Experimental syntax, island effects, and the nature of wh-dependencies in English and Japanese

This paper presents a series of formal acceptability judgment experiments designed to investigate to what extent wh-dependencies in English and Japanese may be formed by the same grammatical operations. These experiments reveal two new effects, the length effect and the relative island effect, that suggest that English wh-displacement and wh-in-situ dependencies form a natural class to the exclusion of Japanese wh-in-situ dependencies. To account for these facts, we argue for a reconception of the offline grammatical operation movement in terms of theories of online sentence processing complexity (Gibson 1998, Alexopoulou and Keller 2007), and for an integration of the barriers/phases analysis of island effects (Chomsky 1986, 2001) with the working memory resource limitation analysis of island effects (Kluender and Kutas 1993, Kluender 2004). These results suggest that formal acceptability judgment experiments can not only further syntactic theories, but may also facilitate a closer interaction between offline grammatical theories and online processing theories.

Keywords: Experimental Syntax, wh-movement, wh-in-situ, magnitude estimation, island effects

1. Introduction

1.1 Wh-dependencies

One of the more salient features of human language is the existence of non-local dependencies. This paper focuses on one class of non-local dependencies in particular, namely those that occur in questions, as exemplified in (1-3):

(1) What does Mary think that John bought __?

(2) Who thinks that John bought what?

(3)

Keiko-wa	Taro-ga	nani-o	katta-to	omottei-masu-ka?
K-TOP	T-NOM	what-ACC	bought-COMP	think-POL-Q

'What does Keiko think that Taro bought?'

We begin with a note on terminology as a way of circumscribing the object of study for the experiments described in this paper. Each of the syntactic structures in (1-3) has several names in the literature (e.g., wh-movement dependencies, covert movement dependencies, multiple wh-questions, wh-in-situ, etc.). The names themselves usually denote a specific assumption about the genesis of the structure (e.g. wh-movement dependency) or a salient feature of the structure (e.g., multiple wh-question). One goal of the experiments in this paper is to determine to what extent these three structures may share common features, and possibly a common genesis, therefore we will use the general term *wh-dependency* to denote a non-local dependency involving at least one wh-word to avoid any assumptions about the features or genesis of the dependency. In that sense, all three of the sentences above contain a wh-dependency.

On the surface, these three sentences appear very different. In sentence (1), the normally local relationship between *what* and *bought* is non-local. Again, in order to avoid assumptions about the source of this non-locality, we will call this a *wh-displacement dependency*. In (2), the local relationship between *what* and *bought* is restored. However, there is an additional relationship between *what* in the embedded clause and *who* in the matrix clause. We will refer to this additional (non-local) dependency as an English *wh-in-situ dependency*. Similarly, the relationship between *what* and *bought* in (3) is local as expected; however there is also a non-local dependency between the wh-word *what* and the licensing particle *ka* in the matrix C position. We will refer to this non-local dependency as a Japanese *wh-in-situ dependency*.

What is interesting about these three sentences is that even though the non-local dependencies appear very different – (1) is between an argument and verb, (2) is between an argument and argument, and (3) is between an argument and licensing particle – various analyses have argued that these three dependencies are created by the same grammatical mechanism. For example, one extensive line of analyses (Huang 1982, Lasnik and Saito 1984, Chomsky 1986, and many others) has proposed a single syntactic operation called *movement* that creates all three of these dependencies: in (1) the wh-word overtly moves to the matrix spec, CP; in (2) and (3), the in-situ wh-word *what* moves to the matrix spec, CP at the level of LF. In the latter case, the effect is covert with respect to the structure that is pronounced as shown in (4-5):

(4) [$_{CP}$ [who what] [$_{IP}$ __ think that John bought __]]?

(5) [$_{CP}$ [nani-o] [$_{IP}$ Keiko-wa Taro-ga _____ katta-to omottei-masu-ka?
what-ACC K-TOP T-NOM _____ bought-COMP think-POL-Q
'What does Keiko think that Taro bought?'

The experiments in this paper are an attempt to refine our understanding of the similarities and differences among these wh-dependencies by extending the set of phenomena that can be used to classify these dependencies.

1.2 Island effects and the displacement/in-situ asymmetry

One prima facie empirical argument against a common source for all three wh-dependencies is the existence of island effects (Ross 1967). As Ross first noted, wh-displacement dependencies in English are theoretically unbounded with respect to distance, but appear to be ungrammatical when the verb is within certain structures. He metaphorically named these structures islands. Example (6a) shows that a wh-displacement dependency is possible when the verb is within a CP headed by *that*, while (6b) shows that a wh-displacement dependency is impossible when the verb is within a CP headed by *whether*:

(6) a. What does Mary think that John bought __ ?
b. *What does Mary wonder whether John bought __ ?

However, as Huang 1982 observed for English and Chinese, and Lasnik and Saito 1984 observed for Japanese, it is grammatical to have a wh-in-situ dependency inside an island:[1]

(7) a. Who wonders whether John bought what __?
b. Tsuma-ga otto-ga nani-o shita-kadooka shirabeta-no
wife-NOM husband-NOM what-ACC did-COMP investigated-Q
'What did the wife investigate whether the husband did?'

In order to maintain the position that all wh-dependencies are created by a single mechanism called *movement*, the theory has to be complicated slightly. Huang (1982) proposes that only overt movement (as in (1)) is constrained by islands, while covert movement (as in (2-3)) is free to violate island constraints. Nishigauchi 1990 takes a different tack, and argues that both overt and covert movement are constrained by islands; however, covert movement allows for large-scale pied-piping of the entire island structure, in effect eliminating the need to move *out* of the island, and therefore the island effect. While the details of these analyses differ, the end result is the same: movement is

[1] Crucially, this is only true of wh-arguments such as *who* and *what*. Wh-adjuncts such as *how* and *why* do not participate in grammatical wh-in-situ dependencies inside of islands. The experiments in this paper focus on wh-arguments only.

the operation underlying all wh-dependencies, but movement is bifurcated into two types (overt and covert).[2]

A second thread of analyses seeks to eliminate the bifurcated nature of movement, but in so doing, bifurcates the class of wh-dependencies with wh-displacement on one hand and both types of wh-in-situ on the other. The guiding idea behind these analyses is that wh-words must be functionally interpreted via one of several grammatical mechanisms. Movement is one such mechanism; however, there are other operations available to the grammar for the interpretation of wh-words, such as unselective binding (Tsai 1994) and choice-functions (Reinhart 1997). Under these analyses, movement still underlies wh-displacement dependencies, but wh-in-situ dependencies are functionally interpreted using one of the non-movement operations. The end result is a bifurcation of wh-dependencies: wh-displacement is created by movement, while wh-in-situ in both English and Japanese is handled by a second interpretive mechanism.[3]

While these two approaches differ significantly at a theoretical level, they capture the same empirical facts: they both account for the existence of wh-displacement and wh-in-situ, and they both account for the grammaticality of wh-in-situ within islands. In order to further the debate between these two approaches, we need new data. To that end, we use formal acceptability judgment collection techniques (sometimes called experimental syntax, Cowart 1997), to define two new effects for wh-dependencies: the length effect and the *relative* island effect. We then use these effects to (re-)classify the three dependencies in (1-3). As we shall see shortly, these two new effects of wh-dependencies suggest that neither of the current classifications is correct. Instead, we will argue that wh-displacement and English wh-in-situ form a natural class to the exclusion of Japanese wh-in-situ. We then suggest an analysis that captures this new classification using the movement operation and successive cyclic covert movement.

1.3 The role of experimental syntax in syntactic theory

This paper operates at two levels: an empirical/theoretical level that seeks to further the debate between movement and non-movement analyses of wh-dependencies, and a methodological level that seeks to answer one of the primary questions facing the burgeoning field of experimental syntax – What is the role of experimental syntax in syntactic theorizing? There have been two general responses to this question in the literature: 1) Experimental syntax can reveal previously unobserved data, especially between categorically grammatical sentences (e.g., Keller 2000, Sorace and Keller 2004, and Featherston 2005), and 2) Experimental syntax can contribute to debates over which

---

[2] Recently, Chomsky (e.g. 2001) proposed a new dependency-forming operation, AGREE. Under the AGREE approach, all wh-dependencies are created by the same operation similar to the older movement approach, with the difference being that CP in English has a single EPP feature that triggers a single instance of wh-movement, whereas Japanese CP has no EPP feature.

[3] Hagstrom 1998 pursues a hybrid approach to Japanese wh-in-situ in which the ka-particle begins adjacent to the wh-word and then moves to the matrix C position. This approach shares some elements of the movement approach (movement of ka), and some elements of the non-movement approach (the wh-word doesn't move).

acceptability effects are due to the grammar, and which may be due to factors outside of the grammar, such as properties of sentence processing (e.g. Kluender and Kutas 1993, Alexopoulou and Keller 2007, and Sprouse 2008). The experiments presented in this paper demonstrate both possibilities: the length effect and relative island effect are defined as a difference between categorically grammatical sentences, and also suggest an increased role for interaction between the syntactic literature and the sentence processing literature, as the analyses of the results presented here require the reconception of offline grammatical operations in terms online parsing processes.

The rest of the paper is organized as follows. Section 2 presents the first of the two new effects, which we call the dependency length effect. A series of experiments in English and Japanese demonstrate that both wh-displacement and wh-in-situ in English show a length effect, but that Japanese wh-in-situ does not. Furthermore, these experiments suggest that the length effect is specific to wh-dependencies, as binding dependencies in English do not show a length effect. Section 3 presents the second effect, which we call the relative island effect (as opposed to the absolute island effect that is standard in the literature). Experiments in English and Japanese demonstrate that English wh-in-situ dependencies show relative island effects for Subject and Complex NP islands, but that Japanese wh-in-situ dependencies do not show any relative island effects. Section 4 presents an analysis of the length effect as a reflex of the grammatical operation movement within theories of complexity from the sentence processing literature (e.g. Gibson 1998). Similarly, section 5 presents an analysis of the relative island effects found for wh-in-situ in English that is base upon an interaction of the barriers/phases approach to islands (Chomsky 1986, 2001) and the working memory resource limitation approach to islands (Kluender and Kutas 1993, Kluender 2004). Section 6 concludes.

2. The dependency length effect

The two new effects at the center of this study are predicated upon the growing body of research suggesting that the comparison of grammatical structures, that is, structures that native speakers say are grammatical in an informal yes/no task, may reveal differences in relative acceptability that are significant for syntactic theory (e.g. Kluender and Kutas 1993, Featherston 2005, Phillips et al 2005, Alexopoulou and Keller 2007, Sprouse 2008). We call the first of these effects the dependency length effect. First reported by Phillips et al. 2005 in an ERP study of the processing of wh-displacement dependencies in English, the dependency length effect is a difference in acceptability that arises by increasing the length of a wh-displacement dependency without changing any of the other structural properties of the sentence (overall length, clause types, grammaticality, etc.). Phillips et al compared a length of one CP-clause to a length of two CP-clauses:

(8) The professor hoped [$_{CP}$ that the committee knew [$_{CP}$ **which researchers** the generous millionaire would fund __ at the institute]]

The professor knew [$_{CP}$ **which researchers** the committee hoped [$_{CP}$ that the generous millionaire would fund __ at the institute]]

In a 7-point rating task, Phillips et al. found that the 1-clause dependencies were rated significantly higher than the 2-clause wh-dependency, despite both structures being categorically grammatical.

While the Phillips et al. study suggests that length affects the acceptability of wh-displacement dependencies in English, it leaves the following questions unanswered:

i. Is the length effect unique to English wh-displacement dependencies or does it extend to English wh-in-situ dependencies?

ii. Does the length effect index a syntactic property of the dependencies, or does it index a semantic/interpretive property of the dependencies?

iii. Does the same pattern of results hold for Japanese wh-in-situ dependencies?

This section reports the results of three acceptability experiments designed to answer these questions and determine to what extent the length effect can be used to probe the nature of wh-dependencies.

2.1 The design

To address the first question, experiment 1 included three structural manipulations arranged as pairs of conditions. The first pair was a straightforward replication of the Phillips et al. wh-displacement conditions: one-clause versus two-clauses. The second was a manipulation of the distance of wh-in-situ dependencies, again one-clause versus two-clauses. The third pair was included to address a possible confound: the manipulation of the length of an embedded wh-dependency in English also entails the manipulation of the size of the embedded question. It is logically possible that longer embedded questions lower acceptability for non-syntactic reasons. Therefore a third manipulation was included to tease apart the contribution of the size of the embedded question from the contribution of the length of the dependency. In this manipulation, the complementizer *whether* was used to hold the length of the dependency constant while varying the length of embedded question. In English, in-situ wh-words cannot form a dependency with *whether*, but must form a dependency with a wh-word in spec,CP. By placing a wh-word in the matrix spec,CP, we can force the dependency to be the entire length of the sentence, while simultaneously varying the size of the embedded question through the position of *whether*. If there is a length effect to wh-in-situ dependencies but not to this manipulation of the size of the embedded question, then we can be reasonably sure that the length effect is not driven by the size of the embedded question. In the examples given in Table 1, the boundaries of the wh-dependencies are marked by wh-words in **bold**. The size of the embedded question for size-manipulation is marked by *italics*:

Table 1: Conditions in Experiment 1

condition	example
displacement-short	Who hoped that you knew **who** the mayor would honor __?
displacement-long	Who knew **who** you hoped that the mayor would honor __?
in-situ-short	Who hoped that you knew **who** would honor **who**?
in-situ-long	Who knew **who** hoped that you would honor **who**?
size-small	**Who** hoped that you knew *whether the mayor would honor* ***who***?
size-large	**Who** knew *whether you hoped that the mayor would honor* ***who***?

While the manipulation of the size of the embedded question in experiment 1 begins to address the second question in (9), it does not rule out the possibility that the underlying cause of the length effect is the semantic/interpretive nature of the wh-dependencies themselves, rather than a syntactic property. For example, if there is a length effect for wh-in-situ dependencies in English, the underlying cause could be the interpretive relationship between the two wh-words in the dependency, which gives rise to the pair-list reading of multiple-wh questions in English (Wachowicz 1974). To test this possibility, experiment 2 looked for length effects in dependencies that also have the same semantic/interpretive properties as wh-in-situ dependencies, but different syntactic properties: namely, binding dependencies.

Standard binding dependencies between an R-expression and a pronoun are similar to wh-in-situ dependencies in that they are both interpretive dependencies that hold between two non-identical lexical items. Bound-variable binding dependencies between an R-expression and a pronoun are even more similar to wh-in-situ dependencies that they both hold between non-identical lexical items and both lead to pair-list readings. However, binding dependencies and wh-in-situ dependencies differ in one important respect: the tail of wh-in-situ dependencies (a wh-word in-situ) in English must be syntactically licensed by the head of the dependency (a wh-word in spec,CP); the tail of binding dependencies (pronouns) can occur without any type of licensor, or in other words, they can appear without being part of a syntactic binding dependency. If the length effect is simply a reflex of the semantic/interpretive properties of these dependencies, then both wh-in-situ and binding dependencies should show a length effect. If, however, the length effect reflects the syntactic properties of these dependencies, then the effect should not arise with binding dependencies. Three types of binding dependencies were tested: standard binding between an R-expression and a possessive pronoun, bound-variable style binding between a quantifier and a possessive pronoun, and bound-variable style binding between a wh-word and a possessive pronoun. Possessive pronouns were used as the tail of the binding dependencies to avoid interfering effects of binding constraints in the distance manipulation (Chomsky 1981).

Table 2: Conditions in Experiment 2

condition	example
binding-short	Who hoped that you knew if **John** found **his** wallet?
binding-long	Who knew if **John** hoped that the detective found **his** wallet?
bv-qr-short	Who hoped that you knew if **everyone** found **their** wallet?
bv-qr-long	Who knew if **everyone** hoped that the detective found **their** wallet?
bv-wh-short	Who hoped that you knew **who** found **their** wallet?
bv-wh-long	Who knew **who** hoped that the detective found **their** wallet?

Experiment 3 was designed to determine to what extent the English results extend to Japanese. It is well know that English wh-in-situ dependencies differ significantly from Japanese wh-in-situ dependencies: Japanese allows for wh-in-situ dependencies that involve only a single wh-word (the dependency holds between the wh-word and the question particle *ka* or *no*), whereas English wh-in-situ dependencies must be between two wh-words. While many syntactic analyses have equated the two (see section 1), it is logically possible that they will behave differently with respect to length effects. Experiment 3 tested three structural manipulations, again in pairs of conditions: the length of a wh-in-situ dependency, the length of a standard binding dependency, and the length of a bound-variable binding dependency.

Table 3: Conditions in Experiment 3

condition	example				
wh-short	Takeshi-wa	Keiko-ga	sensei-ga	**dare**-o	suisensuru-**ka**
	T-TOP	K- NOM	prof.-NOM	who -ACC	recommend-Q
	tazuneta-to	itta-no			
	asked-COMP	said-Q			
	'Who did Takeshi said that Keiko asked that the professor recommend?'				
wh- long	Takeshi-wa	Keiko-ga	sensei-ga	**dare**-o	suisensuru-to
	T-TOP	K- NOM	prof.-NOM	who -ACC	recommend-C
	tazuneta-**ka**	itta-no			
	asked-Q	said-Q			
	'Who did Takeshi said that Keiko asked that the professor recommend?'				
bind-short	Tooru-wa	Makiko-ga	**Tomoya-to-Reiko**-ga		
	T-TOP	M-NOM	**T-and-R**-NOM		
	zibun-tachi-no-kekkonshiki-no-basho-o			eranda-to	itta-to
	self-PL-GEN-wedding-GEN-location-ACC			chose-COMP	said-COMP
	kiita-no				
	heard-Q				
	'Did Tooru hear that Makiko said that Tomoya and Reiko chose the location of their wedding?"				

bind-long	Hitoshi-wa	**Seiji-to-Tomoko**-ga	Fumiko-ga		
	H-TOP	**S-and-T**-NOM	F-NOM		
	zibun-tachi-no-inu-no-namae-o		kimeta-to	itta-to	
	self-PL-GEN-dog-GEN-name-ACC		decide-COMP	said-COMP	
	omotta-no				
	thought-Q				
	'Did Hitoshi think that Seiji and Tomoko said that Fumiko decided the name of their dog?"				
bv-short	Sono-untenshu-wa	basugaido-ga	jookyaku-no-**zen'in**-ga		
	That-driver-TOP	tourguide- NOM	passeger-GEN-**all**-NOM		
	zibun-tachi-no-basu-o	oboeteiru-to	itta-to		
	self-PL-GEN-bus-ACC	remember-COMP	said- COMP		
	omotta-no				
	thought- Q				
	'Did the driver think that the tour guide said that all the passengers remember their own seats?"				
bv-long	Hitoshi-wa	**Seiji-to-Tomoko**-ga	Fumiko-ga		
	H-TOP	**S-and-T**-NOM	F-NOM		
	zibun-tachi-no-inu-no-namae-o		kimeta-to	itta-to	
	self-PL-GEN-dog-GEN-name-ACC		decide-COMP	said-COMP	
	omotta-no				
	thought-Q				
	'Did Hitoshi think that Seiji and Tomoko said that Fumiko decided the name of their dog?"				

2.2 Participants

26 University of Maryland undergraduates participated in experiment 1 (wh-dependency length). 22 Princeton University undergraduates participated in experiment 2 (binding-dependency length). 49 undergraduates at Kansai University in Osaka, Japan participated in experiment 3 (Japanese wh- and binding-dependency length). All were self-reported monolingual speakers of English and Japanese respectively with no prior training in linguistics.

2.3 Materials and Procedure

In all three experiments, 8 lexicalizations of each condition were created, and distributed among 8 lists using a Latin Square procedure. These conditions were mixed with unrelated conditions from another experiment to fill out the entire spectrum of grammaticality. The items in each list were then psuedorandomized such that related conditions were never presented sequentially. Participants in experiment 1 rated 32 items in total, participants in experiment 2 rated 28 items in total, and participants in experiment 3 rated 54 items in total.

The task was magnitude estimation (Stevens 1957, Bard et al 1996, Keller 2000), which has been argued to be an ideal task for detecting differences between categorically

grammatical structures (e.g., Keller 2000, Featherston 2005, Alexopoulou and Keller 2007, and Sprouse 2008). This is because magnitude estimation provides participants with a theoretically unlimited number of potential responses, so they have the ability to report any size difference in acceptability that they think is meaningful (Stevens 1957, Bard et al 1996). The task itself is straightforward: participants are asked to rate the acceptability of a target sentence by using the acceptability of a reference sentences as a unit of measure.

(9) An example of linguistic magnitude estimation

Reference: What do you wonder whether Mary bought?
Acceptability: 100

Target Item: What did Lisa meet the man that bought?
Acceptability: ___

In this example, the acceptability of the reference sentence is assigned an acceptability rating of 100. The participant must then rate the acceptability of the target item using the reference as a unit of measure. If the participant feels the target item is only half as acceptable as the reference, they would assign it a rating of 50. If it is twice as acceptable, they would rate it 200. Because the positive number line is theoretically infinite, participants have an unlimited number of potential responses, and can therefore report any difference in acceptability that they wish. The reference sentences chosen for experiments 1 and 2 was an if-island violation: *What did you ask if your mother bought for your father*? The reference sentence for experiment 3 was an internally headed relative clause containing a wh-word: *Kenji-wa dare-ga ringo-o kattekita-no-o katteni tabeta-no?* ('Did Kenji eat the apple that who bought?'). The goal was to choose reference sentences that are approximately in the middle of the range of acceptability (Keller 2000).

2.4 Results and Discussion

Responses were divided by the score of the reference sentence, and then log transformed prior to analysis to correct for the rightward skew inherent in magnitude estimation, following the standard practices set out in Bard et al. 1996 and Keller 2000. Note that the log transformation converts the score of the reference sentence to 0, which means that positive log-transformed judgments indicate structures that were rated higher than the reference, and negative log-transformed judgments indicate structures that were rated lower than the reference. No notion of grammaticality is implied by the sign of the judgments. Paired-samples t-tests were performed on each pair of short and long dependencies. Following the suggestion of Alexopoulou and Keller 2007, the full details of the statistical analyses have been collated in Appendix A. The graphs in Figures 1-3 report the mean log-transformed judgments as well as the standard error for each experiment. Significant differences are indicated using an asterisk notation: * = $p<.05$, ** = $p<.01$, *** = $p<.001$.

The primary question motivating experiments 1-3 is to what extent the length effect can be used to classify wh-dependencies in English and Japanese, which we broke down into three questions, repeated here for convenience:

i. Is the length effect unique to English wh-displacement dependencies or does it extend to English wh-in-situ dependencies?

ii. Does the length effect index a syntactic property of the dependencies, or does it index a semantic/interpretive property of the dependencies?

iii. Does the same pattern of results hold for Japanese wh-in-situ dependencies?

We now examine each of these questions in turn in Figures 1 and 2.

Figure 1: Length effects for wh and binding dependencies in English

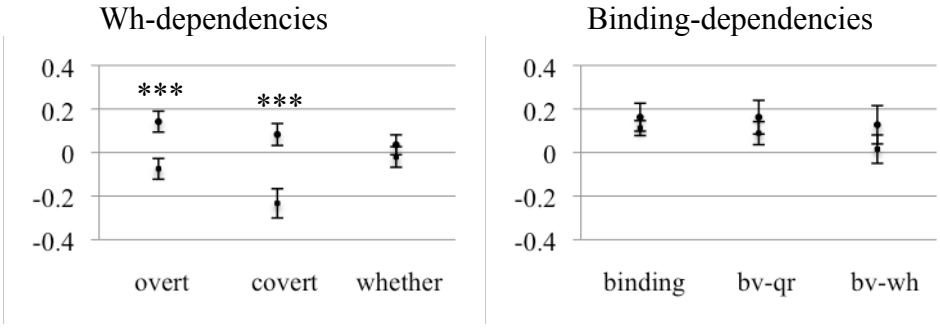


Figure 1 suggests that the length effect does appear to extend to English wh-in-situ dependencies: there is a significant decrease in acceptability for longer dependencies for both wh-displacement and wh-in-situ in English. Figure 1 also indicates that there is no difference between longer and shorter embedded questions when the wh-in-situ dependency length is held constant, suggesting that the length effect is indeed due to the length of the dependency, and cannot be reduced to an effect of the size of the embedded question. Figure 1 also suggests that the length effect indexes a syntactic property of English wh-dependencies: binding dependencies, which share the interpretive properties of wh-in-situ dependencies but not the syntactic properties, do not show length effects.

Figure 2: Length effects on Wh and Binding dependencies in Japanese

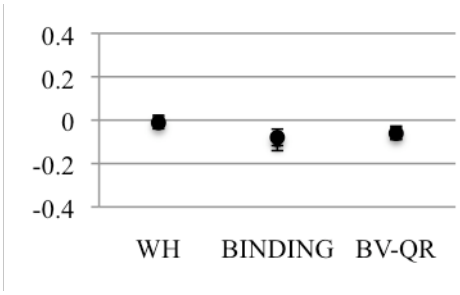


Figure 2 indicates that neither wh-in-situ nor binding dependencies in Japanese demonstrate length effects. This suggests that there is a qualitative difference between the properties of English wh-in-situ dependencies and Japanese wh-in-situ dependencies.

Taken as a whole, these results suggest that there is a qualitative similarity between English wh-displacement dependencies and wh-in-situ dependencies, such that English wh-dependencies form a natural class to the exclusion of Japanese wh-in-situ dependencies. This is surprising given that the two sets of analyses discussed in section 1 argue for either a single class of wh-dependencies consisting all three types, or a class consisting of English and Japanese wh-in-situ to the exclusion of English wh-displacement. An analysis of these results will be presented in section 4, but first, section 3 presents further evidence of this new classification by examining island effects with wh-in-situ.

3. The relative island effect

The results of section 2 suggest that, with respect to length effects, English wh-displacement and wh-in-situ form a natural class to the exclusion of Japanese wh-in-situ. However, the observations of Huang 1982 and Lasnik and Saito 1984 suggest that, with respect to island effects, English and Japanese wh-in-situ form a natural class to the exclusion of English wh-displacement: English and Japanese wh-in-situ do not show island effects, at least for wh-arguments such as *who* and *what*. In this section we would like to argue that this conflict only holds when island effects are defined in absolute terms, using the notion of categorical grammaticality. When island effects are defined in terms of relative acceptability, two island effects emerge for English wh-in-situ: Subject islands and NP islands. Similar to the dependency length effect, this relative island effect, is predicated upon the ability of experimental syntax techniques to reliably distinguish the relative acceptability of categorically grammatical sentences (Keller 2000, Featherston 2005).

3.1 The relative definition of island effects

In order to construct a relative definition of island effects, it is perhaps easiest to take the absolute definition as a starting point. For example, Huang 1982 observes that while English wh-displacement dependencies within island structures are considered categorically ungrammatical by native speakers (as indicated by the traditional asterisk), English wh-in-situ dependencies are considered categorical grammatical (no asterisk).

(10) Whether Island
  a. *What do you wonder whether John bought?
  b. Who wonders whether John bought what?

(11) NP Island
  a. *What did you make the claim that John bought?
  b. Who made the claim that John bough what?

While these comparisons demonstrate that wh-in-situ within islands is *more acceptable*

than wh-displacement out of islands, they do not necessarily indicate that there is no wh-in-situ *island effect* per se. For instance, it could be the case that there is a wh-in-situ island effect but this effect is *weaker* than the wh-displacement island effect. In fact, it could be the case that this island effect is so weak that the sentences are still considered fully acceptable by native speakers, similar to the Superiority effect Featherston 2005 demonstrates in German. To test this possibility, we need a definition of island effect that makes no reference to notions of grammaticality and ungrammaticality. However, eliminating references to grammaticality and relying entirely on relative acceptability admits the possibility that non-island effects, such as the length effect discussed in section 2, could contribute to relative differences in acceptability. In other words, we need a definition that controls for possible confounding factors known to affect the acceptability of wh-dependencies. We can then define island effects as the additional decrease in acceptability above and beyond the factors known to affect wh-dependency acceptability.

There are two primary factors in the literature that are known to affect wh-dependency acceptability. The first factor is STRUCTURE: Kluender and Kutas 1993 demonstrate that potential island structures, such as an embedded CP headed by *if*, are significantly less acceptable than non-island structures, such as an embedded CP headed by *that*. We can control for the acceptability cost inherent in potential island structures by including the factor STRUCTURE with two levels: non-island and island. The second factor is LENGTH: we have seen from the discussion in section 2 that wh-dependencies that span 2 clauses are less acceptable than wh-dependencies that span only a single clause. Many of the island effects investigated here require a wh-dependency that spans 2 clauses. In order to control for the acceptability cost inherent in longer wh-dependencies, the relative definition of island effect includes a second factor LENGTH with two levels: short and long. This yields a 2x2 factorial design, illustrated here for the Whether island:

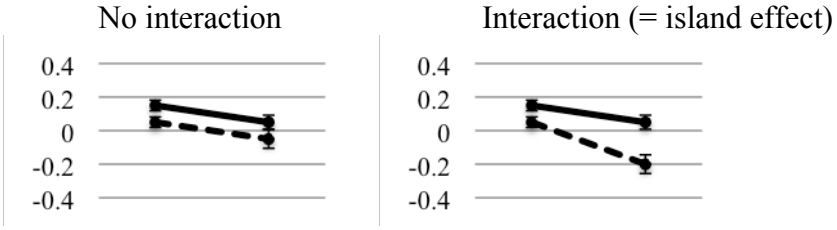
Table 4: Interaction definition of island effect for English Whether islands

	short	long
non-island	Who __ thinks that John bought a car?	What do you think that John bought __?
island	Who __ wonders whether John bought a car?	What do you wonder whether John bought __?

We can visualize the results of this design by plotting the four conditions above as endpoints for two horizontally stacked lines. Given our knowledge of the effects of STRUCTURE on acceptability ratings, we would expect the top line to represent the two non-island structure conditions and the bottom line to represent the two potential island structure conditions (a main effect of STRUCTURE). Given our knowledge of the effect of LENGTH on acceptability ratings, we would expect the lines to have a downward slope from left to right if the leftmost points represented the short movement conditions, and the right most points represented the long movement conditions (a main effect of LENGTH). In other words, if STRUCTURE and LENGTH are the only factors affecting acceptability ratings, we would expect two parallel lines with a slightly downward slope

from left to right as in Figure 3, and we would expect two main effects in a two-way repeated measures ANOVA. If island effects exist, they will appear as lower acceptability in the potential-island/long movement condition (the bottom right point in the figures) as in Figure 3, and will appear as an interaction effect of STRUCTURE x LENGTH in a two-way repeated measures ANOVA, with planned paired-samples t-Tests demonstrating that the effect is in the potential-island/long movement condition.[4]

Figure 3: Visualization of the interaction definition of island effects



Island effects for Japanese wh-in-situ can be similarly defined as an interaction of STRUCTURE x LENGTH, with the only difference being that LENGTH refers to the length of the dependency between the in-situ wh-word and the question particle in matrix C (*ka* or *no*). Examples are given in Table 7 in section 3.3.

Island effects for English wh-in-situ can also be defined as an interaction; however, because wh-in-situ in English only occurs with multiple wh-questions, the definition must be altered slightly. The first factor, STRUCTURE, is exactly as before, and controls for the acceptability cost inherent in potential island structures. The second factor, which we call DEPENDENCY, controls for the acceptability difference between a single wh-question and a multiple wh-question. DEPENDENCY has two levels: the first level is a single wh-question with a matrix wh-word, and the second is a multiple wh-question with the in-situ wh-word in the embedded clause:

Table 5: Interaction definition of island effect for English wh-in-situ Whether islands

	short	long
non-island	Who __ thinks that John bought a car?	Who thinks that John bought what?
island	Who __ wonders whether John bought a car?	Who wonders whether John bought what?

3.2 Participants

42 undergraduates at the University of Maryland and the University of California, Irvine participated in experiment 4. All were self-reported monolingual speakers of English with no previous training in linguistics. Experiment 5 was a subdesign in the survey reported

[4] It should be noted that the interaction definition of island effects neither assumes, nor identifies, a specific cause for the island effect – an issue that will be discussed in detail in section 5.

as experiment 3 in section 2: 49 monolingual Japanese speakers from Kansai University participated in that experiment.

3.3 Materials and Procedure

Once again, 8 lexicalizations of each condition presented in Table 4 through Table 8 were created and distributed among 8 lists using a Latin Square design. The items were then mixed with unrelated items from another experiment and pseudorandomized such that related items did not appear sequentially. The task was magnitude estimation (Stevens 1957, Bard et al 1996, Keller 2000), and the reference sentences were identical to experiments 1-3: *What did you ask if your mother bought for your father*? and *Kenji-wa dare-ga ringo-o kattekita-no-o katteni tabeta-no?* ('Did Kenji eat the apple that who bought?').

Table 6: Materials for English wh-displacement islands

island	example
ADJ	Who thinks that you forgot a briefcase at the office?
	Who worries if you forget a briefcase at the office?
	What do you think that you forgot at the office?
	What do you worry if you forget at the office?
NP	Who doubted that you could finish your homework?
	Who doubted the claim that you could finish your homework?
	What did you doubt that you could finish?
	What did you doubt the claim that you could finish?
RC	Who believes that the lady owns a dog?
	Who believes the lady that owns a dog?
	What do you believe that the lady owns?
	What do you believe the lady that owns?
SUBJ	What do you think the speech interrupted?
	What do you think interrupted the TV show?
	What do you think the speech interrupted the TV show about?
	What do you think the speech about interrupted the TV show?

Table 7: Materials for Japanese wh-in-situ islands

island	example				
ADJ	Dare-ga	Takeshi-ga	kuruma-o	katta-to	itta-no
	Who-NOM	T-NOM	car-ACC	bought-COMP	say-Q
	'Who said that Takeshi bought a car?'				
	Takeshi-ga	kuruma-o	katta-ra	dare-ga	yorokobu-no
	T-NOM	car-ACC	bought-COND	who-NOM	be_happy-Q
	'Who would be happy if Takeshi bought a car?'				
	Keiko-ga	Takeshi-ga	nani-o	katta-to	itta-no
	K-NOM	T-NOM	what-ACC	bought-COMP	say-Q

	‘What did Keiko say that Takeshi bought?’					
	Takeshi-ga	nani-o	katta-ra	Keiko-ga	yorokobu-no	
	T-NOM	what-ACC	bought-COND	K-NOM	be_happy-Q	
	‘Who would be happy if Takeshi bought a car?’					
NP	Dare-ga	Satoshi-ga	shigoto-o	kaeru-to	yososhita-no	
	Who-NOM	S-NOM	job-ACC	change-COMP	predict-Q	
	‘Who predicted that Satoshi would change his job?’					
	Dare-ga	Satoshi-ga	shigoto-o	kaeru-toiu	yoso-o	
	Who-NOM	S-NOM	job-ACC	change-COMP	prediction	
	tateta-no					
	made-Q					
	‘Who made a prediction that Satoshi would change his job?’					
	Kanako-ga	Satoshi-ga	nani-o	kaeru-to	yososhita-no	
	K-NOM	S-NOM	what-ACC	change-COMP	predict-Q	
	‘What did Kanako predict that Satoshi would change?’					
	Kanako-ga	Satoshi-ga	nani-o	kaeru-toiu	yoso-o	
	K-NOM	S-NOM	what-ACC	change-COMP	prediction	
	tateta-no					
	made-Q					
	‘What did Kanako make a prediction that Satoshi would change?’					
RC	Dare-ga	kisha-ga	sono-kiji-o	kaita-to	shinjita-no	
	Who-NOM	reporter-NOM	that-article-ACC	wrote-COMP	believe-Q	
	‘Who believed that the reporter wrote that article?’					
	Dare-ga	sono-kiji-o	kaita	kisha-o	shinjita-no	
	Who-NOM	that-article-ACC	wrote	reporter-ACC	believe-Q	
	‘Who believed the reporter who wrote that article?’					
	Keiko-ga	kisha-ga	nani-o	kaita-to	shinjita-no	
	K-NOM	reporter-NOM	what-ACC	wrote-COMP	believe-Q	
	‘What did Keiko believe that the reporter wrote?’					
	Keiko-ga	nani-o	kaita	kisha-o	shinjita-no	
	K-NOM	what-ACC	wrote	reporter-ACC	believe-Q	
	‘What did Keiko believe the reporter who wrote?’					
WH	Dare-ga	Shingo-ga	ryokoo-o	shita-to	itta-no	
	Who-NOM	S-NOM	trip-ACC	did-COMP	said-Q	
	‘Who said that Shingo made (a/the) trip?’					
	Dare-ga	Shingo-ga	ryokoo-o	shita-kadooka	kiita-no	
	Who-NOM	S-NOM	trip-ACC	did-COMP	asked-Q	
	‘Who asked whether Shingo made (a/the) trip?’					
	Satoko-ga	Shingo-ga	nani-o	shita-to	itta-no	
	S-NOM	S-NOM	what-ACC	did-COMP	said-Q	
	‘What did Satoko say that Shingo did?’					
	Satoko-ga	Shingo-ga	nani-o	shita-kadooka	kiita-no	
	S-NOM	S-NOM	trip-ACC	did-COMP	asked-Q	
	‘What did Satoko ask whether Shingo did?’					

SUBJ	Takeshi-ga	Keiko-ga	dare-o	nagutta-to	kiita-no
	T-NOM	K-NOM	who-ACC	hit-COMP	heard-Q
	'Who did Takeshi hear that Keiko hit?'				
	Takeshi-ga	dare-ga	Hitoshi-o	nagutta-to	kiita-no
	T-NOM	who-NOM	H-ACC	hit-COMP	heard-Q
	'Who did Takeshi hear that hit Hitoshi?'				
	Satoru-ga	yuumei-seijika-nitsuiteno-kiji-ga		dare-niyoru	
	S-NOM	known-politician-about-article-NOM		who-BY	
	hantaiundo-o	okoshita-to	itta-no		
	protest-ACC	caused-COMP	said-Q		
	'Who did Satoru say that the article about a known politician caused the protest by?'				
	Satoru-ga	dare-nitsuiteno-kiji-ga		roodoosha-niyoru	
	S-NOM	who-about-article-NOM		working_class-BY	
	hantaiundo-o	okoshita-to	itta-no		
	protest-ACC	caused-COMP	said-Q		
	'Who did Satoru say that the article about caused the protest by the working class?'				

Table 8: Materials for English wh-in-situ islands

island	example
ADJ	Who thinks that you forgot a briefcase at the office?
	Who worries if you forget a briefcase at the office?
	Who thinks that you forgot what at the office?
	Who worries if you forget what at the office?
NP	Who doubted that you could finish your homework?
	Who doubted the claim that you could finish your homework?
	Who doubted that you could finish what?
	Who doubted the claim that you could finish what?
RC	Who believes that the lady owns a dog?
	Who believes the lady that owns a dog?
	Who believes that the lady owns what?
	Who believes the lady that owns what?
SUBJ	What do you think the speech interrupted?
	What do you think interrupted the TV show?
	Who thinks the speech interrupted the TV show about what?
	Who thinks the speech about what interrupted the TV show?

3.4 Results and Discussion

Responses were divided by the score of the reference sentence, and then log transformed to correct for the rightward skew inherent in magnitude estimation prior to analysis, following the standard practices set out in Bard et al. 1996 and Keller 2000. Island effects were defined as an interaction of structure x length/dependency localized as a decrease in

the fourth condition in each island subdesign, therefore each subdesign was analyzed using two-way repeated measures ANOVAs and planned paired-samples t-Tests. Following the suggestion of Alexopoulou and Keller 2007, the full details of the statistical analyses have been collated in Appendix A. The graphs in Figures ?-? report the mean log-transformed judgments as well as the standard error for each experiment. Solid lines represent the non-island structures and dashed lines represent the potential-island structures. Significant interactions in the ANOVAs (i.e. island effects) are indicated using an asterisk notation after the title of each graph. Significant differences in the planned t-Tests between non-island and potential island structures are indicated with asterisks in the space between the rightmost points on the graph. Asterisk notation: * = $p<.05$, ** = $p<.01$, *** = $p<.001$.

Figure 4 English Wh-displacement islands

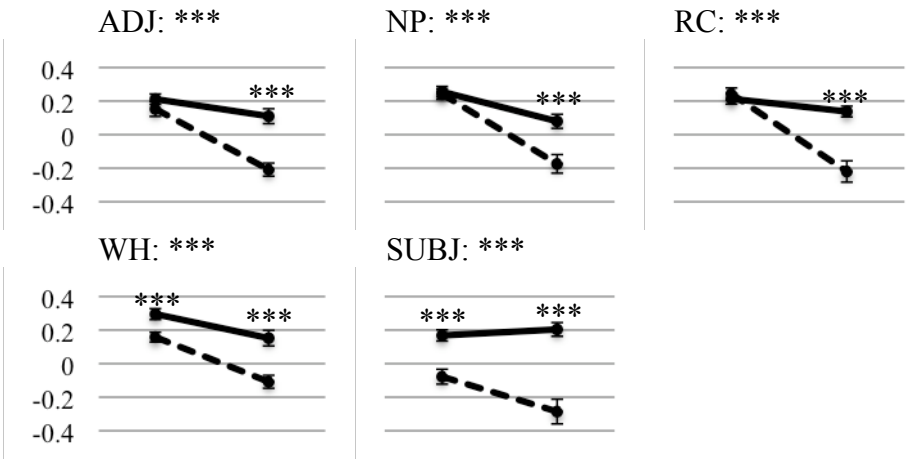
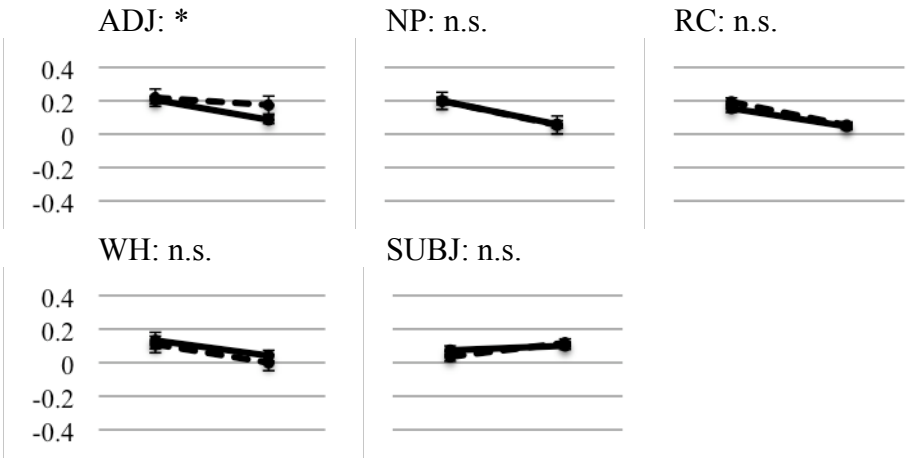


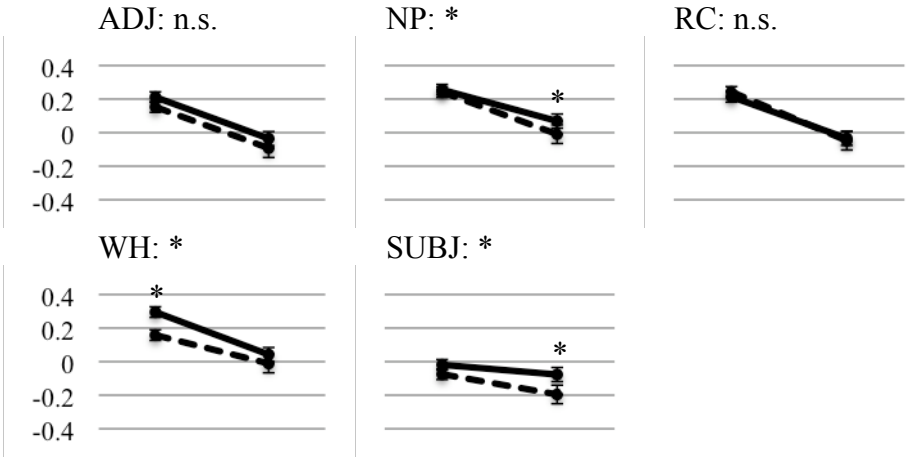
Figure 4 demonstrates the validity of the relative definition of island effect: as expected, there were highly significant interaction effects for each of the islands tested with English wh-displacement.

Figure 5 Japanese Wh-in-situ islands



For Japanese wh-in-situ dependencies, the only significant interaction effect occurred with Adjunct islands. However, the effect is in the opposite direction: the non-island structures are rated higher than the island structures. It is possible that this reflects a general preference for pre-subject phrases over post-subject phrase, as the adjunct phrases were fronted to avoid possible pragmatic complications, while the complement phrases in the control conditions were in their standard post-subject position. As such, it appears that there are no island effects for Japanese wh-in-situ dependencies, even using the relative definition of island effect.

Figure 6 English Wh-in-situ islands



There were three significant interactions for English wh-in-situ dependencies: NP, Whether, and Subject islands. Planned t-Tests confirm that the interactions for NP and Subject islands conform to the definition of island effect laid out in section 3.1: the difference between the non-island and potential-island structures is significant only when the wh-in-situ dependency is within the structures. These effects are not as large as the effects for English wh-displacement dependencies (see the Appendix for effect sizes. The interaction for Whether islands does not conform to the definition of island effect: the difference between the non-island and potential-island is significant only when there is no wh-in-situ dependency within the structures. In other words, a wh-in-situ dependency inside of a Whether island is actually better than we would expect given the relative acceptability of the structure without wh-in-situ dependencies. We will present an analysis of each of these results in section 5.

4. Length Effects and Covert Movement

The results form the experiments presented in sections 2 and 3 suggest four facts that need to be explained:

i. There are length effects for English wh-displacement and English wh-in-situ, but not for English binding dependencies, Japanese wh-in-situ, or Japanese binding dependencies.

ii. There are wh-in-situ island effects for NP and Subject islands, but not Adjunct, Relative Clause, or Whether islands in English.

iii. The wh-in-situ island effects are weaker than wh-displacement island effects

iv. There is a "reverse" island effect for Whether islands

These facts suggests that, at some level, English wh-displacement and wh-in-situ form a natural class to the exclusion of Japanese wh-in-situ: the English wh-dependencies show length and relative island effects, while Japanese wh-in-situ does not. One logical interpretation of these results is that whatever mechanism (or mechanisms) that gives rise to the length and relative island effects is operative in English wh-displacement and wh-in-situ, but that this mechanism is inoperative in Japanese wh-in-situ. We would like to suggest that the mechanism underlying the length and relative island effects is the grammatical operation called movement: English wh-displacement and wh-in-situ are both instances of movement (along the lines of Huang 1982, Lasnik and Saito 1984, Chomsky 1986, and many others), while Japanese wh-in-situ is handled by a separate mechanism (perhaps unselective binding (Tsai 1994) or choice function interpretation (Reinhart 1997)). In what follows, we pursue this hypothesis and its consequences for the nature of movement and island effects. As will soon become apparent, these consequences extend beyond the traditional realm of syntactic theory, as these results suggest a direct interaction with some of the results from psycholinguistic studies of sentence processing.

4.1 Successive cyclic covert movement

One of the more salient characteristics of movement is that it appears to be locally bounded: movement of an element cannot cross a CP node without first stopping in Spec, CP (e.g., Chomksy 1973, 1986). Historically, the bounded nature of movement has often been called successive cyclicity: movement across two CP phrases requires two instances of movement, whereas movement within a single CP phrase requires only one instance of movement. One natural interpretation of the length effect is as an index of the number of movement operations required by the wh-dependency. The longer wh-dependency conditions in this experiment require two instances of movement, while the shorter conditions only require one. The successive cyclic nature of movement can straightforwardly explain the existence of length effects with both wh-displacement and wh-in-situ dependencies if they are both derived from movement. Furthermore, given that binding dependencies are not generally assumed to involve movement operations, this would also explain the absence of length effects with binding dependencies (but see Kayne 2002 for a movement analysis of pronouns, and Hornstein 2000 for a movement analysis of anaphors).

While a movement analysis of both wh-in-situ and wh-displacement provides a straightforward explanation for the parallel length effects found in these experiments, it is rare for transformational theories, for instance those from the Government and Binding framework, to assume that covert movement is successive cyclic (e.g. Epstein 1992). Yet, if movement underlies wh-in-situ, it must be covert, and if covert movement leads to a

length effect, this analysis would suggest that it is successive cyclic. Prima facie, then, the traditional GB assumption that covert movement is not successive cyclic raises a problem for a movement analysis of length effects. However, recent Minimalist theories argue for a so-called single cycle approach to syntax in which there is no covert component. Under these analyses, there can be no asymmetry between overt and covert movement, as they are the same syntactic operation, differing only in the phonetic realization of the head and tail of the dependencies (e.g., Bobaljik 1995, Groat and O'Neil 1996, Pesetsky 1998, Bošković 2002, etc.). At a certain level then, the length effect found in this study can be viewed as corroborating one of the predictions of single cycle theories of syntax.

4.2 Length effects and the processing of wh-dependencies

The wh-displacement length effect has many correlates in the sentence processing literature. Online measures such as reading times and even ERPs appear to be sensitive to the distance between a displaced wh-word and its gap site. As such, several theories of sentence processing complexity may predict the length effect given the plausible assumption that increased processing complexity leads to lower acceptability (e.g. Kluender and Kutas 1993). For example, the theory of wh-processing proposed in Kluender and Kutas 1993, Kluender 1998, and Kluender 2004, which assumes a version of the Capacity Constrained Comprehension theory of Just and Carpenter 1992, argues that there is an increased processing load associated with holding a wh-word in working memory across a (CP) clause boundary. Under this scenario, two (CP) clause boundaries would certainly lead to an increased processing load when compared to one (CP) clause boundary. Similarly, the Syntactic Prediction Locality Theory (Gibson 1998, Gibson and Warren 2004) as extended in Alexopoulou and Keller 2007 suggests that the processing complexity of wh-dependencies increases as a function of the number of discourse referents and syntactic heads that intervene between the displaced wh-word and its gap site. The longer wh-dependencies in the experiments presents in this paper involve both a greater number of discourse referents and a greater number of syntactic heads. As such, it is fairly straightforward to interpret the wh-displacement length effect in this paper as a manifestation of processing complexity in acceptability judgments: the wh-word must be placed in working memory until it can be associated with its gap location; each successive CP clause either directly or indirectly taxes the working memory system by increasing the complexity of processing. In other words, the wh-filler is placed in memory, and then there is a forward search for a gap location. The longer this search goes on, the harder it becomes through a combination of working memory and processing complexity issues.

While the existence of wh-displacement length effects connects directly with length effects found in the sentence processing literature, it is not immediately obvious how these processing theories translate to wh-in-situ dependencies, especially given that there has been relatively little processing research done on wh-in-situ dependencies in English. Robert Kluender (p.c.) suggests that because in-situ wh-words are only licensed by a wh-word in spec, CP, when the parser encounters an in-situ wh-word, it may have to search back up the tree for a licensing wh-word. If we assume that the in-situ wh-word must be stored in working memory during the search, perhaps because of the interpretive

nature of the dependency, then this backward search may affect acceptability similar to the way that the forward search for a gap site with wh-displacement affects acceptability. If true, this would also suggest that the forward-search of the parser for the licensor *ka* in Japanese, which has been shown to affect reading times (e.g., Miyamoto and Takahashi 2000), does not invoke the same processing resources as the forward and backward searches in English since there is no concomitant decrease in acceptability (no length effect). This line of reasoning suggests a future for studies that compare the acceptability and processing of the different directions of wh-in-situ dependencies in English and Japanese.

4.3 Rethinking successive cyclic movement

A study by Frazier and Clifton 1989 suggests that a processing perspective on wh-dependencies may provide a new way of thinking about successive cyclic movement. They show that reading times are faster for sentences involving wh-displacement in a single clause (12a) than for wh-displacement that spans an embedded clause (12b), and that this difference cannot be attributed to the length of the dependency as counted by number of words:

(12) a. What did Katie and Tom mail to New York?
b. What did Sue think Tom mailed to New York?

From the point of view of syntactic theory, the crucial difference between these two is the intermediate trace at the embedded CP in (12b), or in other words, the fact that (12b) involves successive cyclic movement. From the point of view of sentence processing, the crucial difference is that the embedded CP is introduced by an additional verb, or more precisely, that the additional verb introduces a potential gap site between the wh-word and the actual gap site. There is ample evidence that the parser attempts to associate wh-words with the first potential gap site encountered (the active filling strategy: Crain and Fodor 1985, Stowe 1986) regardless of subcategorization restrictions (Pickering and Traxler 2003), which suggests that it may be profitable to think of successive cyclic movement in terms of successive attempts at gap-filling.

Combining the complexity theories from section 4.2 and a gap-filling perspective on successive cyclic movement, Kluender's suggestion (p.c) can be fleshed out as a mapping between offline grammatical operations and online parsing processes. For instance, overt movement could be an offline representation of the online process of placing a wh-word in working memory and searching forward through the structure for a gap site. Building on the ideas in Frazier and Clifton 1989, the successive cyclicity of overt movement could be an offline representation of the online process of actively attempting to fill the potential gap site at each verb encountered. Similarly, if Kluender's suggestion proves correct, covert movement in English could be an offline representation of the online process of placing an in-situ wh-word in working memory and searching backward through the tree for a wh-word licensor in spec, CP. The logical corollary of this might then be that successive cyclicity of covert movement is an offline representation of an online process of actively searching for the wh-word licensor at every CP encountered. In other words, there might be active backward licensor-searching

the same way there is active filling. The real question is whether there is any evidence for active licensor-searching.

As mentioned above, there is little processing data on the wh-in-situ in the literature. However, the mysterious "reverse" island effect for Whether islands may receive a natural interpretation under an active backward licensor-search analysis. The experiment in section 3 suggest that not only is there no island effect for Whether islands, but that wh-in-situ inside of an embedded *whether* clause is actually rated higher than one would expect. While to our knowledge there is no obvious grammatical reason for this amelioration effect, there is a potential licensor-search reason: if licensor-searching is active, the search may be temporarily interpret *whether* as a viable licensor because *whether* shares many morphosyntactic features with wh-words. This illusion may lead to an acceptability increase similar to the well-known grammaticality illusions in the literature (e.g., *More people have been to Russia than I have* (first noted by Montalbetti 1984)). This effect may suggest that online studies of English wh-in-situ may uncover evidence of an active backward search for a licensor.

The mapping between offline grammatical primitives and online processes sketched above is in no way unique: attempts to rectify theories of grammar and theories of online processing form the basis of many research programs in the sentence processing field (see especially the work of Lyn Frazier, Charles Clifton, Robert Kluender, and Colin Phillips). What is interesting is that this discussion arises in the context of a set of experiments that were designed to better understand syntactic theories – experiments that measure only offline judgments of sentence acceptability (see also Keller 2000, Featherston 2005, and Alexopoulou and Keller 2007). This suggests that experimental syntax techniques may allow syntacticians to not only gather data to further syntactic theories, but also to build a new body of data that may help to facilitate interaction between theoretical syntax and sentence processing.

5. Island effects and the processing of wh-dependencies

The discussion in section 4 addressed two of the four unexplained results from the experiments presented in this paper: the length effect for English wh-dependencies and the "reverse" island effect for English wh-in-situ Whether islands. But that leaves two results to be explained:

i. There are wh-in-situ island effects for NP and Subject islands, but not Adjunct, Relative Clause, or Whether islands in English.
ii. The wh-in-situ island effects are weaker than wh-displacement island effects

As before, we would like to argue that the key to understanding these results lies in the nature of covert movement, and suggests a possible interaction between syntactic theories and theories of sentence processing.

5.1 Successive cyclic movement and the nature of islands

The idea that island effects arise due to constraints on the operation of movement has a long and distinguished history in generative grammar (Chomsky 1973, Huang 1982,

Lasnik and Saito 1984, Chomsky 1986, Chomsky 1993, Chomsky 2001, and many others). So at first glance it may seem benign to suggest that movement underlies both wh-displacement and wh-in-situ in English. However, wh-displacement and wh-in-situ do not demonstrate the same set of island effects: wh-in-situ only demonstrates a subset of island effects, specifically Subject islands and NP islands. This suggests that Subject islands and NP islands may be qualitatively similar to each other at a certain level and form a natural class to the exclusion of the other island types tested (Adjunct, Relative Clause, and Whether islands). While it is fairly standard for grammatical theories of island effects to divide the island types into natural classes, we know of no existing classification that groups Subject islands and NP islands together to the exclusion of the other island types. For example, Ross's original (1967) Complex NP Constraint grouped what we call NP islands and Relative Clause islands together as "complex" nouns; Huang's (1982) Condition on Extraction Domains grouped Subject and Adjunct islands together as ungoverned phrases; and Chomsky's original (1973) Subjacency Condition was modified by Rizzi 1982 and Torrego 1984 to group Subject and Whether islands together as islands that arise from IP bounding nodes. If wh-displacement and wh-in-situ are both formed by movement, then we need an explanation for this new classification of islands with respect to covert movement.

Successive cyclic movement may begin to explain the difference between Subject and NP islands on the one hand, and Adjunct, Relative Clause, and Whether islands on the other. The barriers/phases based approach to island constraints (e.g. Chomsky 1986, 2001) argues that islands arise when successive cyclic movement cannot use spec, CP as a landing site, such as when it is filled by another element. In other words, the barriers/phases insight is that islands arise because of a failure of successive cyclic movement. Successive cyclic movement fails for Whether islands and Relative Clause islands under most standard analyses, and it fails for Adjunct islands under some analyses (e.g. Sawada and Larson 2004). However, successive cyclic movement does not fail for Subject and NP islands: both of these island types require additional mechanisms to derive their island status. For example, to account for Subject islands, Huang 1982 posited the Condition on Extraction Domains, Chomsky 1986 posited a constraint against adjunction to arguments, and Uriagereka 1999 reconceived the Culicover and Wexler 1981 freezing principle within a linearization/multiple-spell-out framework. In a way then, the existence of wh-in-situ island effects with Subject and NP islands (but not other islands) is further evidence that an additional mechanism is necessary to account for Subject and NP islands. The real question is what is the nature of that mechanism, and can it account for the fact that wh-in-situ Subject and NP island effects are weaker than wh-displacement Subject and NP island effects.

5.2 Subject and NP islands as a processing phenomenon

It almost goes without saying that there is an active debate about the nature of island effects in the syntactic and sentence processing literatures. While syntactic theories tend to posit grammatical constraints as the cause of island effects, the sentence processing literature has presented evidence that the unacceptability of islands may derive from the processing load of island structures. Most notably, Kluender and Kutas 1993 argue that island effects, defined as a decrease in acceptability for wh-displacement dependencies

within an island structure, may arise because the parser simply does not have enough working memory capacity to both parse the structure and complete the dependency. While the experiments presented in this paper were not designed to bear directly on this question – the definition of island effects used in these experiments is compatible with either a grammatical of processing source of the island effects – the natural classes of island effects suggested by these experiments is reminiscent of the natural classes suggested in Kluender 2004. Kluender 2004 argues that Subject and NP islands arise due to a working memory capacity limitation at the clause boundary that introduces the complex Subjects and complex NPs, similar to the analysis of WH islands in Kluender and Kutas 1993. This analysis, coupled with the active backward licensor-search suggested in section 4, may offer an explanation for the existence of Subject and NP islands with wh-in-situ as well as the fact that they are weaker than their wh-displacement counterparts.[5]

In section 4, we suggested that Adjunct, Relative Clause and Whether island effects for overt wh-movement can be reduced to a failure of successive cyclic movement. Under the mapping hypothesis in section 4, this would translate into a processing problem (perhaps due to active filling) at the bridge verb. For example, Kluender and Kutas 1993 argue that the lexical semantics of *whether* (and *if*), which semantically introduce a set of propositions, cause additional complexity at the clause boundary that interacts with the working memory requirements of maintaining a wh-word in working memory to create the unacceptability that we call an island effect. Similar analyses can be extended to Relative Clause islands, which also involve a wh-word at the clause boundary, and Adjunct islands, which introduce a similar type of lexical semantic complexity (Truswell 2007). This situation may be complicated further by the active filling attempt at the bridge verb immediately before the clause boundary: at the very moment that the parser is recovering from a failed attempt at active filling, it must also parse these additional complexities (the lexical semantics of *whether*, or the new wh-dependency of a relative clause).

On the other hand, covert movement shows no sign of island effects for the successive cyclicity based island types. Recall that under the mapping hypothesis above, covert movement is an offline representation of the backward search for a wh-word licensor. The lack of island effects for these islands may simply reflect the singular nature of this process: during the backward search process the parser has only one goal, to find a licensor. The parser does not have to integrate new dependencies or complex lexical semantics, as these tasks were accomplished during the initial forward parse, therefore successive cyclicity based island effects would not be expected to arise during the backward search. Subject and NP island effects, on the other hand, may still arise because the complex discourse referents (complex NPs) that lead to the island effects must still be re-parsed during the search for a licensor. Furthermore, these island effects may be weaker for wh-in-situ because the backward search process is limited to simply examining the morphosyntactic properties of the lexical items, and does not include additional parsing processes such as active filling and the integration of complex phrase structure.

---

[5] See Phillips 2006 for a dissenting analysis of Subject islands based on the existence of evidence for active filling inside of potential Subject islands.

6. Conclusion

The empirical contribution of this study is fairly straightforward. We sough to determine to what extent the three types of wh-dependencies in (1-3) are created by the same processes. The results of these experiments suggest that there is a cross-linguistic variation: English wh-dependencies appear to be similar (both show length and island effects), while Japanese wh-in-situ dependencies appear to be qualitatively different. However, real value of this study may not lie in the empirical contribution, or even the theoretical proposals, but rather in the nature of the theoretical proposals. In order to account for this data, we were forced to reinterpret offline syntactic analyses in terms of online parsing processes. For example, in order to account for the length effects, we proposed a reconception of the grammatical operations of overt and covert movement as offline representations of the online processes of active filling and a backward search for a licensor respectively (reminiscent of the rightward grammars of Steedman 2002 and Phillips 2003). Similarly, in order to account for the wh-in-situ island effects, we proposed an integration of the barriers/phases analysis of island effects (Chomsky 1986, 2001) with the working memory resource analysis of island effects (Kluender and Kutas 1993, Kluender 2004). Whether this analysis ultimately proves correct or not, we believe that these results (and similar results in Keller 2000, Featherston 2005, Alexopoulou and Keller 2007, and Sprouse 2008) suggest that a closer interaction between the syntactic literature and the sentence processing literature may not only be possible in the years to come, but may be necessary to understand the nature of the effects that are found using experimental syntax techniques.

Appendix of Statistical Results

1. Length effects

Table A1: Length effects for English wh-dependencies (cf. Figure 1)

	Short		Long		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect size (r)
Wh-displacement	0.14	0.24	-0.07	0.24	3.99	.001	.62
Wh-in-situ	0.08	0.26	-0.23	0.34	3.83	.001	.61
Embedded Question	0.04	0.23	-0.02	0.24	0.88	.388	

Table A2: Length effects for English binding-dependencies (cf. Figure 1)

	Short		Long		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect size (r)
Binding	0.16	0.30	0.11	0.16	0.9	.379	
Bound-variable-QR	0.16	0.36	0.09	0.25	0.9	.379	
Bound-variable-WH	0.12	0.41	0.02	0.30	1.39	.180	

Table A3: Length effects for Japanese dependencies (cf. Figure 2)

	Short		Long		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect size (r)
Wh-in-situ	-0.01	0.19	0.00	0.17	-0.31	.762	
Binding	-0.08	0.27	-0.11	0.24	0.57	.569	
Bound variable	-0.06	0.19	-0.06	0.21	-0.07	.943	

2. Relative Island effects

Table A4: Two-way repeated measures ANOVAs for English Wh-displacement islands.

	STRUCTURE			LENGTH			INTERACTION S x L		
	F	*p*	eta^2	F	*p*	eta^2	F	*p*	eta^2
ADJ	43.1	.001	.512	80.6	.001	.663	14.4	.001	.260
NP	15.3	.001	.272	64.4	.001	.611	15.4	.001	.272
RC	23.9	.001	.368	34.7	.001	.458	34.0	.001	.453
WH	54.0	.001	.568	40.32	.001	.496	6.45	.015	.136
SUBJ	56.9	.001	.581	3.24	.079		10.8	.002	.208

Table A5: Two-way repeated measures ANOVAs for English Wh-in-situ islands

	STRUCTURE			NUMBER OF WH			INTERACTION S x N		
	F	*p*	eta^2	F	*p*	eta^2	F	*p*	eta^2
ADJ	3.99	.052	.089	85.9	.001	.672	.001	.974	
NP	5.47	.024	.118	40.7	.001	.498	4.29	.047	.098
RC	.075	.785		92.1	.001	.692	.745	.393	
WH	17.4	.001	.298	34.8	.001	4.59	3.56	.056	.093
SUBJ	10.1	.003	.197	8.40	.006	.170	5.73	.032	.110

Table A6: Two-way repeated measures ANOVAs for Japanese Wh-in-situ islands

	STRUCTURE			POSITION			INTERACTION S x P		
	F	*p*	eta^2	F	*p*	eta^2	F	*p*	eta^2
ADJ	1.13	.293		20.4	.001	.299	4.37	.042	.083
NP	.008	.927		31.6	.001	.397	.009	.923	
RC	.482	.491		20.1	.001	.295	.505	.481	
WH	1.28	.264		12.5	.001	.207	.106	.747	
SUBJ	.154	.697		6.45	.014	.118	2.27	.138	

Table A7: English. Paired-samples t-Tests for matrix wh-words with non-island structures versus matrix wh-words with island structures.

	non-island structure		island structure		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect Size (r)
ADJ	0.21	0.20	0.15	0.29	1.464	.151	
NP	0.26	0.21	0.24	0.20	0.582	.564	
WH	0.30	0.21	0.16	0.19	4.582	.001	.58
RC	0.21	0.20	0.24	0.23	1.099	.278	
SUBJ1	0.17	0.22	-0.08	0.29	5.601	.001	.66
SUBJ2	-0.02	0.26	-0.08	0.25	1.477	.147	

Table A8: English. Paired-samples t-Tests for wh-displacement out of non-island structures versus wh-displacemet out of island structures.

	non-island structure		island structure		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect Size (r)
ADJ	0.11	0.29	-0.21	0.25	6.511	.001	.71
NP	0.08	0.27	-0.17	0.36	4.208	.001	.55
WH	0.15	0.30	-0.11	0.25	6.242	.001	.70
RC	0.14	0.21	-0.22	0.41	5.842	.001	.67
SUBJ1	0.20	0.27	-0.29	0.48	6.549	.001	.72

Table A9: English. Paired-samples t-Tests for wh-in-situ in non-island structures versus wh-in-situ in island structures.

	non-island structure		island structure		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect Size (r)
ADJ	-0.04	0.17	-0.09	0.25	1.727	.092	
NP	0.07	0.20	-0.01	0.24	2.367	.023	.35
WH	0.04	0.27	-0.01	0.30	1.593	.119	
RC	-0.03	0.22	-0.05	0.27	0.305	.762	
SUBJ2	-0.08	0.19	-0.19	0.31	2.756	.010	.40

Table A10: Japanese. Paired-samples t-Tests for matrix wh-words with non-island structures versus matrix wh-words with island structures.

	non-island structure		island structure		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect Size (r)
ADJ	.20	.26	.22	.36	.285	.777	
NP	.20	.19	.20	.17	.009	.993	
RC	.15	.17	.19	.35	.932	.356	
WH	.13	.34	.11	.35	.848	.401	
SUBJ	.07	.18	.10	.15	1.06	.294	

Table A11: Japanese. Paired-samples t-Tests for wh-in-situ in non-island structures versus wh-in-situ in island structures.

	non-island structure		island structure		Paired-samples t-test		
	Mean	SD	Mean	SD	t	*p*	Effect Size (r)
ADJ	.09	.16	.17	.38	1.62	.112	
NP	.06	.22	.05	.14	.113	.911	
RC	.05	.27	.05	.27	.184	.855	
WH	.04	.22	.00	.33	.831	.410	
SUBJ	.10	.15	.12	.17	.579	.565	

References

ALEXOPOULOU, T. and F. KELLER. 2007. Locality, Cyclicity and Resumption: At the Interface between the Grammar and the Human Sentence Processor. Language 83:1, 110-160.

BARD, E. G., D. ROBERTSON, and A. SORACE. 1996. Magnitude Estimation of Linguistic Acceptability. Language 72: 32-68.

BOBALJIK, J. 1995. Morphosyntax: The Syntax of Verbal Inflection. Doctoral Dissertation, Massachusetts Institute of Technology.

BOŠKOVIĆ, Z. 2002. On Multiple Wh-Fronting. Linguistic Inquiry 33: 351-383.

CHOMSKY, N. 1973. Conditions on transformations. In A festschrift for Morris Halle, S. Anderson and P. Kiparsky, 237-286. New York, Holt, Rinehart and Winston.

CHOMSKY, N. 1981. Lectures on Government and Binding. Mouton de Gruyter.

CHOMSKY, N. 1986. Barriers. Cambridge, MIT Press.

CHOMSKY, N. 1995. The minimalist program. Cambridge, MIT Press.

CHOMSKY, N. 2001. Derivation by Phase. In Ken Hale: a Life in Language, M. Kenstowicz, 1-52. Cambridge, MIT Press.

COWART, W. 1997. Experimental Syntax: Applying Objective Methods to Sentence Judgments. Sage Publications.

FEATHERSTON, S. 2005. Universals and grammaticality: Wh-constraints in German and English. Linguistics 43: 667-711.

FRAZIER, L., and C. CLIFTON. 1989. Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *4*(2), 93–126.

GIBSON, E. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition*, *68*, 1-76.

GIBSON, E., and T. WARREN. 2004. Reading time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, *7*(1), 55–78.

GROAT, E. and J. O'Neil. 1996. Spell-Out at the LF Interface. In Minimal Ideas, W. Abraham, S. Epstein, H. Thráinsson, and C. J. Zwart, 113-139. John Benjamins Publishing Company.

HAGSTROM, P. 1998. Decomposing Questions. Doctoral dissertation, MIT.

HORNSTEIN, N. 2000. Move! A minimalist theory of construal. Blackwell.

HUANG, C.-T. 1982. Move wh in a language without wh-movement. The Linguistic Review 1: 369-416.

JUST, M., and P. CARPENTER. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.

KAYNE, R. 2002. Pronouns and their antecedents. In Derivation and Explanation in the Minimalist Program, S. Epstein and D. Seely. Oxford, Blackwell.

KELLER, F. 2000. Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality. Doctoral Dissertation, University of Edinburgh.

KLUENDER, R. 1998. On the distinction between strong and weak islands: A processing perspective. *In Syntax and semantics volume 29: The limits of syntax*, ed. by Peter Culicover and Louise McNally, 241–279. Elsevier.

KLUENDER, R. 2004. Are subjects islands subject to a processing account? In Proceedings of WCCFL 23, ed. Angelo J. Rodriguez Vineeta Chand, Ann Kelleher and Benjamin Scheiser, 475–499. Somerville, MA: Cascadilla Press.

KLUENDER, R., and M. KUTAS. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8:573–633.
LASNIK, H. and M. SAITO. 1984. On the nature of proper government. Linguistic Inquiry 15: 235-289.
MIYAMOTO, E.T. and S. TAKAHASHI. 2000. The processing of wh-phrases and Interrogative complementizers in Japanese. In N. Akatuka & S. Strauss(eds.) Japanese/Korean Linguistics, 10. Stanford: CSLI.
NISHIGAUCHI, T. 1990. Quantification in the Theory of Grammar. Dordrecht: Kluwer.
PESETSKY, D. 1998. Some Optimality Principles of Sentence Pronunciation. In Is the Best Good Enough?, P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky, 337-384. Cambridge, MIT Press.
PHILLIPS, C. 2006. The real-time status of island constraints. Language, 82, 795-823.
PHILLIPS, C., N. KAZANINA and S. ABADA. 2005. ERP effects of the processing of syntactic long-distance dependencies. Cognitive Brain Research 22: 407-428.
REINHART, T. 1997. Quantifier Scope: How Labor is Divided Between QR and Choice Functions. Linguistics and Philosophy 20: 335-397.
ROSS, J. 1967. Constraints on Variables in Syntax. Doctoral dissertation, MIT.
SAWADA, M. and R. LARSON (2004) Presupposition & root transformations in adjunct clauses, in M. Wolf and K. Moulton eds., *Proceedings of NELS 34*, UMASS: Amherst.
SCHÜTZE, C. 1996. The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology. The University of Chicago Press.
SPROUSE, J. 2007. A program for experimental syntax. Doctoral Dissertation, University of Maryland.
SPROUSE, J. 2008. The differential sensitivity of acceptability to processing effects. Linguistic Inquiry. 39(4): 686-694.
SORACE, A. and F. Keller. 2004. Gradience in Linguistic Data. Lingua 115: 1497-1524.
STEEDMAN, M. 2002. The syntactic process. MIT Press.
STEPANOV, A. 2007. The End of CED? Minimalism and Extraction Domains. Syntax10: 80-126.
STEVENS, S. S. 1957. On the psychophysical law. Psychological Review 64: 153-181.
TRUSWELL, R. 2007. Locality of Wh-movement and the Individuation of Events. Ph.D. thesis, University College London.
TSAI, W.-T. 1994. On nominal islands and LF extraction in Chinese. Natural Language and Linguistic Theory 12: 121-75.
URIAGEREKA, J. 1999. Multiple spell-out. In Working minimalism, S. Epstein, and N. Hornstein, 251-282. Cambridge, MIT Press.
WACHOWICZ, K. 1974. Against the universality of a single wh-question movement. Foundations of Language 11: 155-166.
WEXLER, K. and P. CULICOVER. 1981. Formal principles of language acquisition. Cambridge, MIT Press.