

Tendencias en la investigación lexicográfica del español

El diccionario como objeto
de estudio lingüístico y didáctico

Stefan Ruhstaller
Josefina Prado Aragonés
(eds).



EL MÉTODO LEXICOGRÁFICO DE GASTON GROSS

JOSE LUIS HERRERO INGELMO
Universidad de Salamanca

*Como puedes manejar las palabras
desactivarlas solazarte en ellas
como puedes dejar que las palabras
se evadan de tu corazón empecinado*
Mario Benedetti, *El olvido está lleno de memoria*

*Los diccionarios impresos no van a dejar de existir de
un día para otro, ni quizá jamás en vista de lo barato y
duradero que puede ser un libro. En cambio las ventajas
cada vez más claras de un diccionario dinámico infor-
matizado, harán que se prefieran los datos recuperados de
un banco de datos en constante evolución y expi estos en
una pantalla, y no lo fijado para siempre en papel*

W. Steven Dodd (Universidad de veter

1. EL LABORATORIO DE LINGÜÍSTICA INFORMÁTICA: EL PROYECTO DE UN DICCI- ELECTRÓNICO DE GASTON GROSS

Esta comunicación tiene como finalidad dar a conocer los fundamentos teóricos lexicográficos de Gaston Gross y su aplicación metodológica. El profesor Gross dirige el Laboratorio de Lingüística Informática de la Universidad de París XIII, compuesto por un equipo de léxico y gramática, un Centro de Terminología y Neología (CNT) y un Laboratorio de Automática Documentaria y Lingüística (LADL)¹.

1. Es un grupo de investigación diversificada, pero que tiene como objetivo común el estudio del léxico en el sentido amplio: en sus relaciones con la sintaxis, la terminología, la neología, el argot. La utilización de la informática es común a todos estos equipos que elaboran bases de

El primer grupo, que es el que ahora nos interesa, tiene dos objetivos de investigación: 1) El desarrollo de modelos lingüísticos apropiados para el tratamiento automático, en particular de la noción de CLASE DE OBJETOS (teoría); 2) La elaboración de bases de datos lexicográficas concebidas, desde su planteamiento inicial, en la perspectiva de un tratamiento automático de la lengua –diccionarios electrónicos– (práctica). Un proyecto de “longue haleine”... El objetivo último es la elaboración de un diccionario electrónico del francés que se pueda utilizar para la traducción automática y que reúna, bajo la forma de una base de datos, el conjunto de las propiedades sintácticas y semánticas necesarias para la automatización de la lengua.

El profesor Gross señala que existe un cierto pesimismo, en los últimos años, en torno a la posibilidad de la traducción automática (parece que hay como una resignación a la traducción asistida). No hay que olvidar que la traducción (TA o TAO) no es más que el paso de una lengua *natural* o otra lengua *natural*. Es cierto que la descripción integrada de hechos sintácticos, semánticos y léxicos necesita mucho tiempo de investigación (grupos) y muchos recursos técnicos (financiación fundamentalmente asumida por organismos oficiales). Como es sabido, entre los muchos problemas que plantea la traducción automática, hay dos especialmente complejos: las EXPRESIONES FIJAS (*figement*) y la POLISEMIA, que son dos de las propiedades más importantes de las lenguas naturales.

En cuanto al primero, LAS EXPRESIONES FIJAS, la tradición gramatical lo relegaba a un lugar secundario dentro del capítulo sobre las palabras compuestas. Pero en el planteamiento de un diccionario electrónico es fundamental establecer estos tipos de expresiones, puesto que las entradas y la traducción no pueden basarse ya en la palabra aislada, como sucede en la lexicografía convencional. Para el francés se pueden establecer unos 500 tipos de expresiones fijas de sustantivos (varios centenares de miles de palabras), 250 tipos de adjetivos compuestos y unas 35.000 expresiones verbales³.

datos lingüísticas con varios fines: bases de datos léxicos con vistas a la constitución de diccionarios tradicionales o electrónicos; preparación de bases de datos neológicas y terminológicas que tienen como finalidad, entre otras, prestaciones de asesoramiento en ciertas instituciones oficiales nacionales (Delegación General de la Lengua Francesa etc.) e internacionales (Red Internacional de Neología y de Terminología –RINT).

2. Se está trabajando en la traducción al inglés y al alemán. El Centro de Investigaciones Lingüísticas de la Universidad de Salamanca (CILUS) se suma a este proyecto en lo referente a la traducción al español. Este Centro, creado en 1995, trabaja en la creación de materiales lexicográficos y de enseñanza del español como segunda lengua en soporte informático.

3. En la gramática española, las expresiones fijas no han tenido un desarrollo sistemático (Vid., A. Zuloaga, “La fijación fraseológica” *Thesaurus* 30 (1995), pp. 225-248; J. Martínez Marín, “Las expresiones fijas en español: perspectivas teórico-descriptiva y aplicada”, *Gramma-temas* 1, Univ. de León, 1992, pp. 181-196). No existe en español un diccionario como el *Oxford Dictionary of Current Idiomatic English* de A.O. Cowie et alii –sólo en el ámbito de la lingüística aplicada (enseñanza del español como segunda lengua) se han publicado algunas monografías, como la de P. Domínguez, *El español idiomático. Frases y modismos del español*, Ariel, Barcelona, 1988. Uno de los trabajos iniciales para el español es el de sistematizar y censar los tipos de expresiones fijas.

El otro gran problema de la traducción automática es la POLISEMIA. Los diferentes significados de una palabra están condicionados por el contexto sintáctico: las tablas sintácticas construidas en términos de predicados y argumentos por Maurice Gross (1975) intentan dar cuenta del comportamiento sintáctico que el significado de las palabras condiciona. En esas tablas, la semántica está reducida a la subcategorización de los sustantivos sobre la base de rasgos como ‘humano’, ‘concreto’... Son rasgos indispensables, pero –como veremos– en número restringido y con un carácter vago. Estas dos limitaciones pueden superarse añadiendo nuevos rasgos y proponiendo un instrumento más preciso. La aportación teórica fundamental de Gaston Gross al ámbito de la investigación lingüística es el concepto de *clase de objetos*.

2. FUNDAMENTOS TEÓRICOS: DE LA FRASE A LA “CLASE DE OBJETOS”: PREDICADOS (OPERADORES) Y ARGUMENTOS; VERBOS SOPORTE; RASGOS SINÁCTICOS Y CLASES DE OBJETOS; PREDICADOS GENERALES Y ESPECÍFICOS

Antes de entrar en la descripción de los conceptos teóricos que maneja, vamos a citar brevemente la línea de pensamiento en la que se insertan los trabajos del lingüista parisiense⁴. Como telón de fondo ideológico, están los trabajos de la filosofía analítica (Frege, Putnam), que ponderan, como base de su reflexión, el principio de *contextualidad* o *contextualización*: las palabras no tienen sentido (ni referencia) fuera del contexto del enunciado en el que ellas figuran. Es decir, el significado de una palabra puede formularse en términos de su contexto lingüístico, lo que lleva a considerar que la unidad de descripción mínima no es la PALABRA O EL SIGNO en el sentido de Saussure, sino la FRASE U ORACIÓN.

En el ámbito específicamente lingüístico, el profesor Gross parte de los trabajos de Z.S. Harris (1951 y 1976), sobre todo de su modelo de oración elemental, y de los de Maurice Gross (1975) sobre la noción de empleo y el desarrollo de un léxico-gramática: el léxico no puede ser tratado de manera informática si no se dota a cada elemento de un comportamiento previsible en una oración, comportamiento calculado en términos de contexto. Esto tiene como consecuencia que las descripciones no pueden constituir niveles diferentes, sino que deben estar integradas, con vistas al reconocimiento del empleo y a la resolución de las ambigüedades.

Predicados (operadores) y argumentos

El modelo de oración elemental de Z.S. Harris responde al esquema PREDICADO (OPERADOR) / ARGUMENTOS, que parece más operativo, desde el punto de

4. Utilizamos, para la exposición del soporte teórico y del método, la bibliografía que recogimos al final, especialmente el documento interno del LLI titulado “Structuration d’une base de données en vue de l’élaboration d’un dictionnaire bilingue”.

vista léxico, que los cortes binarios heredados de la lógica clásica (sujeto/predicado; tema/rema)⁵:

| | | | | | |
|--------------------------------------|--------|---------|--------------|-----------|---|
| P (x) | dormir | P (x,y) | mirar | P (x,y,z) | d |
| PREDICADOS ⁶ (OPERADORES) | | | ARGUMENTOS | | |
| <i>dese-</i> | | | (N0, phrase) | | |
| <i>desear</i> | | | (N0, phrase) | | |
| <i>deseo</i> ⁷ | | | (N0, phrase) | | |
| <i>deseoso</i> | | | (N0, phrase) | | |

Lo que produce la unidad y la coherencia de la descripción es el hecho de proponer la oración como unidad de base. Esta oración está compuesta de un OPERADOR ACOMPAÑADO DE 5 ARGUMENTOS (definidos semánticamente) y esta combinación está caracterizada por particularidades aspectuales y propiedades transformacionales que le son propias. Pero es un nivel muy abstracto para dar cuenta de oraciones reales: el esquema *admirar* = V (N0, N1) o N0 V N1, así formulado, puede dar lugar a frases aceptables: "Juan admira a María" o a frases aberrantes: *"La explosión admira a María" o *"El bolígrafo admira los fuegos artificiales". La solución (Chomsky 1965) está en la caracterización de los argumentos en términos de rasgos sintácticos semánticos⁸, unidos al predicado por restricciones de selección: (*admir-* N0hum; *comer* N0hum, N1 concreto)⁹.

5. La oposición *sujeto/predicado, tema/rema* establece una preeminencia del sujeto sobre los complementos en la medida en que sitúa a estos últimos en un nivel inferior, bajo el nudo del grupo verbal. Ahora bien, los estudios que se han llevado a cabo en el marco de la gramática transformacional han demostrado que los complementos aportan muchas más informaciones para la solución de la ambigüedad que el sujeto

6. Estos predicados son predicados de primer nivel por oposición a los conectores o relativos entre oraciones, que se pueden considerar como predicados de segundo nivel. De cualquier modo, en esta apretada síntesis, no entraremos en su consideración

7. Los sustantivos tienen argumentos elementales (*María recoge flores* no susceptibles tener argumentos-; y argumentos no elementales (*El general ordena la destrucción del puente*).

8. Se utiliza indistintamente "rasgos sintácticos" "rasgos semánticos" o "rasgos semántico-sintácticos".

9. En los últimos cuatro lustros, han surgido varios modelos de análisis lingüístico en los que el léxico, descrito en función de las características sintagmáticas se ha convertido en el componente central (verbo, palabra clave): la gramática léxica de Hudson la gramática funcional de S.C. Dik, la gramática de estructura sintagmática de control nuclear de Pollard y Sag, la gramática de base unificada y la teoría de la rección y el ligamento. Todos estos modelos tienen como característica común la construcción de las oraciones no a partir de un esquema abstracto, ya sea categorial o funcional, sino de la extracción de un verbo del léxico que trae consigo los elementos con los que se combina, siendo la combinatoria una característica idiosincrásica del léxico verbal de las lenguas individuales.

Los verbos soporte

Al igual que los verbos, los sustantivos y adjetivos predicativos son operadores. Su estructura argumental (operador arg1, arg2, arg3) necesita dos operaciones para convertirse en una frase real: la LINEALIZACIÓN, que coloca correctamente los argumentos de una parte y otra del operador, y la ATRIBUCIÓN, que —en el caso de los verbos— consiste en añadir un morfema temporal. Esta posibilidad no existe para los sustantivos y adjetivos, que no tienen una conjugación morfológica. Por el contrario, un adjetivo es actualizado por un verbo copulativo, un sustantivo predicado es actualizado por verbos predicativamente vacíos que conjugan, por así decir, esos sustantivos. Se trata de verbos como *faire, avoir, etre, etre+ prep, donner, recevoir*, etc., llamados VERBOS SOPORTE ("Luc a décrit la ville" "Luc a fait la description de la ville", "Luc admire cette attitude" "Luc a de l'admiration pour cette attitude", "Luc voyage en Italie" "Luc est en voyage en Italie", "Luc a giflé Max" "Luc a donné une gifle à Max" "Luc a été giflé par Max" "Luc a reçu une gifle de Max").

La introducción del concepto de VERBOS SOPORTE¹⁰ en la descripción de las lenguas es importante, sobre todo, para la traducción. Por una parte, los verbos soporte permiten explicar y resolver la polisemia de los operadores nominales. Tomemos la palabra *charge*. Sus diferentes significados pueden explicarse gracias a los diferentes soportes que se pueden emplear. "Luc a donné à Max la charge de faire ce travail", "Luc a la charge de faire ce travail", "Ce travail est une charge pour Luc", "Luc est en charge de ce travail", "Cette batterie est en charge", "Une notaire a une charge". Por otra parte, la construcción de verbos soporte da cuenta de la constitución de los grupos nominales y, concretamente, de la naturaleza de la preposición: "Luc a du respect pour cette attitude", "Le respect de Luc pour cette attitude", "Luc est en admiration devant cette attitude", "L'admiration de Luc devant cette attitude", "Luc porte un jugement sévère sur Max", "le jugement sévère de Luc sur Max", "Luc a fait un salut à Max", "le salut de Luc à Max".

Gross resume así las principales propiedades de los verbos soporte:

1) actualizan los predicados nominales¹¹.

10. Este concepto tiene su origen en la obra de Harris que los denomina verbos que operan en la nominalización de los verbos

11. No tienen por ellos mismos función predicativa en razón del principio de que no puede haber dos predicados en una oración simple. Se desprende de lo que acabamos de decir que el verbo soporte puede ser borrado de una oración sin que ésta pierda su *status* de oración. Sólo la actualización estará ausente. Esta reducción opera después de la relativación: "Lu a donné une réponse à cette question", "la réponse que Luc a donné à cette question" "la réponse de Luc à cette question". Por el contrario, la supresión de un verbo predicativo supone *de facto* la de la oración, porque no queda más que una supresión de los sustantivos, en el caso de un predicado verbal con dos argumentos.

- 2) No pueden ser el objeto de un cambio de categoría (nominalizaciones, adjetivaciones, verbalizaciones), que es un rasgo propio de los predicados. Los soportes *être, faire, avoir* no tienen forma nominal¹².
- 3) Como los verbos soportes actualizan a los predicados nominales, ellos asumen además las informaciones aspectuales que les conciernen.

Los rasgos sintácticos organizados en las clases de objetos

El lingüista francés ha añadido a la lista de rasgos que se manejan habitualmente el rasgo sintáctico de ACONTECIMIENTO. Tenemos, pues, los siguientes: humano <hum>, animal <ani>, vegetal <veg>, inanimado concreto <concr>, inanimado abstracto <abst>, locativo <loc>, tiempo <temp>, acontecimiento <acon>¹³. Cada uno de estos rasgos se relaciona con predicados u operadores muy generales. Los rasgos sintácticos de selección no son suficientes para delimitar el dominio de los argumentos de un operador¹⁴.

Las informaciones aportadas por los rasgos sintácticos son útiles: con la ayuda de los rasgos *concreto* y *abstracto*, pueden explicarse las diferencias entre “llevar un cubo” y “llevar un negocio”¹⁵. La traducción inglesa de empleo concreto es *to carry*. Sin embargo, esta descripción es demasiado amplia: un *vestido* es un concreto tanto como *cubo*, pero *llevar* se traduce

12. Cuando una forma verbal dada es susceptible de nominalización lo es en su empleo predicativo y no en el de verbo soporte. Así, *donner* es predicativo con un complemento concreto *Luc a donné une pierre précieuse à Lea* y puede estar nominalizado *le don d'une pierre précieuse est un geste symbolique*. Por el contrario, con un predicado nominal el cambio morfológico es imposible: **Luc a donné un conseil à Paul*; **le don d'un conseil n'est que de l'hypocrisie*.

13. Sobre este último rasgo, Gross argumenta lo siguiente: desde este punto de vista *caillou* (piedra) es un concreto exactamente igual que *olor* o *ruido*: dependen respectivamente de la vista, del olfato o del oído. Si el objetivo de estas nociones es solucionar la ambigüedad y seleccionar la buena traducción de los operadores entre otros posibles, se ve inmediatamente que esta definición de naturaleza filosófica no nos es de gran ayuda. En efecto, el sustantivo *olor* puede figurar con un gran número de verbos que no caracterizan los concretos ‘evidentes’ como *piedra* (‘Un olor agrio se eleva de los rosales’, ‘Este objeto exhala un olor nauseabundo’, ‘Emana un buen olor este plato’, ‘Este ha provocado un mal olor’, ‘Hay un buen olor en esta cocina’, ‘Un olor extraño subía por las escaleras’). Verbos como *hacer*, *provocar* no pueden aplicarse a concretos, puesto que son específicos de nombres de ACONTECIMIENTO. Este análisis es confirmado por la posibilidad de ciertos adjetivos que se aplican también no a concretos, sino a ACONTECIMIENTOS: *olor sibilante, fugitivo, reiterado, persistente*, etc. (“Structuration...”, p. 5)

14. Considerados insuficientes, los rasgos sintácticos que sirven para caracterizar a los sustantivos en el léxico-gramática de los verbos han sido completados por otras propiedades. Por ejemplo, el recurso a los nombres clasificadores léxicos o morfológicos que pueden representar toda una clase de complementos para un verbo dado (EJEMPLO: CLASIFICADOR LÉXICO de los nombres de deportes (*María practica deporte [esquí, tenis...]*: *El esquí [tenis] es un deporte* [a priori, limitados]); ADORNO-CLASIFICADOR MORFOLÓGICO de todos los nombres que pueden aparecer en la misma posición (*Pedro adorna la casa con [adornos] [guirnalda, flores, fotos...]* [a posteriori, ilimitados] *mujeres empujadas* [dibujo]). Cf. J. Giry Schneider, ‘Sélection et sémantique. Problèmes et modèles’, *Langages*, 115, pp. 5-14.

15. Lo mismo sucede en “Juan toma un te...” “Juan toma una descripción”

por *to wear* (“Lleva un vestido”, “He wears a jacket”). Si el sustantivo *vestido* no llevara más que el código *concreto*, como por ejemplo, *cubo*, el sistema no estaría en condiciones de seleccionar la buena traducción. Para ello, es necesario añadir el rasgo *vestido* (por lo tanto, dentro del rasgo sintáctico *concreto*, habrá una serie de palabras con el rasgo *vestido*). Llamamos a estas INFORMACIONES SEMÁNTICAS, ESTABLECIDAS SOBRE LA BASE DE COMPATIBILIDADES ENTRE PREDICADOS (OPERADORES) Y ARGUMENTOS, CLASES DE OBJETOS. Este código, anotado para cada elemento de la clase, permite a un analizador prever el reconocimiento (o la generación) del empleo y su buena traducción.

Cada clase identificada da lugar a la elaboración de varios ficheros complementarios, que recogen los argumentos (sustantivos comunes y propios) y los predicados apropiados (verbos, adjetivos, sustantivos predicativos). Al final, veremos un ejemplo concreto con la clase de objetos *voies*.

Predicados generales y específicos

La existencia de dos niveles de informaciones (el de los rasgos y el de las clases de objetos) presenta una gran ventaja si se quieren hacer activas las informaciones del diccionario. En efecto, estos dos niveles no seleccionan los mismos tipos de predicados. Por ejemplo, los sustantivos marcados con el rasgo *concreto* comparten todos un número determinado de predicados, aquellos –entre otros– que tienen el rasgo de *color, peso, materia...* Se les puede llamar PREDICADOS OPERADORES GENERALES. Se aplican por igual a términos como *martillo* o *autobús*. La traducción automática ha demostrado que estos operadores generales no plantean verdaderos problemas de reconocimiento (corresponden a los rasgos sintácticos).

No sucede lo mismo con los PREDICADOS ESPECÍFICOS de las clases de objetos. El recuento de los predicados específicos u OPERADORES APROPIADOS constituye la puesta a punto de gramáticas locales, propias de cada clase de palabras. El verbo *manejar* se aplica a un *concreto* de la clase de *utensilios*¹⁶ y no a la *de medios de transporte*. Esta última, a su vez, tiene predicados que le son propios y que no se reducen a su propiedad de *concreto*: *ver* o *comprar* se pueden aplicar tanto a los *medios de transporte* como a los *alimentos* o los *libros*¹⁷.

16. No se consideran aquí los empleos metafóricos como “*maneja* conceptos que no comprende”.

17. “*Prendre* es diferente y puede ser parafraseado por *aller en* (‘J’ai pris le train pour y aller’, ‘J’y suis allé en train’). *Prendre* es un operador específico de *los medios de transporte* y es sinónimo de *emprunter* y no de *saisir*, como un concreto ordinario. *Voiture* designa otro medio de transporte y se construye también con *prendre*: ‘Je prendrai la voiture’. Pero en este caso la sinonimia no opera ya con *emprunter*. Debemos, pues, desagrupar: medio de transporte individuales y medios de transporte en común. Estas dos clases tienen una sintaxis diferente: la red sinónima no es la misma, como se acaba de ver y los transportes en común tienen predicados específicos: ‘Cet autobus, (*cette voiture) dessert la capitale’; tienen complemento de tiempos específicos: ‘l’autobus de 8h47’, ‘*la voiture de 8h47’; tienen complementos locativos específicos: ‘l’autobus de Marseise’ / ‘*la voiture de Marseille’” (“Structuration...”, p. 7).

Parece claro, pues, que la noción de clases de objetos compete a la vez a la semántica, a la sintaxis y al léxico. Para cada uno de los ocho rasgos sintácticos que hemos analizado más arriba, hay que establecer la lista de clases de objetos necesarias para la descripción monosémica de los predicados. Este trabajo no está acabado para todos los rasgos. Se piensa que para la descripción automática del francés debe “censarse” un millar de clases. Gross comenta que es un trabajo inmenso, pero no es una utopía: actualmente se tienen 300 clases y unas 150.000 palabras. Como ejemplo, presentamos las clases de objetos humanos, cuyo “censo” es más o menos definitivo.

- adepto: protestante taoísta
- edad colectiva: juventud, vejez
- colectivo: muchadumbre, cuadrilla <col>
- defecto físico: deforme cojo <défphy>
- defecto moral: mentiroso
- escrito: libro, Biblia
- ser mítico: sifide
- grado: capitán
- instrumentista: pianista violinista
- lugar de habitación: casa
- locativo por metonimia: calle
- lugar de espectáculo: cine, teatro
- enfermo físico: tuberculoso
- movimiento político: maoísmo
- predicado dativo: beneficiario
- partidario filosófico: platonico, racionalista
- personaje literario: Romeo
- declaración: frase discurso
- cualidad moral: piadoso, virtuoso
- cualidad psíquica: equilibrado
- pariente: padre, hermano
- religioso: bernardino, cura
- religioso (función): vicario, cura
- soldado: zuavo
- deportista: futbolista
- edad: niño, viejo
- apelativo: señor, sire
- defecto intelectual: imbecil, estúpido
- defecto psicológico: perezoso, nervioso
- doctrina: catolicismo
- etnia: Sioux
- función: diputado, presidente
- humano (propriamente dicho): hombre, gente, tipo
- lugar de culto: iglesia, templo
- locativo: parisino, alemán
- lugar escolar: escuela, facultad
- lugar de trabajo: oficina, fábrica
- enfermo psíquico: neurótico, loco
- predicado activo: fumador, conductor
- predicado pasivo: pensionista, elegido
- partidario político: comunista, socialista
- profesión: carpintero, chófer
- cualidad intelectual: inteligente, sagaz
- cualidad física: rápido, forzudo
- relación: vecino, condiscípulo
- relación colectiva: familia
- religioso (apelativo): padre, monseñor
- religioso (título): prelado, arzobispo
- especialista: jurista, espeleólogo
- título: duque, príncipe

Gaston Gross señala una serie de aplicaciones interesantes del concepto de clase de objetos:

- a) Es necesario para la definición de los argumentos en el reconocimiento y la generación de los predicados.
- b) Las clases de objetos son el lugar del tratamiento de la sinonimia. Por ejemplo, “tomar <alimento>: comer”, “tomar <bebida>: beber”; “grave <enfermedad>: seria”, “grave <nota>: baja”, “grave <voz>: profunda”...
- c) Uno de los objetivos de estas búsquedas es la traducción automática. Como los elementos léxicos forman parte de clases no “disjuntas” (un *bosque* puede ser considerado como un *locativo* o como un *colectivo de vege-*

tal) las clases constituyen una descripción más flexible que los árboles de descripción semánticos cuyos grupos son por definición “disjuntos”.

- d) La noción de operador apropiado permite señalar lo que, en una lengua, es idiosincrásico y no puede ser deducido por el conocimiento de otra lengua. En francés, los cambios de dirección (*viraje*, etc.) ‘se négocient’. Esos modismos deben ser sistemáticamente buscados, por que aquí se originan gran parte de los errores en la traducción.
- e) Como todos los sustantivos figuran en el diccionario con su rasgo sintáctico, su clase de objetos y su dominio, se puede establecer un sistema de documentación automática
- f) Finalmente se puede considerar esta base de datos como la materia prima de un diccionario papel.

3. EL MÉTODO: ESTRUCTURA DE UN DICCIONARIO ELECTRÓNICO

El tratamiento de las lenguas naturales bajo la forma de base de datos se ha convertido en estos últimos años en una necesidad que se ha impuesto a todos. La estructuración de las informaciones presenta grandes ventajas. El número constante de campos permite describir los elementos léxicos de manera unificada y coherente. Se tiene la garantía de este modo de disponer, para todas las unidades léxicas del mismo tipo, de las mismas informaciones. Las comparaciones son posibles y estamos en condiciones de proceder a controles que garanticen la exhaustividad y la uniformidad.

3.1. La estructura de una base de datos de los argumentos sustantivos

Λ) Campos morfológicos¹

El primer campo está constituido por la entrada, ya se trate de un nombre simple o de un compuesto. Tenemos a continuación los siguientes campos.

CAMPO M representa la *forma morfológica*. En particular se anota, para los nombres compuestos, el tipo del que se trata: Nadj., N de N., AdjN, VN, V Dét N, etc. En francés se han recopilado más de 500 clases¹⁹.

CAMPO G indica la *categoría gramatical*. Se señala la categoría y las indicaciones de género y número, ex.: nm, nf, nms, nmp, nfs, nfp²⁰.

¹⁸ Todos los sustantivos con argumento o predicados comparten las mismas propiedades morfológicas

¹⁹ Esta anotación es importante para el reconocimiento de los compuestos en los textos. En efecto, para un sistema informático una palabra es definida como una serie de caracteres separados por dos blancos. Esta definición excluye series discontinuas como los nombres compuestos. Es, pues, indispensable anotar los blancos que no son rupturas sintácticas

²⁰ Figura así para ciertos nombres el carácter obligatorio de plural (*aguas residuales les eaux usées*) o de singular (*utilité a boveda celeste*). Son lo tradición, les *pluralia tantum*

- CAMPO F describe las *variaciones morfológicas* producidas por la flexión. Los nombres comprenden en francés aproximadamente 20 tipos de flexiones diferentes (anotados, por ejemplo, N1, N2, etc.)²¹.
- CAMPO V recoge las *variantes gráficas*.

b) *Campos sintáctico-semánticos (vid. supra: Los rasgos sintácticos organizados en las clases de objetos.):*

- CAMPO T. *Rasgos sintácticos*.
- CAMPO C. *Clases de objetos*.

c) *Campos sociolingüísticos*

- CAMPO D informa del *dominio* en el cual la palabra en cuestión toma su significación²².
- CAMPO R anota *los niveles de lengua* como familiar, popular, literaria, arcaísmo, regional...; en resumen, todas las informaciones necesarias para una buena utilización del elemento léxico en cuestión.

d) *Campos que conciernen a la traducción (multilingües)*

- CAMPOS E Y D reciben las traducciones inglesa y alemana²³.

3.2. La estructura de una base de datos de los predicados²⁴

Presentamos ahora la estructura de una base de datos que permite describir los PREDICADOS VERBALES, NOMINALES O ADJETIVOS. Como para los argumentos, hay un conjunto de campos que conciernen a la morfología (cuya estructuración depende por supuesto de la naturaleza del predicado). Vienen a continuación los campos que llevan la indicación de dominio de argumentos (de su naturaleza semántica). Estos últimos se describen con la ayuda de rasgos sintácticos y, cuando son insuficientes, con la indicación de clases de objetos. Los campos siguientes son dedicados a la sinonimia y a la antonimia. Las últimas secciones indican los dominios y registro, las transformaciones características del empleo así como la traducción.

Precisaremos ahora la forma particular de la base en lo que concierne a cada tipo de predicado.

1. Gaston Gross sigue los tipos de flexiones que han sido elaboradas en el LADL (en particular en los diccionarios DELAS y DELAF). Estas indicaciones son necesarias para el reconocimiento y la generación automática de todas las formas nominales que aparecen en el discurso.

22. Gaston Gross y su equipo han elaborado, mediante una comparación sistemática de todos los diccionarios disponibles, una lista de aproximadamente 500 dominios.

23. Esperamos que pronto el grupo de investigación del CILUS esté en condiciones de añadir a estos campos de traducción la del español.

24. Incluimos los verbos y los sustantivos. Los adjetivos tienen una estructura semejante a la de estos últimos.

3.2.1. *Predicados verbales*

Se incluyen sólo los verbos predicativos²⁵.

a) *Campos morfológicos*

CAMPO M. En el caso de las “locuciones verbales” o en general de lo compuestos, presenta la secuencia calculada en términos de categorías gramaticales V DÉT N, V PRÉP N, etc. Se indica también las secuencias verbales discontinuas con el lugar de posibles inserciones: “*prendre (cette information) en consideration*”

CAMPO F. Describe las variaciones morfológicas de los verbos según su tipo de conjugación (V1, V2, V3, etc.)²⁶.

b) *Campos sintáctico-semánticos (describen los argumentos)*

– CAMPOS DE ARGUMENTOS (N1,N2,N3). Se respeta el principio según el cual *EL VERBO Y SU PREDICADO ES DEFINIDO POR LA SUCESIÓN MÁS LARGA DE SUS ARGUMENTOS*. Cada argumento será anotado en función de su posición respecto al predicado verbal: N1, N2, N3. Como se ha indicado más arriba, los argumentos son definidos en un doble nivel. El primero está constituido por los ocho rasgos sintácticos a los que ya nos hemos referido. Cuando se trata de un predicado “general” como *ver* o *ser de color x*, será suficiente anotar el rasgo concreto en posición de sujeto. Por el contrario, cuando el predicado tiene un espectro más restringido, es necesario añadir, al lado de la indicación del rasgo, la clase de objetos que subcategoriza el rasgo. Así, el complemento de un verbo como *manejar* no está suficientemente definido con la ayuda del rasgo concreto (no se pueden *manejar* todos los concretos): es preciso dar una indicación suplementaria de que se trata de *instrumento*. Habrá, pues, en lo que concierne a los campos, la configuración siguiente:

prendre/N1:hum/N2 inc (prendre une cigarette)
prendre/N1:hum/N2.inc<alim (prendre un steak)
prendre/N1:hum/N2.inc<méd> (prendre de l'aspirine)
prendre/N1:hum/N2.inc<mti (prendre sa voiture)
prendre/N1:hum/N2.inc<mtc (prendre le bus)
prendre/N1:hum,mt/N2 loc<voie> (prendre l'autoroute)

25. Se excluyen todas las otras formas que tienen una morfología verbal y que no tienen función predicativa como los verbos auxiliares, los verbos soporte, las formas verbales que entran en la composición de verbos compuestos (como el verbo *prendre* en *prendre la mouche*). Las propiedades que vamos a estudiar ahora sólo son pertinentes para ellos.

26. Hay en francés un centenar de conjugaciones diferentes; esta diversidad se explica esencialmente por razones fonéticas e históricas. La conjugación francesa es, pues, mucho más compleja que los tres grupos que la tradición escolar enseña.

CAMPO DE NOMINALIZACIÓN Y ADJETIVACIÓN. Aquí se indica si el empleo de un verbo predicativo dado (definido por su dominio de argumentos en términos de clases de objetos) es susceptible de nominalización y de adjetivación. Estos cambios morfológicos se hacen con la ayuda de VERBOS SOPORTES, como hemos visto más arriba.

CAMPO DE TRANSFORMACIONES SINTÁCTICAS. Se precisa si el empleo verbal en cuestión responde o no a grandes transformaciones: pasiva, interrogación parcial referida a uno u otro de sus argumentos, pronominalización

c) *Campos de sinonimia y antonimia*

Se anotan en EL CAMPO S los equivalentes sinonímicos y en EL CAMPO A los antónimos. Es evidente en el cuadro de la subcategorización de los verbos por las clases de objetos que se puede dar cuenta de la sinonimia. He aquí algunos ejemplos:

prendre <aliment>. manger e to have; eat/d essen, nehmen
prendre <boisson> boire/e to have; drink/d trinken, nehmen
prendre <moyen de transport> emprunter, voyager en/e to take, go by d-benützen

d) *Campos de dominio y registro*

Vid *Supra* Base de datos de Argumentos sustantivos.

e) *Campos que concierne a la traducción (multilingües)*

Vid *Supra* Base de datos de Argumentos sustantivos.

3.2.2. Predicados nominales

A) *Campos morfológicos*

Los diferentes campos que tratan de la morfología son comunes a todos los sustantivos (ya los hemos visto en 3.1. argumentos sustantivos).

B) *campos sintáctico semánticos*

- CAMPOS DE ARGUMENTO (N1, N2, N3). Los argumentos son descritos de la misma manera que los de los verbos. Figuran en campos anotados N1, N2, N3. Se conforman igualmente al principio de la secuencia más larga y a la indicación de las clases de objetos. Como la lista de los elementos de cada clase está establecida, se está en condiciones de generar automáticamente el conjunto de frases que se pueden construir alrededor del predicado.
- CAMPO DE ACTUALIZACIÓN DE LOS PREDICADOS NOMINALES. El CAMPO W está reservado para la actualización de los predicados nominales.
- CAMPO DE TRANSFORMACIONES MORFOLÓGICAS. Se anota si el predicado nominal puede tener una forma verbal o adjetiva. Así, el predicado

deseo, actualizado por verbos soporte como *tener* o *experimentar*, está en relación morfológica con el verbo *desear* y el adjetivo *deseoso*. 'Lucas tiene el deseo de hacer el bien'; 'Lucas desea hacer el bien', 'Lucas está deseoso de hacer el bien'.

- CAMPO DE TRANSFORMACIONES SINTÁCTICAS.

c) *Campos léxicos. campo de sinonimia. / campo de antonimia*

d) *Campos de dominio y registro*

e) *Campos que conciernen a la traducción (multilingües)*

4. UN EJEMPLO: LA CLASE DE OBJETOS "VOIES"

Cuando se ha constituido una clase de objetos, se elaboran seis fichas diferentes

- un primer fichero elabora la lista de los *nombres comunes* de la clase, provistos de su rasgo sintáctico y de la indicación de la clase de objetos.
- el segundo fichero comprende los *nombres propios*, con las mismas informaciones.
- los tres ficheros siguientes describen los PREDICADOS u OPERADORES apropiados de la clase de *voies*: LAS VERBOS, LOS ADJETIVOS Y LOS SUSTANTIVOS PREDICATIVOS. Cada uno de estos predicados lleva la indicación de la naturaleza de sus argumentos.
- el sexto fichero recoge expresiones fijas en las que intervienen los elementos de la clase

I. DICCIONARIO DE NOMBRE COMUNES.

autoroute/G:n/T:loc/C:voie/D:transp/E motorway;highway
avenue/G:nf/T:loc/C:voie/D:transp/E:avenue
route/G:nf/T:loc/C:voie/D:transp/E:road
rue/G:nf/T:loc/C:voie/D:transp/E:street
 .../...

II. DICCIONARIO DE NOMBRES PROPIOS:

Nationale 7/G:nf/T:loc/C:voie/D:transp
Francilienne/G:nf/T:loc/C:voie/D:transp
Nationale card/G:nf/T:loc/C:voie/D:transp
A1/G:nf/T:loc/C:voie/D:transp
 .../...

III. DICCIONARIO DE OPERADORES VERBALES APROPIADOS:

SUJET <VOIE>

bifurquer/NO:loc<voie>/E:to fork, to branch off
conduire à/NO:loc<voie>/NI:loc/E:to lead to; to go to
desservir/NO:loc<voie>/NI:loc/E:to serve
grimper/NO:loc<voie>/E:to climb

*mener à/NO: loc<voie>/NI:loc/E:to go to; to lead to
serpenter/NO: loc<voie> E:to snake; to wind*

..
SUJET <HUMAIN> OBJET <VOIE>

bitumer/NO:hun NI:loc<voie>/E:to asphalt, to tarmac

emprunter/NO:hun/NI:loc<voie>/E:to take

goudronner NO:hun/NI:loc<voie>/F:to tar

prendre/NO:hun/NI:loc<voie>/E:to take

..
IV. DICCIONARIO DE LOS ADJETIVOS APROPIADOS:

boueux/NO:loc<voie>/D:transp/E:muddy

caboteux/NO:loc<voie> D:transp/E:bumpy

congestionné/NO:loc<voie>/D:transp/E:congest

défoncé/NO:loc<voie>/D:transp/E:broken up

embouteillé NO:loc<voie>/D:transp/E:congest

fréquenté/NO:loc<voie>/D:transp/E:busy

glissant/NO:loc<voie>/D:transp/E:slippery

impraticable/NO:loc<voie>/D:transp/E:impassable

passant/NO:loc<voie>/D:transp/E:busy

roulant/NO:loc<voie>/D:transp/E:fast

..
V. DICCIONARIO DE PREDICADOS NOMINALES APROPIADOS:

asphaltage/NO:hun/NI:loc<voie>/W:procéder a/E:asphalting

goudronnage/NO:hun/NI:loc<voie> /W:procéder a/E:tarring

pavage/NO:hun/NI:loc<voie>/W:procéder a/E:paving

perçement/NO:hun/NI:loc<voie>/W:procéder a/T:building

..
VI. DICCIONARIO DE LAS PREPOSICIONES FIJAS:

rebrousser chemin/E:to turn back turn round and go back

se frayer un chemin/E:to cut a path through

se mettre en chemin/T:to set out set off

tous les chemins vont à Rome/E:all roads lead to R

tarder en chemin/E:to loiter on the way

être par voies et chemins/E:to be always on the move

se mettre en route/E:to start set out

..

Hemos intentado resumir el pensamiento teórico que sustenta el método lexicográfico de G. Gross²⁷ Hay que reconocer la sólida armazón teórica del

27. Sólo nos hemos referido, dado el poco espacio de este tipo de trabajo, al ámbito de la oración simple. La oración compuesta para el profesor francés responde también a una estructura de predicados y argumentos el núcleo es el predicado y las oraciones relacionadas son sus argumentos. Tampoco nos hemos referido al adverbio que en ocasiones es considerado como predicado (en "Lucas trabaja bien" se puede eliminar, su argumento sujeto es la frase "el trabajo de Lucas es bueno") y en otras como argumento (en "Lucas ha tomado bien que se la haya hecho una sugerencia" no se puede eliminar tomar es el predicado y bien es un argumento).

proyecto del diccionario electrónico y la operatividad de esos conceptos en el diseño de bases de datos que den cuenta del reconocimiento y generación de oraciones posibles en la lengua y de su capacidad de ser traducidas por un diccionario electrónico. Gross utiliza como corpus las gramáticas y los diccionarios (y su propia competencia como hablante), pero quizás fuera conveniente utilizar algún corpus textual del que poder extraer información para la elaboración de las bases de datos y en el que poder comprobar ("validar", según algunos) la efectividad de la asignación de los diferentes tipos de argumentos a los predicados concretos. Por otro lado, la noción CLASE DE OBJETOS y de VERBOS SOPORTE son conceptos dinámicos abiertos, que se va construyendo poco a poco a medida que avanza la investigación (las gramáticas de las clases de objetos). Estamos, en definitiva, ante un reto claro y bien estructurado para dotar al léxico de gramáticas parciales que, con mucho trabajo por delante –eso sí–, puede dar lugar (a medio plazo quizás) a un diccionario electrónico que antes de lograr la traducción automática ha servido, sin duda, para hacer avanzar la teoría lingüística desde una fructífera e inevitable sinergia de la gramática y el léxico.

BIBLIOGRAFÍA DE G. GROSS²⁸

- "Structuration d'une base de données en vue de l'élaboration d'un dictionnaire bilingue", 1-20, Rapport interne
- (1988): "Réflexions sur la notion de l'union conjonctive", *Langue Française*, , pp. 19-36.
- (1989): *Les constructions converses du français*, Droz, Genève
- (1990) "Definition des noms composés dans un lexique-grammaire", *Langue française*, 87 pp.84-90
- (1991) "Typologie des adjectivaux", *Analyse et synthèse dans les langues romanes et slaves*, pp.163-188, G. Narr Verlag, Tübingen.
- (1992) "Forme d'un dictionnaire électronique", en Clas, A y Safar, H., *L'environnement traductionnel*, Silley, Univ. de Québec, pp.255-261
- (1993) "Trois applications de la notion de verbe support", *L'Information grammaticale*, 59, pp.6-11.
- (1994) "Connecteurs et traitement automatique", *TA-TAO: Recherches de pointe et applications immédiates*, Aupelf-Uref, Montréal, pp. 287-306
- (1994) "Classes d'objets et description des verbes", *Langages*, 115, pp.15-31.
- (1994) "Dictionnaires électroniques et traduction automatique", *Langages*, 116, pp. 48-59
- (1994) "Un outil pour le FLE: les classes d'objets", *Actes du colloque du FLE*, Presses Universitaires de Lille, Lille, 1994.
- (1995) "À quoi sert la notion de partie de discours?", *Les classes de mots. Traditions et perspectives*, Presses Universitaires de Lyon, pp.217-231, Lyon.

28. Se puede consultar la bibliografía generada en el Laboratorio de Lingüística Informática en la página de Internet www-lli.univ-paris13.fr.

- (1995): "Une sémantique nouvelle pour la traduction automatique: les classes d'objets", *La Tribune des Industries de la Langue et de l'Information électronique*, 17-18-19, Paris.
- (1995): "À propos de la notion d'humain", *Linguisticae Investigationes Supplementa* 17, J. Benjamins B.V. Amsterdam, pp. 1-80.
- (1995): "Enseignement des connecteurs", *Les Cahiers de l'Asdifle ndeg.6 Lexique et didactique du français langue étrangère*, Paris.
- (1995): "La structure événementielle des substantifs", *Folia Linguistica*, XXIX/1-2, pp. 43-65, Mouton De Gruyter, Berlin.
- (1996) *Le figement en français des noms composés aux locutions*, Ophrys.
- (1996), "Une catégorie hétérogène L'adjectif", *Studia Italici di Linguistica Teoretica e applicata*, Rome, 3, pp 44 -460

BIBLIOGRAFÍA COMPLEMENTARIA

- COURTOIS, B. y M. SILBERZTEIN, eds. (1990): *Les dictionnaires électroniques du français*, *Langue Française*, 8
- CHOMSKY, N. (1965): *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Mass.
- FREGE, G. (1971): *Estudios sobre Semántica*, Barcelona, Ariel.
- GROSS, M. (1975): *Méthodes en syntaxe*, Paris, Hermann.
- (1994), "Dictionnaires électroniques et traduction automatique", *Langages*, 116, pp.48-59.
- HARRIS, Z.S. (1951): *Methods in structural linguistics*, Chicago University Press.
- (1976): *Notes du cours de syntaxe*, Paris, Le Seuil.
- SILBERTZTEIN, M. (1989) "Dictionnaires électroniques et reconnaissance lexicale automatique", Ph. D. dissertation, Laboratoire d'Automatique Documentaire et Linguistique, Université Paris
- (1990), "Le dictionnaire électronique des mots composés", *Langue Française*, 87, pp.71-83.