

An Illusion of Grammaticality in *Wh*-Questions*

Nayoun Kim^{**}, Jiayi Lu
(Sungkyunkwan University, Stanford University)

Kim, Nayoun & Lu, Jiayi. 2021. **An Illusion of Grammaticality in *Wh*-Questions.** *Korean Journal of Linguistics*, 46-3, 699-716. This paper investigates the lingering effects of *wh*-question contexts, involving in particular *wh*-adjuncts and *wh*-arguments. We present novel evidence that initial interpretations of adjunct *wh*-phrases can potentially linger in the subsequent parse, thereby ameliorating the acceptability of apparent island violations. Results from the study's Q/A pair formal acceptability rating experiment reveal that *why*-questions with attachment ambiguity allow the grammatical alternative parse to "linger", augmenting acceptability judgments even when a disambiguating answer sentence forces an ungrammatical or less preferred interpretation. From a methodological perspective, this study also illustrates how the linkage between string acceptability and structural well-formedness could potentially fail. (Sungkyunkwan University, Stanford University)

Key words: lingering effect, illusion of grammaticality, experimental syntax, *wh*-questions, islands

1. Introduction

In this study, we investigate the lingering effects in *wh*-question contexts, involving *wh*-adjuncts and *wh*-arguments. It has been shown that during sentence processing, the structural representations that the readers temporarily entertained could remain in memory even after structural reanalysis happens, giving rise to a "lingering effect" (Christianson, Hollingworth, Halliwell, Ferreira, 2001, Ferreira, Christianson, Hollingworth. 2001, Slattery et al. 2013, Sturt 2007, Van Gompel et al. 2006). Through the study of *wh*-questions, this study aims to show that the lingering effect can lead to an illusion of grammaticality.

Why-questions in English give rise to structural ambiguities that we can

* We are grateful to the anonymous reviewers and the members in the Syntax, Semantics, and Sentence Processing Lab at Northwestern University for valuable comments and suggestions.

** First & Corresponding Author

use to test for lingering effects. In *why*-questions containing embedded clauses, the gap of *why* could be situated in the matrix clause (an “upstairs” interpretation) or inside the embedded clause (a “downstairs” interpretation). We employed a Question-Answer pair acceptability rating paradigm to test what happens when a disambiguating answer to the ambiguous *why*-question forces an illicit structure. We show that even when the disambiguating answer forces an ungrammatical island-violating parse, the alternative grammatical parse that the readers built prior to reanalysis seems to linger, boosting acceptability ratings.

2. Background

2.1 Attachment Ambiguity in *Why*-Question

Unlike *wh*-questions with argument *wh*-elements (e.g. *who*, *what*) that leave easily identifiable gaps, *why*-questions can be ambiguous with regard to the origination site for *why*. Consider the cases where a *why*-question contains an embedded clause. In (1), the gap for *why* could reside either in the matrix clause, yielding an “upstairs interpretation” (the location of the gap is upstairs to the embedded clause in the structure), or in the embedded clause, yielding a “downstairs interpretation” (as the gap has a low structural position, downstairs to the matrix clause). The two possible gap positions are shown in (1). If the upstairs interpretation is adopted, *why* then modifies the matrix verb. A plausible answer for the upstairs interpretation would be (1a) where the question asks about the reason for thinking. If the downstairs interpretation is adopted, then *why* modifies the embedded verb, e.g., *quit*. A plausible answer for the downstairs interpretation would be (1b) where the question asks about the reason for quitting the job (Shlonsky & Soare 2011). We will refer to this ambiguity in the positions of *why*-gaps as the “attachment ambiguity” of *why*-questions.¹

¹ We do not discuss dual properties of *whys* (reason and purpose *why*) because this study is concerned with the attachment ambiguity captured in bi-clausal *why*-questions. Some studies argue for the two different *whys* in English yielding different interpretations with respect to two distinct syntactic dependency relations (see Chapman & Kučerová 2016,

- (1) Q: [_{CP} *Why do you* (1) *think* [_{CP} *that she* (2) *quit the job*]]?
 a. Because she seems to be sad.
 b. Because she did not like the job.

2.2 Lingering Effect in *Why*-Question

In the present study, we investigate whether the attachment ambiguity of *why*-questions can give rise to the “lingering” effect. Our research question is motivated by the well-documented garden-path effect in the resolution of syntactic ambiguities. In reading sentences like (2), *the deer* is likely to be initially parsed as an object of the transitive verb (*hunted*) in the subordinate clause. At the region, *ran into*, the transitive analysis where *the deer* is an object turns out to be globally incompatible, which necessitates a reanalysis. Slower reading times at the disambiguation region (*ran*) has been ascribed to the reanalysis process (Ferreira & Henderson 1991, Frazier & Rayner 1982, Pickering & Traxler 1998, Sturt, Pickering, Crocker 1999).

- (2) While the man hunted the deer ran into the woods.
 (Christianson et al. 2001: 369)

Although the final interpretation is successfully disambiguated at the word *ran*, readers seem to nevertheless keep entertaining the initial erroneous parse in memory (Christianson et al. 2001, Ferreira et al. 2001, Slattery et al. 2013, Sturt 2007). Christianson et al. (2001) tested whether readers can completely reanalyze the structure when the initial structure turns out to be globally incompatible with the input during sentence processing. When asked to answer comprehension questions like ‘*Did the hunter shoot the deer?*’ after reading sentences like (2), around 70% of the participants gave incorrect answers. The high rate of the incorrect answers can be explained by the wrong parse persisting in spite of reanalysis. On this view, readers were still impacted by their initial parse of the sentence where *the deer* serves as a direct object of the verb, *hunted*. The structure that is globally erroneous yet locally

Kim, Wellwood, & Yoshida under revision, and Shlonsky & Soare 2011).

permissible is still preserved by the readers even after structural reanalysis, leading to a “lingering effect”².

Now, let us look at how lingering effects could manifest themselves in the processing of *wh*-questions. Consider when one of the possible attachment sites of *why* in a *why*-question is ruled out by island constraints. In the long-distance dependencies like (3), the distance between the dependent element (*why*) and the controlling element (gap) which governs the interpretation and the grammatical function of the dependent element could, by and large, be infinitely far apart.

- (3) Why did he meet the girl at the party?

Despite such unboundness nature in a long-distance dependencies, *why* cannot be linked to the gap inside the relative clause as in (4) nor inside a *whether*-clause as in (5). Certain structures that constrain the long-distance dependencies are referred to as islands (Chomsky 1973, 1977, 1986, Ross 1967; refer to Phillips 2013, Sprouse, Wagers, & Phillips 2012, and Yoshida et al. 2014 for experimental studies on syntactic accounts of islands).³

- (4) Q: [_{CP} *Why did he* √ __ *meet* [_{NP} *the girl* [_{CP} *who* * __ *quit the job*]]?

- (5) Q: [_{CP} *Why did he* √ __ *meet* [_{island} *whether the girl* * __ *quit the job*]]?

In (4), the relative clause constitutes an island that blocks *wh*-dependencies. The interpretation where *why* originates upstairs (in the matrix clause) is possible, but the interpretation where *why* originates downstairs (inside the relative clause) is not possible. This is because the downstairs interpretation of *why*

² Similar results have been widely reported in the sentence processing literature (Fanselow & Frisch 2006, Ferreira & Henderson 1991, Fujita & Cummings 2020, Frazier & Rayner 1982, Pickering & Traxler 1998, Ferreira & Patson 2007, Sturt 2007, Slaterry et al. 2013).

³ Note that there are alternative approaches to account for the sources of island effects. Island effects could be attributed to the processing overload arising from holding the dependent element in an open dependency while processing the complex domain which is itself difficult (Christensen & Nyvad 2014, Deane 1991, Hofmeister & Sag 2010, Kluender & Kutas 1993) or to the extra-syntactic considerations such as pragmatic/referential mechanisms (Abrusán 2011, Ambridge & Goldberg 2008, Deane 1991, Erteschik-Shir 1973).

involves a *wh*-dependency that crosses the relative clause, which constitutes an island. Same applies in (5) where only the answer targeting the upstairs interpretation is possible, but the answer targeting the downstairs interpretation is not possible because the downstairs interpretation of *why* involves a *wh*-dependency that crosses the *whether*-clause.

Now consider when the answer sentence forces the downstairs interpretation of the *why*-question (in the case of A2 in (6) and (7)).

(6) Q: [_{CP} *Why did he* ✓__ *meet* [_{NP} *the girl* [_{CP} *who* *__ *quit the job*]]?⁴

A1: ✓Because he wanted to ask something.

A2: #She quit because she didn't like the job.

(7) Q: [_{CP} *Why did he* ✓__ *wonder* [_{Island} *whether the girl* *__ *quit the job*]]?

A1: ✓Because he cares about her.

A2: # Because she is moving to Chicago.

In these examples, the question involves two attachment sites (high-attachment and low-attachment structures). The interpretation yielded by the high-attachment structure is totally acceptable but that of the low-attachment violates islands. In this case, the grammatical high-attachment structure might still “linger” to ameliorate the island violation even when the answer sentence forces the low-attachment structure. If such a lingering effect does exist, readers would face an illusion that the *why*-questions in (6) and (7) are acceptable even with the low-attachment structural analyses. Specifically, although the answer targeting the downstairs interpretation is not allowed due to islands, the initial syntactic analysis from upstairs structure may remain in memory, subsequently ameliorating the later parse with island violations. Note that such a boost of acceptability is not possible in argument *wh*-questions such as *who*-questions or *what*-questions that do not give rise to attachment ambiguities.

In the present study, we conducted an experiment to test whether the attachment ambiguity of *why* yields lingering effects that can potentially boost

⁴ The unacceptability of the particular sentence is indicated by an asterisk, *.

sentence acceptability. In particular, we tested if island violations can be ameliorated in *why*-questions with attachment ambiguities. If so, we expect smaller island sensitivity in potentially ambiguous yet disambiguated *why*-questions compared to argument *wh*-questions without attachment ambiguities.

2.3 Rationale behind the Factorial Design

Formal methodologies used in experimental syntax allow for more systematic investigations on islands (Sprouse 2007, Sprouse et al. 2012, Dillon & Hornstein 2013, Sprouse & Hornstein 2013, Kush, Lohndal, & Sprouse 2018). Recent studies on experimental syntax have demonstrated that two factors contribute to islands: the dependency length (the structural/linear distance between the dependent element and the gap) and the availability of islands (the presence and absence of islands). In general, sentences with longer dependencies crossing the clausal boundary are judged less acceptable compared to shorter dependencies without crossing the clausal boundary. Furthermore, sentences with island structures are judged to be less acceptable compared to those that do not violate such constraints. However, noteworthy recent findings reveal that longer dependencies with island structures lead to even lower acceptability ratings than short dependencies with island structures (Sprouse et al. 2012, Dillon & Hornstein 2013, Sprouse & Hornstein 2013, Kush et al. 2018). This so-called “super-additive” effect can be accounted for by assuming that additional degradation in acceptability occurs when the long-distance extraction and island structure combine together.

In our study, we test for island sensitivity as the super-additive interaction of *Dependency distance* (crossing clausal boundary/long *wh*-dependency vs. not crossing boundary/short *wh*-dependency) and *Embedded structure* (island vs. non-island), which isolates island effect from acceptability degradations due to the main effects of syntactic dependency distance and the types of embedded clause structures used (Sprouse 2007, Sprouse et al. 2012, Dillon & Hornstein 2013, Sprouse & Hornstein 2013, Kush et al. 2018).

3. The Experiment

3.1 Participants

Eighty undergraduate students from Northwestern University were recruited as participants. They participated in the study as a reward for course credits offered at Northwestern University. All of them self-reported as native speakers of English and provided informed consent.

3.2 Design

A total of twenty-four items were prepared. Experimental items were presented as Q-A pairs (Table 1). We employed a 2X2X2 design with *Wh-category* (argument vs adjunct), *Distance* (crossing clausal boundary/long *wh*-dependency vs. not crossing boundary/short *wh*-dependency), and *Structure* (island vs non-island). Argument-*wh* items are *who*- and *what*-questions, and adjunct-*wh*-items are *why*-questions. The dependency distance of the adjunct-*wh*-items is forced by the answer sentences that accompany the target sentences. *Whether*-clauses are used as the island construction, and complement clauses are used as the non-island construction. In *wh*-adjunct items, the answer sentences always begin with a *Because*-clause (requesting for the reason behind an agent performing a particular action), and are used to disambiguate attachment ambiguity. In *wh*-argument items, there is no attachment ambiguity.

The experimental stimuli were pseudo-randomized based on a Latin square design in order to inhibit experimental stimuli of the same condition from appearing adjacent to one another. These stimuli were intertwined with another 24 fillers which were not relevant to the manipulations used in the current study and matched in complexity as well as length. A sample set of stimuli with a total of eight conditions is shown in Table 1.

Table 1. Example Set of Stimuli for the Experiment

	island	non-island
wh-argument		
Short-Dependency	Q: Who speculates whether the musician sang a jazz song? A: A drummer.	Q: Who speculates that the musician sang a jazz song? A: A drummer.
Long-Dependency	Q: What does the drummer speculate whether the musician sang? A: A jazz song.	Q: What does the drummer speculate that the musician sang? A: A jazz song.
wh-adjunct		
Short-Dependency	Q: Why does the drummer speculate whether the musician sang a jazz song? A: Because the drummer was looking for a singer.	Q: Why does the drummer speculate that the musician sang a jazz song? A: Because the drummer was looking for a singer.
Long-Dependency	Q: Why does the drummer speculate whether the musician sang a jazz song? A: Because the musician only knows jazz.	Q: Why does the drummer speculate that the musician sang a jazz song? A: Because the musician only knows jazz.

3.3 Procedure

Subjects participated in this experiment in a sound-proof lab, and were asked to read the sentences presented on a desktop PC using Linger software (Rohde 2003). They saw Q/A pair sentences, and were asked to rate how natural the conversation (between two interlocutors, A and B) is. The rating of 1 indicates totally unnatural and the rating of 7 indicates totally natural. They were also advised that there is no right or inadmissible answer, and to judge the conversation between two interlocutors based on the naturalness. They were given 7 practice sentences prior to the actual experiment to get familiarized with the experiment procedure. The whole experiment took approximately 20 minutes to complete,

and a short break was given to participants if needed.

3.4 Analysis

Data were analyzed using linear mixed effect regression models carried out by means of the *lme4* package in R version 3.2.3 (Baayen 2008, Baayen et al. 2008, Bates et al. 2015, Jaeger 2008). All models contained the maximal random effects structure (Barr et al. 2013), which involved random intercepts for participants and items together with random slopes for fixed effects and their interaction, supposing that the model successfully converged.

3.5 Predictions

Island effects are captured as a super-additive interaction between (Dependency) *Distance* and *Structure* (island vs. non-island) as shown in previous studies (Sprouse et al. 2012). If the attachment site ambiguity of *why* serves to ameliorate island effects, but the lack of ambiguity with *who* does not, we would expect a super-additive interaction between *Distance* and *Structure* such that *why* should mitigate island sensitivity more than *who* in island conditions. The difference in island sensitivity between the two different *wh*-questions should be characterized as a significant super-additive interaction between *Distance*, *Structure*, and *Wh-category*. However, if no amelioration arises from ambiguous attachment sites, we expect to observe no three-way interaction between *Distance*, *Structure*, and *Wh-category* but only a main effect of *Distance*. In non-island conditions, *why* should exhibit higher island sensitivity compared to *who* with obvious gaps (Lu et al. 2020).

3.6 Results

The mean acceptability ratings of each condition are shown in Figure 1. A main effect of *Structure* was found such that non-island constructions were rated higher than constructions involving islands ($\beta = -1.63$, $SE = 0.19$, $t = -8.43$). A marginal main effect of *Distance* was also observed in a direction where

long dependencies were rated lower than short dependencies ($\beta = -0.82$, $SE = 0.13$, $t = -6.31$). A main effect of *Wh-category* was also observed such that arguments were rated higher than adjuncts in general ($\beta = -0.15$, $SE = 0.08$, $t = -1.97$). A significant super-additive *Distance X Wh-category* interaction effect was found ($\beta = 0.90$, $SE = 0.13$, $t = 7.21$); a further subset analysis shows that the *Distance X Wh-category* is significant in the island items ($\beta = 1.52$, $SE = 0.23$, $t = 6.65$) but not in the non-island items ($\beta = 0.29$, $SE = 0.30$, $t = 0.97$).

A significant *Distance X Structure* interaction effect was found ($\beta = -1.11$, $SE = 0.14$, $t = -8.85$). A subset analysis shows that the *Distance X Structure* interaction is significant in the *wh*-argument items ($\beta = -1.72$, $SE = 0.25$, $t = -7.00$) but not in the *wh*-adjunct items ($\beta = -0.52$, $SE = 0.31$, $t = -1.67$). This suggests that the *who*-questions show island sensitivity, yet the *why*-questions do not show island sensitivity, which is compatible with our hypothesis that the *why*-questions with attachment ambiguities, despite being disambiguated by the answer sentences, still induce a lingering effect which ameliorates the island violation. Further corroborating this finding was a significant three-way interaction ($\beta = 1.21$, $SE = 0.25$, $t = 4.82$), showing that *wh*-arguments have greater island sensitivity than *wh*-adjuncts.

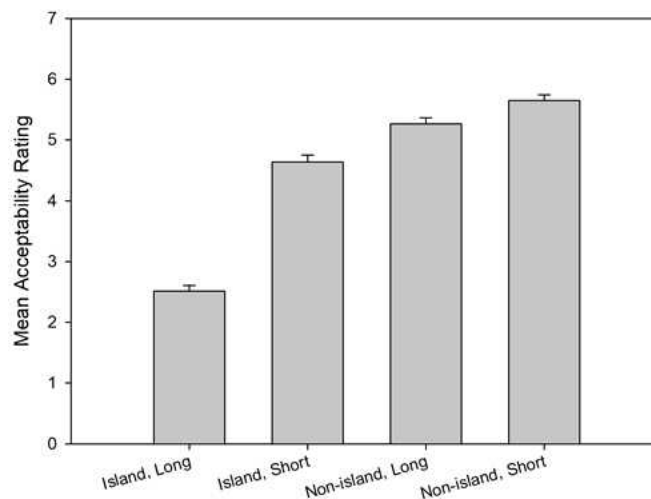
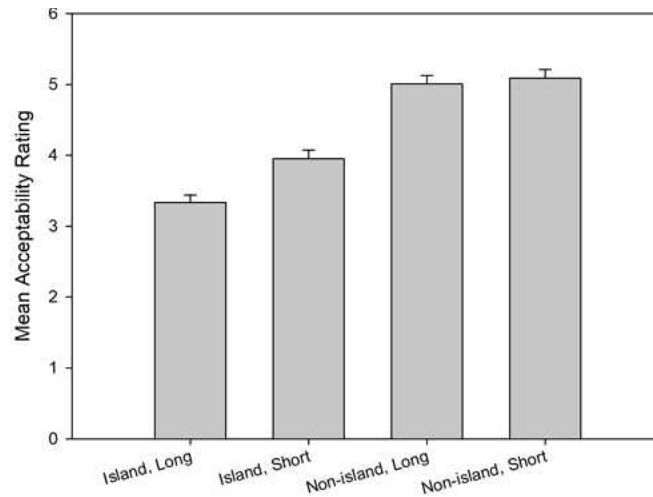


Figure 1. Acceptability ratings of *Wh*-argument items

Figure 2. Acceptability ratings of *Wh*-adjunct itemsTable 2. Statistical Analysis: Estimates, Standard Error, and t-value from a linear mixed effect model for Experiment ($|t| > 2$ indicating significance)

	<i>Estimate</i>	<i>SE</i>	<i>t-value</i>
(Intercept)	4.43	0.11	42.11
Structure	-1.63	0.19	-8.43
Distance	0.82	0.13	-6.31
Wh-category	-0.15	0.08	-1.97
Structure x Distance	-1.11	0.13	-8.85
Distance x Wh-category	0.90	0.13	7.21
Structure x Wh-category	0.46	0.13	3.66
Structure x Distance x Wh-category	1.21	0.25	4.82

4. Discussion

In the experiment, we tested whether the attachment ambiguity in *wh*-questions can lead to the lingering effect. To address this question, we employed a Question-Answer pair paradigm to force the island-violating

downstairs interpretations in *why*-questions. Our results show that the island effect, detected as a super-additive interaction of *Distance X Structure*, is significant in the *wh*-argument conditions (i.e. *who*-questions) but not in the *wh*-adjunct conditions (i.e. *why*-questions). We also find a significant three-way interaction of *Distance X Structure X Wh-category*, where the island sensitivity of *wh*-arguments is greater than that of *wh*-adjuncts, further corroborating the finding that the island effect in *why*-questions is ameliorated.

The amelioration of the island effect in *why*-questions but not *who*-questions supports our hypothesis that the attachment ambiguity in *why*-questions causes a lingering effect. Since the position of origination of *why* is ambiguous between the matrix clause and the embedded clause in our test sentences, the parse with the matrix gap (the upstairs interpretation of *why*) remains a possibility to the reader before reaching the disambiguating answer sentence. In fact, when the embedded clause constitutes an island, the parse with the embedded gap (the downstairs interpretation of *why*) yields an island violation and should be strongly dispreferred, while the upstairs interpretation is the only grammatical parse and thus should be the strongly preferred. When the disambiguating answer sentence coerces the island-violating downstairs interpretation of *why*, the readers would reanalyze the sentence structure as containing an embedded gap rather than a matrix gap. Since the forced downstairs interpretation is island-violating, the sentence should be rated as unacceptable by the readers. However, since the previous parse with the upstairs interpretation of *why* has been previously entertained by the readers before reanalysis, it may still “linger” in the readers’ memory. Since this lingering parse does not give rise to any island violation and is perfectly grammatical, its lingering would boost the acceptability of the entire sentence.

On the other hand, *who* in *who*-questions has a unique base position marked by the overt gap. When the *wh*-dependency crosses an island boundary, an island violation would arise. Since there is no structural reanalysis, there would be no lingering parse to ameliorate the island violation. Therefore, compared to *why*-questions, *who*-questions should show greater island sensitivity. These predictions are supported by our results: the island effect is only detected in *who*-questions but not *why*-questions, and the significant *Distance X Structure X Wh-category* interaction suggests that *who* is more island sensitive than

why in this experiment.

Notably, the Empty Category Principle (ECP) (Chomsky 1981, 1986, Huang 1982, Lasnik and Saito 1984, 1992, Cheng 2009) also predicts a difference between the island sensitivity of certain *wh*-arguments and *wh*-adjuncts. However, the direction of the effect predicted by ECP is the opposite of what we observed. The difference between argument and adjunct asymmetry in terms of the island effect has been attributed to their configurational/structural positions. Direct objects are analyzed as the sister of a lexical head but the adjuncts are not the sister of the lexical head (they are sister of the X-bar projection). According to the ECP, the trace left by *wh*-movement must be licensed and the licensing condition crucially refers to these structural differences between argument and adjunct. Under ECP accounts, a *wh*-trace must be licensed. A *wh*-trace is licensed either when the trace is the sister of the lexical head or the trace is locally bound by the *wh*-phrase or another *wh*-trace. In the case of direct object trace, it is licensed because it is the sister of the lexical head. On the other hand, the adjunct trace must be bound by the nearby *wh*-phrase or *wh*-trace. In *why*-questions with a downstairs interpretation of *why*, the intermediate trace of *why* (a *wh*-adjunct) is not properly governed. This leads to violation of the ECP, yielding degradation in acceptability. In argument *wh*-questions like *who* questions, there is no violation of the ECP. This predicts that the island effect should be more robust in *why*-questions than *who*-questions, which is the opposite of what we observed. Therefore, our findings cannot be simply reduced to the ECP effect⁵.

On a broader level, one important implication of our findings concerns the use of string acceptability in experimental syntax research. The existence of lingering effects like the one we observed suggests that the acceptability rating of a string is affected by not only the final parse of the string that the reader commits to, but also the well-formedness of other parses that are previously entertained by the reader. Ideally, acceptability judgment experiments should inform the well-formedness of a particular structural analysis of a string. Yet when readers' judgments are affected by the previously considered parses

⁵ Note that we are not making any claim against the ECP itself: it is likely that the ECP effect still stands, but the lingering effect outweighs the ECP effect in determining the direction of the *Distance X Structure X Wh-category* interaction in this particular study.

of a string (as in the case of this experiment), it is unclear what the final judgment scores can tell us about the grammaticality of the final parse. Ott (2017) also raised similar concerns about the use of string acceptability in the study of Generative Grammar: he pointed out that the acceptability of strings does not constitute valid empirical evidence for the study of strong generative capacity, because string acceptability cannot be directly employed to study structural well-formedness. Our findings in this study supports Ott (2017)'s point by providing a concrete example of how the linkage between string acceptability and structural well-formedness is challenged: we showed that string acceptability may be affected by multiple structures, and thus cannot trivially map to the well-formedness of any single parse.

Lastly, we want to point out this study raises an intriguing follow-up question. It is possible that the answer sentences in the question-answer pairs are not successfully disambiguating the question sentence. In particular, the downstairs interpretation of *why*-questions with an embedded island structure might be so dispreferred due to its ungrammaticality that the answer sentence fails to force such a parse. As a result, there would be no structural reanalysis happening, and the participants would be giving ratings solely based on the upstairs interpretation of the *why*-questions. However, this does not suggest that the lingering effect is not happening. On the contrary, this would be an extreme case of the lingering effect, one where the original parse is so committed to linger that the reader ignores the disambiguating information and considers only the original parse of a sentence when giving ratings. One way to examine this possibility is to test what happens when neither of the possible interpretations of a *why*-question is forced (e.g., when no answer is provided). If the disambiguating answers in the current study truly fail to disambiguate, there should be no change in acceptability when no answer sentences are provided. Alternatively, we can test the reliability of disambiguating answer sentences using comprehension questions. By directly probing the reader's interpretation of the critical sentence, we can know whether the disambiguating sentences are effective. We leave this for future research.

5. Conclusion

In this study, we aim to show that when readers are processing *wh*-questions with attachment ambiguities for the *wh*-phrases, the temporarily entertained interpretations of the sentence can potentially linger even when the readers commit to a different parse after disambiguation. When the final parse forced by the disambiguating information results in ungrammaticality while the previously entertained parse is grammatical, the lingering of the previous parse leads to amelioration of sentence unacceptability.

Results from the Q/A pair acceptability rating experiment support our hypothesis. Our findings reveal that *why*-questions containing island structures allow the grammatical matrix *wh*-gap parse to linger even when the disambiguating sentence forces the island-violating embedded *wh*-gap parse.

References

- Abrusán, Márta. 2011. "Presuppositional and negative islands: A semantic account," *Natural Language Semantics* 19(3), 257-321.
- Ambridge, Ben and Adele E. Goldberg. 2008. "The island status of clausal complements: Evidence in favor of an information structure explanation," *Cognitive Linguistics* 19(3), 357-389.
- Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. Harald, Douglas J. Davidson, and Douglas M. Bates. 2008. "Mixed-effects modeling with crossed random effects for subjects and items," *Journal of Memory and Language* 59(4), 390-412.
- Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *Journal of Memory and Language* 68(3), 255-278.
- Bates, Douglas M., Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software* 67(1), 1-48.
- Chapman, Cassandra and Ivona Kučerová. 2016. "Structural and semantic ambiguity of *why*-questions: An overlooked case of weak islands in English," *Proceedings of the Linguistic Society of America* 1(15), 1-15.
- Cheng, Lisa Lai-Shen. 2009. "Wh-in-situ, from the 1980s to now," *Language and*

Linguistics Compass 3(3), 767–791.

- Chomsky, Noam. 1973. "Conditions on transformations," in Stephen R. Anderson and Paul Kiparsky, eds., *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1977. "On wh-movement," in Peter Culicover, Thomas Wasow, and Adrian Akmajian, eds., *Formal syntax*, 71–132. New York: Academic Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Christensen, Ken Ramshøj and Anne Mette Nyvad. 2014. "On the nature of escapable relative islands," *Nordic Journal of Linguistics* 37(1), 29–45.
- Christianson, Kiel, Andrew Hollingworth, John F. Halliwell, and Fernanda Ferreira. 2001. "Thematic roles assigned along the garden path linger," *Cognitive Psychology* 42(4), 368–407.
- Deane, Paul. 1991. "Limits to attention: A cognitive theory of island phenomena," *Cognitive linguistics* 2(1), 1–64.
- Dillon, Brian and Norbert Hornstein. 2013. "On the structural nature of island constraints," in Jon Sprouse and Norbert Hornstein, eds., *Experimental syntax and island effects*, 208–222. Cambridge: Cambridge University Press.
- Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: MIT dissertation.
- Fanselow, Gisbert, and Stefan Frisch. 2006. "Effects of processing difficulty on judgements of acceptability," in Gisbert Fanselow, Caroline Féry, Ralf Vogel, and Matthias Schlesewsky, eds., *Gradience in grammar: Generative perspectives*, 291–316. Oxford: Oxford University Press.
- Ferreira, Fernanda and John M. Henderson. 1991. "Recovery from misanalyses of garden-path sentences," *Journal of Memory and Language* 30(6), 725–745.
- Ferreira, Fernanda, Kiel Christianson, and Andrew Hollingworth. 2001. "Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis," *Journal of Psycholinguistic Research* 30(1), 3–20.
- Ferreira, Fernanda and Nikole D. Patson. 2007. "The 'good enough' approach to language comprehension," *Language and Linguistics Compass* 1, 71–83.
- Frazier, Lyn and Keith Rayner. 1982. "Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences," *Cognitive Psychology* 14(2), 178–210.
- Fujita, Hiroki and Ian Cunnings. 2020. "Reanalysis and lingering misinterpretation of linguistic dependencies in native and non-native sentence comprehension," *Journal of Memory and Language*, 115.
- Hofmeister, Philip and Ivan A. Sag. 2010. "Cognitive constraints and island effects," *Language* 86(2), 366–415.
- Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: MIT dissertation.
- Jaeger, T. Florian. 2008. "Categorical data analysis: Away from ANOVAs

- (transformation or not) and towards logit mixed models," *Journal of Memory and Language* 59(4), 434-446.
- Kim, Nayoun, Alexis Wellwood, and Masaya Yoshida. under revision. "Processing wh-phrases: how *who*, *how*, and two *whys* interact with memory."
- Kluender, Robert and Marta Kutas. 1993. "Subjacency as a processing phenomenon," *Language and Cognitive Processes* 8(4), 573-633.
- Kush, Dave, Terje Lohndal, and Jon Sprouse. 2018. "Investigating variation in island effects," *Natural Language & Linguistic Theory* 36(3), 743-779.
- Lasnik, Howard and Mamoru Saito. 1984. "On the nature of proper government," *Linguistic Inquiry* 15(2), 235-289.
- Lasnik, Howard and Mamoru Saito. 1992. *Move α : Conditions on its application and output*. Cambridge, MA: MIT Press.
- Lu, Jiayi, Cynthia K. Thompson, and Masaya Yoshida. 2020. "Chinese wh-in-situ and islands: A formal judgment study," *Linguistic Inquiry* 51(3), 611-623.
- Ott, Dennis. 2017. "Strong generative capacity and the empirical base of linguistic theory," *Frontiers in Psychology* 8, 1617.
- Phillips, Colin. 2013. "On the nature of island constraints I: Language processing and reductionist accounts," in Jon Sprouse and Norbert Hornstein, eds., *Experimental syntax and island effects*, 64-108. Cambridge: Cambridge University Press.
- Pickering, Martin J. and Matthew J. Traxler. 1998. "Plausibility and recovery from garden paths: An eye-tracking study," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(4), 940-961.
- Rohde, Doug. 2003. Linger: A flexible platform for language processing experiments, version 2.94. Available at <http://tedlab.mit.edu/~dr/Linger/>.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation.
- Shlonsky, Ur and Gabriela Soare. 2011. "Where's 'Why'?", *Linguistic Inquiry* 42(4), 651-669.
- Sturt, Patrick. 2007. "Semantic re-interpretation and garden path recovery," *Cognition* 105(2), 477-488.
- Sturt, Patrick, Martin J. Pickering, and Matthew W. Crocker. 1999. "Structural change and reanalysis difficulty in language comprehension," *Journal of Memory and Language* 40(1), 136-150.
- Slattery, Timothy J., Patrick Sturt, Kiel Christianson, Masaya Yoshida, and Fernanda Ferreira. 2013. "Lingering misinterpretations of garden path sentences arise from competing syntactic representations," *Journal of Memory and Language* 69(2), 104-120.
- Sprouse, Jon. 2007. *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge*. College Park, MD: University of Maryland dissertation.
- Sprouse, Jon and Norbert Hornstein, eds. 2013. *Experimental syntax and island effects*. Cambridge: Cambridge University Press.

- Sprouse, Jon, Matt Wagers, and Colin Phillips. 2012. "A test of the relation between working-memory capacity and syntactic island effects," *Language* 88(1), 82-123.
- Van Gompel, Roger P.G., Martin J. Pickering, Jamie Pearson, and Gunnar Jacob. 2006. "The activation of inappropriate analyses in garden-path sentences: Evidence from structural priming," *Journal of Memory and Language* 55(3), 335-362.
- Yoshida, Masaya, Nina Kazanina, Leticia Pablos, and Patrick Sturt. 2014. "On the origin of islands," *Language, Cognition and Neuroscience* 29(7), 761-770.

Nayoun Kim, Assistant Professor
Department of English Language and Literature, Sungkyunkwan University
25-2 Sungkyunkwan-ro, Jongro-gu, Seoul 03063, Republic of Korea
E-mail: nayoun@skku.edu

Jiayi Lu, Ph.D. Student
Department of Linguistics, Stanford University
Margaret Jacks Hall, Building 460, CA 94305-2150, United States
E-mail: jiayi.lu@stanford.edu

Received: July 27, 2021
Revised: September 12, 2021
Accepted: September 29, 2021