A Grammar of Spanish in HPSG

Luis Pineda & Ivan Meza
Department of Computer Science
IIMAS, UNAM, México.
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with se
  - Passive reflexive (*pasiva refleja*)
  - Active impersonal (* impersonal activa*)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with *se*
  - Passive reflexive (*pasiva refleja*)
  - Active impersonal (*impersonal activa*)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions
Antecedents

- The DIME Project (Diálogos Inteligentes Multimodales en Español)
- A Collaboration with James Allen through the NSF/Conacyt bilateral program for research in computer science
- Objective: To build a task oriented conversational system in Spanish in a simple design domain (Kitchen design)
- Main tasks:
  - The DIME Corpus
  - Dimex (Speech recognition system in mexican Spanish)
  - A Spanish grammar and parser
  - A Multimodal dialogue manager

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The development process

1999, 2000

Preliminary ideas and problems:
clitics and word-order

Proclitics are words
Word order: slash cat.

Spanish Syntax:
Gili Gaya, 1973

Allen’s NLP:
TRAINS & TRIPS

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The development process

Preliminary ideas and problems: clitics and word-order

Proclitics are words
Word order: slash cat.

Initial work with HPSG & LKB

1999, 2000

2001

Spanish Syntax: Gili Gaya, 1973

Allen’s NLP: TRAINS & TRIPS

HPSG: Sag & Wasow, 1999

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The development process

1999, 2000

Preliminary ideas and problems:
oclitics and word-order

Proclitics are words
Word order: slash cat.

2001

Auxiliaries &
Preliminary clitic model

Spanish Syntax: Gili Gaya, 1973

Allen’s NLP: TRAINS & TRIPS

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The development process

- Spanish Syntax: Gili Gaya, 1973
- Allen’s NLP: TRAINS & TRIPS
- Preliminary ideas and problems: clitics and word-order
- Proclitics are words
  Word order: slash cat.
- Auxiliaries & Preliminary clitic model
- Miller & Sag, 1995
  Monachesi, 1993

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The development process

First clitic model:
The clitic word and Head-Proclitic Rule

Periphrasis with subject raising and control verbs

Auxiliaries & Preliminary clitic model

Gili Gaya, 1973
Sag & Wasow, 1999

2001

2002

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The development process

2001

Auxiliaries & Head-Proclitic Rule

Periphrasis with subject raising and control verbs

2002

First clitic model: The clitic word and Head-Proclitic Rule

Gili Gaya, 1973

Sag & Wasow, 1999

Periphrasis with object control verbs and clitic operations

2003

Current grammar and implementation

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The development process

2001
Auxiliaries & Head-Proclitic Rule

Periphrasis with subject raising and control verbs

2002
First clitic model: The clitic word and Head-Proclitic Rule

Periphrasis with object control verbs and clitic operations

2003
Current grammar and implementation

Ohio-State: Jirka Hana & Detmar Meurers
The Structure of the grammar

- Lexicon & Basic Grammatical Combination schemes
- Verbal system
- Clitic system
- Other phenomena
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with se
  - Passive reflexive (pasiva refleja)
  - Active impersonal (impersonal activa)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The Spanish Periphrasis

- General notion of auxiliaries
- Copulative *ser* & *estar* (to be)
  - Attributive sentences
  - The passive sentence
- Auxiliaries in predicative sentences
- Other periphrasis
The Spanish Periphrasis

- General notion of auxiliaries
- Copulative *ser & estar* (to be)
  - Attributive sentences
  - The passive sentence
- Auxiliaries in predicative sentences
- Other periphrasis
Auxiliaries

- Traditional grammar:
  - Auxiliaries: verbs that appear in periphrasis
  - Periphrasis: constructions with auxiliary verbs!

- General taxonomy:
  - “Real” auxiliaries: *ser*, *estar* (to be) and *haber* (have)
  - Verbal phrases
  - Semi-auxiliaries: modals

- Little agreement on *haber*:
  - Part of the composite conjugation: The link between *haber* and the participle cannot be broken
**haber**

- Part of the complex conjugation:
  - *he comido*  
    (I have eaten)
  - *habré caminado*  
    (I will have waked)

- In perihrasis with other verbs the participle agrees with the direct object:
  - *tengo leídos los libros*  
    (I have read the books)
  - *tengo leídas las cartas*  
    (I have read the letters)

- but:
  - *he leído los libros*  
    (I have read the books)
  - *he leído las cartas*  
    (I have read the letters)

- **haber** is an inflexion, and the only dummy verb!
Auxiliaries: General notions

- Do occur conjugated in periphrasis
- Have a grammatical function (i.e. have lost their lexical meaning in the context)
- Express temporal and aspectual information
- Have no agent but help to determine the agent of verbs of content in the periphrasis
- But, there are periphrasis without auxiliaries...
Auxiliaries: a formal notion

- Auxiliaries are subject-raising verbs

$\begin{bmatrix}
\text{srv} - \text{lxm} \\
\text{ARG} - \text{ST}
\end{bmatrix}
\begin{cases}
\#1, \\
\begin{bmatrix}
\text{phrase} \\
\text{SPR} \langle \#1 \rangle
\end{bmatrix}
\end{cases}$

- Have a grammatical function: have lost their lexical meaning in the periphrasis

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The Spanish Periphrasis

- General notion of auxiliaries
- Copulative *ser* & *estar* (to be)
  - Attributive sentences
  - The passive sentence
- Auxiliaries in predicative sentences
- Other periphrasis
The Spanish Periphrasis

- General notion of auxiliaries
- Copulative *ser* & *estar* (to be)
  - Attributive sentences
  - The passive sentence
- Auxiliaries in predicative sentences
- Other periphrasis
Two copulatives: ser & estar

- **Ser**: permanent attributions
  - *El cielo es infinito* (the sky is infinite)
  - *El cielo está infinito*

- **Estar**: contingent attributions
  - *La fruta está madura* (The fruit is mature)
  - *La fruta es madura*
Two copulatives: ser & estar

- Properties that can be permanent and contingent
  - *El niño es alto* (the boy is tall)
  - *El niño está alto*

- *Estar*: permanent attributions!
  - *Juan está muerto* (Juan is dead)

- *Ser*: cannot go with a permanent attribution!!
  - *Juan es muerto*
Aspectual character

• Aspect of copulatives
  – *ser*: imperfective
  – *estar*: perfective

• The aspect of the attribute must agree with the aspect of the copulative
  – Aspect of attributes?
  – *ser*: things that are known
  – *estar*: things that need to be experienced
Two copulatives: ser & estar

- Properties that can be permanent and contingent
  - *El niño es alto* (The state of being tall)
  - *El niño está alto* (The result of the process of growing up)

- *Estar*: permanent attributions!
  - *Juan está muerto* (a perfective state, resulting from the
    imperfective state of being alive)

- To know versus to experience
  - *El niño es alto* (I know that the boy is tall)
  - *El niño está alto* (I can see that the boy is now tall)
Two copulatives: *ser* & *estar*

- Complementary sense of copulatives:
  - The process of *ser* becomes into the state of *estar*

- Perfect tenses of the imperfective *ser* correspond to the imperfective tenses of the perfective *estar*
  - *La casa está construida* (the house is built) when
  - *La casa ha sido construida* (The house has been built)
The HPSG Model

- *Ser* and *estar* are subject raising verbs with an aspectual attribute
- Attributes have an aspectual feature:
  - Always perfective
  - Always imperfective
  - Underdetermined
The HPSG Model

- *Ser* and *estar* are subject raising verbs with an aspectual attribute

- Attributes have an aspectual feature:
  - Always perfective
  - Always imperfective
  - Underdetermined
The HPSG Model

• Lexical entry for *alto*

\[
\langle \text{alto}, \left[ \begin{array}{c}
\text{adj - lxm} \\
\text{SYN} \\
\text{HEAD} \\
\text{AGR} \\
\text{SPR} \\
\langle \#1 \rangle \\
\text{sing - masc}
\end{array} \right] \rangle
\]
The HPSG Model

• Structure of attributive sentence:

```
#1[NP]

[SPR
[COMPS

Juan

es

alto

[SPR
[COMPS

#2

[HEAD AGR sing-masc

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
```
The HPSG Model

- Semantics of the attributive with *ser*:

$$\begin{align*}
\text{INDEX} & \quad i \\
\text{MODE} & \quad \text{prop} \\
\text{RESTR} & \quad \begin{cases} 
\text{RELN} & \quad \text{name} \\
\text{SIT} & \quad u \\
\text{NAME} & \quad \text{juan} \\
\text{NAMED} & \quad i \\
\text{THM} & \quad i \\
\text{TYPE} & \quad \text{ess} \\
\end{cases}
\end{align*}$$

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The Spanish Periphrasis

- General notion of auxiliaries
- Copulative *ser* & *estar* (to be)
  - Attributive sentences
  - The passive sentence
- Auxiliaries in predicative sentences
- Other periphrasis
The passive sentence

- The aspectual value of the attribute must agree with aspectual value of the passive predication:
  - María es querida  (Maria is loved)
  - María está querida
  - La puerta está abierta  (the door is opened)
  - La puerta es abierta
  - La puerta es abierta todas las mañanas
Passive sentence

- Lexical rule for the passive construction

\[
\langle #2, \left[ \text{ARG-ST} \left\langle \text{NP} \quad \text{NP}_j \quad \text{NP}_k \right\rangle \right] \Rightarrow \left\langle F_{\text{pass}}(#2), \left[ \text{SYN|HEAD} \left[ \text{FORM} \quad \text{pass} \right] \right] \right. \\
\left. \quad \left[ \text{ARG-ST} \left\langle \text{NP}_j \quad \text{NP}_k \right\rangle \right] \right\rangle
\]
Structure of the passive sentence

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The Spanish Periphrasis

- General notion of auxiliaries
- Copulative *ser* & *estar* (to be)
  - Attributive sentences
  - The passive sentence
- Auxiliaries in predicative sentences
- Other periphrasis
Auxiliaries: a formal notion

- Auxiliaries are subject-raising verbs

\[
\begin{bmatrix}
    srv - lxm \\
    ARG - ST \\
    #1, \\
    phrase \\
    SPR \\
    \langle #1 \rangle
\end{bmatrix}
\]

- Have a grammatical function: have lost their lexical meaning in the periphrasis

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Five test for subject-raising

1. Appear in periphrasis
   \( \text{puedes mostrar el catalogo a Juan} \)

2. Its complement can be an impersonal verb
   \( \text{puede llover} \)
   \( \text{tiene llovido} \) *

3. To ask for the direct object both verbs are required as it is a single predication
   \( \text{puedes mostrar el catalogo} \)
   \( \text{¿qué puedes mostrar?} \Rightarrow \text{el catalogo} \)
   \( \text{¿qué puedes?} \) *
Five test for subject-raising

4. The dummy *haber* is required to ask for the action:

   *puedes mostrar el catálogo*

   ¿qué puedes hacer? ⇒ *mostrar el catálogo*

   ¿qué puedes? *

5. The content verb can appear in passive voice but the auxiliary verb cannot (it is not agentive)!

   *El catálogo puede ser mostrado*

   *El catálogo es podido ser mostrado* *
Our list of auxiliares

- Followed by an infinitive:
  - *Poder*  *puedes comer* (possibility)
  - *Ir a*  *vas a comer* (incoative)
  - *Venir a*  *vienes a comer* (aproximative, terminative)
  - *Volver a*  *vuelves a comer* (reiterative)
  - *Haber de*  *habrás de comer* (obligative)
  - *Tener que*  *tienes que comer* (obligative)
  - *Deber de*  *debes de comer* (hypothetical)
  - *Llegar a*  *llegas a comer* (approximative)
  - *Acabar de*  *acabas de comer* (perfective)
  - *Alcanzar a*  *alcanzas a comer* (perfective)
Our list of auxiliares

- Followed by gerund (all durative):
  - **estar**  **estar comiendo**  (no locative meaning)
  - **Ir**  **vas comiendo**  (no eating meaning)
  - **Venir**  **vienes comiendo**  (no coming meaning)
  - **seguir**  **seguir comiendo**  (no following meaning)
  - **andar**  **andas comiendo**  (no walking meaning)

- Subject-raising:
  - **El FBI sigue visitando a Pedro**
  - **Pedro sigue siendo visitado por el FBI**
Our list of auxiliares

- Followed by participle:
  - copulative and passive
    - *ser*: *es querido* (is loved)
    - *estar*: *esta abierta* (is opened)
  - complex conjugation
    - *haber*: *has comido* (you have eaten)
The results of the test...

- Differs from traditional lists based on the notion of periphrasis and the lost of lexical meaning (i.e. Gily Gaya)
- Our additions:
  - _poder_ (can)
  - _haber_ (have)
- But we don’t include verbal phrases that although have periphrastic meaning require an agent and cannot appear with impersonal verbs:
  - _pasar a:_ *pasa a llover*
  - _echar a:_ *echa a llover*
  - _traer:_ *trae llovido*
  - _quedar:_ *queda llovido*
The Spanish Periphrasis

- General notion of auxiliaries
- Copulative *ser* & *estar* (to be)
  - Attributive sentences
  - The passive sentence
- Auxiliaries in predicative sentences
- Other periphrasis
Periphrasis without auxiliares

• Followed by an infinitive: the modals
  - Poder  puedes comer  (you are able to eat)
  - querer   quieres comer  (you want to eat)
  - deber    debes comer  (you want to eat)
  - osar     osas venir  (you dare to come)
  - soler    sueles venir  (you use to come)
  - pensar   piensas venir  (you think to come)
  - esperar  esperas venir  (you expect to come)
  - lograr   logras venir  (you manage to come)
  - saber    sabes venir  (you know how to come)

• Periphrasis with many verbs:
  - quiere poder comer  (he want to be able to eat)
Periphrasis without auxiliaries

- Followed by an infinitive:
  - Usually preserve their lexical meaning but can have periphrastic meanings as well
  - Subject-control verb:
    \[
    ARG-ST \leftarrow \begin{array}{c}
    NP_i, \left[ \text{phrase} \right]
    \end{array}
    \begin{array}{c}
    FORM \quad \text{inf}
    \end{array}
    \begin{array}{c}
    SPR \left\langle NP_i \right\rangle
    \end{array}
    \]
  
  - *Expect?*  In English this is an object-raising verb!
Periphrasis without auxiliaries

- Followed by an infinitive: Object-control verbs
  - The object of the verb is the subject of the infinitival verb

- Verbs of the senses:
  - ver: *ve caminar a María*        (sees María to walk)
  - oir: *oye cantar a los pajaros*  (hears to sing the birds)

- Verbs of will:
  - mandar: *manda marchar a los soldados*
    combains the soldiers to march
  - permitir: *permite jugar a los niños*
    *allows the children to play*
  - prohibir: *prohibe estacionar al coche*
    forbids the car to park
Periphrasis without auxiliaries

- Followed by an infinitive:
  - Preserve their lexical meaning
  - object-control verb (version 1):

\[
\begin{align*}
\text{ARG} - ST & \langle \langle \text{#1}, \text{#2}, \text{#3} \rangle \\
\text{COMPS} & \langle \langle \text{#2}, \text{#3} \rangle \\
\text{SPR} & \langle \langle \text{#1} \rangle \\
\text{ocv} - \text{lxm} & \end{align*}
\]
Periphrasis without auxiliaries

- Followed by gerund: Verbs of movement
  - *Ir*  *vas comiendo*  *(going meaning)*
  - *Venir*  *vienes comiendo*  *(coming meaning)*
  - *andar*  *andas comiendo*  *(walking meaning)*
  - *seguir*  *sigues corriendo a*  *(following meaning)*
Periphrasis without auxiliaries

- Followed by gerund: the stative verbs
  - *continuar*  *continua visitando*  (he stays visiting)
  - *vivir*  *vive jugando*  (lives playing)

- Followed by participle:
  - *llevar*:  *llevo andados muchos km*
    *I have walked many km*
  - *dejar*:  *dejaron encargado al niño*
    *they left the children in care*
Periphrasis

• Auxiliaries (subject-raising):
  – *ser*, *estar* and *have*
  – auxiliary verbal phrases

• Followed by infinitive:
  – Modals (imperfective): subject-control
  – Senses and will (perfective): object-control

• Followed by gerund (imperfective):
  – Verbs of movement: subject-raising
  – Stative verbs: subject-raising

• Followed by participle (imperfective):
  – Stative verbs: subject-raising
Content

• Antecedents
• The Spanish periphrasis
• The clitic system
• Reflexives and reciprocal
• Impersonal constructions with se
  – Passive reflexive (*pasiva refleja*)
  – Active impersonal (*impersonal activa*)
• Ergative constructions
• Pronominal questions
• Word order
• Implementation and verification
• Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The clitic system

- Clitic sentences
- Definition and properties of clitics
- The model
- The HPSG Model
Clitic sentences

- Basic enclitic and proclitic forms
  - muestra*lo  (show it to him)
  - *se lo muestras

- Clitic climbing:
  - puede mostrarse*lo  (he can show it to him)
  - *se lo puede mostrarse

- Clitic climbing with several verbs:
  - puede haber querido mostrarse*lo
  - *se lo puede haber querido mostrar
  - puede haberselo querido mostrar
  - he could have wanted to show it to him
Clitic sentences

• Object-control verbs:
  – ve a María comerla (he sees Maria eat it)
  – *ve a María la comer
  – *vela la comer
  – la ve comerla
  – *la la ve comer
  – *la a María ve comer
The ethical dative

• The very common se:
  – *Ve a María comersela la manzana
  – he sees Maria eat the apple for herself

• Other clitic forms:
  – ve a María comersela
  – la ve comersela
  – *se la a María ve comer
  – le pude ver poneselo

• Pleonastic forms:
  – a María se la ve comer                (Maria, he sees her eat it)
  – al niño la ve poneselo               (the boy, he sees Maria to put it on him)
  – la ve poneselo al niño

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The offending participle

• The participle does not like enclitics:
  – *ha visto a María puestoelo
  – he has seen Maria tu put herself it (the hat) on
  – se lo ha visto puesto
  – *ha vistoelo puesto

• With nominalized verbs:
  
  *El mostrando el cuadro yo → El mostrándolo yo
  
  *My showing of the picture → My showing of it

• With gerunds:
  
  *Mostrando el cuadro → mostrando
  
  *The showing of the picture → The showing of it

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The clitic system

- Clitic sentences
- Definition and properties of clitics
- The clitic system
- The HPSG Model
The clitics

- Untressed form (*átonos*)
  - Go without preposition
  - Represent direct, indirect and circumstancial complements

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td><em>me</em></td>
<td><em>nos</em></td>
</tr>
<tr>
<td>2nd</td>
<td><em>te</em></td>
<td><em>os</em></td>
</tr>
<tr>
<td>3rd</td>
<td><em>lo, la, le (se)</em></td>
<td><em>los, las, les (se)</em></td>
</tr>
<tr>
<td>neuter</td>
<td><em>lo</em></td>
<td></td>
</tr>
<tr>
<td>reflexive and impersonal</td>
<td><em>se</em></td>
<td><em>se</em></td>
</tr>
</tbody>
</table>
Properties of clitics

- Convey case information:
  - *lo, la, los, las*: accusative
  - *le, les, se*: dative

- Pleonastic use:
  - *a mi no me han visto* (they haven´t seen me)
  - *a ti estoy mirandote* (I´m looking at you)
  - *les prestaban los libros a los estudiantes*
  - *les prestaban los libros a ellos*
  - *a los estudiantes les prestaban los libros*
  - *they use to lend the books to the students*
Clitic order

- Impersonal forms (except participle): enclitic
  - Imperative: *muestramelom* (show it to me)
  - Gerund: *mostrandolombo*
    - mostrando el cuadro → mostrandolombo
    - *The showing the picture* → The showing of it
  - Infinitive: *mostrarlo*
    - *El mostrarel cuadro yo* → *El mostrarlo yo*
    - *My showing of the picture* → My showing of it

- Personal forms: proclitic
  - *me lo muestra* (she show it to me)
Clitic order

• In complex form, always proclitic (Gili Gaya):
  – In the complex conjugation:
    
    me lo ha mostrado  (he has show it to me)
    
    *ha mostradomelo

• But with infinitives, no problem:
  
  puede mostrar el catálogo a María
  
  puede mostrarselo
  
  se lo puede mostrar
  
  he shows the catalog to María
Clitic order

● With gerunds is also ok:

*El FBI continua visitando a Pedro*

*El FBI continua visitandolo*

*El FBI lo continua visitando*

*The FBI continues to visit Pedro*

● The restriction seems to apply to participles:

– *El FBI ha visitado a Pedro*

– *lo ha visitado*

– *ha visitadolo*

● Otherwise, there seems to be very little restrictions!
Clitic order

- The order is strict (with qualifications):
  - A verb can have one, two or even three clitic
  - They are all either enclitic of proclitic
    
    muestramelo (show it to me)
    me lo muestras
    *me muestras lo
  
- The order is the same in proclitic and enclitic:
  - se must be in front always
  - Second persons in front of the first always
  - Third persons (but se) are always last

- There are dialectal exceptions (Gili Gaya):
  - me se cae el sombro (the hat drops from I myself)
  - te se ve la oreja (The ear can be seen from you yourself)

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Properties of atonic clitics

- A lot of dialectal variation!
  - Spain: leísimo (le accusative), laísimo (la dative)
  - Latin American: leísimo (less than in Spain!)

- Personal versus reflexive se:
  - Personal se is a duplication of le
    From Latin illi: dio lle lo → dio ge lo → dio se lo
  - Reflexive: from reflexive Latin se:
    Can occur as direct and indirect object!
Clitics in Mexican Spanish

• The sequences used in Mexican Spanish (singular):
  – One word: “se” “te” “me” “le” “lo” “la”
  – Two words:
    “se te” “se me” “se le” “se lo” “se la”
    “te me” “te la” “te lo”
    “me le” “me lo” “me la”
  – Three words (quite artificial!):
    “te me lo”
  – Duplicating the entries for different cases:
    “le” “le”

• Adding plurals, combinations of singular and plurals and duplication for clitics with different cases does not increase the size of the list significantly (about 70 cases)!

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Status of clitics

● Phonological factors:
  – The verb is most *tensive* part of the sentence (Gili Gaya & Tomas Navarro Tomas):
    ● Proclitic: every thing in front
    ● Enclitic: every thing after
  – Proclitic acquire lexical independence: they are words!
  – Enclitic are attracted to the *tensive* part: *inflexions*
are proclitics words?

- The standard tests to tell inflexions from words
  - Inflexion attach to one kind:
    - *Juan es fiel a María* → *le es fiel* (host = adjective)
  - Scope over coordination:
    - *María comprará y vestirá el abrigo*
      → *lo comparará y vestirá*
    - *Maria comprará y pondrá las botas*
      → *se las comparará y pondrá*
Proclitic

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitic repair and enclitic

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Proclitics are *words*!

- **The Spanish intuition and spelling:**
  - Proclitic: *me lo dices*
  - Enclitic: *dímelo*

- **Pleonasm breacking in the middle of the verbal phrase:**
  - *ve a María poner el sombrero* (en la mesa)
  - he sees Maria put the hat (in the table)
  - *se lo ve poner*
  - *a María [se lo ve poner]_{vp}*
  - *[se lo ve a María poner]_{vp}*

- **Clitic repairs:**
  - “*me <sil> muestrame*”  (DIME corpus)
The clitic system

- Clitic sentences
- Definition and properties of clitics
- The Model
- The HPSG Model
The model

- Should account for clitic distribution
  - Proclitic and enclitic
  - Simple and composite forms:
    - Composite conjugation
    - Periphrastic constructions (with and without auxiliares)
    - Independently of a main verb:
      - *El decirlo tú* (Your saying of it)
      - *El general diciéndolo* (the general saying it)

- Should account for rigid order (with qualifications)

- Pleonastic use

- Subordinate constructions

- Should be flexible enough to accommodate dialectal variation
The model (version 1)

- Clitics sequences are words in the lexicon, including case information:
  - One clitic: “se” “te” “me” “le” “le” “lo” “la”
  - Two clitics:
    “se te” “se me” “se le” “se lo” “se la”
    “te me” “te la” “te lo”
    “me le” “me lo” “me la”
  - These are the only orders permitted!

- Dual behavior:
  - Proclitics: independent lexical units
  - Enclitics: inflexional affixes, but must be clitic words!

- Clitic constructions undergo ONE cancelation operation:
  - Proclitics: Head-Proclitic Rule (a full syntactic operation)
  - Enclitics: Lexical cancelation (through lexical rules)
Clitic insertion and cancelation

- Explicit arguments:
  - muestra el catálogo a María (he shows the catalog to Maria)

- Clitic forms:
  - muestraselo
  - se lo muestra

- Lexical Cancelation level-0 (LC-0):
  - muestraselo (through a lexical rule)
  - 1 clitic insertion: “selo” (it is a clitic word)
  - 1 Cancelation: the arguments are canceled by the same operation

- Syntactic cancelation (SC):
  - se lo muestra (Head-proclitic rule)
  - 1 insertion: the clitic word “se lo”
  - Cancelation postponed over one verb: muestra
Clitic climbing

- Explicit arguments:
  - *puede mostrar el catálogo a María*
  - he can show the catalog to Maria

- Clitic forms:
  - *puede mostrarselo* (LC-0)
  - *se lo puede mostrar* (SC)
  - *puedeselo mostrar* (LC-1)

- In all three cases one insertion and one cancelation!

- Clitic climbing:
  - The argument of *mostrar* seems to be attached to *poder*
  - Successfull interpretation depends on been able to establish the right bindings!
Clitic climbing

- Explicit arguments:
  - *puede haber querido mostrar el catálogo a María*
  - he could have wanted to show the catalog to Maria

- Clitic forms:
  - *puede haber querido mostraselo* (LC-0)
  - *se lo puede haber querido mostrar* (SC)
  - *puedeselo haber querido mostrar* (LC-3)
  - *puede haberselo querido mostrar* (LC-2)
  - *puede haber queridoselo mostrar* (LC-1)

- In all fives cases there one insertion and one cancelation!

- Clitic can climb over any number of subject raising and subject control verbs in the periphrasis!
Clitics in *ver* transitive

- Explicit arguments:
  - *ve a María*
  - he sees Maria

- Clitic forms:
  - *vela* (LC-0: Imperative)
  - *la ve* (SC)

- Again one insertion and one cancelation!

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitics climbing with *ver* transitive

- **Explicit arguments:**
  - *puede ver a María*
  - he can see Maria

- **Clitic forms:**
  - *puede verla* (LC-0)
  - *la puede ver* (SC)
  - *puedela ver* (LC-1, but personal form)

- **Clitic climbing is ok.**

- **Again one insertion and one cancelation!**

---

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitics in object-control verbs

- Explicit arguments:
  - *ve comer la manzana* (grammatical but a bit odd...)
  - he sees to eat the apple

- It looks like in Spanish
  - Complement of *ver*: *comer la manzana*
  - Direct object of *comer*: *la manzana*

- Clitic forms:
  - *ve comerla* (LC-0)
  - *la ve comer* (SC)
  - *vela comer* (LC-1: Imperative)

- But... *ver* (to see) in an object-object control so:
  - object of *ver* (to see) should be the subject of *comer* (to eat)!
  - *The apple* cannot be the subject of *to eat!*

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitics in object-control verbs

- Explicit arguments:
  - \textit{ve} \textit{comer} \textit{la manzana} a \textit{María}
  - he sees \textit{Maria} to eat the apple (accusative accusative!)

- But in Spanish
  - Indirect object of \textit{comer}: \textit{Maria} (\textit{dativo ético})
  - Direct object of \textit{ver}: \textit{comer} \textit{la manzana} a \textit{María}?

- Clitic forms:
  - \textit{ve} \textit{comersela} \quad (LC-0)
  - \textit{se} \textit{la} \textit{ve} \textit{comer} \quad (SC)
  - \textit{vesela} \textit{comer} \quad (LC-1: Imperative)

- But again, sees in an object-object control so:
  - object of \textit{ver} (to see) should be the subject of \textit{comer} (to eat)!

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitic’s scope

- Sentence with explicit arguments:
  - *ve a María comer la manzana (also a bit odd...)
  - he sees Maria eat the apple

- Clitic forms:
  - *ve a María la comer
  - *vela la comer
  - *la la ve comer (“la la” is not a clitic word)
  - *la a María ve comer

- Clitics have local scope: the clitic domain

- Clitic climbing is not allowed if there are conflicts between local clitic domains

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitic’s scope

- The idiomatic form:
  - *Ve a María comerse la manzana*
  - sees Maria eat for herself the apple
- Where does this sentence come from?
  - *ve a María comer la manzana para María*
  - he sees Maria to eat the apple for Maria
- Clitic forms:
  - *Ve a María comerse la manzana* (1 insertion and LC-0)
  - *ve a María comersela* (1 insertion and LC-0)
  - *la ve comersela* (Inst. 1: SC; Inst. 2:LC-0 & Inf.)
  - *vela comersela* (Inst. 1: LC-0 & Imp; Inst. 2:LC-0 & Inf.)
  - *a María se la ve comer* (ok:1 insertion and SC)
  - *se la a María ve comer*
  - *a María se la ve comer* (1 insertion and SC)

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The model (version 2)

- Clitics sequences are words in the lexicon, including case information
- Dual behavior:
  - Proclitics: independent lexical units
  - Enclitics: inflexional affixes, but must be clitic words!
- Clitics form local domains:
  - Clitic order is local to clitic domains (e.g. *la ve comersela*)
  - If domains conflict they impose barriers to clitic climbing
- There is one insertion and cancelation for each local domain
Clitic subsumption

- The idiomatic form again:
  - *Ve a María comersela la manzana*
  - sees Maria eat for herself the appel

- Clitic forms:
  - *ve a María comersela* (LC-0)
  - *sela la ve comer*
  - *la sela ve comer*

- But these are ok!:
  - *se la ve comer* (2 insertions, 1 SC!)
  - *vesela comer* (2 insertions, 1 complex LC: Imperative)

- Clitic subsumption: *se = la* (María) + *se* (María)

- Complex lexical cancelation: Two insertions
Clitic’s subsumption

- Clitic subsumption: the *se* subsumes the two corref. inst.
  - *se* can pass when *se* in *sela* and *la* (the dir. object of *ver*) correfer!
  - *se la ve comer*: Two insertions but only one SC
  - *vesela comer*: Two insertions and a Composite-LC

- Clitic forms:
  - *Ve a María comerse la manzana*  
    (1 insertion and LC-0)
  - *ve a María comersela*  
    (1 insertion and LC-0)
  - *vela+sela comer*
  - *vesela comer*  
    (subsumption and Comp-LC)
  - *se la ve comer*  
    (subsumption and SC)
The model (version 3)

- Clitics sequences are words in the lexicon, including case information, but present dual behavior:
  - Proclitics: independent lexical units
  - Enclitics: inflexional affixes, but must be clitic words!

- Clitics form local domains:
  - There is one insertion and cancelation for each local domain
  - Clitic order is local to clitic domains
  - Conflicting domains create barriers to clitic climbing

- If clitics of different domains corefer:
  - The coreferential forms can be subsummed into one
  - Clitic climbing is permitted: the intermediate offending clitic is made transparent
Clitic composition

- Another idiomtactic form:
  - *se lo pude ver poner*
  - I could see her put it out there

- Origin of the sentence
  - *pude ver a María poner el sombrero*
  - *pude ver a María ponerlo*
  - *puede verse ponerlo* (leísmo le = se almost compulsory!)
  - *puede verse + lo poner* (Hidden composition step)
  - *(se lo) pude ver poner* (There is no “la lo” in the clitic lexicon)
Ambiguity of *se*

- Another idiomatic form:
  - *se la pude ver comer*
  - I could see her eating it

- Origin of the sentence
  - *pude ver a María comer la manzana*
  - *pude ver a María comerla*
  - *pude verle comerla* (leísimo, almost compulsory!)
  - *pude verse+la comer* (*se*=le is an arg. of *ver* and *la* of *comer*)
  - *pude verse la comer* (Composition)
  - *(se la)pude ver comer* (2 insertions, composition and SC)

Like subsumption but with accusative *se*!

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Blocking composition

- Another idiomtactic form:
  - le pude ver ponesela
  - I could see her put it (the hat on the boy)

- Origin of the sentence
  - pude ver a María poner el sombrero al niño
  - pude ver a María poneselo
  - pude verle ponerlo
  - *pude verle+selo poner (“le se lo” is not a cletic)
  - le pude ver ponerlo (the clitics have to stay in their local domains)
Blocking combinations!

- Another idiomatic form:
  - se la pude ver poner
  - I could see her put it (the hat on the child)

- Origin of the sentence
  - pude ver a María poner el sombrero al niño
  - pude ver a María ponerselo
  - pude verse ponerselo (leísmo se = le; almost compulsory!)
  - *pude verse+selo poner (composition?)
  - *pude verse+lo poner (subsumption and composition and LC?)
  - se lo pude ver poner (2 insertions, subsumption and SC)

The last one looks ok but it has a different meaning: no longer reference to María or to the child!
Clitic composition

- Composition of clitics of different verbs:
  - Sum of the individual clitics to form a clitic word
  - One insertion per clitic word in the composition
  - Must be in accessible local domains

- Composite Lexical Cancelation
  - *pude verse+lo poner*

- Syntactic cancelation:
  - *se la pude ver poner*
The model (version 4)

- Clitics sequences are words in the lexicon, including case information, but they present dual behavior:
  - Proclitics: independent lexical units
  - Enclitics: inflexional affixes, but must be clitic words!
- Clitics form local domains:
  - There is one insertion and cancelation for each local domain
  - Clitic order is local to clitic domains
  - Conflicting domains create barriers to clitic climbing
- Clitic subsumption
  - Correferential forms can be subsummed into one (dative wins!)
  - Clitic climbing is permitted: the intermediate offending clitic is made transparent
The model (version 4... cont)

- Clitic composition
  - Clitics of different verbs (insertions) in accessible local domains can be added on to form a single word of the clitic lexicon
  - Composite clitics are canceled in a single operation
Pleonastic clitics

• Sentence with explicit arguments:
  – *ve a María poner el sombrero al niño
  – he sees Maria to put on the child the hat

• Clitic forms:
  – *ve a María ponerle el sombrero al niño (dative duplication)
  – *ve a María ponerlelo al niño
  – *ve a María poneselo al niño (le is a duplication of se)
  – la *ve poneselo al niño
  – al niño la *ve poneselo

• In general clitic insertion can cancel the arguments, but it can also leave the arguments explicit, as long as the insertion and the argument are co-indexed
The offending participle

- A idiomatic form:
  - *se lo ha visto puesto
  - he has seen it put on to her

- Source sentence:
  - *ha visto a María puesto el sombrero a María  (Not Grammatical)
  - *ha visto a María puestoselo  (Not Grammatical)
  - se lo ha visto puesto  (Clitic subsumption)
  - *ha vistoselo puesto  (Clitic subsumption, but offending participle)

- In participle, clitics must go to the front:
  - LC-0 is not allowed for participles at the end of the periphrasis, but pleonastic clitics are ok.
  - se lo ha visto puesto a María
  - se lo ha visto puesto el sombrero
  - se lo ha visto puesto el sombrero a María

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The final model (version 5)

- Clitics sequences are words in the lexicon, including case information, but they present dual behavior:
  - Proclitics: independent lexical units
  - Enclitics: inflexional affixes, but must be clitic words!

- Clitics form local domains:
  - There is one insertion and cancelation for each local domain
  - Clitic order is local to clitic domains
  - Conflicting domains create barriers to clitic climbing

- Clitic subsumption
  - Correferential forms can be subsummed into one (dative wins!)
  - Clitic climbing is permitted: the intermediate offending clitic is made transparent

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The final model (version 5... cont)

- Clitic composition
  - Clitics of different verbs (insertions) in accessible local domains can be added on to form a single word of the clitic lexicon
  - Composite clitics are canceled in a single operation
- Clitic duplication and offending participle:
  - Insertion can leave the argument explicit and clitic and argument are co-indexed
  - LC-0 is not allowed for participles at the end of the periphrasis
- First low of clitics dynamics: clitics are never created or destroyed, they are only transformed

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The clitic system

- Clitic sentences
- Definition and properties of clitics
- The Model
- The HPSG Model
The HPSG model

• A set of clitic words in the lexicon
  – Orders allowed
  – Case information

• Clitization rules:
  – For enclitics with LC-0
  – For proclitics:
    • Postponed cancelation (Syntactic Cancelation)
    • Definition of a clitic list (Valence feature)
  – For auxiliaries, with LC-\(i\), where \(i > 0\)
  – For pleonastic constituents:
    • Clitization leaves the argument in the complement list.
    • Complement and clitic are coindexed

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The HPSG model

- Sintactic cancelation:
  - Head-Proclitic Rule

- Clitic climbing:
  - Simple:
    - Subject raising and subject control verbs
  - With clitic composition:
    - Object control verbs
  - With Subsumption:
    - Object control verbs + subsumption lexical rule

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitic word

\[ \langle se \  lo, \ \begin{bmatrix} \text{clitic} \\
\text{CL–LIST} \ \langle \text{noun[CASE \ dat]}, \ \text{noun[CASE \ dat]} \rangle \end{bmatrix} \rangle \]
Lexical rule for enclitics

\[
\left\langle \#1, \begin{bmatrix}
\text{verb} \\
\text{SPR} \\
\text{COMPS} \\
\left\langle \text{NP} \right\rangle \\
\left\langle \text{[]} \right\rangle \\
\left\langle \text{[]} \right\rangle \\
\end{bmatrix} \right\rangle \Rightarrow \left\langle F_{\text{clit}}(\#1), \begin{bmatrix}
\text{SPR} \\
\text{COMPS} \\
\left\langle \text{NP} \right\rangle \\
\end{bmatrix} \right\rangle
\]
Enclitic sentence

mostrarmelo

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Lexical rule for proclitics (postponed cancelation)

\[ \langle #1, \begin{bmatrix} SPR \\ COMPS \end{bmatrix} \langle #1 \rangle \rangle \Rightarrow \langle #1, \begin{bmatrix} SPR \\ COMPS \\ CL-LIST \end{bmatrix} \langle #1NP \rangle \rangle \]
Head-Proclitic Rule
(Syntactic cancelation)
Proclitic sentence

se lo

muestras

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Basic definitions of \( srv \) and \( scv \)

- **Subject raising:** \( poder, haber, \) etc.

\[
\begin{bmatrix}
\text{srv} \rightarrow \text{lkm} \\
\text{SPR} \left\langle \#1 \right\rangle \\
\text{COMPS} \left\langle \text{phrase} \right\rangle \\
\text{SPR} \left\langle \#1 \right\rangle
\end{bmatrix}
\]

- **Subject control:** \( querer, pasar a, \) etc.

\[
\begin{bmatrix}
\text{scv} \rightarrow \text{lkm} \\
\text{SPR} \left\langle \#1\text{NP}_i \right\rangle \\
\text{COMPS} \left\langle \text{phrase} \right\rangle \\
\text{FORM inf} \\
\text{SPR} \left\langle \#1\text{NP}_i \right\rangle
\end{bmatrix}
\]
Lexical entry for verbs in periphrasis

- The clitic list is a head feature
- Clitic lists of the verb and its complement are coindexed

\[
\begin{bmatrix}
SPR & \langle #1 \rangle \\
CL-LIST & #2 \\
COMPS & \begin{bmatrix}
SPR & \langle #1 \rangle \\
COMPS & \langle \rangle \\
CL-LIST & #2 \\
\end{bmatrix}
\end{bmatrix}
\]

- Clitic list “climbs” by the Head-feature principle

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Simple clitic climbing

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Simple clitic climbing with several verbs (raising and control)
Lexical rule for auxiliaries
(complex lexical cancellation)

\[
\langle \#2, \text{[COMPS \#1]} \rangle \Rightarrow F_{\text{clit-selo}}(\#2), \quad \text{COMPS} \quad \langle \text{CL-LIST} \quad \text{CL-LIST} \quad \langle [\text{NP-dat}][\text{NP-acc}] \rangle \rangle
\]

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Sentence with clitized auxiliaries

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Clitic composition

\[
\begin{align*}
&\text{ocv} - \text{lxm} \\
&\text{CL} - \text{LIST} \left< [] \oplus \# a \right> \\
&\text{SPR} \left< \# 1 \right> \\
&\text{COMPS} \left< \# 2, \# 3 \right. \\
&\left. \text{CL} - \text{LIST} \left< \# a \right> \right> \\
&\text{SPR} \left< \# 2 \right> \\
&\text{ARG} - \text{ST} \left< \# 1, \# 2, \# 3 \right>
\end{align*}
\]

In the definition of object-control verbs

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Object-control verb after clitization

\[
\begin{align*}
\text{CL–LIST} & \langle #2 \rangle \oplus \langle #1 \rangle \\
\text{SPR} & \langle #3 \rangle \\
\text{COMPS} & \left[ \begin{array}{c}
\text{phrase} \\
\text{CL–LIST} & \langle #1 \rangle \\
\text{SPR} & \langle #2 \rangle 
\end{array} \right]
\end{align*}
\]
Sentence with clitic composition
Clitic subsumption

- Modification of object-control verbs

\[
\begin{align*}
\langle \#1, \left[ ocv-lxm \right. \\
CL-\text{LIST} \langle [] \rangle \\
COMP \langle [\text{COMP} \langle CL-\text{LIST} \#a \{[],[]\}\rangle],\rangle \rangle \Rightarrow \#1, \left[ CL-\text{LIST} \#a \\
COMP \langle [CL-\text{LIST} \#a]\rangle \rangle
\end{align*}
\]
Sentence with clitic subsumption

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Lexical rule for pleonastic constructions

\[
\langle #1, \ [COMPS \ \langle #2 \rangle \rangle \Rightarrow \langle #1, \ \begin{bmatrix}
CL - LIST & #2_i \\
COMPS & \langle #2NP_i \rangle
\end{bmatrix}\rangle
\]
Pleonastic sentence

se la ve a María comer

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
But...

- Local clitic domains:
  - *pude verle+selo poner ("le se la" is not a clitic)
  - le pude ver ponerselo

- Conflicting domains
  - *ve a María la comer (There is material between the verb and the inflexion)
  - *vela la come ("la la" is not a clitic word)
  - vela a María comer (Ok!, Imperative)
Summary of the model

• The clitic word:
  – Accounts for word order and constrains possible clitic sequences, either simple or composite
  – Clitic syntactic cancelation

• Identification of local domains

• New clitic operations:
  – Composition
  – Subsumption

• Account of pleonastic use

• Dialectal variation:
  – alternative sets of clitic words?
  – Order parameter in clitic composition?
Further work

• Clitics in composite sentences
  – Yuxtaposition
  – Coordination
  – Subordination
    • Substantival
    • Adjetival
    • Adverbial

• Dialectal variation:
  – Alternative sets of clitic words?
  – Order parameter in clitic composition?

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with se
  - Passive reflexive (pasiva refleja)
  - Active impersonal (impersonal activa)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Reflexive pronouns

- Reflexives use the clitic system coindexing arguments: all clitics can appear in reflexive construction

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>me</td>
<td>nos</td>
</tr>
<tr>
<td>2nd</td>
<td>te, os</td>
<td></td>
</tr>
<tr>
<td>3rd</td>
<td>se</td>
<td>se</td>
</tr>
<tr>
<td>neuter</td>
<td>lo</td>
<td></td>
</tr>
<tr>
<td>reflexive and impersonal</td>
<td>se</td>
<td>se</td>
</tr>
</tbody>
</table>

- The reflexive se comes from the Latin reflexive se

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Reflexives and reciprocals

- **Direct reflexive:**
  - *me baño*  
    (I bath himself)
  - *te vistes*  
    (you dress yourself up)
  - *se moja*  
    (he wets himself)

- **Indirect reflexive:**
  - *me lavo las manos*  
    (he washes his hands)
  - *se lava las manos*  
    (he washes his hands)

- **Dative reflexive (clitic subsumption):**
  - *se la ve comer*  
    (he sees her to eat it)

- **Reciprocals:**
  - *se tutean los niños*  
    (the children talk by the tú to each other)
Reflexives and reciprocals

• The pronoun *se* cannot be repeated, and in that case has to be changed by *uno*:
  
  – *uno se peina*  
  (one combes himself)  

  – *se tutean unos a otros*  
  (They talk to eachother by the *tú*)
The model

- Use the verbal and the clitic system
- An additional lexical rule to co-indexed arguments standing in a reflexive relation
  - Lexical rule for direct reflexive:
    \[
    \langle #1, \left[ \begin{array}{c} SPR \\ \text{COMPS} \langle [\ldots] \rangle \end{array} \right] \rangle \Rightarrow F_{\text{conj}}(\#1), \begin{array}{c} \text{CONJ} \\ \text{SEM} \end{array} \left[ \begin{array}{c} \text{HEAD|CL-LIST} \langle \#1 \rangle \\ \text{SPR} \\ \text{ACT} \langle i \rangle \\ \text{UND} \langle i \rangle \end{array} \right] \right]
    \]
  - Similar rules for indirect, dative reflexive (ethical dative) and reciprocals
  - Scope constrained to clitic local domains in case of domains conflicts

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The reflexive construction

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Pseudo-reflexives

- Reflexive construction with intransitive verb!
  - *me muero* (I died myself)
  - *se hace el listo* (I makes myself the smart one)
  - *nos hacemos tontos* (we are making fools ourselves)

- There is no recipient of the action!

- Very idiosyncratic: defined directly in the lexicon

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Content

• Antecedents
• The Spanish periphrasis
• The clitic system
• Reflexives and reciprocal
• Impersonal constructions with se
  – Passive reflexive (pasiva refleja)
  – Active impersonal (impersonal activa)
• Ergative constructions
• Pronominal questions
• Word order
• Implementation and verification
• Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Impersonal constructions with se

- Non-referential use of se
- Evolved from the Latin reflexive se
- Passive reflexive (*pasiva refleja*):
  - *se abre la puerta* (the door is opened)
  - *se* is just a syntactic marker for the form
- Active impersonal (*impersonal activa*):
  - *se vive tranquilo en Columbus*
  - one lives quietly in Columbus
  - *se* is a non-referential pronoun

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with *se*
  - Passive reflexive (*pasiva refleja*)
  - Active impersonal (*impersonal activa*)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Passive reflexive

• The reflexive *se*:
  – Only form for the 3rd. person, singular and plural and express accusative and dative case

• Direct reflexive (accusative):
  – *se lava* (he washes himself)

• Indirect reflexive (dative):
  – *se lava las manos* (he washes his hands)

• Ethical dative (participation in the action)
  – *María se comió las manzanas*
  – *María ate the apples (in her own benefit)*

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Passive reflexive

- Ethical dative (with intransitive verbs!)
  - *se queda en su casa todo el día*
  - *she stays in her house all day*

- Ethical dative (with stative verbs!)
  - *ella se creía que no era cierto*
  - She herfsself believed that it wasn´t true

- Still reflexives?
  - There is no action
  - Then, there can be no recipient!
  - The subject is not really an agent!
Passive reflexive

- Ethical dative (with stative verbs!)
  - se ha muerto un vecino mio
  - a neighbour of main has died.

- Still reflexives?:
  - The subject is a patient!

- By analogy:
  - La puerta se abre (the door is opened)
  - Los pajaros se alborotaron por el ruido
  - The birds were made restless by the noise
Passive reflexive

- **Active form**
  - *Juan abre la puerta*  
    (John opens the door)

- **First of passive (primera de pasiva)**
  - *La puerta es abierta por Juan*  
    - the door is opened by John

- **Second of passive (segunda de pasiva)**
  - *La puerta es abierta*  
    (The door is opened)

- **Reflexive passive:**
  - *se abre la puerta*  
    (The door is opened)
  - *la puerta se abre*  
    (The door is opened)

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Passive reflexive

- Reflexive passive:
  - *se abre la puerta* (The door is opened)
  - *la puerta se abre* (The door is opened)

- Syntactic properties:
  - The verb is *transitive* in active form
  - Agreement between the passive subject and the verb
  - Agreement between the *se* form and the verb (third person)

- The form *se* is no longer a pronoun:
  - non-referential
  - a syntactic marker for the form

- In written and spoken Spanish this is the form most commonly used (the normal passive communicates a foreign and heavy tone to the language)
The HPSG model

- Use the verbal and the clitic system
- A lexical rule:

\[
\begin{align*}
\langle #1, \left[ \begin{array}{c}
SPR \langle NP \rangle \\
COMPS \langle #2, #3... \rangle
\end{array} \right] \rangle & \rightarrow \langle F_{3rd}(#1), \left[ \begin{array}{c}
SPR \\
COMPS \\
CL\text{-}LIST
\end{array} \right] \rightarrow \langle #2 & agr \rangle \\
\langle #3, . . . \rangle \\
\langle se & agr \rangle
\end{align*}
\]
Structure of the reflexive passive
Semantics

- No referential se:

\[
\begin{bmatrix}
\text{INDEX} & \text{none} \\
\text{MODE} & \text{none} \\
\text{RESTRICTION} & \langle \rangle \\
\end{bmatrix}
\]

La puerta se abre

\[
\begin{bmatrix}
\text{INDEX} & i \\
\text{MODE} & \text{prop} \\
\text{RESTRICTION} & \begin{bmatrix}
\text{RELATION} & \text{name} \\
\text{SIT} & j \\
\text{NAME} & \text{door} \\
\text{NAMED} & v \\
\end{bmatrix} \\
& \begin{bmatrix}
\text{RELATION} & \text{open} \\
\text{SIT} & i \\
\text{ACT} & \text{name} \\
\text{UND} & v \\
\end{bmatrix}
\end{bmatrix}
\]
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with se
  - Passive reflexive (pasiva refleja)
  - Active impersonal (impersonal activa)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Active Impersonal

- Also from the reflexive *se*:
  - Only form for the 3rd. person, singular and plural and express accusative and dative case

- There are limitation to the use of the passive reflexive if the verbs are reflexive or reciprocal:
  - *se martirizan los cristianos*
  - *se tutean los niños*

- Ambiguity:
  - Is there some one that punish the cristians, or do we that to each other?
  - The children talked to each other by the *tú*, or there is some one else, the mother, that *tutea* the children?

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Active Impersonal

- Ambiguity resolution:
  - Use the passive (impersonal)
  - Spanish developed the use of the *a* preposition for direct objects of persona and determined abstract objects

- The passive (impersonal):
  - *Los cristianos son martirizados*
  - *Los niños son tuteados*

- The active impersonal:
  - *se martiriza a los cristianos*
  - *se tutea a los niños*
Active Impersonal

- The form is no longer a passive and can be used with intransitive and stative verbs:
  - se queda en su casa (one stays at his home!)
  - se vive tranquilo (one lives quietly)

- Properties:
  - Active sentences with **intransitive** verbs and **transitive** verbs that take as a direct object an animate or determined abstract object
  - The form *se* is a non-referential pronoun, and must agree with the verb in the 3rd. person!

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The HPSG model

- Use the verbal and the clitic system
- A lexical rule:

\[
\langle #1, \left[ \begin{array}{c} SPR \\ COMPS \end{array} \right] \rangle \Rightarrow \langle #1, \left[ \begin{array}{c} SPR \\ COMPS \end{array} \right] \rangle \left[ \begin{array}{c} CL-LIST \\ [se] \end{array} \right] \langle FORM \ mod] \rangle \ #2 \rangle
\]
Structure of the active impersonal

\[
\begin{array}{c}
\text{se} \\
\text{vive} \\
\text{tranquilo} \\
en \text{Cuernavaca}
\end{array}
\]
Semantics

- No referential *se*:

\[\begin{array}{c}
\text{INDEX} \\
\text{MODE} \\
\text{RESTRICTION}
\end{array}
\begin{array}{c}
i \\
\text{prop} \\
\text{RELN live} \\
\text{SIT i} \\
\text{ACT none} \\
\text{UND} \\
\text{NAMED} \\
\text{v}
\end{array}
\begin{array}{c}
\text{RELN name} \\
\text{SIT j} \\
\text{NAME cuernavaca} \\
\text{v}
\end{array}
\begin{array}{c}
\text{RELN tranquil} \\
\text{SIT k} \\
\text{THM i}
\end{array}\]

\[\text{Se vive tranquilo en Cuernavaca}\]
Relation between the *a* preposition in direct objects and *leísmo*

- The active impersonal:
  - se martiriza a los cristianos
  - se tutea a los niños

- The use of pronouns (*leísmo* is obligatory):
  - se les martiriza
  - se les tutéa
Convergence between passive reflexive and active impersonal

- But if the direct object is a thing, there is no *a* preposition
- The passive in plural:
  - *se venden libros*
- The singular:
  - *se vende libros* (Passive: books are sold)
  - *se vende libros* (Impersonal: one sells books)
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with *se*
  - Passive reflexive (*pasiva refleja*)
  - Active impersonal (*impersonal activa*)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions
Ergative constructions

- Active sentence with passive subject!
- Singular
  - El chocolate le gusta a María
  - María likes (enjoys) the chocolate
- Plural
  - Los chocolates le gustan a María
  - María likes (enjoys) chocolates
- Has nothing to do with word order:
  - a María le gustan los chocolates
  - le gustan a María los chocolates
  - le gustan los chocolates a María

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Two ways to think about it

- **Active sentence with passive subject!**
  - Agreement between passive subject and verb
  - There is no direct object (accusative)
  - There is an indirect object (dative)
  - The subject stands for the patient
  - Almost always with clitics and pleonastic complement!

- **Ergative sentence:**
  - Agreement between verb and direct object
  - Direct object in nominative

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Semantics

- The dative complement stands is the actor (although passive) of the action
- The subject stands for the recipient or undergoer of the action
- Semantics:
  - gusta(María, los chocolates)
HPSG Model

- Lexial entry for *gustar*:

\[
\begin{bmatrix}
SYN \\
SPR \\
COMPS
\end{bmatrix}
\begin{bmatrix}
\langle NP_j \rangle \\
ACT \ i \\
\langle NP_i -dat \rangle \\
UND \ j
\end{bmatrix}
\]
HPSG Model

- Lexical rule for clitization of *gustar*:

  \[
  \langle gustar \left[ \begin{array}{c} SPR \\ COMPS \\ \langle NP \rangle \end{array} \right] \rangle \Rightarrow \langle gustar \left[ \begin{array}{c} HEAD \\ SPR \\ COMPS \\ \langle NP \rangle \end{array} \right] \rangle \\
  \left[ \begin{array}{c} CL-\text{LIST} \\ \#1_i \end{array} \right]
  \]

- For a pleonastic complement!

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Structure of the ergative sentence

Los chocolates le gustan a María

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
HPSG Model

- Semantics:

Los chocolates le gustan a María

\[
\begin{align*}
\text{INDEX} & \quad \text{prop} \\
\text{MODE} & \quad \text{prop} \\
\text{RESTR} & \quad \text{RESTR}
\end{align*}
\]

\[
\begin{pmatrix}
\text{RELN} & \text{name} \\
\text{SIT} & \text{chocolate} \\
\text{NAME} & \text{NAME} \\
\text{NAMED} & \text{NAMED}
\end{pmatrix}
\begin{pmatrix}
\text{RELN} & \text{name} \\
\text{SIT} & \text{v} \\
\text{ACT} & \text{j} \\
\text{UND} & \text{i}
\end{pmatrix}
\begin{pmatrix}
\text{RELN} & \text{name} \\
\text{SIT} & \text{u} \\
\text{NAME} & \text{maria} \\
\text{NAMED} & \text{j}
\end{pmatrix}
\]
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with se
  - Passive reflexive (pasiva refleja)
  - Active impersonal (impersonal activa)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions
Pronominal questions

- Spanish *wh-questions*
- Pronouns:
  - *quién*  who  (nominative)
  - *qué*  what  (accusative)
  - *a quién, para quién*  whom  (dative)
  - *de quién*  whose  (genitive)
  - *dónde*  where  (locative)
  - *cuándo*  when  (temporal)
  - *cómo*  how  (instrumental)
  - *por qué*  why  (causal)
- Full arguments (subject or complements)
Pronominal questions

- Unbounded dependencies: gap categories
- Ample use of clitic but
- The gap constructions and the clitic system are orthogonal phenomena, so they interact very smoothly!
HPSG Model

- Lexical rule to introduce a *gap* argument:

\[
\langle #3, \left[ \text{\textit{SPR}}, \langle \text{\textit{COMP}}, \langle [\text{\textit{1}}, \text{\textit{#1}}, \text{\textit{#2}}, \text{\textit{...}}] \rangle \rangle \right] \Rightarrow \#3, \left[ \text{\textit{SPR}}, \text{\textit{COMP}}, \langle \text{\textit{#1}}, \text{\textit{wh}}, \text{\textit{CASE acc}} \rangle \rangle \right] \]

- Substitutes complement by *gap*
- Include case information: to cancel the *gap* with the corresponding pronoun
- Can be pipe-lined with a clitization rule

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Structure *what*-question

```
wh
#3

<table>
<thead>
<tr>
<th>SPR</th>
<th>COMPS</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>cliti</th>
<th>CL–LIST</th>
<th>#2</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SPR</th>
<th>COMPS</th>
<th>CL–LIST</th>
<th>GAP</th>
<th>#3</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SPR</th>
<th>COMPS</th>
<th>CL–LIST</th>
<th>GAP</th>
<th>#3</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SPR</th>
<th>COMPS</th>
</tr>
</thead>
</table>

qué le muestras a María

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Structure *when*-question

cuando se lo muestra a Maria

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with se
  - Passive reflexive (pasiva refleja)
  - Active impersonal (impersonal activa)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions
Word order

- Linear order (English):
  - Analytic: Determinant follows the determined
  - Oriented towards the listener
- Grapping (*envolvente*) order (Spanish):
  - Synthetic: Relevant things in front
  - Centered on the speaker
- There is no ideal linear or grapping order!
- The problem of word order: The inference by means the interpreter associate constituents in the input string (written or spoken) with the verbal arguments

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The problem of word order

- The inference performed by the interpreter to associate constituents in the input string (written or spoken) with their correspondent verbal arguments
- In ideal linear order:
  - Position of sentence´s constituents
- In Spanish:
  - Meaning of prepositions of constituents (considering lexical meanings of their complements, and verb)
  - Intonation
  - Position of sentence constituents!
The wanders of a

- In reflexives the preposition a is introduced to resolve ambiguity between reflexive and passive
- In active sentences:
  - A nominal complement with a is always the DO
  - Subjects are normally agents
  - Inanimate objects cannot be subjects in active sentences
Preposition for all other complements

- Indirect object: *a, para*  (beneficiary)
- Circumstantial complement:
  - *a:* movement, time
  - *de:* ownership, belonging, material, quantity, origin, mood, time
  - *en:* state, location
  - *para:* path
  - *por:* path, agent in passives, mean, mood, cause

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The chocolate sentences

- Juan dio los chocolates a María a las ocho
- Juan dio a María los chocolates a las ocho
- a María dio Juan los chocolates a las ocho
- a maría dio los chocolates Juan a las ocho
- a las ocho Juan dio los chocolates a María
- a los ocho a María Juan dio los chocolates
The chocolate sentences

- *se los dio los chocolates a las ocho*
- *se los dio los chocolates a las ocho a María*
- *a María se los dio*
- *se los dio a María a las ocho*
- *se los dio a las ocho*
- *se los dio*
- *
- Total of 87 sentences forms, without taking into account different clitic cases!

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Taming the beast!

- The usual order: SVO
- The verb cannot really go beyond the second position in simple sentences (contained in a phonic group) in idiomatic Spanish, unless there clitics involved!
- There are many local phenomena that produce different orders, but these are orthogonal to real word-order problems:
  - The clitic system (including pleonastic constructions)
  - Forms of unbounded dependecies (pronomial questions)
Our strategy

• Deal with local phenomena with lexical rules:
  – Inversion of direct and indirect complement

• Deal with complements in front of their normal position as unbounded dependencies
  – The ablative complements are more conspicuous!

• Augment variations of the main combination schemes, but with care to avoid over-generation, and also unwanted interactions:
  – Head-Specifier rule with verb in front of subject
Local phenomena

- Lexical rule for direct and indirect object inversion:
Using inversion of complements
Additional combination scheme

- Head-Specifier Rule with inversion:

\[
\begin{bmatrix}
\text{phrase} \\
\text{SYN}
\end{bmatrix}
\begin{bmatrix}
\text{PR} \\
\langle \rangle
\end{bmatrix}
\xrightarrow{H}
\begin{bmatrix}
\text{phrase} \\
\text{SYN}
\end{bmatrix}
\begin{bmatrix}
\text{PR} \\
\text{COMPS}
\end{bmatrix}
\begin{bmatrix}
\langle #1 \rangle
\end{bmatrix}
\]

#1
Using gaps

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Using gap and Inversion of comp.

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The chocolate sentences

- \textit{Juan dio los chocolates a María a las ocho}
The chocolate sentences

● Juan dio a María los chocolates a las ocho

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The chocolate sentences

● a maría dio los chocolates Juan a las ocho

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The chocolate sentences

- *a las ocho Juan dio los chocolates a María*
The chocolate sentences

• a las ocho a María Juan dio los chocolates

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with *se*
  - Passive reflexive (*pasiva refleja*)
  - Active impersonal (* impersonal activa*)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Analysis with the DIME corpus

- 15 subjects and 31 dialogues

Rate per dialogue:       Total (31 dialogues):
886 words               27459 words
185 expressions         6121 expressions
115 turns               3606 turns
14 minutes              7:10 hours

- Vocabulary 1122 word types

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
Analysis with the DIME corpus

● Spontaneous speech phenomena:
  – noises
  – interjections
  – stutters
  – speech-repairs
  – simultaneous speech
  – interrupts
  – long silences

● Full “syntactic” sentences: 45% (aprox. 2750)

● Sentences including one or more clitics: 748 (27%)
Clitic’s statistics

- Sentences with clitics: 748
  - Clitic system: 688 (Aprox. 2/3 proclitic and 1/3 enclitics)
    - One clitic pronoun: 677
    - Two clitic pronouns: 11
  - Other forms with se: 60

- Form with se
  - Reflexive passive (*pasiva refleja*): 39
    - 16 different verbs
    - Incidentally there are no instances of the normal passive
  - Active impersonal (*impersonal activa*): 2 (*se le llama*)
  - Pseudo reflexives (*pseudo reflejas*): 17
  - Unidentified forms: 2

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The HPSG grammar

- The Grammar rules:
  - HCR, HSR, HMR, HFR, ...
- The Lexical rules:
  - ARP, lexical rules, ...
- The inflection rules:
  - Conjugations, derivational rules, ...
- The lexicon

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The HPSG grammar

• The irregular forms:
  – Exceptions to the inflections rules

• The roots:
  – Features description of saturated structures:
    sentences and noun phrases

• The type hierarchy
The grammar rules

- Head specifier rule (1)
- Head complement rule (3)
- Head filler rule (2)
- Head modifier rule (2)
- Head postponed subject (1)
- Head proclitic rule (1)
- Subject pro-drop rule (2)
- Coordination rule (2)
The implementation

- 818 of types
- 199 words in the lexicon
- 14 grammar rules
- 30 lexical rules
- 49 inflectional rules
Content

- Antecedents
- The Spanish periphrasis
- The clitic system
- Reflexives and reciprocal
- Impersonal constructions with *se*
  - Passive reflexive (*pasiva refleja*)
  - Active impersonal (*impersonal activa*)
- Ergative constructions
- Pronominal questions
- Word order
- Implementation and verification
- Conclusions
The Structure of the grammar

Lexicon, lexical rules &
Type hierarchy
Basic Combination schemes

Verbal system

Clitic system

Other phenomena
Orthogonal phenomena

Clitic system

Argument structure

Word order

L. Pineda & I. Meza, IIMAS, UNAM, Mex., 2003
The end