Subjects are islands for TOP constructions, WHQs, and RCs, despite the information structural differences across the constructions. Our results show that the ban on subject sub-extraction cannot be reduced to construction specific discourse-based preferences, and is best attributed to the grammatical operation underlying said constructions – movement.

Selected bibliography:

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.

Chomsky, N. (1973). Conditions on transformations. *A Festschrift for Morris Halle*. Anderson and P. Kiparsky (eds.). New York: Holt, Rinehart & Winston, 232-286.

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

Diesing, M. (1992). *Indefinites*. Linguistic Inquiry Monograph 20. Cambridge, MA: MIT Press.

Erteschik-Shir, N. (1973). *On the nature of island constraints*. MIT dissertation.

Goldberg, A. (2006). Constructions at work. *Cognitive Lingusitics* 20 (1).

Huang, C. J., (1982). *Logical relations in Chinese and the theory of grammar*. MIT dissertation.

Milsark, G. (1974). *Existential sentences in English*. MIT dissertation.

Privoznov, D. (2021). *A theory of two strong islands*. MIT dissertation.

Sprouse, J. (2007). *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge*. UMCP Dissertation.

Sprouse, J., Wagers, M., & Phillips, C. (2012). A test of the relation between working-memory capacity and syntactic island effects. *Language*, 82-123.

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.