

UNA ENFERMEDAD INFANTIL DEL ANALISIS DE DATOS TEXTUALES EN LAS CONSTRUCCIONES CIENTIFICAS EN HISTORIA, LA "TEORIA DEL DISCURSO". (*)

Por M. Borillo y J. Virbel

3.- Análisis del Análisis Automático del Discurso aplicado al estudio del discurso martinista

3.1. Sobre la delimitación del problema estudiado y la constitución del corpus

a) El tema de investigación definido por los autores se articula en varios niveles de precisión; en el primer nivel, se pasa de una:

Interpretación de la función del discurso iluminista a finales del siglo XVIII

a una



Interpretación relativa a St Martin

El paso (la inclusión de este problema general, respecto de un nivel global de conocimiento de procesos culturales históricos) a una interrogación específica, está fundado en razones que pueden parecer accidentales:

- facilidad de acceso
- silencio sobre el pensamiento político de St. Martin
- antigüedad de los estudios concernientes

o más fundamentales:

- influencia decisiva en ciertos autores del s. XIX

En un segundo nivel, se pasa de:

Obras completas de St. Martin

a:



"Mon portrait historique et philosophique"

por dos razones de orden totalmente diferente:

- representatividad y extensión, temática y cronológica del "Portrait" con respecto al conjunto de la Obra
- aptitud del "Portrait" para prestarse al análisis que GP tienen en proyecto

(*) La primera parte de este artículo apareció en nuestro Boletín anterior, n° 25.

Se constata que los autores han pasado de una preocupación general, propia de un aspecto del pensamiento filosófico y político, a una investigación más precisa, sobre un autor, y de una obra de este autor. Las razones que fundamentan esta especificación, en este nivel, constituyen el primer paso de un conjunto de hipótesis, que deberían precisarse y ramificarse con la continuación de la investigación, y que por consiguiente sólo servirán mientras que la construcción ulterior, fundamentada ya en parte sobre ellas, esté sometida a verificación. Sin embargo se puede apreciar una gran heterogeneidad en estas justificaciones.

b) Anclado en una investigación sobre un autor y un documento, el problema inicial se redefine con más precisión. También así las razones que motivan estas precisiones corresponden solamente a elecciones, cuyas justificaciones pueden ser variables, pero cuya precisión sólo reside en la adecuación de las operaciones ulteriores que dichas elecciones autoricen y hagan posibles. La cuestión inicialmente planteada, que deriva como se ha visto, de un conjunto de hipótesis anteriores y adyacentes, conduce finalmente, al término de una reflexión dirigida según dos ejes paralelos de análisis del problema (cf. Esquema adjunto), a una investigación referente a la "duplicidad" en el Portrait: existencia, manifestaciones, función.

c) Nos parece fundamental subrayar que:

1/ Las decisiones hacen referencia cada vez a justificaciones de naturaleza sumamente variada (ver más abajo) pero cuya necesidad no es nunca demostrada o cuya coherencia con relación al conjunto del proyecto no es nunca establecida.

2/ Se pasa así por elecciones sucesivas de la interrogación inicial al estudio del término "circunstancia" en el Portrait. Estas elecciones tienen de común: que son siempre una posibilidad entre otras igualmente concebibles. Si se trata de manifestar efectivamente el rigor científico que reclama, está claro que cada una de estas elecciones deberá:

- o ser planteada como uno de los "axiomas" sobre los que se apoya el razonamiento
- o ser introducida como conocimiento ya establecido anteriormente
- o tener el statut de hipótesis secundaria, con la condición de validarla inmediatamente, o a posteriori, con vistas a la construcción global, por el examen de la función lógica de cada elemento del discurso.

No es cierto, teniendo en cuenta la naturaleza de los fenómenos analizados, que estas precauciones aseguren a los resultados un statut claramente definido del punto de vista científico. Pero sin ellas está asegurada la no científicidad.

Si se refiere a las justificaciones adelantadas por los autores relativas a las diferentes decisiones, se constata que dependen de niveles y de consideraciones muy diversas:

I

II

Conocimiento Histórico
de la obra de St Martín

Teoría del Discurso

- | | | |
|--|---|--|
| <p>① combate por la</p> <p>② lo que define una "duplicidad" inherente al pensamiento de St Martín</p> <p>③ el Portrait es un discurso-monólogo</p> | <p>promoción y liberación del individuo</p> <p>imputación imperativa a la sumisión a Dios</p> | <p>① Sin semántica general. Sin categorización a priori</p> <p>② Un discurso es el resultado de un proceso de producción definido por las condiciones de producción</p> <p>③ Este proceso de producción se apoya en el sistema de la lengua pero resulta de las condiciones de producción exteriores a la lengua</p> <p>④ Los rasgos visibles son fenómenos de selección-combinaciones lingüísticas.</p> |
|--|---|--|

- ① El Historiador define un corpus como algo dependiente de un mismo estado de las condiciones de producción
- ② El Portrait depende de un estado definido de condiciones de producción

Cuestiones para resolver

- 1 - Rasgos de duplicidad en el Portrait?
- 2 - Condición de posibilidad de la duplicidad en la organización del discurso
- 3 - Función de la duplicidad

Figura 1

- a) - de un conocimiento histórico "general", relativo al iluminismo como corriente de pensamiento, y el lugar de St Martin en esta corriente,
- b) - de un conocimiento histórico "particular", relativo al contenido de la obra de St Martin, y en particular del "Portrait",
- c) - de una concepción de la representatividad de una obra particular con respecto a una obra completa, fundamentada aquí sobre:
 - la extensión en el tiempo de redacción
 - la anarquía en el orden de edición
 - la amplitud de los temas abordados
- d) - de una teoría del discurso, sobreentendida por consideraciones lingüísticas, de donde resulta, a través del concepto de "selección-combinación" una visión del discurso "Portrait" y un procedimiento de examen,
- e) - de un análisis del pensamiento de St Martín del que se extrae el concepto de "duplicidad",
- f) - de un análisis del Portrait del que proviene la elección del estudio del término "circunstancias".

Ninguno de estos elementos está explícitamente dado para lo que es, pero el statut de alguno de ellos es suficientemente claro para que su afectación pueda efectuarse sin ambigüedad. Es así que a), b) y c) constituyen "conocimientos previos", por diversas razones: b), en particular, es de naturaleza perfectamente no-homogénea lo que GP pretenden poner en práctica en su construcción, y depende más bien de la historia tradicional; c) dependería más bien de cierta "sociología" de la producción literaria. El punto d) podría constituir uno de los "axiomas" mencionados más arriba; pero de todas formas está estrechamente relacionado con los puntos e) y f) que representan la elección central, la del término "circunstancias", alrededor de la cual se ordenará el corazón mismo de lo que se presenta, hay que decirlo, como una demostración. Intentar hacer desaparecer el equívoco que se une a la selección de este término aparece claramente como condición necesaria para el examen de la consistencia lógica del conjunto del discurso.

Se nos da una indicación sobre lo que los autores no desean: prohíben toda clasificación o categorización a priori. Esto supone pues, que ellos van a referirse únicamente a la obra misma para evitar el riesgo de encontrar como resultado del análisis el contenido mismo introducido y organizado por esta categorización; por el contrario, comparan sistemáticamente secuencias que pongan de manifiesto un sistema de sustitución de términos con relación a puntos invariantes. El punto invariante se define a partir de numerosas y atentas lecturas a continuación de las cuales cierto número de palabras, sustantivos y adjetivos, aparecen acompañando regularmente las manifestaciones de "duplicidad": apología o denigración del hombre, presencia inmediata o lejana de Dios. Entre estos términos, - que parecen constituir el "arsenal semántico de la duplicidad" - el término "circunstancia" se impone a la atención por las siguientes razones:

- no aparece episódicamente
- está en relación con todas las preocupaciones de St Martin, y por tanto sobrecargado de significaciones
- convoca la historia individual y colectiva.

De hecho, el camino que conduce a la elección del término "circunstancia" corresponde exactamente a aquel que GP rechazan inmediatamente antes, con la particularidad de que es negado como tal, y que por consiguiente queda totalmente implícito.

Si en efecto se hace el inventario de las condiciones que permiten definir esta elección, se constata que es preciso:

- a) - haber definido y saber aislar en el texto las manifestaciones de "duplicidad", lo que supone por lo tanto
- b) - haber definido un modo de selección de los términos "que acompañan" estas manifestaciones, ya sea en término de categoría gramatical, de contenido, de aparición en contextos de entornos definidos, etc.
- c) - haber definido las "preocupaciones" de St Martin, y saber identificarlas y aislarlas en el texto, para verificar la relación de un término de la lista seleccionada con ellas
- d) - haber definido y saber identificar los elementos que denotan o connotan "la historia individual" y la historia "colectiva" para verificar también ahí la presencia de una relación.

Lo cual es lo mismo que hacer, sin decirlo, o sea sin imponer las restricciones de explicitación y de regularidad absolutamente necesarias, "una categorización a priori de los términos encontrados en el material sometido a análisis".

Hay que estar de acuerdo con GP cuando afirman que la operación de constitución del corpus no es neutra, o exterior a la problemática, pero que necesita por el contrario la intervención de una teoría extra-lingüística, en nuestro caso histórica; pero esto sólo concierne al contenido del corpus (es decir, la elección de "circunstancia" mejor que otro término) y no a la forma en que éste se ha obtenido, la cual, en el presente caso se abstrae del dominio de las hipótesis, y por consiguiente peca de arbitraria, en el sentido de no-justificada.

Así, referente a la fase central de la constitución del corpus, no solamente el statut, con respecto a la demostración, de las diferentes operaciones que conducen a la elección de "circunstancia" no está especificado, sino que además, su naturalidad efectiva no está explicitada. Por esto los autores, sin darse cuenta, se ponen en contradicción con sus afirmaciones (explícitas) sobre lo que no son. De hecho son lo que no deberían ser; en particular, es imposible que sean regulares.

3.2. Sobre los procedimientos

Las observaciones que se pueden hacer sobre las operaciones efectuadas son del mismo tipo que las que han sido formuladas respecto a la constitución del corpus, es decir, que muchas de ellas están dadas, tácitamente, como "naturales", en el sentido de que no reciben otra justificación que la de ser utilizadas, mientras que, también aquí, encarnan en cierto modo las hipótesis, totalmente implícitas, sobre el análisis efectuado.

Se debe observar inmediatamente que no se da ninguna regla referente a la segmentación del texto y que defina la dimensión de las secuencias, que varían "de la frase al párrafo". Se puede suponer que existen estas reglas, pero en este caso nada garantiza que dos analistas que persiguen los mismos fines logren la misma descomposición.

Después se ha visto que el análisis sintáctico persigue a partir de las secuencias los siguientes resultados:

- segmentación en frases y después en proposiciones.
- regularización del orden de los términos.
- caracterización de las relaciones inter-proposicionales.
- segmentación de las proposiciones en UMA.

Pero no existe "camino real" único que nos conduzca de las secuencias a los UMA, como se puede ver analizando el mismo ejemplo dado por GP, sino más bien una red compleja de correspondencias mal definidas, a través de la cual ningún camino particular recibe caracterización explícita.

Desde el punto de vista general, se constata que las operaciones no están ni justificadas, ni son regulares en su utilización, lo cual plantea nuevamente el problema de su reproducibilidad.

Por ello la frase (secuencia) siguiente (29): "J'ai senti dans cette circonstance combien les hommes étaient aveugles de solliciter des emplois publics, puisque cela les plaçait dans des positions qui les condamnaient à ne pas croire à l'honnêteté", da lugar (cf. esquema siguiente) a las diversas operaciones reseñadas más abajo:

1 - segmentación en 5 proposiciones.

2 - transformación de ciertos morfemas de tiempo:

- étaient → sont (1)
- solliciter → sollicitent (2)
- plaçait → sont placés (3)
- condamnaient → condamne (4)
- ne pas croire → ne croient pas (5)

pero: j'ai senti → j'ai senti

Estas transformaciones son de tipo diferente:

- desde el punto de vista del contenido:
 - (1), (3), (4) corresponden a: tiempo pasado→tiempo presente
 - (2), (5) corresponden a: infinitivo→modo personal
 - (3) corresponden a: activo→pasivo
- desde el punto de vista de su función:
 - (1), (3), (4) corresponden a una regularización por supresión de los morfemas específicos de tiempo
 - (2), (3), (5) corresponden a transformaciones más generales en el nivel de la proposición

3 - cambios de subordinantes

combien → que
 de → si
 puisque → parce que

Los dos grupos de modificaciones aportadas a las proposiciones iniciales no plantean tanto la cuestión de su verdad o de su falsedad, sino la de su legitimidad. Ahora bien, lo que fundamenta la legitimidad de estas operaciones, es el empleo validable que se hace de ellas. Según el dominio de regencia, y el objetivo del análisis, la reducción combien→que, el aumento de→si, la transformación condamner a V inf→condamner à ce que Phrase serán justificables o no; también aquí lo que hay que validar es el conjunto de las operaciones efectuadas, incluidas éstas.

El paso desde las proposiciones así tratadas a los UMA presenta los mismos aspectos.

Constatamos:

1 - una nueva segmentación:

sollicitent des emplois publics → { sollicite emploi
 emploi public

2 - sustituciones lexicales:

je → ego
 hommes → homme

3 - reducciones lógicas, ortográficas, o sintácticas:

les → L
 ces → cs
 circonstance → circ
 une → U
 sont aveugles → être aveugles
 sont placés → être placés
 condamne → condamner
 ne croient pas → ne pas croire
 être → E

pero:

sollicitent → sollicite
 ai senti → avait senti
 être aveugle → être aveugle

donde "sollicite" queda en presente, "sentir", al mismo tiempo que el paso de la 1ª a la 3ª persona, pasa del pretérito perfecto al pluscuamperfecto y "être aveugle" no se transforma en "E aveugle", como "être public" → "E public", por ejemplo.

4 - supresiones: "ainsi".

5 - nuevas relaciones entre UMA, marcadas por →, que son categorizadas.

Estas observaciones sólo afectan a la fase inicial (manual) del proceso, pero podrían ser ampliadas a las demás operaciones, y en particular en lo que concierne al cálculo de la proximidad de los UMA, apoyado en una ponderación de los elementos morfosintácticos de los enunciados. También en este caso, ya se trate de la elección de los pesos respectivos de cada tipo de unidades, o del umbral que acepta o rechaza la asociación de UMA, la única justificación que puede existir es la práctica, es decir, la experimentación. Será preciso además para ello que el papel de la lógica de estos elementos en la construcción intelectual esté claramente explicitado. En caso inverso, ¿cómo podrían ser "controlables"?

De la misma manera, las relaciones que pueden unir los UMA, y a partir de las cuales se efectúan estos cálculos, no tienen apoyo en justificaciones externas aparte de los resultados que permiten establecer. Para su eventual justificación interna (¿coherencia formal?) se puede interrogar acerca del sistema de relaciones utilizado, sobre todo si se constata que son de orden muy diferente.

Las seis relaciones obtenidas son las siguientes (30):

- 01 : que
- 25 : composición de más de 2 argumentos
- 10 : si
- 06 : porque
- 51 : relativa cuyo antecedente está en la 1ª UMA
- 26 : composición de 2 predicados

01 y 26 marcan de hecho una construcción sintáctica del tipo verbo operador, y hasta cierto punto son redundantes; en el ejemplo de la frase analizada, no se ven las razones que pueden motivar la distinción (sólo suponiendo comportamientos sintácticos de superficie) entre sentir que (P) y condamner à (ce que P→Inf). (01) (26)

10 y 06 indican relaciones de determinación en el sentido lógico más que en el gramatical, como también se puede ver

en el ejemplo dado, en la medida en que el "si" y el "parce que" introducidos en el análisis de la frase provienen de un tratamiento de la frase inicial en el que no figuran.

3.3. Sobre la interpretación de los resultados

El paso de estos resultados a su sentido real (es decir, la interpretación que se les asocia) no es en absoluto inmediato, sino que consiste por el contrario en agregarles un discurso cuya relación con los datos que lo justifican, - estos mismos resultados - no es una consecuencia definida de éstos. En otras palabras, estos resultados sólo tienen el sentido que GP les dan, haciendo intervenir otros datos que no habían sido tomados en consideración al comienzo ni señalados en ninguna otra parte, como definidores del campo semántico del problema propuesto. Los resultados son arbitrarios en sí mismos, en la medida en que no tienen un status lógico definido, en lo que respecta a las hipótesis y a las cuestiones formuladas en el origen. Las intervenciones exteriores, referibles en el discurso de GP parecen al menos de tres tipos:

a/- una extrapolación y una generalización semántica, a partir de los dominios o reagrupamientos obtenidos.

b/- una generalización de las observaciones efectuadas a partir de los resultados del análisis en el conjunto del Portrait, así como en la obra de St Martin. A veces este aspecto es indisoluble de su simétrico que consiste por el contrario en hacer intervenir un conocimiento del Portrait o de la obra completa para esclarecer el sentido de estos resultados, - y sin que frecuentemente se pueda decidir la justificación de su empleo.

c/- una generalización a (o por el contrario como anteriormente, una intervención de) un conocimiento colateral del pensamiento iluminista, y del período revolucionario.

Un ejemplo simple del caso a/ lo suministra el comentario siguiente: "Dans ce grand domaine, les circonstances sont aperçues/rencontrées ... les circonstances bonnes ou mauvaises sont donc subies, jamais provoquées" (31), a propósito del dominio así constituido:

les circonstances	repossantes commodes agréables flatteuses	que j'ai	aperçues rencontrées
-------------------	--	----------	-------------------------

El ejemplo anterior permite también precisar el aspecto b/. En efecto, entre las dos frases citadas más arriba está inserta la siguiente: "le philosophe inconnu se désigne toujours dans son Portrait comme un effet de l'histoire, jamais comme un agent historique; la débilité de ses intestins ou la mort de Robespierre, l'incompréhension des femmes ou l'irruption de Bonaparte sont des circonstances qui lui échappent ... ou des échappatoires" (32).

Este ejemplo es relativamente simple con relación a los fenómenos que deseamos destacar. Podrían darse muchos otros. Así: "Dans le combat pour l'"esprit" engagé par Saint Martin, les circonstances semblaient avoir pour fonction principale de freiner la diffusion de sa philosophie. Les conditions propices à l'usage de la parole ont fait cruellement défaut au Philosophe inconnu (D 23, D 26)" (33).

Por último, el siguiente comentario ilustra suficientemente el tercer aspecto señalado más arriba (c/): "L'illumination qui réservait, avant 1789, son enseignement aux seuls initiés et affirmait la nécessité de taire certaines vérités, souhaite devenir bavard et obtenir un succès égal à celui de Rousseau ... L'usage de la parole apparaît donc comme une possibilité d'accès à l'univers spirituel" (34).

Nos encontramos pues con que los datos necesarios para la explotación de los resultados son de la misma naturaleza que los que han sido ya intervenidos, frecuentemente de forma implícita, en el momento de la definición y del análisis del problema. La razón de esta situación es simple: como estos datos no han sido todos introducidos como conocimientos establecidos o hipótesis previas, es preciso recurrir de nuevo a ellos para utilizarlos, es decir, interpretar aquí los resultados. Con el riesgo de circularidad lógica - hay que subrayarlo - a que está expuesta la ausencia de distinción explícita entre las diferentes funciones que una información puede desempeñar en la elaboración de una "demostración". Otra forma de decir consiste en observar que nunca se ha partido, en los conocimientos utilizados a lo largo del análisis, de aquellos que están ligados a la construcción, y que solamente pueden servir como hipótesis, ni de aquellos que son independientes, y que sólo pueden concurrir en la verificación de las hipótesis propuestas, y la inserción de los resultados obtenidos en el conocimiento anterior, modificado y enriquecido de esta forma.

Dicho de otro modo, la formulación última del problema, bajo la forma de triple interrogación respecto a la duplicidad en St Martin no está suficientemente especificada, de suerte que la solución depende de la calidad (imaginación, intuición, cultura, etc.) del "interpretador", y por tanto sin duda alguna variable respecto con él, sin que existan por otra parte medios definidos de guiar y verificar cada interpretación particular.

NOTAS (*)

- | | |
|---|--|
| (1) por Gérard GAYOT y Michel
PECHEUX, <u>Annales ESC</u> , 26,
3-4, mayo/agosto 1971,
p. 681-704. | (16) id. |
| (2) p. 682. | (17) id. |
| (3) p. 682. | (18) p. 683. |
| (4) id. | (19) id. Subrayado por los
autores. |
| (5) id. | (20) p. 683. |
| (6) p. 683. | (21) id. |
| (7) id. | (22) id. |
| (8) id. | (23) id. |
| (9) p. 686. | (24) p. 688. |
| (10) id. Subrayado por los
autores. | (25) p. 685. |
| (11) p. 686-687. | (26) id. Subrayado por los
autores. |
| (12) p. 687. Subrayado por
los autores. | (27) p. 688. |
| (13) p. 687. | (28) p. 705. |
| (14) id. | (29) p. 688. |
| (15) id. | (30) p. 691. |
| | (31) p. 694. |
| | (32) id. |
| | (33) p. 695. |
| | (34) id. |

REFERENCIAS

- BORILLO, M. (1971). Construction d'un raisonnement déductif au moyen de la simulation d'un raisonnement archéologique traditionnel, in Archéologie et Calcul. 10x18 UGE. Paris.
- BORILLO, M., FERNANDEZ DE LA VEGA, W., GUENOCHÉ, A., JANON, M., VIRBEL, J. (1974). "Analyse des textes et raisonnement en histoire. Une expérience de recherche historique à partir de l'analyse d'un corpus d'inscriptions funéraires romaines", in Archéologie et Calcul, 10x18, UGE. Paris.
- PECHEUX, M. (1969). "Analyse automatique du discours" (Dunod).
- POPPER, K.R. (1973). "La logique de la découverte scientifique" (Payot).
- URADCA-GRIL (1974). "Procédures d'analyse et méthodes de validation dans l'étude des données textuelles", Prj/64. URADCA. Marseille.
- VIRBEL, J. (1974). "Contribution à une réflexion méthodologique sur les conditions d'emploi de moyens informatiques pour les données textuelles" (A paraître), URADCA. Marseille.

(*) Las 28 primeras notas corresponden a la parte de este mismo artículo aparecida en nuestro Boletín n° 25.