1. Introduction

This study is part of a research program seeking to unveil the general processes underlying the production of non-canonical word order constructions in non-native grammars (see WOSLAC project webpage: www.uam.es/woslac). An in-depth investigation into word order in non-native grammars will reveal interesting results regarding, for instance, the relative difficulty of acquiring lexicon-syntax and syntax-discourse interface properties, as well as issues to do with transfer and the role of input which are very much at the centre of debate in second language acquisition (SLA) theory.

We focus on the analysis of V(erb) S(ubject) structures using experimental data from L2 Spanish, as well as corpus data from L2 English. One of the most important conclusions of the research presented here is that a full account of word order acquisition needs to take into account properties at three interfaces: lexicon-syntax, syntax-discourse and syntax-phonology. In this paper, we show that learners of Romance (L1 English/Greek – L2 Spanish) as well as Romance learners of English (L1 Spanish/L1 Italian – L2 English) are sensitive to discourse status (focus) in their L2, but they show residual (yet persistent) problems when encoding information status syntactically. Our studies suggest that learners’ deficits are not located at the interfaces, but rather in the computational system (narrow syntax).

2. Word Order in neutral and non-neutral contexts

Spanish and Italian have ‘free’ word order: VS order is found ‘freely’ with all classes of V’s, with or without preverbal elements and requiring no special intonation pattern. This is illustrated in (1) for Spanish, with a transitive V like comprar ‘to buy’ (1a) and both types of intransitive V’s: unergative (1b) and unaccusative (1c).

---

* This work has been supported by research grants from the Spanish Ministry of Education and Science (HUM2005-0127/FILO) and from the Autonomous Region of Madrid & Universidad Autónoma de Madrid (09/SHD/016).
(1) Ha comprado un libro María (transitive)  
    has-3sg bought a book María  
    ‘María has bought a book.’

(2) Habló Juan. (unergative)  
    spoke-3sg Juan  
    ‘Juan spoke’

(3) Llegaron tres niñas (unaccusative)  
    Arrived-3pl three girls  
    ‘Three girls arrived.’

    English, however, is a ‘fixed’ word order language: VS order can be found in certain contexts, as in XP V S structures in (4) (corpus examples from Biber et al. 1999: 912-913), which are highly restricted by properties operating at the lexicon-syntax interface: they typically contain unaccusative verbs expressing existence and appearance (see Levin & Rappaport-Hovav 1995).

(4) a. On one long wall hung a row of Van Goghs.  
    b. Then came the turning point of the match.  
    c. With incorporation, and the increased size of the normal establishment came changes which revolutionized office administration

    Since Perlmutter’s (1978) *Unaccusative Hypothesis* two classes of intransitive Vs have been distinguished, which differ in the position occupied by their only argument in the base structure (or D-Structure): unergatives in (5a) have an external argument in the base structure (or D-Structure): unergatives in (5a) have an external argument (in <Spec, VP>, after Koopman & Sportiche’s (1991) VP-internal subject hypothesis), but lack an internal argument. Unaccusatives in (5b), on the other hand, have an internal argument, but no external argument.

(5) a. *unergative*  
    ‘John spoke’

    b. *unaccusative*  
    ‘Three girls arrived’

![Diagram](image-url)
The syntactic distinction between the two classes of verbs is systematically related with a semantic distinction. Unergatives typically denote activities controlled by an agent (*speak* in (5a), and also *cry, cough, sweat, jump, run, dance, work, play*, etc.), while unaccusatives are associated with themes (*arrive* in (5b), and also *blossom, appear, exist, deteriorate, arrive, come*, etc.). The rule of NP-movement applies to promote both NPs in (5) to <Spec, IP> at S-Structure (see arrowed lines), in order to satisfy their Case requirements, or, more recently, the requirement that <Spec, IP> must be filled by an overt element (roughly Chomsky’s (1981) Extended Projection Principle (EPP)).

While in the English sentences in (5) there is no surface manifestation of unaccusativity, after the application of NP-movement, in languages like Spanish and Italian, the argument of unaccusative Vs need not raise to <Spec, IP>. In discourse-neutral contexts (e.g., in answer to questions like ¿Qué pasó? ‘What happened?’), a distinction is found between unergative and unaccusative clauses in Spanish, (6): SV is the favoured structure for unergatives (((6a) vs. (6b)), but VS for unaccusatives (((6a’) vs. (6b’))).

(6) a. María gritó (unergative)  a.’ # María llegó. (unaccusative)  
María shouted-3sg María arrived-3sg  
b. # Gritó María. b.’ Llegó María  
‘María shouted’ ‘María arrived.’

The possibility of having VS structures like (6b’) is associated with ‘free inversion’, which is among the cluster of properties characterising languages positively marked for the Null Subject Parameter (e.g. Spanish/Italian/Greek vs. English) (*inter alia*, Rizzi 1997, Zagona 2002). It has been recently argued that ‘free’ inversion is to be governed by features operating at the syntax-discourse interface. Indeed, it seems to be the case that inversion in Romance serves a focalization purpose (*e.g.*, Belletti 2001, Domínguez 2004, Vallduví 1993, Zubizarreta 1998). Thus, in sequences like those in (7) where the wh-element *quién* ‘who’ triggers narrow focus in the reply, a preverbal S, commonly interpreted as topic in Spanish, is regarded as pragmatically odd (7Bii) (see also similar examples for Italian in Belletti 2001, 2004).

(7) A: ¿Quién gritó/llegó?  
‘Who shouted/arrived?’  
B: i. Gritó/llegó Juan  ii. #Juan shouted/arrived  
shouted/arrived Juan ‘Juan shouted/arrived’

Under the assumption that there is a Focus Phrase which is located between VP and TP (Belletti 2004), the order VS with both unergatives and unaccusatives is the result of raising S to FocusP (as well as V to T), (8).

(8) a. \[TP [T [T gritó] [Focus P Juan [VP Juan [V’ [v gritó]]]]]]

b. \[TP [T [T llegó] [Focus P Juan [VP [V’ [v llegó] [NP Juan]]]]]]
The above data are confirmed by empirical studies on native speakers (e.g., Hertel 2003, Lozano 2006a, 2006b for Spanish; Pinto 1997 for Italian). Verb choice (lexicon-syntax interface) interacts with information structure (syntax-discourse interface) to determine word order: SV is favoured with unergatives and VS with unaccusatives in neutral contexts like (6). But VS is preferred to SV with both unaccusatives and unergatives in narrow-focus contexts, (7). In contrast, in English, as well as in Greek (e.g., Georgiazentis 2004), presentational focus is marked in situ (<Spec,TP>), so the resulting order is SV.

3. L2 experimental evidence: previous studies

At the lexicon-syntax interface, it is well known that English-speaking learners of Spanish are sensitive to the syntactic effects of the Unaccusative Hypothesis from early stages of development (Hertel 2003) and show native-like behaviour at advanced stages (Lozano 2006a). Learners significantly prefer SV to VS with unergatives, but VS to SV with unaccusatives in neutral contexts, as Spanish natives do. But, at the syntax-discourse interface, narrow focused Ss in sentence-final position are acquired late in both L2 Spanish and L2 Italian. Typically, native-like knowledge is never attained, as learners show optionality by simultaneously preferring both SV and VS with both verb types, while Spanish natives clearly prefer the pragmatically adequate order: VS, irrespective of verb type. Similar findings are reported for Greek learners of non-native Spanish (Lozano 2006b) and learners of L2 Italian (Belletti & Leonini 2004).

4. L2 corpus evidence

4.1. Preliminary remarks

In the corpus study, we analyse in detail the production of VS in the writing of Spanish learners of L2 English (Lozano & Mendikoetxea, submitted 2009). Additional evidence comes from the corpus data of Spanish and Italian learners of English (Lozano & Mendikoetxea 2008). The purpose of these studies was to determine whether the properties that govern the occurrence of VS order in the non-native grammars of Spanish and Italian learners of English are the same as those operating in native English.

Though, as pointed out in section 2, English has fixed word order determined by lexicon-syntactic properties, VS order is found in restricted contexts, such as inversion constructions with an opening adverbial (XP V S), as illustrated in (4) above, and there-constructions like those in (9), (from Biber et al. 1999:945).

(9) a. Somewhere deep inside there arose a desperate hope that he would embrace her.
     b. In all such relations there exists a set of mutual obligations in the instrumental and economic fields.
     c. There came a roar of pure delight.
XP V S structures like those in (4) and there-constructions in (9) are very similar in their properties and are governed by features operating at three interfaces: (i) lexicon-syntax, (ii) syntax-discourse and (iii) syntax-phonology.

Regarding (i), at the **lexicon-syntax interface**, Vs in inversion structures are always unaccusative and they typically express existence and appearance (*exist, come, appear, arrive,...*), never unergative, as shown by the contrast between the there-constructions in (10a) and (10b).

(10) a. There arrived four girls at the station (unaccusative)
b. *There sang four girls at the opera (unergative)

At the **syntax-discourse interface** (ii), the S in inversion structures and there-constructions is often focus or relatively unfamiliar information (End-Focus Principle, Quirk *et al.* 1985), while the preverbal material often links the S to previous discourse (relatively familiar information) (see Birner 1994, 1995; Bresnan, 1994), as in (11), an example used in Birner (1995: 237).

(11) We have complimentary soft drinks and coffee. Also complimentary is red and white wine.

Finally, at the **syntax-phonology interface** (iii), choice of ordering is also influenced by phonological or syntactic heaviness: Long and complex elements tend to be placed towards the end of the clause. This is the End-Weight Principle (Quirk *et al.* 1985), as illustrated in (12) (a corpus example from Birner 1994: 254, underlining is ours) and other examples given here (e.g. (4c), (9a-b)).

(12) Michael puts loose papers like class outlines in the large file-size pocket. He keeps his checkbook handy in one of the three compact pockets. The six pen and pencil pockets are always full and in the outside pocket go his schedule book, chap stick, gum, contact lens solution and hair brush.
[Land’s End March 1989 catalog. p. 95]

The conclusion is that subjects which are focus, long and complex tend to occur postverbally in those structures which allow them, i.e., with the right (sub)class of unaccusatives. In contrast, as we have seen, VS order in Spanish and Italian is not restricted by properties at the **lexicon-syntax** interface, as all Vs allow VS structures (though only unaccusatives appear in VS constructions in neutral contexts), but it seems to be governed by properties at the **syntax-discourse** interface, particularly by the End-Focus Principle. As for the **syntax-phonology** interface, we are aware of no specific studies in Spanish/Italian regarding heaviness, but we are assuming that in these languages heavy subjects tend to be postponed, as this has been shown to be a universal processing mechanism: placing long and complex elements towards the end of the sentence reduces the processing burden and, thus, eases comprehension by the receiver and/or facilitates planning and production (Hawkins 1994, Wasow 1997, 2002).
If this is so, the only difference between Italian and Spanish, on the one hand, and English, on the other hand, concerns the lexicon-syntax interface: in English VS is found only with unaccusative Vs, while in Spanish and Italian no such restriction applies. If transfer applied in this area, then we would expect learners to produce VS structures with all verb classes.

4.2. The phenomenon in SLA

VS order has received considerable attention in the literature, which focuses mostly on the production of ungrammatical VS structures in L2 English by speakers from a variety of L1’s, but the issue has been dealt with in quite a scattered, unsystematic and rather intuitive fashion (see, among others, White 1986, Zobl 1989, Rutherford 1989, Oshita 2004). These studies, however, show a remarkably consistent pattern: postverbal subjects are produced only with unaccusatives, but never with unergatives, (13).

(13)a. Sometimes comes a good regular wave. (L1 Japanese; Zobl 1989: 204)
   b. On this particular place called G... happened a story which now appears on all Mexican history books... (L1 Spanish; Rutherford 1989: 178-179)
   c. The bride was very attractive, on her face appeared those two red cheeks... (L1 Arabic; Rutherford 1989: 178-179)

This adds to other types of evidence which point towards the psychological reality of the Unaccusative Hypothesis in L2 acquisition (inter alia, Balcom 1997, Hertel 2003, Hirakawa 1999, Ju 2000, Lozano 2006a, 2006b, Montrul 2004, Oshita 2000, Sorace 1993). This is so despite the fact that English lacks overt marking for unaccusatives, rendering the unergative-unaccusative distinction inaccessible: unaccusatives overwhelmingly appear in SV constructions and there is no auxiliary or morphological marker for unaccusatives. Thus, although inversion structures with unaccusatives are found in English (see § 2 above), the rarity of the construction (Biber et al. 1999: 945) makes it unlikely that such VS order is sufficiently represented in the input data to learners to count as positive evidence.

These findings provide the theoretical basis for our analysis and will be confirmed by our results but there is a crucial difference between our study and previous studies: previous studies focused on the errors and the differences between natives and learners, but our study emphasises the similarities between them. Furthermore, while former studies have focused on the nature of V (lexicon-syntax), we turn our attention to the properties of S in VS structures (syntax-discourse; syntax-phonology). In previous work, we conducted a corpus analysis of VS constructions by Spanish and Italian learners of English as an L2 and found all these conditions to be significant (see Lozano & Mendikoetxea 2008). We now expand our analysis of the Spanish learners data and compare our results with those obtained from a comparable native English corpus.
4.3. Hypotheses

Our general hypothesis is that the conditions licensing VS in L2 English are the same as those in native English, despite differences in syntactic encoding: that is, learners are sensitive to the licensing mechanisms for VS structures, but do not know how to encode them syntactically, thus producing both grammatical and ungrammatical VS structures. Furthermore, we predict that unaccusativity is a necessary, but not a sufficient condition for the occurrence of VS structures, as there are conditions on S, as well as on V. Thus, we formulate the following three hypotheses, (14).

(14)a. **H1**: at the lexicon-syntax interface, VS order will be produced with unaccusatives only (never with unergatives), as predicted by the Unaccusative Hypothesis.

b. **H2**: at the syntax-phonological form interface, postverbal Ss tend to be syntactically long and heavy (while preverbal Ss tend to be light), as predicted by the End-Weight Principle.

c. **H3**: at the syntax-discourse interface, postverbal Ss tend to be new information, focus (but preverbal Ss tend to be topic), as predicted by the End-Focus Principle.

4.4. Method

We took as the basis for our search the inventory of intransitive Vs proposed by Levin (1993) and Levin & Rappaport-Hovav (1995). In total, 73 lemmas (word types) (i.e., 32 unaccusatives and 41 unergatives) were searched in the Spanish subcorpus of ICLE (Granger et al. 2002), and a comparable corpus of L2 English we are collecting (WriCLE), making a total of over a quarter of a million words (Table 1). We compared them against an equivalent English native corpus (LOCNESS). We used the software WordSmith (Scott 2004) to extract nearly 2000 usable concordances.

Table 1: Corpora details

<table>
<thead>
<tr>
<th>Learner corpus</th>
<th>Words</th>
<th>Native corpus</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICLE-Spanish</td>
<td>200,376</td>
<td>LOCNESS USarg</td>
<td>288,177</td>
</tr>
<tr>
<td>WriCLE</td>
<td>63,836</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total no. of words:</strong></td>
<td><strong>264,212</strong></td>
<td><strong>288,177</strong></td>
<td></td>
</tr>
</tbody>
</table>

4.5. Results and discussion

Figure 1 shows the production of postverbal Ss with unergatives and unaccusatives (H1: lexicon-syntax). Neither the natives nor the learners produced VS structures with unergatives like *cough, cry, shout*. By contrast, a proportion of postverbal Ss were produced with unaccusatives like *exist, appear,*
arise: 2.3% in the English native corpus and 7.1% in the learner corpora. This difference is small but statistically significant (χ² = 19.7, p < 0.01), but the crucial fact is that both groups produce postverbal Ss only with unaccusatives, and never with unergatives. This finding clearly replicates previous research and clearly confirms H1 in (14a), as predicted by the Unaccusative Hypothesis.

Figure 1: SV and VS produced with intransitives

The most frequent structure for the Spanish group was ungrammatical insertion of expletive *it*, followed by ungrammatical zero (Ø) insertion, (15a,b). By contrast, English natives produced grammatical structures, mostly loco-temporal XP-insertion and there-insertion, (16a,b). The conclusion is that both natives and learners behave alike in the sense that they produced VS only with unaccusatives, but our learners produce mostly ungrammatical structures (64%) and there is overproduction of VS (with a particular overproduction bias for the unaccusative V exist).

(15) a. *In the name of religion it had occurred many important events.
   b. *...because Ø exist the science technology and the industrialisation.

   b. Certainly there exists a demand for this work to be done.

Regarding H2 (syntax-phonology) in (14b), postverbal Ss are more frequently heavy than light in both corpora (Figure 2). Importantly, learners’ production rates do not significantly differ from natives’ (p = 0.65). In other words, both natives and learners tend to produce VS order with unaccusative Vs when S is heavy, yet SV order when S is light. This finding supports H2, as predicted by the End-Weight Principle.
Consider now H3 (syntax-discourse) in (14c). Postverbal Ss with unaccusative Vs are overwhelmingly focus (that is, new or relatively unfamiliar information), 100% in the native corpus and around 98% in the Spanish corpus (Figure 3). By contrast, preverbal Ss with unaccusative Vs tend to be topic (that is, old information), around 83% for the natives and 89% for the learners. Importantly, learners’ production rates do not significantly differ from natives’ ($\chi^2=1.49$, $p=0.22$). These results confirm H3: learners produce VS structures with unaccusative Vs when S is focus, yet SV structures are produced when S is topic, as dictated by the End-Focus Principle.
A final question is whether all these results are a result of L1 influence, or rather, whether they are universal. In previous work, we have compared L1 Italian vs. L1 Spanish learners of L2 English (Lozano & Mendikoetxea 2008). Both groups behaved identically and followed the same three general principles discussed above, as English natives do, though learners produce ungrammatical structures. We also have additional (but still preliminary) data from L1 French learners of L2 English who behave similarly to the Italian and Spanish group. Then, learners’ behavior seems to stem from general and developmental principles.

5. Conclusion

While information status notions such as topic and focus are universal, different languages realize them differently via syntax, morphology, intonation or via a combination of these (Vallduví 1993, Vallduví & Engdahl 1996). Our experimental results for adult L1 English learners of L2 Spanish suggest that learners are sensitive to information status but are unable to encode it syntactically with the pragmatically most adequate order, so they choose optionally both orders.

Regarding our corpus studies, both Spanish and Italian learners of L2 English appear to be sensitive to discourse status (and weight effects) but show persistent problems in the syntactic encoding of the construction and overuse the construction (particularly, *it*-insertion structures). It can then be concluded that L2 learners deficits are syntactic (and not discursive) in nature, as they show persistent problems with the syntactic encoding of information status, yet they are aware of the discursive principles governing information packaging.

Our studies also support a substantial body of research that considers the Unaccusative Hypothesis to be psychologically real in L2: L2 learners are aware of the argument structure distinction between unaccusative and unergative Vs and use this as a guiding principle to construct L2 mental grammars. But unaccusativity is a necessary but not a sufficient condition for the acceptability and/or the production of VS structures in non-native English. Crucially, properties operating at both the syntax-discourse and the syntax-phonology interfaces, relevant for a variety of word order phenomena in native English, also play a crucial role in constituent ordering in L2 English.

References


