Annual Review of Linguistics

Structural, Functional, and Processing Perspectives on Linguistic Island Effects

Yingtong Liu,[1,2,*] Elodie Winckel,[3,*] Anne Abeillé,[4] Barbara Hemforth,[4] and Edward Gibson[2]

[1]Department of Linguistics, Harvard University, Cambridge, Massachusetts, USA

[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA; email: egibson@mit.edu

[3]Department of English and American Studies, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

[4]Laboratoire de Linguistique Formelle, Université de Paris, LLF, CNRS, Paris, France

Annu. Rev. Linguist. 2022. 8:495–525

The *Annual Review of Linguistics* is online at linguistics.annualreviews.org

https://doi.org/10.1146/annurev-linguistics-011619-030319

*These authors contributed equally to this article

Keywords

syntactic islands, long-distance dependencies, filler-gap dependencies, discourse constraints, usage-based grammar, syntactic constructions, Focus–Background Conflict, linguistic interference, linguistic encoding

Abstract

Ross (1967) observed that "island" structures like "Who do you think [$_{NP}$ the gift from __] prompted the rumor?" or "Who did you hear [$_{NP}$ the statement [$_{S}$ that the CEO promoted __]]?" are not acceptable, despite having what seem to be plausible meanings in some contexts. Ross (1967) and Chomsky (1973) hypothesized that the source of the unacceptability is in the syntax. Here, we summarize how theories of discourse, frequency, and memory from the literature might account for such effects. We suggest that there is only one island structure—a class of coordination islands—that is best explained by a syntactic/semantic constraint. We speculate that all other island structures are likely to be explained in terms of discourse, frequency, and memory.

1. INTRODUCTION

Going back to Ross (1967) and Chomsky (1973), the discovery of unacceptable long-distance dependency structures such as "Who do you think [$_{NP}$ the gift from __] prompted the rumor?" or "Who did you hear [$_{NP}$ the statement [$_{S}$ that the CEO promoted __]]?" has played a major role in theories of the potential innate structure of syntax in the human mind. In this article, we summarize the evidence and arguments, focusing on experimental research in English over the past 15 years. Current evidence provides little support for the innate syntax view. We speculate that most of these "island" structures are likely to be explained in terms of discourse, frequency, and memory.

1.1. Filler-Gap Constructions

Filler-gap constructions are structures that involve a displaced constituent—a "filler"—that appears in a position other than its canonical position in a declarative clause. The canonical position of the filler in a declarative is known as the "gap" site, which we will indicate with an underscore (__). Such constructions include *wh*-questions, relative clauses, exclamatives, clefts, and topicalizations in English and many other languages. For example, the declarative form of a simple clause is provided in example 1a, along with a *wh*-question version of this clause in example 1b, where the fronted filler (*what*) is the patient (object). A corresponding relative clause is provided in example 1c, an exclamative in example 1d, an *it*-cleft in example 1e, and a topicalization structure in example 1f:[1]

(1a) Mary bought the apple.
(1b) *Wh*-question: What did Mary buy __?
(1c) Relative clause: I like the apple that Mary bought __.
(1d) Exclamative: What an apple Mary bought __!
(1e) *It*-cleft: It was the apple that Mary bought __.
(1f) Topicalization: The apple, Mary bought __.

Some variants of these constructions do not involve an overt filler (as in the English relative clause *I like the apple Mary bought*). In some languages filler-gap dependencies have a bound pronoun instead of a gap and do not display the same constraints. We mostly focus on English here and so we will not discuss such alternatives further.

1.2. The Definition of Syntactic Island

Here we define the notion of syntactic "island" as an unacceptable filler-gap dependency, which has been traditionally interpreted as ungrammatical: not generated by the grammar of the language in question.

[1]The notation of "filler" and "gap" originated from the movement-based generative theories (Ross 1967, Chomsky 1977), which we use for ease of exposition in this article. An element is not necessarily "fronted" as a result of an operation that has moved that element: Some frameworks simply assume that different constructions are characterized with different word orders and that the dependency is between a "filler" and a head ("buy" or "bought" in the sentences in example 1) (Müller 2016), for instance, Lexical Functional Grammar (Bresnan et al. 2015), Head-driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994), Construction Grammar (Goldberg 1995). The current review does not settle this question, and the structural, functional, and processing approaches can be adapted to different formal analyses of filler-gap dependencies.

1.2.1. The behavioral observation: an unacceptable filler-gap dependency. While the filler-gap relationships in example 1 are acceptable, some others are less so. Originally, syntactic islands were defined as syntactic environments that block filler-gap dependencies (Ross 1967).[2] Some commonly discussed types of islands are provided in example 2 for *wh*-questions: (*a*) a complex NP island, (*b*) a subject island, (*c*) a *wh*-island, (*d*) a coordination island, and (*e*) an adjunct island (Ross 1967, Chomsky 1973):[3]

(2a) *Who did you hear [NP the statement [S that the CEO promoted __]]?
(2b) *Who do you think [NP the gift from __] prompted the rumor?
(2c) *What did you wonder [S whether John bought __]?
(2d) *What did John buy [NP a shirt and __]?
(2e) *What did you worry [S if John bought __]?
(Sprouse et al. 2016)

The acceptability of these examples contrasts with examples with no filler-gap dependency, but a similar structure (example 3), and with examples that include an "in-situ" *wh*-phrase (example 4), which are licensed in particular contexts, such as trying to assess what someone said (an "echo-question"):[4]

(3a) Did you hear [NP the statement [S that the CEO promoted Elizabeth]]?
(3b) Do you think [NP the gift from Elizabeth] prompted the rumor?
(3c) Did you wonder [S whether John bought that jacket]?
(3d) Did John buy [NP a shirt and a jacket]?
(3e) Did you worry [S if John bought a jacket]?

(4a) You heard [NP the statement [S that the CEO promoted who]]?
(4b) You think [NP the gift from who] prompted the rumor?
(4c) You wondered [S whether John bought what]?
(4d) John bought [NP a shirt and what]?
(4e) You worried [S if John bought what]?

1.2.2. The traditional theoretical interpretation: ungrammaticality. The motivation for discussing these kinds of unacceptable sentences was that they were originally assumed not to be generated by the grammar, and thus they were deemed ungrammatical. This is why we have prefixed each example in example 2 with an asterisk, indicating that the researcher judged that the source of the unacceptability is in the grammar.

The notion of "grammaticality" is a theoretical notion: It means that a sentence is hypothesized by a researcher not to be generated by the grammar of the target language. The behavioral

[2]The name "island" derives from a movement or displacement metaphor for long-distance dependencies between two positions. The idea of an "island" is a location from which we cannot move easily (perhaps because, in the metaphor, we need to be on land to move from one place to another).
[3]Some studies divide islands into two categories—strong and weak islands—though the distinction is not always sharp (see Szabolcsi & Lohndal 2017, Szabolcsi & den Dikken 2014).
[4]Some linguists have argued that island constraints may apply to similar constructions in languages without filler-gap dependencies in the syntax, such as Mandarin (e.g., Huang 1982). Experimental evidence for this approach is provided by Lu et al. (2020), who found a locality constraint on "why" questions in Mandarin. But as shown by Cheng (2009), all other Chinese *wh*-words (e.g., "who," "what," "where") allow dependencies that would be unacceptable in English fronted *wh*-questions as in example 2 (Chaves & Putnam 2020), suggesting a difference between "why" and the other *wh*-words. The same appears to hold for Japanese (Nishigauchi 1990, p. 99), which has been argued since Ross (1967) to allow extractions from subjects (contrary to English), so the Japanese equivalent of example 4b is acceptable (see recent experimental evidence in Omaki et al. 2020).

dependent measure that syntax researchers usually work with is "acceptability": whether a sentence sounds acceptable ("good," "okay") or not in a particular context. Acceptability is a continuous, gradable notion: Sentences can be fully acceptable or partially acceptable, all the way down to completely unacceptable.

The sentences in examples 1, 3, and 4 are all relatively acceptable to many speakers of English. We might measure this with a rating scale, such as a "Likert" scale from 1 to 5 or from 1 to 7. In practice, it does not matter much what kind of scale we use when measuring acceptability: Binary scales (acceptable versus unacceptable) and scales from 1 to 5, from 1 to 7, and from 1–10 are all commonly used in the literature (Weskott & Fanselow 2011, Sprouse et al. 2013). Ratings on these scales are highly correlated across materials. It might seem that one might get more precision with a wider scale, but in practice this is not true. Hence any kind of scale works equally well, such that a participant simply evaluates the item against their language model, and gives it the required rating for each trial. The limitation in ratings is being consistent across many materials. If an experimental participant is rating many materials (say, 10–100 items), it is difficult to be consistent across all items.

Factors that we know affect acceptability ratings include lexical frequency and world knowledge. See, for instance, the sentences in example 5:

(5a) The horse bothered the donkey.
(5b) The zebu aggressed the zonkey.[5]
(5c) The dog bit the boy.
(5d) The boy bit the dog.
(5e) The girl ate the pizza.
(5f) The pizza ate the girl.

Examples 5a, c, and e all describe plausible events and use frequent English words to do so. Example 5b is comparable in meaning to example 5a but uses three low-frequency words (zebu, zonkey, and aggress), and people will rate this sentence correspondingly lower than example 5a. The sentences in examples 5d and f describe implausible events, and are rated low on an acceptability scale accordingly.

When world knowledge (including the local context) and lexical frequency cannot explain the unacceptability of a sentence form, syntacticians argue that the grammar of the language may be the source. For example, incorrect verbal agreement between the subject noun phrase and the verb (example 6a), or between an auxiliary verb and its following verb (example 6c), results in materials that English speakers will rate as unacceptable:

(6a) *The horses bothers the donkey.
(6b) The horses bother the donkey.
(6c) *The horse seems to bothering the donkey.
(6d) The horse seems to be bothering the donkey.

When the noun phrase *horses* is plural, it needs plural agreement on the verb as in example 6b, not singular agreement as in example 6a. And when the infinitival form of *be* is missing in example 6c, an ungrammatical form results, even though the meaning is probably clear in

[5]Thanks to Kyle Mahowald for informing us about zebus.

both examples 6a and c. Such materials are usually rated as unacceptable, even though they can be interpreted easily.[6]

In order to parsimoniously account for acceptability judgments across syntactic constructions, Chomsky (1965) proposed that each sentence has two levels of representation—a deep structure and a surface structure. The sentence we produce and hear is the surface form, which is transformed from the deep structure. In the case of long-distance dependency constructions, it was proposed that such constructions are transformed from their corresponding canonical declarative order. For instance, the *wh*-question *What did Mary eat?* (example 7b) is transformed from example 7a, via fronting the filler *what*; similarly, the deep structure of the topicalization construction *The apple, the girl ate* (example 8b) is transformed from example 8a, and the object *the apple* is moved to the beginning of the whole clause during transformation:

(7a)	Mary ate what?	(Deep structure)
(7b)	What did Mary eat __?	(Surface structure)
(8a)	The girl ate the apple.	(Deep structure)
(8b)	The apple, the girl ate.	(Surface structure)

Based on this movement hypothesis, Chomsky (1973, 1977, 1981, 1986a) argued for a structural account of island effects, originally called "Subjacency." According to the Subjacency constraint, movement (equivalently filler-gap extraction) is disallowed between two positions when there are two or more intervening bounding nodes, for instance, S(entence) (equivalently IP in more modern versions) and NP (equivalently DP) in English (see also Huang 1982, Rizzi 1990). Thus, the unacceptability of examples 2a–c and 2e can be attributed to the filler *who/what* moving across at least two bounding nodes in the transformation.

1.2.2.1. Construction-independence of unacceptability. An important property of the syntactic approach to the explanation of islands is that island effects are proposed to be similarly impossible across all sorts of filler-gap constructions with different meanings, for instance, *wh*-questions, relative clauses, clefts, topicalization (Chomsky 1964, 1973; see also Schütze et al. 2015). In other words, island effects are proposed to be independent of constructions and meaning. However, we see that this assumption of similar judgments across different constructions is incorrect. The

[6]There are also materials that are more acceptable than grammatical controls, but not generated by the grammar, such as the "missing-verb-phrase" examples (Gibson & Thomas 1999, Futrell et al. 2020a; example attributed to J. Fodor) in sentence ii:

(i) The patient who the nurse who the clinic had hired admitted met Jack.

(ii) *The patient who the nurse who the clinic had hired met Jack.

Sentence ii is missing a verb phrase associated with one of the preceding subject noun phrases, yet is often perceived as more acceptable than its grammatical control (sentence i). There are also examples where the compositional meaning is absurd, but which seem perfectly plausible (so-called "depth-charge" materials; Paape et al. 2020, Zhang et al. 2021):

(iii) No head injury is too trivial to be ignored.

These and other grammatical "illusion" materials—which exist only in complex meaning environments—potentially provide evidence for how the language processor works in constructing meaning. One proposal for which there is gathering evidence is the communication-based "noisy-channel" hypothesis, whereby people are guessing what was meant, given what was said, and in complex environments, they may not notice errors by the speaker (Shannon 1948, Levy 2008b, Gibson et al. 2013, Futrell et al. 2020a).

canonical example of unacceptable long-distance extractions is in *wh*-questions, but often similar extractions from relative clauses are acceptable (see Section 3, where we develop this further).

1.2.2.2. Innateness and learnability issues. In the Minimalist Program and its precursors, constraints on long-distance dependencies are unlearnable and hence innate, because of the classic poverty of the stimulus argument as proposed by Chomsky (1973, 1981, 1986b): Since these constraints are purely structural and hold across various constructions, children are unlikely to be exposed to the right input across all those different constructions. They are only exposed to examples of acceptable sentences, and there is no instruction of direct negative evidence for them to learn which long-distance dependencies are ungrammatical/unacceptable. Thus, island constraints (and constraints on other long-distance dependencies, such as anaphor resolution) must be innate (Hoekstra & Kooij 1988, Newmeyer 1991; for a critical view, see Ambridge et al. 2014).[7]

1.2.3. Sprouse's definition of syntactic islands: a superadditive interaction in complexity between two factors. More recently, Sprouse (2007) has suggested that island effects can be defined more quantitatively, such that the unacceptability of an island sentence goes beyond the additive badness of two components of complexity of a sentence (see also Sprouse et al. 2016, among others). Specifically, Sprouse et al. suggest that there is a superadditive interaction between the components that contribute to the processing difficulty of the island structure. For example, they consider two factors that might contribute to the unacceptability of an extraction from an NP, as in example 2a: (*a*) the extraction site: subject versus object position; and (*b*) the presence of the complex NP *the statement that*. They refer to extractions from subject position as "short" extractions, and those from object position as "long" extractions. Thus this 2 × 2 comparison would be as in examples 9a–d, with possible acceptability results for such a comparison in **Figure 1**:

(9a) Short, simple: Who heard [that the CEO promoted the manager]?
(9b) Short, complex: Who heard [the statement [that the CEO promoted the manager]]?
(9c) Long, simple: Who did you hear [that the CEO promoted __]?
(9d) Long, complex: Who did you hear [the statement [that the CEO promoted __]]?

Under Sprouse et al.'s definition, the unacceptability of example 9d (repeated from example 2a) is not explained by the additive weights of the two factors. The superadditivity is what makes it an "island" effect. Sprouse et al. interpret this superadditivity as evidence for syntactic constraints making such structures syntactic islands. Following Sprouse et al., some additional constraints beyond the complexity of the constructions must be underlying the superadditivity, and they claim that they are syntactic in nature.[8]

[7] Pearl & Sprouse (2013) propose a computational model trained with trigram frequencies of nonterminal syntactic categories (e.g., NP, CP, IP) calculated from child-directed speech corpora to simulate adult acceptability judgment data reported by Sprouse et al. (2012). This model shows superadditive effects for extraction phenomena on certain *wh*-dependencies. However, as the authors acknowledge, their proposal does not extend to other constructions such as relative clauses or clefts, etc., and it makes the wrong predictions for other kinds of constructions, such as parasitic gaps and across-the-board constructions. Thus, although the model provides an interesting idea of how to approach the learnability problem of unacceptable long-distance phenomena, its applicability is restricted.

[8] Almeida (2014) proposes the idea that there may be superadditivity in a 2 × 2 comparison with one condition much the worst, even if the worst of the four is still quite acceptable. He refers to such interactions as "subliminal" islands. But these are not traditional syntactic islands, because the examples are fully acceptable, and thus cannot be ungrammatical.

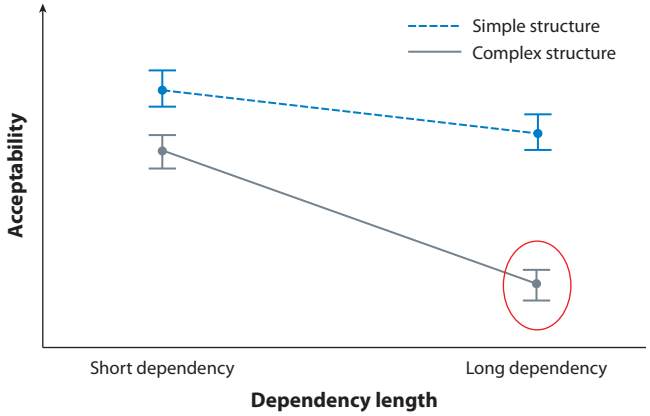


Figure 1

Illustration of an island effect as defined by Sprouse et al. (2016): a superadditive interaction between dependency length (short versus long) and complexity of the structures (complex versus simple), such that the long-dependency/complex structure is much the worst of the four conditions (*red circle*).

1.2.4. Problems with the assumption that syntactic islands are ungrammatical. According to the original claim, filler-gap constructions in island configurations are ungrammatical, independent of lexical, plausibility, and construction differences. However, many acceptable examples have been provided in the linguistic literature (including by Ross himself) and are found in well-edited corpora for most of the islands above, such as examples 10–14. These various kinds of exceptions to each type of island cast doubt on a purely syntactic explanation of unacceptable filler-gap constructions:

Counterexamples to the complex NP island

(10a) The funds that I have [hopes [the bank will squander __]] amount to more than a billion.
(Ross 1967, p. 139)

(10b) Which Middle East country did you hear [rumors [that we had infiltrated __]]?
(Pollard & Sag 1994, p. 206)

(10c) Violence is something that there are [many Americans [who condone __]].
(McCawley 1981, p. 108)

Counterexamples to the subject island

(11a) Of which cars were [the hoods __] damaged by the explosion?
(Ross 1967, p. 242)

(11b) In his bedroom, which [to describe __ as small] would be a gross understatement, he has an audio studio setup.
(Chaves 2013, p. 303)

Counterexamples to the *wh*-island

(12a) He told me about a book which I can't figure out [whether to buy __ or not].
(Ross 1967, p. 27)

(12b) How many points are the judges arguing about [whether to deduct __]?
(Kroch 1998, p. 29)

Counterexamples to the coordination island

(13a) How much can you [drink __ and still stay sober]?
(Lakoff 1986, example 2)

(13b) How many lakes can we [destroy __ and not arouse public antipathy]?
(Pollard & Sag 1994, p. 201)

Counterexamples to the adjunct island

(14a) What are you working so hard [in order to achieve __]?
(Boeckx 2012, p. 24; cf. Truswell 2007)

(14b) That's the symphony that Schubert died [without finishing __].
(Pollard & Sag 1994, p. 201)

The main problem with the notion of "syntactic island" is that its definition typically presumes that the source of the unacceptability is due to syntactic constraints. Such a definition is contradicted by the existence of the counterexamples above, along with many others.

Let us consider Sprouse's (2007) definition of syntactic island specifically, because it has received some attention in the recent generative syntactic literature. As with the original definition, the primary problem with Sprouse's definition is that it presumes that we can interpret the source of the superadditivity between two factors as coming from the syntax. However, this is not necessarily the case. We have no prior reason to think that the source of any superadditivity might be coming from violations of syntactic constraints, from violations of discourse constraints, or from overworking the processing system. Finding such an interaction simply means there is some additional factor contributing to the additional complexity of a specific combination of conditions, but we do not know what it is: It does not differentiate among the possible explanations. Many variants of each kind of explanation are equally possible after such an observation. We need independent reasons to interpret an observed superadditive interaction as coming from any source, including the syntax.[9]

Moreover, it should be kept in mind that no experiment controls all factors: All an experiment can do is control the factors that the experimenter was aware of. For example, the design in example 9 is intended to control for processing difficulty as a potential source of the observed interaction, but it actually does not. A potential confound is the number of NPs between the *wh*-filler and its dependent position (the gap site). In the short conditions in examples 9a and b, the *wh*-filler is adjacent to the gap site, corresponding to a length of zero, whereas in the long, simple condition, the distance between the *wh*-filler and the gap site is two NPs (*you*, *the CEO*), but in the long, complex condition, the distance between the *wh*-filler and the gap site is three NPs (*you*, *the statement*, *the CEO*).[10] Thus, processing complexity is an equally good explanation for the

[9]Sprouse et al. (2012) provide a published example of this fallacy. They test a resource theory of acceptability of island effects, and find no evidence for that theory. They then conclude, "We believe that the results of the experiments presented in this article provide strong support for grammatical theories of island effects because we can find no evidence of a relationship between processing resource capacity and island effects" (Sprouse et al. 2012, p. 118).

[10]No matter how one counts distance, the distance is confounded in Sprouse et al.'s design in example 9. If we instead count words rather than NPs, there are more words between the *wh*-filler and gap site for the long, complex versions relative to the long, simple ones, but no difference for the simple conditions. If we only count NPs or words between the embedded verb hear and the gap site, then the distance is longer for the long, complex condition relative to the long, simple condition, but again there is no difference for the short conditions.

observed interaction. Moreover, aspects of discourse function are also potentially confounded in this design (cf. Goldberg 2006, Abeillé et al. 2020a). Thus, interpreting such an interaction as coming from the syntax is an inferential error. For example, Keshev & Meltzer-Asscher (2019) showed that an interaction obtained in this paradigm for *wh*-islands is probably due to unmatched processing factors across conditions, not likely to grammatical rules. We come back to this specific example in Section 4.1.2.

This is not to say that the experimental design in example 9 is inadequate: It controls some factors but not others. We need independent reasons to interpret an observed superadditive interaction as coming from syntax. Finding such an interaction does not determine the source of the effect. Each source is equally well on the table after such an observation.

1.2.5. A simpler definition of island configuration: an unacceptable filler-gap dependency. As a consequence of the uncertainty of understanding the source of an island configuration, we will adopt the simplest definition of island structure: an unacceptable filler-gap dependency, with no claim about the source of the unacceptability. We adopt this simpler definition so that we can continue to refer to the relevant configurations as "island configurations" independently of their theoretical interpretations. The problem with Sprouse's definition is that one needs to understand why a filler-gap configuration is unacceptable in order to label it correctly, and our understanding may change over time. It is therefore simpler to use a definition of island configuration that does not depend on our understanding of the source of the unacceptability.

1.3. An Island Effect that Researchers Agree on: Extractions of Full Conjuncts

Whereas there is disagreement in the literature about how to explain most island effects, there is no disagreement with respect to certain conjunct islands as in examples 15 (Ross 1967) and 16 (Chaves 2012, Chaves & Putnam 2020), where one or several of the full conjuncts have been extracted:

(15a) *Who did you invite Mark and __?
(15b) *Who did you invite __ and Mark?
(15c) *Who did you invite __ and __?

(16) *What did the market crash wipe out the whole investment let alone __?

It does not seem possible to extract one or more full conjuncts, in any language. Consequently, researchers unanimously explain these phenomena in terms of syntax, sometimes called the "conjunct constraint" (Sag 2010). In an analysis without movement (a "traceless" analysis), the definition of coordination as a construction that necessarily implies (at least) two conjuncts can account for the ill-formedness of examples 15 and 16: The coordination in examples 15a and b and example 16 has only one conjunct, and in example 15c it has no conjunct at all (Sag 2010, p. 511; Chaves 2012, pp. 505–7).[11]

[11]The label "coordination island" also sometimes refers to a second kind of case, namely, fronting part of a conjunct (Grosu 1973). This type of fronting is unacceptable in many cases, as in example i:

(i) *The lute which Henry [plays __] and [sings madrigals] is warped.

This constraint on extraction is called the Element constraint and has two types of counterexamples in English (and in head-final languages as well) (Goldsmith 1985, Lakoff 1986, Kehler 2002, Chaves 2012). First,

In the remainder of this review, we summarize three major types of theories to account for all other island phenomena: (*a*) syntactic/structural accounts in Section 2[12], (*b*) functional/discourse accounts in Section 3, and (*c*) processing accounts in Section 4.[13]

2. PURE SYNTACTIC/STRUCTURAL ACCOUNTS: THE "GENERATIVE" APPROACH

2.1. Superadditivity Effects in Acceptability Ratings

As discussed above, Sprouse et al. (2016) investigated the potential sources of the unacceptability of four kinds of island structures (*whether*, complex NP, subject, and adjunct islands) in matrix/embedded *wh*-questions and relative clauses in English and Italian. To do so, they manipulated two factors in each: dependency length (short versus long) and the existence of some feature of the island structure in the sentence (island versus nonisland structures).

For English *wh*-dependencies, significant interactions were found for *whether*, complex NP, and adjunct islands; as for subject islands, *wh*-questions formed by bare-*wh*-fillers (*who/what*) did not yield a significant interaction ($p = 0.062$), while the interaction for *wh*-questions formed by complex fillers (*which car*) was significant. As for English relative clauses, the authors found a significant interaction for *wh*-islands and complex NP islands, while no significant interaction was observed for adjunct islands. The results for subject islands were mixed—the interaction was significant in the first experiment, but not so in the replication. The authors concluded that (*a*) the results could be captured by (variants of) the structural accounts (e.g., Chomsky 1986a), and (*b*) island effects vary across dependency types—*wh*-dependencies and relative clauses should not be treated the same in analyses.

Like Sprouse et al. (2016), Dillon & Hornstein (2013) investigated a variant of a complex NP island. Each experiment was designed to manipulate construction type (declarative versus interrogative) and gap location (object of the verb *open* versus inside a complex NP structure, *a clumsy attempt to open* __). The results showed a significant interaction between construction type and gap location, thus the authors ascribed the unacceptability of example 17d to the syntactic category of

---

Across-the-Board (ATB) extractions like example ii are possible if the fronted element corresponds to a gap in all conjuncts (Ross 1967, de Vries 1992):

(iia) What did Peter [buy __last week] and [throw away __yesterday]?

(iib) Which famous scientist did Peter read [a book by __] and [a newspaper article about __]?

Second, extraction out of one conjunct is ameliorated when the two conjuncts stand in an asymmetrical discourse relation (their ordering cannot be reversed with the same meaning) (Kehler 2002), as in example iii:

(iiia) Here's the whisky which I [went to the store] and '[bought __].

(iiib) What was the maximum amount that I can [contribute __] and [still get a tax deduction]?

ATB extractions and extractions out of an asymmetrical conjunct are unproblematic provided the structure complies with the discourse, processing, and performance factors. For example, Goldberg (2013) argues that in asymmetrical conjuncts one of the two conjuncts is backgrounded.

[12]For a review of formal semantic accounts of islands, see Szabolcsi & Lohndal (2017).

[13]Another approach is that of Hawkins (1999), which focuses on how and why language processing shapes grammar that includes constraints on filler-gap constructions/island phenomena. For example, it is argued that English subject-extracted relative clauses must include a lexical complementizer (unlike object-extracted relative clauses), because without the lexical complementizer there would always be difficult ambiguity with a main clause.

the extraction domain—a variant of complex NP structures:

(17a) Mary heard someone clumsily attempt to open the door.
(17b) Mary heard [a clumsy attempt to open the door].
(17c) What did Mary hear someone clumsily attempt to open __?
(17d) What did Mary hear [a clumsy attempt to open __]?

But as discussed above, the observation of such interaction effects leaves open the source: It could potentially originate in any of syntax, discourse, or processing. These researchers do not provide independent reasons that these effects might be due to structural constraints.

2.2. Mismatch Between Acceptability Ratings and Reading Times for Subject Islands

Phillips (2006) provides evidence that is claimed to support a structural account of the unacceptability of subject islands. First, he found that extraction out of an infinitival clause modifying a subject (example 18a) was rated as similarly unacceptable to extraction out of a finite clause modifying a subject (example 18b) (Phillips 2006, p. 805). Second, he observed longer reading times on the embedded verb for implausible compared to plausible subjects for the infinitival clause modifiers (e.g., for *expand* in example 19b versus example 19a), but there was no such difference for extraction from the finite clause modifiers (examples 19c and d) (Phillips 2006, p. 808). Because people are sensitive to plausibility differences in online reading times for the infinitival versions, he concluded that the unacceptability in the rating experiment could not be due to processing difficulty. He therefore concluded that the unacceptability of the materials in example 18 must be due to their syntactic structure:

Sample item for the rating task

(18a) Infinitival clause modifier
The outspoken environmentalist worked to investigate what the local campaign to preserve__ had harmed the annual migration.
(18b) Finite clause modifier
The outspoken environmentalist worked to investigate what the local campaign that preserved __had harmed the annual migration.

Sample item for the online reading experiment
Infinitival clause modifier

(19a) The school superintendent learned which schools the proposal to expand dramatically and innovatively upon the current curriculum would overburden __ during the following semester. [expand a school = plausible]
(19b) The school superintendent learned which high school students the proposal to expand dramatically and innovatively upon the current curriculum would motivate __ during the following semester. [expand a student = implausible]

Finite clause modifier

(19c) The school superintendent learned which schools the proposal that expanded dramatically and innovatively upon the current curriculum would overburden __ during the following semester. [expand a school = plausible]
(19d) The school superintendent learned which high school students the proposal that expanded dramatically and innovatively upon the current curriculum would motivate __ during the following semester. [expand a student = implausible]

Unfortunately, Phillips's conclusion does not follow. First, Chaves & Dery (2019, pp. 503–4) observe that several of the finite clause modifiers in Phillips's materials are ill-formed irrespective of the extraction, making the conclusions potentially less compelling.[14] Second, the sets of materials in the acceptability judgment experiment and those in the online experiment were not matched in an important way known to affect complexity: Bare-*wh*-fillers (*what/who*) were used in the acceptability judgment task (example 15), whereas *which*-NPs (*which schools*) were used in the online experiment (example 19). Hofmeister & Sag (2010) have shown higher acceptability ratings for extractions involving materials with *which*-NPs compared to bare-*wh*-fillers, so it is possible that the low acceptability of the materials in the ratings experiment was in part due to the use of bare-*wh*-fillers. An underspecified syntactic source for the unacceptability of the materials in example 18 is therefore not evident. And third, even if Phillips could rule out a specific processing component, that would not necessarily imply a syntactic source: The source could be a discourse effect or some other processing effect.

Phillips (2006) also evaluated versions of his materials with an extra "parasitic" gap, and found that these materials were rated as quite acceptable. The source of the acceptability of parasitic gap materials is an open question in the literature; we discuss this question with respect to islandhood in Section 4.3.

3. FUNCTIONAL/DISCOURSE ACCOUNTS

As we have seen in the sections above, syntactic approaches to islands focus on the purely syntactic aspect of filler-gap constructions (the "movement" from the gap site to the filler position). However, filler-gap constructions (e.g., *wh*-questions, clefts) are characterized not only by word order, but also by their specific discourse status. This has led some linguists, as early as in the 1970s, to study discourse factors that might explain island phenomena.

One of the most important differences with structural accounts is that functional accounts assume that fronting is licensed by syntax, even in island configurations.[15] What creates the island effect is not fronting, but the fact that the pragmatic requirements that typically lead to the use of fronting are not met. In other words, there is no good reason to use this filler-gap construction in the particular context. Sentences showing an island effect are therefore not ungrammatical, but rather infelicitous, because they are not in adequation with the specific context of the construction.

3.1. Information Structure and Extraction

Approaches based on information structure (or discourse status) predict that acceptability of extraction should be gradient and should depend on the discourse status of the gap site, and for some proposals, on the discourse status of the extracted element as well.

[14]Chaves & Dery (2019) cite problematic cases from Phillips's materials, such as the supposedly plausible *the struggle that battled the deadly disease*, which is not very plausible. The fact that the nonextracted counterparts were not always plausible may have led to low ratings from the participants.

[15]It is possible to integrate some of the theoretical spirit of the functional accounts into the generative syntactic framework via the use of covert functional phrases (i.e., FocP, TopP). But it remains unclear (*a*) why children are born with these covert functional heads rather than learn the functions of these constructions via language exposure and social interaction, and (*b*) how these covert functional phrases can capture the gradience in sentence acceptability.

3.1.1. The focus approach (Erteschik-Shir). In response to Ross's seminal work (Ross 1967), Erteschik-Shir argues that only certain elements are accessible for long-distance dependencies (gaps) and that their accessibility mainly depends on discourse factors. She states, "Extraction can occur only out of clauses or phrases which can be considered dominant in some context" (Erteschik-Shir 1973, p. 22). The notion of "dominance" should be understood here as an equivalent of the more modern notion of focus, that is, the element that carries the main (and usually new) information of the utterance.

Let us illustrate her argument with the example of the subject island exemplified in example 20 (repeated from example 2b): In general,[16] the main subject is the topic of the utterance. The relevant information will thus typically not be the subject, but what is said about it. Since the subject is not "dominant," fronting part of it is not allowed. This would explain why a subject island effect arises:

(20) *Who do you think [$_{NP}$ the gift from __] prompted the rumor?

3.1.2. The topic approach (Kuno). In Kuno's opinion, "Only those constituents in a sentence that qualify as the topic of the sentence can undergo extraction processes [. . .]" (Kuno 1987, p. 23). He calls this rule the Topichood Condition for Extraction. Note that his definition of NP topichood is more that of a discourse topic: It means that the NP is a good candidate to be the subject of the next utterance (hence a continuation topic).[17]

For instance, in example 21a, the continuation topic *she* could potentially refer to Mary as well as to Marilyn Monroe. However, according to Kuno, it is more likely to be interpreted as an anaphor to Mary, evidence that Marilyn Monroe is an unlikely (discourse) topic in this context. The Topichood Condition for Extraction would then explain why an extraction of the corresponding element, as in example 21b, is unacceptable:

(21a) I want to buy Mary's portrait of Marilyn Monroe. She's such a great artist.
(21b) *It is Marilyn Monroe who I want to buy [Mary's portrait of __].

Kuno's approach is quite different from Erteschik-Shir's: It is usually assumed that topic and comment (i.e., what is said about the sentence topic) are in complementary distribution, and that the focus domain is part of the latter. In both approaches, however, the relationship between discourse status and the extraction constraint is not spelled out explicitly: If a specific discourse criterion is not met, then fronting is not licensed, but there is no explanation about where the interface between pragmatics and syntax lies.

3.1.3. An attempt to reconcile the focus and topic approaches: salience. Deane (1991) attempts to resolve the differences between Erteschik-Shir's and Kuno's approaches by appealing to the notion of "salience." According to Deane, both focus and topic are cognitively salient. The focused element is salient because it is relevant and it is marked as such by the addresser (e.g., through prosodic stress). The topic, on the other hand, is salient because it is central to the discourse (we need to know what we are talking about). Furthermore, since the topic has most probably been

[16]That is, in a context that Erteschik-Shir (2006, 2007) calls the "canonical f[ocus]-structure."
[17]This is not the usual definition of a sentence topic: The sentence topic that contributes to information structure is usually defined as what the utterance is about.

mentioned before in the discourse, it has been cognitively activated. Salient elements imply some cognitive costs. Yet, fronting is an operation that also poses a significant cognitive cost. This operation is facilitated when the fronted elements are already salient, because they are more easily accessible. So it is easier to extract a subject (*Who left early?*), which is usually a sentence topic, than an adjunct (*When did your neighbor leave?*), which is (by default) less salient. Since the complement of the subject in example 20 and the complement of the object in example 21b are not salient, they are not easily accessible for extraction.

A similar idea leads Goldberg (2013) to assert that elements that are neither focused nor topical—she calls them backgrounded elements—are islands to extraction: "Backgrounded Constituents are Islands (BCI)." Since she assumes backgroundedness to be a gradient property, the violation of the BCI is gradient accordingly: that is why fronting out of an "island" can be more or less acceptable depending on the context.

This approach, however, as well as the preceding ones, predicts no difference across filler-gap constructions. As long as the fronting does not target a focus or topic, a penalty by extraction is expected.[18]

3.1.4. The discourse-clash approach. The solution may lie in completely dispensing with linking island phenomena to fronting, in order to keep only their discourse function (which may or may not involve fronting). This is the approach proposed by Abeillé et al. (2020a), who define their Focus-Background Conflict (FBC) constraint as follows: "A focused element should not be part of a backgrounded constituent" (Abeillé et al. 2020a, p. 3) (here, "backgrounded" should be understood as presupposed or nonfocus).

Indeed, the inconsistency in defining an element as both focus and nonfocus is underlined by several linguists (Simonenko 2016, Rizzi 2017). The case of direct questions is suggestive: The fronted element is the one about which the inquirer defines herself as being ignorant and in search of information; she cannot therefore introduce it as part of a presupposition at the same time (Simonenko 2016). By analogy, the same rationale can be extended to any form of focus, as suggested by Abeillé et al. (2020a). As a consequence, types of fronting that do not involve focusing should not be affected by the FBC constraint. This is the case, for example, for relatives in which the fillers are not (necessarily) focused. Empirical studies corroborate this dichotomy between focusing and topicalizing fronting constructions in English [experimental studies by Sprouse et al. (2016) and Abeillé et al. (2020a)], French [corpus studies by Abeillé & Winckel (2020); experimental studies by Abeillé et al. (2020a) and Abeillé et al. (2020b)], Italian [experimental studies by Sprouse et al. (2016)], and Norwegian [experimental studies by Kush et al. (2018, 2019)].

The corpus studies by Abeillé & Winckel (2020) reveal that fronting the complement of a subject noun (claimed to lead to a subject island effect) are very common in relative clauses in French (actually even the most common usage in written corpora). However, they did not find a single occurrence of such fronting for interrogatives.

Abeillé et al. (2020a) aimed to replicate Sprouse et al.'s (2016) experiments on subject islands for English (see Section 2.1) with some important differences in the materials tested: Their material is more similar across conditions (all fillers inanimate, same preposition, same semantic content), they added ungrammatical controls to their design, and most importantly, they compared

[18]Note that backgroundedness/salience in Goldberg's BCI account is gradient, so it is possible that differences in salience of the domain of extraction in distinct constructions (e.g., *wh*-question versus relative clause) can explain some of the cross-construction variation in island effects, although the theory is not yet fully developed with respect to this question.

extraction of the whole nominal complement (so-called pied-piping; example 22a) with extraction of the embedded NP (preposition stranding; example 22b), while Sprouse et al. only tested preposition-stranded versions:

(22a) Subject, pied-piping
The dealer sold a sportscar, of which [the color __] delighted the baseball player because of its surprising luminance.

(22b) Subject, P-stranded
The dealer sold a sportscar, which [the color of __] delighted the baseball player because of its surprising luminance.

The results of Abeillé et al. (2020a) for English show a clear contrast between *wh*-questions and relative clauses with pied-piping, such that there is a penalty when extracting out of the subject in *wh*-questions but not in relative clauses.[19] Sprouse et al. (2016) obtained the same contrast in Italian, and Abeillé et al. (2020a) found the same contrast in French. Furthermore, Abeillé et al. (2020b) compared relative clauses with another focalizing construction, *it*-clefts in English and *c'est*-clefts in French: They reproduced their previous results on relative clauses (no interaction), whereas clefts in both languages showed interaction effects. In Norwegian, Kush et al. (2018, 2019) tested a series of potential island environments first as *wh*-questions, then as topicalizations. They found that extractions out of adjuncts show the same contrast across constructions, with interaction effects in *wh*-questions but not in topicalizations. Kush et al. (2018, 2019) showed further results on extractions out of adjuncts in English (similar results can be found in Sprouse et al. 2016, Gibson et al. 2021). **Figure 2** illustrates many of these findings.

Finally, Tollan & Palaz (2021) looked at so-called *that*-trace effects. For *wh*-questions, it is unacceptable to front the subject of an embedded clause introduced by *that*, whereas it is fine to front the subject when the embedded clause has no lexical complementizer, as illustrated in example 23:

Wh-question

(23a) *Which family member did Lucy think that __ could drive grandad home?
(23b) Null *wh*: Which family member did Lucy think __ could drive grandad home?

Tollan & Palaz (2021) observed that this effect is greatly reduced for relative clauses, as in example 24, and they posit an explanation based on information structure following a line of explanation similar to that of the FBC:

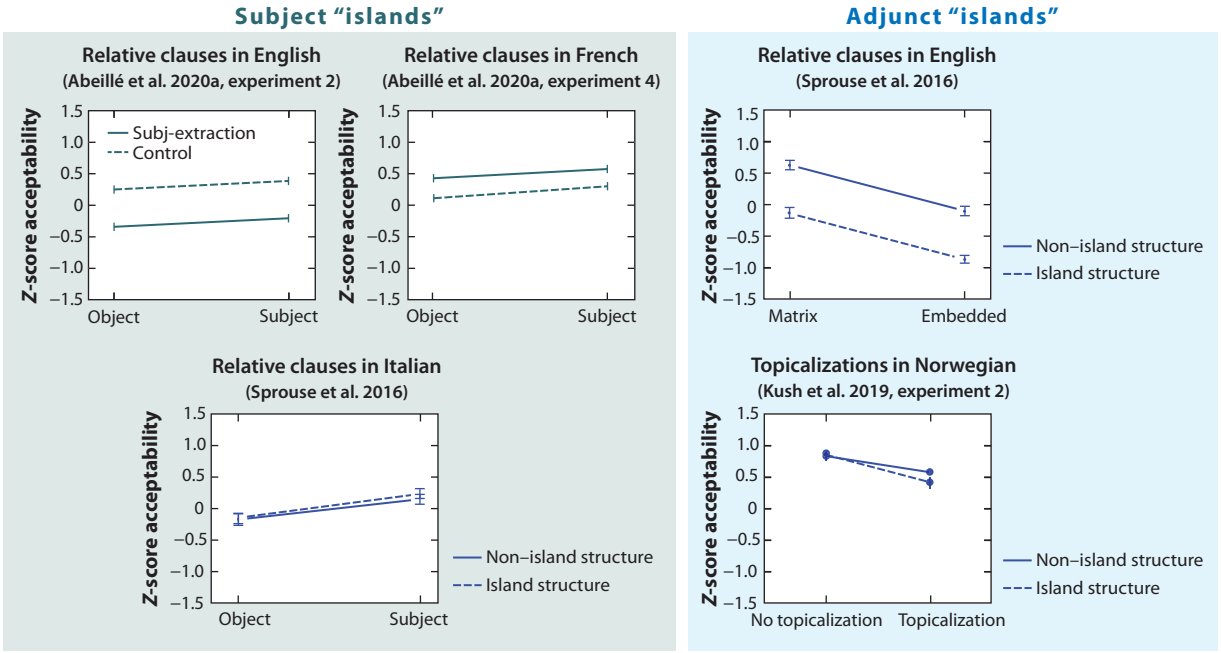
Relative clause

(24a) The family member who Lucy thought that __ could drive grandad home knew Pat.
(24b) Null RC: The family member who Lucy thought __ could drive grandad home knew Pat.

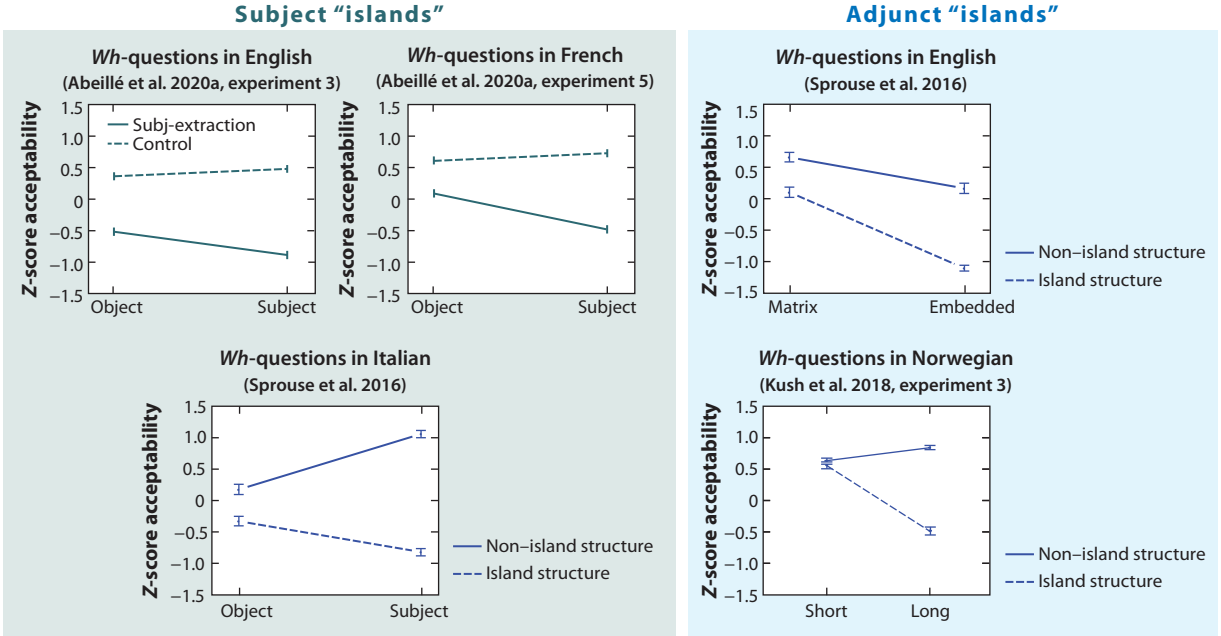
The FBC may also be able to account for some contrasts among focalizing constructions that are not accounted for in the syntactic literature, for example, the preference for indefinites when questioning the complement of a noun (Erteschik-Shir 1973, Davies & Dubinsky 2003; for experimental evidence, see Keller 2000): Because indefinite NPs introduce new entities (unlike definite NPs), the questioned element more likely belongs to the focal domain in example 25a than to

[19]With preposition stranding, they reproduce the results of Spouse et al. with a penalty when extracting out of the subject in both constructions. The authors attribute this to processing factors that have nothing to do with islands.

a Nonfocalizing constructions



b Focalizing constructions



(*Caption appears on following page*)

Figure 2 (*Figure appears on preceding page*)

Z-scored acceptability ratings for focalizing and nonfocalizing constructions in English, French, Italian, and Norwegian. Two-by-two comparisons crossing two factors [e.g., for subject-island structures (subject, object) × (filler-gap, non-filler-gap)] show a clear contrast such that the two factors interact for focalizing constructions but do not interact for nonfocalizing constructions. Norwegian figures adapted with permission from Kush et al. (2018, 2019). Italian figures and English adjunct-extraction figures adapted with permission from Sprouse et al. (2016). French data and all other English data from Abeillé et al. 2020a.

that in example 25b, which results in example 25a being more acceptable than example 25b. As expected under the FBC, no such contrast holds in example 25c, which is a relative clause:

(25a) Which actress did you buy [a picture of __]?
(25b) # Which actress did you buy [that picture of __]?
(25c) That is the actress who I bought [a/that picture of __].

3.2. Relevance

An element of a sentence can be more or less related to the main question under discussion. This property, called relevance, depends on our world knowledge. Kuno (1987) has convincingly illustrated the importance of relevance by using the following contrast:

(26a) What did you see [pictures of __]?
(Kuno 1987, p. 23)
(26b) *What did you see [a book about __]?
(Kuno 1987, p. 23)
(26c) What did you read [a book about __]?

Although example 26b is less natural than example 26a, the syntactic environment cannot be the source of the acceptability difference, since it is the same in both sentences. However, as Kuno (1987) points out, seeing a picture is synonymous with seeing what is in the picture. It is difficult to imagine under what circumstances one could see a photograph without seeing the person or object it represents. The same is not true for the subject of a book: One can see an object that one identifies as a book without seeing what the book is about. Thus, the contrasts in example 26 originate in the fact that the subject of the image is more relevant for the action of seeing an image than the subject of a book is for the action of seeing a book. On the other hand, example 26c is fine, where the verb for the direct object *book* is *read* rather than *see*.

Kluender (2004, pp. 121–22) takes a similar position and says that the relationship between the filler and the main-clause predicate is as important as the relationship between the filler and the gap. Chaves & King (2019) tested Kuno's hypothesis on examples similar to example 26. In a first norming experiment, they asked English speakers to judge how relevant an item was for a given action (example 27a). In a second experiment, participants were asked to provide acceptability judgments using the same material, this time as an interrogative (example 27b). The authors found a strong correlation between the relevance score and the acceptability judgments: The more relevant an element was considered to be for a given action, the more acceptable its fronting was judged to be:

(27a) How much does the topic of a comment matter when posting/misreading a comment?
(27b) What did Kayla post/misread a comment about?

Chaves & Putnam (2020, p. 286) define what they call the Relevance Presupposition Condition as follows: "[T]he referent that is singled out for extraction in [an Unbounded Dependency Construction] must be highly relevant (e.g., part of the evoked conventionalized world knowledge)

relative to the main action that the sentence describes. Otherwise, extraction makes no sense from a Gricean perspective, as there is no reason for the speaker to draw attention to a referent that is irrelevant for what the sentence contributes to the discourse."

4. PROCESSING ACCOUNTS

The earliest processing-based accounts of island phenomena were provided by Pritchett (1991), Kluender (1991), and Kluender & Kutas (1993a,b), who proposed that the memory load associated with processing adjunct islands (example 29b) and *wh*-islands (example 29c) can end up exceeding a limited capacity (see also Kluender 2004). Kluender & Kutas (1993a,b) provided event-related potential (ERP) and reading-time data for materials including nonextracted main-clause versions like those in example 28 and corresponding extracted main-clause versions like those in example 29:

(28a) Simple-clause-embedded, no-extract-main-clause
Have you forgotten [that you faxed a copy of that contract to the corporate office on Friday]?

(28b) *If*-clause-embedded, no-extract-main-clause
Have you forgotten [if he dragged her to the movie that weekend]?

(28c) Object-extraction, no-extract-main-clause
Have you forgotten [who he dragged __ to the movie that weekend]?

(29a) Simple-clause-embedded, *wh*-extract-main-clause
Who has she forgotten [that the boss referred that matter to __ for further study]?

(29b) Adjunct island: *if*-clause-embedded, extract-main-clause
What have you forgotten [if he dragged her to __ that weekend]?

(29c) *Wh*-island: object-extraction, extract-main-clause
What_i have you forgotten [who_j he dragged $\_\__j$ to $\_\__i$ that weekend]?

These experiments were 2 × 2 designs, comparing simple clauses to *if*-clauses and object extraction, similar to those in the study by Sprouse et al. (2016), as discussed in Section 2. And similar to Sprouse et al., Kluender & Kutas (1993a,b) found interactions (superadditivity) between the construction type and extraction. However, Kluender and Kutas interpreted these interactions as support for a working memory account.

Sprouse et al. (2012) correctly observed that Kluender and Kutas had no independent measure of working memory, so it was difficult to evaluate the claim that the observed interactions were due to working memory constraints. In an attempt to evaluate Kluender and Kutas's claim, Sprouse et al. (2012) defined verbal working memory as working memory scores measured by *n*-back tasks or serial recall tasks. Sprouse et al. found that individual working memory resources as measured by these tasks did not correlate with acceptability judgments of island constructions. Based on these results, they argued that working memory was not a likely explanation for the unacceptability of island structures. Rather, they argued that the unacceptability of gaps inside of various island structures is best explained by structural accounts.

One concern with this interpretation is that the working memory tasks used by Sprouse et al. (2012) may not reflect comprehension difficulties due to working memory capacity in sentence processing (Hofmeister et al. 2012a,b, 2013). Relatedly, there is no evidence that these working memory tasks predict comprehension difficulties of sentences that are known to be difficult due to working memory limitations (e.g., multiply nested structures) (Gibson & Scontras 2013). If Sprouse et al. could show that their memory tasks predict parsing difficulties in other phenomena that have been shown to be due to working memory, this would be stronger evidence for their interpretation. It would obviously also be more convincing for a working memory–based argument

to show that tasks that have been shown to correlate with comprehension difficulty also predict interindividual differences in island effects.

4.1. Working Memory–Based Accounts

Two aspects of cognitive costs that are associated with processing sentences are thought to contribute to the complexity of some island effects:[20]

1. Encoding aspects of a linguistic structure in memory (Vasishth & Lewis 2006; Hofmeister 2007, 2011; Hofmeister & Sag 2010).
2. Retrieving aspects of a linguistic structure at the end points of syntactic dependencies, in order to integrate the meaning in memory (Gibson 1998, 2000; Fiebach et al. 2001; Gordon et al. 2001; Warren & Gibson 2002; Van Dyke & Lewis 2003; Grodner & Gibson 2005; Lewis & Vasishth 2005; Lewis et al. 2006; Van Dyke & McElree 2006; Acheson et al. 2010; Boston 2012; Hsiao et al. 2014; Futrell et al. 2020b). The current evidence suggests that the difficulty of retrieving an earlier dependency site is most affected by interference of potentially similar elements in the interim (Gordon et al. 2001, Lewis et al. 2006). The existence of such retrieval difficulty gives rise to many behavioral effects, including the difficulty of processing nested syntactic structures (Gibson 1998, 2000); the difficulty of processing object-extracted relative clauses compared to subject-extracted relative clauses in English (King & Just 1991, Grodner & Gibson 2005, Pozniak & Hemforth 2015); and preferences for short dependencies in examples of temporary ambiguity in various constructions across languages (e.g., Stowe 1986, Frazier 1987, Clifton & Frazier 1989, Hawkins 1999, Futrell et al. 2015).

4.1.1. Island effects explained by weak encoding associated with a bare pronoun *wh*-word.

Hofmeister (2007, 2011) observed that semantically rich fillers, such as *the ruthless military dictator* (example 30b), led to faster reading times at the verb (*encouraged*) than semantically simple fillers like *the dictator* in (example 30a):

(30a) The diplomat contacted **the dictator** [who the activist looking for more contributions **encouraged** __] to preserve natural habitats and resources.

(30b) The diplomat contacted **the ruthless military dictator** [who the activist looking for more contributions **encouraged** __] to preserve natural habitats and resources.

Based on these findings, Hofmeister proposed that semantically informative fillers can help in parsing long-distance dependencies and can increase acceptability. Semantically informative fillers can facilitate downstream retrieval of the corresponding linguistic representations from memory (at the gap site), due to increased activation and resistance to interference (Hofmeister & Sag 2010). Specifically, processing constituents that syntactically depend on or modify a representation in memory reactivates (or preactivates) that representation, leading to a boost to the activation level and making the representation easier to access. In addition, semantically rich fillers provide more distinguishing features, which help to reduce interference with other discourse representations in memory at that time (Anderson & Reder 1979, Reder 1980, Bradshaw & Anderson 1982, Anderson 1983, Wiseman et al. 1985, Reder et al. 1986, McDaniel et al. 1988).

[20]There is evidence that maintaining aspects of the structure in memory also contributes to the cognitive complexity of a linguistic structure (Chomsky & Miller 1963; Gibson 1998, 2000), but this maintenance cost has not been proposed to affect the complexity of island structures.

Three kinds of island phenomena have been claimed to be explainable in terms of differences in encoding complexity on the *wh*-filler. In particular, Hofmeister & Sag (2010) and Hofmeister (2011) showed that processing costs (as measured by reading times) and acceptability ratings of complex NP islands, *wh*-islands, and adjunct islands can be greatly improved by employing complex fillers (*which*-NP versus *what/who*).[21]

For instance, in *wh*-islands, reading times were significantly shorter in the WHICH condition (example 31b) than in the BARE condition (example 31a) in the spillover PP region (*after the annual*. . .). There was no significant difference in reading time between the WHICH condition (example 31b) with island violation and the BASELINE condition (example 31c) without island structures, suggesting that the unacceptability of the BARE condition could be due to parsing difficulties associated with the semantic features of the filler (cf. Donkers et al. 2011, Tollan & Heller 2016):

(31) Albert learned that the managers dismissed the employee with poor sales after the annual performance review.
(31a) BARE: Who did Albert learn [whether they dismissed __ after the annual performance review]?
(31b) WHICH: Which employee did Albert learn [whether they dismissed __ after the annual performance review]?
(31c) BASELINE: Who did Albert learn [that they dismissed __ after the annual performance review]?

In addition, Hofmeister et al. (2013) showed that Superiority effects as in example 32 can be reduced to processing difficulty that arises from memory retrieval and similarity-based interference. The authors found that as the semantic richness/informativity of *wh*-elements increases from example 32a to d, acceptability ratings increase and reading times decrease at the verb *signed* and the following spillover region (example 33) in parallel ways. Sentences with only bare *wh*-words (example 32a) were rated the lowest and read the slowest, while sentences including two *which*-NPs (example 32d) were judged the most acceptable and read the fastest. The two conditions including both a bare *wh*-word and a *which*-NP (examples 32b and c) received intermediate ratings and reading times.[22] The authors further demonstrated that the acceptability amelioration effect holds not only for *which*-NPs, but also for other complex *wh*-phrases, such as *what book*:[23]

(32a) ??Mary wondered what who read.
(32b) Mary wondered which book who read.
(32c) Mary wondered what which boy read.
(32d) Mary wondered which book which boy read.

[21]Besides filler type, Hofmeister & Sag (2010) showed that having an indefinite NP (*a report that*... versus *the report that*...) inside the filler-gap dependency improves acceptability for *wh*-questions, consistent with the discourse-clash approach of Abeillé et al. (2020a).

[22]The condition in example 32b was judged significantly more acceptable than the condition in example 32a in experiments 2 and 4, but not in experiment 1, in Hofmeister et al. (2013). The authors concluded that the results of experiment 1 were spurious null results.

[23]A related proposal for why semantically rich fillers lead to shorter retrieval times at the gap site of *wh*-dependencies is the D-linking hypothesis (Chung 1994, Pesetsky 2000), according to which semantically restrictive *wh*-phrases narrow down the list of candidate answers/focus alternatives—a question starting with *which article* limits its answer to the set of articles, whereas a *what* question can target all nonhuman entities. It is claimed that this reduces the computational effort for answering a question for interrogatives with semantically richer *wh*-phrases.

(33) Ashley disclosed (what/which agreement) (who/which diplomat) **signed after receiving permission** from the president.

4.1.2. Island effects explained by difficult retrieval from memory. Interference-based accounts attribute the unacceptability of extractions out of some island structures to processing difficulties due to retrieval across interfering constituents in the sentence. The magnitude of the interference effects may depend on factors such as the prominence of the interfering element and its similarity to the target constituent to be retrieved (Gordon et al. 2001, Van Dyke 2003, Lewis et al. 2006, Van Dyke & McElree 2006, Vasishth & Lewis 2006, Friedmann et al. 2009, Villata et al. 2018, Keshev & Meltzer-Asscher 2019).

4.1.2.1. Interference effects in filler-gap and cataphora dependencies across island structures. Sprouse et al. (2016) demonstrated an interaction in acceptability ratings for *wh*-islands (example 34) relative to controls. They attributed this interaction to a syntactic factor, but as observed above, it could also result from other uncontrolled factors in the experimental design:

(34) What do you wonder [whether John bought __]?

Keshev & Meltzer-Asscher (2019) suggest that one uncontrolled factor is the complexity of the material between the filler and the gap: The extra *wh*-item between the filler and the gap could result in extra processing difficulty for the *wh*-island materials. In order to test this idea, Keshev & Meltzer-Asscher (2019) compared (Hebrew) materials with a long-distance filler-gap dependency (example 35) to matched materials with a long-distance anaphoric dependency as in example 36:

Wh-island

(35) ha-safranit mekira et ha-student ha-mitkaše še-ha-profesor ha-kašuax hisik **matai** ha-metargelet telamed (oto). [oto = her = optional resumptive pronoun]
'The librarian knows the weak student$_i$ that the strict professor gathered **when** the assistant will teach (her$_i$).'

Cataphora

(36) axrey še-ha-safraniyot hikiro ota, ha-profesor ha-kašuax hisik **matai** ha-metargelot yelamdu et ha-studentit ha-mitkaša.
'After the librarian met her$_i$, the strict professor gathered **when** the assistants will teach the weak student$_i$.'

Keshev and Meltzer-Asscher found a similar interaction in both filler-gap and cataphora dependencies. Thus, the unacceptability of extraction out of *wh*-islands may be best captured by encoding or retrieval interference, rather than by ungrammaticality (cf. Yoshida et al. 2014). These findings reveal that island phenomena may not be as special as initially claimed by structural accounts—instead, they seem to be intrinsically similar to other types of filler-gap dependencies and even cataphora.

4.1.2.2. Interference effects in (Featural) Relativized Minimality. Rizzi (2013) proposes a structural account of interference effects. Relativized Minimality (Rizzi 1990) requires that no relation can hold between the extracted *X* and its trace *Y* if there is an intervening element *Z* that possesses some syntactic characteristics/features with *X*. Rizzi (2013) explains the unacceptability of *wh*-islands and relative clause islands through this same constraint ("Featural" Relativized Minimality; see also Friedmann et al. 2009, Villata et al. 2018). Example 37 illustrates this hypothesis:

In example 37a, *whether* is the intervening element between *what* and its trace, and both *whether* and *what* are syntactically similar—[+*wh*] specifiers in A′-positions. In example 37b, *who* is the intervening element:

(37a)	*Wh*-islands:	*What$_{+wh}$ do you wonder	[whether$_{+wh}$ John bought	__]?
		X	*Z*	*Y*
(37b)	RC-islands:	What$_{+wh}$ do you look for the man	[who$_{+wh}$ bought	__]?
		X	*Z*	*Y*

Rizzi's account cannot explain observed interference effects in cataphora (Keshev & Meltzer-Asscher 2019; see above). Furthermore, interference effects were attested even when the interfering element/marking did not appear between the filler and the gap, but merely in the same sentence (Koesterich et al. 2021).

Another piece of evidence that cannot easily be explained with Relativized Minimality comes from Atkinson et al. (2016), who tested interrogatives with extraction out of an embedded question (*wh*-islands). Their results show that participants give lower acceptability judgments to sentences with an intervening *who* (example 38a) than to those with an intervening *which* + N (example 38b):

(38a) **Which$_{+wh}$ athlete$_{+N}$** did she wonder [**who$_{+wh}$** would recruit __]?
(38b) **Which$_{+wh}$ athlete$_{+N}$** did she wonder [**which$_{+wh}$ coach$_{+N}$** would recruit __]?

Atkinson et al.'s results are unexpected under Featural Relativized Minimality, since there is more syntactic similarity between the filler and the intervening *which* + N in example 38b than between the filler and the intervening *who* in example 38a. Instead, these results corroborate Hofmeister et al.'s (2013) hypothesis that the semantic richness of the *wh*-elements increases acceptability ratings.

4.2. Lexical and Construction Frequency Effects in Some Islands

It has been claimed that factive and manner-of-speaking verbs block *wh*-dependencies (examples 39b and c)—so-called factive and manner-of-speaking islands—whereas verbs like *say* allow them (example 39a):

(39a) Bridge verb
What did John say/think that Mary bought?
(39b) Factive verb
?? What did John know/notice that Mary bought?
(39c) Manner-of-speaking verb
?? What did John whisper/mutter that Mary bought?

Some previous studies attributed the unacceptability of sentences like examples 39b and c to discourse (Ambridge & Goldberg 2008), syntactic (e.g., Kiparsky & Kiparsky 1970, Stowell 1981, Snyder 1992), or semantic factors (e.g., Kiparsky & Kiparsky 1970). But Liu et al. (2021) found that the observed sentence acceptability ratings are most simply explained by two factors: (*a*) verb-frame frequency, the joint probability of the verb and it taking a sentence complement P(matrix verb, sentence complement), such that both filler-gap dependencies and their corresponding declaratives formed by verbs of lower verb-frame frequencies are less acceptable (cf. Hale 2001, 2003; Jurafsky 2003; Verhagen 2007; Dąbrowska 2008; Kothari 2008; Levy 2008a); and (*b*) construction type, such that *wh*-questions and *it*-clefts are less acceptable than canonical

declaratives. Liu et al. did not find any evidence of interactions between verb-frame frequency and construction type (*wh*-question or *it*-cleft versus declarative), and hence no evidence for an independent factor that would cause acceptability degradation solely in filler-gap constructions but not in declaratives. Thus, the authors concluded that the low acceptability of filler-gap constructions formed by certain sentence complement verbs is due to infrequent linguistic exposure (cf. Hsu & Griffiths 2016, Kidd et al. 2010, Navarro et al. 2012, Voorspoels et al. 2015, Xu & Tenenbaum 2007).[24]

4.3. "Parasitic" Gaps

Some unacceptable filler-gap dependencies seem to improve when there is a second dependency position ("gap"): The second gap has been called a "parasitic gap" (Engdahl 1982). The competing theories discussed in this review have different approaches to these phenomena.

According to the structural approach, the gap inside the island structure is proposed to be something other than a gap—a silent pronoun—and is acceptable only because it takes advantage of the other (licensed) gap (e.g., Cinque 1990):

(40) This is a bill that [the senators who objected to __] would probably not benefit from __.
(Chaves & Dery 2014, p. 103)

This view is challenged by the fact that materials often improve when both gaps are in two constructions considered as islands. For instance, example 41 has a gap as part of a subject (a subject island) and in an adjunct (an adjunct island), and example 42 has a gap in both adjuncts:

(41) What kind of books do [the authors of __] argue about royalties [after writing __]?
(Levine & Hukari 2006, p. 256)

(42) [Which AC unit] did you drive Alex crazy [complaining about __] yesterday [after buying __ from Craigslist]?
(Chaves & Putnam 2020)

Discourse-based approaches often explain examples with one gap by some kind of pragmatic infelicity (see Section 3.2 on relevance): The version with two gaps makes the extracted element more relevant for the proposition as a whole, since it plays two roles at the same time, hence making the sentence more acceptable.

The double gap effect may be best explained by a processing approach. For example, Chaves proposes that the gap-filling process reactivates the referent in the comprehender's memory, such that accessing this piece of information is later facilitated (Chaves 2012; Chaves & Dery 2014, 2019; Culicover & Winkler 2021). This is supported by independent evidence from processing in the work of Vasishth & Lewis (2006). Furthermore, having a second gap where the gap is most expected, as in example 39, avoids a potential filled-gap effect.

[24]Building on Liu et al. (2019) (an earlier version of Liu et al. 2021), Richter & Chaves (2020) suggest that frequency is perhaps not a good explanation for island effects like those in example 39. In their study, Richter and Chaves found that verb-bias toward an S or NP-complement did not predict the acceptability of their materials very well. But they did not investigate Liu et al.'s verb-frame frequency account: They only looked at verb-bias. In comparing the two approaches, Liu et al. (2021) found that verb-frame frequency is a better predictor of acceptability than verb-bias for their experiment 2, with 45 verbs of Richter and Chaves's verb set.

Table 1 The three approaches discussed in this review and their predictions with respect to six major unresolved issues about the source of island effects

Issue	Structural accounts (e.g., Chomsky 1977, 1986a)	Functional accounts (e.g., Goldberg 2006, 2013; Abeillé et al. 2020a)	Processing accounts (e.g., Hofmeister & Sag 2010, Liu et al. 2021)
The source of the island	Structural rules governing movement, as part of the innate language faculty (e.g., Subjacency)	Inaccessibility of the gap site, or clash of function between the filler-gap construction and the domain containing the gap	Processing difficulties, due to factors such as high working memory load or low linguistic exposure
Is gradience of island effects predicted?	Not straight-forwardly, but see Chomsky 1986a, Müller 1998, Uriagereka 2012	Yes	Yes
Where grammar comes from	Innate language faculty	Exposure and statistical generalizations	Not at issue in these approaches
Prediction of cross-construction variation	No	Only for the discourse-clash version, due to distinct functions of different constructions	There is no explicit account predicting cross-construction variation, but such variation is possible, depending on the processing difficulty associated with each specific construction.
Prediction of cross-linguistic variation	Yes (e.g., bounding nodes, the core concept of Subjacency, vary across languages)	No explicit account. Human communication is expected to use a strategy based on salient and backgrounded information cross-linguistically. However, the details might differ cross-linguistically.	To the extent that constructions vary in their usage across languages, this kind of account is consistent with cross-linguistic variation (e.g., topicalization is more frequent in Norwegian than in English (Kush et al. 2019). In addition, differences in word order predict differences in processing difficulty across languages.
Learnability of the relevant grammatical knowledge	Some constraints are unlearnable.	Grammatical knowledge is mostly learnable.	Not at issue in these approaches

5. SUMMARY AND CONCLUSION

Ever since Ross (1967) first noticed the unacceptability of a variety of filler-gap structures in English, there have been numerous attempts to explain the unacceptability of these materials.

We have discussed three major types of approaches to island structures in this review: structural accounts, functional/discourse accounts, and processing accounts. These approaches differ in the answers they provide to a number of general questions about the human capacity of language processing as summarized in **Table 1**.

Chomsky's (1973, 1977, 1986a) original approach was to suggest that there are possibly innate constraints on syntactic "movement," applying across constructions, possibly parameterized across languages. While this account was elegant in its simplicity and for its initial coverage of the original examples, it ended up not being able to account for many kinds of examples across many

constructions. Furthermore, it had the additional weakness of requiring the assumption of innate structure, because, if the judgments were correct, then the system would not be learnable. The current state of the art suggests that this structure-based approach may be incorrect in its assumption of cross-construction uniformity of acceptability judgments. This may lead to a somewhat less elegant system than what Chomsky originally proposed, but it has the huge advantages of (*a*) empirical coverage and (*b*) learnability.

Whereas the conjunct islands discussed in Section 1.3 are still understood to be explained by a structural/meaning constraint, there is no strong evidence that structure is the source of the unacceptability of any other island structure. Rather, the current set of results from corpus studies and experiments suggests a more nuanced view, with gradient acceptability and cross-construction variation, suggesting an important role of discourse, frequency, and memory constraints in explaining island phenomena as well as their counterexamples. We speculate that all of these island structures may eventually be fully explained in terms of discourse, frequency, and memory constraints.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank the following for their constructive comments on earlier versions of this article: Rui Chaves, Peter Culicover, Brian Dillon, Evelina Fedorenko, Richard Futrell, Danièle Godard, Adele Goldberg, Maayen Keshev, Robert Levine, Mark Liberman, Philip Miller, Colin Phillips, Rachel Ryskin, and Suzanne Winkler.

The work was supported by a grant from the National Science Foundation Linguistics Program (BCS-2020840) to E.G. This work was also supported by a public grant overseen by the French National Research Agency (ANR) as part of the program "Investissements d'Avenir" (ANR-10-LABX-0083). It contributes to the IdEx Université de Paris (ANR-18-IDEX-0001). E.W. was supported by the German Research Foundation (DFG) as part of the research program "Long-distance Dependencies in French: Comparative Analyses (HPSG and the Minimalist Program)," G. Mensching, Göttingen (ME 1252/14-1)/S. Müller, Berlin (MU 2822/9-1).

LITERATURE CITED

Abeillé A, Hemforth B, Winckel E, Gibson E. 2020a. Extraction from subjects: differences in acceptability depend on the discourse function of the construction. *Cognition* 204:104293. **https://doi.org/10.1016/j.cognition.2020.104293**

Abeillé A, Hemforth B, Winckel E, Gibson E. 2020b. *The subject island as a case of focus-background conflict*. Talk presented at Linguistic Evidence 2020, Eberhard Karls Univ. Tübingen, Tübingen, Ger.

Abeillé A, Winckel E. 2020. French subject island? Empirical studies of *dont* and *de qui*. *J. French Lang. Stud.* 30(3):275–300. **https://doi.org/10.1017/S0959269520000137**

Acheson DJ, Postle BR, MacDonald MC. 2010. The interaction of concreteness and phonological similarity in verbal working memory. *J. Exp. Psychol.: Learn. Mem. Cogn.* 36(1):17–36. **https://doi.org/10.1037/a0017679**

Almeida D. 2014. Subliminal wh-islands in Brazilian Portuguese and the consequences for syntactic theory. *Rev. ABRALIN* 13(2). **https://revista.abralin.org/index.php/abralin/article/view/1187**

Ambridge B, Goldberg AE. 2008. The island status of clausal complements: evidence in favor of an information structure explanation. *Cogn. Linguist.* 19(3):357–89. **https://doi.org/10.1515/COGL.2008.014**

Ambridge B, Pine JM, Lieven EVM. 2014. Child language acquisition: why universal grammar doesn't help. *Language* 90(3):e53–90. **https://doi.org/10.1353/lan.2014.0051**

Anderson JR. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard Univ. Press

Anderson JR, Reder L. 1979. An elaborative processing explanation of depth of processing. In *Levels of Processing in Human Memory*, ed. L Cermak, F Craik, pp. 385–404. Hillsdale, NJ: Lawrence Erlbaum

Atkinson E, Apple A, Rawlins K, Omaki A. 2016. Similarity of *wh*-phrases and acceptability variation in *wh*-islands. *Front. Psychol.* 6:2048. **https://doi.org/10.3389/fpsyg.2015.02048**

Boeckx C. 2012. *Syntactic Islands*. Cambridge, UK/New York: Cambridge Univ. Press

Boston MF. 2012. *A computational model of cognitive constraints in syntactic locality*. PhD Diss., Cornell Univ., Ithaca, NY

Bradshaw GL, Anderson JR. 1982. Elaborative encoding as an explanation of levels of processing. *J. Verbal Learn. Verbal Behav.* 21(2):165–74. **https://doi.org/10.1016/S0022-5371(82)90531-X**

Bresnan J, Asudeh A, Toivonen I, Wechsler S. 2015. *Lexical-Functional Syntax*. Malden, MA/Chichester, UK: Wiley-Blackwell. 2nd ed.

Chaves RP. 2012. On the grammar of extraction and coordination. *Nat. Lang. Linguist. Theory* 2(30):465–512

Chaves RP. 2013. An expectation-based account of subject islands and parasitism. *J. Linguist.* 49(2):285–327

Chaves RP, Dery JE. 2014. Which subject islands will the acceptability of improve with repeated exposure? In *Proceedings of the 31st West Coast Conference on Formal Linguistics (WCCFL 31)*, ed. RE Santana-LaBarge, pp. 96–106. Somerville, MA: Cascadilla

Chaves RP, Dery JE. 2019. Frequency effects in subject islands. *J. Linguist.* 55(3):475–521. **https://doi.org/10.1017/S0022226718000294**

Chaves RP, King A. 2019. A usage-based account of subextraction effects. *Cogn. Linguist.* 30(4):719–50. **https://doi.org/10.1515/cog-2018-0135**

Chaves RP, Putnam MT. 2020. *Unbounded Dependency Constructions: Theoretical and Experimental Perspectives.* Oxford, UK: Oxford Univ. Press

Cheng LL-S. 2009. *Wh*-in-situ, from the 1980s to now. *Lang. Linguist. Compass* 3(3):767–91

Chomsky N. 1964. Current issues in linguistic theory. In *The Structure of Language: Readings in the Philosophy of Language*, ed. JA Fodor, JJ Katz, pp. 50–118. Englewood Cliffs, NJ: Prentice Hall

Chomsky N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press

Chomsky N. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, ed. S Anderson, P Kiparsky, pp. 232–85. New York: Winston

Chomsky N. 1977. On WH-movement. In *Formal Syntax*, ed. PW Culicover, T Wasow, A Akmajian, pp. 71–132. New York: Academic

Chomsky N. 1981. *Lectures on Government and Binding*. Berlin: Mouton de Gruyter

Chomsky N. 1986a. *Barriers*. Cambridge, MA: MIT Press

Chomsky N. 1986b. *Knowledge of Language: Its Nature, Origin, and Use*. Westport, CT: Greenwood

Chomsky N, Miller GA. 1963. Introduction to the formal analysis of natural language. In *Handbook of Mathematical Psychology*, Vol. 2, ed. R Duncan Luce, RR Bush, E Galanter, pp. 269–321. New York: John Wiley

Chung S. 1994. *Wh*-agreement and 'referentiality' in Chamorro. *Linguist. Inq.* 25(1):1–44

Cinque G. 1990. *Types of Ā-Dependencies*. Cambridge MA: MIT Press

Clifton C, Frazier L. 1989. Comprehending sentences with long distance dependencies. In *Linguistic Structure in Language Processing*, ed. GN Carlson, MK Tannenhaus, pp. 272–317. Dordrecht, Neth.: Kluwer

Culicover P, Winkler S. 2021. Parasitic gaps aren't parasitic: the case of the uninvited guest. *Linguist. Rev.* In press

Dąbrowska E. 2008. Questions with long-distance dependencies: a usage-based perspective. *Cogn. Linguist.* 19(3):391–425. **https://doi.org/10.1515/COGL.2008.015**

Davies WD, Dubinsky S. 2003. On extraction from NPs. *Nat. Lang. Linguist. Theory* 21(1):1–37

de Vries G. 1992. *On coordination and ellipsis*. PhD Thesis, Catholic Univ. Brabant, Tilburg, Neth.

Deane P. 1991. Limits to attention: a cognitive theory of island phenomena. *Cogn. Linguist.* 2(1):1–63

Dillon B, Hornstein N. 2013. On the structural nature of island constraints. In *Experimental Syntax and Island Effects*, ed. J Sprouse, N Hornstein, pp. 208–20. Cambridge, UK: Cambridge Univ. Press. **https://doi.org/10.1017/CBO9781139035309.011**

Donkers J, Hoeks JCJ, Stowe LA. 2011. D-linking or set-restriction? Processing *Which*-questions in Dutch. *Lang. Cogn. Process.* 28(1–2):9–28. **https://doi.org/10.1080/01690965.2011.566343**

Engdahl E. 1982. Restrictions on unbounded dependencies in Swedish. In *Readings on Unbounded Dependencies in Scandinavian Languages*, ed. E Engdahl, E Ejerhed, pp. 151–74. Stockholm: Almquist & Wiksell Int.

Erteschik-Shir N. 1973. *On the nature of island constraints*. PhD Diss., MIT, Cambridge, MA

Erteschik-Shir N. 2006. What's what? In *Gradience in Grammar: Generative Perspectives*, ed. G Fanselow, C Féry, M Schlesewsky, R Vogel, pp. 317–35. Oxford, UK: Oxford Univ. Press

Erteschik-Shir N. 2007. *Information Structure: The Syntax-Discourse Interface*. Oxford, UK: Oxford Univ. Press

Fiebach CJ, Schlesewsky M, Friederici AD. 2001. Syntactic working memory and the establishment of filler-gap dependencies: insights from ERPs and fMRI. *J. Psycholinguist. Res.* 30:321–38

Frazier L. 1987. Syntactic processing: evidence from Dutch. *Nat. Lang. Linguist. Theory* 5:519–59

Friedmann N, Belletti A, Rizzi L. 2009. Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1):67–88. **https://doi.org/10.1016/j.lingua.2008.09.002**

Futrell R, Gibson E, Levy RP. 2020a. Lossy-context surprisal: an information-theoretic model of memory effects in sentence processing. *Cogn. Sci.* 44(3):e12814. **https://doi.org/10.1111/cogs.12814**

Futrell R, Levy R, Gibson E. 2020b. Dependency locality as an explanatory principle for word order. *Language* 96(2):371–413

Futrell R, Mahowald K, Gibson E. 2015. Large-scale evidence of dependency length minimization in 37 languages. *PNAS* 112(33):10336–41. **https://doi.org/10.1073/pnas.1502134112**

Gibson E. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition* 68(1):1–76

Gibson E. 2000. The dependency locality theory: a distance-based theory of linguistic complexity. In *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium*, ed. A Marantz, Y Miyashita, W O'Neil, pp. 95–126. Cambridge, MA: MIT Press

Gibson E, Bergen L, Piantadosi ST. 2013. Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *PNAS* 110(20):8051–56. **https://doi.org/10.1073/pnas.1216438110**

Gibson E, Hemforth B, Winckel E, Abeillé A. 2021. *Acceptability of extraction out of adjuncts depends on discourse factors*. Poster presented at the 34th Annual CUNY Conference on Human Sentence Processing, Mar. 4–6 (virtual)

Gibson E, Scontras G. 2013. *Working memory and syntactic islands revisited*. Poster presented at the 26th Annual CUNY Conference on Human Sentence Processing, Gainesville, FL, Mar. 21–23

Gibson E, Thomas J. 1999. Memory limitations and structural forgetting: the perception of complex ungrammatical sentences as grammatical. *Lang. Cogn. Process.* 14(3):225–48

Goldberg AE. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Univ. Chicago Press

Goldberg AE. 2006. *Constructions at Work*. Oxford, UK: Oxford Univ. Press

Goldberg AE. 2013. Backgrounded constituents cannot be extracted. In *Experimental Syntax and Island Effects*, ed. J Sprouse, N Hornstein, pp. 221–38. Cambridge, UK: Cambridge Univ. Press

Goldsmith JA. 1985. A principled exception to the Coordinate Structure Constraint. In *CLS 21: Papers from the General Session at the Twenty-First Regional Meeting*, pp. 133–43. Chicago: Chicago Linguist. Soc.

Gordon PC, Hendrick R, Johnson M. 2001. Memory interference during language processing. *J. Exp. Psychol.: Learn. Mem. Cogn.* 27(6):1411–23. **https://doi.org/10.1037/0278-7393.27.6.1411**

Grodner D, Gibson E. 2005. Consequences of the serial nature of linguistic input for sentential complexity. *Cogn. Sci.* 29(2):261–90. **https://doi.org/10.1207/s15516709cog0000_7**

Grosu A. 1973. On the nonunitary nature of the Coordinate Structure Constraint. *Linguist. Inq.* 4:88–92

Hale JT. 2001. A probabilistic Earley parser as a psycholinguistic model. In *NAACL '01: Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies*, pp. 1–8. New York: Assoc. Comp. Mach. **https://doi.org/10.3115/1073336.1073357**

Hale JT. 2003. *Grammar, uncertainty and sentence processing*. PhD Thesis, Johns Hopkins Univ., Baltimore, MD

Hawkins JA. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(2):244–85. **https://doi.org/10.2307/417261**

Hoekstra T, Kooij JG. 1988. The innateness hypothesis. In *Explaining Language Universals*, ed. JA Hawkins, pp. 31–55. Oxford, UK: Blackwell

Hofmeister P. 2007. *Representational complexity and memory retrieval in language comprehension*. PhD Thesis, Stanford Univ., Stanford, CA

Hofmeister P. 2011. Representational complexity and memory retrieval in language comprehension. *Lang. Cogn. Process.* 26(3):376–405. **https://doi.org/10.1080/01690965.2010.492642**

Hofmeister P, Casasanto LS, Sag IA. 2012a. Misapplying working-memory tests: a reductio ad absurdum. *Language* 88(2):408–9. **https://doi.org/10.1353/lan.2012.0033**

Hofmeister P, Casasanto LS, Sag IA. 2013. Islands in the grammar? Standards of evidence. In *Experimental Syntax and Island Effects*, ed. J Sprouse, N Hornstein, pp. 42–63. Cambridge, UK: Cambridge Univ. Press

Hofmeister P, Sag IA. 2010. Cognitive constraints and island effects. *Language* 86(2):366–415

Hofmeister P, Staum LC, Sag IA. 2012b. How do individual cognitive differences relate to acceptability judgments? A reply to Sprouse, Wagers, and Phillips. *Language* 88(2):390–400

Hsiao Y, Gao Y, MacDonald MC. 2014. Agent-patient similarity affects sentence structure in language production: evidence from subject omissions in Mandarin. *Front. Psychol.* 5:1015. **https://doi.org/10.3389/fpsyg.2014.01015**

Hsu A, Griffiths TL. 2016. Sampling assumptions affect use of indirect negative evidence in language learning. *PLOS ONE* 11(6):e0156597. **https://doi.org/10.1371/journal.pone.0156597**

Huang C-TJ. 1982. *Logical relations in Chinese and the theory of grammar*. PhD Thesis, MIT, Cambridge, MA

Jurafsky D. 2003. Probabilistic modeling in psycholinguistics: linguistic comprehension and production. In *Probabilistic Linguistics*, ed. R Bod, J Hay, S Jannedy, pp. 39–96. Cambridge, MA: MIT Press

Kehler A. 2002. *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: CSLI Publ.

Keller F. 2000. *Gradience in grammar*. PhD Diss., University of Edinburgh, Edinburgh, UK

Keshev M, Meltzer-Asscher A. 2019. A processing-based account of subliminal wh-island effects. *Nat. Lang. Linguist. Theory* 37(2):621–57

Kidd E, Lieven EVM, Tomasello M. 2010. Lexical frequency and exemplar-based learning effects in language acquisition: evidence from sentential complements. *Lang. Sci.* 32(1):132–42. **https://doi.org/10.1016/j.langsci.2009.05.002**

King J, Just MA. 1991. Individual differences in syntactic processing: the role of working memory. *J. Mem. Lang.* 30(5):580–602. **https://doi.org/10.1016/0749-596X(91)90027-H**

Kiparsky P, Kiparsky C. 1970. FACT. In *Progress in Linguistics*, ed. M Manfred, KE Heidolph, pp. 143–173. Berlin: De Gruyter

Kluender R. 1991. *Cognitive constraints on variables in syntax*. PhD Diss., Univ. Calif., San Diego

Kluender R. 2004. Are subject islands subject to a processing account? In *Proceedings of the WCCFL*, Vol. 23, ed. B Schmeiser, V Chand, A Kelleher, A Rodriguez, pp. 101–25. Somerville, MA: Cascadilla

Kluender R, Kutas M. 1993a. Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *J. Cogn. Neurosci.* 5(2):196–214

Kluender R, Kutas M. 1993b. Subjacency as a processing phenomenon. *Lang. Cogn. Process.* 8:573–633

Koesterich N, Keshev M, Meltzer-Asscher A. 2021. *Encoding interference in filler-gap and filler-resumptive dependencies*. Talk presented at the 34th Annual CUNY Conference on Human Sentence Processing, Mar. 4–6 (virtual)

Kothari A. 2008. Frequency-based expectations and context influence bridge quality. In *Proceedings of the Thirty-Eighth Western Conference on Linguistics (WECOL 2008)*, ed. M Grosvald, D Soares, pp. 136–49. Davis: Univ. Calif. Davis, Dep. Linguist.

Kroch AS. 1998. Amount quantification, referentiality, and long *wh*-movement. *Univ. Pa. Work. Pap. Linguist.* 5(2):21–36

Kuno S. 1987. *Functional Syntax: Anaphora, Discourse and Empathy*. Chicago/London: Univ. Chicago Press

Kush D, Lohndal T, Sprouse J. 2018. Investigating variation in island effects: a case study of Norwegian wh-extraction. *Nat. Lang. Linguist. Theory* 36(3):743–79. **https://doi.org/10.1007/s11049-017-9390-z**

Kush D, Lohndal T, Sprouse J. 2019. On the island sensitivity of topicalization in Norwegian: an experimental investigation. *Language* 95(3):393–420

Lakoff G. 1986. Frame semantic control of the Coordinate Structure Constraint. In *Papers from the 22nd Regional Meeting of the Chicago Linguistic Society*, pp. 152–67. Chicago: Chicago Linguist. Soc.

Levine R, Hukari T. 2006. *The Unity of Unbounded Dependency Constructions*. Chicago: Univ. Chicago Press

Levy R. 2008a. Expectation-based syntactic comprehension. *Cognition* 106(3):1126–77. **https://doi.org/10.1016/j.cognition.2007.05.006**

Levy R. 2008b. A noisy-channel model of rational human sentence comprehension under uncertain input. In *EMNLP '08: Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 234–43. New York: Assoc. Comp. Mach. **https://doi.org/10.3115/1613715.1613749**

Lewis RL, Vasishth S. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cogn. Sci.* 29:375–419

Lewis RL, Vasishth S, Van Dyke JA. 2006. Computational principles of working memory in sentence comprehension. *Trends Cogn. Sci.* 10(10):447–54. **https://doi.org/10.1016/j.tics.2006.08.007**

Liu Y, Ryskin R, Futrell R, Gibson E. 2019. Verb frequency explains the unacceptability of factive and manner-of-speaking islands in English. In *Proceedings of the 41st Annual Conference of the Cognitive Science Society*, ed. AK Goel, CM Seifert, C Freksa, pp. 685–91. Montreal: Cogn. Sci. Soc.

Liu Y, Ryskin R, Futrell R, Gibson E. 2021. A verb-frame frequency account of constraints on long-distance dependencies in English. *Cognition*. In press. **https://doi.org/10.1016/j.cognition.2021.104902**

Lu J, Thompson CK, Yoshida M. 2020. Chinese *wh*-in-situ and islands: a formal judgment study. *Linguist. Inq.* 51(3):611–23

McCawley JD. 1981. The syntax and semantics of English relative clauses. *Lingua* 53(2–3):99–149. **https://doi.org/10.1016/0024-3841(81)90014-0**

McDaniel MA, Dunay PK, Lyman BJ, Kerwin MLE. 1988. Effects of elaboration and relational distinctiveness on sentence memory. *Am. J. Psychol.* 101(3):357–69. **https://doi.org/10.2307/1423084**

Müller G. 1998. *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Dordrecht, Neth.: Kluwer

Müller S. 2016. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches*. Berlin: Lang. Sci. 4th rev. ext. ed.

Navarro DJ, Dry MJ, Lee MD. 2012. Sampling assumptions in inductive generalization. *Cogn. Sci.* 36(2):187–223. **https://doi.org/10.1111/j.1551-6709.2011.01212.x**

Newmeyer FJ. 1991. Functional explanation in linguistics and the origins of language. *Lang. Commun.* 11(1–2):3–28. **https://doi.org/10.1016/0271-5309(91)90011-J**

Nishigauchi T. 1990. *Quantification in the Theory of Grammar*. Dordrecht, Neth.: Springer

Omaki A, Fukuda S, Nakao C, Polinsky M. 2020. Subextraction in Japanese and subject-object symmetry. *Nat. Lang. Linguist. Theory* 38(2):627–69

Paape D, Vasishth S, von der Malsburg T. 2020. Quadruplex negatio invertit? The on-line processing of depth charge sentences. *J. Semant.* 37(4):509–55. **https://doi.org/10.1093/jos/ffaa009**

Pearl L, Sprouse J. 2013. Syntactic islands and learning biases: combining experimental syntax and computational modeling to investigate the language acquisition problem. *Lang. Acquis.* 20(1):23–68

Pesetsky DM. 2000. *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press

Phillips C. 2006. The real-time status of island phenomena. *Language* 82(4):795–823

Pollard CJ, Sag IA. 1994. *Head-Driven Phrase Structure Grammar*. Stanford, CA: Cent. Stud. Lang. Inf.

Pozniak C, Hemforth B. 2015. *Processing subject and object relative clauses in French and Mandarin Chinese*. Poster presented at the 28th Annual CUNY Conference on Human Sentence Processing, Los Angeles, CA, Mar. 19–21

Pritchett BL. 1991. Subjacency in a principle-based parser. In *Principle-Based Parsing*, ed. RC Berwick, SP Abney, C Tenny, pp. 301–45. Dordrecht, Neth.: Springer. **https://doi.org/10.1007/978-94-011-3474-3_12**

Reder LM. 1980. The role of elaboration in the comprehension and retention of prose: a critical review. *Rev. Educ. Res.* 50(1):5–53. **https://doi.org/10.3102/00346543050001005**

Reder LM, Charney DH, Morgan KI. 1986. The role of elaborations in learning a skill from an instructional text. *Mem. Cogn.* 14(1):64–78. **https://doi.org/10.3758/BF03209230**

Richter S, Chaves R. 2020. Investigating the role of verb frequency in factive and manner-of-speaking islands. In *Proceedings for the 42nd Annual Meeting of the Cognitive Science Society*, pp. 1771–77. n.p.: Cogn. Sci. Soc.

Rizzi L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press

Rizzi L. 2013. Locality. *Lingua* 130:169–86. **https://doi.org/10.1016/j.lingua.2012.12.002**

Rizzi L. 2017. Uniqueness of left peripheral focus, "further explanation", and Int. In *Order and Structure in Syntax I: Word Order and Syntactic Structure*, ed. LR Bailey, M Sheehan, pp. 333–43. Berlin: Lang. Sci.

Ross JR. 1967. *Constraints on variables in syntax*. PhD Thesis, MIT, Cambridge, MA

Sag IA. 2010. English filler-gap constructions. *Language* 86(3):486–545. **https://doi.org/10.1353/lan.2010.0002**

Schütze CT, Sprouse J, Caponigro I. 2015. Challenges for a theory of islands: a broader perspective on Ambridge, Pine, and Lieven. *Language* 91(2):e31–39. **https://doi.org/10.1353/lan.2015.0014**

Shannon CE. 1948. A mathematical theory of communication. *Bell. Syst. Tech. J.* 27(3):379–423

Simonenko A. 2016. Semantics of DP islands: the case of questions. *J. Semant.* 32(4):661–702

Snyder W. 1992. Wh-*extraction and the lexical representation of verbs*. Work. Pap., MIT, Cambridge, MA

Sprouse J. 2007. *A program for experimental syntax: finding the relationship between acceptability and grammatical knowledge*. PhD Thesis, Univ. Md., College Park

Sprouse J, Caponigro I, Greco C, Cecchetto C. 2016. Experimental syntax and the variation of island effects in English and Italian. *Nat. Lang. Linguist. Theory* 34(1):307–44. **https://doi.org/10.1007/s11049-015-9286-8**

Sprouse J, Wagers M, Phillips C. 2012. A test of the relation between working-memory capacity and syntactic island effects. *Language* 88(1):82–123. **https://doi.org/10.1353/lan.2012.0004**

Sprouse J, Schütze CT, Almeida D. 2013. A comparison of informal and formal acceptability judgments using a random sample from *Linguistic Inquiry* 2001–2010. *Lingua* 134:219–48

Stowe LA. 1986. Parsing WH-constructions: evidence for on-line gap location. *Lang. Cogn. Process.* 1(3):227–45. **https://doi.org/10.1080/01690968608407062**

Stowell TA. 1981. *Origins of phrase structure*. PhD Thesis, MIT, Cambridge, MA

Szabolcsi A, Den Dikken M. 2014. Islands. In *The Second* Glot International *State-of-the-Article Book*, ed. L Cheng, R Sybesma, pp. 213–40. Berlin: Mouton de Gruyter

Szabolcsi A, Lohndal T. 2017. Strong versus weak islands. In *The Wiley Blackwell Companion to Syntax*, ed. M Everaert, HC van Riemsdijk. Hoboken, NJ: Wiley. 2nd ed. **https://doi.org/10.1002/9781118358733.wbsyncom008**

Tollan R, Heller D. 2016. Elvis Presley on an island: wh dependency formation inside complex NPs. In *Proceedings of the 46th Meeting of the North East Linguistic Society (NELS 46)*. Amherst, MA: Grad. Linguist. Stud. Assoc.

Tollan R, Palaz B. 2021. *The dual nature of subjecthood: unifying subject islands and* that-*trace effects*. Talk presented at the 34th Annual CUNY Conference on Human Sentence Processing, Mar. 4–6 (virtual)

Truswell R. 2007. Tense, events, and extraction from adjuncts. In *Proceedings from the Panels of the Forty-third Annual Meeting of the Chicago Linguistic Society*, Vol. 43, Part 2, pp. 233–47. Chicago: Chicago Linguist. Soc.

Uriagereka J. 2012. *Spell-Out and the Minimalist Program*. Oxford, UK: Oxford Univ. Press

Van Dyke J. 2003. Distinguishing effects of structure and decay on attachment and repair: a cue-based parsing account of recovery from misanalyzed ambiguities. *J. Mem. Lang.* 49(3):285–316. **https://doi.org/10.1016/S0749-596X(03)00081-0**

Van Dyke JA, Lewis RL. 2003. Distinguishing effects of structure and decay on attachment and repair: a cue-based parsing account of recovery from misanalyzed ambiguities. *J. Mem. Lang.* 49(3):285–316. **https://doi.org/10.1016/S0749-596X(03)00081-0**

Van Dyke JA, McElree B. 2006. Retrieval interference in sentence comprehension. *J. Mem. Lang.* 55(2):157–66. **https://doi.org/10.1016/j.jml.2006.03.007**

Vasishth S, Lewis RL. 2006. Argument-head distance and processing complexity: explaining both locality and antilocality effects. *Language* 82(4):767–94

Verhagen A. 2007. *Constructions of Intersubjectivity*. Oxford, UK: Oxford Univ. Press. **https://doi.org/10.1093/acprof:oso/9780199226702.001.0001**

Villata S, Tabor W, Franck J. 2018. Encoding and retrieval interference in sentence comprehension: evidence from agreement. *Front. Psychol.* 9:2. **https://doi.org/10.3389/fpsyg.2018.00002**

Voorspoels W, Navarro DJ, Perfors A, Ransom K, Storms G. 2015. How do people learn from negative evidence? Non-monotonic generalizations and sampling assumptions in inductive reasoning. *Cogn. Psychol.* 81:1–25. **https://doi.org/10.1016/j.cogpsych.2015.07.001**

Warren T, Gibson E. 2002. The influence of referential processing on sentence complexity. *Cognition* 85(1):79–112. **https://doi.org/10.1016/S0010-0277(02)00087-2**

Weskott T, Fanselow G. 2011. On the informativity of different measures of linguistic acceptability. *Language* 87(2):249–73

Wiseman S, MacLeod CM, Lootsteen PJ. 1985. Picture recognition improves with subsequent verbal information. *J. Exp. Psychol.: Learn. Mem. Cogn.* 11(3):588–95. **https://doi.org/10.1037/0278-7393.11.3.588**

Xu F, Tenenbaum JB. 2007. Word learning as Bayesian inference. *Psychol. Rev.* 114(2):245–72. **https://doi.org/10.1037/0033-295X.114.2.245**

Yoshida M, Kazanina N, Pablos L, Sturt P. 2014. On the origin of islands. *Lang. Cogn. Neurosci.* 29(7):761–70. **https://doi.org/10.1080/01690965.2013.788196**

Zhang Y, Ryskin R, Gibson E. 2021. *A noisy-channel explanation for depth-charge illusions*. Talk presented at Architectures and Mechanisms for Language Processing (AMLaP) 2021, Paris, Fr., Sept. 2–4