Age-Related Effects on Constraints on *Wh*-movement

Boyoung Kim and Grant Goodall
University of California, San Diego

1. Introduction

One of the standard ways to study the effect of age on second language acquisition is to explore the relation between learners' age of arrival (AoA) in the new language environment and their ultimate attainment in that language. Many studies of this type have used a battery of morphosyntactic tests to measure ultimate attainment (e.g. Johnson and Newport (1989), Birdsong and Molis (2001), DeKeyser, Alfi-Shabtay and Ravid (2010)), while others have used more holistic or impressionistic measures (e.g. Flege (1999), Hakuta, Bialystok and Wiley (2003)). There is no reason to expect that all areas of language, or even all areas within a domain such as syntax, will be uniformly affected by age of arrival, however, and syntactic theory gives us a meaningful way to look for particular properties that might be differentially affected. Moreover, recent techniques in experimental syntax allow us to detect possible differences with great precision (Cowart (1997), Myers (2009)). These techniques include the use of a factorial design in the construction of sentence stimuli, a counterbalanced and randomized stimuli list, and perhaps most importantly, a response method that allows subjects to express the full range of perceived differences in acceptability.

In this study, we look at two different areas of syntax, island constraints and the *that*-trace effect, and we employ the techniques of experimental syntax to explore differences among particular sentence types and speaker populations. The specific island constraints that we examine are the *Wh*-island Constraint and the Complex Noun Phrase Constraint (CNPC), shown in (1) and (2), respectively.

(1) *Who do you wonder [whether [Ann saw __]]?

(2) *Who do you believe [the claim that [Ann saw __]]?

The *that*-trace effect is exemplified in (3).

(3) *Who do you think [that [__ will see Mary]]?

In much recent work (e.g. Kluender (1998, 2004), Hofmeister and Sag (2010)), the idea has been pursued that some island phenomena may be the result of capacity constraints on processing, rather than grammatical constraints per se. Although the issue is far from settled, we will tentatively adopt this view here. With regard to the *Wh*-island Constraint and the CNPC, the core idea is that these island domains are particularly difficult for the processor to handle, and that when combined with the filler-gap dependency produced by *wh*-movement, the result overwhelms the processor's resources and the sentence is perceived as unacceptable.

The *that*-trace effect appears to be much less amenable to an account along these lines, and most analyses assume that it is at heart a grammatical effect. For concreteness, we adopt one particular account of this type, that proposed in Rizzi and Shlonsky (2007). In this analysis, the Subject Criterion requires that elements moving into SPEC of TP be "frozen" in that position (i.e. unavailable for further movement), so *wh*-movement out of SPEC of TP should be impossible. For a language to permit a

* We thank the audience at GASLA 11 for their many stimulating questions and suggestions. Thanks also to members of the Experimental Syntax Lab at UCSD, and to audiences at WECOL and LSA, where related work was presented. All errors remain our own.

 Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference (GASLA 2011), ed. Julia Herschensohn and Darren Tanner, 54-62. Somerville, MA: Cascadilla Proceedings Project.

subject *wh*-question, then, it must avoid moving the subject through this position, and the particular way that English accomplishes this is by truncating the upper part of the embedded clause (including both C and SPEC of TP), allowing the subject to move directly out of a lower position. This is illustrated in (4), where the CP and TP domains (in italics) are truncated.

(4) Who do you think ***[$_{CP}$ that [$_{TP}$*** __ [$_{VP}$ __ will see Mary]]?

This results in the celebrated contrast between (3) and (5).

(5) Who do you think [__ will see Mary]?

(3) is clearly ungrammatical under this account, because the presence of *that* indicates that truncation could not have occurred, so the *wh*-phrase would have moved through SPEC of TP, violating the Subject Criterion.

2. Implications for acquisition and AoA

We have just sketched an account where the *Wh*-island Constraint and the CNPC result from processing limitations and the *that*-trace effect follows from a grammatical constraint. If this basic division of labor is correct, there are clear implications for acquisition. With regard to the island constraints, for instance, we expect L2 speakers to be no better, and probably worse, than native speakers in their ability to process difficult filler-gap dependencies with very high capacity demands (see Michael and Gollan (2005) for related discussion). We thus expect L2 speakers to display island effects similar to those of native speakers. Moreover, we expect these effects to show up relatively quickly, i.e. as soon as they have acquired *wh*-movement and the relevant structures (embedded *wh*-questions and complex NPs), so major age-related differences are not predicted. Anyone who has had enough exposure to the language to acquire the structures should thus display island effects, regardless of the age of arrival.

With regard to the *that*-trace effect, the situation is very different. Assuming that the Subject Criterion follows from the basic architecture of the syntax, learners will know that extraction from SPEC of TP is impossible. Subject *wh*-questions should thus be disallowed, and to the extent that examples like (4) are noticed, these should pose a puzzle. Learners must posit some additional mechanism to allow (4), and this mechanism must be compatible with the rest of their grammar (thus excluding options such as extraction from a postverbal subject position, as occurs in other languages - see Rizzi and Shlonsky (2007)). The evidence that truncation is the appropriate mechanism is indirect and amounts essentially to a process of eliminating the other possibilities. We thus expect the emergence of the *that*-trace effect in L2 speakers to be slow (see Gathercole (2002)). After acquiring the appropriate structures (embedded *that*-clauses and *wh*-movement), learners must notice examples like (4) in the input and then consider the various ways that (4) might be generated without violating the Subject Criterion. If they conclude that truncation is the relevant mechanism, they should then exhibit the classic *that*-trace effect, allowing embedded subject extraction only when *that* is absent but being indifferent to the presence or absence of *that* otherwise. It is reasonable to expect that this would be a very extended developmental process, and that different AoAs might be differentially able to complete it.

In this article, we explore these predictions by means of two experiments: one which investigates the extent to which AoA affects whether learners display island effects, and one which does the same with regard to the *that*-trace effect.

3. Experiment 1: *Wh*-islands and CNPC

3.1. Subjects

Three groups participated in this experiment. The first two groups were born in Korea and immigrated to the U.S. later in life. The *Early Arrival* group arrived in the U.S. between ages 6 and 10 inclusive, and the *Late Arrival* group arrived between ages 12 and 15 inclusive. All had a minimum of

7 years of residence in the U.S. The third group consisted of U.S.-born monolingual native speakers of English. There were 36 participants in the Early Arrival group, 36 in the Late Arrival group, and 72 native controls (though one participant from this latter group was excluded from analysis because he/she gave the same two responses to all stimuli). All were students at the University of California, San Diego participating for course credit. Table 1 gives further information about the two L2 English groups. None spoke a language other than Korean and English.

	Age of Arrival (year)		Length of Residence (years)		Education in U.S. (years)		% English use now		% L1 (Korean) use now	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Early	8.1	(1.64)	11.8	(2.07)	11.6	(1.88)	62.66	(21.45)	36.34	(20.80)
Late	13.1	(1.10)	8.9	(1.97)	8.5	(1.69)	42.81	(16.51)	57.19	(16.60)

Table 1. Language Background and Experience Information of L2 groups.

3.2. Method

Participants were given a 9-point scale acceptability judgment task on paper. Prior to the experiment, participants had a practice session with 3 sentences which received varying judgment ratings by native speakers. Subjects were instructed not to analyze the sentences, but to give their first reaction by rating how good or bad the sentences sounded to them on a scale from 1 (very bad) to 9 (very good). They indicated their response by circling the appropriate number on the scale. Upon completion of the experimental session, participants filled out a language background and experience questionnaire (results summarized in Table 1).

3.3. Materials

This experiment followed a 3 x 2 design, with factors for extraction domain (*that*-complement clause / *wh*-island / CNPC) and question type (yes/no vs. *wh*-question). Sample items for the resulting 6 conditions are given in (6). Each list consisted of 4 tokens for each of the 6 conditions, distributed among 36 lists using a Latin Square design and in pseudo-randomized order. Fillers consisted of the 20 experimental items from Experiment 2, plus 57 additional fillers (e.g. violations of subject-verb agreement, misplaced particles, etc.).

(6) a. Yes/no question with *that*-clause: Did Lisa check *that* you arrested the artist?
b. *Wh*-question with *that*-clause: Who did Lisa check *that* you arrested?
c. Yes/no question with *wh*-island: Did Lisa check *whether* you arrested the artist?
d. *Wh*-question with *wh*-island: Who did Lisa check *whether* you arrested?
e. Yes/no question with *CNPC*: Did Lisa check *the report that* you arrested the artist?
f. *Wh*-question with *CNPC*: Who did Lisa check *the report that* you arrested?

3.4. Results

Results are presented in Figures 1-3. A repeated measures ANOVA was carried out with the between-subjects factor AoA (native, Early, Late) and within-subjects factors extraction domain (*that*-complement clause / *wh*-island / CNPC) and question type (yes/no vs. *wh*-question) with alpha at .05. Significant main effects of extraction domain ($F1 (2, 280) = 87.137$, $p < .0001$, $F2 (2, 210) = 50.272$, $p < .0001$), and of question type ($F1 (1, 140) = 501.455$, $p < .0001$, $F2 (1, 105) = 562.812$, $p < .0001$) were found, as was a significant interaction of these two factors ($F1 (2, 280) = 26.878$, $p < .0001$, $F2 (2, 210) = 26.971$, $p < .0001$). Turning to between-subjects factors, the AoA effect was significant ($F1 (2, 140) = 4.997$, $p = .008$, $F2 (2, 105) = 15.486$, $p < .0001$), as well as its interaction with extraction domain ($F1 (4, 280) = 6.679$, $p < .0001$, $F2 (4, 210) = 4.096$, $p = .003$) and with question type ($F1 (2, 140) = 3.618$, $p = .029$, $F2 (2, 105) = 5.818$, $p = .004$). A three-way interaction of question type, extraction domain, and AoA was also significant ($F1 (4, 280) = 5.113$, $p = .001$, $F2 (4, 210) = 3.121$, $p = .016$).

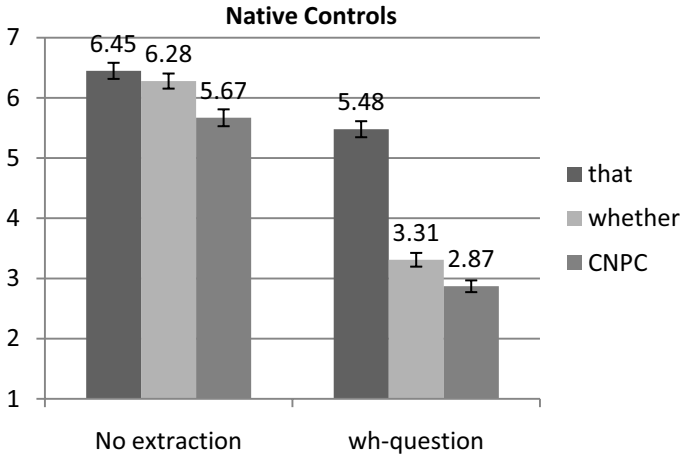

Figure 1. Native controls: Acceptability in three environments with and without extraction

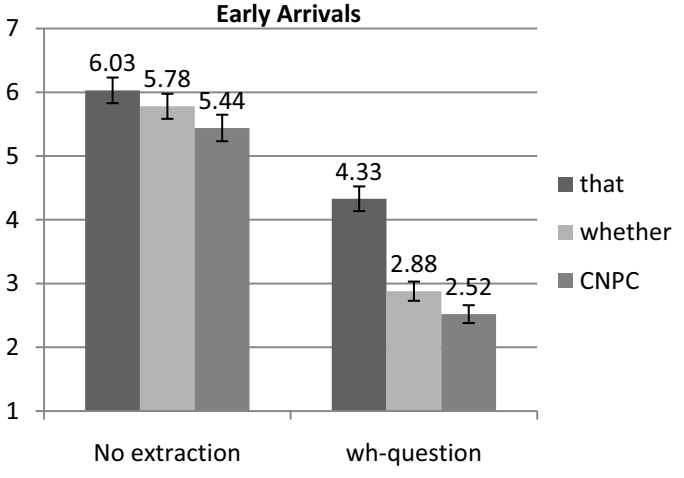

Figure 2. Early Arrivals: Acceptability in three environments with and without extraction

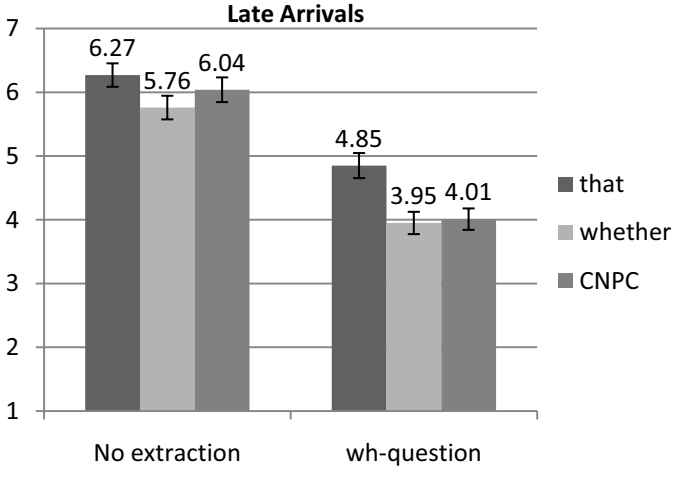

Figure 3. Late Arrivals: Acceptability in three environments with and without extraction

A series of separate repeated measures ANOVAs was run for each group. In both native controls and the Early Arrival group, the ANOVA yielded significant main effects for question type (by subjects only in natives: F1 (1, 70) = 383.895, $p < .0001$; in Early Arrivals: F1 (1, 35) = 182.073, $p < .0001$), F2 (1, 35) = 183.942, $p < .0001$; in Late Arrivals: F1 (1, 35) = 65.246, $p < .0001$, F2 (1, 35) = 180.022, $p < .0001$) and for extraction domain (in natives: F1 (2, 140) = 96.901, $p < .0001$, F2 (2, 70) = 23.757, $p < .0001$; in Early Arrivals: F1 (2, 70) =26.061, $p < .0001$, F2 (2, 70) = 4.873.701, $p = .01$; in Late Arrivals: F2 (2, 70) = 8.479, $p = .001$, F2 (2, 70) = 9.269, $p = .001$). A significant interaction between question type and extraction domain was also shown for the native and Early Arrival groups (in natives: F1 (2, 140) = 38.412, $p < .0001$, F2 (2, 70) = 187.151, $p < .0001$; in Early Arrivals: F1 (2, 70) = 8.519, $p < .0001$, F2 (2, 70) = 14.494, $p < .0001$). Post-hoc paired t-tests also showed that among the *wh*-questions, the *that*-clause was significantly better than the other two extraction domains for all three groups. In addition, the *wh*-question with the *wh*-island was better than the CNPC for the native group alone (t(283)= -43.07, $p < .0001$). These results suggest that all three groups display an island effect for *wh*-islands and the CNPC, though only the natives show a more fine-grained difference in acceptability between the two. A fuller discussion of these results will be given in section 5.

4. Experiment 2: *That*-trace effect

4.1. Participants and method

The participants and method were identical to those in Experiment 1.

4.2. Materials

Materials were created following a 2 x 2 design, with factors for extraction site (subject/object) and presence/absence of *that*. 20 sentences were created per test condition using 20 different lexical items. Sentences were distributed using a Latin Square among 36 lists consisting of 5 tokens for each of the 4 test conditions, for a total of 20 experimental sentences in each list. These sentences were pseudo-randomized. Each list also had 81 fillers (the 24 experimental items from Experiment 1 and 57 additional items). (7) shows sample experimental items for the four conditions.

(7) a. Subject extraction with *that* : Who did Bill think that saw you?
b. Object extraction with *that* : Who did Bill think that you saw?
c. Subject extraction without *that* : Who did Bill think saw you?
d. Object extraction without *that* : Who did Bill think you saw?

4.3. Results

Results for Experiment 2 are presented in Figures 4-6. A repeated-measures ANOVA was run with the between-subjects factor AoA (native, Early, Late) and within-subjects factors extraction site (subject, object) and presence of *that* (+that, that). There was a significant main effect for extraction site (subject/object) (F1 (1, 140) = 84.219, $p < .0001$; F2 (1, 57) = 40.292, $p < .0001$) and for the presence of *that* (F1 (1, 140) = 4.498, $p = .036$). AoA was also found to be significant (F1 (2, 140) = 11.572, $p < .0001$, F2 (2, 57) = 8.766, $p < .0001$). It showed a two-way interaction with presence of *that* (F1 (2, 140) = 18.128, $p < .0001$, F2 (2, 57) = 9.817, $p < .0001$) and a three-way interaction with presence of *that* and extraction site (F1 (2, 140) = 13.073, $p < .0001$, F2 (2, 57) = 5.868, $p = .005$).

Each group was also separately tested by repeated-measures ANOVA with the within-subjects factors extraction site (subject/object) and presence of *that* (+that/-that). In native controls, the ANOVA yielded significant effects of extraction site (F1 (1, 70) = 94.944, $p < .0001$, F2 (1, 19) = 35.083, $p < .0001$) and of presence of *that* (F1 (1, 70) = 19.124, $p < .0001$, F2 (1, 19) = 6.442, $p = .020$). Crucially, a significant interaction of extraction site and presence of *that* was found (F1 (1, 70) = 30.211, $p < .0001$, F2 (1, 19) = 21.255, $p < .0001$), signaling the *that*-trace effect. Planned paired sample t-tests confirmed the *that*-trace effect: the subj/+that condition was significantly less acceptable than either obj/+that (t (354) = -43.25, $p < .0001$) or subj/-that (t (354) = 24.30, $p < .0001$).

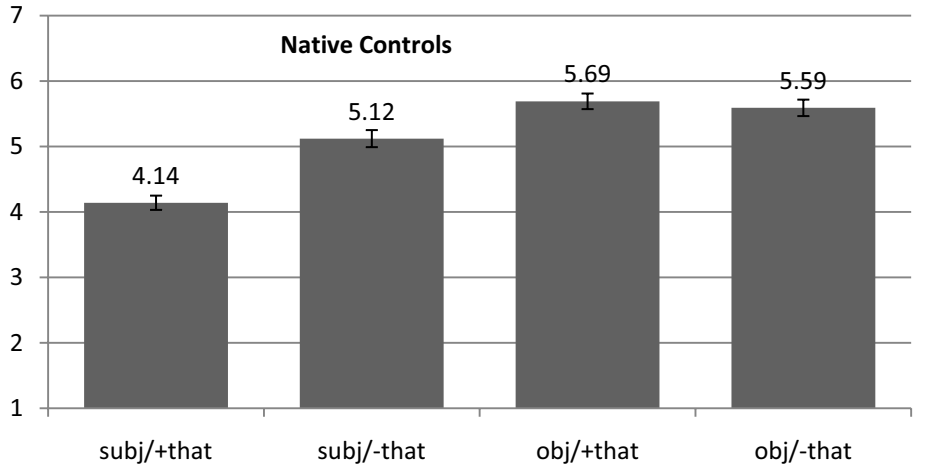

Figure 4. Native controls: Acceptability of subject/object *wh*-questions with/without *that*

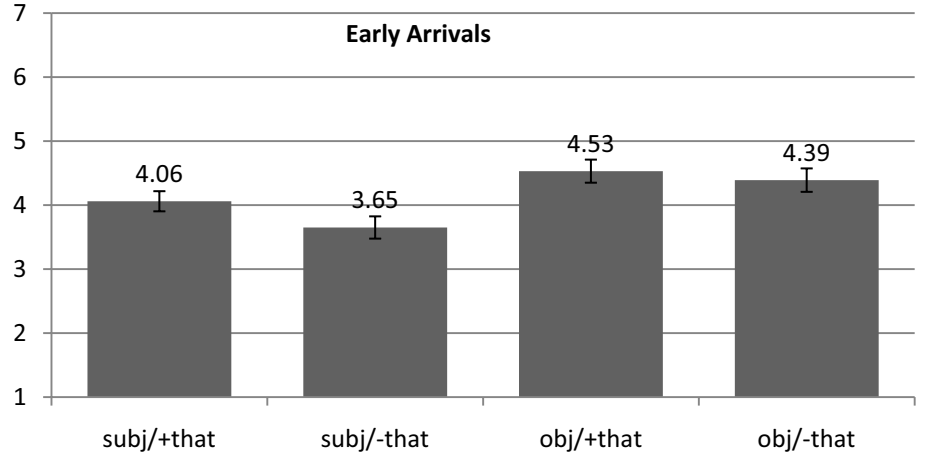

Figure 5. Early Arrivals: Acceptability of subject/object *wh*-questions with/without *that*

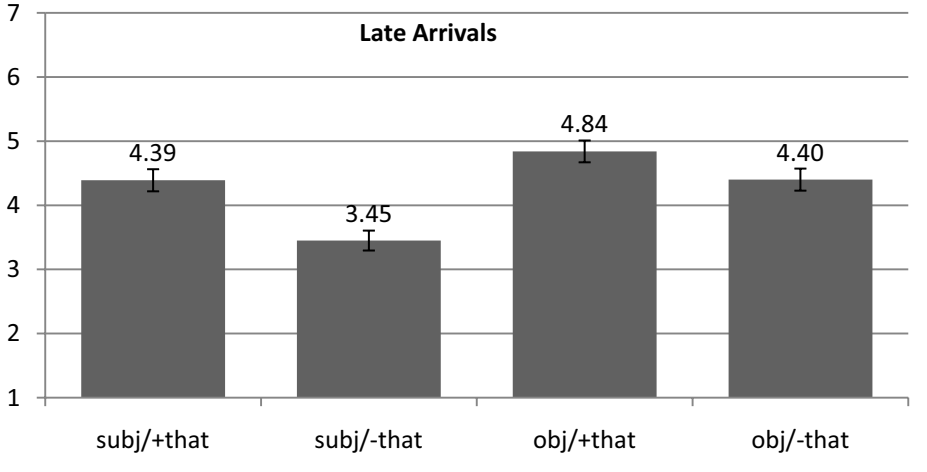

Figure 6. Late Arrivals: Acceptability of subject/object *wh*-questions with/without *that*

In the Early Arrival group, only a main effect of extraction site was found ($F1\ (1, 35) = 12.788$, $p = .001$, $F2\ (1, 19) = 4.652$, $p = .044$), with a preference for object extraction. In the Late Arrival group, a main effect of extraction site ($F1\ (1, 35) = 16.084$, $p < .0001$, $F2\ (1, 19) = 19.846$, $p < .0001$) was found, with an object advantage. In addition, a significant main effect of presence of *that* was found ($F1\ (1, 35) = 15.493$, $p < .0001$, $F2\ (1, 19) = 27.618$, $p < .0001$), with higher acceptability for the conditions with *that* than those without. In neither L2 group was there an interaction between extraction site and presence of *that*, and thus no *that*-trace effect was observed.

5. Discussion

Experiments 1 and 2 largely confirm our expectations with regard to the effects of AoA. In the case of the island constraints examined in Experiment 1, we saw earlier that there was no reason to expect large age-related differences. If these island phenomena are the result of exceeding the processor's capacity to manage filler-gap dependencies in the context of complex syntactic configurations, we would expect all three groups to be affected and to display the consequent drop in acceptability in extraction out of island contexts. This is in fact exactly what we observe in Experiment 1: all three groups show that *wh*-movement out of *wh*-islands and CNPC contexts is significantly worse than out of *that*-clauses. In short, all three groups show island effects, just as our account of islands would lead us to expect.

The case of the *that*-trace effect in Experiment 2 is somewhat different. We saw earlier that under Rizzi and Shlonsky's account, the effect follows from the truncation process, a strategy used by English to circumvent the Subject Criterion. As discussed in section 2, the evidence for this process in the input is sparse, and it is thus easy to imagine that acquiring truncation would be slow and that it could be slower still in older learners who might not as readily attend to the crucial data in the input (i.e. the fact that (4) is possible). We therefore expect to find AoA effects with regard to *that*-trace phenomena, and Experiment 2 shows that this expectation is fulfilled. The native controls show a clear *that*-trace effect (i.e. they find extraction of an embedded subject in the presence of *that* significantly worse than the same without *that* or than extraction of an object with *that*), but neither of the L2 groups do. Moreover, both of the L2 groups show a significant decline in acceptability for embedded subject extraction in general, just as we would expect if they were obeying the Subject Criterion but had not yet fully developed a truncation process.

6. Conclusion

The analyses of island constraints and the *that*-trace effect given in the introduction clearly predict that these two sets of phenomena will be affected differently by increasing the learner's AoA, and taken jointly, Experiments 1 and 2 show that this appears to be true. The island constraints are observed in all learner groups (with an average AoA ranging from 0, in the case of the native controls, to 13.1, in the case of the Late Arrivals), while the *that*-trace effect occurs only in the native controls. The results of these two experiments have interesting implications for the notion of a critical period and for the interrelatedness of studies in second language acquisition, theoretical syntax, and experimental syntax, and it is to these implications that we now turn.

6.1. Critical period

If we take a critical period in second language acquisition to refer to a time of high sensitivity early in life, followed by a sharp decline and then low sensitivity through the rest of the lifespan, we find no evidence in support of this notion in the results here. This is particularly clear in Experiment 1, where all 3 groups exhibit straightforward island constraints, regardless of the differing AoA. Experiment 2 might seem to be more consistent with the notion of a critical period, but upon closer inspection, this does not appear to be true. The Late Arrivals show a significant preference for the +*that* condition in the case of subject extraction, the Early Arrivals show no significant preference, and the native controls show a significant preference in the other direction (i.e. for the -*that* condition). The Early Arrivals thus occupy an intermediate position between the natives and the Late Arrivals,

suggesting that we are not observing a sharp cutoff between those who exhibit the *that*-trace effect and those who do not, but rather a gradual decline in sensitivity to the type of input that matters for the *that*-trace effect, at least for the range of ages examined here. This conclusion is in line with other studies that have argued, on the basis of very different types of evidence, for a gradual decline in L2 sensitivity throughout the lifespan (e.g. Birdsong 2005, Hakuta, Bialystok and Wiley (2003)).

6.2. Second language acquisition, theoretical syntax and experimental syntax

We have showed in this study some specific ways in which AoA does and does not affect ultimate attainment in second language acquisition. Though we have been able to offer an explanation for why some phenomena are susceptible to AoA effects and others are not, we have not attempted to pinpoint the ultimate cause in the case of those that are. That is, we have not attempted to decide whether the AoA effects observed are due to something internal to the individual (e.g. cognitive changes due to aging that affect sensitivity to the input) or something external (e.g. reduced amount of input in the Late Arrival group compared to the Early Arrivals, and in both groups compared to the native controls). Both of these possibilities seem plausible, and future work will have to tease them apart and determine the contribution of each (see, e.g., Birdsong (2006) and Hakuta, Bialystok and Wiley (2003) for discussion).

Beyond its specific empirical findings, the present study offers some strategies that may be useful as research into the effects of age on language acquisition proceeds. Much previous work has focused on relatively large domains within language, such as pronunciation or morphosyntactic accuracy. Nothing that we know about language leads us to expect that these domains should behave in monolithic fashion, however, and indeed, it seems very plausible that within various domains, separate components would be affected differentially by increased AoA. Within syntax in particular, this is especially clear: if syntax is a system of interacting principles and mechanisms, there is no reason to expect them all to behave identically with regard to the effects of AoA. Age-related declines in sensitivity could thus be highly variable across different areas of syntax, and this is exactly what the present study has found, using insights and techniques from both theoretical and experimental syntax. Importantly, the type of results that we have found here informs not only our understanding of second language acquisition, but also our theory of syntax itself. As we have seen, specific syntactic analyses make specific predictions for AoA effects, and exploring these effects empirically can in principle help decide among competing syntactic hypotheses.

References

Birdsong, David. (2005). Interpreting age effects in second language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 109–127). New York: Oxford University Press.

Birdsong, David. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56*, 9–49.

Birdsong, David and Molis, Michelle. (2001). On the Evidence for Maturational Constraints in Second Language Acquisition. *Journal of Memory and Language*, *44*, 235-249.

Cowart, Wayne. (1997). *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks: Sage Publications.

DeKeyser, Robert M., Alfi-Shabtay, Iris, and Ravid, Dorit. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics, 31*(3), 413-438.

Flege, James E. (1999). Age of learning and second-language speech. In David P. Birdsong (Eds.) *New perspectives on the critical period hypothesis for second language acquisition.* Hillsdale, NJ: Erlbaum.

Gathercole, Virginia M. (2002). Monolingual and bilingual acquisition: Learning different treatments of *that*—Trace phenomena in English and Spanish. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 220–254). Clevedon: Multilingual Matters.

Hakuta, Kenji, Bialystok, Ellen, and Wiley, Edward. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science, 14,* 31–38.

Hofmeister, Philip and Sag, Ivan. (2010). Cognitive constraints and island effects. *Language*. 86:366–415.

Johnson, Jacqueline S. and Newport, Elissa L.. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 1, 60-99.

Kluender, Robert. (1998). On the distinction between strong and weak islands: a processing perspective. In: Culicover, P., McNally, L. (Eds.), *Syntax and Semantics*: The Limits of Syntax. Academic Press, New York.

Kluender, Robert. (2004). Are subject islands subject to a processing account? In Vineeta Chand, Ann Kelleher, Angelo J. Rodriguez & Benjamin Schmeiser (Eds.), *23rd West Coast Conference on Formal Linguistics (WCCFL 23)*, 514–527. Somerville, MA: Cascadilla Press.

Myers, James. (2009). Syntactic judgment experiments. *Language and Linguistics Compass, 3*(1), 406–423.

Michael, E. B.and Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 497–515). New York: Oxford University Press.

Rizzi, Luigi and Shlonsky, Ur. (2007). Strategies of Subject Extraction. in H.-M.Gärtner and U. Sauerland (Eds). *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*. 115-16. Berlin: Mouton de Gruyter.

Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference (GASLA 2011)

edited by Julia Herschensohn
and Darren Tanner

Cascadilla Proceedings Project Somerville, MA 2011

ISBN 978-1-57473-445-4 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Kim, Boyoung and Grant Goodall. 2011. Age-Related Effects on Constraints on *Wh*-movement. In *Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference (GASLA 2011)*, ed. Julia Herschensohn and Darren Tanner, 54-62. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2545.