

Las construcciones con clíticos en la interlengua de aprendientes de español como L2. Tratamiento automático.

Carolina Tramallino, Viviana Valenti.

Grupo INFOSUR

Universidad Nacional de Rosario

Rosario, Argentina

Abstract

In the Spanish grammar, clitics are considered intermediate elements that are situated somewhere between word formation and the syntactic processes that are established between words. The complexity that characterizes them becomes still more evident in the interlanguage (IL) constructions produced by learners of Spanish as a second language. The objective of the present work will be to achieve a description and formalize the constructions with clitics that were observed in the IL, analyzing both the constructions that coincide with those of the standard L1 and those that deviate from it. To carry out the automatic analyses we will use a software tool called MPS that works in a modular way with the Smorph software. Finally, we will evaluate the results obtained.

Keywords: clitics, interlanguage constructions with clitics, automatic analysis.

Resumen

En la gramática del español, los clíticos son considerados elementos intermedios que se ubican entre la formación de palabras y los procesos sintácticos que se establecen entre palabras. La complejidad que los caracteriza se hace aún más evidente en las construcciones de interlengua (IL) producidas por aprendientes de español como segunda lengua. El objetivo del presente trabajo será lograr una descripción y realizar la formalización de las construcciones con clíticos que observamos en la IL, analizando tanto las construcciones coincidentes con la lengua estándar, así como las que se desvían de la misma. Para el análisis automático utilizaremos la herramienta informática MPS que trabaja de manera modular con el software Smorph y finalmente, evaluaremos los resultados alcanzados.

Palabras clave: clíticos, construcciones con clíticos de la interlengua, análisis automático.

1. INTRODUCCIÓN

En el presente trabajo, intentaremos una primera aproximación a la producción de construcciones con pronombres átonos por parte de aprendientes de español como segunda lengua. El objetivo es lograr el análisis automático de dichas construcciones.

El interés en los pronombres átonos o clíticos surge en parte, por la dificultad que éstos presentan durante su aprendizaje. Teniendo en cuenta que los clíticos presentan un comportamiento diferente en las distintas lenguas, es de esperar que surjan construcciones de interlengua que se diferencian de aquellas encontradas en la L1.

Para poder llevar a cabo una formalización de las mismas, que permita la posterior implantación en máquina, será necesario entonces, estudiar el comportamiento de los clíticos en la lengua meta, en este caso el español, y luego realizar una descripción de la interlengua, para poder clasificar y distinguir tanto a las construcciones que coinciden en superficie con la lengua estándar, como a aquellas que se apartan de la misma.

El corpus está conformado por treinta textos escritos con un total de 4609 palabras, pertenecientes a estudiantes de un nivel inicial que poseen variadas lenguas de origen. Los datos proceden del grupo Infosur. Siendo que la investigación está enmarcada en la lingüística computacional, el objetivo es el de emplear herramientas informáticas para lograr el análisis automático, no sólo de las construcciones estándares sino también de las estructuras que presenta la interlengua.

Para ello, en primer lugar, explicaremos a qué se llama clíticos y con qué estructuras pueden aparecer en el español; en segundo lugar describiremos la interlengua, deteniéndonos particularmente en las construcciones donde el clítico se encuentre antepuesto o pospuesto al verbo, y en oraciones imperativas, dejando de lado el análisis semántico en cuanto a los tipos de verbos que exigen clíticos y a los roles temáticos que pueden desempeñar estos últimos.

A continuación, mostraremos brevemente el funcionamiento de Smorph [1] y de MPS [2], que son las herramientas que emplearemos para el análisis automático y por último, explicaremos la implantación en máquina y las modificaciones pertinentes que deberán ejecutarse en los archivos que componen el software Smorph, para el reconocimiento de las construcciones desviadas.

2. CONSIDERACIONES TEÓRICAS SOBRE LOS CLÍTICOS EN EL ESPAÑOL.

Para clasificar y definir a los clíticos, objeto del presente trabajo, nos basamos en el estudio realizado por Fernández Soriano [3].

El español cuenta con dos series de pronombres: los átonos o clíticos, que se colocan adjuntos al verbo, y los tónicos, que aparecen en posición de objeto y alternan con los sintagmas nominales plenos. Mientras que los segundos se refieren exclusivamente a personas, los primeros pueden referir tanto a objetos como a personas.

Los pronombres clíticos de 3ra persona son los únicos que presentan distinciones morfológicas de caso y se dividen en: clíticos de acusativo (la-las-lo-los) y de dativo (le- les).

Si tuviéramos que explicar el estudio de los clíticos, diríamos que éstos se ubican entre la formación de palabras y los procesos sintácticos que se establecen entre palabras. Fonológicamente, dependen del verbo con el que aparecen, es decir, no portan acento. Tampoco pueden formar parte de una coordinación:

Esteban compró gaseosas y vinos.

* Esteban las y los compró.

La propiedad que los diferencia de los artículos o posesivos antepuestos, es que deben aparecer adjuntos al verbo, no pudiendo intervenir ningún elemento, excepto otro clítico:

* Lo sí quiero.

Además no se aceptan secuencias de más de tres clíticos. Por lo tanto, podemos afirmar que forman una unidad morfológica entre sí:

* me puede decirlo

Con verbos conjugados, el pronombre átono aparece adelante (proclisis), mientras que la posición posterior al verbo (enclisis) puede obtenerse con infinitivos (mirarla), gerundios (mirándola) e imperativos afirmativos (mirála). Los participios en el caso del español no admiten la adjunción de un clítico.

En cuanto a la adjunción de clíticos, ésta se da en bloque, es decir, no puede adjuntarse al verbo sólo uno de los clíticos, sino todos los de la secuencia:

Ej: quiero robárselo

* lo quiero robarse

* se quiero robarlo

3. DESCRIPCIÓN DE LAS CONSTRUCCIONES

En este apartado, se mostrarán las construcciones con clíticos que pueden aparecer en español y luego, las halladas en el corpus.

3.1. Las construcciones con clíticos en el español.

Para poder abordar la descripción de las construcciones de interlengua, fue necesario determinar todas aquellas posibles estructuras del español, en donde los clíticos aparecen en posición pospuesta y antepuesta. De ello se obtuvieron las siguientes combinaciones, que serán tomadas como base para la creación de reglas en MPS:

Construcciones en posición antepuesta:

Verbo conjugado	1 clítico	cl+verbo	te quiero
	2 clíticos	cl+cl+verbo	se lo dio

Construcciones en posición pospuesta:

Verbo conjugado	1 clítico	Imp+cl	dame
(solo imperativo)	2 clíticos	Imp+cl+cl	dámelo
Gerundio	1 clítico	ger+cl	poniéndolo
	2 clíticos	ger+cl+cl	poniéndoselo
Infinitivo	1 clítico	Inf+cl	darle
	2 clíticos	Inf+cl+cl	dárselo

3.2. Las construcciones con clíticos en la interlengua.

De la observación de los datos, se encontraron la totalidad de las estructuras con clíticos antepuestos, propias de la lengua meta. En cuanto a la posición pospuesta, solo se hallaron en la muestra, tres de las seis posibles construcciones. En algunos casos, las construcciones coincidieron con la lengua estándar y en otros presentaron desviaciones. (Ver tablas 1 y 2).

Tabla 1: ESTRUCTURAS COINCIDENTES CON LA L1 Y DESVIADAS. POSICIÓN ANTEPUESTA

Verbo conjugado	<u>1 clítico:</u> Te quiero: Cl+verbo	- saber si él aún la amaba - Me gustaba - Él lo olvidó cuando vino. - le dieron dos 2 cosas - Te quiero y te llama - la persona que lo encuentre lo lee -me dijiste -me gusta muchísimo -te permitiría -te quiero cantar	<u>Ausencia de clíticos:</u> - todavía puede tocar este órgano (se puede) - rompió una photo y el documento de identidad y (lo) escondió debajo un armatorio
	<u>2 clíticos:</u> Cl+cl+verbo	-la casa se le vino encima	

Tabla 2: ESTRUCTURAS COINCIDENTES CON LA L1 Y DESVIADAS. POSICIÓN POSPUESTA

Verbo Conjugado (IMP)	<u>1 clítico:</u> Imp+cli	-Llamame	
	<u>2 clíticos:</u> Daselo /ponetelo: Imp+cl+cl	-----	
Gerundio	<u>1 clítico:</u> Escuchándolo: Ger+cl	-----	
	<u>2 clíticos:</u> moviéndoselo ger+cl+cl	-----	
Infinitivo	<u>1 clítico:</u> Inf+cl	- darle gracias - Espero verte pronto - corresponderles - informarte - mendarte (mandarte) - mendarme (mandarme) - enviarte - queria divorciarse quisieron casarse	puedo dar ti mis fotos enviar ti usa pronombre personal objeto inf+probj
			Quiero ti decir voy a ti enviar Antepone un pronombre personal objeto probj+inf
			Dar a vos (darte) Agrega preposición + pronombre. personal sujeto Inf+prep+ppsuj
			Prefiero me preocupar Te enviar Antepone clítico al infinitivo cl+inf
	<u>2 clíticos:</u> Inf+cl+cl	-Quiero robárselo	

4. ANÁLISIS AUTOMÁTICO

El análisis automático se realizará con los Programas Smorph y Módulo Post Smorph (MPS), ya que ambos constituyen una herramienta declarativa separada de la maquinaria algorítmica, lo cual permite al lingüista trabajar con cualquier lengua y adaptar la herramienta a su propósito investigativo. Por lo tanto, la información lingüística obtenida tendrá que ser debidamente formalizada para que el autómata los interprete.

El primero de los programas mencionados, Smorph, es un analizador y generador morfosintáctico desarrollado en el GRIL por Sala Aït- Mokhtar, bajo la dirección de Gabriel Bès, que realiza la tokenización y el análisis morfológico en una sola etapa, produciendo las formas correspondientes a un lema, con los valores adecuados.

MPS trabaja sobre la salida de los datos de Smorph. Éste es un tokenizador de estado finito que segmenta textos sobre la base de la información semántica recibida y realiza un análisis morfosintáctico, reconociendo sintagmas y etiquetándolos.

Dicho tratamiento, que es previo al de la sintaxis general de la oración, convierte al software en una herramienta adecuada para el tratamiento automático de los clíticos.

MPS ejecuta dos funciones principales que son la *recomposición* y la *correspondencia*. Dentro de la recomposición tenemos dos tipos de reglas: de *reagrupamiento* y *división*.

La que usaremos en este caso, es la función de reagrupamiento, a partir de la cual la herramienta mediante determinadas reglas, expresará los valores de entrada, sobre dos o más estructuras de datos de salida de Smorph, y los valores de salida, sobre la estructura agrupada.

4.1. Implementación en Smorph

Se emplean los siguientes procedimientos:

Modificaciones en el archivo “rasgos”

Debemos crear la etiqueta morfosintáctica que corresponda a pronombres personales con caso objeto, para distinguir a *mí, ti* y *sí* de los pronombres personales con caso sujeto como *yo, vos, él*.

TPRON	tippron
prpers	pronpers
probj	pronobj
prindf	pronindf
prdem	prondem
prpos	pronpos .

Modificaciones en el archivo “entradas”

Se declaran los pronombres personales objeto en las entradas, con el rasgo proobj de la siguiente manera:

mí /probj .

ti /probj .

sí /probj .

4.2. Análisis morfológico a partir de Smorph de textos de aprendientes de español:

“*Te escribo para informarte* de un evento cultural muy original”

“...siempre *me dijiste* que querías compartir tus libros con desconocidos...”

“*Quiero ti decir* que mi viaje en la Argentina fue muy interesante.”

“Yo voy a *ti enviar* las fotos tomadas durante mi estadía en Rosario”.

“Si tu quieres, yo *puedo dar ti* mis fotos.”

“*Prefiero me preocupar* de los derechos humanes, antes los derechos de los animales.”

“*Te escribo para te enviar* fotografías tomadas durante nuestro estadía”

“Es possible *de me le enviar*”

“*Me encanta* esta fiesta y quiero un recuerdo”.

“Gustavo dijo que *la amaba* y *se fueron* a vivir juntos.”

“Él *lo olvidó* cuando vino a Argentina.”

“La casa *se le vino* encima...”

'Prefiero'.

['preferir', 'EMS','v', 'MODOV','ind', 'PERS','1a', 'NUM','sg', 'TPO','pres', 'TR','irr', 'TC','c2'].

'me'.

['lo', 'EMS','cl', 'TPCRF','rflnse'].

'preocupar'.

['preocupar', 'EMS','v', 'MODOV','infin', 'TR','r', 'TC','c1'].

'de'.

['de', 'EMS','prep'].

'los'.

['el', 'EMS','det', 'TDET','art'].

'derechos'.

['derechos', 'EMS','nom', 'GEN','masc', 'NUM','pl'].

'humanos'.

['humano', 'EMS','adj', 'GEN','masc', 'NUM','pl'].

'yo'.

['yo', 'TPRON','prpers'].

'puedo'.

['poder', 'EMS','v', 'MODOV','ind', 'PERS','1a', 'NUM','sg', 'TPO','pres', 'TR','hi'].

'dar'.

['dar', 'EMS','v', 'MODOV','infin', 'TR','r', 'TC','c1'].

'ti'.

['ti', 'TPRON','probj'].

'la'.

['el', 'EMS','det', 'TDET','art'].

'casa'.

['casar', 'EMS','nom', 'GEN','fem', 'NUM','sg',,].

'se'.

['lo', 'EMS','cl', 'TPCRF','rflse'].

'le'.

['lo', 'EMS','cl', 'TPCL','nrfl'].

'vino'.

['venir', 'EMS', 'v', 'MODOV', 'ind', 'PERS', '3a', 'NUM', 'sg', 'TPO', 'prets', 'TR', 'hi'].

4.3. Reglas de reagrupamiento en MPS para las construcciones de la lengua estándar

Se crean reglas en MPS para analizar automáticamente construcciones con clíticos existentes en el español estándar. Las siguientes corresponden tanto a estructuras en las que el clítico se encuentra antepuesto al verbo, como pospuesto al infinitivo. Para eso, las formamos en el archivo rcm.txt de la siguiente manera:

Regla N° 1: %infin + cl da infcl%

S1 [L1, 'EMS', 'v', 'EMS', 'infin'] S2 [L2, 'EMS', 'cl'] --> S1+S2 [L1+L2, 'EMS', 'infcl'].

“Preocuparme”

Regla N° 2: % cl+v da svn%

S1 [L1, 'EMS', 'cl'] S2 [L2, 'EMS', 'v'] --> S1+S2 [L1+L2, 'EMS', 'svn'].

“me gustaba”

Regla N°3: % cl+cl+v da svn%

S1 [L1, 'EMS', 'cl'] S2 [L2, 'EMS', 'cl'] S3 [L3, 'EMS', 'v'] --> S1+S2+S3 [L1+L2+L3, 'EMS', 'svn'].

“se le vino”

Luego, en el archivo Smorph, la herramienta arroja como resultado la etiqueta morfosintáctica correspondiente a los lemas agrupados:

'enviar te'.

['enviar te', 'EMS', 'infcl'].

'me gustaba'.

['me gustaba', 'EMS', 'svn'].

4.4. Reglas para las construcciones propias de interlengua

Finalmente, se procede a crear cuatro reglas de reagrupamiento para poder detectar las construcciones con clíticos propias de la interlengua. En éstas se especifica en el resultado, que la construcción se desvía de la lengua meta a partir de los caracteres desv o dsv:

a) Crearemos una regla propia para la interlengua de aprendientes de español que pueda analizar los casos en los que los clíticos anteceden al infinitivo. Ej: *te enviar, lo cual arrojará como resultado la etiqueta morfosintáctica: “infcldesv”.

$S1 [L1, 'EMS', 'cl'] S2 [L2, 'EMS', 'v', 'EMS', 'infin'] \rightarrow S1+S2 [L1+L2, 'EMS', 'infcldesv']$.

Te enviar

El resultado entonces, distinguirá a la construcción desviada:

'te envió'.

['lo enviar', 'EMS', 'vercl'].

'te enviar'.

['lo enviar', 'EMS', 'infcldesv'].

b) Para casos en que se reemplaza el clítico por un pronombre personal, tenemos la siguiente regla:

$S1 [L1, 'EMS', 'infin'] S2 [L2, 'EMS', 'prep'] S3 [L3, 'EMS', 'prpers'] \rightarrow S1+S2+S3 [L1+L2+L3, 'EMS', 'prspdesv']$.

Dar a vos

c) Para las estructuras en donde se utiliza un pronombre personal con caso objeto en lugar del clítico correspondiente, creamos otra regla:

$S1[L1, 'EMS', 'v', 'MODOV', 'infin'] S2[L2, 'TPRON', 'probj'] \rightarrow S1+S2[L1+L2, 'EMS', 'prbjdsv']$.

Dar ti

d) Para las estructuras en las cuales el pronombre personal objeto se halla antepuesto al infinitivo, creamos la siguiente regla:

$S1[L1, 'TPRON', 'obj'] S2[L2, 'EMS', 'v', 'MODOV', 'infin'] \rightarrow S1+S2[L1+L2, 'EMS', 'aprbjds']$.

Ti decir

Obtenemos los siguientes resultados :

'te enviar'.

['lo enviar', 'EMS', 'infcldesv'].

'dar a vos'.

['dar a vos', 'EMS', 'prspdesv'].

'dar ti'.

['dar ti', 'EMS', 'prbjds'].

'ti decir'.

['ti decir', 'EMS', 'aprbjds'].

5. CONSIDERACIONES FINALES

En este trabajo, Las descripciones provistas por la gramática tradicional fueron importantes ya que contribuyeron a ver el comportamiento de los pronombres átonos y poder confrontarlos con las desviaciones halladas.

Al analizar el corpus nos encontramos con la ausencia de clíticos en la posición antepuesta. Estos casos no pudieron ser formalizados para su posterior automatización, ya que las herramientas informáticas no nos permitirían evidenciar la falta de clíticos en determinadas construcciones verbales que deberían llevarlos. Con respecto a las construcciones en posición pospuesta, sólo se hallaron desviaciones cuando intervenía un verbo en infinitivo.

Se agruparon entonces, dichas construcciones en cuatro casos desviados respecto de la lengua estándar, para lo cual en primer lugar, debimos etiquetar los pronombres personales, distinguiendo el caso objeto. Para eso modificamos el archivo rasgos y el de entradas de Smorph. Luego creamos reglas de reagrupamiento en MPS para lograr la automatización.

Como proyección de la presente investigación, esperamos contar con un corpus más amplio que nos proporcione otras construcciones con clíticos propias de interlengua, las cuales puedan ser detectadas y analizadas mediante herramientas informáticas.

Referencias

- [1] Aït- Mokhtar Salah y Rodrigo Mateos José Lázaro 1995. “Segmentación y análisis morfológico de textos en español utilizando el sistema SMORPH”. SEPLN, 17, 29-41.
- [2] MPS ha sido especificado en el GRIL por Caroline Hagège, José Rodrigo, Gabriel G. Bès y Faiza Abacci, e implantado en C++ en un contexto Windows por Faiza Abacci.
- [3] Fernández Soriano, O, 1993. Los pronombres átonos. Gramática del español. Director: Ignacio Bosque. Taurus Universitaria.