UN ANALIZADOR AUTOMÁTICO ORIENTADO AL APRENDIZAJE DE LAS CONSTRUCCIONES VERBALES

María Concepción Álvarez Lebredo
Universidad de Santiago de Compostela

1. Introducción

En el proceso de aprendizaje de las destrezas comunicativas, especialmente de la expresión escrita, es necesario que el alumno extranjero lleve a cabo tareas intermedias que le permitan desarrollar su competencia gramatical. Uno de los aspectos más importantes de la gramática es la valencia o régimen verbal. Cuando utilizamos un verbo ponemos en juego al menos la siguiente información: a) número de argumentos con que se construye el verbo (valencia cuantitativa); b) funciones que desempeñan esos argumentos; obligatoriedad u opcionalidad de los mismos; cuál es la preposición regida, si la hay (valencia cualitativa); c) categorías gramaticales que pueden realizar esas funciones (rasgos de subcategorización); d) rasgos semánticos que el verbo exige a sus argumentos (restricciones de selección).
Para adquirir estos conocimientos y usarlos adecuadamente proponemos el apoyo de una herramienta informática con la que el alumno podrá formular sus propias hipótesis, comprobarlas, modificarlas en vista de los resultados y obtener muestras de lengua real que le permitan cotejar sus propias intuiciones y le ayuden a obtener una idea del uso de los verbos en español ajustada a la realidad.

2. Componentes y funcionamiento del sistema

La herramienta de la que hablamos consta de dos partes: un analizador *sintáctico (parser) y un diccionario electrónico. Ambas son complementarias y están* perfectamente integradas.

El parser es un analizador sintáctico sencillo que proporciona el análisis de construcciones verbales del español. Con él, el usuario puede hacer consultas sobre las secuencias lingüísticas que le planteen dudas. Cumple una doble función: a) función de reconocedor; b) función de analizador.

Centro Virtual Cervantes

Su primera dimensión se pone de manifiesto porque sólo admite construcciones correctas. Si la secuencia escrita por el usuario está mal formada o es incorrecta por algún aspecto, el analizador la rechaza. El usuario deberá hacer una nueva propuesta. Utilizaremos como ejemplo el verbo GUSTAR, verbo particularmente difícil para un extranjero porque la correspondencia entre funciones sintácticas y papeles semánticos es inversa a la mayoría de las lenguas (cf. ing. *I like something*; fr. *J'aime quelque chose*; esp. **yo gusto algo; (a mí) me gusta algo*)

(1) `yo gusto algo`
`all: 0`[1]
(2) `>>a él me gusta`
`all: 0`

El sistema rechaza (1) y (2) porque son agramaticales. En (1) no se han escogido las funciones sintácticas adecuadas. En (2) no existe la concordancia debida entre el COMPLEMENTO INDIRECTO (*a él*) y el clítico *me* y por eso no se obtiene análisis. El sistema no las reconoce como cadenas válidas en español.

La malformación también puede ser de tipo semántico, como en (3)

(3) `>> le gusta a algo`
`all: 0`

Esta secuencia también es rechazada por el sistema porque el verbo GUSTAR exige que el CI sea animado.

En cambio, si la secuencia está bien construida, tanto desde el punto de vista sintáctico como semántico, el sistema la admite y proporciona su análisis sintáctico en forma de estructura arbórea:

(4) `>> me gusta ella`

```
frase verbal (independiente, indicativo, primera)
  PREDICADO con OI (suj oi, animado subj, animado, tercera, singular,
indicativo, primera, singular)
    clitico (dativo, primera, singular, GENERO)
      "me"
    verbo (suj oi, animado, animado, tercera, singular, indicativo)
      "gusta"
  SUJETO (animado, tercera, singular)
    fn (animado, tercera, singular)
      "ella"
```

El analizador automático está construido sobre la base de una gramática formal escrita en el formalismo AGFL[2].

---

[1] El significado del mensaje `all: 0` significa que el *parser* no encuentra ningún análisis para esta secuencia.

[2] *AGFL* fue desarrollado en la Universidad de Nimega para escribir y generar gramáticas formales de

Centro Virtual Cervantes

Además de la información sintáctica, el usuario obtiene también la definición del verbo adecuada al contexto específico de la secuencia dada. De esta forma, puede comprobar que la frase verbal que él mismo ha construido se corresponde con el sentido esperado, y no con otra acepción diferente del mismo verbo. También le sirve para conocer el significado del verbo en esa construcción, en el caso de que lo ignore. Las definiciones nos las proporciona el segundo módulo del sistema: el diccionario electrónico, que contiene todas las acepciones y sentidos de los verbos más frecuentes del español.

Veamos cómo opera el sistema. La relación entre gramática y diccionario se establece mediante la línea que define al verbo. Un programa escrito en el lenguaje PERL lee en el análisis la forma verbal; la lematiza, es decir, obtiene su infinitivo; y busca en el diccionario electrónico la entrada correspondiente a ese verbo. Dentro de esa entrada selecciona la acepción o acepciones que respondan a las características del verbo tal como se muestran en el árbol de análisis. En el ejemplo (4) de más arriba veíamos la siguiente línea:

```
verbo (suj oi, animado, animado, tercera, singular, indicativo)
    "gusta"
```

El sistema obtiene a partir de la forma "gusta" el infinitivo; a continuación busca, en la entrada correspondiente a GUSTAR del diccionario electrónico, las acepciones que respondan a las características verbo (`suj oi, animado, animado`) que se leen en la pantalla. Es decir, busca los sentidos del verbo GUSTAR que se construyan con sujeto y complemento indirecto, cuyo primer argumento, el sujeto, sea animado, y cuyo segundo argumento, el c.indirecto, también sea animado. Estos son los resultados que se obtienen:

```
verbo (suj oi, animado, animado)

A alguien le gusta una persona, si siente atracción por ella, o le cae bien o le
resulta simpática o agradable porque coincide con su modo de pensar o actuar.

    <Tampoco a los hombres les gustaban las chicas con complejos.>
    <Me gustas, he vivido un verano en tus brazos, me resisto a  perderte.>
    <Estas señoritas que vienen a confesar los pecados de los demás, no me
gustan, señorita.>
    <Nos gusta bastante porque es de esas personas, que cuando hay que decir
algo, lo dice sin dar vueltas.>
    <De la lectura de los textos, se desprende que al autor de "El aleph" o
"Mirando hacia atrás con ira" no le gustaban los  directores alemanes ni los
soviéticos.>
```

Inmediatamente después de la definición el sistema ofrece un conjunto de ejemplos extraídos de ARTHUS (Archivo de Textos Hispánicos de la Universidad de Santiago) para que el aprendiz disponga de ejemplos del uso de ese verbo en la lengua real.

---

lenguas naturales. Para aproximarse a este sistema el lector puede consultar [Koster 91] y [Derksen 97].

Centro Virtual Cervantes

3. Algunos ejemplos más.

Otros verbos que pueden ofrecer dificultad al alumno extranjero son la pareja ACORDAR-RECORDAR. Estos verbos, en sus acepciones más frecuentes son prácticamente sinónimos pero exigen construcciones sintácticas diferentes:

alguien recuerda algo
alguien se acuerda de algo

El usuario del sistema puede comprobar si las suposiciones respecto a la sintaxis de estos verbos son las correctas. Puede darse el caso de que confunda las construcciones y que asigne a un verbo la estructura de otro. Por ejemplo, si el input fuese **acordamos algo*, en lugar de *recordamos algo* (cf. ing. *we remember something*), el sistema reconocería esa secuencia como válida en español. En este caso son las definiciones que siguen al análisis las que le indicarían al usuario que no se trata del sentido que esperaba.

```
(5)  >>acordamos algo
   frase verbal (independiente, indicativo, primera)
      PREDICADO (suj od, animado, inanimado, primera, plural, indicativo)
         verbo (suj od, animado, inanimado, primera, plural, indicativo)
           "acordamos"
      C DIRECTO (inanimado, tercera, singular)
         fn (inanimado, tercera, singular)
           "algo"

          all:  1

    Si varias personas acuerdan una cosa, llegan a un trato sobre lo que se
hará con esa cosa, adoptan una decisión común en relación con esa cosa,
deciden por mayoría el futuro de esa cosa.
```

Esta información se completa con ejemplos, como el siguiente:

```
    <Los vecinos acordaron el presupuesto. [GDLE]
```

Si el usuario confundiese las construcciones en el sentido contrario y escribiese como input **me recuerdo de algo* (cf. fr. **je me souviens de quelque chose*), no obtendría ningún análisis. El sistema no lo reconocería como cadena válida en español. Es agramatical y por eso la rechaza. Las construcciones que darán resultados satisfactorios son *recuerdo algo* y *me acuerdo de algo*. Las definiciones que nos ofrece el *output* corresponderán a la idea inicial de que partía el alumno y que pretendía expresar utilizando el español:

```
(6) >> recuerdo algo

  frase verbal (independiente, indicativo, primera)
    PREDICADO (suj od, animado, inanimado, primera, singular, indicativo)
       verbo (suj od, animado, inanimado, primera, singular, indicativo)
         "recuerdo"
```


Centro Virtual Cervantes

```
    C DIRECTO (inanimado, tercera, singular)
       fn (inanimado, tercera, singular)
         "algo"
```

Alguien recuerda a una persona, una cosa o un hecho si lo guarda en la memoria o si le viene repentinamente a la mente esa persona, cosa o hecho.

```
(7) >> me acuerdo de algo

frase verbal (independiente, indicativo, primera)
 PREDICADO(pron suj cp, de, animado, inanimado, primera, singular, indicativo)
   verbo (pronominal suj cp, de, animado, inanimado, primera, singular,
indicativo)
          "me"
           "acuerdo"
   C PREPOSICIONAL (de, inanimado)
      preposicion (de)
         "de"
      fn (inanimado, tercera, singular)
         "algo"
```

Alguien se acuerda de una persona, de una cosa o de un hecho si lo guarda en la memoria o si le viene repentinamente a la mente esa persona, cosa o hecho.

4. Las definiciones

Las definiciones que proporciona el diccionario electrónico pretenden ser lo más útiles posible para el estudiante no nativo. Comienzan siempre por una fórmula en la que se usa el verbo que se está definiendo tal como se hace en un contexto real:

Alguien se acuerda de algo, de alguien o de hacer algo si...

Alguien recuerda una cosa, persona o hecho si/cuando...

La fórmula retoma la construcción que el usuario ha propuesto al sistema (el *input*), a veces algo más completa. Esto refuerza sus hipótesis y le sirve de ayuda para fijar la construcción. Es un sistema semejante al que utiliza para el inglés el COLLINS COBUILD. En la segunda parte, la definición propiamente dicha, se hace referencia a los argumentos del verbo mediante pronombres tónicos, clíticos, posesivos...

... si *lo* tiene en la memoria.

... si guarda *su* recuerdo en la mente...

Además se procura que las definiciones no se limiten a una enumeración de sinónimos. A menudo los proporcionan, pero además contiene alguna paráfrasis con vocabulario básico, accesible al estudiante no nativo.

5. El análisis

El análisis también cumple una función en el proceso de aprendizaje. Pone en juego los conocimientos gramaticales del alumno encaminándole a adquirir una mayor competencia respecto al uso del verbo en cuestión.

Centro Virtual Cervantes

En (6) el análisis de *recuerda algo* revela la existencia de un C. DIRECTO. La observación del *output* puede poner en marcha nuevas hipótesis acerca de las posibilidades de pronominalización del verbo RECORDAR. El alumno que sepa que un C.DIRECTO puede ser sustituido por un clítico acusativo podrá escribir en la línea de comandos *lo recuerda* y obtendrá resultados satisfactorios. En cambio, el sistema no admitirá **se lo acuerda* y sólo aceptará *lo acuerda* con un sentido muy diferente al del verbo RECORDAR.

Otros aspectos gramaticales que puede ensayar el usuario son la clitización del complemento indirecto; las concordancias (sujeto y predicado; dativo y c.indirecto...); elipsis del SUJETO y de otros actantes no obligatorios, etc.

6. Los ejemplos

Pretenden ser lo más representativos dentro de lo posible. Por ello, en muchas ocasiones se buscan ejemplos que contengan los complementos más frecuentes con los que aparece el verbo. También se intenta que sean variados, que no procedan todos de la misma obra o autor. Otro requisito es que no sean complejos, se seleccionan ejemplos sencillos, si es posible.

Los ejemplos se extraen de ARTHUS. En el caso de que este corpus no contenga ejemplos para una acepción que se considera frecuente, o los ejemplos sean demasiado complejos o insuficientes para ilustrar el sentido que se está considerando, recurrimos a los ejemplos de los diccionarios. Se usan cinco diccionarios: DUE, CLAVE, GDLE, SALAMANCA y DALE (este último en menor medida porque en raras ocasiones contiene ejemplos con el verbo conjugado, casi todos los ejemplos llevan en verbo en infinitivo).

7. Características del sistema

7.1. Ventajas

I. el protagonista del proceso de aprendizaje es el alumno. Tiene la ocasión de comprobar sus propias hipótesis y corregir sus errores por su cuenta en pasos sucesivos.

II. el proceso es interactivo. Esto conduce a un aprendizaje significativo. En este sentido el sistema aventaja como utensilio de consulta a los diccionarios convencionales en varios aspectos:

a) más fácil consulta porque el propio sistema selecciona del conjunto total de acepciones y sentidos del verbo las que corresponden a la secuencia propuesta por el usuario.

b) la información gramatical no está expresada mediante abreviaturas ni otros símbolos de difícil comprensión, sino usando el verbo en la propia definición mediante las fórmulas ya comentadas. Estas contienen todo lo que el alumno necesita saber, expresado de la forma más natural: número de actantes; función

ASELE. Actas IX (1998). MARÍA CONCEPCIÓN ÁLVAREZ LEBREDO. Un analizado...
Centro Virtual Cervantes

de los mismos; categoría gramatical y rasgos semánticos.

c) los ejemplos son reales frente a lo que ocurre en la mayoría de los diccionarios y gramáticas, donde se confeccionan ejemplos *ad hoc* para ilustrar una determinada acepción o concepto gramatical. Los ejemplos ayudan a que el alumno se familiarice con la lengua tanto coloquial como literaria.

III. puede servir de complemento a las prácticas de aula, reforzando lo aprendido por el alumno en ella en lo referente a las destrezas comunicativas y, en menor medida, la producción oral.

7.2. Problemas

Quizá la desventaja más clara son sus propios límites. El sistema no puede abarcar todas las dimensiones del enunciado lingüístico. La gramática no es un área *independiente, sino que es inseparable de la pragmática, de la cohesión textual o* fluidez del discurso (elisión de elementos consabidos, enlaces extraoracionales...), de aspectos socioculturales, etc. Estos elementos no están presentes en esta herramienta. Si se utiliza para fines pedagógicos, como proponemos, hay que tener en cuenta ese hecho. Por ello, no se debe desligar del trabajo del aula. Su utilidad será mucho mayor si está guiada de forma prudente por el profesor que deberá comentar y analizar los resultados, aclarar los ejemplos, etc. No se trata de ningún método de aprendizaje gramatical. Se trata, como hemos dicho, de una herramienta, un utensilio más que puede utilizar el aprendiz de español. En este sentido, se sitúa en la línea de las obras de consulta, como los diccionarios convencionales, pero pretende superarlos, aportando ventajas y aspectos nuevos. Concretamente, pretende ser más completo, más cómodo y estimulante.

Dentro de lo estrictamente gramatical, también falta en este sistema todo aquello que rebasa el ámbito de la valencia o régimen verbal, como los tradicionalmente llamados CIRCUNSTANCIALES, la negación, las cláusulas relativas... De ahí que el profesor deba incidir en estos aspectos para no dar lugar a *confusiones.*

La selección de los ejemplos es otra tarea del profesor. Los que hemos seleccionado inicialmente para el sistema no pasan de ser una propuesta inicial, basada el el corpus con que contamos. Pero quizá no todos sean adecuados para el aprendizaje de la lengua, o quizá sólo lo sean para un nivel avanzado. Para niveles inferiores, se deben buscar ejemplos alternativos adaptados al grado de familiaridad de la lengua de un alumno o grupo de alumnos.

7.3. Límites

Hemos comentado en el apartado anterior algunas de las limitaciones más sobresalientes del sistema. También es necesario señalar que se trata de un trabajo en curso, no es una herramienta terminada. La primera fase del trabajo acabará cuando se completen tanto en la gramática como en el diccionario todos los aspectos que se refieren a los 100 verbos más frecuentes del español (según los datos de ARTHUS). En este momento, no estaremos ante un producto final; será sólo un

Centro Virtual Cervantes

"pre-producto", que no se podrá utilizar todavía con alumnos, a no ser en pequeños experimentos o ejercicios muy concretos. Cuando se acabe esta fase del trabajo, será un objeto de investigación que quizá pueda dar frutos interesante en el terreno de la enseñanza de E/LE en un plazo no demasiado largo.

REFERENCIAS

CLAVE: *Diccionario de Uso del Español Actual*, Madrid, Ediciones SM, 1996

COLLINS COBUILD: *English Language Dictionary,* London, Collins, 1987.

DALE*: Diccionario Actual de la Lengua Española*, Barcelona, Bibliograf, 1993.

Derksen,C. (1997) *Manual for the AGFL System.* (Disponible en la dirección www.cs.kun.nl.agfl.Papers.html)

DUE: María Moliner, *Diccionario de Uso del Español*, Madrid, Gredos, 1996.

GDLE: *Gran Diccionario de la Lengua Española*, Barcelona, Larousse Planeta S.A., 1996.

Koster, C.H.A. (1991), "Affix Grammars for Natural Languages", *in Attribute Grammars, Apllications and Problems*. International Summer School SAGA, Praga, Junio 1991. *Lecture Notes in Computer Science,* vol. 545, Springer Verlag. (Disponible en la dirección http://www.cs.kun.nl/agfl/Papers.html).

SALAMANCA: *Diccionario Salamanca de la Lengua Española*, Universidad de Salamanca, Santillana, 1996.

Centro Virtual Cervantes