**AIM OF THE PRESENTATION**

- To show that:
  - Learners of Romance (L1 English/Greek – L2 Spanish) as well as Romance learners of English (L1 Spanish/Italian – L2 English) are sensitive to discourse status (focus) in L2 subject-inversion structures, **BUT**: they show residual (yet persistent) problems when encoding information status syntactically.
  - Experimental and corpus data support claim that deficits at syntax-discourse interface are **syntactic** in nature.
    - Experimental data of **L2 Spanish** (Lozano 2006a, 2006b),
    - Corpus analysis studies in **L2 English** (Lozano & Mendikoetxea, 2007, 2008, in preparation)

- A full account of word order acquisition needs to take into account properties at three interfaces: **lexicon-syntax, syntax-discourse and syntax-phonology**.

---

**Word Order L1 Spanish/Italian (1)**

**Lexicon-syntax: Unaccusative Hypothesis**

- Neutral (non-focus) contexts: Discourse-initial
  - a. **María gritó** (unacc)
  - b. #**Gritó María.** (‘María shouted’)

- Unaccusatives: VS

  - Greek: like Spa.
  - English: strictly SV (no surface syntactic effects)

---

**Word Order L1 Spanish/Italian (2)**

**Syntax-discourse**

- Postverbal subjects are produced freely with all verb classes
  - a. Ha **telefoneado María al presidente.** (transitive), Has phoned Mary the president
  - b. Ha **hablado Juan.** (unergative)
  - c. Ha **llegado Juan** (unaccusative)

- Inversion as ‘**focalisation**’:
  - preverbal subjects are **topics** (given information)
  - and postverbal subjects are **focus** (new information) (Belletti 2001, 2004, Zubizarreta 1998)

- Neutral (non-focus) contexts: Discourse-final
  - a. **María llegó** (unacc)
  - b. #**Llegó María.** (‘María arrived.’)

- Inversion as ‘focus’:
  - a. **¿Quién ha llegado/hablado?** (Sp) Who has arrived/spoken?
  - b. #**Juan ha llegado/hablado**
  - c. #**Quién ha llegado/hablado Juan?** (Sp)
  - d. #**Gianni ha arrivate/parlato** (It)
  - e. #**Gianni è arrivato/parlato Gianni**
**Syntax-discourse: Narrow focus**

Greek & Eng: pres focus in situ: SV (pres. focus subject checked in Spec, TP)

**Empirical evidence: Spanish/Italian natives**

- **Empirical studies** on Spanish native speakers show that verb choice may determine word order (Hertel 2003, Lozano 2003, 2006a, b; see also Pinto 1999):
  - **Neutral**: [lexicon-syntax]
    - SV (unerg) – VS (unacc)
  - **Narrow focus contexts**: [syntax-discourse]
    - VS (unerg, unacc)

**Previous findings**

- **Lexicon-Syntax** (Unacc. Hypothesis):
  - English-speaking learners of Spanish are sensitive to the syntactic effects of the Unacc. Hypothesis from early stages of development: SV with unergatives but VS with unaccusatives (De Miguel, 1993; Hertel, 2000, 2003; Hertel & Pérez-Leroux, 1999).

- **Syntax-Discourse**: under-researched area:
  - Hertel (2003): presentationally focused subjects in final position are acquired late in L2 Spa. Same finding for L2 Italian (Belletti & Leonini, 2004).
  - Distinct word orders to mark focus is acquired late or perhaps never acquired in native-like fashion.
Method (Lozano 2006a, 2006b)

<table>
<thead>
<tr>
<th>Groups</th>
<th>Spanish native controls</th>
<th>L1 Greek</th>
<th>L2 Spa</th>
<th>L1 English</th>
<th>L2 Spa</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Proficiency   | Mean = 92% (range 80% - 100%) | Mean = 90% (range 80% - 100%) |

Advanced proficiency (Univ Wisconsin College Placement Test)

- Instrument:
  - Contextualised acceptability judgement test (Hertel, 2000)

  Tú estás en una fiesta con tu amiga Laura. Laura sale de la habitación y en ese momento llega la policía porque hay mucho ruido en la fiesta. Cuando Laura vuelve, te pregunta: ¿Quién llegó? Tú contestas:

(a) La policía llegó.   – 2   – 1   0   +1   +2
(b) Llegó la policía.   – 2   – 1   0   +1   +2

Translation: “You are at a party with your friend Laura. Laura leaves the room and at that moment the police arrive because the party is too noisy. When Laura comes back, she asks you: ‘Who arrived?’ You answer:

(a) The police arrived  – 2   – 1   0   +1   +2
(b) Arrived the police  – 2   – 1   0   +1   +2

Results: neutral contexts (Unaccusative Hypothesis)

Unergatives (SV) Unaccusatives (VS)

<table>
<thead>
<tr>
<th>0</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>Greek</td>
<td>English</td>
<td>Spanish</td>
<td>Greek</td>
<td>English</td>
</tr>
<tr>
<td>sig</td>
<td>sig</td>
<td>sig</td>
<td>[graph]</td>
<td>[graph]</td>
<td>[graph]</td>
</tr>
</tbody>
</table>

Convergence with natives (native-like knowledge)

Results: focused contexts (Narrow focus at syntax-discourse interface)

¿Quién gritó / llegó? “Who shouted / arrived?”

Unergatives (VS) Unaccusatives (VS)

<table>
<thead>
<tr>
<th>0</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>Greek</td>
<td>English</td>
<td>Spanish</td>
<td>Greek</td>
<td>English</td>
</tr>
<tr>
<td>sig</td>
<td>n.s.</td>
<td>n.s.</td>
<td>[graph]</td>
<td>[graph]</td>
<td>[graph]</td>
</tr>
</tbody>
</table>

Divergence with natives (subtype: optionality)

Conclusion of Experimental Study

- Features at the syntax-discourse interface are persistently problematic for advanced learners, e.g. work by Sorace, Tsimpli, Montrul → pragmatic deficit.

- Lozano (2006a, 2006b):
  - There is no reasons to think that learners are not sensitive to the topic/focus distinction, as it is present in L1 (in fact, it is universal: Vallduví 1993, 1995, Vallduví & Engdahl 1996).
  - Learners are sensitive to discourse status but are unable to encode it syntactically with the pragmatically most adequate word order → syntactic deficit

- In line with findings by Dominguez and Arche (2008):
  - Availability of optional forms can be accounted for by a purely syntactic deficit, probably due to apparently ambiguous input occurs
**CORPUS STUDY #1**

**V-S structures in:**

English natives  
L1 Spa – L2 Eng

Lozano & Mendikoetxea (2007, in prep.)

---

**RESEARCH QUESTIONS**

- **Main question:**
  - What are the conditions governing the production of VS structures in L2 English?
  - Do learners of English produce inverted subjects (VS) under the same conditions as English natives do, regardless of problems to do with syntactic encoding (grammaticality)?

---

**Word Order in native English (1)**

- **Fixed SV(O) order - Restricted use of postverbal subjects:**
  - **(a) XP V S (Inversion structures with an opening adverbial)**
    1. [On one long wall] hung a row of Van Goghs. [FICT]
    2. [Then] came the turning point of the match. [NEWS]
    3. [Within the general waste type shown in these figures] exists a wide variation. [ACAD]
    
    **[Biber et al. 1999: 912-3]**
  - **(b) There-constructions**
    1. Somewhere deep inside [there] arose a desperate hope that he would embrace her. [FICT]
    2. In all such relations [there] exists a set of mutual obligations in the instrumental and economic fields [ACAD]
    3. [There] came a roar of pure delight as... [FICT]

    **[Biber et al. 1999: 945]**

---

**Word order in native English (VS order)**

- **Lexicon-syntax interface** (Levin & Rappaport-Hovav, etc):
  - Unaccusative Hypothesis (Burzio 1986, etc) [existence and appearance]
  - (8) "There sang four girls at the opera. [unergative verb]
  - (9) There arrived four girls at the station. [unaccusative verb]

- **Syntax-discourse interface** (Biber et al, Birner 1994, etc):
  - Postverbal material tends to be focus/relatively unfamiliar information, while preverbal material links S to previous discourse: Principle of End-Focus.
  - (10) We have complimentary soft drinks and coffee. Also complimentary is red and white wine.

- **Syntax-Phonological Form (PF) interface** (Arnold et al 2000, etc)
  - Heavy material is sentence-final (Principle of End-Weight, Quirk et al. 1972) – general processing mechanism (reducing processing burden)
  - (11) One Sunday morning the warm sun came up and - pop! out of the egg came a tiny and very hungry caterpillar. [Eric Carle, The very hungry caterpillar, London: Penguin]

Subjects which are focus, long and complex tend to occur postverbally in those structures which allow them (unaccusative VS).
Word Order in native Spa / Ital (VS order)

- Lexicon-syntax interface
  No restrictions: postverbal subjects occur with all verb classes

- Syntax-discourse interface
  Postverbal subjects in Spanish and Italian are focus

- Syntax-Phonological Form (PF) interface
  Heavy subjects show a tendency to be postposed – a universal language processing mechanism: placing complex elements at the end of a sentence reduces the processing burden (J. Hawkins 1994).

Subjects which are focus, long, and complex tend to occur postverbally, with no restrictions at the lexicon-syntax interface.

---

The phenomenon in SLA


- Only with unaccusative verbs (never with unergatives).
  - Unaccusatives: arrive, happen, exist, come, appear, live...
  - Unergatives: cry, speak, sing, walk ...

- L1 Spanish/Italian/Arabic — L2 English:
  12) ...it arrived the day of his departure...
  13) And then at last comes the great day.
  14) In every country exist criminals
  15) ...after a few minutes arrive the girlfriend with his family too.

- Explanation: lexicon-syntax interface (Psychological reality of the Unaccusative Hypothesis)

CRUCIAL DIFFERENCE: These previous studies focused on ERRORS, thus emphasising the differences between native and non-native structures. By contrast, our study emphasises the similarities between native and non-native structures.

---

Hypotheses

GENERAL HYPOTHESIS:
Conditions licensing VS in L2 Eng are the same as those in native Eng, DESPITE differences in syntactic encoding.

- H1 [LEXICON]: Lexicon-syntax interface:
  Postverbal subjects with unaccusatives (never with unergatives)

- H2 [WEIGHT]: Syntax-PF interface:
  Postverbal subjects: heavy (but preverbal light)

- H3 [FOCUS]: Syntax-Discourse interface:
  Postverbal subjects: focus (but preverbal topic)

---

METHOD (1)

Based on Levin (1993) and Levin & Rappaport-Hovav (1995):

<table>
<thead>
<tr>
<th>UNACCUSATIVES</th>
<th>VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXISTENCE</td>
<td></td>
</tr>
<tr>
<td>exist</td>
<td></td>
</tr>
<tr>
<td>flow</td>
<td></td>
</tr>
<tr>
<td>grow</td>
<td></td>
</tr>
<tr>
<td>hide</td>
<td></td>
</tr>
<tr>
<td>live</td>
<td></td>
</tr>
<tr>
<td>remain</td>
<td></td>
</tr>
<tr>
<td>rise</td>
<td></td>
</tr>
<tr>
<td>settle</td>
<td></td>
</tr>
<tr>
<td>spread</td>
<td></td>
</tr>
<tr>
<td>survive</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UNACCUSATIVES</th>
<th>VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMERGENCE</td>
<td></td>
</tr>
<tr>
<td>appear</td>
<td></td>
</tr>
<tr>
<td>arise</td>
<td></td>
</tr>
<tr>
<td>awake</td>
<td></td>
</tr>
<tr>
<td>begin</td>
<td></td>
</tr>
<tr>
<td>break</td>
<td></td>
</tr>
<tr>
<td>develop</td>
<td></td>
</tr>
<tr>
<td>emerge</td>
<td></td>
</tr>
<tr>
<td>flow</td>
<td></td>
</tr>
<tr>
<td>follow</td>
<td></td>
</tr>
<tr>
<td>happen</td>
<td></td>
</tr>
<tr>
<td>occur</td>
<td></td>
</tr>
<tr>
<td>rise</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UNACCUSATIVES</th>
<th>VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISAPPEARANCE</td>
<td></td>
</tr>
<tr>
<td>die</td>
<td></td>
</tr>
<tr>
<td>disappear</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UNACCUSATIVES</th>
<th>VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>INHERENTLY DIRECTED MOTION</td>
<td></td>
</tr>
<tr>
<td>arrive</td>
<td></td>
</tr>
<tr>
<td>come</td>
<td></td>
</tr>
<tr>
<td>drop</td>
<td></td>
</tr>
<tr>
<td>enter</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UNACCUSATIVES</th>
<th>VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>BODILY PROCESSES</td>
<td></td>
</tr>
<tr>
<td>breathe</td>
<td></td>
</tr>
<tr>
<td>cough</td>
<td></td>
</tr>
<tr>
<td>cry (*)</td>
<td></td>
</tr>
<tr>
<td>sweat (**)</td>
<td></td>
</tr>
</tbody>
</table>
**Corpora**

- **Corpora:**
  - L1 Spa – L2 Eng (2 corpora)
  - Eng natives

<table>
<thead>
<tr>
<th>Learner corpora</th>
<th>Native corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words</td>
<td>LOCNESS USarg</td>
</tr>
<tr>
<td>LOCNESS USmixed</td>
<td>18,826</td>
</tr>
<tr>
<td>LOCNESS Alevels</td>
<td>60,209</td>
</tr>
<tr>
<td>LOCNESS BRsur</td>
<td>68,586</td>
</tr>
</tbody>
</table>

| Total no. of words | 288,177 |

**Query software:** WordSmith v. 4.0 (Scott 2004)

---

**H1: Unaccusative: grammatical vs ungrammatical VS**

**GRAMM.**
- There-insertion:
  - Natives: there exists a demand for this work to be done...
  - Learners: There exist positive means of earning money.
- AdvP-insertion:
  - Natives: Thus began the campaign to educate the public...
  - Learners: ...and here emerges the problem.
- Locative inversion:
  - Natives: [no production]
  - Learners: In the main plot appear the main characters: Volpone and Mosca.

**UNGRAM.**
- *it*-insertion:
  - Learners: *In the name of religion it had occurred some important events.
- *Ø*-insertion:
  - Learners: *because exist the science technology and the industrialisation.
- *XP*-insertion:
  - Learners: *In 1760 occurs the restoration of Charles II in England.

---

**Table 1: Corpora details**

<table>
<thead>
<tr>
<th>Learner corpora</th>
<th>Native corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words ICLE-Spanish</td>
<td>LOCNESS USarg</td>
</tr>
<tr>
<td>WriCLE</td>
<td>149,574</td>
</tr>
<tr>
<td>LOCNESS USmixed</td>
<td>18,826</td>
</tr>
<tr>
<td>LOCNESS Alevels</td>
<td>60,209</td>
</tr>
<tr>
<td>LOCNESS BRsur</td>
<td>68,586</td>
</tr>
</tbody>
</table>

| Total no. of words | 264,212 |

**H1 results: syntax-lexicon**

**Table 1: Frequency of postverbal subjects produced**

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Verb type</th>
<th>Postverbal subjects</th>
<th>Usable concordances</th>
<th>% frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner</td>
<td>Unerg</td>
<td>0</td>
<td>181</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Unac</td>
<td>58</td>
<td>820</td>
<td>7.1%</td>
</tr>
<tr>
<td>Native</td>
<td>Unerg</td>
<td>0</td>
<td>185</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>Unac</td>
<td>16</td>
<td>719</td>
<td>2.2%</td>
</tr>
</tbody>
</table>

---

**H1: Unaccusative: grammatical vs ungrammatical VS**

**According to structure type**

- According to structure type

---

**Table 1: Frequency of preverbal material**

<table>
<thead>
<tr>
<th>Type of preverbal material</th>
<th>Learners</th>
<th>Natives</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>it</em>-insertion</td>
<td>41.4%</td>
<td>36.2%</td>
</tr>
<tr>
<td>Locative inversion</td>
<td>13.3%</td>
<td>15.6%</td>
</tr>
<tr>
<td>XP-insertion</td>
<td>10.6%</td>
<td>10.4%</td>
</tr>
<tr>
<td>There-insertion</td>
<td>31.4%</td>
<td>32.1%</td>
</tr>
<tr>
<td>AdvP-insertion</td>
<td>13.8%</td>
<td>13.6%</td>
</tr>
<tr>
<td>Ø-insertion</td>
<td>6.6%</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

---

**Graphs and charts**

- Bar charts for frequency of production
- Pie charts for comparison between learners and natives
H2 results: syntax-phonology

**Examples H2:** syntax-phonology

**SV:** typically LIGHT (Pronoun, D + N)

Learners: ...but they may appear everywhere.

...since the day eventually came...

Natives: These debates began over two decades ago.

...a great controversy exists over the topic.

**VS:** typically HEAVY (postmodification)

Learners: Against this society drama emerged an opposition headed by Oscar Wilde and Bernard Shaw.

...exists yet in Spain a group of people who are supposed to be professional soldiers.

Natives: With this theory also came the area of quantum mechanics.

Thus began the campaign to educate the public on how one contracts aids.

H3: syntax-discourse

**Examples H3:** syntax-discourse

**VS:** FOCUS

Learners: ...there also exists a wide variety of optional channels which have to be paid.

Natives: With this theory also came the area of quantum mechanics.

**SV:** typically TOPIC

Learners: I use the Internet ... I find windows ... if they press on any of these windows ... these windows cannot appear because a child could enter easily...

Natives: However, Hugo is not prepared ... Louis took such exception to Hugo ... Hugo came from a bourgeois family.
Conclusion of corpus study (1)

Unacc Focus Heavy

Interfaces:
- Lexicon-syntax
- Syntax-discourse
- Syntax-phonology

Conclusion of corpus study (2)

- These results confirm that Spanish L2 learners of English produce postverbal subjects under exactly the same 3 interface conditions as in L1 English (unaccusativity being a necessary but not a sufficient condition).
  - Unaccusativity Hypothesis: postverbal subjects appear with unaccs.
  - End-weight principle: postverbal subjects tend to be long and complex.
  - End-focus principle: postverbal subjects tend to be focus.

- So, learners do not show a pragmatic deficit at the syntax-discourse interface.

- Learners show rather a persistent problems in the syntactic encoding of the construction
  - High production of ungrammatical examples (it-insertion, Ø-insertion, wrong XP).
  - Spanish learners overuse the construction and show a lexical bias for the V exist.

Example
* ... it will not exist a machine or something able to imitate the human imagination.

General conclusions (1)

- **L1 English-L2 Spanish**: learners are sensitive to discourse status (Topic, Focus) BUT are unable to encode it syntactically with the pragmatically most adequate word order.

- **L1 Spanish – L2 English**: learners are sensitive to discourse status (and weight effects) but show persistent problems in the syntactic encoding of the construction and overuse the construction.
Our studies:
- Support substantial body of research: Unaccusative Hypothesis to be psychologically real in L2. BUT
- Unaccusativity is a necessary but not a sufficient condition for the acceptability and/or production of VS structures.
- Properties operating at both the syntax-discourse and the syntax-phonology interfaces, relevant for a variety of word order phenomena in native English (see e.g., Arnold et al. 2000), also play a crucial role in constituent ordering in L2 English.

Language specific or universal pattern?
- L1 Italian – L2 Eng vs L1 Spa – L2 Eng (Lozano & Mendikoetxea 2008): same results.
- Unaccusativity Hypothesis: postverbal subjects appear with unacs.
- End-weight principle: postverbal subjects tend to be long and complex.
- End-focus principle: postverbal subjects tend to be focus.
- Also: evidence from L1 French – L2 Eng (unpublished results yet).

In short: syntactic (not pragmatic) deficit in L2.

REFERENCES (1)

REFERENCES (2)
- Pinto, M. (1997), Licensing and Interpretation of Inverted Subjects in Italian, Utrecht: LEID.
Corpora used in the study

- **ICLE**: *International Corpus of Learner English*

- **LOCNESS**: *Louvain Corpus of native English Essays*, UCL/CECL, Louvain-la Neuve

- **WriCLE**: *Written Corpus of Learner English*; Universidad Autónoma de Madrid (Rollinson, O'Donnell, Mendikoetxea, in progress)
  [http://www.uam.es/woslac](http://www.uam.es/woslac)