

# Especificaciones para la anotación de SPinTX con información lingüística para fines pedagógicos

Martí Quixal

Diciembre 2012 – hoy

# Contents

<b>I</b>	<b>Definición de LIST y SET</b>	<b>4</b>
<b>1</b>	<b>LIST y SET basados en información morfosintáctica</b>	<b>5</b>
1.1	Categorías gramaticales . . . . .	5
1.2	Rasgos morfosintácticos . . . . .	5
<b>2</b>	<b>LIST y SET basados en lemas</b>	<b>6</b>
2.1	Los LIST . . . . .	6
2.2	Los SET . . . . .	6
<b>3</b>	<b>LIST y SET basados en palabras (literales)</b>	<b>7</b>
<b>4</b>	<b>LIST y SET basados en etiquetas pedagógicas</b>	<b>8</b>
<b>II</b>	<b>Secciones previas</b>	<b>9</b>
<b>5</b>	<b>Reglas para rectificar errores del análisis morfosintáctico</b>	<b>10</b>
<b>III</b>	<b>Secciones principales</b>	<b>12</b>
<b>6</b>	<b>Uso de los determinantes</b>	<b>14</b>
6.1	Usos de los artículos definidos . . . . .	14
6.2	Usos de los determinantes demostrativos . . . . .	14
6.3	Usos de los determinantes posesivos . . . . .	15
<b>7</b>	<b>Género gramatical y artículos</b>	<b>16</b>
7.1	Género masculino en palabras terminadas en “a” . . . . .	16
7.2	Cambio de artículo femenino a masculina en palabras con a tónica inicial . . . . .	16
<b>8</b>	<b>Uso de las preposiciones</b>	<b>17</b>
8.1	Usos de <i>para</i> . . . . .	17
8.1.1	El uso de <i>para</i> para expresar el destinatario o audiencia de una acción . . . . .	17
8.1.2	El uso de <i>para</i> para expresar una finalidad . . . . .	17
8.1.3	El uso de <i>para</i> para expresar una opinión . . . . .	18
8.1.4	Usos de <i>para</i> no clasificados . . . . .	18
8.2	Usos de <i>por</i> . . . . .	18
8.2.1	Usos de <i>por</i> en locuciones . . . . .	18
8.2.2	Usos de <i>por</i> no clasificados . . . . .	19

8.2.3	Usos de preposiciones no clasificados . . . . .	19
<b>9</b>	<b>Uso de los pronombres</b>	<b>20</b>
9.1	Pronombres de sujeto . . . . .	20
9.2	Pronombres de objeto directo o complemento atributivo . . . . .	20
9.3	Pronombres de objeto directo, indirecto, o posesivos . . . . .	20
9.4	Pronombres de objeto indirecto . . . . .	20
9.5	Pronombres de objeto doble . . . . .	21
9.6	Pronombre se (solo) . . . . .	21
<b>10</b>	<b>Formas verbales conjugadas</b>	<b>22</b>
10.1	Formas verbales para tiempos de indicativo . . . . .	22
10.1.1	Pretérito imperfecto . . . . .	22
10.1.2	Pretérito perfecto . . . . .	22
10.2	Verbos psicológicos: gustar, encantar, sorprender y otros . . . . .	23
10.2.1	Usos de <i>gustar</i> para expresar gustos propios . . . . .	23
10.2.2	Usos de <i>gustar</i> para expresar gustos de tu interlocutor . . . . .	23
10.2.3	Usos de <i>gustar</i> para expresar gustos de otras personas . . . . .	23
10.3	Expresar sorpresa . . . . .	23
10.3.1	Usos de <i>sorprender</i> . . . . .	24
10.4	Formas verbales para tiempos de subjuntivo . . . . .	25
10.5	Verbos copulativos . . . . .	25
10.5.1	Verbo ser . . . . .	25
10.5.2	Verbo estar . . . . .	26
10.6	Verbo haber . . . . .	26
10.6.1	Verbo haber en su uso com existencial . . . . .	26
10.7	Verbos auxiliares: ser, estar y haber . . . . .	26
10.7.1	Verbo estar en su uso como auxiliar . . . . .	26
10.7.2	Verbo haber en su uso como auxiliar en tiempos verbales compuestos . . . . .	26
10.7.3	Verbo ser en su uso como auxiliar . . . . .	26
<b>11</b>	<b>Funciones comunicativas</b>	<b>28</b>
<b>12</b>	<b>Pragmática</b>	<b>29</b>
12.1	Marcadores discursivos . . . . .	29
12.1.1	Marcadores discursivos indicadores de revisión . . . . .	29
12.1.2	Marcadores discursivos indicadores de causa . . . . .	30
12.1.3	Marcadores discursivos indicadores de igualdad . . . . .	31
12.1.4	Elementos discursivos usadors para ganar tiempo . . . . .	32
<b>13</b>	<b>Estructuras características del habla en hablantes de herencia</b>	<b>33</b>
13.1	Estructuras derivadas del contacto entre lenguas . . . . .	33
13.1.1	Expresiones con cambio de código . . . . .	33
<b>14</b>	<b>Vocabulario multpalabra</b>	<b>34</b>
14.1	Nombres propios . . . . .	34
14.1.1	Nombres de localizaciones geográficas . . . . .	34
14.1.2	Combinaciones de verbo y preposición de régimen . . . . .	34

# Introducción

Este documento describe el proceso de anotación de las transcripciones de los clips del corpus SPinTX (Spanish in Texas) con información pedagógica para facilitar el uso de los mismos en la enseñanza y el aprendizaje de español como lengua extranjera.

DOCU

DRAFT

## Part I

# Definición de LIST y SET

DRAFT

## Chapter 1

# LIST y SET basados en información morfosintáctica

1.1 Categorías gramaticales

1.2 Rasgos morfosintácticos

DRAFT

## Chapter 2

# LIST y SET basados en lemas

### 2.1 Los LIST

### 2.2 Los SET

DRAFT

## Chapter 3

# LIST y SET basados en palabras (literales)

DRAFT



## Chapter 4

# LIST y SET basados en etiquetas pedagógicas

DRAFT

## Part II

### Secciones previas

DRAFT

## Chapter 5

# Reglas para rectificar errores del análisis morfosintático

- ETIQ: CC
- EJEM: *pero*
- DESC: Regla para cambiar a *pero* la lectura de nombre por la de conjunción. Ahora se aplica indiscriminadamente, pero podría ser conveniente controlar que antes del *pero* no hay un artículo. Por ejemplo: “el *pero* de la cuestión es que podría romperse el motor”.
- ETIQ: Cambia Noun NC Singular Masculine por Verb VLadj
- EJEM: *hecho*
- DESC: Regla para cambiar a *pero* la lectura de nombre por la de conjunción. Ahora se aplica indiscriminadamente, pero podría ser conveniente controlar que antes del *pero* no hay un artículo. Por ejemplo: “el *pero* de la cuestión es que podría romperse el motor”.
- ETIQ: Cambia Verb VLfin Pres Indi Sing 1st por Noun NC Singular Masculine en la palabra “rancho”.
- EJEM: *rancho*
- DESC: Lo hace de forma indiscriminada. Quizá en el futuro habrá que cambiarla, refinarla.
- ETIQ: Cambia lectura de “dicho” como determinante por lectura como adjetivo/participio.
- EJEM: *dicho*
- DESC: Lo hace de forma indiscriminada. Quizá en el futuro habrá que cambiarla, refinarla.
- ETIQ: Cambia lectura de “oye” como presente de subjuntivo por lectura de presente de indicativo/imperativo.
- EJEM: *oye*
- DESC: Lo hace de forma indiscriminada. Quizá en el futuro habrá que cambiarla, refinarla.

- ETIQ: Reglas que cambian la lectura de "ser" por "ir" en "fueron" cuando no va precedida de un pronombre se o seguida la preposición con o un gerundio.
- EJEM: *se fueron juntos, se fueron a la escuela*
- DESC: Quizá en el futuro habrá que cambiarlas, refinarlas.
- ETIQ: Cambia lectura de "maestro" de adjetivo sustantivo.
- EJEM: *maestro, maestra, maestros, maestras*
- DESC: Lo hace de forma indiscriminada. Quizá en el futuro habrá que cambiarla, refinarla. Por ejemplo, en "obras maestras", que no ocurre en la versión actual del corpus.
- ETIQ: Cambia lectura de "eras" y "eran" de lema ambiguo entre erar y ser a sólo lema ser.
- EJEM: *eras, eran*
- DESC: Consideramos erar un verbo muy poco frecuente. Si hiciera falta, se podría refinar la regla.

DRAFT

## Part III

### Secciones principales

DRAFT

Las reglas de identificación de fenómenos pedagógicos están ordenadas de más concretas a más generales. Así, por ejemplo, se detectan primero, cuando es posible, los distintos usos de *para* (opinión, intención, destino, etc. como *@Gram:Prep:Para:Opinion...*) y luego se marcan como simplemente *@Gram:Prep:Para* el resto de ocurrencias de *para*. Luego, se hace lo propio con otras preposiciones, hasta el punto que si para una preposición no hay reglas especiales, sus ocurrencias se marcan simplemente como *@Gram:Prep*.

Esta estrategia se utiliza para evitar que una misma palabra reciba por partida doble informaciones de igual granularidad. Para ello, antes de asignar una etiqueta, se comprueba que no haya otra asignada del mismo nivel. A efectos prácticos, etiquetas como *@Gram:Prep:Para:Opinion...* y *@Gram:Prep:Para* se consideran del mismo nivel.

Para evitar este tipo de duplicidades también se podría usar una estrategia basada en la distinción entre ADD y MAP (véase la documentación de vislcg3). Por ahora no se ha hecho, porque esto impide la asignación de múltiples etiquetas a un mismo token/una misma palabra.

NOTA: Ahora mismo sólo estamos asignando hasta tres niveles de descripción pedagógica. Véase el Google Doc Pedagogical typologies <https://docs.google.com/spreadsheet/ccc?key=0AlyBsKogCRnwdDd3ZWRjejZlQU5XZWl6WVVUZmtKcEE&usp=sharing>.

DRAFT

## Chapter 6

# Uso de los determinantes

### 6.1 Usos de los artículos definidos

Identificaremos los usos de los artículos definidos *el, la, los, las*, y el artículo (en algunas gramáticas considerado pronombre) *lo* cuando va seguido de un adjetivo.

Rule

```
ADD (@Gram:Det:Art) Articulo IF (0 DetArtDefinidosLiterales) (NOT 1  
FiniteVerbForms OR PronombreDeObjetoNoAmbiguos OR SeLiteral);
```

- ETIQ: Asigna @Gram:Det:Art
- EJEM: *el, la, los, las*
- DESC: Cualquier ocurrencia no seguida de un verbo finito o un pronombre clítico, porque treetagger no distingue entre su lectura como determinante y la pronominal.

Rule

```
ADD (@Gram:Det:Art) Articulo IF (0 LoLiteral) (*1 Adj BARRIER  
FiniteVerbForms OR Conjunction OR LimiteOracionSimple) ;
```

- ETIQ: Asigna @Gram:Det:Art
- EJEM: *lo* + ADJETIVO
- DESC: Cualquier ocurrencia de LO seguida de un adjetivo (con un adv opcional en medio).

### 6.2 Usos de los determinantes demostrativos

Rule

```
ADD (@Gram:Det:Demo) ArtPronDemo IF (-1 Prep OR FiniteVerbForms) (0  
EsteEseAquel) (NOT 0 EstoEsoAquello) (1 Nombre) ;
```

- ETIQ: Asigna @Gram:Det:Demo
- EJEM: Prep/Verbo *este/ese/aquel* (LEMMA) + Nombre
- DESC: Cualquier ocurrencia de un demostrativo precedida de una preposición o un verbo conjugado y seguido de un nombre.

### 6.3 Usos de los determinantes posesivos

Rule

```
ADD (@Gram:Det:Pos) DetPosSing IF (*1 Nombre + SG BARRIER FiniteVerbForms  
OR Conjunction OR LimiteOracionSimple) ;
```

Rule

```
ADD (@Gram:Det:Pos) DetPosPlu IF (*1 Nombre + PL BARRIER FiniteVerbForms  
OR Conjunction OR LimiteOracionSimple) ;
```

Rule

```
ADD (@Gram:Det:Pos) DetPosFemPlu IF (*1 Nombre + PL BARRIER  
FiniteVerbForms OR Conjunction OR LimiteOracionSimple) ;
```

Rule

```
ADD (@Gram:Det:Pos) DetPosMascPlu IF (*1 Nombre + PL BARRIER  
FiniteVerbForms OR Conjunction OR LimiteOracionSimple) ;
```

Rule

```
ADD (@Gram:Det:Pos) DetPosFemSing IF (*1 Nombre + SG BARRIER  
FiniteVerbForms OR Conjunction OR LimiteOracionSimple) ;
```

Rule

```
ADD (@Gram:Det:Pos) DetPosMascSing IF (*1 Nombre + SG BARRIER  
FiniteVerbForms OR Conjunction OR LimiteOracionSimple) ;
```

- ETIQ: Asigna @Gram:Det:Poss
- EJEM: Prep/Verbo *mi/tu/su/sus/nuestro/nuestros/vuestro/vuestros* (LEMMA) + Nombre
- DESC: Cualquier ocurrencia de un determinante posesivo seguido de un nombre (se admiten ADV y ADJ intercalados).



## Chapter 7

# Género gramatical y artículos

### 7.1 Género masculino en palabras terminadas en “a”

Rule

```
ADDRELATION (Gram:Genero:ExcepMasc) DetAll IF (0 ElDetLiteral + SG OR  
UnDetLiteral + SG OR ArtPronDemo + SG) T0 (1 SustMascAcabadosEnA + Nombre  
+ SG) ;
```

- ETIQ: Asigna R:Gram:Genero:ExcepMasc
- EJEM: *el/un/ese* + (SUST acabado en -a, *tema, día, problema, trauma*)
- DESC: No se aplica con todos los determinantes porque TreeTagger no asigna género a todos y lematiza con criterios curiosos.

### 7.2 Cambio de artículo femenino a masculina en palabras con a tónica inicial

Rule

```
ADDRELATION (Gram:Genero:ExcepArtFem) DetAll IF (0 ElDetLiteral) T0 (1  
SustFemConATonicaInicial + Nombre + SG) ;
```

- ETIQ: Asigna R:Gram:Genero:ExcepArtFem
- EJEM: *el* + (SUST con a- tónica inicial, *arma, aula, agua*)
- DESC: No se aplica con todos los determinantes porque TreeTagger no asigna género a todos y lematiza con criterios curiosos.

## Chapter 8

# Uso de las preposiciones

### 8.1 Usos de para

Veáse <http://dev.coerll.utexas.edu/spanishtx/spe/html/adv11.html>.

#### 8.1.1 El uso de *para* para expresar el destinatario o audiencia de una acción

Rule

```
ADD (@Gram:Prep:Para) Prep IF (-1 Ser0Servir + Verbo) (0 Para) (NOT 0 GramPrep) (1 DetAll) (2 Personas);
```

- ETIQ: Asigna @Gram:Prep:Para
- EJEM: Fue escrito *para* los estudiantes.
- EJEM: Es fundamental *para* una persona.
- DESC: Requiere *para* + DET + NOMBRE < +humano >

Rule

```
ADD (@Gram:Prep:Para) Prep IF (-1 Ser0Servir + Verbo) (0 Para) (NOT 0 GramPrep) (1 Personas);
```

- ETIQ: Asigna @Gram:Prep:Para.
- EJEM: Ropa *para* niños.
- DESC: Requiere *para* + NOMBRE < +humano >

#### 8.1.2 El uso de *para* para expresar una finalidad

OBSERVACIÓN: Dudo de si el uso de “Juan es muy bueno para el fútbol” se parece o no al de “hemos venido antes para no quedarnos sin asiento”. De momento los pongo juntos.

Rule

```
ADD (@Gram:Prep:Para) Prep IF (0 Para) (NOT 0 GramPrep) (1 Infinitive);
```

- ETIQ: Asigna @Gram:Prep:Para.
- EJEM: Soy muy mala *para* recordar.
- DESC: Requiere *para* + INFINITIVO

Rule

ADD (@Gram:Prep:Para) Prep IF (0 Para) (NOT 0 GramPrep) (1 PronPersFuerte OR No) (2 Infinitive);

- ETIQ: Asigna @Gram:Prep:Para.
- EJEM: Había que saber hablar inglés *para* uno desarrollar su trabajo.
- DESC: Requiere *para* + PRONOMBRE PERSONAL NOMINATIVO + INFINITIVO

Rule

ADD (@Gram:Prep:Para) Prep IF (0 Para) (NOT 0 GramPrep) (1 Que);

- ETIQ: Asigna @Gram:Prep:Para.
- EJEM: Me daba clases *para* que aprendiera inglés.
- DESC: Requiere *para* + QUE

### 8.1.3 El uso de *para* para expresar una opinión

Rule

ADD (@Gram:Prep:Para) Prep IF (0 Para) (NOT 0 GramPrep) (1 PronombrePersonal) (NOT 1 YoLiteral OR TuLiteral);

- ETIQ: Asigna @Gram:Prep:Para.
  - EJEM: *Para* mi.
  - DESC: Requiere *para* + PRONOMBRE PERSONAL ACUSATIVO
- C: Esta regla es muy floja. Habrá que revisarla.

### 8.1.4 Usos de *para* no clasificados

Rule

ADD (@Gram:Prep:Para) Prep IF (0 Para) (NOT 0 GramPrep);

- ETIQ: Asigna @Gram:Prep:Para.
- EJEM: *para* (cualquier otra ocurrencia)
- DESC: Requiere *para*

## 8.2 Usos de *por*

Veáse <http://dev.coerll.utexas.edu/spanishtx/spe/html/adv11.html>.

### 8.2.1 Usos de *por* en locuciones

Rule

ADD (@Gram:Prep:Por) Prep IF (0 Por) (NOT 0 GramPrep) (1 LemasDePalabrasEnLocucionesConPor OR FormasDePalabrasEnLocucionesConPor);

- ETIQ: Asigna @Gram:Prep:Por
- EJEM: *por ejemplo, por eso, por tanto*
- DESC: Requiere *por* seguida por alguna palabra en la lista de palabras predefinidas *LemasDePalabrasEnLocucionesConPor*

Rule

ADD (@Gram:Prep:Por) Prep IF (0 Por) (NOT 0 GramPrep) (1 LoLiteral) (2 LemasDePalabrasEnLocucionesConPor);

- ETIQ: Asigna @Gram:Prep:Por.
- EJEM: *por lo tanto, por lo regular*
- DESC: Requiere *por* seguida por LO y alguna palabra en la lista de palabras predefinidas *LemasDePalabrasEnLocucionesConPor*

### 8.2.2 Usos de *por* no clasificados

Rule

ADD (@Gram:Prep:Por) Prep IF (0 Por) (NOT 0 GramPrep);

- ETIQ: Asigna @Gram:Prep:Por.
- EJEM: *por* (cualquier ocurrencia)
- DESC: Requiere *por*

### 8.2.3 Usos de preposiciones no clasificados

Rule

ADD (@Gram:Prep) Prep (NOT 0 GramPrep);

- ETIQ: Asigna @Gram:Prep.
- EJEM: *PREP* (cualquier ocurrencia)
- DESC: Requiere *PREP*

DRAFT

## Chapter 9

# Uso de los pronombres

### 9.1 Pronombres de sujeto

Rule ADD (@Gram:Pron:Sujeto) PronPersFuerte;

- ETIQ: Asigna @Gram:Pron:Sujeto.
- EJEM: *yo, tú, él, ella, ud., nosotros, vosotros, ellos, ellas, uds.* (cualquier ocurrencia)
- DESC: La ocurrencia de uno de ellos

### 9.2 Pronombres de objeto directo o complemento atributivo

Rule ADD (@Gram:Pron:Objeto) ElLema IF (NOT -1 Prep) (0 Artículo) (1 FiniteVerbForms) ;

- ETIQ: Asigna @Gram:Pron:Objeto
- EJEM: *mí nunca me la* hicieron pero que *estabamos*
- DESC: *lo, la, los, las* VERBO CONJUGADO

### 9.3 Pronombres de objeto directo, indirecto, o posesivos

Rule ADD (@Gram:Pron:Objeto) PronPers IF (0 PronombreDeObjetoNoAmbiguos);

- ETIQ: Asigna @Gram:Pron:Objeto
- EJEM: *me, te, os, nos* (cualquier ocurrencia)
- DESC: La ocurrencia de uno de ellos.

### 9.4 Pronombres de objeto indirecto

Rule ADD (@Gram:Pron:Objeto) PronPersClitico IF (0 LeLema);

- ETIQ: Asigna @Gram:Pron:Objeto
- EJEM: *le, les*
- DESC: La ocurrencia de uno de ellos.

## 9.5 Pronombres de objeto doble

Rule ADDRELATION (Gram:Pron:ObjetoDoble) PronPers (0 PronombreDeObjetoNoAmbiguos)  
T0 (1 ElLema + Artículo);

Rule ADDRELATION (Gram:Pron:ObjetoDoble) PronPersSE (0 SeLiteral) T0 (1 ElLema  
+ Artículo);

- ETIQ: Asigna @Gram:Pron:ObjetoDoble.
- EJEM: *se lo, se la, se los, se las, me lo, me la, me los, me las, te lo, te la, te los, te las*, etc.
- DESC: La ocurrencia de cualquier combinación de dos pronombres clíticos.

## 9.6 Pronombre se (solo)

Rule ADD (@Gram:Pron:Otros) PronPersSE (0 SeLiteral) (NOT 0 GramPron) (NOT 1  
FiniteVerbForms);

- ETIQ: Asigna @Gram:Pron:Otros.
- EJEM: *se lo, se la, se los, se las, me lo, me la, me los, me las, te lo, te la, te los, te las*, etc.
- DESC: La ocurrencia de cualquier combinación de dos pronombres clíticos.

# Chapter 10

## Formas verbales conjugadas

### 10.1 Formas verbales para tiempos de indicativo

#### 10.1.1 Pretérito imperfecto

Rule

```
ADD (@Gram:Imperf) FiniteVerbForms IF (0 FiniteVerbForms +  
PretImperfSimple) (NOT 0 Haber) ;
```

- ETIQ: Asigna @Gram:Imperf.
- EJEM: *cantaba, corría, temía, partía* (cualquier ocurrencia)
- DESC: Requiere una forma identificada como pretérito simple que no sea del verbo haber (auxiliar).

#### 10.1.2 Pretérito perfecto

##### Pretérito perfecto simple

Rule

```
ADD (@Gram:Indefinido) FiniteVerbForms IF (0 FiniteVerbForms +  
PretPerfSimple) (NOT 0 Haber) ;
```

- ETIQ: Asigna @Gram:PretPerf.
- EJEM: *cantó, corrió, anduvo, partió* (cualquier ocurrencia)
- DESC: Requiere una forma identificada como pretérito perfecto simple que no sea del verbo haber (auxiliar).

##### Pretérito perfecto perifrástico o compuesto

Rule

```
ADDRELATION (Gram:PerfPerif Gram:Aux:Haber) FiniteVerbForms (0 Haber +  
FiniteVerbForms + Presente) (NOT 0 Subjuntivo) TO (1 Part);
```

- ETIQ: Asigna R:Gram:PerfPerif a la forma del verbo HABER
- EJEM: *ha cantado, ha corrido, ha andado dormido* (cualquier ocurrencia)
- DESC: Requiere una forma finita de HABER seguida de un participio

## 10.2 Verbos psicológicos: gustar, encantar, sorprender y otros

Veáse <http://dev.coerll.utexas.edu/spanishtx/spe/html/adv13.html>.

Veáse <http://dev.coerll.utexas.edu/spanishtx/spe/html/beg05.html>.

### 10.2.1 Usos de *gustar* para expresar gustos propios

**Rule** ADDRELATION (Gram:Gustar:Propios) PronPers (0 Me OR NosLiteral) TO (1 Gustar-Like + 3aPersAmbig);

- ETIQ: Asigna @Gram:Gustar:Propios.
- EJEM: *me gusta* ayudar a mis amigos
- DESC: Requiere la forma ME seguida de una forma de GUSTAR en tercera persona (posiblemente ambigua)

**Rule** ADDRELATION (Gram:Gustar:Propios) PronPers (0 Me OR NosLiteral) (1 Haber) TO (1 Gustar-Like + Part);

- ETIQ: Asigna R:Gram:Gustar: Propios a la forma del verbo HABER
- EJEM: *me ha gustado* ayudar a mis amigos
- EJEM: *me hubieras gustado* con un vestido verde
- DESC: Requiere la forma ME seguida de una forma de HABER en tercera persona (posiblemente ambigua)

### 10.2.2 Usos de *gustar* para expresar gustos de tu interlocutor

**Rule** ADD (@Gram:Gustar:Interlocutor) FiniteVerbForms IF (-1 Tu) (0 Gustar-Like + 3aPersAmbig);

- ETIQ: @Gram:Gustar:Interlocutor
- EJEM: Dime, *te gusta* vivir en Estados Unidos.]
- DESC: Secuencia del lema TE y un forma de GUSTAR en tercera persona del singular.

### 10.2.3 Usos de *gustar* para expresar gustos de otras personas

**Rule** ADDRELATION (Gram:Gustar:Ajenos) FiniteVerbForms (0 Gustar-Like + 3aPersAmbig) TO (-1 LeLema);

- ETIQ: R:Gram:Gustar:Ajenos a la forma del verbo GUSTAR
- EJEM: Dime, *le gusta* vivir en Estados Unidos a tu hijo.]
- DESC: Secuencia del lema LE y un forma de GUSTAR en tercera persona del singular.

## 10.3 Expresar sorpresa

Veáse <http://dev.coerll.utexas.edu/spanishtx/spe/html/adv13.html>.



### 10.3.1 Usos de *sorprender*

- Rule** ADDRELATION (Gram:Gustar:Otros) FiniteVerbForms (0 Sorprender-Like + 3aPersAmbig) TO (-1 Yo);
- ETIQ: Asigna Gram:Gustar:Otros (Gram:ExpresarSorpresas:Propias) a la forma de SORPRENDER
  - EJEM: Siempre *me sorprende* su forma de hablar.
  - DESC: [Espacio para describir con palabras lo que hace la regla.]
- Rule** ADDRELATION (Gram:Gustar:Otros) FiniteVerbForms (0 Sorprender-Like + 3aPersAmbig) TO (-1 Tu);
- ETIQ: Asigna Gram:Gustar:Otros (antes era Gram:ExpresarSorpresas:Interlocutor) a la forma de SORPRENDER
  - EJEM: ¿Y *te sorprende* su forma de hablar?
  - DESC: [Espacio para describir con palabras lo que hace la regla.]
- Rule** ADDRELATION (Gram:Gustar:Otros) Part (0 Sorprender-Like) TO (\*-1 Tu BARRIER GrupoNominal OR LimiteOracionSimple LINK 1 Haber + 3aPersAmbig);
- ETIQ: Asigna Gram:Gustar:Otros (antes era Gram:ExpresarSorpresas:Interlocutor) a la forma de HABER
  - EJEM: ¿Hay alguna palabra mexicana que *te haya sorprendido*?
  - DESC: [Espacio para describir con palabras lo que hace la regla.]
- Rule** ADDRELATION (Gram:Gustar:Otros) FiniteVerbForms (0 Sorprender-Like + 3aPersAmbig) TO (-1 LeLema);
- ETIQ: Asigna Gram:Gustar:Otros (antes era Gram:ExpresarSorpresas:Ajenas) a la forma de SORPRENDER
  - EJEM: ¿Y *le sorprende* su forma de hablar?
  - DESC: [Espacio para describir con palabras lo que hace la regla.]
- Rule** ADDRELATION (Gram:Gustar:Otros) Part (0 Sorprender-Like) TO (\*-1 LeLema BARRIER GrupoNominal OR LimiteOracionSimple LINK 1 Haber + 3aPersAmbig);
- ETIQ: Asigna Gram:Gustar:Otros (antes era Gram:ExpresarSorpresas:Ajenas) a la forma de HABER
  - EJEM: ¿Hay alguna palabra mexicana que *le haya sorprendido*?
  - DESC: [Espacio para describir con palabras lo que hace la regla.]

## 10.4 Formas verbales para tiempos de subjuntivo

**Rule** ADD (@Gram:MSub:Pres) Verbo IF (0 Presente + Subjuntivo) (NOT -1 De) (NOT 1 Part);

**Rule** ADD (@Gram:MSub:Pret) FiniteVerbForms IF (0 Verbo + Pasado + Subjuntivo) (NOT -1 De) (NOT 1 Part);

**Rule** ADDRELATION (Gram:MSub:Pres) FiniteVerbForms (0 Haber + Verbo + Presente + Subjuntivo) (NOT -1 De) TO (1 Part);

**Rule** ADDRELATION (Gram:MSub:Pret) FiniteVerbForms (0 Haber + Verbo + Pasado + Subjuntivo) (NOT -1 De) TO (1 Part);

- ETIQ: Asigna Gram:MSub:Pres a verbos en presente o pretérito de subjuntivo
- EJEM: ¿Hay alguna palabra mexicana que *le haya sorprendido*?
- DESC: No se aplica si la forma verbal no esta precedida por una "de" para evitar formas incorrectamente marcadas como subjuntivo por el tagging básico. Si se trata de una forma compuesta, es decir, HABER + PART, se marca todo como tal.

## 10.5 Verbos copulativos

### 10.5.1 Verbo ser

**Rule** ADD (@Gram:SerEstar:Ser) Verbo (0 VerboSer OR VerboSerAmbiguoConErar) (NOT -1 Haber);

**Rule** ADD (@Gram:SerEstar:Ser) Verbo (0 VerboSerAmbiguoConIr) (NOT 1 Prep) (NOT -1 Haber);

**Rule** ADDRELATION (Gram:SerEstar:Ser) Verbo (0 Haber + FiniteVerbForms) TO (1 VerboSer + Part OR VerboSerAmbiguoConErar + Part) (NOT 1 Part);

- ETIQ: @Gram:SerEstar:Ser
- EJEM: *soy, eres, ha sido,* pero no *ha sido comprado*etc.
- EJEM: *fui, fuimos...* NOT Prep
- DESC: Cualquier forma de *ser* del pretérito indefinido no seguido de una preposición (porque entonces seguramente será del verbo ir). También las formas perifrásticas con haber que no vayan seguidas de un participio

## 10.5.2 Verbo estar

Rule ADD (@Gram:SerEstar:Estar) Verbo (0 VerboEstar);

- ETIQ: @Gram:SerEstar:Estar
- EJEM: *estoy, estás, estaban*, etc.
- DESC: Cualquier forma de *estar*

## 10.6 Verbo haber

### 10.6.1 Verbo haber en su uso com existencial

Rule ADD (@Gram:Haber:Exist) Verbo (0 Haber) (NOT 1 Part);

- ETIQ: @Gram:Haber:Exist
- EJEM: *Había* tres personas en el ascensor.
- DESC: Cualquier forma de *haber* que no vaya seguida de un participio

## 10.7 Verbos auxiliares: ser, estar y haber

### 10.7.1 Verbo estar en su uso como auxiliar

Se etiqueta con RGram:SerEstar:EstarAux las secuencias de estar seguido de un gerundio.

Rule ADDRELATION (Gram:SerEstar:EstarAux) VerboEstar IF (0 FiniteVerbForms) TO (1 Gerundio);

Rule ADDRELATION (Gram:SerEstar:EstarAux) VerboEstar IF (0 Infinitive) TO (1 Gerundio);

- ETIQ: R:Gram:SerEstar:EstarAux
- EJEM: y me *estuvo platicando* allí
- DESC: Actualmente sólo se aplica sobre formas conjugadas y en secuencias estar (infinitivo) más gerundio. Habrá que ver cuán necesario es aplicar reglas para "hemos estado mirando" o "estando comprando?" (el interrogante los uso para marcar que dudo de la gramaticalidad de esta oración.

### 10.7.2 Verbo haber en su uso como auxiliar en tiempos verbales compuestos

Se etiqueta con @Gram:Aux:Haber, pero se hace en las reglas que sirven para etiquetar los tiempos verbales correspondientes. Véase por ejemplo la anotación (y la regla) del Presente Perfecto en la sección 10.1.2.

### 10.7.3 Verbo ser en su uso como auxiliar

```

821: cosas así saben que siempre <soy reservada> , callada y que hablo
1177: niña siempre ... siempre he <sido callada> y reservada y siempre he
1629: . ¿ Cómo te gustaría <ser recordada> ? Mmm ... como una
2000: que es bueno , que <es divertido> , uno dice que es
2276: o mexicano , que ya <fue criado> en Estados Unidos . Este
16201: u opina que el español <sea enseñado> primero ? Pues van ...
17374: Unidos ? Si , yo <soy nacida> aquí . ¿ Alguna vez
19250: sentido ... aunque mi hermana <fue nacida> aquí , nosotros sentimos que
20857: que no toda la población <es educada> al nivel que debe ser
26684: sé cómo los escogerían pero <era jalado> y la gente de a
29420: llamarles la atención , nunca <fueron impresionados> con el barro . Siempre
37601: trabajos siendo bilingüe , que <siendo monolingüe> . [ Palabras inaudible de
46133: es una ciudad muy , <es considerada> patrimonio en ... en México
46710: era psicóloga y la otra <era licenciada> en administración de empresas ;
46893: conscientes y unos ciudadanos que <sean agradecidos> de la oportunidad que tienen
47265: , por su cultura , <es golpeado> y alguien te pide o
49938: Pipino Cuevas , y ella <era aficionada> a mirar esas peleas ,
51931: siguió estudiando una licenciatura y <es licenciada> en administración de empresas .
53157: hijos ? No , no <soy casada> . Muy bien . ¿
56029: a noviar , mi esposa <es nacida> en Estados Unidos , cuando
57098: ... que hasta la fecha <son malcriados> , rezongones , pero todavía
61044: estudiante era ? Pues , <era dedicado> , no ... desgraciadamente ,
63285: ... una familia humilde . <Son nacidos> en la sierra de Chihuahua
67513: pero por la mayoría todavía <soy visto> como , guises , soy
67519: visto como , guises , <soy visto> como algo que no pertenece
68082: como le llaman muchos ? <Es visto> de cierta manera en luz
68196: que eres raro , que <eres mojado> , que eres inmigrante ,
70858: o sea , viste que <soy reservado> , con cuidado , con
73204: , no importa si yo <era calificado> o no ... dijo ,
74128: , uno de mis hermanos <era casado> entonces tenía ya dos sobrinas
80950: spanglish ? Eso viene a <ser spanglish> , ¿ verdad ? Si
SPINTXLOC>

```

Figure 10.1: Ocurrencias de ser seguido de participio en Batch 1.

## Chapter 11

# Funciones comunicativas

DRAFT

# Chapter 12

## Pragmática

### 12.1 Marcadores discursivos

Véase <http://www.cs.famaf.unc.edu.ar/~laura/shallowdisc4summ/index.html>.

#### 12.1.1 Marcadores discursivos indicadores de revisión

**Rule** ADD (@Prag:MarcadoresDisc:Revision) MarcDiscRevision;

- ETIQ: @:Prag:MarcadoresDisc:Revision
- EJEM: *aunque, excepto o pero*
- DESC: [Espacio para describir con palabras lo que hace la regla.]

**Rule** ADDRELATION (Prag:MarcadoresDisc:Revision) Sin TO (1 Embargo);

- ETIQ: R:Prag:MarcadoresDisc:Revision (relación entre el primer i el último elemento de la locución)
- EJEM: *sin embargo*
- DESC: [Espacio para describir con palabras lo que hace la regla.]

**Rule** ADDRELATION (Prag:MarcadoresDisc:Revision) De TO (1 Hecho);

- ETIQ: R:Prag:MarcadoresDisc:Revision (relación entre el primer i el último elemento de la locución)
- EJEM: *de hecho*
- DESC: [Espacio para describir con palabras lo que hace la regla.]

**Rule** ADDRELATION (Prag:MarcadoresDisc:Revision) En TO (1 Realidad);

- ETIQ: R:Prag:MarcadoresDisc:Revision (relación entre el primer i el último elemento de la locución)
- EJEM: *en realidad*
- DESC: [Espacio para describir con palabras lo que hace la regla.]

- Rule**    ADDRELATION (Prag:MarcadoresDisc:Revision) APrep (1 Pesar) T0 (2 De);
- ETIQ: R:Prag:MarcadoresDisc:Revision (relación entre el primer i el último elemento de la locución)
  - EJEM: *a pesar de*
  - DESC: [Espacio para describir con palabras lo que hace la regla.]

- Rule**    ADDRELATION (Prag:MarcadoresDisc:Revision) De (1 Todos) T0 (2 Modos);
- ETIQ: R:Prag:MarcadoresDisc:Revision (relación entre el primer i el último elemento de la locución)
  - EJEM: *de todos modos*
  - DESC: [Espacio para describir con palabras lo que hace la regla.]

- Rule**    ADDRELATION (Prag:MarcadoresDisc:Revision) Es (1 Cierto) T0 (2 Que);
- ETIQ: R:Prag:MarcadoresDisc:Revision (relación entre el primer i el último elemento de la locución)
  - EJEM: *s cierto que*
  - DESC: [Espacio para describir con palabras lo que hace la regla.]

### 12.1.2 Marcadores discursivos indicadores de causa

- Rule**    ADD (@Prag:MarcadoresDisc:Causa) Porque;
- ETIQ: @Prag:MarcadoresDisc:Causa
  - DESC: [Espacio para describir con palabras lo que hace la regla.]

- Rule**    ADDRELATION (Prag:MarcadoresDisc:Causa) Prep IF (0 Por) (1 LoLiteral) T0 (2 Tanto);

- Rule**    ADDRELATION (Prag:MarcadoresDisc:Causa) Prep IF (0 Por) T0 (2 Tanto OR FormasDePalabrasEnLocucionesCausalesConPor);
- ETIQ: Asigna @Prag:MarcadoresDisc:Causa.
  - EJEM: *por lo tanto*
  - EJEM: *por eso*
  - DESC: Requiere *por* seguida por LO y TANTO

- Rule** ADDRELATION (Prag:MarcadoresDisc:Causa) Debido T0 (1 APrep);
- ETIQ: R:Prag:MarcadoresDisc:Causa (relación entre el primer i el último elemento de la locución)
  - DESC: [Espacio para describir con palabras lo que hace la regla.]
- Rule** ADDRELATION (Prag:MarcadoresDisc:Causa) Gracia (0 PL) (NOT 2 Dios) T0 (1 APrep);
- ETIQ: R:Prag:MarcadoresDisc:Causa (relación entre el primer i el último elemento de la locución)
  - DESC: [Espacio para describir con palabras lo que hace la regla.]
- Rule** ADDRELATION (Prag:MarcadoresDisc:Causa) Gracia (0 PL) (1 APrep) T0 (2 Dios);
- ETIQ: R:Prag:MarcadoresDisc:Causa (relación entre el primer i el último elemento de la locución)
  - EJEM: y de ahí fui a la ceremonia y *gracias a dios* me hice ciudadano
  - DESC: [Espacio para describir con palabras lo que hace la regla.]

### 12.1.3 Marcadores discursivos indicadores de igualdad

- Rule** ADD (@Prag:MarcadoresDisc:Igualdad) Adv IF (0 Ademas) (NOT 1 De);
- Rule** ADD (@Prag:MarcadoresDisc:Igualdad) Adv IF (0 Tambien);
- OBSERVACIÓN: No añado *precisamente* porque es ambiguo entre su lectura literal (preciso, justo, cierto) como modificador verbal y su lectura como marcador discursivo. Para anotararlo con fiabilidad habría que tener en cuenta su posición respecto al principio de la oración y el verbo principal. Ejemplos: *sí muchas pero no sé precisamente cuál cuál sería una, allí vivimos como un año y precisamente yo me tomé me tomaron fotos en un un estudio de fotografía*
- ETIQ: @:Prag:MarcadoresDisc:Igualdad
  - EJEM: *además, también*
  - DESC: Requiere *además* o *también* no seguido de DE

- Rule** ADDRELATION (Prag:MarcadoresDisc:Igualdad) Prep IF (0 Por) T0 (1 Ejemplo);
- ETIQ: R:Prag:MarcadoresDisc:Igualdad (relación entre el primer i el último elemento de la locución)
  - EJEM: *por ejemplo*
  - DESC: Requiere *por* seguida por alguna palabra en la lista de palabras predefinidas *Lemas-DePalabrasEnLocucionesConPor*



**Rule**    ADDRELATION (Prag:MarcadoresDisc:Igualdad) ConjLetra0 TO (1 Sea);

- ETIQ: R:Prag:MarcadoresDisc:Igualdad
- EJEM: *o sea*
- DESC: [Espacio para describir con palabras lo que hace la regla.]

**Rule**    ADDRELATION (Prag:MarcadoresDisc:Igualdad) No TO (1 SoloAcento);

- ETIQ: R:Prag:MarcadoresDisc:Igualdad
- EJEM: porque uno *no sólo* puede ayudar en el aspecto de inglés pero en el aspecto del español también
- DESC: [Espacio para describir con palabras lo que hace la regla.]

**Rule**    ADDRELATION (Prag:MarcadoresDisc:Igualdad) Tal TO (1 Vez);

- ETIQ: R:Prag:MarcadoresDisc:NoTipificado
- EJEM: noto que *tal vez* sus hijos hablan el spanglish más unos más que otros
- DESC: [Espacio para describir con palabras lo que hace la regla.]

#### 12.1.4 Elementos discursivos usadors para ganar tiempo

**Rule**    ADD (@Prag:RecuDiscurso:GanarTiempo) ArtPronDemo IF (NOT -1 Prep) (0 EsteLiteral) (NOT 1 Nombre);

- ETIQ: @Prag:ElementosDiscurso:GanarTiempo
- EJEM: *Este* en mi niñez era un poco más soñadora mucho más soñadora de hecho (...)
- DESC: Anota ocurrencias de ESTE que no vayan precedidas de una preposición ni seguidas de un sustantivo.

## Chapter 13

# Estructuras características del habla en hablantes de herencia

### 13.1 Estructuras derivadas del contacto entre lenguas

#### 13.1.1 Expresiones con cambio de código

Incluimos aquí expresiones que mezclan palabras en ambas lenguas, en este caso el español y el inglés. En la bibliografía especializada se utiliza el término inglés “code switching”.

Rule

ADDRELATION (Herit:Contacto:Mixtas) Verbo IF (0 ListaVerbosCorpusParaContactoMixtos) TO (*1 PalabraEnIngles BARRIER PalabraEnEspanol);
---

- ETIQ: R:Herit:Contacto:Mixtas
- EJEM: *hacer request, comer turkey*
- DESC: Anota secuencias de un verbo seguid de una o más palabras en inglés. Por ahora lo limitamos a ser, hacer, decir y comer porque TreeTagger anota como verbo algunas palabras que no lo son y no queremos generar ruido para los usuarios.

## Chapter 14

# Vocabulario multipalabra

### 14.1 Nombres propios

#### 14.1.1 Nombres de localizaciones geográficas

Rule      ADDRELATION (Vocab:Entities:Geo) Artículo IF (0 ("<El>")) TO (1 ("<Paso>"));

Rule      ADDRELATION (Vocab:Entities:Geo) Nombre IF (0 ("<Estados>")) TO (1 ("<Unidos>"));

Rule      ADDRELATION (Vocab:Entities:Geo) Nombre IF (0 ("<Ciudad>")) TO (1 ("<Juárez>"));

- ETIQ: R:Vocab:Entities:Geo
- EJEM: *Ciudad Juárez, El Paso, Estados Unidos*
- DESC: Reglas para anotar entidades frecuentes en el corpus. En este caso de nombres geográficos.

#### 14.1.2 Combinaciones de verbo y preposición de régimen

Rule      ADDRELATION (Vocab:Fraseol:VerbPrep) Verbo IF (0 ("hablar")) TO (\*1 Prep BARRIER LimiteOracionSimple OR GrupoNominal OR Verbo);

Rule      ADDRELATION (Vocab:Fraseol:VerbPrep) Verbo IF (0 ("pensar")) TO (\*1 Prep BARRIER LimiteOracionSimple OR GrupoNominal OR Verbo);

Rule      ADDRELATION (Vocab:Fraseol:VerbPrep) Verbo IF (0 ("empezar")) TO (\*1 Prep BARRIER LimiteOracionSimple OR GrupoNominal OR Verbo);

- ETIQ: R:Vocab:Fraseol:VerbPrep
- EJEM: *hablar PREP, pensar PREP, empezar PREP*

- DESC: Ahora no se distingue entre usos lexicalizados y usos composicionales (p.e., “empezar en San Antonio y terminar en Austin” se trata igual que “empezé con la tradición en la preparatoria”)

DRAFT