Revisiting Wh-movement: The Availability of an Uninterpretable [wh] Feature in Interlanguage Grammars

Roger Hawkins
University of Essex

1. The nature of the language faculty

A recent conception of the language faculty is that it provides an inventory of linguistic features (person, number, Case, finiteness, tense, telicity, etc) and a set of computational devices for assembling the features into lexical items, combining them into expressions, interpreting those expressions semantically and assigning them phonological form. The computations are devices like Merge, Agree, Move, Match (e.g. matching the features of phonological forms to the features of syntactic terminal nodes). The computations operate under restrictive principles like `Movement must be the shortest possible', `Agree involves a relation between an interpretable feature and a cognate uninterpretable feature', `Match involving two items A and B requires that the features of one are exhaustively identical to the features of the other' (Chomsky 1998, 2001; Harley & Noyer 1999; Adger 2003).

By hypothesis, the computations and the principles constraining them are cross-linguistically invariant. They are operative in all mature native grammars, are available for the lifetime of the organism, hence are unaffected by any possible critical period that might affect language acquisition (Hale 1996). The inventory of linguistic features contains more features than any specific language makes use of in the assembly of lexical items. Languages differ in the selections they make from this inventory, and this is a major aspect of cross-linguistic variation.

2. Computational devices and features in SLA

Given this conception of the language faculty, there is every reason to think that the computations and the principles constraining them are necessary components of interlanguage grammars (ILGs). In fact, considerable work in SLA in recent years has provided evidence to support this assumption. Work on `poverty of the stimulus' cases such as scope interactions (Dekydtspotter, Sprouse and Anderson 1997, Dekydtspotter and Sprouse 2001), the Overt Pronoun Constraint (Kanno 1997, Pérez-Leroux & Glass 1999), the Empty Category Principle (Kanno 1996), representations triggered neither by the L1 nor L2 input (Schwartz & Sprouse 2000, among others), suitably re-interpreted within the conception of grammar described above, suggest that the computations of ILGs are just those of the language faculty.

There has been considerable debate, however, about the availability of the inventory of features for the construction of ILGs in (adult) SLA. Where a speaker's primary language (or languages) has not selected a feature for the assembly of lexical items, there are two possibilities for what might happen in later second language acquisition: (a) the feature in question is still available for selection, and just needs input to trigger its selection; (b) the feature is no longer available; there is a critical period for availability after which unused features of a certain type are cleared from the cognitive architecture. This possibility in syntax is analogous to the claim in phonology that features that are unused for phonological contrasts in primary language acquisition disappear beyond a critical period (Brown 2000, Larson-Hall 2004).

3. Features present `despite appearances' vs features missing `despite appearances'

The debate in syntax has arisen because of considerable observed variability in the performance of L2 speakers with mature ILGs by comparison with native speakers of the target language, and in some cases divergence from native speaker grammars (Sorace 1993, 2003; White 2003, chapter 6). Is such variability/divergence the result of some problem unconnected with the selection of features (such as a problem in mapping between syntactic output and phonological form), or does it occur because certain features are no longer available for selection, having disappeared at the end of a critical period, with variability/divergence the effect of the ILG coping without this feature?

A number of influential studies have argued for the former view. For example, Haznedar & Schwartz 1997, Lardiere 2000, Prévost & White 2000 suggest that syntactic representations in mature ILGs can be complete, operating not only with the computations of the language faculty, but having full access to features even when they are not instantiated in primary language acquisition. Variability of the kind displayed by the L2 English speaker studied by Lardiere (Patty) in producing past tense and 3rd person present tense S-V agreement in speech (illustrated in table 1) is not a reflection of the syntactic representation. The syntactic representation is target-like, as illustrated by performance on Case marking of pronouns and absence of verb raising. The problem has to do with mapping between levels of representation, and is exacerbated in speech where communication pressure is involved.

Table 1: Production of appropriate forms by Patty in conversation at three sampling points (after 10 years of immersion (time 1) and 18+ years of immersion (times 2 and 3))

	Past (reg + irreg)	Pron Case	3p -s	No V raising
Time 1	24/69 (35%)	49/49 (100%)	2/42 (5%)	69/69 (100%)
Time 2	191/548 (35%)	378/378 (100%)	0/4 (0%)	122/123 (99%)
Time 3	46/136 (34%)	76/76 (100%)	1/22 (5%)	42/42 (100%)

On this view (which will not be discussed in detail here) it might be said that the **full resources of the language faculty are available** for the construction of ILGs **despite appearances** to the contrary in performance.

The second view - that variability/divergence in L2 performance is a consequence of the unavailability of certain features as the result of a critical period – argues the opposite: that syntactic representations in mature ILGs may be incomplete, and that while the computations of the language faculty are operating, some features may be missing (Smith & Tsimpli 1995, Hawkins & Chan 1997, Tsimpli 2003). This view faces the problem that in performance L2 speakers can appear to have acquired properties that should be problematic because of the unavailability of the relevant features. For example, although Patty fails to produce past tense in a target-like way, she does produce it in one-third of required cases; and Lardiere reports that she hardly misuses past tense. In this case the claim has to be that **features are missing from ILGs despite appearances** to the contrary in performance. This claim will be discussed in the present paper in the context of the apparent L2 acquisition of obligatory wh-phrase movement to CP by speakers of L1s without obligatory movement (Japanese/Chinese/Korean).

4. Hawkins & Chan: a `missing feature despite appearances' case revisited

First, consider a claimed `missing-feature-despite-appearances' case examined in Hawkins & Chan's (1997) study of the acquisition of English restrictive relative clauses (RRCs) by native speakers of Cantonese. Hawkins & Chan considered a number of the characteristic properties of English RRCs like the one illustrated in (1):

1   The patient who I visited was very sick

Such constructions are assumed to involve the movement of an overt wh-pronoun from a first-merged position to the specifier of CP, as in (2), where the first-merged position is indicated by < >:

2   The patient [$_{CP}$ who [$_{IP}$ I visited <who>]] was very sick

The reason why movement is assumed to be involved is that violations of `shortest move' occur when the relative pronoun *who* moves too far (traditionally referred to as `subjacency violations'):

3a.   *This is the patient [who [Mary told me [when [she will visit <who>]]]]
  b.   *This is the patient [who [Mary heard [the news [that the doctor had sent <who> to hospital]]]]

In (3a) the dependency between *who* and its copy is interrupted by another intervening wh-word *when* (a wh-island violation), while in (3b) *who* has crossed an adjunct boundary, assuming a complement to a noun to be an adjunct (a complex NP violation). Secondly, resumptive pronouns are ungrammatical in the normal case, though in languages which lack relative operator movement they are often possible:

4   The patient who I visited *him was very sick

In addition, but not directly connected with wh-movement, there is a constraint against a doubly-filled C in English, ruling out examples like (5):

5   The patient *who that I visited was very sick

    Using a grammaticality judgement task (59 relevant items out of a total of 101 sentences) involving 20 grammatical and 39 ungrammatical RRCs (wh-island violations, complex NP violations, resumptive pronouns and doubly-filled Cs), Hawkins and Chan tested L1 speakers of Cantonese at 3 proficiency levels in English, and compared them with native speaker controls. Results are presented in tables 2 and 3.

Table 2: Accuracy in judging (i) grammatical RRCs; (ii) ungrammatical doubly-filled C; (iii) ungrammatical resumptive pronouns (%)

	Gramm RRCs	*Doubly-filled C	*Resumptive Prons
Elementary Chinese (n=47)	56	50	38
Intermed. Chinese (n=46)	67	68	55
Advanced Chinese (n=54)	79	83	90
NS controls (n=32)	96	99	98

Table 3: Accuracy in rejecting *wh-island violations and *complex NP violations

	*wh-island	*complex NP
Elementary Chinese (n=47)	63	71
Intermed. Chinese (n=46)	54	61
Advanced Chinese (n=54)	41	38
NS controls (n=32)	98	85

On the basis of the results in table 2, Cantonese speakers of English appear, as proficiency increases, to approximate to the target language. The advanced speakers accept grammatical RRCs, and reject doubly-filled C and resumptive pronouns. This might lead one to the conclusion that they have established a movement operation for RRCs in their ILGs. However, the results in table 3 are problematic for this view. Cantonese speakers appear to become less target-like with proficiency in detecting the ungrammaticality of violations of `shortest move'. The elementary speakers apparently know about constraints on wh-movement while advanced proficiency speakers apparently allow a high proportion of such violations. Hawkins & Chan explain the results as follows. The low proficiency speakers reject sentences violating `shortest move' not because they involve shortest move, but because

they lack resumptive pronouns. These are speakers who do not typically reject resumptive pronouns in English sentences (62% of responses accept them) and judge grammatical relative clauses around chance level (56%). Since RRCs in Cantonese do not have wh-phrase movement, and allow resumptive pronouns, there is the strong possibility that the performance of these speakers in English derives from the transfer of L1 properties into their ILGs.

The advanced proficiency speakers, by contrast, recognise the impossibility of overt resumptive pronouns (90% of responses reject them), treat grammatical relative clauses largely as grammatical (79% acceptance rate), and additionally know about the restriction on doubly-filled C (83% rejection rate). However, despite the appearance of target-like performance on these properties, the subjects are not moving wh-phrases to the specifier of the relative CP. If they were, they would be sensitive to violations of shortest move. Hawkins & Chan claim that they have arrived at an ILG where a wh-phrase first-merges with relative C and binds an obligatory null resumptive pronoun in the embedded clause. (Obligatory null resumptive pronouns are not unusual. They are found in subject relative clauses in Cantonese and other languages lacking movement in relative clauses).

Why do Cantonese speakers establish such a grammar, when input leads native speakers of English to a grammar involving wh-movement? Following Rizzi (1990), Hawkins & Chan assume that English relative C has the features [+predicative] and [±wh]. The presence of the [wh] feature requires the presence of a wh-word/operator in the embedded clause and drives movement to the specifier of CP. By contrast, Hawkins & Chan claim that Cantonese lacks the [wh] feature of C. Following Smith & Tsimpli (1995), they then claim that [wh] is a type of feature that is subject to a critical period. The relevant `type' is one associated with a functional category that `determine[s] parametric differences between languages' (1997: 216). This is extended to the more general claim that features of functional categories that determine parametric options are subject to a critical period.

5. Three observations about Hawkins & Chan's claim

There are three things to note about the claim:
(a) the description `features that determine parametric differences between languages' is vague;
(b) the proposal allows for the availability of all the resources of the language faculty other than the features that have disappeared;
(c) there is no claim that post-critical-period L2 speakers cannot select new features from the UG inventory or establish new categories. In particular, if functional categories necessarily involve at least one interpretable feature, and interpretable features are not those `determining parametric differences', L2 speakers should encounter no difficulty acquiring these new features/categories.

6. Refining what counts as a feature determining parametric differences - Adger's (2003) account of wh-movement in interrogatives

It is important, then, to define more carefully what counts as a feature that `determine[s] parametric differences'. Recent developments in linguistic theory from the perspective of the minimalist program allow a finer definition to be given. Consider the account of wh-movement in English interrogatives offered by Adger (2003). Interrogatives involve an agreement dependency between an interrogative complementizer with the interpretable feature [Q] and a constituent in the clause with the interpretable feature [wh], e.g.

6	C[Q]	...	D[wh, human]	(= who)	
			D[wh, non-human]	(= what)	
			P[wh, location]	(= where)	etc

The agreement dependency is established through an uninterpretable [$u$wh] feature which is part of C: C[Q, $u$wh:]. This invokes the operation Agree, whose function is to value and delete the uninterpretable feature. Agree takes the following form (Adger 2003: 169):

7   **Agree**
In a configuration:

X[F: value] ... Y[*u*F:] or X[*u*F:] ... Y[F: value]
where ... represents c-command and F = a feature

F values *u*F

In the case of a wh-question, the operation of Agree is illustrated in (8):

8   C[Q, *u*wh:] ... D[wh] → C[Q, ~~*u*wh~~: wh] ... D[wh]

By hypothesis, wh-questions take the form in (8) in all languages. The difference between English and wh-in-situ languages is that [*u*wh] is strong in English, but weak in wh-in-situ languages. Adger represents strength as an asterisked feature:

8'   English: C[Q, *u*wh*:] ... D[wh]

Strength is interpreted as a requirement for the local operation of Agree. This forces the wh-phrase to raise to the specifier of CP. In wh-in-situ languages there is no requirement for valuing to occur locally.

7. Extending Adger's account to restrictive relative clauses

Consider an extension of this account to restrictive relative clauses. The complementizer involved is predicative rather than interrogative. Minimally, then, C has the interpretable feature [pred]. In English, RRCs involve wh-operator movement, as discussed above. This implies that C[pred] in English also has a strong uninterpretable [*u*wh*:] feature which agrees with an interpretable [wh] feature on a constituent in the relative clause, and requires that constituent to move to the specifier of CP:

9   C[pred, *u*wh*:] ... D[wh] → [$_{CP}$ D[wh] [$_C$ pred, ~~*u*wh*~~: wh]] ... <D[wh]>

In Cantonese, there is no D[wh] in relative clauses, unlike interrogatives. There is no agreement between C[pred] and a position in the relative clause. The relationship is rather one of binding between the head and a pronoun, which may be null, in the relative clause:

10   [[Ngo jungyi pro$_i$] ge] ni go neuijai$_i$
          I   like (her) C this CL girl   (CL = classifier)
         `This girl who I like'

The learnability problem for the Cantonese speaker of English is to establish that English has a strong [*u*wh*:] feature associated both with C[Q] and C[pred], whereas in Cantonese a weak [*u*wh:] feature is associated with C[Q], and no [*u*wh:] at all is associated with C[pred].

8. The `failed functional features hypothesis' updated: uninterpretable features are subject to a critical period

Hawkins & Chan's claim that `feature[s] which determine parametric differences are subject to a critical period' can now be reformulated as:

11   Uninterpretable features not selected from the UG inventory of features during the critical period disappear (cf Tsimpli 2003).

Observe that this claim is not reducible to the claim of the Local Impairment Hypothesis (Beck 1998). Local Impairment proposes that the strength property alone becomes impaired in post-critical-period L2 acquisition. The claim in (11), by contrast, predicts that there will be cases where languages differ in their selection of [*u*F] independently of strength. Adger's account of past tense in English is that it involves an agreement relation between an interpretable tense feature on T and an uninterpretable [*u*tense:] feature on v:

12   T[tense: past] ... v[*u*tense:] → T[tense: past] ... v[~~*u*tense~~: past]

Cantonese and Mandarin do not have a [*u*tense:] feature on v. If [*u*F]s are subject to a critical period, by (11) it is predicted that speakers of Cantonese/Mandarin who are post-critical-period learners of English will not have access to [*u*tense:]. Strength is not involved here. Furthermore, local impairment claims that the strength of features becomes inaccessible whether the relevant feature is strong or weak in the L1. However, (11) proposes that where a feature has been selected during the critical period, it will be available later on. If [*u*wh*:] has been selected, it will be available in the development of ILGs. This can be illustrated from the results in the Hawkins & Chan study of a comparator group of French speakers at the same levels of proficiency in English. French, like English, has a strong [*u*wh*:] feature on C[pred].

Table 4: Accuracy in judging (i) grammatical RRCs; (ii) ungrammatical doubly-filled C; (iii) ungrammatical resumptive pronouns (%) (French informants)

	Gramm RRCs	*Doubly-filled C	*Resumptive Prons
Elementary French (n=47)	81	91	81
Intermed. French (n=46)	88	95	90
Advanced French (n=54)	92	98	96
NS controls (n=32)	96	99	98

Table 5: Accuracy in rejecting *wh-island violations and *complex NP violations (French informants)

	*wh-island	*complex NP
Elementary French (n=47)	59	72
Intermed. French (n=46)	66	79
Advanced French (n=54)	85	90
NS controls (n=32)	98	85

9. Acquisition of wh-movement in English interrogatives by Japanese speakers

One consequence of the claim in (11) is that native speakers of languages with wh-in-situ interrogatives who learn a language with wh-movement interrogatives beyond the critical period will not have access to the strength part of the [*u*wh*:] feature in constructing their ILGs. There is plenty of apparent counterevidence to this claim in the literature (Martohardjono 1993, Li 1998, White & Juffs 1998, Ohba in progress), The remainder of this paper will look at apparent counterevidence more closely, and will suggest that there are still grounds for maintaining the claim in (11). A lot of the counterevidence may just be apparent, derived from the generative power of the computations and accessible features of the language faculty.

9.1 Relevant properties of Japanese interrogatives

Consider the acquisition of wh-movement in English interrogatives by native speakers of Japanese, a wh-in-situ language. In Japanese, vP, TP and CP are all head-final. Furthermore, in interrogatives, C[Q] is mapped onto a free-form phonological exponent:

13   [$_{CP}$ [$_{TP}$ John-ga [$_{vP}$ kinou hon-o   kaimashita $_{vP}$]   -ta $_{TP}$] ka $_{CP}$]?
          John-Nom   yesterday book-Acc buy   past   Q?
          `Did John buy a book yesterday?'

In wh-questions, wh-phrases may remain in situ, as in (14):

14   John-ga   kinou   nani-o   kaimashita ka?
          John-Nom yesterday what-Acc   buy+past Q?
          `What did John buy yesterday?'

But Japanese also allows optional leftward scrambling, as in (15):

15   Nani-o John-ga kinou <nani-o> kaimashita ka?

Crucially, scrambling does not involve movement driven by a strong [$u$wh*:] feature. If such a feature were involved, in multiple wh-questions scrambling should give rise to superiority/subjacency effects, as English does. A comparison between English and Japanese in (16) and (17) shows that such effects do not arise in parallel cases in Japanese:

16a.   Did John say [who ate what]?
   b.   John-wa [dare-ga   nani-o   tabeta ka] itta no?
          John-Top who-Nom what-Acc   ate C   said Q?

17a.   *What did John say [who ate <what>]?
   b.   Nani-o   John-wa   [dare-ga   <nani-o> tabeta ka] itta no?
          What-Acc John-Top who-Nom   <what-Acc> ate C say Q?

Although *nani-o* has crossed *dare-ga* in (17b), there is no violation of superiority/subjacency. This would follow if scrambling is the effect of a different kind of movement from that required by [$u$wh*:]. (See the discussion section for consideration of what kind of movement might be involved.)

9.2 Apparent evidence that Japanese speakers of English have acquired wh-movement

There is quite a lot of apparent evidence in the existing literature to suggest that when Japanese speakers acquire English they acquire the 'strength' part of [$u$wh*:] that drives wh-movement in interrogatives. Yusa (1999) tested 25 intermediate-proficiency Japanese speakers of English (JSE) on a range of phenomena, and found impressively target-like performance on a number of properties:

(a) In a question formation task, 25/25 of the subjects produced target-like wh-fronting and subject-auxiliary inversion.
(b) 20/25 were sensitive to the impossibility of contracting *is* to *'s* in cases like: *Do you know where John is/*'s ___ now?* On the assumption that *is*-contraction is blocked when a following wh-phrase has been moved, this suggests that 20/25 speakers have wh-movement.
(c) 20/25 were sensitive to a strong crossover violation in sentences like: *I know who$_i$ he$_{*i/j}$ said has the best smile*. On the assumption that the impossible binding case results from the pronoun binding a wh-variable, 20/25 speakers have acquired wh-movement.

Ohba (in progress), using a grammaticality judgement test involving scalar judgements from -2 (impossible) to +2 (possible sentence of English) found that 8 advanced proficiency JSE distinguished grammatical long-distance wh-questions (*Who does the woman know that Janet loved <who>?*) from ungrammatical wh-questions with weak and strong violations of subjacency (weak violation: **What did Rebecca wonder who would believe <what>?*; strong violation: **What did your parents visit a restaurant which served <what>?*) Mean ratings of grammaticality by the two groups are given in table 6:

Table 6: Mean ratings of grammaticality on a scale from -2 (impossible) to +2 (possible sentence of English)

	Grammatical	Weak		Strong	
		*wh-island	*CNPC	*RRC	*Adjunct
JSE=8	1.3	-0.8	-1.2	-1.4	-1.1
NS=15	1.7	-1.9	-1.8	-2.0	-1.9

The evidence presented above appears to point strongly to the possibility that the JSE have established the [*u*wh*:] feature of English C[Q]. If they have, we would expect them to be sensitive not only to subjacency violations but also to another corollary of movement forced by the [*u*wh*:] feature: superiority effects. Superiority effects arise in the configuration in (18):

18   C[Q, *u*wh*:] ... [wh]1 ... [wh]2

[wh]2 cannot move to the CP crossing [wh]1 because this would be a violation of shortest move. Hence, cases like the following are ungrammatical in English:

19a.   *What did who buy <what>?
   b.   *When did who leave <when>?
   c.   *What did John buy <what> where?

(In (19c), *where* (a modifier of vP) c-commands *what*, giving rise to a superiority violation, a fact confirmed by the responses of the native speakers in the experiment reported below).
        In Japanese there are no superiority effects in cases parallel to those in (19) (as illustrated in (17)) because there is no [*u*wh*:] feature attracting a constituent with an interpretable [wh] feature.

9.3 Superiority/subjacency in multiple wh-questions: a test case

        Hawkins & Hattori (in progress) investigated the sensitivity of high proficiency JSE to superiority in biclausal sentences like the following:

20a.   When did Sophie's brother warn [Sophie would phone who <when>]?
   b.   *Who did Sophie's brother warn [Sophie would phone <who> when]?

There were two reasons for choosing biclausal sentences rather than single clause sentences. Firstly, to avoid the possibility that JSE reject violations like *What did who buy?* because they have a conscious strategy excluding *do*-support with subject wh-words (apparently taught as such in some English classrooms in Japan). Secondly, to allow for two possible scope readings of the wh-word in the matrix clause: one in the matrix clause and the other in the embedded clause, as illustrated in (21):

21a.   When did Sophie's brother warn <when> [Sophie would phone who <when>]?
   b.   Who did Sophie's brother warn <who> [Sophie would phone *<who> when]?

This allowed us to exploit a tension between the pragmatic plausibility of an embedded scope reading and its syntactic impossibility because of superiority or subjacency. We used this tension in the construction of a version of a truth value judgement task (Crain & Thornton 1998) which took the form: story + question + 3 possible answers. All answers in the experimental items are pragmatically `true' given the story, but syntactic constraints restrict whether they are grammatically possible answers or not. To illustrate:

Example test item

Sophie was angry. Her holiday had been ruined because the hotel she had booked through a travel agency was full, and she had to sleep in a tent. Sophie's brother was a friend of Norman who owned the travel agency. He spoke to Norman on Thursday and told him that Sophie would be phoning his manager, Mrs Smith, the following day to ask for her money back.

Question: Who did Sophie's brother warn Sophie would phone when?

Answer 1: He warned Norman that Sophie would phone on Friday.
Answer 2: He warned that Sophie would phone Mrs Smith on Friday.
Answer 3: He warned Norman on Thursday that Sophie would phone.

Answer 1 corresponds to *who* having scope in the matrix clause (*He warned who?*), and *when* in the embedded clause (*Sophie would phone when?*), and is both pragmatically plausible, given the story, and grammatically possible. Answer 2 corresponds to both *who* and *when* having scope in the embedded clause (*Sophie would phone who when?*). This is pragmatically plausible given the story, but violates superiority (*when* being superior to object *who*). Answer 3 corresponds to *when* having scope in the matrix clause (syntactically impossible), but again the sentence is pragmatically plausible, given the story.

The test took the following form:

(a) The wh-words used in the questions were: *who, what, where, when*.
(b) All embedded clauses in the questions were finite.
(c) 8 test items served as a `syntax test'; these involved questions with a single wh-word in the matrix clause of a biclausal structure. In 4 of these, the scope of the wh-word was unambiguously the embedded clause:
      Who did Tom discover [Mary had met <who>]?

In the other 4, scope was ambiguously the matrix or embedded clause:
      Who did the specialist promise <who> [Mary would see <who>]?
(A context in which this question is plausible is where Mary has an eye problem (e.g. cataracts) and where a response to the scope of *who* in the matrix clause might be: `The specialist promised her husband that she would see (again)', and a response to the scope of *who* in the embedded clause might be: `The specialist promised that Mary would see her husband (again)').

The purpose of the syntax test was to ensure that informants selected for analysis on the experimental items could appropriately interpret long-distance wh-moved structures. Informants were selected for analysis on the experimental items if they chose appropriate answers and rejected inappropriate answers on at least 6 out of 8 items of the syntax test.

There were 14 main experimental items involving multiple wh-questions, distributed as follows:

(a) The matrix wh-word was interpretable only in the embedded clause, with no violation of superiority/subjacency. Example:
      Who did the headteacher suspect [<who> had taken what]? (k=3)

(b) The matrix wh-word was interpretable either in the main or the embedded clause, with no violation of superiority/subjacency. Example:
      When did Henry remember <when> [Louise had lost what <when>]? (k=4)

(c) The matrix wh-phrase was pragmatically plausible either in the main clause or the embedded clause, but the embedded clause interpretation was blocked by a superiority violation. Example:
      Who did Sophie's brother warn <who> [Sophie would telephone *<who> when]? (k=3)

d) The matrix wh-phrase was pragmatically plausible either in the main or embedded clause, but the embedded clause interpretation was blocked by an intervening wh-word in the embedded clause, but without a superiority violation (`subjacency'). Example:

When did Rupert discover <when> [who Nora had met <who> *<when>]? (k=3)

(e) The matrix wh-phrase was pragmatically plausible either in the main clause or the embedded clause, but the embedded clause interpretation was blocked by a superiority violation and a subjacency violation. Example:

Who did the weather office warn <who> [when the hurricane might strike *<who> <when>] (k=1)

(There were more intended tokens of this type in the test instrument, but due to an error in design which emerged after administration of the test, only 1 token was countable).

The 8 `syntax test' items and 14 experimental items were randomised with 16 other items to produce a 38-item instrument.

Subjects
An initial cohort of 40 Japanese speakers of English (JSE) and a control group of 19 native speakers of English (NSE) took the test. Those whose performance was analysed were then selected as follows:

(i) Any JSE who had lived in an English-speaking country before puberty was excluded.
(ii) Any speaker (JSE or NSE) who failed the syntax test was excluded.
(iii) Any speaker (JSE or NSE) who chose all 3 answers in more than 5 of the 29 items where only 2 answers were grammatically possible was excluded (to exclude those with a tendency to response `yes').

This left 19 experimental JSE (age range: 22-43, length of residence in an English-speaking country: 9 months-18 years), and a control group of 11 NSE.

The task was presented to informants as a paper and pencil test. Each item (that is, story + question + 3 answers) was on a separate sheet (with no possible return to earlier answers). For the JSE, the story component was in Japanese (following a design used by Dekydtspotter, Sprouse and Anderson 1997, Dekydtspotter and Sprouse 2001) to reduce responses which might be due to a misunderstanding of the story, but the question and the answers were in English. The test was untimed.

Each answer chosen by an informant was given a score of 1 and each unchosen answer a score of 0. Group means were then calculated for responses to each of the three answers. The non-parametric Mann-Whitney U test was used to test for significant differences between and within the two groups of informants. The alpha level was set at p<.05. A non-parametric test (rather than a parametric test like ANOVA) was chosen because the native group were categorical in their responses to two of the test conditions, which meant that their results on these items were not normally distributed. (In fact a one-way ANOVA was run just for the purposes of comparison with the Mann-Whitney U test, producing similar values).

9.4 Results

Results for answers corresponding to the matrix wh-word having matrix or embedded scope are presented in table 7. (The number of choices of the third response, which typically corresponded to the embedded wh-word having scope in the matrix clause – always ungrammatical in English – were almost never chosen, so results for this are not reported.)

Table 7: Multiple wh-questions: mean choice of answers corresponding to the scope of a matrix wh-word (significant differences indicated by arrows)

		Embedded scope		Matrix scope	
		JSE	NSE	JSE	NSE
Qu type					
(a)	k=3	0.96	0.97		
	SD	0.105	0.100		
(b)	k=4	0.78	0.75	0.92	0.91
	SD	0.262	0.194	0.119	0.169
(c)	k=3	0.75 ↔	0.33	0.88	0.85
	SD	0.291	0.298	0.165	0.273
(d)	k=3	0.58 ↔	0.21	0.93	1.00
	SD	0.366	0.308	0.140	0.00
(e)	k=1	0.58 ↔	0.00	0.95	0.91
	SD	0.507	0.00	0.229	0.302

Key: (a) = embedded scope only, no syntactic violation
      (b) = matrix and embedded scope, no syntactic violation
      (c) = matrix and *embedded scope, violation of superiority
      (d) = matrix and *embedded scope, violation of subjacency
      (e) = matrix and *embedded scope, violation of superiority and subjacency

Rows (a) and (b) in table 7 report the frequency with which subjects chose answers where the matrix wh-word in the question has scope in the embedded clause and there is no superiority/subjacency violation. For the (a) answers there was no matrix scope possibility; for the (b) answers scope was either in the matrix or the embedded clause. JSE and NSE responded very similarly in these cases, with no significant difference between them. This is an important baseline. It shows that the experimental subjects are able to interpret long-distance *wh-word ... gap* dependencies and recognise scope ambiguities.

Row (c) reports answers where the matrix wh-word in the question can be interpreted either in the matrix or embedded clause in terms of pragmatic plausibility, but where a superiority violation should block the embedded reading. Such a reading is blocked for the NSE who are significantly different in their responses here to embedded scope than in the (b) case. The JSE, however, are not significantly different in their responses to (b) and (c). Furthermore, they are significantly different from the NSE. The JSE are not apparently sensitive to superiority effects.

Row (d) reports answers where again the matrix wh-word in the question can be interpreted both in the matrix and embedded clause in terms of pragmatic plausibility, but where a subjacency violation should block the embedded reading. This is the case for the NSE. They are significantly different in their responses to embedded scope compared with their answers in the (b)-type cases. Their responses to subjacency violations are not significantly different from their responses to superiority violations. The frequency of choices of embedded scope by the JSE are lower in this case than in cases (b) and (c), but this difference does not reach significance. They are again significantly different from the NSE. Statistically, then, the JSE are not sensitive to subjacency violations. There is, nevertheless, a tendency to reject embedded readings when a subjacency-like violation occurs.

Finally, in the one case where an embedded scope reading for a matrix wh-word violates both superiority and subjacency, this has a strengthening effect on the judgements of the NSE. None of them accept the embedded reading, and this response is significantly different from the cases involving superiority or subjacency violations alone. Again the JSE are not significantly different in their responses to this case from the earlier cases. But they are significantly different from the NSE.

The results suggest that despite the awareness of the JSE in this study of the interpretive possibilities of long distance *wh-word ... gap* dependencies, and despite reports from earlier studies that Japanese speakers of English can acquire a range of properties associated with English wh-movement to interrogative C, they are nevertheless not sensitive to constraints that apply to movement: superiority

and subjacency. By implication, they have some other representation for *wh-word ... gap* dependencies in English than movement to C[Q] driven by a [*u*wh*:] feature. Their ILGs are missing a feature, despite appearances.

10. A compensatory representation for the missing [*u*wh*:]: [*u*Foc*:]?

How might the JSE in the Hawkins & Hattori study be compensating for a missing [*u*wh*:]? A number of the world's languages appear to require leftward movement of wh-phrases for reasons other than local valuing of [*u*wh*:]. Bošković (2002) (following Stjepanović 1999) argues that the driving force for the movement of the second wh-phrase in multiple wh-fronting languages like Bulgarian, Serbo-Croatian and Russian, as illustrated in (22), is Focus and not a [wh] feature on C[Q]:

22a.	Koj kogo običa	Bulgarian
	Who whom loves	
	`Who loves whom?'	
b.	Ko koga voli?	Serbo-Croatian
	Who whom loves	
c.	Kto kogo ljubit?	Russian
	Who whom loves	

Kobayashi (2002) and Bailyn (2001) suggest that scrambling in Japanese is leftward movement to a Focus position. Kawamura (2004) suggests that scrambling involves movement to satisfy a Σ-feature, and that such movement has different properties from wh-movement.

Suppose that the JSE in the present study have established ILGs for English where a wh-phrase moves leftwards obligatorily to satisfy a Focus requirement rather than the [*u*wh*:] feature of C? How might this account for the observed behaviour? Assume that the JSE have established a Focus projection in the left periphery of C[Q] sentences with an interpretable feature relating to the identification of non-presupposed information in the clause - call it [ident] for convenience - and an uninterpretable feature [*u*Foc*:]. Further assume that an interpretable [Foc] feature is assigned to all wh-words, on the grounds that wh-words inherently identify non-presupposed information (E Kiss 2002, Kobayashi 2002). In multiple wh-questions, one of the wh-words must move to value [*u*Foc*:]. But it does not matter which wh-word moves. Unlike C[Q, *u*wh*:], the Focus projection does not establish an operator-variable construction; it merely identifies non-presupposed information. Hence there is no interpretively-relevant variable to bind and the intervention of another wh-word with a [Foc] feature does not give rise to a superiority effect:

23 [$_{FocP}$ D[wh, Foc]$_1$ [ident, ~~*u*Foc*~~: Foc]] ... D[wh, Foc]$_2$ ... <D[wh, Foc]$_1$> - no superiority violation

One might explain the mild subjacency effect that arises where one wh-word crosses another that has already been moved to the left periphery as follows. Although wh-words have a [Foc] feature, there is no requirement on them to move. Movement arises because of a [*u*Foc*:] feature in the left periphery of a clause. Movement can be cyclic, with a wh-word moving first to a local Focus position and then on up to a higher Focus position, but one wh-word crossing another already in a Focus position violates `shortest move'.

An alternative possibility that won't be pursued here is that there is no independent Focus projection in the ILGs of the Japanese speakers, but a [*u*Foc*:] feature is added to C[Q] to give C[Q, *u*Foc*:]. Wh-movement is then to C, but driven by a different feature from [*u*wh*:].

11. Conclusion

Much recent SLA work has focussed on the indirect relationship between the suppliance of forms by L2 speakers in performance, and their underlying competence. In particular the claim has been that performance under-represents competence. The present study suggests that, conversely, there may be cases where apparent target-like performance conceals non-target-like underlying competence. I have

argued that although high proficiency L2 speakers of English with Japanese as their L1 might look like they have acquired wh-movement driven by a [*u*wh*:] feature, it turns out that they have a different representation, one consistent with the claim that uninterpretable features not present in a speaker's L1 may be inaccessible in later second language acquisition.

References

Adger, David 2003: *Core syntax: a minimalist approach*. Oxford: Oxford University Press.

Bailyn, John F. 2001: On scrambling: a reply to Bošković and Takahashi. *Linguistic Inquiry* 32, 635-657.

Beck, Maria-Luise 1998: L2 acquisition and obligatory head movement. English-speaking learners of German and the local impairment hypothesis. *Studies in Second Language Acquisition* 20, 311-348.

Brown, Cynthia 2000: The interrelation between speech perception and phonological acquisition from infant to adult. In John Archibald (ed) *Second language acquisition and linguistic theory*. Malden, MA: Blackwell, 4-63.

Bošković, Željko 2002: On multiple wh-fronting. *Linguistic Inquiry* 33, 351-383.

Chomsky, Noam 1998: Minimalist inquiries: the framework. *MIT Working Papers in Linguistics* 15, 1-56.

Chomsky, Noam 2001: Derivation by phase. In Michael Kenstowicz (ed) *Ken Hale: a life in language*. Cambridge, MA: MIT Press, 1-52.

Crain, Stephen & Thornton, Rosalind 1998: *Investigations in Universal Grammar: a guide to experiments on the acquisition of syntax*. Cambridge: Cambridge University Press.

Dekydtspotter, Laurent & Sprouse, Rex 2001: Mental design and (second) language epistemology: adjectival restrictions of wh-quantifiers and tense in English-French interlanguage. *Second Language Research* 17, 1-35.

Dekydtspotter, Laurent, Sprouse, Rex & Anderson, Bruce 1997: The interpretive interface in L2 acquisition: the process-result distinction in English-French interlanguage grammars. *Language Acquisition* 6, 297-332.

E Kiss, Katalin 2002: *The syntax of Hungarian*. Cambridge: Cambridge University Press.

Hale, Ken 1996: Can UG and the brain be distinguished in L2 acquisition? *Brain and Behavioral Sciences* 19, 728-730.

Harley, Heidi & Noyer, Rolf 1999: Distributed morphology. *Glot International* 4, 3-9.

Hawkins, Roger & Chan, Cecilia 1997: The partial availability of Universal Grammar in second language acquisition: the `failed functional features hypothesis'. *Second Language Research* 13, 187-226.

Hawkins, Roger & Hattori, Hajime in progress: Acquisition of English multiple wh-questions by Japanese speakers. University of Essex.

Haznedar, Belma & Schwartz, Bonnie D. 1997: Are there optional infinitives in child L2 acquisition?. In Hughes, E., Hughes, M. & Greenhill, A. (eds) *Proceedings of the 21st annual Boston University conference on language development*. Somerville, MA: Cascadilla Press.

Johnson, Jacqueline S. & Newport, Elissa L. 1991: Critical period effects on universal properties of language: the status of subjacency in the acquisition of a second language. *Cognition* 39, 215-258.

Kanno, Kazue 1996: The status of a non-parametrized principle in the L2 initial state. *Language Acquisition* 5, 317-332.

Kanno, Kazue 1997: The acquisition of null and overt pronominals in Japanese by English speakers. *Second Language Research* 13, 265-287.

Kawamura, Tomoko 2004: A feature-checking analysis of Japanese scrambling. *Journal of Linguistics* 40, 45-68.

Kobayashi, Akiko 2000: The third position for a wh-phrase. *Linguistic Analysis* 30, 177-215.

Lardiere, Donna 2000: Mapping features and forms in second language acquisition. In John Archibald(ed) *Second language acquisition and linguistic theory*. Malden, MA: Blackwell, pp 102-129.

Larson Hall, Jenifer 2004: Predicting perceptual success with segments: a test of Japanese speakers of Russian. *Second Language Research* 20, 32-76.

Li, Xiaoli 1998: Adult L2 accessibility to UG: an issue revisited. In Suzanne Flynn, Gita Martohardjono & Wayne O'Neil (eds) *The generative study of second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Martohardjono, Gita 1993: *Wh-movement in the acquisition of a second language: a cross-linguistic study of three languages with and without movement*. Unpublished doctoral dissertation. Cornell University.

Ohba, Hiromasa in progress: *The acquisition of wh-movement in English questions and relative clauses by native speakers of Japanese*. Doctoral dissertation. University of Essex.

Pérez-Leroux, Ana Teresa & Glass, William 1999: Null anaphora in Spanish second language acquisition: probabilistic versus generative approaches. *Second Language Research* 15, 220-249.

Prévost, Philippe & White, Lydia 2000: Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16, 103-133.

Rizzi, Luigi 1990: *Relativized minimality*. Cambridge, MA: MIT Press.

Schachter, Jacqueline 1990: On the issue of incompleteness in second language acquisition. *Second Language Research* 6, 93-124.

Smith, Neil & Tsimpli, Ianthi-Maria 1995: *The mind of a savant: language learning and modularity*. Oxford: Blackwell.

Sorace, Antonella 1993: Incomplete vs. divergent representations of unaccusativity in near-native grammars of Italian. *Second Language Research* 9, 22-47.

Sorace, Antonella 2003: Near-nativeness. In Catherine J. Doughty & Michael H. Long (eds) *Handbook of second language acquisition*. Malden, MA: Blackwell, 130-151.

Stjepanović, Sandra 1999: *What do second position cliticization, scrambling, and multiple wh-fronting have in common*? Unpublished doctoral dissertation. University of Connecticut, Storrs.

Tsimpli, Ianthi-Maria 2003: Interpretability of features in first and second language acquisition: Greek clitics and determiners. Ms. Aristotle University of Thessaloniki.

Tsimpli, Ianthi-Maria & Roussou, Anna 1991: Parameter resetting in L2? *University College London Working Papers in Linguistics* 3, 149-169.

White, Lydia 2003: *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

White, Lydia & Juffs, Alan 1998: Constraints on wh-movement in two different contexts of non-native language acquisition: competence and processing. In Suzanne Flynn, Gita Martohardjono, & Wayne O'Neil (eds) *The generative study of second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Yusa, Noriaki 1999: Multiple-specifiers and wh-island effects in L2 acquisition. In Elaine Klein & Gita Martohardjono (eds) *The development of second language grammars: a generative approach*. Amsterdam: John Benjamins.

Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)

edited by Laurent Dekydtspotter,
Rex A. Sprouse, and Audrey Liljestrand

Cascadilla Proceedings Project   Somerville, MA   2005

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:

Hawkins, Roger. 2005. Revisiting Wh-movement: The Availability of an Uninterpretable [wh] Feature in Interlanguage Grammars. In *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*, ed. Laurent Dekydtspotter et al., 124-137. Somerville, MA: Cascadilla Proceedings Project.

or:

Hawkins, Roger. 2005. Revisiting Wh-movement: The Availability of an Uninterpretable [wh] Feature in Interlanguage Grammars. In *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)*, ed. Laurent Dekydtspotter et al., 124-137. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1161.