Korean Journal of English Language and Linguistics, **Vol 26, August 2026, pp. 1053-1080**
DOI: 10.15738/kjell.26..202608.1053

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Processing Syntactic Islands in L2 English and Working Memory: Evidence from Vietnamese Learners*

Thi Quyen Nguyen (International University – Vietnam National University, Ho Chi Minh City)

Received: April 5, 2026
Revised: July 11, 2026
Accepted: August 13, 2026

Nguyen, Thi Quyen
Lecturer, School of Languages
International University, Vietnam National University, Ho Chi Minh City
VNU Township, Quarter 33, Linh Xuan Ward, Ho Chi Minh City, Vietnam
Tel: +84-28-372-44-270 (ext: 3234)
Email: ntquyen@hcmiu.edu.vn

* This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number C2025-28-05. Part of this work was developed from the author's doctoral dissertation.

ABSTRACT

Nguyen, Thi Quyen. 2026. Processing syntactic islands in L2 English and working memory: Evidence from Vietnamese learners. *Korean Journal of English Language and Linguistics 26*, 1053-1080.

This article examines island effects in L2 grammar, focusing on Vietnamese learners of English whose L1 is a *wh*-in-situ language with few island restrictions. Two questions are addressed: (a) Are these learners sensitive to English island constraints that are not instantiated in Vietnamese, and (b) can island effects in L2 English be reduced to sole processing demands? Two acceptability judgment experiments were conducted with Vietnamese EFL learners at three proficiency levels and a native English control group, using factorial designs and difference-in-differences (DD) measures following Sprouse et al. (2012). Study 1 manipulated the island structure, dependency length and D-linking across four island types (*Wh*-, Complex NP, Subject and Adjunct islands). Study 2 reduced the processing load and added an operation span task to assess working memory. Across both studies, island-violating sentences received substantially lower ratings and DD scores increased with proficiency. Processing factors modulated acceptability, especially for weaker islands, but did not eliminate island effects, and working memory did not reliably predict island sensitivity. These findings indicate that processing factors modulate, but do not determine, island effects, which are better explained as reflecting UG-constrained grammatical representations in interlanguage. The developmental trajectory observed across proficiency levels is therefore more consistent with the predictions of the Full Transfer/Full Access hypothesis than with the Fundamental Difference Hypothesis, providing evidence that adult L2 learners can acquire fine-grained island constraints despite limited support from the first language.

KEYWORDS

island effects, interlanguage grammar, Vietnamese learners, *wh*-movement, working memory, Universal Grammar

1. Introduction

A defining feature of human language is its capacity to establish long-distance dependencies between elements separated by multiple clauses. In English, a *wh*-phrase may move from its original position to the left edge of a clause, leaving behind an implicit gap, as illustrated in (1) below.[1] It has been proposed that the dependency between a *wh*-phrase and its gap can extend across any number of embedded clauses.

(1) Who did John think that Paul suspected that Mary killed _?

Although long-distance dependencies are not limited by conventional measures of length, not all such dependencies are well-formed. Crucially, when the gap position occurs within certain structures, the sentence becomes unacceptable, as illustrated in (2).

(2) *Who did John make the claim that Paul suspected that Mary killed _?

In (2), extraction of the *wh*-phrase from the embedded clause headed by the noun phrase *the claim* yields a sharply degraded sentence. Structures that resist such extractions are known as *islands*, and the pattern of unacceptability that results from illicit extraction is termed an *island effect* (Ross 1967). Since Ross's (1967) original work, island phenomena have been central to debates about the architecture of grammar and the status of universal grammar (UG). In one influential view, island effects reflect abstract, innately specified constraints on movement, such as Subjacency or the Condition on Extraction Domain (Chomsky 1973, 1977, 1986, Huang 1982). These constraints capture the intuition that certain structural domains are opaque to *wh*-movement in a way that goes beyond simple linear distance or surface complexity, and they have often been taken as paradigmatic cases of the poverty-of-the-stimulus problem: speakers converge on fine-grained locality conditions for which the input provides little direct evidence.

An alternative line of research attributes at least some island effects to general processing limitations. In this view, highly complex dependencies overtax limited cognitive resources, leading to degraded acceptability even in the absence of categorical grammatical violations (Hofmeister and Sag 2010, Kluender 1998, 2004, Kluender and Kutas 1993). Factors such as dependency length, embedding, discourse accessibility and retrieval difficulty are assumed to influence how easily a *wh*-dependency can be constructed and maintained. D-linked *wh*-phrases and supportive contexts, which improve discourse accessibility and predictability, are known to ameliorate the acceptability of many island-violating sentences, especially on "weak" islands such as *wh*-islands (Alexopoulou and Keller 2013, Kim and Goodall 2016, Kush et al. 2013, 2018, Pesetsky 1987).

Recent experimental work has sought to disentangle these grammatical and processing perspectives via factorial acceptability designs and difference-in-differences (DD) measures (Sprouse et al. 2012). By independently manipulating island structure and dependency length and by computing superadditive island effects as DD scores, Sprouse et al. (2012) and subsequent studies have shown that native speakers exhibit robust, positive DD scores for a range of island types and that these effects remain even when D-linking and contextual support reduce the processing load (Kush et al. 2018). Importantly, these studies generally report little or no systematic relationship between individual working-memory measures and island sensitivity, which poses challenges for strong resource-

[1] In this study, I use theory-neutral terminology and refer to the endpoint of this dependency as the gap.

limitation accounts. More recently, Aldosari et al. (2024) extended this methodology to Najdi Arabic learners of English and reported that L2 learners patterned closely with native speakers in both the magnitude and hierarchy of island effects, again without robust working-memory correlations.

These questions are particularly important in second language acquisition (SLA). If island effects are indeed rooted in UG-constrained syntax, then the presence of robust island sensitivity in adult L2 learners, especially when the relevant constraints are not clearly instantiated in the L1, supports the view that UG remains available beyond childhood (Belikova and White 2009, Hawkins and Chan 1997, White 2003). Conversely, if island phenomena can be reduced to processing constraints and domain-general resource limitations, then the argument from island effects to UG access becomes weaker. SLA also provides a useful testing ground for dissociating the roles of grammar and processing: learners' interlanguage systems might instantiate UG-constrained representations, yet their performance can still be shaped by L1 transfer and by limitations in L2 processing efficiency (Clahsen and Felser 2006).

The present study investigates these issues in the interlanguage grammar of Vietnamese learners of English. Vietnamese is essentially an SVO language whose *wh*-elements typically remain in situ, as illustrated in (3).

(3) a. Bình uống sữa.
Bình drink milk
'Bình drinks milk.'
b. Bình uống gì (thế/vậy)?
Bình drink what Q-PRT
'What does Bình drink?'

In Vietnamese, *wh*-phrases such as *gì* 'what' generally appear in their canonical theta position, and clauses containing *wh*-in-situ do not generally exhibit the same pattern of island restrictions observed with overt *wh*-movement in English. Many *wh*-dependencies that violate English island constraints are acceptable in Vietnamese, especially those involving argument *wh*-phrases and certain *wh*-adjuncts (Bruening and Trần 2006, Tsai 2009). This typological profile raises important questions for SLA: Can Vietnamese learners acquire English island constraints that are underdetermined by their L1? If so, what does their performance reveal about the relative contributions of grammatical knowledge and processing pressures? Vietnamese–English pairing is particularly informative when viewed against the broader L2 island literature. Work with other L1s has shown that learners can be sensitive to island violations in both offline judgments and online processing (e.g., Aldwayan et al. 2010, Boxell and Felser 2017, Li 1998, Martohardjono 1993, Omaki and Schulz 2011, Perpiñán 2020, White and Juffs 1998). Aldosari (2015), for example, reported that Najdi Arabic learners, whose L1 differs from English in the realization of *wh*-dependencies and island constraints, behave similarly to native speakers in terms of both island hierarchies and the absence of systematic working-memory effects. At the same time, it remains unclear whether these findings generalize across typologically distinct first languages and whether island effects reflect the same underlying grammatical representations in L1 and L2.

This paper addresses these questions in a comprehensive way through two formal acceptability judgment experiments and an individual-differences measure of working memory. Two research questions guide the study. First, are Vietnamese learners of English sensitive to island violations in their acceptability judgments, and how does their sensitivity compare to that of native speakers across different island types? Second, to what extent can L2 island effects be attributed to processing factors such as dependency length, D-linking and working memory, as opposed to UG-constrained grammatical principles?

To answer these questions, we adopted the factorial design and DD logic of Sprouse et al. (2012) and situated our work within the recent native and L2 literature on the source of island effects (Aldosari 2015, Kush et al. 2018, Sprouse et al. 2012). Study 1 manipulated the island structure, dependency length and D-linking across four classic island types (*wh*-islands, complex NP islands, subject islands and adjunct islands) in Vietnamese learners at three proficiency levels and a native English control group. Study 2 reduced the processing load by using D-linked *wh*-phrases throughout and embedding target sentences in short contexts; an operation span task was employed to gauge individual working-memory capacity. Together, the two studies examine how island effects in Vietnamese–English interlanguage can inform the grammar–processing debate and key L2 theories, especially the Full Transfer/Full Access hypothesis (Schwartz and Sprouse 1996) and the Fundamental Difference Hypothesis (Bley-Vroman 1990, 2009).

2. Literature Review

2.1 Grammar-based Accounts of Island Effects

In generative grammar, island effects are taken to reflect abstract grammatical principles that limit where constituents can be extracted. Ross (1967) first systematically documented a range of island constraints by observing that certain structural configurations do not permit extraction of *wh*-phrases or other moved constituents. Among the most influential constraints are the *Wh*-Island Constraint, the Complex NP Constraint, the Subject Island Constraint, and the Adjunct Island Constraint.

Wh-Islands arise when extraction takes place out of an embedded interrogative clause. Ross observed that while extraction out of a declarative complement is generally acceptable, extraction out of an embedded *wh*-clause yields ungrammaticality, as in (4), in which the contrast in acceptability reflects the fact that extraction from a [+wh] clause (e.g., introduced by *whether*) is blocked.

(4) a. Who did Carl think that John met _?
b. *Who did Carl wonder whether John met _?

The Complex NP Constraint (CNPC) prohibits extraction from noun phrases that contain a clause (Ross 1967, p. 127), as illustrated in the contrast in (5). In (5b), the gap is embedded inside a clause dominated by a noun (*claim*), forming a complex NP island.

(5) a. Who did Carl claim that John met _?
b. *Who did Carl make the claim that John met _?

Subject Islands involve extraction from subjects, which are generally resistant to subextraction, as shown in (6) where part of the subject NP [*your interest in* _] cannot be extracted and questioned.

(6) *Who does your interest in _ seem strange to me?

The details of Ross's examples aside, subsequent work has generalized the observation: complex subjects, whether nominal or clausal, typically resist extraction.

Finally, the Adjunct Island Constraint (Huang 1982) captures the well-known contrast between extraction from complement clauses and extraction from adjunct clauses such as temporal or causal clauses. While *wh*-movement from a complement CP is usually permitted, movement out of an adjunct CP is strongly degraded, as in (7b) where the *wh*-phrase is extracted from a *before*-clause adjunct, yielding an island violation.

(7) a. Which song did Carl say that John sang _?
b. *Which song did Carl leave before John sang _?

Early generative work sought to unify these patterns under broader locality principles. For example, Chomsky's (1977) Subjacency condition restricts movement across multiple bounding nodes, providing a unified account of several island effects. To address remaining cases, Huang (1982) proposed the Condition on Extraction Domains (CED), which links extractability to whether a domain is properly governed, thereby capturing both Subject and Adjunct Islands.

Subsequent developments in government and binding theory and the Minimalist Program recast these ideas in terms of barriers and locality conditions, including Relativized Minimality (Chomsky 1986, 1995, Rizzi 1990). Despite theoretical refinements, a core intuition has persisted: island effects arise because the grammar forbids certain long-distance dependencies on the basis of abstract structural configurations.

Within this grammar-based tradition, island phenomena have been widely cited as strong evidence for Universal Grammar. Many island constraints are infrequent, opaque, or arguably absent in the primary linguistic data, yet native speakers converge on highly similar and fine-grained judgments about which extractions are acceptable and which are not. The abstractness and underdetermination of island patterns have therefore been taken as classic cases of the poverty-of-the-stimulus problem (e.g., Boeckx 2008, Sprouse 2007) as speakers appear to know more about the locality conditions of movement than the input alone can plausibly support.

At the same time, attempts to subsume island phenomena under a single, unified grammatical principle have faced persistent empirical challenges. Moreover, much of the traditional evidence rests on informally collected judgments, raising questions about whether the observed patterns reflect categorical grammatical constraints, processing difficulty, or some combination of the two. Nonetheless, if at least some UG-constrained locality conditions are part of the mental grammar and govern the formation of *wh*-dependencies, the poverty-of-the-stimulus argument extends naturally to acquisition. Such constraints must be available to learners if they come to exhibit systematic island sensitivity on the basis of underdetermined input, particularly when their L1 does not overtly instantiate the relevant restrictions.

2.2 Processing-based Accounts of Island Effects

In parallel with grammar-based accounts, a robust line of research has argued that at least some island effects reflect properties of the sentence processing system rather than, or in addition to, categorical grammatical constraints (Hofmeister and Sag 2010, Kluender 1998, 2004, Kluender and Kutas 1993). On these accounts, complex dependencies impose heavy demands on working memory and general cognitive resources; when these demands exceed a certain threshold, comprehenders rate sentences as unacceptable, even if the underlying structure is grammatically well-formed.

Kluender and Kutas (1993) and Kluender (1998, 2004) develop a fairly elaborate processing-based account in which island effects arise from the cumulative processing demands associated with complex configurations. Processing a long-distance *wh*-dependency across an island involves several operations: maintaining the *wh*-filler in working memory, searching for the appropriate gap position, and carrying out referential and lexical processing at the left edge of an embedded clause or complex nominal. Each of these processes, taken in isolation, imposes a manageable burden on working memory. However, when they occur in parallel, as in extraction out of complex subjects, adjuncts, or *wh*-complements, the total processing demands may exceed the capacity of the system. Kluender (1998) suggests that once limits are approached, the processing system "cuts back" on ongoing activation to keep the system stable, with the result that the representation of the *wh*-filler can be lost. The resulting failure to maintain the dependency during processing may contribute to the strong impression that such sentences are "ungrammatical".

Building on this work, Hofmeister and Sag (2010) refine the processing-based approach by identifying specific processing factors that contribute cumulatively to island effects. They highlight, in particular, (i) the complexity of the syntactic structure (e.g., clausal embeddings, complex subjects, adjuncts, relative clauses) and (ii) the length of the *wh*-dependency. Longer dependencies require the *wh*-filler to be held active in memory across more material and to survive more potential interference; more complex structures require additional resources to build and maintain. When long dependencies arise from complex domains, the combined cost is predicted to be superadditive, leading to sharply reduced acceptability. A key virtue of such accounts is that they promise to simplify the grammar by deriving island effects from independently motivated properties of the processing system, thereby reducing or eliminating the need for stipulative island constraints.

Within this framework, much attention has been given to *island amelioration*, namely cases where sentences that would traditionally be classified as island violations receive improved acceptability under certain conditions. Hofmeister and Sag (2010) and Goodall (2015) show that island-violating sentences become more acceptable when the extracted phrase is made more semantically rich or discourse accessible, for example, by adding a restrictive modifier or by using a complex *wh*-phrase. These manipulations are thought to facilitate the maintenance and retrieval of the *wh*-filler: a more informative NP has a more distinctive memory representation and is therefore easier to retrieve at the gap site. The well-known effect of D-linked *wh*-phrases, such as *which of the students*, can be understood in similar terms (Pesetsky 1987): D-linked phrases pick out a salient, contextually given set of individuals, reducing uncertainty about the intended referent and easing retrieval in long-distance dependencies. Importantly, such amelioration does not necessarily imply that the underlying grammatical violation has disappeared, since acceptability may reflect both grammatical and nongrammatical factors. Nevertheless, the systematic modulation of acceptability by D-linking and contextual support provides clear evidence that processing and discourse factors contribute to the severity of island effects.

A related strand of research investigates whether individual working-memory capacity predicts island sensitivity. If island effects are fundamentally driven by resource limitations, individuals with greater working-memory capacity should be better able to process complex wh-dependencies and, in principle, should exhibit weaker island penalties. Studies have therefore used complex-span tasks such as operation span (Conway et al. 2005, Turner and Engle 1989) to index working-memory capacity and related these scores to measures of island sensitivity. The evidence, however, has been mixed, with some studies reporting relationships under particular measures or conditions and others finding little systematic association between working-memory capacity and island sensitivity (Aldosari 2015, Aldosari et al. 2024, Hofmeister et al. 2012, Sprouse et al. 2012).

From an SLA perspective, processing-based accounts have important consequences. If island effects can be fully reduced to processing complexity and domain-general resource limitations, demonstrating island sensitivity in L2 learners would not by itself provide straightforward evidence for UG access. Such patterns might instead reflect learners' increasing ability to process complex L2 structures and maintain long-distance dependencies. However, grammar and processing need not be mutually exclusive explanations. Processing factors may modulate the magnitude of island penalties, while structural constraints independently restrict the dependencies that the grammar permits. Distinguishing these possibilities therefore requires experimental designs that separate the effects of structural complexity and dependency length and test whether their combination produces a superadditive penalty beyond their independent processing costs.

2.3 Universal Grammar Accessibility in SLA

Beyond the debate over whether island effects arise from grammatical constraints or processing limitations lies a broader question concerning the nature of adult second language (L2) acquisition: namely, whether adult learners retain access to Universal Grammar (UG). Two influential hypotheses addressing this issue are the Full Transfer/Full Access (FT/FA) Hypothesis (Schwartz and Sprouse 1996) and the Fundamental Difference Hypothesis (FDH) (Bley-Vroman 1990, 2009). Although both hypotheses seek to explain how interlanguage grammars develop, they make fundamentally different assumptions regarding the role of UG and, consequently, different predictions concerning the acquisition of grammatical phenomena that are underdetermined by the input, such as syntactic island constraints.

The FT/FA Hypothesis proposes that the initial state of L2 acquisition corresponds to the final state of the learner's L1 grammar (Full Transfer), while subsequent development is guided by continued access to UG (Full Access) (Schwartz and Sprouse 1996). Under this account, learners are expected to exhibit L1 influence during the early stages of acquisition, but they are nevertheless capable of restructuring their interlanguage grammars to accommodate target-language properties that are absent from the L1. Consequently, if the L1 does not instantiate the relevant island restrictions, FT/FA allows for initial L1-based patterns followed by restructuring toward target-like constraints as learners gain experience with the L2.

By contrast, the FDH argues that adult L2 acquisition is fundamentally different from child L1 acquisition because learners no longer have the same access to the domain-specific acquisition mechanisms available during first language acquisition and instead rely more heavily on previously acquired L1 knowledge and domain-general learning mechanisms (Bley-Vroman 1990, 2009). Accordingly, properties that are absent from the L1 and underdetermined by L2 input present a particular learnability problem. Applied to island phenomena, the FDH therefore leads to the expectation that learners whose L1 does not instantiate English-like island constraints may have persistent difficulty developing fully target-like sensitivity to such violations.

Since the formulation of the FT/FA Hypothesis and the FDH, research on adult second language acquisition has expanded considerably. More recent research has broadened the theoretical landscape considerably, emphasizing the roles of usage, experience, input processing, and general cognitive mechanisms in L2 development (e.g., Ellis 2006, MacWhinney 2005, VanPatten 2015), as well as potential differences between native and nonnative syntactic processing (Clahsen and Felser 2006). These perspectives have enriched our understanding of second language acquisition by moving beyond a simple dichotomy between grammatical competence and processing limitations.

Nevertheless, the present study adopts the FT/FA and FDH as its principal theoretical frameworks for two reasons. First, both hypotheses make explicit and contrasting predictions concerning the accessibility of Universal Grammar in adult L2 acquisition, which directly bears on the acquisition of syntactic island constraints. Second, these hypotheses address questions of UG accessibility that have been central to previous research on L2 island phenomena (e.g., Martohardjono 1993, White and Juffs 1998), providing an appropriate basis for interpreting the present findings and facilitating comparison with the existing literature. Rather than attempting to evaluate all contemporary models of interlanguage development, the present study focuses specifically on whether the acquisition of English island constraints by Vietnamese learners is more consistent with the predictions of continued UG accessibility or with an account in which adult L2 acquisition relies primarily on L1-based knowledge and domain-general learning mechanisms.

In the present study, L1 transfer is not treated as an independent alternative to grammatical or processing accounts. Within the FT/FA framework, L1 transfer characterizes the initial state of interlanguage development, while subsequent restructuring may reflect learners' continued access to UG. Accordingly, L1 transfer is considered primarily as a potential source of developmental differences across proficiency levels, rather than as an alternative explanation for the ultimate source of English island constraints.

2.4 Previous Studies on Island Effects in SLA

Building on the theoretical perspectives outlined in the previous section, a substantial body of SLA research has examined whether adult L2 learners can acquire syntactic island constraints that are absent or only weakly instantiated in their first language. Much of the early work on L2 island sensitivity focused on learners from *wh*-in-situ languages such as Chinese, Japanese, and Korean (Bley-Vroman et al. 1988, Johnson and Newport 1991, Li 1998, Martohardjono 1993, Schachter 1990, White and Juffs 1998). This work forms part of a broader literature using the acquisition of English *wh*-dependencies to investigate UG accessibility in learners whose L1 does not instantiate English-style overt *wh*-movement (e.g., Hawkins and Chan 1997, Hawkins and Hattori 2006).

Early studies typically employed binary grammaticality judgment tasks using relatively small sets of constructed sentences. Island-violating structures were compared with non-island controls across several island types, including relative clauses, sentential subjects, and *wh*-islands, and learner performance was generally measured by the proportion of correctly rejected violations. Although the findings were mixed, a broad tendency was for more advanced learners to show greater sensitivity to island violations. Advanced learners often demonstrated sensitivity to island violations, particularly for extraction from relative clauses and sentential subjects, whereas judgments of weaker island types, especially *wh*-islands, remained more variable (Martohardjono 1993, White and Juffs 1998). These findings suggested that learners could acquire at least some abstract restrictions on long-distance dependencies, but methodological limitations made it difficult to determine whether such performance reflected genuine grammatical knowledge or alternative explanations such as L1 transfer, distributional learning, or task-specific strategies.

Subsequent research refined this picture by distinguishing between strong and weak island domains. For example, Martohardjono (1993) reported that L2 learners were consistently successful in rejecting extractions from relative clauses and sentential subjects but showed greater variability for noun-complement clauses. Similarly, Belikova and White (2009) argued that learners are more likely to acquire universally constrained structures, such as Subject and Adjunct Islands, than language-specific constraints such as whether-islands, whose grammatical status varies across languages (see also Alexopoulou and Keller 2013). Together, these studies suggest that island sensitivity develops gradually and is influenced by both the nature of the island constraint and learners' stage of interlanguage development.

Methodological developments in the broader experimental island literature have provided new tools for addressing these questions in SLA. Rather than relying on categorical grammaticality judgments, these studies employ Likert-scale ratings or magnitude estimation, fully crossed factorial designs, and difference-in-differences (DD) measures to quantify island effects more precisely (Almeida 2014, Kim and Goodall 2016, Kush et al. 2013, 2018, Sprouse et al. 2012). In these designs, island structure (island vs. non-island) and dependency length (short vs. long) are manipulated independently, allowing island effects to be identified as superadditive interactions between the two factors. A nonzero superadditive DD effect therefore provide stronger evidence for genuine island effects than simple decreases in acceptability, which may instead reflect processing costs associated with longer dependencies.

Recent L2 research has increasingly applied these methodological insights while also examining how grammatical constraints interact with processing-related factors. Aldosari et al. (2024), for example, building on the factorial methodology of Sprouse et al. (2012), found robust island effects in both native speakers and Najdi Arabic learners of English, together with little systematic relationship between working-memory capacity and island sensitivity. Complementary evidence from online processing studies suggests that L2 learners can also become sensitive to structural restrictions on filler–gap dependency formation during real-time comprehension (Aldwayan et al. 2010, Boxell and Felser 2017, Omaki and Schulz 2011). Collectively, these findings indicate that processing factors may influence the magnitude of island effects without necessarily accounting for their existence.

Despite these advances, important questions remain. Relatively few L2 studies have systematically examined grammatical and processing factors within the same research design, and learners from typologically relevant wh-in-situ backgrounds remain underrepresented. Vietnamese is particularly informative because overt wh-movement is generally absent and dependencies that would violate island constraints in English may be acceptable in Vietnamese (Bruening and Trần 2006, Tsai 2009). The present research therefore examines whether Vietnamese learners develop sensitivity to English island constraints and whether that sensitivity can be accounted for by processing-related factors. Across two studies, island structure and dependency length are systematically manipulated: Study 1 additionally manipulates D-linking, whereas Study 2 examines island effects under processing-facilitating conditions using D-linked wh-phrases and supportive contexts and tests whether individual working-memory capacity predicts island sensitivity.

2.5 Vietnamese *wh*-in-situ and the Island Phenomena

Vietnamese is a *wh*-in-situ language in which *wh*-elements typically remain in their canonical position within the clause, with interrogative force signaled by the presence of a *wh*-item and, optionally, a sentence-final question particle. *Wh*-fronting is marginally possible but is generally associated with focus or contrastive readings and is plausibly distinct from the A′-movement observed in English (Bruening and Trần 2006, Tsai 2009). Vietnamese also distinguishes between argument *wh*-phrases (e.g., *ai* 'who', *gì* 'what') and certain adjunct *wh*-phrases (e.g., *vì sao* 'why', *khi nào* 'when'). Argument *wh*-phrases typically remain in situ and can take wide scope without overt movement. Adjunct *wh*-phrases show more restricted behavior. For example, causal *vì sao* 'why' tends to occur clause-initially when interpreted with wide scope, as illustrated in (8).

(8) a. Vì sao anh mua chiếc xe (thế)?

vì sao	anh	mua	chiếc.xe	thế
why	you	buy	CL.car	Q-PRT

'Why did you buy the car?'

LF: 'What is the reason x such that you bought the car because of x?'

b. *Anh mua chiếc xe vì sao?

anh	mua	chiếc.xe	vì sao
you	buy	CL.car	why

LF: 'Why did you buy the car?'

In (8a), the *wh*-phrase *vì sao* appears clause-initially and yields a natural question, whereas its clause-final counterpart in (8b) is sharply degraded.

With respect to island phenomena, previous work suggests an important asymmetry between argument and adjunct *wh*-phrases in Vietnamese (Bruening and Trần 2006, Trần 2009, Tsai 2009). Argument *wh*-phrases can occur inside domains corresponding to English *wh*-, complex NP, subject, and adjunct islands while taking scope outside those domains. Causal *vì sao*, by contrast, exhibits considerably stronger locality restrictions. Representative examples are given in (9)–(12).

(9) *Wh*-islands

a. John tưởng $[_{CP}$ ai đã mua gì vậy]?
John think who ASP buy what Q-PRT
'Who is x and what is y such that John thought that x bought y?'

b. ??John muốn biết $[_{CP}$ ai vì sao chết thế]?
John want know who why die Q-PRT
??'What is the reason x and who is y such that John wanted to know y died because of x?'
Wide-scope: ??Why did John want to know who died.
Narrow-scope: Who did John want to know why died.

(10) Complex NP islands

a. John nghe $[_{NP}$ tin đồn $[_{CP}$là An yêu ai (vậy)]]?
John hear rumor that-$_{COMP}$ An love who Q-PRT
'Who is x such that John heard a rumor that An loves x?'

b. ??John thích đọc [NP sách $[_{CP}$ An viết vì sao (thế)]]?
John like read book An write why Q-PRT
?? 'What is the reason x such that John likes to read books that An writes because of x?'

(11) Subject islands

a. $[_{NP}$ Bộ phim của ai] đạt giải Oscars (vậy)?
Movie of who get prize Oscars Q-PRT
'Who is x such that the movie of x received Oscars?'

b. *$[_{CP}$An vì sao mất tích] làm gia đình lo lắng (thế)?
An why disappear make family worry Q-PRT
'What is the reason x such that that An disappeared because of x makes her family worried?'

(12) Adjunct islands

a. John vui $[_{PP}$ khi làm gì (vậy)]?
John happy when do what Q-PRT
'What is x such that John is happy when he does x?'

b. *John khóc $[_{PP}$ khi Lan vì sao lấy chồng (thế)]?
John cry when Lan why take husband Q-PRT
'What is the reason x such that John cried when Lan got married because of x?'

As shown in (9a–12a), argument *wh*-phrases can occur within *wh*-, complex NP, subject, and adjunct islands while receiving scope outside these domains. By contrast, the causal *wh*-adjunct *vì sao* 'why' exhibits island sensitivity: sentences are degraded when *vì sao* is interpreted as taking scope outside the relevant island domain.

Previous studies therefore suggest that Vietnamese *wh*-argument questions exhibit little or no sensitivity to canonical island constraints, whereas causal *vì sao* displays substantially stronger locality effects (Bruening and

Trần 2006, Trần 2009, Tsai 2009). Although existing analyses differ concerning the mechanisms that license Vietnamese *wh*-in-situ and the role of sentence-final question particles, they converge on the observation that Vietnamese argument *wh*-dependencies are substantially less restricted by islands than English *wh*-movement. This asymmetry makes Vietnamese particularly informative for L2 acquisition. The relative freedom of argument *wh*-dependencies could lead to L1 transfer, resulting in weaker sensitivity to English island violations, particularly at lower proficiency levels. If UG remains available, however, learners should in principle be able to develop target-like restrictions on English wh-movement despite the absence of directly corresponding restrictions on argument *wh*-dependencies in their L1.

2.6 Research Questions

Given this theoretical and typological background, the current study addresses two research questions (RQ):

RQ1. Are Vietnamese learners of English sensitive to English island constraints that are not robustly instantiated in their L1, and how does this sensitivity compare with that of native speakers across different island types?

RQ2. To what extent can island effects in L2 English be explained by processing factors rather than grammatical constraints?

RQ2 is examined from two complementary perspectives: (a) whether island effects can be accounted for by processing costs associated with dependency length and whether they are attenuated under processing-facilitating conditions, including D-linking and contextual support; and (b) whether individual differences in proficiency and working memory predict learners' sensitivity to island violations. To address these questions, the present study combines factorial acceptability judgment designs with a working-memory measure, focusing on Vietnamese L1–English L2 learners whose L1 provides limited evidence for English-like island restrictions on argument *wh*-dependencies.

3. Study 1

3.1 Research Design

Study 1 addresses the two research questions introduced in Section 2.5 by examining whether Vietnamese learners exhibit sensitivity to English island constraints and by evaluating the relative contributions of grammatical structure and processing factors to these effects using a factorial acceptability judgment design. Specifically, Study 1 investigates the effects of structural constraints, dependency length, and D-linking on the acceptability of English *wh*-dependencies across four island types within a 2 × 2 × 2 factorial design. The findings establish the empirical basis for evaluating competing grammatical and processing-based accounts of island effects. Study 2 builds on these findings by examining island effects under reduced processing demands and incorporating an individual measure of working memory to determine whether proficiency and cognitive resources account for the observed patterns of island sensitivity.

3.2 Participants

Study 1 included four groups of participants. The control group comprised 25 native speakers of English recruited via an online crowdsourcing platform (Amazon Mechanical Turk). The three learner groups together consisted of 115 Vietnamese university students majoring in English or related disciplines. They were categorized into High Advanced (N = 42), Advanced (N = 45), and High Intermediate (N = 28) groups based on a combination of proficiency indicators, including standardized test scores (IELTS) and performance on a 40-item cloze test adapted from O'Neill (1981). The test comprised a roughly 500-word passage with 40 blanks, each accompanied by three response options, and participants received one point for each correct answer (maximum score = 40). The mean age of the native speakers was in the middle twenties, while the Vietnamese learners were mostly in their late teens or early twenties.

On average, the High Advanced group had IELTS scores ranging from 7.5 to 8.5 and achieved the highest cloze scores, close to the ceiling level. The Advanced group scored between 6.0 and 7.0 on the IELTS and slightly lower on the cloze test, whereas the High Intermediate group scored between 5.0 and 5.5 on the IELTS and had the lowest cloze scores. The age of first exposure to English ranged from approximately 5–11 years for the advanced learners and somewhat later for the high intermediate group. The demographic information and proficiency scores for Study 1 are summarized in Table 1 below:

Table 1. Participant Demographics and Proficiency Scores in Study 1

Group	N	Age (M)	Age of first exposure to English (M)	IELTS score range	Cloze score range (out of 40)
Native Speakers	25	27.0	N/A	N/A	N/A
High Advanced	42	19.0	7.1	7.5–8.5	37–40
Advanced	45	19.0	7.2	6.0–7.0	34–36
High Intermediate	28	20.0	8.2	5.0–5.5	<34

3.3 Materials

Building on the framework proposed by Sprouse et al. (2012) for evaluating processing-based accounts of island effects, Study 1 adopted a factorial design to disentangle grammatical violations from processing-related costs. The design targeted two processing factors that can independently lower acceptability: (i) the presence of complex island structures and (ii) long-distance *wh*-dependencies, both of which impose greater processing demands. In contrast to Sprouse et al. (2012), this study introduced a third factor, namely *wh*-phrase complexity (bare vs. D-linked), given that D-linked *wh*-phrases are known to mitigate processing difficulty in island contexts.

Study 1 thus investigated three binary factors: LENGTH (short vs. long [±L]), STRUCTURE (absence vs. presence of island structure [±IS]), and D-LINK (bare vs. D-linked [±DL]), resulting in eight conditions (see Table 2). Each condition was realized with two lexicalizations, yielding a total of 64 test items (8 × 2 × 4) distributed across four island types: *Wh*-Islands, Complex NP Islands, Subject Islands, and Adjunct Islands. While the first two have been extensively examined in processing-based research, the latter two remain comparatively underexplored, thereby enabling a broader test of the processing-based account.

Table 2. Eight Conditions in the Main Test

<table>
<tr><td></td><td></td><td colspan="4">Dependency length</td></tr>
<tr><td></td><td></td><td colspan="2">[-L]</td><td colspan="2">[+L]</td></tr>
<tr><td>Island structure</td><td>[-IS]</td><td>[-IS; -L; -DL]</td><td>[-IS; -L; +DL]</td><td>[-IS; +L; -DL]</td><td>[-IS; +L; +DL]</td></tr>
<tr><td></td><td>[+IS]</td><td>[+IS; -L; -DL]</td><td>[+IS; -L; +DL]</td><td>[+IS; +L; -DL]</td><td>[+IS; +L; +DL]</td></tr>
<tr><td></td><td></td><td>[-DL]</td><td>[+DL]</td><td>[-DL]</td><td>[+DL]</td></tr>
<tr><td></td><td></td><td colspan="4">D-link</td></tr>
</table>

For STRUCTURE, [-IS] conditions lacked island configurations, whereas [+IS] conditions contained the relevant structures: *whether*-clauses (*Wh*-Islands), definite complex NPs containing *that*-clauses (Complex NP Islands), complex subjects (NP + PP) for Subject Islands, and adjunct clauses introduced by *before/after/when* for Adjunct Islands. For LENGTH, [-L] conditions involved extraction from matrix subject position to matrix [Spec, CP], while [+L] conditions involved long extraction from embedded positions to matrix [Spec, CP] (from embedded objects in *Wh*-, Complex NP, and Adjunct Islands; from embedded subjects in Subject Islands). For D-LINK, [-DL] conditions used bare *wh*-words (*who* for subjects, *what* for objects), whereas [+DL] conditions used D-linked phrases (*which* + singular NP, matched to animate/inanimate roles). In lexicalization, matrix verbs were non-factive (e.g., *think, believe*) in [-IS] conditions; in [+IS] conditions, non-factives were used for Subject and Adjunct Islands, *wonder* for *Wh*-Islands, and intransitive verbs for Adjunct Islands. Embedded clauses always contained transitive verbs. Complex NP Islands used nouns such as *lie, claim, news, rumor, fact, story, statement, suspicion, suggestion*, all taking *that*-clausal complements. Main test items are exemplified in (9-12) below:

(13) *Wh*-Islands
Who/which neighbor thought that Bill broke the vase? [-IS; -L; ±DL]
What/which vase did the neighbor think that Bill broke? [-IS; +L; ±DL]
Who/which neighbor wondered whether Bill broke the vase? [+IS; -L; ±DL]
What/which vase did the neighbor wonder whether Bill broke? [+IS; +L; ±DL]

(14) Complex NP Islands
Who/which professor heard that Ann failed the exam? [-IS; -L; ±DL]
What/which exam did the professor hear that Ann failed? [-IS; +L; ±DL]
Who/which professor heard the story that Ann failed the exam? [+IS; -L; ±DL]
What/which exam did the professor hear the story that Ann failed? [+IS; +L; ±DL]

(15) Subject Islands
Who/which teacher believed the student told a lie? [-IS; -L; ±DL]
Who/which student did the teacher believe told a lie? [-IS; +L; ±DL]
Who/which teacher believed the student from the local school told a lie? [+IS; -L; ±DL]
Where/which school did the teacher believe that the student from told a lie? [+IS; +L; ±DL]

(16) Adjunct Islands
Who/which lawyer thought that the client filed a report? [-IS; -L; ±DL]
What/which report did the lawyer think the client filed? [-IS; +L; ±DL]
Who/which lawyer laughed when the client filed a report? [+IS; -L; ±DL]
What/which report did the lawyer laugh when the client filed? [+IS; +L; ±DL]

Alongside the 64 main items, 32 ungrammatical fillers were included to balance grammaticality. Assuming that roughly one quarter of the main items would receive the lowest acceptability ratings (a 3:1 acceptable-to-unacceptable ratio), the addition of these fillers brought the ratio closer to 1:1, thereby minimizing judgment bias and discouraging response strategies (Cowart 1997). The fillers were similar in length to the test items. Half consisted of *tough*-predicates and degree adjectival constructions containing resumptive pronouns, while the other half consisted of relative-clause constructions containing resumptive pronouns or *that*-trace violations. Examples of fillers are provided in Table 3 below.

Table 3. Filler Items Used in Study 1

Structure	Example	Number of items
tough + RP	*The scientist was hard for the students to understand her.*	8
too + adj + RP	*The mountain was too steep to climb it.*	8
RP in matrix subject	*There was a teacher who she believed that the students admired the principal.*	4
RP in embedded subject	*There was a musician who the reporter said that he impressed the audience.*	4
RP in embedded object	*There was a painter who the critic claimed that the gallery rejected him.*	4
that-trace	*There was a boy who the coach said that won the race.*	4

Fillers and test items were distributed across four lists using a Latin square design so that each participant saw each lexicalization in only one experimental condition. Each of the four lists was pseudo-randomized into two presentation orders, resulting in eight survey versions. Acceptability was rated on a 5-point Likert scale (1 = lowest, 5 = highest), with intermediate points left unlabeled.

3.4 Procedures

Study 1 was conducted online via Amazon Mechanical Turk for the native speaker group and face-to-face using pencil-and-paper in quiet classrooms for the three learner groups of Vietnamese EFL learners. The recruited native speakers first provided background information (nationality, residency, age, first and second language, education level), then rated sentence acceptability on a 1–5 Likert scale after completing five practice items; they typically finished within 15–20 minutes, and only completed surveys were recorded and financially rewarded. For the learner groups, informed consent and background questionnaires were collected first, followed by a timed cloze test (maximum 20 minutes). The participants then completed the same acceptability judgment task (instructions provided in Vietnamese) and were encouraged to complete all items. They generally spent about 15 minutes on the cloze test and 30 minutes on the judgment task, and all of them received financial compensation.

3.5 Data Analysis

Data from the three learner groups and the native speaker group were analyzed using a common analytical procedure. First, participants' raw acceptability ratings were z-score transformed to minimize individual scale-use bias (e.g., restricted response ranges or extreme-response tendencies). Standardized ratings were obtained by subtracting each participant's mean rating from each observation and dividing by the participant's standard deviation, yielding scores that represented deviations from individual rating baselines.

To examine the contributions of grammatical structure, processing factors, and learner proficiency to island sensitivity, the z-transformed ratings were analyzed using linear mixed-effects models (LMEs). Separate models were fitted for each island type using the *lme4* package (Bates et al. 2015) in R (R Core Team 2024). Fixed effects included STRUCTURE (Island vs. Non-Island), LENGTH (Short vs. Long), D-LINK (Present vs. Absent), GROUP (Native, High Advanced, Advanced, and High Intermediate), and their interactions. Participants and items were included as crossed random effects. Following Barr et al. (2013), maximal random-effects structures justified by the design were initially specified and simplified where necessary in response to convergence or singularity diagnostics. The final model was: Rating ~ STRUCTURE × LENGTH × D-LINK × GROUP + (1 + STRUCTURE × LENGTH × D-LINK | Participant) + (1 + GROUP | Item). Statistical significance of fixed effects was evaluated using Type III analysis of variance with Satterthwaite's approximation, implemented in the *lmerTest* package (Kuznetsova et al. 2017). Particular attention was paid to the STRUCTURE × LENGTH interaction, which constitutes the critical test of the superadditive island effect.

Difference-in-differences (DD) scores were also computed for each participant and island type following Sprouse et al. (2012). Specifically, D1 = [−IS; +L] − [+IS; +L] quantified the STRUCTURE effect under long dependencies, D2 = [−IS; −L] − [+IS; −L] quantified the STRUCTURE effect under short dependencies, and DD = D1 − D2 represented the superadditive island effect. DD scores were computed separately for the bare and D-linked conditions. Positive DD scores indicate a superadditive island penalty, with larger positive values reflecting stronger island effects. DD scores served as participant-level measures of island sensitivity and were used both descriptively and in complementary inferential analyses comparing the magnitude of island effects across groups and island types.

All analyses were conducted in R (Version 4.4.1; R Core Team 2024) using the packages *lme4* (Bates et al. 2015), *lmerTest* (Kuznetsova et al. 2017), *emmeans* (Lenth 2024), and *broom.mixed* (Bolker and Robinson 2024).

3.6 Results

3.6.1 General patterns

The mean acceptability ratings for each experimental condition across the four island types, along with the mean difference-in-differences (DD) scores for each group, are presented in Table 4. In this table, the asterisk (*) marks the critical [+IS; +L] conditions, in which a long-distance dependency crosses an island boundary. Native speakers and the High Advanced group pattern similarly across all island types and both bare and D-linked conditions, with low ratings for [+IS; +L] and positive DD scores (0.39–1.58), indicating a clear descriptive pattern of superadditivity. Effects are largest for Complex NP, Subject, and Adjunct Islands and smallest for *Wh*-Islands. The Advanced group shows smaller but generally positive island sensitivity, most evident for Subject, Complex NP, and D-link Adjunct Islands, whereas the High Intermediate group shows the least consistent pattern: DD scores are near zero or negative in several conditions, although positive effects remain for D-linked Complex NP and Adjunct Islands. Across groups, D-linked *wh*-phrases generally increased acceptability in the critical [+IS; +L] conditions relative to bare *wh*-phrases, although this amelioration in raw ratings did not necessarily correspond to smaller DD effects.

To visualize the overall proficiency and island-type patterns, Figure 1 plots mean DD scores collapsed across D-LINK; the effects involving D-LINK are examined separately in the mixed-effects analyses below. The figure shows a clear proficiency-related increase in DD scores. Complex NP and Subject Islands generally produced the largest DD effects, whereas *Wh*-islands tended to produce smaller effects. Participant-level DD scores were analyzed using a 4 × 4 mixed-design ANOVA, with GROUP as a between-participants factor and ISLAND TYPE as a within-participants factor. Because the assumption of sphericity was violated, Greenhouse–Geisser-corrected degrees of freedom are reported. There were significant main effects of GROUP, $F(3, 136) = 34.21$, $p < .001$, $\eta p^2 = .43$, and ISLAND TYPE, $F(2.68, 364.87) = 8.45$, $p < .001$, $\eta p^2 = .06$, but no significant GROUP × ISLAND TYPE interaction, $F(8.05, 364.87) = 1.06$, $p = .388$. Holm-adjusted comparisons showed that Native speakers and High Advanced learners had similar DD scores, and both groups scored significantly higher than the Advanced and High Intermediate groups. The Advanced group also scored higher than the High Intermediate group. Across groups, Complex NP and Subject Islands produced larger DD scores than *Wh*-Islands, whereas Adjunct Islands did not differ significantly from *Wh*-Islands.

Table 4. Mean z-scores and DD Scores for all Conditions and Island Types

Group	Condition	*Wh-*		Complex NP		Subject		Adjunct	
		[-DL]	[+DL]	[-DL]	[+DL]	[-DL]	[+DL]	[-DL]	[+DL]
Native Speakers (N=25)	[-IS; -L]	0.86	0.80	0.75	0.64	0.64	0.63	0.64	0.59
	[-IS; +L]	0.36	0.56	0.24	0.55	0.22	0.29	0.39	0.73
	[+IS; -L]	0.27	0.29	0.81	0.76	0.60	0.44	0.49	0.69
	*[+IS; +L]	-1.10	-0.33	-1.27	-0.91	-1.27	-1.17	-0.86	-0.46
	DD score	0.88	0.39	1.58	1.58	1.44	1.01	1.10	1.29
High Advanced (N=42)	[-IS; -L]	0.87	0.67	0.75	0.64	0.74	0.60	0.68	0.59
	[-IS; +L]	0.44	0.55	-0.08	0.44	0.15	0.30	0.04	0.58
	[+IS; -L]	0.46	0.47	0.74	0.75	0.61	0.55	0.53	0.67
	*[+IS; +L]	-0.98	-0.24	-1.18	-0.75	-0.99	-0.97	-1.09	-0.49
	DD score	1.01	0.59	1.09	1.30	1.01	1.23	0.98	1.15
Advanced (N=45)	[-IS; -L]	0.85	0.69	0.72	0.62	0.61	0.60	0.62	0.31
	[-IS; +L]	0.15	0.62	-0.47	0.26	-0.38	-0.19	-0.22	0.57
	[+IS; -L]	0.44	0.17	0.74	0.60	0.47	0.39	0.35	0.52
	*[+IS; +L]	-0.66	-0.14	-0.87	-0.52	-1.19	-0.95	-0.66	-0.19
	DD score	0.40	0.24	0.43	0.76	0.66	0.54	0.17	0.96
High Intermediate (N=28)	[-IS; -L]	0.85	0.48	0.60	0.50	0.77	0.49	0.70	0.40
	[-IS; +L]	0.08	0.51	-0.48	0.49	-0.57	-0.04	-0.38	0.50
	[+IS; -L]	0.55	0.16	0.64	0.38	0.50	0.22	0.25	0.41
	*[+IS; +L]	-0.34	0.18	-0.61	-0.34	-0.79	-0.43	-0.61	0.05
	DD score	0.11	0.00	0.17	0.71	-0.05	0.12	-0.23	0.46

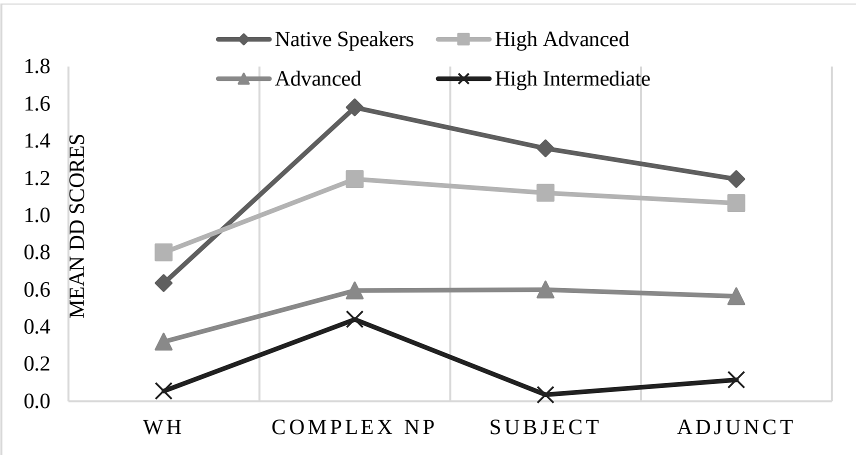

Figure 1. Mean DD Scores Across Groups and Island Types

3.6.2 Mixed-effects analyses

To evaluate the contributions of grammatical structure, processing factors, and learner proficiency to island sensitivity, separate linear mixed-effects models were fitted for each island type. Fixed effects included STRUCTURE (Island vs. Non-Island), LENGTH (Short vs. Long), D-LINK, GROUP (Native Speakers, High Advanced, Advanced, and High Intermediate), and their interactions, with participants and items included as crossed random effects. Because the critical prediction of the study concerns the existence of island effects, particular attention was paid to the STRUCTURE × LENGTH interaction, which corresponds to the superadditive island effect measured by the DD score.

Across all four island types, the mixed-effects models revealed highly significant main effects of STRUCTURE and LENGTH (all ps < .001), indicating that island configurations and longer dependencies independently reduced acceptability. The critical STRUCTURE × LENGTH interaction was significant for every island type (*Wh*-: $\beta = -0.113$, SE = 0.017, $F(1, 17.05) = 43.14$, $p < .001$; Complex NP: $\beta = -0.238$, SE = 0.027, $F(1, 16.51) = 79.18$, $p < .001$; Subject: $\beta = -0.195$, SE = 0.022, $F(1, 17.09) = 81.11$, $p < .001$; Adjunct: $\beta = -0.184$, SE = 0.027, $F(1, 16.78) = 46.07$, $p < .001$). These interactions demonstrate robust overall superadditive island effects beyond the additive effects of STRUCTURE and LENGTH.

The models also demonstrated consistent effects of D-linking. Significant LENGTH × D-LINK interactions were observed for all four island types, indicating that that D-linking modulated the acceptability cost associated with longer dependencies. More importantly, the STRUCTURE × LENGTH × D-LINK interaction, which directly tests whether D-linking modulated the superadditive island effect, was significant for *Wh*-Islands, $F(1, 17.04) = 4.63$, $p = .046$, but not for Complex NP islands, $F(1, 16.63) = 1.60$, $p = .223$, or Subject islands, $F(1, 16.82) = 0.02$, $p = .889$. For adjunct islands, the interaction was borderline, $F(1, 16.79) = 4.48$, $p = .050$. Thus, although D-linking systematically affected the acceptability of long-distance dependencies, there was little evidence that it attenuated the superadditive island effect for the stronger island types. Its effect on superadditivity was clearest for *Wh*-Islands.

With respect to learner development, the main effect of Group reached significance only for Subject Islands ($F(3, 139.92) = 4.47$, $p = .005$), indicating overall differences in acceptability ratings across proficiency groups for this island type. More importantly, significant STRUCTURE × GROUP interactions were observed for all four island types (*Wh*-: $F(3, 139.96) = 6.30$, $p < .001$; Complex NP: $F(3, 139.56) = 9.11$, $p < .001$; Subject: $F(3, 139.90) = 5.73$, $p = .001$; Adjunct: $F(3, 139.65) = 5.62$, $p = .001$), demonstrating that the effect of island structure differed systematically across learner groups. Critically, all four models yielded significant STRUCTURE × LENGTH × GROUP interactions (*Wh*-: $F(3, 1662.57) = 7.90$, $p < .001$; Complex NP: $F(3, 1665.59) = 18.17$, $p < .001$; Subject: $F(3, 1943.87) = 16.56$, $p < .001$; Adjunct: $F(3, 1944.35) = 11.79$, $p < .001$), indicating that the superadditive island effect varied significantly as a function of learner group. These interactions reflect a clear developmental pattern, with island sensitivity progressively increasing from the High Intermediate to the Advanced and High Advanced groups, approaching native-speaker performance.

In summary, Study 1 yielded four principal findings. First, the mixed-effects analyses demonstrated robust superadditive island effects across all four island types, as evidenced by significant STRUCTURE × LENGTH interactions. Second, D-linking generally facilitated the processing of long-distance dependencies, although its ameliorating effect on island violations was limited to *Wh*-Islands. Third, learner proficiency significantly modulated island sensitivity, with significant STRUCTURE × LENGTH × GROUP interactions observed for every island type. Finally, the DD analyses showed that High Advanced learners approached native-speaker performance, whereas High Intermediate learners exhibited substantially weaker island effects, particularly for *Wh*-Islands. Together, these findings suggest that Vietnamese learners progressively acquire grammatical sensitivity to English island constraints, while processing factors modulate, but do not determine, their acceptability judgments.

4. Study 2

Building on the findings of Study 1, Study 2 further investigates the same research questions under conditions designed to minimize processing demands. Specifically, by incorporating D-linked *wh*-phrases, supportive discourse contexts, and an individual measure of working memory, Study 2 examines whether the island effects observed in Study 1 remain robust when processing difficulty is reduced and whether learner proficiency and working memory explain developmental variation in island sensitivity.

4.1 Participants

Participants included 20 native English speakers recruited from an expatriate community in Hanoi and 99 Vietnamese EFL learners from several universities in Hanoi. Selection criteria were identical to Study 1, and the same cloze test was administered to provide an additional measure of L2 proficiency. Participant information is summarized in Table 5 below.

Table 5. Participant Demographics and Proficiency Scores in Study 2

Group	N	Age (M)	Age of first exposure to English (M)	IELTS score range	Cloze score range (out of 40)
Native Speakers	20	38.0	N/A	N/A	N/A
High Advanced	42	21.0	8.0	7.5–8.5	37–40
Advanced	37	21.0	8.0	6.0–7.0	34–36
High Intermediate	20	19.0	8.1	5.0–5.5	< 34

4.2 Stimuli Design, Procedures and Data Analysis

Study 2 combined a formal acceptability judgment task with a working memory task. The judgment task largely followed Study 1 but incorporated modifications intended to reduce processing difficulty, motivated by previous findings concerning D-linking, discourse context, and working memory (Hofmeister and Sag 2010, Hofmeister et al. 2012). First, each *wh*-question was preceded by a declarative background sentence to provide discourse context and avoid pragmatic oddity. For example, an island-violating question such as *Which car do you wonder whether John bought?* was paired with a background like *I wonder whether John bought a car*. The context introduced the relevant discourse referent and provided a supportive discourse context for the target question (Hofmeister and Sag 2010, Kluender 1998). Second, only D-linked *which*-NP phrases were used instead of bare *wh*-words, as D-linking narrows referential possibilities and facilitates processing (Hofmeister et al. 2012; Pesetsky 1987).

Study 2 thus employed a factorial design crossing STRUCTURE and LENGTH (two levels each) across four island types (*Wh*-, Complex NP, Subject, Adjunct), yielding four conditions per island. With four lexicalizations per condition, this produced 64 main test items, along with 32 ungrammatical fillers similar to Study 1. Survey construction otherwise followed the same procedures as in Study 1.

Working memory was measured using a computerized version of the Operation Span task developed by Unsworth et al. (2005), implemented in OpenSesame (Mathôt et al. 2012). This task was chosen over the serial-recall and n-back measures used by Sprouse et al. (2012) for two reasons. First, serial-recall and n-back tasks lack a secondary processing component and therefore measure short-term memory rather than working memory, whereas complex span tasks interleave storage with processing demands (Hofmeister et al. 2012; Kane et al. 2004). Second, evidence shows that the Operation Span task better predicts comprehension of syntactically complex sentences; for example, O'Rourke (2013) found that only Operation Span scores correlated with comprehension accuracy for garden-path sentences.

In the Operation Span task, participants solved a math problem (e.g., $2 \times 5 + 3 = ?$), judged whether a displayed number was correct, and then memorized a letter. After sequences of three to seven operation–letter pairs, they recalled the letters in order. Each participant completed three trials at each set size (3–7 items), producing 75 letters and 75 math operations, with five practice trials provided beforehand.

Both native speakers and learner groups followed the same procedure. Participants first completed the acceptability judgment task on paper, reading the declarative background sentence for context but rating only the test sentence. They then performed the Operation Span task on a computer. Before the main task, participants completed three practice sessions: letter recall only, math operations only, and the combined task. During the real task, participants had limited time for each math problem and were required to maintain at least 80% accuracy; math accuracy was continuously displayed to discourage strategic trade-offs. Scores were automatically calculated at the end.

Judgment data were analyzed as in Study 1. The Operation Span task yielded an absolute score (the sum of the set sizes for perfectly recalled sets), a letter score (the total number of letters recalled in the correct serial position), and math-error measures. Participants with more than 20% math errors were excluded. To examine whether individual working-memory capacity predicted island sensitivity, DD scores were regressed on OSPAN performance. Separate simple linear regression models were fitted for the absolute and letter scores for each participant group and island type using R (Version 4.4.1; R Core Team 2024).

4.3 Results

4.3.1 Basic island effects under a reduced processing load

Table 6 summarizes the mean z-transformed acceptability ratings and DD scores for Study 2. All four structural conditions contained D-linked, contextually supported *wh*-dependencies: non-island short ([−IS; −L]), non-island long ([−IS; +L]), island short ([+IS; −L]), and island long ([+IS; +L]).

Table 6. Mean z-Scores (SDs) and DD Scores by Group and Condition

Group	Condition	*Wh*-	Complex NP	Subject	Adjunct
Native Speakers (N=20)	[-IS; -L]	0.94 (.36)	0.96 (.34)	0.84 (.47)	0.85 (.45)
	[-IS; +L]	1.03 (.34)	0.85 (.59)	0.64 (.46)	0.85 (.57)
	[+IS; -L]	0.86 (.30)	0.92 (.36)	0.72 (.56)	0.72 (.61)
	*[+IS; +L]	-0.18 (.52)	-0.58 (.72)	-0.91 (.58)	-0.90 (.59)
	DD scores	1.12	1.39	1.43	1.62
High Advanced (N=42)	[-IS; -L]	0.80 (.48)	0.91 (.34)	0.71 (.56)	0.76 (.50)
	[-IS; +L]	0.83 (.48)	0.67 (.59)	-0.01 (.92)	0.64 (.69)
	[+IS; -L]	0.70 (.46)	0.80 (.46)	0.62 (.61)	0.67 (.59)
	*[+IS; +L]	0.38 (.73)	-0.40 (.86)	-0.91 (.59)	-0.80 (.74)
	DD scores	0.35	0.95	0.82	1.35
Advanced (N=37)	[-IS; -L]	0.45 (.70)	0.53 (.63)	0.37 (.69)	0.40 (.59)
	[-IS; +L]	0.63 (.61)	0.45 (.72)	-0.28 (.97)	0.37 (.87)
	[+IS; -L]	0.38 (.68)	0.43 (.65)	0.40 (.68)	0.52 (.65)
	*[+IS; +L]	0.27 (.84)	-0.29 (1.0)	-0.85 (.89)	-0.57 (.99)
	DD scores	0.28	0.63	0.60	1.06
High Intermediate (N=20)	[-IS; -L]	0.27 (.74)	0.37 (.85)	0.14 (.72)	0.22 (.81)
	[-IS; +L]	0.45 (.86)	0.53 (.55)	-0.23 (.91)	0.35 (.84)
	[+IS; -L]	-0.06 (.80)	0.26 (.85)	0.09 (.72)	0.16 (.75)
	*[+IS; +L]	0.17 (.80)	0.04 (.93)	-0.44 (.88)	-0.09 (.97)
	DD scores	-0.05	0.38	0.17	0.38

Despite these processing-facilitating conditions, positive DD effects persisted for most island types and groups. Native speakers showed substantial positive DD scores across all four island types (*Wh*- = 1.12, Complex NP =

1.39, Subject = 1.43, Adjunct = 1.62). High Advanced and Advanced learners also showed positive DD effects throughout, although these were generally smaller than those of native speakers. The High Intermediate group exhibited the weakest pattern, with relatively small DD scores for Complex NP (.38), Subject (.17), and Adjunct Islands (.38) and essentially no *Wh*-Island effect (−.05). Comparison with the D-linked conditions of Study 1 revealed no uniform reduction in DD scores in Study 2: effects were numerically smaller for some learner groups and island types but comparable or larger in others.

As in Study 1, participant-level DD scores were analyzed using a 4 × 4 mixed-design ANOVA, with GROUP as a between-participants factor and ISLAND TYPE as a within-participants factor. The analysis revealed significant main effects of GROUP, $F(3, 117) = 37.72$, $p < .001$, and ISLAND TYPE, $F(2.81, 329.22) = 21.54$, $p < .001$. The GROUP × ISLAND TYPE interaction was not significant, $F(8.44, 329.22) = 1.24$, $p = .275$, indicating that the pattern of differences among island types did not vary reliably across groups.

Separate linear mixed-effects models were fitted for each island type. The critical STRUCTURE × LENGTH interaction was significant for every island type (*Wh*-: $\beta = -0.107$, SE = 0.020, $F(1, 20.49) = 28.07$, $p < .001$; Complex NP: $\beta = -0.189$, SE = 0.027, $F(1, 24.76) = 48.40$, $p < .001$; Subject: $\beta = -0.210$, SE = 0.024, $F(1, 21.05) = 77.86$, $p < .001$; Adjunct: $\beta = -0.276$, SE = 0.026, $F(1, 16.73) = 109.74$, $p < .001$), demonstrating robust overall superadditive island effects under the processing-facilitating conditions of Study 2. Main effects of STRUCTURE were significant for every island type, whereas LENGTH significantly affected *Wh*-, Complex NP-, and Subject Islands (all $ps < .001$) but not Adjunct Islands ($p = .103$).

Concerning whether these island effects changed as learner proficiency increased, significant main effects of GROUP indicated overall proficiency-related differences in acceptability ratings across all four island types. More importantly, significant STRUCTURE × GROUP interactions (*Wh*-: $F(3, 120.89) = 8.18$, $p < .001$; Complex NP: $F(3, 120.78) = 5.72$, $p = .001$; Subject: $F(3, 120.78) = 8.67$, $p < .001$; Adjunct: $F(3, 120.70) = 5.53$, $p = .001$) demonstrated that the effect of island structure differed systematically across proficiency groups. Critically, the STRUCTURE × LENGTH × GROUP interaction was significant for all four island types (*Wh*-: $F(3, 120.73) = 13.16$, $p < .001$; Complex NP: $F(3, 120.75) = 6.72$, $p < .001$; Subject: $F(3, 120.72) = 7.14$, $p < .001$; Adjunct: $F(3, 1558.55) = 16.42$, $p < .001$), indicating that the superadditive island effect varied reliably with learner proficiency. Consistent with the DD-score analyses, these interactions show a clear developmental trajectory, with island sensitivity increasing progressively from the High Intermediate to the Advanced and High Advanced groups and approaching the native-speaker pattern.

4.3.2 Working memory and island sensitivity

To examine whether individual working-memory capacity predicted island sensitivity, separate linear regression models were fitted with participant-level DD scores as the dependent variable and Operation Span performance as the predictor for each group and island type. The primary analyses used the absolute OSPAN score, with parallel analyses using the letter score as a robustness check. Since 16 regressions were conducted for each OSPAN measure (four groups × four island types), Holm-adjusted *p*-values were additionally calculated to control for multiple comparisons. The regression results are summarized in Table 7.

The analyses provided little evidence that working-memory capacity systematically predicted island sensitivity. For the absolute OSPAN score, 15 of the 16 relationships were nonsignificant even before correction (all $ps \geq .071$). The only nominally significant association occurred for Adjunct Islands in the High Intermediate group, $\beta = .043$, SE = .016, 95% CI [.010, .075], $t(18) = 2.74$, $p = .013$, $R^2 = .294$. However, this association did not survive Holm correction (p adjusted = .216). No other relationship approached significance, and effect directions varied across groups and island types.

Table 7. Regression Results for DD Scores and Operation Span

Group	Island type	β	SE	95% CI	*t* value	*p* value	R^2
Native speaker (N=20)	*Wh-*	0.0001	.0111	[−.0232, .0234]	0.01	.993	.000
	Complex NP	0.0133	.0070	[−.0013, .0280]	1.92	.071	.170
	Subject	−0.0005	.0105	[−.0226, .0216]	−0.05	.960	.000
	Adjunct	0.0022	.0114	[−.0218, .0263]	0.20	.846	.002
High Advanced (N=42)	*Wh-*	0.0047	.0062	[−.0078, .0172]	0.76	.450	.014
	Complex NP	0.0081	.0081	[−.0082, .0244]	1.00	.322	.025
	Subject	−0.0125	.0096	[−.0319, .0069]	−1.30	.201	.041
	Adjunct	0.0058	.0087	[−.0118, .0233]	0.66	.510	.011
Advanced (N=37)	*Wh-*	−0.0010	.0103	[−.0219, .0198]	−0.10	.920	.000
	Complex NP	−0.0010	.0118	[−.0249, .0228]	−0.09	.930	.000
	Subject	0.0074	.0133	[−.0195, .0344]	0.56	.578	.009
	Adjunct	−0.0092	.0126	[−.0349, .0164]	−0.73	.471	.015
High Intermediate (N=20)	*Wh-*	−0.0082	.0143	[−.0382, .0218]	−0.58	.572	.018
	Complex NP	−0.0166	.0173	[−.0530, .0199]	−0.96	.352	.048
	Subject	−0.0250	.0199	[−.0668, .0169]	−1.25	.226	.080
	Adjunct	0.0426	.0156	[.0099, .0753]	2.74	.0135*	.294

Analyses using the letter score produced the same overall pattern. Again, the only nominally significant relationship occurred for Adjunct Islands among High Intermediate learners, β = .086, SE = .029, 95% CI [.026, .147], $t(18) = 2.99$, $p = .008$, $R^2 = .332$, but this effect likewise did not survive Holm correction (p adjusted = .126). No other letter-score regression was significant.

Overall, neither measure of Operation Span provided reliable evidence that individual working-memory capacity predicted the magnitude of island effects. The absence of systematic relationships across four island types and four participant groups suggests that the substantial DD effects observed in Study 2 cannot be straightforwardly attributed to individual differences in working-memory capacity.

5. Discussion

The present study addressed two overarching questions concerning the acquisition of English island constraints by Vietnamese learners: whether learners exhibit sensitivity to island violations, and the extent to which such sensitivity is shaped by grammatical constraints, processing factors, and interlanguage development. The following discussion considers these questions in turn.

5.1 Sensitivity to English Island Violations

The first research question asked whether Vietnamese learners of English are sensitive to English island violations and how their judgments compare with those of native speakers across different island types. Across both studies, the results provide clear evidence of systematic island sensitivity in the learner groups, but also reveal a strong proficiency-related developmental pattern. High Intermediate learners generally showed small or inconsistent superadditive effects, whereas these effects became larger at higher proficiency levels. High Advanced learners showed substantially more target-like patterns and, in Study 1, did not differ significantly from native speakers in the overall DD analysis. Thus, sensitivity to the interaction between dependency length and island structure becomes increasingly robust as L2 proficiency increases.

The results further revealed differences across island types. Complex NP and Subject Islands generally elicited robust island effects, particularly among the higher-proficiency learners, whereas *Wh*-Islands showed weaker effects and greater susceptibility to ameliorating factors such as D-linking. Adjunct Islands also showed substantial effects, although their magnitude varied across groups and experimental conditions. This pattern is broadly consistent with previous findings that the strength of island effects varies across island types (e.g., Huang 1982; Sprouse et al. 2012). Nevertheless, advanced learners consistently distinguished island-violating from grammatical dependencies across island types, indicating that increased acceptability under facilitative conditions does not eliminate sensitivity to structural constraints.

Taken together, these findings demonstrate that Vietnamese learners progressively acquire sensitivity to English island constraints despite substantial typological differences between English and Vietnamese. This developmental pattern provides the empirical basis for evaluating the theoretical accounts of interlanguage grammar discussed in the following subsections.

5.2 Processing Factors and Limits of Reduction

We now turn to the second research question: whether these effects can be explained solely by processing factors or instead reflect grammatical constraints represented in learners' interlanguage. The two studies provide converging evidence that processing contributes substantially to the acceptability of syntactic island configurations, but that it cannot by itself account for the full pattern of island effects. Consistent with processing-based accounts (e.g., Hofmeister and Sag 2010, Kluender 1998, 2004, Kluender and Kutas 1993), longer dependencies were generally judged less acceptable than shorter ones, and manipulations designed to reduce processing difficulty, including D-linked *wh*-phrases and supportive discourse contexts, affected the acceptability of long-distance dependencies. In Study 1, D-linking attenuated the superadditive island effect for *Wh*-Islands but not for Complex NP or Subject Islands; for Adjunct Islands, the observed modulation was not ameliorative in direction. Moreover, the processing-facilitating conditions of Study 2 did not produce a uniform reduction in DD scores relative to Study 1. These findings confirm that processing demands influence acceptability while also showing that their effects on superadditive island sensitivity vary across island types.

More importantly, several findings indicate that processing alone is insufficient to explain the observed patterns. First, robust island effects remained for Complex NP and Subject Islands even under the highly facilitative conditions of Study 2. Despite minimizing processing demands through D-linking and contextual support, these island types continued to exhibit large positive DD scores and reliable STRUCTURE × LENGTH interactions in both native speakers and high-proficiency learners. If island effects merely reflected the cumulative costs of dependency length, embedding, and memory retrieval, these superadditive effects would be expected to disappear or substantially diminish once processing demands were reduced (Hofmeister and Sag 2010, Kluender 1998). Their persistence instead suggests that these island types are constrained by grammatical principles beyond general processing difficulty.

Second, the working memory results provide little support for a simple resource-limitation account. Under capacity-based approaches, individuals with greater working memory resources are expected to exhibit smaller island penalties (e.g., Aldosari 2015, Baddeley 1992, Kluender 1998). In the present study, however, operation span scores showed no consistent relationship with island sensitivity across learner groups or island types. Of the 16 analyses using absolute operation span scores, only one yielded a nominally significant relationship, and this effect did not survive correction for multiple comparisons. Parallel analyses using letter scores produced the same overall pattern. These results provide no reliable evidence that individual differences in general working memory capacity predict learners' sensitivity to island violations.

Finally, the relative ordering of island types showed some consistency across studies and proficiency levels, with Complex NP and Subject Islands generally producing robust effects and *Wh*-Islands tending to produce weaker effects. Adjunct Islands showed substantial effects but did not occupy a uniformly intermediate position across groups and conditions. This variation is broadly consistent with previous experimental findings that island effects differ in magnitude across structural configurations (Alexopoulou and Keller 2013, Almeida 2014, Sprouse et al. 2012).

Taken together, the findings indicate that processing factors alone cannot fully account for island effects. Processing pressures clearly influence the ease with which long-distance *wh*-dependencies are comprehended and evaluated, thereby modulating the severity of island violations and explaining why weaker islands are particularly susceptible to D-linking and contextual support. However, processing does not determine which syntactic configurations constitute islands in the first place. Rather, the persistence of robust superadditive effects under facilitative conditions points to underlying grammatical constraints on *wh*-dependencies, with processing influencing the degree to which those constraints are manifested during comprehension and judgment (Almeida 2014, Chomsky 1973, 1977, Huang 1982, Ross 1967, Sprouse et al. 2012).

5.3 The Nature of Island Effects in Interlanguage Grammar

The findings from the previous two sections raise a broader theoretical question concerning the nature of island effects in interlanguage grammar. Across both studies, island sensitivity increased systematically with learner proficiency, whereas individual differences in working memory showed no reliable relationship with learners' island sensitivity. Combined with the persistence of robust island effects under facilitative processing conditions, these findings are consistent with grammatical restructuring in learners' interlanguage rather than an account based solely on limitations in general cognitive resources. They further suggest that L1 influence does not determine learners' ultimate grammatical representations. These findings therefore provide a basis for evaluating two influential accounts of adult L2 acquisition discussed in Section 2.3, namely the FT/FA Hypothesis (Schwartz and Sprouse 1996) and the FDH (Bley-Vroman 1990, 2009), which make contrasting predictions regarding the acquisition of grammatical properties that are underdetermined by the input, including syntactic island constraints.

The FT/FA Hypothesis predicts that although the initial state of interlanguage reflects the learner's L1 grammar (Full Transfer), continued access to Universal Grammar enables learners to restructure their grammars when the target language requires properties absent from the L1 (Full Access). Applied to the present study, this account predicts that Vietnamese learners, whose L1 exhibits comparatively weak island effects and lacks overt *wh*-movement (Bruening and Trần 2006, Tsai 2009), should initially show limited sensitivity to English island constraints but gradually acquire native-like representations as proficiency increases. The present findings closely match these predictions. Across both studies, island sensitivity increased systematically with proficiency, and the High Advanced learners exhibited DD scores and STRUCTURE × LENGTH interactions that closely paralleled those of native speakers. Moreover, robust island effects for Complex NP and Subject Islands persisted even under the facilitative conditions of Study 2, indicating that learners had internalized structural restrictions on *wh*-dependencies rather than relying solely on processing strategies. Taken together, these findings suggest that English island constraints become progressively represented in learners' interlanguage grammars despite their comparatively weak manifestation in Vietnamese.

By contrast, the FDH attributes adult L2 acquisition primarily to domain-general learning mechanisms operating over previously acquired linguistic knowledge, without continued access to UG (Bley-Vroman 1990, 2009). Under this account, learners whose L1 provides little evidence for English-style island constraints are expected to experience persistent difficulty acquiring these abstract grammatical restrictions, particularly because island constraints are underdetermined by the linguistic input and rarely receive explicit instruction. The present findings are less readily reconciled with this prediction. Although one could argue that advanced learners approximate native-like performance through accumulated exposure or distributional learning, it remains difficult to explain why Vietnamese learners converge on the same hierarchy of island strength observed in native speakers and continue to exhibit robust superadditive island effects despite the reduction of processing demands. Similar findings have been reported in previous studies of L2 island phenomena and have likewise been interpreted as consistent with continued UG accessibility (e.g., Arnon et al. 2005, Bardovi-Harlig and Sprouse 2018, Belikova and White 2009).

Taken together, the present findings are more consistent with the predictions of the FT/FA Hypothesis than with those of the FDH. While they do not by themselves constitute definitive evidence for continued access to UG, they suggest that adult learners can acquire abstract movement constraints that extend beyond those instantiated in their first language. In conjunction with the findings of Section 5.1, the results support the view that island effects reflect UG-constrained grammatical representations whose manifestation is modulated, but not entirely determined, by processing factors.

5.4 Remaining Issues and Future Directions

Despite the strong support for UG-constrained island representations in Vietnamese interlanguage grammar, the present findings also point to several directions for future research. First, although the cross-sectional data suggest that lower-proficiency learners initially rely more heavily on L1 representations before gradually converging on target-like island constraints, longitudinal studies are needed to determine how this developmental trajectory unfolds and how L1 transfer and UG interact over time (Hawkins and Chan 1997, White 2003). Second, the absence of robust correlations between working memory and island sensitivity does not imply that processing resources are unimportant. Rather, global operation span measures may be too coarse to capture the online mechanisms underlying island processing. Future research combining acceptability judgments with online methods such as self-paced reading, eye tracking, or EEG could provide a more fine-grained account of how processing interacts with grammatical constraints (Kluender and Kutas 1993, Wagers and Phillips 2009). Finally, cross-linguistic comparisons involving other *wh*-in-situ languages (e.g., Chinese and Korean) would help determine whether the developmental trajectory observed here generalizes beyond Vietnamese, thereby clarifying the respective roles of UG, L1 transfer, and processing in the acquisition of syntactic locality constraints (Huang 1982; Tsai 2009).

6. Conclusion

This article has examined island effects in the interlanguage grammar of Vietnamese learners of English through two experiments combining factorial acceptability judgment tasks with a working memory measure. The study addressed two central questions: whether Vietnamese learners acquire sensitivity to English island constraints that are only weakly instantiated in their first language, and to what extent island effects in L2 English reflect grammatical constraints and processing factors.

The findings demonstrate that Vietnamese learners are indeed sensitive to English island violations and that this sensitivity increases systematically with proficiency, approaching native-speaker performance at higher proficiency levels. Across both studies, the classic hierarchy of island strength was replicated, and robust superadditive island effects were observed across four island types. Processing manipulations, such as D-linking and supportive contexts, modulated the severity of violations but did not eliminate island penalties, and individual working memory capacity did not predict island sensitivity. These results support a view of island phenomena as arising from the interaction of UG-constrained grammatical principles with general processing limitations. From an SLA perspective, the developmental trajectory observed across proficiency levels is more consistent with the predictions of the FT/FA hypothesis than with the FDH, suggesting that adult learners can acquire abstract movement constraints despite limited support from their first language.

More broadly, the Vietnamese case highlights the importance of combining formal experimental methods with careful typological and theoretical considerations in SLA research. Island effects, far from being a purely esoteric topic in generative syntax, offer a powerful lens on the interplay between grammar, processing, and acquisition, and they continue to illuminate the nature of interlanguage grammar in multilingual speakers.

References

Aldosari, S. 2015. *The Role of Individual Differences in the Acceptability of Island Violations in Native and Non-native Speakers.* Doctoral dissertation, University of Kansas, Lawrence, KS, USA.

Aldosari, S., L. Covey and A. Gabriele. 2024. Examining the source of island effects in native speakers and second language learners of English. *Second Language Research* 40(1), 51-77.

Aldwayan, S., R. Fiorentino and A. Gabriele. 2010. Evidence of syntactic constraints in the processing of *wh*-movement: A study of Najdi Arabic learners of English. In B. VanPatten and J. Jegerski, eds., *Research in Second Language Processing and Parsing*, 65-86. Amsterdam: John Benjamins.

Alexopoulou, T. and F. Keller. 2013. *What* vs. *Who* and *Which*: Kind-denoting fillers and the complexity of *whether*-islands. In J. Sprouse and N. Hornstein, eds., *Experimental Syntax and Island Effects*, 310-340. Cambridge: Cambridge University Press.

Almeida, D. 2014. Subliminal *Wh*-Islands in Brazilian Portuguese and the consequences for syntactic theory. *Revista da ABRALIN* 13(2), 55-93.

Arnon, I., B. Estigarribia, P. Hofmeister, F. Jaeger, J. Pettibone, I. A. Sag and N. Snider. 2005. Long-distance dependencies without island constraints. Poster presented at *HOWL 3: Hopkins Workshop on Language*.

Baddeley, A. 1992. Working memory. *Science* 255(5044), 556-559.

Bardovi-Harlig, K. and R. A. Sprouse. 2018. Negative versus positive transfer. In J. I. Liontas, ed., *The TESOL Encyclopedia of English Language Teaching*, 1-6. Hoboken, NJ: Wiley Blackwell.

Barr, D. J., R. Levy, C. Scheepers and H. J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3), 255-278.

Bates, D., M. Mächler, B. Bolker and S. Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67(1), 1-48.

Belikova, A. and L. White. 2009. Evidence for the Fundamental Difference Hypothesis or not?: Island constraints revisited. *Studies in Second Language Acquisition* 31(2), 199-223.

Bley-Vroman, R. W. 1990. The logical problem of foreign language learning. *Linguistic Analysis* 20(1-2), 3-49.

Bley-Vroman, R. W. 2009. The evolving context of the Fundamental Difference Hypothesis. *Studies in Second Language Acquisition* 31(2), 175-198.

Bley-Vroman, R. W., S. W. Felix and G. L. Ioup. 1988. The accessibility of Universal Grammar in adult language learning. *Second Language Research* 4(1), 1-32.

Boeckx, C. 2008. *Bare Syntax.* Oxford: Oxford University Press.

Bolker, B. and D. Robinson. 2024. *broom.mixed*: Tidying methods for mixed models (version 0.2.9.6) [R package]. Available online at https://CRAN.R-project.org/package=broom.mixed.

Boxell, O. and C. Felser. 2017. Sensitivity to parasitic gaps inside subject islands in native and non-native sentence processing. *Bilingualism: Language and Cognition* 20(3), 494-511.

Bruening, B. and T. Trần. 2006. *Wh*-questions in Vietnamese. *Journal of East Asian Linguistics* 15(4), 319-341.

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

Chomsky, N. 1977. On *wh*-movement. In P. Culicover, T. Wasow and A. Akmajian, eds., *Formal Syntax*, 71-132. New York, NY: Academic Press.

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

Chomsky, N. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Clahsen, H. and C. Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27(1), 3-42.

Conway, A. R. A., M. J. Kane, M. F. Bunting, D. Z. Hambrick, O. Wilhelm and R. W. Engle. 2005. Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review* 12(5), 769-786.

Cowart, W. 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments.* Thousand Oaks, CA: SAGE Publications.

Ellis, N. C. 2006. Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics* 27(2), 164-194.

Goodall, G. 2015. The D-linking effect on extraction from islands and non-islands. *Frontiers in Psychology* 5, 1493

Hawkins, R. and C. Y.-H. Chan. 1997. The partial availability of Universal Grammar in second language acquisition: The 'Failed Functional Features Hypothesis'. *Second Language Research* 13(3), 187-226.

Hawkins, R. and H. Hattori. 2006. Interpretation of English multiple *wh*-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research* 22(3), 269-301.

Hofmeister, P. and I. A. Sag. 2010. Cognitive constraints and island effects. *Language* 86(2), 366-415.

Hofmeister, P., L. S. Casasanto and I. A. Sag. 2012. How do individual cognitive differences relate to acceptability judgments?: A reply to Sprouse, Wagers and Phillips. *Language* 88(2), 390-400.

Huang, C.-T. J. 1982. *Logical Relations in Chinese and the Theory of Grammar.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA.

Johnson, J. S. and E. L. Newport. 1991. Critical Period effects on universal properties of language: The status of Subjacency in the acquisition of a second language. *Cognition* 39(3), 215-258.

Kane, M. J., D. Z. Hambrick, S. W. Tuholski, O. Wilhelm, T. W. Payne and R. W. Engle. 2004. The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General* 133(2), 189-217.

Kim, B. and G. Goodall. 2016. Islands and non-islands in native and heritage Korean. *Frontiers in Psychology* 7, 134.

Kluender, R. 1998. On the distinction between strong and weak islands: A processing perspective. In P. W. Culicover and L. McNally, eds., *Syntax and Semantics 29: The Limits of Syntax*, 241-279. San Diego, CA: Academic Press.

Kluender, R. 2004. Are subject islands subject to a processing account? In *Proceedings of WCCFL 23*, 475-499.

Kluender, R. and M. Kutas. 1993. Subjacency as a processing constraint. *Language and Cognitive Processes* 8(4), 573-633.

Kush, D., A. Omaki and N. Hornstein. 2013. Microvariation in islands? In J. Sprouse and N. Hornstein, eds., *Experimental Syntax and Island Effects*, 239-264. Cambridge: Cambridge University Press.

Kush, D., T. Lohndal and J. Sprouse. 2018. Investigating variation in island effects. *Natural Language & Linguistic Theory* 36(3), 743-779.

Kuznetsova, A., P. B. Brockhoff and R. H. B. Christensen. 2017. *lmerTest* package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13), 1-26.

Lenth, R. V. 2024. *emmeans*: Estimated Marginal Means, aka Least-Squares Means (version 1.10.5) [R package]. Available online at https://CRAN.R-project.org/package=emmeans.

Li, X. 1998. Adult L2 accessibility to UG: An issue revisited. In S. Flynn, G. Martohardjono and W. O'Neil, eds., *The Generative Study of Second Language Acquisition*, 89-110. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

MacWhinney, B. 2005. Extending the Competition Model. *International Journal of Bilingualism* 9(1), 69-84.

Martohardjono, G. 1993. *Wh-Movement in the Acquisition of a Second Language: A Cross-Linguistic Study of Three Languages with and without Overt Movement.* Doctoral dissertation, Cornell University, Ithaca, NY, USA.

Mathôt, S., D. Schreij and J. Theeuwes. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2), 314-324.

Omaki, A. and B. Schulz. 2011. Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition* 33(4), 563-588.

O'Neill, R. 1981. *Kernel Lessons Plus*. Harlow, UK: Longman.

O'Rourke, P. 2013. Working memory capacity and the cognitive underpinnings of syntactic processing. Poster presented at the *87th Annual Meeting of the Linguistics Society of America*.

Perpiñán, S. 2020. *Wh*-movement, islands, and resumption in L1 and L2 Spanish: Is (un)grammaticality the relevant property? *Frontiers in Psychology* 11, 395.

Pesetsky, D. 1987. *Wh*-in-situ: Movement and unselective binding. In E. Reuland and A. G. B. ter Meulen, eds., *The Representation of (In)definiteness*, 98-129. Cambridge, MA: MIT Press.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing* (version 4.4.1) [Computer software]. Available online at https://www.R-project.org/

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

Ross, J. R. 1967. *Constraints on Variables in Syntax.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA.

Schachter, J. 1990. On the issue of completeness in second language acquisition. *Interlanguage Studies Bulletin (Utrecht)* 6(2), 93-124.

Schwartz, B. D. and R. A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12(1), 40-72.

Sprouse, J. 2007. Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1, 123-134.

Sprouse, J., M. Wagers and C. Phillips. 2012. A test of the relation between working-memory capacity and syntactic island effects. *Language* 88(1), 82-123.

Trần, T. 2009. *Wh-quantification in Vietnamese*. Doctoral dissertation, University of Delaware, Newark, DE, USA.

Tsai, C.-Y. E. 2009. *Wh-Dependency in Vietnamese and the Syntax of Wh-in-Situ.* Master's thesis, National Tsing Hua University, Hsinchu, Taiwan.

Turner, M. L. and R. W. Engle. 1989. Is working memory capacity task dependent? *Journal of Memory and Language* 28(2), 127-154.

Unsworth, N., R. P. Heitz, J. C. Schrock and R. W. Engle. 2005. An automated version of the operation span task. *Behavior Research Methods* 37(3), 498-505.

VanPatten, B. 2015. Input processing in adult second language acquisition. In B. VanPatten and J. Williams, eds., *Theories in Second Language Acquisition: An Introduction*, 113-134. New York: Routledge.

Wagers, M. and C. Phillips. 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45(2), 395-433.

White, L. 2003. *Second Language Acquisition and Universal Grammar.* Cambridge: Cambridge University Press.

White, L. and A. Juffs. 1998. Constraints on *wh*-movement in two different contexts of non-native language acquisition: Competence and processing. In S. Flynn, G. Martohardjono and W. O'Neil, eds., *The Generative Study of Second Language Acquisition*, 111-129. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Examples in: English, Vietnamese
Applicable Languages: English, Vietnamese
Applicable Level: Tertiary

