

# Verb Second in Advanced L2 English: A Learner Corpus Study

**Tom Rankin**

Vienna University of Economics and Business

## 1. Introduction

The present study examines the interface between syntax and discourse-pragmatics in the production of verb second (V2) structures in a corpus of English texts by advanced L1 German and Dutch speakers. The evidence shows that the residual V2 produced by the learner groups studied is the result of a deficit at the interface rather than the transfer of narrow V2 syntax per se. The analysis offered suggests that the nature of V2 in the learners' L1 combined with evidence in the English input make it difficult for learners to definitively lose the V2 constraint, which remains a persistent option after the fronting of certain constituents.

This builds on the study of V2 transfer by Robertson & Sorace (1999). Their study looked only at patterns of inversion in the learner corpus and expanded this by eliciting judgements on various different inversion structures. In order to test the hypothesis that V2 transfer is connected to a deficit at the interface between syntax and discourse-pragmatics, the present study investigates various diagnostics for V2, i.e. verb movement past adverbs and negation and lexical verb movement in questions in the L1 German and Dutch components of the International Corpus of Learner English (ICLE) (Granger et al 2002).

## 2. Generative SLA and Learner Corpus Linguistics

The research reported here also seeks to exploit a methodological interface between generative approaches to SLA and learner corpus linguistics. As generative SLA and corpus-based research into learner language have moved along parallel tracks, without any major points of contact or mutual influence, I begin by explaining why just such an interface should be beneficial for both fields. Because of the nature of learner corpora, which tend to be made up of written production data by advanced students (usually of English), the field has traditionally been concerned with the nature of advanced grammatical competence and in particular has sought to study the locus of residual 'non-nativeness' in the (usually written) production of advanced language students. In addition to studying the differences between learner and native language on the basis of the frequency of occurrence of specific lexical items or grammatical structures, an assumption underlying much corpus-based research is that non-nativeness is to be found at the level of discourse or information structure in learner texts even after the grammatical system at the level of sentence grammar has been mastered. This in effect mirrors recent developments in generative SLA, where the interfaces between narrow syntax and the domains of discourse-pragmatics and lexical-semantics have been proposed as the locus of residual optionality and L1 transfer effects long after the narrow syntax of the target language appears to have been mastered by very proficient second language learners (e.g. Sorace 2003).

Despite the apparent convergence of interests in both fields, the generative SLA and learner corpus communities have remained disconnected (cf. Granger 2009). While researchers in the field of learner corpus linguistics have often emphasised the concrete pedagogical implications of their work (see for example papers in Aijmer 2009) generative SLA researchers have tended to concentrate on theoretical questions pertaining to the role and nature of UG and the L1 in the process of acquiring a second language and have not studied the ramifications of generative findings for language pedagogy. However, as Alan Juffs (2009) suggests in his plenary talk, detailed theory-based linguistic analyses could make significant contributions to the field of language pedagogy.

I would argue that a rapprochement between the fields of learner corpus linguistics and generative SLA could begin by applying generative theoretical analyses to learner corpus data. In this way we gain a new viewpoint from which to study the interfaces in SLA by investigating how learners deploy discourse-pragmatic competence in context and we will perhaps begin to see whether and how theoretical analyses of learner production might be exploited to help learners improve their mastery of information structuring in a second language.

### 3. The Interface Hypothesis

The interfaces between syntax and other domains have been the focus of intense research recently in generative SLA. The Interface Hypothesis can be defined as the generalization that “narrow syntactic properties are completely acquirable in a second language, even though they may exhibit significant developmental delays, whereas interface properties involving syntax and another cognitive domain may not be fully acquirable” (Sorace and Filiaci 2006, 340). This is in fact a rather strong version of the hypothesis. It has been suggested that interface properties, while apparently posing more problems than narrow syntax for L2 learners, may in fact be ultimately acquirable (cf. Rothman 2009).

Another important factor usually assumed in the Interface Hypothesis is that residual optionality involves not just non-mastery of the interfaces but also the continued influence of the L1. In particular, where the L1 has a more economical narrow syntactic structure for the relevant target language feature to be acquired, this will be evident in continued transfer even at very advanced proficiency levels. The definition of “residual optionality” offered above, as a particular pattern of preferences over time, can now be further refined as below, again drawing on Sorace’s words (2006: 111-112 my emphasis):

“One of the characteristics of L2 advanced grammars... is residual optionality, that is *unsystematic L1 effects* surfacing in the L2 speakers’ production.”

Consider the Italian data in (1) and (2) (from Sorace 200: 112-113).

- (1)
- a. Perche Maria e andata via?  
Why Maria is gone away?
  - b. (perche) lei ha trovato un altro lavoro  
(because) she has found another job.
  - c. (perche) \_\_\_\_ ha trovato un altro lavoro.  
(because) has found another job.

Sorace argues that, in these examples, the answer in (1b) would be typical of English L1 speakers of Italian, even at near-native stages of acquisition, where they optionally realise the subject overtly whereas a native speaker of Italian would more naturally produce (1c) in response to the question, with a null subject pronoun. A similar pattern can be observed for the position of the subject in relation to the verb in the examples in (2) (Sorace 2006: 113).

- (2)
- a. Che cosa e successo? What happened?
  - b. Gianni e partito  
Gianni is left.
  - c. E partito Gianni.  
Is left Gianni.

Again, the answer in (2b) would be characteristic of an English learner of Italian, where the subject is optionally realised in preverbal position. In native Italian, the natural placement in response to this sort of all-focus question would be postverbal. The relevant point is that the null-subject parameter syntactically licences null and post-verbal subjects, but that the distribution of these forms in discourse is regulated by the discourse-pragmatics interface. Sorace (2006: 113) notes that there is a striking asymmetry in the sort of pattern observed with advanced English-speaking learners of Italian with regard to these phenomena. A near-native (English L1) learner of Italian as a second language may optionally produce overt subject pronouns and preverbal subject pronouns in those contexts where native Italian would require null subjects and postverbal subjects. However, it is claimed that the

reverse does not hold, English learners of Italian do not overgeneralize in the opposite direction, i.e. they do not produce null subjects and postverbal subjects in pragmatically inappropriate contexts. The fact that these options, which are also not instantiated in English, are seemingly mastered, leads Sorace (2006: 113) to conclude that “these speakers therefore have acquired a null-subject grammar. The optionality in their grammar does not affect the syntactic licensing of null subjects, but is at the level of discourse conditions on the distribution of pronominals and on the placement of subjects.”

Where English has a more economical pattern of always overtly realising pronominal subjects in preverbal position, this is still transferred into the L2. There is an ongoing debate over how exactly concepts such as economical and complex should be understood. It may mean either in terms of processing resources or in terms of representations. This is an issue that is not addressed in the present paper and for the moment I rely on a linguistic definition of markedness proposed by Henry and Tangney (1999) as the less complex grammar being that which instantiates movement in all contexts (see Discussion below).

### *3.1. An Interface Account of V2 Transfer*

Robertson and Sorace (1999) studied patterns of transfer of V2 in the English of German speaking learners using corpus and magnitude estimation evidence. The variable empirical findings for which they develop an interface account of V2 transfer are summed up as follows (Robertson & Sorace 1999: 343).

- i. Learners who use the residual V2 constraint are never a majority at any level. The proportion declines steadily over the course of development but there remains a significant minority who use the construction at advanced levels.
- ii. The use of the V2 constraint by those learners who do use it is never categorical, even in the earliest stages of IL grammar, and it declines in frequency among those who use it over the course of development.
- iii. The use of the V2 constraint is restricted to non-thematic verbs; there are no examples of main verb raising in the corpus.
- iv. The use of the V2 constraint is clearly due to the influence of the L1 grammar.
- v. The use of the V2 constraint is not motivated by functional, pragmatic or stylistic considerations.

The analysis Robertson and Sorace develop to account for this rests on two main theoretical assumptions: (1) that V2 is a reflex of the checking of a strong C feature, rather than a strong V-feature on C, and (2) that the values of these strong C-features ([+affect], [+wh], [-neg], etc) are abstract lexical items, which are available for selection from the lexicon and can therefore be included in the Numeration like any other lexical item. The selection of one of these features for inclusion in the Numeration is determined by the pragmatically motivated choices of a speaker. In German, for example, where any constituent is fronted for discourse-pragmatic reasons, the presence of a strong [+affect] C-feature then forces overt movement to C for checking purposes. Where the feature is not part of the Numeration it plays no role in the derivation.

Applying this reasoning to the case of German speaking learners acquiring English, the assumption is made that, analogously to the ‘false friends’ phenomenon, the lexical entries of the strong L1 C-features may be copied from German into the IL lexicon. These lexical entries are then available for inclusion in the Numeration and where one is included for specific discourse-pragmatic reasons, it will force movement to C. If it is not included, there will be no overt verb movement. “Optionality then derives from the exercise of choice at the point of Numeration” (Robertson and Sorace 1999: 353).

The developmental pattern is explained by the fact that in response to TL input an initial obligatory V2 grammar must be restructured as there is ample evidence in English that fronting a constituent in declarative clauses does not give rise to movement to C. The learner therefore learns that the relevant feature should not be included in the Numeration and this is a self-reinforcing learning mechanism whereby lack of use leads to attrition in feature strength (p 354). The learning process for German speakers targeting English is to establish which C-features are strong ([neg] and [wh]) and which are not.

The locus of the problem on Robertson and Sorace's (1999) account is therefore syntactic interfaces. One could argue that it is either at the interface of syntax with the lexicon, i.e. the projection of lexical items in the syntax, or at the interface with discourse-pragmatics as the relevant strong features are connected to the exercise of discourse-pragmatic options. Sorace (2005: 64) suggests an interpretation which casts the optionality in the light of the syntax-discourse interface. Thus it is suggested that "V2 phenomena are related to the specification of the illocutionary force of an utterance, and ultimately to the speaker's pragmatically motivated choice, for example, the decision to put a constituent in focus, or to topicalise it" (Sorace 2005: 64). English diverges from German in that V2 is, in Sorace's terms "lexically conditioned" and therefore restricted to certain types of fronted constituents, while German has a generalised V2 pattern. The difficulties for German speaking learners reside at the discourse interface and mastering the less consistent syntactic consequences of pragmatic choices in English.

## 4. Linguistic Background

### 4.1. V2 in German and Dutch

Verb movement in general and in particular the V2 phenomenon in the Germanic languages have been the subject of intense study in formal and theoretical linguistics for some time. In spite of the attention paid to the phenomenon, there is no definite consensus on the detailed mechanics of V2 movement. Given the wealth of possible analyses for V2, it would be prudent at this point to highlight a caveat suggested by Bohnacker (2006: 451): "In syntactic acquisition research it is generally advisable to keep the – often ephemeral – formal syntactic apparatus to a minimum."

In order to keep the analysis as simple as possible, and translatable into whatever syntactic machinery one might prefer, the 'classic' analysis of V2 by den Besten (1983) will be assumed. Thus it is assumed that V2 is derived by movement of the finite verb from an underlyingly head-final VP through I to C. This is accompanied by topicalisation of some other constituent to sentence-initial position (the prefield).

Ideas abound on what the motivation for this movement might be. Among the ideas proposed in the literature, V2 languages may be analysed as hosting either tense and finiteness features on C (Platzack 1986a, b) or discourse-relevant features such as topic/focus (Frey 2005) or as having a Split-CP with different projections each of which encode the illocutionary force of the clause (Westergaard & Vangsnes 2005). For the sake of simplicity, I will adopt here a Minimalist update of the classic analyses as outlined in Adger (2003: 329-330). On this view, matrix C values the uninterpretable [Decl] clause-type feature on T as strong and T, to which the finite verb has moved, raises to C in declaratives giving rise to V2. C likewise bears a strong topic feature which must be checked by movement of some constituent which is interpreted as the sentence topic. Any constituent can bear a [top] feature in German and Dutch and be moved to sentence initial position. The sentence topic is by default the subject when there is no other information structural motivation to front a different sentence constituent.

### 4.2. Mixed V2 in English

While English is a non-raising language in which finite verbs always remain in situ in VP, there are a number of instances where a (residual) V2 constraint is in evidence. Thus there is evidence that 'be' and 'have' raise, in the case of 'be' apparently in line with a V2 constraint (3).

- (3) Next to the English Department is the administration building.

Furthermore, Rizzi (1996) characterises aux-movement in questions and negative inversion as instances of a residual V2 constraint. Westergaard (2007) suggests that, in addition to interrogatives, English shows evidence of a V2 constraint in various clause types when certain information structural conditions are in place. In this analysis, stylistic inversion (SI) as in (4) is derived by VP remnant movement when the verb is informationally lighter than the subject.

- (4) On the table lay some papers.

This sort of structure has traditionally been analysed in terms of non-movement of a subject base generated as the complement of an unaccusative verb. This is not the place to venture into the intricacies of various analyses of SI in English. It has been suggested, however, that the occurrence of unaccusatives in this structure is an artefact of a more general discourse-pragmatic constraint on the verb that it cannot introduce into the discourse information which is relatively newer than the fronted constituent (Birner 1995; Levin & Rappaport-Hovav 1995). Whatever the syntactic derivation, the relevant fact is that this is one of a number of structures which are identical on the surface in English and German/Dutch, e.g. (5) and (6).

- (5) Naast het instituut voor Anglistiek is het administratiegebouw.  
Neben dem Institut für Anglistik ist das Verwaltungsgebäude  
Beside the English Department is the administration building
- (6) Op de tafel liggen enkele papieren.  
Auf dem Tisch lagen einige Papiere.  
On the table lay some papers

The task facing a V2-L1 learner of English is therefore twofold. They must firstly realise that lexical verbs in English do not move out of VP. The second, and perhaps more difficult, task is to realise that while English does license surface V2, this is constrained by a number of factors including (1) the nature of the fronted constituent as fronted operators trigger aux T-to-C movement, (2) the type of verb in inversion, (3) the relative discourse status of the subject and fronted constituent. While detailed explanations of these discourse-pragmatic factors tend not to be covered in language classes, the learners in this study, who are all university students of English Language and Literature will undoubtedly have experience of the these sorts of structures in the type of input they are exposed to, including literary works and more formal written texts.

Building on Sorace (2005), it is hypothesised that resetting the verb movement parameter for English will be relatively straightforward, i.e. that lexical verbs never raise. On the other hand, mastering structures at the interfaces with discourse-pragmatics will prove persistently problematic. It is therefore predicted that evidence confirming the Interface Hypothesis would be if there is a distinction between V2 in narrow syntactic versus interface contexts. Narrow syntactic contexts are operationalised as (1) subject-initial clauses where evidence for V2 would be lexical verb movement over adverbs and negation and (2) in questions, where transfer in narrow syntax would be movement of lexical verbs to second position. Interface contexts are operationalised as those clauses which have fronting or topicalisation of a non-subject XP, given that mastering discourse-pragmatics will be more problematic, it is to be expected these contexts may prove more difficult to master and may show optional transfer of V2 from the L1.

## 5. The Study

### 5.1. Learner Corpora

The German and Dutch L1 components of ICLE were used, henceforth ICLE-DU and ICLE-GE. Each of the corpora was tailored for the present study by using the filtering criterion that only those texts produced by monolingual native L1 speakers of German and Dutch were included, any participants who listed second or third languages spoken at home were not included. This resulted in the final make up of the ICLE corpora used in the study as outlined in Table 1.

There are no independent measures of proficiency of the learners in ICLE. They are defined as advanced on the basis of their educational level (see Discussion). All students are in the final year of a university course in English Language and Literature.

	<i>ICLE-DU</i>	<i>ICLE-GE</i>
Word Count	218555	221621
Text Types <sup>1</sup>	237 Argumentative 10 Literary	393 Argumentative 14 Literary
Task Setting	174 Non-exam 33 Exam	213 Non-exam 158 Exam
Timing	207 Untimed 14 Timed	163 Timed 211 Untimed
<b>Total</b>	<b>247</b>	<b>407</b>

Table 1: Make up of the learner corpora

As will become obvious, while the results for the V2 diagnostics are qualitatively similar in both learner corpora, the L1 German learners consistently produce more instances of apparent V2. This can perhaps be explained by the fact that a greater proportion of the German learners' texts were produced under pressure, i.e. in examination conditions or with a time limit

The LOCNESS corpus of native English essays was used as a point of reference for the learner corpora. The make up of the corpus is outlined in Table 2.

<i>Word Count</i>	<i>Text Types</i>
149,574	Argumentative Texts (American University Students)
18,826	Literary Texts (American University Students)
59,568	Argumentative and literary texts (British University Students)
60,209	Argumentative Texts (British A-Level Students) <sup>2</sup>
<b>288,177</b>	<b>Total</b>

Table 2: Make up of LOCNESS Native English Corpus

## 5.2. Method

The subcorpora of ICLE were all tagged using the CLAWS7 tagset in WMatrix (Rayson 2008). The potential diagnostics for verb movement were then extracted semi-automatically by searching for tag sequences which represent the surface orders  $V_{LEX}\text{-Adv-Obj}$ ,  $V_{LEX}\text{-Neg-Obj}$ ,  $XP\text{-V-Subj}$ . In addition, all questions were identified using searches for question punctuation. The search results were then sorted manually to identify the linear sequences of interest here.

The sorting criteria excluded all instances of VAO orders where the adverb intervening between a lexical verb and its object could be interpreted as taking narrow scope of the object, as in (7).

- (7) Such a stay abroad is not only a guarantee for a high level of proficiency in the foreign language, but it **gives also** evidence of adventurousness and independence. ICLE-DU

<sup>1</sup> For each of variable Text Type, Setting and Timing, there are a small number of texts which for which there is no information, thus the numbers here do not equal the total number of texts in the corpora.

<sup>2</sup> Final year of high school, 17-18 years old.

Once the linear orders of interest were extracted from the corpus, various factors were coded for each of the tokens allowing a qualitative comparison between the corpora. For inversion, the type of the initial XP was coded as well as the type of verb. Verbs in inversion structures were coded as either copula *be*, auxiliary, passive or lexical intransitive/transitive. It is normally assumed that verbs which may occur in stylistic or locative inversion in English are unaccusative, but as inversion may be also used as a test for unaccusativity, the transitive/intransitive distinction was used as transitive verbs are generally ungrammatical in stylistic inversion in English and so should prove more revealing. It can be noted already that the vast majority of inverted intransitive verbs were unaccusatives (see below). For VAO the semantic type of the adverb and the finiteness of the verb were coded.

The results reported here exclude negative inversion and quotative inversion in declarative clauses. In all the examples quoted from the corpora, the relevant portions of the clause are in bold.

## 6. Results

### 6.1. Inversion in Declaratives

Overall there were 102 instances of subject-verb inversion in declaratives (44 ICLE-DU, 58 ICLE-GE). In order to contextualise this, it can be estimated that there are at least 4600 sentences beginning with a non-subject XP in ICLE-GE. This is based on counts of sentence-initial constituents in a random sample of three batches of 100 sentences extracted from the corpus.<sup>3</sup> On average 43% of sentences started with a non-subject XP, and the corpus contained 10,926 sentences. Inversion therefore occurs in only a fraction of possible contexts. What is more relevant is how inversion in the learner corpora diverges qualitatively from inversion patterns in English.

The learners also produced inversion which conforms to target language norms for locative and stylistic inversion. However, the most obvious divergence between the patterns of inversion in the native corpus and inversion in the learner corpora is the type of verb which occurs in the structure (Table 1<sup>4</sup>). Also, even where certain verb types, such as copula *be*, would be licit in inversion in native English, the learners show divergence indicative of residual V2 in terms of the fronted constituent (see 6.1.1).

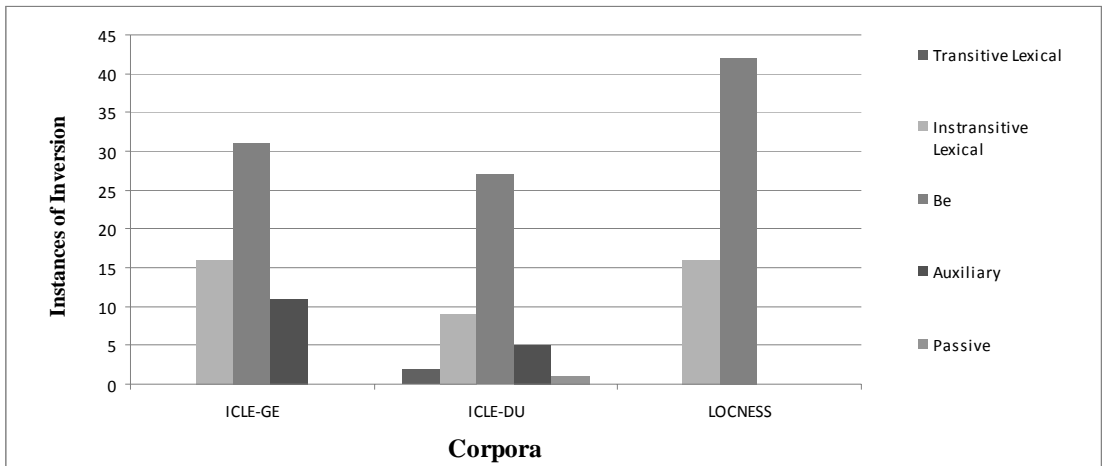


Figure 1: Types of Verbs in Inversion

Intransitive lexical verbs in the native corpus are overwhelmingly prototypical unaccusatives. 12 of the 15 (80%) instances of inversion around intransitive lexical verbs in LOCNESS involve the lexical item ‘to come’. As in the Robertson & Sorace study, the overall pattern here would seem to indicate that residual V2 is confined on the whole to the movement of auxiliary verbs, note however that there are also two instances of V2 with transitive lexical verbs in the Dutch L1 corpus. It is striking that

<sup>3</sup> Extracted using sentence splitter ‘SentParBreaker’ [http://text0.mib.man.ac.uk:8080/scottpioa/sent\\_detector](http://text0.mib.man.ac.uk:8080/scottpioa/sent_detector)

<sup>4</sup> This excludes the results for negative inversion where auxiliaries would be raised to C.

expletive do-insertion is used to maintain V2 order in some instances where there would otherwise be no modal or aspectual auxiliary available for movement (8). This would suggest that while there is a residual V2 constraint, it is not purely syntactic in nature. It seems clear that English narrow syntax has been acquired to the extent that the learners do not permit thematic verbs to move out of VP.

- (8) **Nowadays do not only students** cry out loud, but it's especially their parents who run amok among schools and ministries in order to get them ban this "cruel, inhumane institution" forever. (ICLE-GE)  
**And still do they** have a very powerful army. (ICLE-DU)

The question that suggests itself is whether there are other regularities with those instances of inversion of auxiliaries in declaratives; in particular, whether the type of fronted constituent might play a role. Robertson & Sorace (1999: 336) observe that there is no particular pattern in their corpus data and a range of constituents occur in initial position without any specific prominence in their discourse status. The patterns in ICLE also show little sign of regularity in terms of the syntactic type of initial constituent. One argument DP and adverbials of various types occur in initial position where auxiliaries are inverted. See for example the representative range of initial XPs in (9). While there is little syntactic regularity between the types of constituents, it will be argued later that there may be similarities in terms of what it is possible to topicalise in German and Dutch as opposed to English and that initial position in a V2 sentence provides an economical means for V2-L1 learners to structure information (see Discussion).

- (9) **Only has this place** become smaller.  
**This excellent example of this principle, can we find** whenever two countries or nations went to war.  
**And secondly, can the government** effectively use television to control and influence the public opinion.  
**Already then did America** see itself as a kind of global cop (they used the term very frequently in post-war months).

### 6.1.1. Inversion of copula 'be'

Copula *be* in English patterns like a lexical verb in a V2 language. It inverts in questions and raises over negation. Similarly, fronting the complement of 'be' gives rise to inversion in declaratives. This is a pattern which the learners on the whole seem to have acquired, and if anything have overgeneralised an even more consistent V2 pattern for 'be' than native English allows (see interrogatives below).

However, again the difficulty seems to be one of discourse-pragmatics rather than syntax. This can be illustrated by patterns where an adjectival complement is fronted. In these contexts in the native corpus, there is only inversion after comparatives (10), or after constituents which establish an overt link with the immediately preceding discourse through the use of 'also' (11).

- (10) **More importantly then is the question** of 'Is it right to deprive a child of life because of genetic defects'.  
**Even more alarming is the fact** that heroine, LSD, and peyote are other Schedule 1 drugs.
- (11) **Also taught in his curriculum was that women** who have abortions are more "prone to suicide" and up to 10 percent of them will never again conceive.  
**Also related to this topic is nuclear waste.**

Similar patterns are to be found in the learner corpora but the learners also permit inversion after bare fronted adjectival complements, i.e. where there is no comparison or no explicit link to the previous discourse (12).

- (12) **Important for today is the positive acknowledgement** of the each opposite sex's qualities.  
**Essential is just who decides** what we can watch and why. (ICLE-GE)  
**Ironic is Jim's remark** about this.  
**Striking is the absence of self-esteem** in the black community in the story. (ICLE-DU)



Yet again the most satisfactory explanation here would be that the narrow syntax has been acquired, i.e. the learners know that 'be' has V2 distribution in English. The residual non-nativeness is to be found in discourse-pragmatics, i.e. in what constituents can be felicitously fronted giving rise to inversion around 'be' in declaratives.

## 6.2. Verb-Adverb-Object

The results for VAO orders in ICLE-GE would seem to suggest that there is in fact a more widespread transfer of verb movement from the L1 than could be found based on the inversion data. In ICLE-GE 18 instances of adverbs intervening between finite lexical verbs and their objects account for 3% of all occurrences of adverbs preceding main lexical verbs ( $n=463 \text{ Adv-V}_{\text{LEX,FIN}}$ ). This is rather surprising in the context of overall results for inversion, see representative examples in (13).

- (13) At 6.30 I arrived in Mark's prison cell accompanied by a warder who **closed immediately** the door behind us. (ICLE-GE)  
 Another man **saved always** a part of his earnings in order to be able to fulfill his most cherished dream. (ICLE-GE)

In ICLE-DU there is only one example of an adverb intervening between a finite lexical verb and its object. There are, however, a further 10 examples where auxiliaries and participles in periphrastic tenses and non-finite verb forms appear to raise over adverbs (14).

- (14) In her egoism she is unable **to see objectively** where she is wrong.  
 It is important to make sure that all people remain **to have equally** access to all communication techniques.  
 People today go on trying to find new solutions for new problems, **using as well** all of their imagination

Similar patterns are to be found in the German L1 corpus. In addition to the 18 instances of  $V_{\text{LEX,FIN}}\text{-Adv-Obj}$ , there are a further 9 examples where an adverb intervenes between a non-finite verb form and its object as in (15).

- (15) From one moment to another the silent crowd is turning to an exploding and errupting vulcano **applauding frenetically the runner**.  
 But one has **to take sensitively care** that nobody is hurt thereby!

There is no discernible pattern with regard to the type of adverb which may occur in VAO sequences in either of the learner corpora (see Table 3). This, together with the fact that both finite and non-finite verbs occur in VAO sequences, will be taken as evidence that linear VAO order is indicative of a problem with adverb placement rather than verb movement (see Discussion).

<i>Adverb Type</i>	<i>VAO in ICLE-DU</i>	<i>VAO in ICLE-GE</i>
Additive	1	1
Aspectual	0	5
Degree	2	3
Frequency	1	5
Manner	2	7
Modal	1	2
Stance	0	1
Temporal	4	2
Total	11	26

Table 3: Types of Adverbs in VAO sequences

### 6.3. Inversion in Questions

The results for verb movement in interrogatives or over negation can both be summarised succinctly as they seem to indicate that there was no true V2 in these contexts. There were 421 interrogatives in ICLE-GE and 303 in ICLE-DU. All of these conform overwhelmingly to target English syntax. The isolated instances of lexical verb movement (n=7 ICLE-GE, n=2 ICLE-DU) are not indicative of L1 transfer but rather seem to be the result of overextending T-to-C to the lexical equivalents of auxiliaries (16). Indeed, the examples with lexical *have* below would also be perfectly acceptable in various varieties of English.

- (16) **Has television** as much influence on people as religion had in former days? ICLE-DU  
**What have Serbia, Croatia and Bosnia** to do with happiness? ICLE-GE

Similarly, there is overgeneralisation of inversion with ‘be’ in contexts where participial or infinitival forms are raised (17) (n=1 ICLE-DU, n=3 ICLE-GE)

- (17) However, if it hadn't been for some strong women who enchained this emancipation movement, **how would have been the women's situation** today?  
**Shouldn't be there** a speed limit on German motorways in order to avoid such photos in future? (ICLE-GE)

There is only one example that could be taken to be representative of ‘true’ lexical verb movement in either of the corpora (18).

- (18) **Exclude this two things** themselves mutually? (ICLE-GE)

### 6.4. Verb Movement over Negation

The results for negation also show that the non-movement of lexical verbs in narrow syntactic contexts has been mastered. Patterns of negation with do-support, copula/auxiliary *be* and other auxiliary verbs are illustrated in Table 4 with the tag sequences for each structure.

<i>Tag Sequence</i>	<i>ICLE-DU</i>	<i>ICLE-GE</i>
&VD* &XX ( <i>do</i> + Neg)	764	807
&VH* &XX &V* ( <i>have</i> + Neg + V <sub>LEX</sub> )	65	71
&VB* &XX ( <i>be</i> + Neg)	658	626
&VM &XX 654 ( <i>Modal Aux</i> + Neg)	527	654
<b>Total</b>	<b>2014</b>	<b>2158</b>

Table 4: Patterns of Negation in the Learner Corpora

There are only 15 instances of thematic verbs raising over negation (12 ICLE-GE; 3 ICLE-DU) and of these, all examples in ICLE-DU involve *have*-raising, which again may be licit in native English, although there are no occurrences in LOCNESS. ‘Have’ also occurs in 25% of V<sub>LEX</sub>-Neg sequences in ICLE-GE. It is questionable whether the remaining examples in ICLE-GE are instances of V2. They are either formulaic as in (19) or due to allowing negation to take narrow scope over a particular constituent as in (20), rather than instances of the verb raising over sentential negation.

- (19) **I think not** as there are too many different backgrounds, languages, traditions and religious beliefs. (ICLE-GE)
- (20) Alice Walker **leaves not doubt** that black people's traditions are different from that of their fellow white citizens. (ICLE-GE)

## 7. Discussion

Assuming that there is no evidence of V2 movement over sentential negation or in questions, it is a challenge to reconcile the evidence from subject-verb inversion with the comparatively widespread occurrence of VAO. While inversion suggests that residual V2 is confined to auxiliaries and copula *be* in interface contexts, the VAO sequences would seem to indicate that the learners (at least the L1 German learners in the study) still permit lexical verb movement in narrow syntactic contexts. There are, however, two possible explanations which can account for the anomaly. Firstly, as already suggested, it is possible that the linear order VAO represents a general problem with mastering adverb placement in English, rather than the transfer of a verb movement operation. Chu & Schwartz (2005) have argued that the production of VAO sequences in L2 English by L1 Chinese speakers can be explained as a reflex of the nature of the input as Chinese is also a non-raising language so transfer is not possible. The “irregularity of the input” where English adverbs may occur in a range of surface positions might be misleading for (especially beginner) learners of English and make them unsure about adverb placement in general (Chu & Schwartz 2005: 83). As we have seen, the occurrence of VAO sequences with non-finite verbs would suggest that a similar adverb-placement problem is at the root of the production of VAO in ICLE. Further evidence comes in the form of the range of adverb types in the VAO structures. In ICLE-GE 30% (8 of 26) adverbs in VAO sequences would more naturally occur clause-finally after the verb and its complements. The occurrence of these adverbs pre-verbally again suggests that adverb placement rather than verb placement gives rise to VAO.

Secondly, it has been suggested that the production VAO order in English by OV/V2-L1 speakers could indicate the transfer of scrambling (Weerman & Neelman 1997). While this analysis will not be pursued in detail here, it suffices to note that there are possible independent explanations for the adverb data. This would lead to the conclusion that residual V2 is confined mainly to inversion of auxiliaries and copula *be* in declarative clauses by the ICLE learners. Narrow syntax has been acquired in that they ‘know’ that English lexical verbs do not move out of VP.

What is being suggested here then is that there is a disassociation between syntactic verb movement and residual V2 in the learner corpora. The V2 patterns can be analysed as the transfer of information structural preferences into L2 English, rather than the transfer of a verb movement operation per se. This would thus provide evidence in support of the Interface Hypothesis.

The argument goes as follows. Fronting of constituents to the prefield in German and Dutch and thus subject-verb inversion, provides a straightforward mechanism by which different discourse-pragmatic effects can be achieved, i.e. topicalisation, contrast etc. (see e.g. Frey 2005 on pragmatic consequences of movement to the prefield in German). Where English has a more complicated mechanism for achieving similar discourse-pragmatic effects, the learners optionally fall back on the straightforward V2 mechanism. This is achieved by raising auxiliaries or copula *be* to C, a licit syntactic operation in English, to create a unique prefield position to host discourse-relevant material. Notice that the infelicity of many of the examples of V2 transfer rests not necessarily in the fact that there is inversion per se, but is caused rather by the fronting of a constituent which could not normally be fronted in English (21)

- (21) **Essential is just who** decides what we can watch and why. (ICLE-GE)  
**Important for today is** the positive acknowledgement of the each opposite sex's qualities. (ICLE-GE)  
**Dead are only those** whom their family and friends forget. (ICLE-GE)  
**Ironic is Jim's remark** about this. (ICLE-DU)  
 We might never be able to realize them, but **most important is that** we keep them alive. (ICLE-DU)

In examples such as these, it would be necessary to use clefting or extraposition to communicate the emphasis on the initial adjective, ‘what is essential is ...’. The learners however can optionally rely

on the more economical structure transferred from the L1, without breaking any ‘syntactic rules’ of English as *be* can raise.

In the examples below, again, the source of the divergence from native English is the nature of the fronted constituent. Because of the nature of the adverbs ‘only’ and ‘already’, they can serve to focus or emphasise the constituents they modify without fronting so the more natural order, with the same discourse-pragmatic effect would be ‘This place has become only smaller’ and ‘America saw itself already then as...’. Yet again, however, there are no non-permissible narrow syntactic mechanisms as auxiliaries can in principle raise to C.

- (22) **Only has this place** become smaller. (ICLE-DU)  
**Already then did America** see itself as a kind of global cop (they used the term very frequently in post-war months). (ICLE-DU)

This, however, raises the question of what it means to say that inversion is more economical for these learners. For the examples in (22), it cannot be argued straightforwardly in the same way as with copula *be* that fronting and inversion are more economical than the clefting/extraposition mechanism as this would also not be possible in these contexts. However, if we expand the notion of economy to include markedness, there is a ready explanation in terms of the whole L1 system versus the whole L2 system, which states that V2 in German and Dutch is less marked than the English system and therefore might persist into advanced stages of acquisition. This is based on Henry and Tangney’s (1999) argument that a grammar is more marked or complex if it has movement only in certain contexts. By comparison, a language which instantiates the same movement in all contexts is less complex. This is a departure from the more traditional view that a strong feature which forces movement is marked while a weak feature is unmarked. The corollary of this for second language acquisition is, according to Henry and Tangney (1999: 240), that an L2 learner “will only adopt a more complex grammar ... where there is *overwhelming evidence* in the input in its favour” (my emphasis).

Assuming a Full Access/Full Transfer (Schwartz & Sprouse 1996) model of the initial state in SLA, the evidence in the input which would force restructuring of a transferred V2 grammar is decidedly underwhelming. As we noted above, there are other instances, specifically with copula *be*, but also with other verbs in stylistic and locative inversion where the surface order in English and in Dutch and German are identical. Westergaard (2003: 91) suggests that the relevant restructuring cues for V2-L1 learners of English are *do*-support as a cue for V-in situ and topicalisation as a cue for V3. She goes on to point out, however, that topicalisation is infrequent in English and that *do*-support is often avoided at the early stages of instructed language learning as it is felt to be a complicated structure. The learners in ICLE will no doubt have had ample exposure to *do*-support, but this is in itself ambiguous. While *do*-support in questions and with negation provides unambiguous evidence that lexical verbs never move, it could be analysed as evidence that English in fact has quite an extensive V2 constraint. S-V<sub>AUX</sub>-Neg-V<sub>LEX</sub> and Wh-V<sub>AUX</sub>-Subj-V<sub>LEX</sub> can be accommodated by a V2 grammar and might be taken to indicate that V2 in English is simply fulfilled by a mechanism other than lexical verb movement.

In spite of this ambiguous evidence, the learners in ICLE appear to have successfully acquired a V-in situ, V3 grammar. Residual V2, however, persists as an option at the interface between syntax and discourse-pragmatics where the V2 option in the form of XP-V<sub>AUX</sub>-Subj, by creating a unique clause-initial position, provides a more straightforward means for encoding topicalisation/focus/contrast than the structures used in English. If this analysis is on the right lines, it predicts that it is more likely that V2 will persist after topicalisation of constituents which are licit in a fronted position in the L1 but not in English, e.g. object DPs in certain structures, bare adjectival complements, etc. The unparsed nature of the corpora does not permit a straightforward test of this hypothesis with the current data set, but it could be a jumping off point for further experimental research.

## 8. Conclusions and Caveats

To sum up, I have proposed, in line with the Interface Hypothesis, that information structural patterns are more difficult to acquire in a second language than narrow syntax. Thus the L1 German and Dutch learners in ICLE continue to optionally transfer V2 into English in certain discourse-pragmatic contexts, even after it seems that they have acquired the narrow syntax of English. In addition to

problems acquiring the interfaces, it has been suggested that ambiguous input evidence in English may make it difficult to definitively lose the V2 constraint and thus the nature of the input conspires with difficulties in acquiring the interfaces to make V2 a persistent residual option for these learners.

To come back to the methodological issues about the use of learner corpora for research in generative SLA, some caveats are in order. It is undoubtedly the case that the nature of the corpora, i.e. written production data, produces a somewhat restricted picture of the competence of the learners in the study. In the absence of conscious monitoring processes, the same learners would no doubt produce more structures which diverge from target English syntax; although whether they would also differ qualitatively would be an interesting question. However, the fact that those V2 structures which remain in the corpora have presumably also passed the filter of conscious monitoring of the written production would suggest that they are representative of the learners' competence rather than a purely production error.

A further caveat is the fact that there is no independent measure of the proficiency of the learners in ICLE. I have accepted the definition of the compilers of the corpus that the fact that the learners are all in the final year of university courses in English Language and Literature is indicative of their 'advanced' status. It would of course be preferable to have some independent measure, but the lack of such a measure is not a weakness of learner corpus linguistics per se but rather is a feature of this particular learner corpus as the aim was to collect samples of language from students at similar educational levels.

The added value of learner corpora is that it allows the researcher to get an idea of what learners *do in context* in the L2. This is especially important for studies of features related to the interfaces as it might be necessary to see how a particular interface phenomenon occurs within discourse (see for example Lozano 2009 for null pronoun distribution in L2 Spanish). This permits the formulation of research questions which might then be tested experimentally. This is in fact a more rigorous version of the approach already taken in much generative work. Much research takes as its starting point the observation that speakers of language X tend to produce feature Y when using language Z as L2, and this is then studied experimentally. Large samples of corpus data can lead to insights which might be a more fruitful starting point, cf. Lozano's (2009) suggestion, based on corpus evidence, that L1 English speakers' difficulties with pronoun distribution in L2 Spanish is confined to 3<sup>rd</sup> person pronouns. These are the sort of insights which might be missed if the research is based on judgements or interpretations of sentences constructed by the researcher.

## References

- Adger, David (2003) *Core Syntax* Oxford: Oxford University Press
- Aijmer, Karin (ed.) *Corpora and Language Teaching* Amsterdam: John Benjamins
- Birner, Betty (1995) "Pragmatic Constraints on the Verb in English Inversion" *Lingua* 97, 233-256
- Bohnacker, Ute (2006) "When Swedes begin to learn German: from V2 to V2." *Second Language Research* 22(4), 443-486
- Chu, Wei and Bonnie Schwartz (2005) "Another Look at 'Verb Raising' in the L2 English of Chinese Speakers." In L. Dekydtspotter et al. (eds.) *Proceedings of the 7<sup>th</sup> Generative Approaches to Second Language Acquisition Conference (GASLA 2004)* Sommerville, MA: Cascadilla, 68-85
- Frey, Werner (2005): 'Pragmatic properties of certain German and English left peripheral constructions' *Linguistics* 43. 1, p. 89-129.
- Granger, Sylviane (2009) "The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation" in Karin Aijmer (ed.) *Corpora and Language Teaching* Amsterdam: John Benjamins, 3-32
- Granger Sylviane, E. Dagneaux and F. Meunier (2002) *The International Corpus of Learner English. Handbook and CD-ROM*. Louvain-la-Neuve: Presses Universitaires de Louvain
- Henry, Alison. and D. Tangney (1999) "Functional categories and parameter setting in the second-language acquisition of Irish in early childhood". In A. DeGraff (ed.) *Language Creation and Language Change: Creolization, Diachrony and Development*, Cambridge MA: MIT Press, 139-253
- Juffs, Alan (2009) 'Second Language Processing Failures and directions for some solutions.' Plenary Talk at GASLA-10, University of Illinois at Urbana-Champaign
- Levin, Beth and M. Rappaport-Hovav (1995) *Unaccusativity. At the Syntax-Lexical Semantics Interface*. Cambridge, Mass.: MIT Press

- Lozano, Cristobal. (2009) "Selective deficits at the syntax-discourse interface: Evidence from the CEDEL2 corpus". In: Snape, N., Leung, Y.-I. & Sharwood-Smith, M. (eds). *Representational Deficits in SLA*. Amsterdam: John Benjamins, pp. 127-166.
- Neeleman, Ad and Fred Weerman (1997). L1 and L2 Word Order Acquisition. *Language Acquisition* 6: 125-170
- Platzack, Christer (1986a) 'Comp, Infl and Germanic word order.' In Lars Hellan & Kirsti Koch Christensen (eds) *Topics in Scandinavian Syntax* Dordrecht: Reidel, 185-234.
- Platzack, Christer (1986b) 'The Position of the Finite Verb in Swedish.' In Hubert Haider & Martin Prinzhorn (eds.) *Verb Second Phenomena in Germanic Languages* Dordrecht: Foris, 27-47.
- Rayson, Paul (2008) Wmatrix: a web-based corpus processing environment, Computing Department, Lancaster University. <http://ucrel.lancs.ac.uk/wmatrix/>
- Robertson, Dan and A. Sorace (1999) "Losing the V2 Constraint." In: Klein, E. and Martohardjono, G. (eds.) *The Development of Second Language Grammars: A Generative Approach*. Amsterdam: Benjamins, pp.317-361
- Rizzi, Luigi (1996) "Residual Verb Second and the WH Criterion" In: A. Belletti and L. Rizzi (eds.) *Parameters and Functional Heads* Oxford: OUP, pp. 63-90
- Rothman, Jason (2009) "Pragmatic deficits with syntactic consequences?: L2 pronominal subjects and the syntax-pragmatics interface." *Journal of Pragmatics* 41
- Schwartz, Bonnie and Rex Sprouse (1996) "L2 cognitive states and the Full Transfer/Full Access model." *Second Language Research* 12(1), 40-72
- Sorace, Antonella (2003) "Near-Nativeness." In: C. Doughty and M. Long (eds.) *The Handbook of Second Language Acquisition* Oxford: Blackwell, (130-152)
- Sorace, Antonella (2005) "Selective Optionality in Language Development." In: L. Cornips and K. Corrigan (eds.) (2005) *Syntax and Variation. Reconciling the Biological and the Social*. Amsterdam: Benjamins, 55-80
- Sorace, Antonella (2006) "Gradedness and Optionality in Mature and Developing Grammars." In: G. Fanselow, C. Fery, R. Vogel, M. Schlesewsky (eds.) *Gradience in Grammar. Generative Perspectives* Oxford: OUP, 106-123
- Sorace, Antonella and F. Filiaci (2006) "Anaphora resolution in near-native speakers of Italian." *Second Language Research* 22(3), 339-368
- Westergaard, Marit (2003) 'Unlearning V2: Transfer, Markedness, and the Importance of Input Cues in the Acquisition of Word Order in English by Norwegian Children' *EUROSLA Yearbook* 3, 77-101
- Westergaard, Marit (2007) "English as a Mixed V2 Grammar: Synchronic Word Order Inconsistencies from the Perspective of First Language Acquisition." *Poznan Studies in Contemporary Linguistics* 43(2), 107-131
- Westergaard, Marit and O. Vangsnes (2005) "Wh-questions, V2 and the Left Periphery of Three Norwegian Dialect types." *Journal of Comparative Germanic Linguistics* 8, 117-158

# Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)

edited by Melissa Bowles, Tania Ionin,  
Silvina Montrul, and Annie Tremblay

Cascadilla Proceedings Project Somerville, MA 2009

## Copyright information

Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)  
© 2009 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-434-8 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://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, e-mail: [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://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:

Rankin, Tom. 2009. Verb Second in Advanced L2 English: A Learner Corpus Study. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed. Melissa Bowles et al., 46-59. Somerville, MA: Cascadilla Proceedings Project.

or:

Rankin, Tom. 2009. Verb Second in Advanced L2 English: A Learner Corpus Study. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed. Melissa Bowles et al., 46-59. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #2256.