Acquisition of Wh-movement in L2 Learning: A Cross-linguistic Analysis

Syed Shahrier Rahman

1. Assistant Professor, Department of Linguistics, DU
   Email: shahrier120@yahoo.com

Abstract
Cross-linguistically different human languages bear different patterns of wh-constructions. This evokes a common quarry: what happens when L2 learners of wh-construction are exposed to a language that allows distinct type of wh-patterns (wh-movement or wh-in-situ either) from their L1? Will they show the equal competency like other L2 learners of wh-questions who experience the same type in their L1? In this paper, we take English L2 learners in our concern and argue that despite of having a common platform of UG, L2 learners of English with wh-in-situ L1 background generally face more difficulties than other L2 learners who experience wh-movement in their L1. We shall try to figure out a proposal for a gradient way to acquire wh-movement by saying that English L2 learners who have strong wh-movement features in their L1 perform better than L2 learners with wh-in-situ background who bear some kind of movement properties in their L1. In the same way, this weaker group shows better performance than the English L2 learners who do not experience any sort of wh-movement in their L1. We shall use empirical evidence from different secondary sources to test this hypothesis and finally we shall try to sketch out our results.
share the same level of competence as well. In this paper, we take English as a model for second language learning. Eventually we shall try to figure out that to what extent the L2 learners of English with the background of wh-in-situ L1 acquire wh-movement in spite of the access of UG. In this regard, our primary hypothesis is that there is a gradient way to acquire wh-movement in English and L2 learners of English who experience movement frequently in their L1 will perform better than L2 learners with a wh-in-situ background. Moreover, L2 learners with wh-in-situ background who bear some kind of movement property in their L1 should be able to acquire wh-movement in English more competently than other learners with lesser movement or no movement in their L1. We shall use empirical evidence from different secondary sources to test our hypothesis and will try to draw an inference.

The paper is organized as follows. Section 2 will provide necessary base to understand wh-constructions in UG framework along with the guideline of P&P and Minimalism. This section will show us how the language specific parameters and the features and operators are able to influence L2 acquisition. Section 3 illustrates application of the argument to cross-linguistic data. Section 4 summarizes the findings and tries to draw some conclusions.

2. Theoretical Background

2.1 The Construction of Wh-questions

There are two kinds of wh-constructions are commonly available: wh-in-situ and wh-movement.

(1) Wh-in-situ:
Hufei chi-le shenme (ne)? (Chinese)
Hufei eat-ASP what
‘What did Hufei eat?’

(2) Wh-movement:

What, did Tom eat t₁?

In example (1) the wh-phrase shenme ‘what’ remains in the base location while in English example (2), a wh-phrase leaves a trace in its original position and moves towards the clause initially. However, semantic expressions are same in both the examples. In a wh-question, the root clause (C) carries the question force, and this C, or more precisely, the relevant feature on this C, is associated with the wh-phrase to form an information question. One view is that this “association” in a wh-movement language is accomplished by moving the wh-phrase into the specifier of CP headed by the “question” C. On the contrary, the wh-phrase in a wh-in-situ language moves at LF, so the movement is not phonologically detected. However, we can also extend our thinking and can assume that the wh-phrase moves at the same level uniformly in all languages; the difference arises from whether one pronounces the head of the chain (overt movement) or the tail of the chain (wh-in-situ). This problem in fact is related with the movement versus non-movement phenomenon.

2.2 L2 acquisition of constraints on movement

In wh-constructions, the question morpheme Q (Complementizer; no phonetic value) specifies that there is a clause with question. In English, the wh-phrase moves to the specifier position of the CP with the head Q (Adger 2003). Extraction of wh-phrases from embedded clauses is also commonly available to form questions. However, this movement through extraction does not necessarily produce a grammatical result in every instance. We may see the following example:

(3) *[CP,What, did [IP₁ Bob meet [DP a journalist [CP₂ who, e [IP₂ t₁ writes t₂]]]]] (Hawkins 2001)

Now we can raise a question how and to what extent an L2 learners with an L1 background of wh-in-situ understand this type of ungrammaticality. We can assume that if the ungrammatical examples are absent in the input, it would be easier to show the inviolable access to UG in L2 acquisition. Nevertheless, the fact is that the ungrammaticality of the wh-constructions needs to be acquired. The following constraints will help us to understand the ungrammaticality checking procedure.

Subjacency

According to Chomsky, in movement enabled languages, the wh-phrase can leave a trace in its base position and can move to the
nearest unoccupied CP specifier position. In a cyclic manner it can move again by following the same procedure. However, traces must be "subjacent" to their antecedent as they should be separated by no more than one bounding node (like DP and IP in English). On the contrary, if a constituent moves across more than one bounding node in a single movement, it violates the subjacency principle and the result is ungrammatical.

**Subjacency islands**

We can see the following instances to understand how subjacency works in different contexts:

**wh-island**: If a clause is headed by a wh-phrase (e.g. indirect questions), its CP specifier position is filled. A second wh-phrase cannot move because the nearest unoccupied CP specifier position lies two (bounding) IPs away.

(4) *[CP1 What, did [IP1 Freda discover [CP2 who, e [IP2 ti bought ti]]]]? (Hawkins 2001)

**Complex DP**: wh-phrases cannot be extracted from a clausal complement to a DP because both the DP and its IP would lie between the moved wh-phrase and its trace.

(5) *[CP1 What, did [IP1 Freda hear [DP the news [CP2 ti that [IP2 Janice had bought ti]]]]? (Hawkins 2001)

In some cases there is a clear difference between “slightly ungrammatical” and “very ungrammatical” extractions from the same initial configuration across the same bounding nodes. Subjacency does not explain this. Example (wh-island):

(6) a They wondered [CP whether e [IP she could mend the puncture very quickly]]

b ?[CP1 What, did [IP1 they wonder [CP2 whether e [IP2 she could mend ti very quickly]]]]?

c *[CP1 How quickly, did [IP1 they wonder [CP2 whether e [IP2 she could mend the puncture ti]]]]? (Hawkins 2001)

This example will recall later to explain the facts related to SLA.

**Barriers**

In the light of Chomskian approach (1986), barriers are similar to bounding nodes, but unlike in the bounding nodes approach, any maximal projection can be a barrier depending on its syntactic situation. This correct some problems in the bounding nodes approach, but it also needs several exceptions to produce the correct grammaticality judgments for certain sentence types.

**Empty Category principle (ECP) and Relativized Minimality**

According to the ECP hypothesis, traces are classified as one kind of empty category and they need to be licensed and identified. A trace is licensed as well properly head-governed (Rizzi 1990). Moreover, it is a sister to a lexical head. We can assume that a trace is identified by its theta-role assignment as well it is sister to a theta-role assigner, or if an antecedent governs it. In the second case, the distance between antecedent and trace must be minimal, i.e. no category of the same type may intervene.

**2.3 Wh-construction in MP framework**

In Minimalist Program (1995), Chomsky proposes that wh-movement is generated by a strong operator feature of the functional C-head: “the natural assumption is that C may have an operator feature and that this feature is a morphological property of such operators as wh-. For an appropriate C, the operators raise for feature checking to the checking domain of C: [Spec, CP]” (1995: 199) in that way satisfying their scopal properties. If the operator feature on C is strong, movement is overt (e.g. English), and, on the contrary, if the operator feature is weak, wh-movement is postponed until LF (e.g. Chinese).

In Minimalist Inquiry (2000), Chomsky modifies the proposal, providing with LF movement: all movement operations must happen before the point of Spell-Out. Wh-movement in this framework has the following mechanism: “the wh-phrase has an uninterpretable feature [wh-] and an interpretable feature
[Q], which matches the uninterpretable probe [Q] of a complementizer” (2000: 44). The uninterpretable probe [Q] on C seeks the goal, a wh-phrase, and once the probe locates the goal, the uninterpretable features (on both probe, F[Q], and goal, F[wh]) are checked and deleted (Zavitnevich-Beaulac 2002). This feature checking is done by means of Agree, no movement is involved. According to Chomsky, the uninterpretable [wh-] feature of a wh-phrase is “analogous to structural Case for nouns”, consequently it does not have an independent status, but is a reflex of certain properties of Q. Chomsky suggests that Q is realized on a wh-phrase. Being interpretable, Q determines the semantics of a sentence (and of a wh-element) marking it as interrogative; moreover, the operator’s properties are associated with the feature. It is logical to assume that Q should be the trigger of wh-movement. However, in Chomsky’s scheme Q is, in fact, a ‘free-rider’ which lands in an appropriate operator position, [Spec, CP] not for its own need, but due to some properties of the C-head that need to be satisfied (Zavitnevich-Beaulac 2002).

We have already stated that the P&P and the MP believe in the access of UG in SLA. Moreover, in a broad sense, they also propose a common manner for acquiring wh-construction with any kind of L1 background. However, researches on SLA reveal that resetting parameters from the learner’s L1 setting in accordance with P&P or enabling operators and features within MP framework are not equally possible for every group of L2 learners of English. In the next section, in terms of acquiring wh-movement, we shall see that despite of access of UG in SLA, L2 learners with L1 wh-in-situ background have face more difficulties than L1 speakers with wh-movement knowledge. We shall try to investigate this with the help of empirical supports.

3. Acquisition of Wh-movement

Learning English as a second language is a popular demand in all over the world. One of the important parts of this learning process is wh-acquisition. Much research has been conducted to investigate the acquisition of wh-movement by L2 learners of English. Hawkins (2001) presents several studies about L1 speakers of Korean, Chinese and Indonesian learning English as an L2. We may mention here that Korean language has no wh-movement at all where as in Chinese and Indonesian languages bear wh-movement in some cases, but not in questions. In these studies, generally the restrictions related to wh-movement in English prove to be difficult to acquire, accuracy is significantly above 50% in some cases, learners are sensitive to the difference between strong and weak islands, and individuals seem able to reach accuracy similar to that of an L1 speaker. Hawkins argues that UG is accessible to L2 learners, but it may be difficult for them to reset certain parameters from their L1 setting; in this case, the parameter determining whether wh-phrases move at all.

Johnson and Newport (1991) in their study, showed how subjacency restricts wh-extraction in questions. In English, the subjacency principle is used as a guard against various kinds of illicit long-distance wh-extraction. Therefore, if subjacency constrains L2 grammars, and provided that wh-movement has been acquired, as for example Chinese L2 speakers of English should observe restrictions on wh-extraction, even though such restrictions are not demonstrated in the L1. In Johnson and Newport's study subjects were chosen from native speakers of Chinese. It is important to mention here that the subjects first came in touch with the L2 environment at different ages. It can be ranging from age 4 to adulthood and at the time of testing they had lived in the USA. The following Table (source: White; 2003: 247) shows necessary details and the statistics of the test.

| Languages: L1 = Chinese, L2 = English. Task: Grammaticality judgments. Sample stimuli: |
|---|---|
| Declarative (grammatical): | The policeman who found Cathy should get a reward. |
| wh-question (grammatical): | What should the policeman who found Cathy get? |
In the rest of the study, Johnson and Newport also found a continuous decline in accurate rejections of subjacency violations (White 2003). Moreover a correlation between performance and age of arrival in the USA was also significant in this study. According to the table, L2 learners are comparatively weaker in wh-question than the declaratives. Moreover, in the cases of *No inversion and *Subjacency violations, they show much lower performance. Since the study also shows that rejection of subjacency violations are subject to a maturational decline, it provides support for underlying competence of the learners (White 2003).

From our theoretical discussion, we know that according to MP framework, the C-head has only an uninterpretable Q feature. It will be plausible to view the above mentioned experiment in the light of MP framework. According to this syntactic guideline, the uninterpretable probe [Q] on C cannot be an operator, as it is checked and deleted. The interpretable [+Q] feature, which is presumably a question operator, is assigned to a wh-phrase. Since uninterpretable features are checked without triggering movement, in order to account for displacement of a wh-phrase, Chomsky postulates an EPP-feature on a C head. He suggests that the EPP-feature of C is similar to the EPP-feature of T. It requires [Spec, CP] to be filled which results in the displacement of a wh-phrase. (Zavitnevich-Beaulac 2002)

In ‘Beyond Explanatory Adequacy’ (2001), Chomsky explains the nature of the EPP feature, attributing it as OCC (former EPP) which has ‘semantic function’ is available only when “it contributes to an outcome at SEM that is not otherwise expressible”. And further, “we can think of OCC as having the ‘function’ of providing new interpretation”. Thus OCC now is not just an uninterpretable feature of C, but a feature which indirectly contributes to the semantics of a sentence. Wh-in-situ languages posit another problem for the approach. The interpretable Q feature with its operator’s properties is realized on a wh-phrase. The uninterpretable Q of C is checked in Agree configuration. Since no wh-movement is observed in wh-in-situ languages it implies that the C-head does not have the OCC feature. Covert movement is no longer an option in this approach.

Therefore, we can say that despite the C-head in wh-in-situ language like Chinese does not have the OCC feature, Chinese L2 learners of English are able to acquire wh-movement because of the accessibility of UG. However, they face more difficulties than declarative sentential acquisition.

Another study compared native speakers of Korean, Chinese, Indonesian with Dutch speakers and native controls (Schachter 1989, 1990). The Korean, Chinese, Indonesian subjects had long contact with English environment. However, the study found strong L1 influence in those immersed beyond childhood, in that speakers of languages without movement (Korean) or with partial movement (Chinese, Indonesian) were less likely to detect island violations than speakers of languages with similar movement possibilities to English (Dutch). The following Table (source: Hawkins; 2001:280-281) shows this precisely:
Table 2

Mean correct responses of L2 learners to three types of sentence in English

<table>
<thead>
<tr>
<th>18 Dutch speakers</th>
<th>21 Indonesian speakers</th>
<th>20 Chinese speakers</th>
<th>20 Korean speakers</th>
<th>19 native speaker controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 grammatical sentences treated as grammatical</td>
<td>24 ungrammatical sentences treated as ungrammatical</td>
<td>6 grammatical sentences with wh-movement treated as grammatical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native English</td>
<td>21.6</td>
<td>21.2</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td>22.2</td>
<td>21.9</td>
<td>5.8</td>
<td></td>
</tr>
<tr>
<td>Indonesian</td>
<td>21.2</td>
<td>15.2</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>Chinese</td>
<td>19.8</td>
<td>12.4</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td>19.8</td>
<td>12.4</td>
<td>4.6</td>
<td></td>
</tr>
</tbody>
</table>

Result of this study presents: (i) mean score for each informant-set in evaluating as grammatical 24 grammatical declarative sentences. The sentences were consists of complex subjects, complex DPs, relative clauses and embedded questions; (ii) mean score for each informant-set in judging as ungrammatical 24 sentences which have subjacency violations involving complex subjects, complex DPs, relative clauses and embedded questions; (iii) mean score for each group in judging as grammatical 6 grammatical sentences exhibiting movement of wh-phrases from a direct object position (two cases of movement to the nearest CP, two cases of movement to the next higher CP, and two displaying movement to a yet higher CP) (Hawkins 2001). In this study, Schachter found that while there was a high correlation between all subjects' scores on the grammatical declarative sentences and the sentences involving grammatical wh-phrase movement, the Dutch speakers were significantly better to detect ungrammaticality of the sentences involving subjacency violations than all the other L2 speaker groups. Even they were almost indistinguishable from the native speaker controls. However, the Koreans 'performed quite poorly' (Schachter 1990) on subjacency violations. Though the performance of the Indonesian and Chinese speakers were better than the Koreans, their level of accuracy was significantly lower than the Dutch.

More recent studies have assumed that constraints on movement are a function of the kind of constituent involved, the status of the position from which it moves and how far it has to move in a single step - hence that violations of constraints can vary in strength (Hawkins 2001). Given this assumption, one account of the results from L2 studies suggests that if L2 learners show discrepant sensitivity to strong and weak islands, then their grammars are constrained by principles of UG, even if their ability to detect island violations is significantly worse than that of native speakers; differences between L2 speakers and native speakers are the result of other factors, for example difficulty with processing language in real time (Martohardjono 1993). The following Table (source: Hawkins; 2001:298) will show us the result of the study done by Martohardjono:

Table 3

Relative rejection rates of 'strong' and 'weak' constraints on movement in English by speakers of different L1s

<table>
<thead>
<tr>
<th>Language group</th>
<th>Strong violations (%)</th>
<th>Weak violations (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>99</td>
<td>78</td>
</tr>
<tr>
<td>Italian</td>
<td>91</td>
<td>62</td>
</tr>
<tr>
<td>Indonesian</td>
<td>88</td>
<td>46</td>
</tr>
<tr>
<td>Chinese</td>
<td>75</td>
<td>44</td>
</tr>
</tbody>
</table>

Martohardjono constructed a grammatically judgement task and selected sentences with several kinds of movement violation. She put emphasis in this experiment specially on 'strong violations' (extractions from relative clauses and adjunct clauses) and 'weak violations' (extractions from wh-islands and DPs with clausal complements). She applied the task to L1 speakers of Chinese, Indonesian and Italian, as well as a control group of native speakers.
of English. All the non-native speakers were estimated to be 'advanced'. The results (presented in the table 3) show that while overall the mean rate at which non-native informants rejected grammatical sentences differed between the language groups, the relative rejection of strong versus weak violations is consistent across all groups. On the basis of these results, we may argue on accuracy rates that although L2 learners may need more efforts than native speakers, their judgements of wh-structures may still derive from their knowledge of UG principles.

X. Li (1998) develops a similar idea. To test sensitivity to strong and weak islands, Li chose a Chinese group, a graduate group and the native speaker of English. She used a 34-item grammaticality judgment task in which there were 11 'strong' island violations (wh-phrases extracted from relative clauses and sentential subjects), 11 'weak' islands (wh-phrases extracted from wh-islands and DPs with PP complements and 12 control sentences, 6 of which involved grammatical wh-phrase extractions from structurally parallel sentences. (Hawkins 2001). The results show that the Chinese group, who has less exposure to English than the graduate group, performs less well overall than the native speaker controls. However, they discard strong island violations to a greater extent than weak island violations, stating that the principles of UG involved are nevertheless operative. The graduate group appears to be native-like, with the exception that the native controls reject wh-islands more strongly. These results conflicts with the earlier findings of Schachter (1990) and Johnson and Newport (1991). However, we can understand a common impression from every SLA research that L2 learners who do not have wh-movement in their respective L1 face more difficulties in wh-acquisition than others. Along with this, they develop their performance in a gradient manner. An example with different grade of ungrammaticality will help us to comprehend this thing. Therefore, we recall here the example 6 and paste it here again as example 7:

(7) a. They wondered [CP whether e [IP she could mend the puncture very quickly]]

b. *[CP1 How quickly did [IP1 they wonder [CP2 whether e [IP2 she could mend t1 very quickly]]]]

c. *[CP1 What did [IP1 they wonder [CP2 whether e [IP2 she could mend the puncture t1]]]]? (Hawkins 2001)

In example 7(b), the trace t1 is sister to V mend (a lexical category and theta-role assigner) and has a theta-role. It is therefore licensed and identified. The sentence is still slightly ungrammatical because the wh-phrase moves to CP1specifier directly rather than via CP2 specifier (which is already occupied by the wh-phrase creating the wh-island). In example 7(c), t1 can not be theta-governed because its sister (N puncture) is not a theta-role assigner. It must therefore be antecedent-governed and Relativized Minimality (Rizzi 1990) actives here properly. Since a wh-phrase intervenes between antecedent and trace, the trace is not antecedent-governed either, thus not identified at all. This is a more serious kind of ungrammaticality. The L2 learners of English with wh-in-situ L1 experience this relative acquiring process more frequently.

4. Conclusion

In this paper, previously conducted different studies on wh-acquisition have been presented to show that L1 speakers with wh-in-situ background experience difficulties while they undergo L2 acquisition of English. In the light of MP, we may assume that some uninterpretable features are not available in SLA. The hypothesis within the MP framework requires that the uninterpretable features should be deleted before LF. They do not have any semantic imports or any cues for L2 learners. That is, there exists no visible evidence for learners to know how the grammar computes in terms of the features. The present study put some light on not only the acquisition of wh-movement, but also the difference in experiencing difficulties by L2 learners with wh-in-situ background.

References


