

## EL PROCESAMIENTO DEL TEXTO COMO LENGUAJE NATURAL

Martín J. FERNÁNDEZ ANTOLÍN  
*Universidad Europea Miguel de Cervantes*

### 1. INTRODUCCIÓN:

Los Estudios de Traducción han situado como hipótesis de trabajo la determinación de los límites de su disciplina y, como resultado de esta búsqueda, la caracterización de las unidades lingüísticas, de sentido o pragmáticas que se habrían de convertir en las herramientas con las que evaluar dichos límites. Esta corriente ha provocado que las unidades de traducción objeto de estudio hayan participado de una característica común: su naturaleza analítica, opuesta a genérica según la distinción que realiza De Beaugrande (1980).

Los modelos genéricos, por su parte, han estado sustentados fundamentalmente por metodologías propias del análisis holístico que apuntaban a la taxonomización simplemente como una más de las aplicaciones. Las taxonomías existentes, en consecuencia, se han llevado a cabo de un modo subsidiario, como método para ofrecer al traductor una asistencia en la transferencia interlingüística. No obstante, esta voluntad instrumental delata, a su vez, una falta de interés por la consideración de la sistematización y distribución de las funciones semánticas y pragmáticas que se contienen en el texto como unidad lingüística en sí misma y, más allá del ámbito intralingüístico o intracultural, como fenómeno interlingüístico.

En cualquiera de las taxonomías existentes podemos observar que sus bases metodológicas se han situado en planteamientos que, aunque pudieran parecer opuestos, podremos asimilar tomando como base la característica que señalan como central. Tanto en el caso de las denominadas tipologías unidimensionales (Reiss y Vermeer, 1984; Delisle, 1988 o Nord, 1991) como en el de las multidimensionales (Roberts, 1995, Biber, 1989 o Snell y Crampton, 1983) el término **función** ha cobrado un papel fundamental. En el primer grupo la función se convertía en un parámetro más que habría de asignar validez al resto; en el segundo, la función se sitúa en un teórico plano externo, posición que no es real, ya que vuelve a incluirse dentro de la interacción del resto de parámetros y, de nuevo, como elemento rector de los demás. Es decir, que pese a su teórica localización como resultado de un conjunto de variables que incluyese los ámbitos pragmático y semántico, el valor funcional del texto como unidad lingüística y de traducción se ve relegado a la ordenación de valores extrapolables de un análisis holístico; en otras palabras, se continúa sin prestar atención a las repercusiones que dichas metodologías analíticas observan en los planos semántico y pragmático.

A la luz de este aparente abandono del concepto de tipología textual como elemento integrador de las dimensiones lingüística, pragmática y semántica tanto para los Estudios de

Traducción como para la lingüística aplicada, creemos justificado que nuestra defensa gire, precisamente, en torno al concepto de función; desde ahí, articularemos la propuesta de aplicación tipológica al procesamiento de lenguaje natural que apuntamos en este artículo.

## 2. LA FUNCIÓN COMO PARÁMETRO TIPOLOGICO

No obstante, creemos necesario integrar una noción previa a la de función ya que, de otro modo, no sería posible formular una hipótesis veraz de trabajo sobre la que construir una metodología empírica. Así, la primera cuestión que debemos plantearnos debe tener su origen en la esencia misma del presente trabajo, y podría formularse como ¿qué representa un tipo textual? Para responder a estas cuestión, que implicaría en cierto sentido la reconsideración de la ya conocida acerca de la funcionalidad, creemos conveniente partir de una afirmación de Juan Carlos Sager (1993: 51):

*The form of the communication itself is meaningful (...) each society has its sets of forms and appropriate codes, the deviation of which are also meaningful.*

Según estos postulados, nuestra propuesta empírica podría cifrarse del siguiente modo: el análisis y posterior procesamiento de la unidad textual como objeto de estudio lingüístico –o instrumento interlingüístico en el caso de los Estudios de Traducción– que desempeña una función social debe representar un marco de referencia dentro del que podamos señalar no sólo las actuaciones que se ajusten al mismo sino también las desviaciones que nos proporcionen el análisis holístico y el análisis genérico; en este sentido, dicha centralidad del modelo textual deberá capacitar a los lingüistas y a los traductores para la cuantificación de su carácter nuclear o periférico respecto a la norma establecida. Consecuentemente, si la forma significa socialmente porque se convencionaliza, el texto será un elemento deíctico, ya que gozará de una posición como instrumento de comunicación en una sociedad determinada, y como tal podrá ser descompuesto en diferentes parámetros que nos permitan procesar su naturaleza simbólica.

## 3. LOS CONCEPTOS DE TEXTO VIRTUAL, VARIACIÓN Y CAMBIO

Sin embargo, al considerar al texto como lo que Fillmore (1968, 1975) y Lyons (1977) denominan “deíctico social” nuestra hipótesis se sitúa ante una nueva dimensión que también podemos deducir de la afirmación de Sager: si el modelo textual es un deíctico social, será posible analizar empíricamente aquellos parámetros que lo convierten en un marco de referencia; pero ¿cómo podemos dar cabida funcional y significativa a sus desviaciones?

Para responder a esta pregunta creemos necesario el estudio de dos conceptos: el de texto virtual, por un lado, y el de variación y cambio por otro. La virtualidad del texto nos permite hablar de las condiciones de normalización de los tipos –modelos– textuales. Frente al concepto de sistema del código podemos justificar la existencia de otro sistema convencionalizado socialmente –la norma lingüística de uso– en el que se sitúan los tipos textuales. Si el concepto de tipo textual es una noción artificial que nos permite procesar los textos según sus características holísticas y

genéricas, la posición intermedia del texto virtual habilita la dicotomía entre los conceptos de variación y cambio, que darán respuesta de este modo a la afirmación de Sager acerca de la funcionalidad del componente normativo y de sus desviaciones. En otras palabras, el concepto de texto virtual como ente lingüístico normalizado por el uso nos permite procesar sobre el modelo sistémico sus actualizaciones sociales sin que éstas representen un cambio; al mismo tiempo, el contraste de las desviaciones que se produzcan respecto al modelo sistémico obtendrán una representación gráfica que dará pie, ulteriormente, a su descripción y clasificación científica. De este modo se nos abre la perspectiva de que los prototipos virtuales pueden ser tan diversos en su actualización como una cultura determinada haya estimado normalizar para satisfacer funcionalmente los requerimientos de sus instancias comunicativas.

No obstante, esta funcionalidad no puede reducirse al procesamiento del aparato formal del texto por el que parece abogar Sager (*sets of forms and appropriate codes*). Si los textos desempeñan una función en sociedad es porque son algo más que significantes que encierran un significado unívoco; por esta razón los hemos reconocido como deícticos sociales y por esa misma razón debe llevarse a cabo el procesamiento de los parámetros que socializan el significado referencial y el significado social de las formas lingüísticas.

Pero en este planteamiento no hemos integrado todavía la cuestión traductológica; como hemos afirmado, pese a que los textos son iconos sociales formalmente reconocibles, su función social no habría de justificarse únicamente desde el plano formal. De hacerlo, nos aproximaríamos a la paradoja de Osgood (1957) –sólo a igualdad situacional podrá producirse igualdad significativa–, lo que anularía cualquier posibilidad de procesar el lenguaje natural a no ser por medio de parámetros de desambiguación de instancias comunicativas como contextos de situación; ésta cuestión, sin embargo, se encuentra todavía en estado embrionario y su aplicación real deberá esperar un número de años significativo.

La solución, en el fondo, parece encontrarse en representar la función del lenguaje tal y como se define en su uso matemático: asimilando para nuestra metodología un verdadero carácter multidimensional en el que se pueda contener diferentes ámbitos de una misma realidad. Esta consideración de la funcionalidad como un sumatorio de propiedades textuales nos servirá, a su vez, para establecer una invariante de comparación –o hipótesis metodológica– de cara al procesamiento del intercambio traductológico partiendo de bases lingüísticas; para ello partiremos del concepto de equivalencia (Rabadán 1991; Koller, 1995) sobre el que habrá de superponerse el dinamismo del concepto de similitud (Sovran, 1992; Chesterman 1996). Dado que nuestro objetivo es considerar el texto desde la doble perspectiva normativa y de actualización, las nociones de similitud convergente y divergente dentro del concepto de equivalencia nos permitirán afrontar el desarrollo del modelo de procesamiento taxonómico de los textos con mayores garantías.

#### 4. LA DESCOMPOSICIÓN DEL TEXTO COMO MARCO METODOLÓGICO

Si, como afirmamos, el carácter funcional no puede circunscribirse únicamente al valor formal que presenta el texto como icono, cabe formularnos la pregunta: ¿cómo podemos asignar un valor funcional al texto sin reducirlo a la forma? Para ello hemos establecido como metodología la descomposición de los textos en tres grandes ámbitos: pragmático, semántico y lingüístico, que

hemos analizado desde la lingüística aplicada a los Estudios de Traducción. Si bien esta decisión puede parecer un contrasentido respecto a nuestras críticas sobre el excesivo énfasis en la dimensión holística que achacábamos a tipologías previas, el procedimiento analítico que aquí proponemos no se centrará en unidades textuales aisladas, sino en el valor que aportan al significado social del texto en su conjunto. De este modo, será posible separar la función textual de estos parámetros para situarla como resultado del procesamiento del sumatorio de los grupos de variables.

Si volvemos a las bases de Sager, las respuestas a la dimensión social del texto como modelo de referencia nos llegarán desde la pragmática. Consecuentemente, hemos optado por situar el grupo de variables pragmáticas en posición de prevalencia sobre las formales y las semánticas, ya que el procesamiento del significado dependerá de las condiciones no lingüísticas en las que se produce la instancia comunicativa. Cabría aquí la contestación a nuestra decisión con el argumento de que el sistema lingüístico es únicamente un mediador para la interpretación de las instancias comunicativas; sin embargo, esta posición etnocéntrica y tendente a la perpetuación prescriptiva del sistema lingüístico como marco de actuación irrefutable no debe ocultar el hecho de que también es esa misma cultura la que actúa como mediadora en la asignación de un valor real a los significados lingüísticos según el contexto humano –y utilizamos aquí “humano” en el sentido cultural de “concreto” como antónimo de “discreto”– en el que se produzcan. En consecuencia, al asumir esa posición nuclear para el posterior procesamiento de los códigos semántico y formal, estas variables deberán responder al modo en que se produce la convencionalización del significado del texto como deíctico social: ¿cuál es la intención del texto?, ¿qué grado de conocimiento resulta necesario en la instancia comunicativa?, ¿qué tipo de relación se establece entre emisor, receptor y texto? Y, por último, una vez resueltas estas cuestiones, nos podremos preguntar ¿qué posición adopta el texto dentro de cada una de las culturas como deíctico social?

Si bien el detalle de los parámetros que se integran en cada de estas variables se presenta de forma más gráfica en el apéndice que acompaña a este artículo, para una mayor claridad expositiva y una fluidez en la lectura del artículo simplemente enumeraremos a continuación dichos parámetros mencionando las fuentes de las que parten; las constricciones de espacio de un artículo como el presente nos evitarán, lamentablemente, el desarrollo de los motivos que nos llevaron a seleccionar estos y no otros parámetros o a su descripción detallada. La primera de las preguntas, aquella relacionada con la intención del texto, se aborda en nuestra metodología a través de cuatro pilares: **referencial**, **conativa**, **expresiva** y **textual** (Bühler, 1965; Reiss, 1974; Halliday y Hasan, 1976; Halliday, 1978; Werlich, 1976); la segunda de las cuestiones, que buscaba el grado de conocimiento necesario para la comunicación, se basa en los conceptos de implicatura (Grice, 1975; Levinson, 1983), presuposición (Russell, 1905; Karttunen y Peters, 1975; Wierzbicka, 1975; Gazdar, 1979) y conocimiento compartido (Fraser, 1997; Wierzbicka, 1996) y ha observado cuatro productos paramétricos: **básica**, **paralela**, **comentario** o **discursiva**. La tercera parada en nuestra determinación de variables y parámetros tenía como destino la relación de base icónica que, a través del texto, se establece entre emisor, receptor y modelo textual; para satisfacer esta cuestión, los parámetros integrados en esta variable atienden a la primacía asignada a cada uno de estos participantes: **interpersonal**, **textual** y **metacomunicativa** (Halliday, 1978). Por último, la variable que habría de dar cabida a la posición del modelo textual como deíctico social en una cultura y una lengua determinadas ha asumido los condicionantes propios de las teorías lingüísticas (Holub, 1972) y de los Estudios de Traducción (Even-Zohar, 1979, 1990; House, 1981; Roberts, 1995; Toury, 1995), y el resultado ha sido una tríada paramétrica que se desglosa en: **convencional**, **particular** y **general**.

Dado que la posición social del texto como significado queda cubierta por las variables pragmáticas, el código semántico argumentará sobre ese significado social. Si empleábamos el término de deíxis social siguiendo a Fillmore y Lyons, deberemos centrarnos ahora en lo que representa la deíxis textual. La metodología argumentativa que hemos seguido nos permitirá procesar sobre el ámbito pragmático –a modo de función matemática– una serie de argumentos de orden semántico sobre aquellos aspectos de la deíxis social que ya se encontraban definidos. Así, las preguntas que hemos situado como fundamentales dentro del ámbito semántico han sido: ¿qué tipo de relaciones semánticas son las más apropiadas para las relaciones emisor, receptor, texto? ¿cómo ha de ordenarse el código semántico para soslayar las diferencias en el conocimiento compartido? ¿qué tipo de referencialidad debe aparecer en el texto? En resumen, podemos preguntarnos ¿qué implica afirmar que el texto es significado?

Las respuestas a estas preguntas las hemos buscado a través de la argumentación sobre los conceptos que señalábamos en las variables pragmáticas; de este modo, garantizaremos que la deíxis social del texto parte de la deíxis semántica que se ajusta a las condiciones de la instancia comunicativa en la que se produce la actuación lingüística. Esta posición de la dimensión semántica como ámbito intermedio entre la pragmática y la lingüística nos ha permitido analizar desde un primer plano la noción de función del texto desde unos postulados que permitan su procesamiento como unidad lingüística de significado social. Frente a los modelos previos que ordenaban la función como una suma aleatoria de componentes semánticos y pragmáticos y que, consecuentemente, renunciaban a la posibilidad de argumentar escalarmente sobre el plano pragmático del lenguaje, la argumentación que se produce sobre el componente de referencia social del texto por medio de la semántica nos permite eludir el postulado de que todo significado es, en el fondo, significado social. En otras palabras, sentamos las bases para poder garantizar que únicamente la deíxis textual que facilite la comunicación será funcional, ya que habremos representado las variables de significado y de significado social dentro de un mismo modelo de procesamiento argumental.

Al igual que en el caso de las pragmáticas, procederemos a continuación, y de manera muy breve, a enumerar los parámetros que se han integrado en las variables semánticas indicando, igualmente, la base bibliográfica de la que parten. Los parámetros en los que se ha desglosado la respuesta a la primera pregunta, aquella que se cuestionaba acerca de la base semántica existente entre emisor, receptor y texto –y que enlaza, consecuentemente, con la última pregunta reseñada en las variables pragmáticas–, se centran en la recuperabilidad semántica a partir del tipo –modelo– textual; en nuestro modelo de análisis se distingue tres tipos de acceso al contenido semántico desde el significante textual: **visible**, **semivisible** e **invisible** (Chafe, 1976; Halliday y Hasan, 1976). La segunda de las cuestiones semánticas que debían responderse por medio de variables atendía a la distribución de la carga semántica en el texto teniendo en cuenta también las variables y los parámetros pragmáticos anteriormente señalados; en este sentido, las bases teóricas que han contribuido a formalizar esta respuesta se han basado, fundamentalmente, en las propuestas relacionadas con la tematización en unidades textuales superiores al entorno sintáctico. El resultado se puede cifrar en los siguientes parámetros: **interpersonal básica**, **textual básica**, **textual de comentario**, **textual paralela**, **metacomunicativa de comentario** y **metacomunicativa discursiva** (Danes, 1970; Horn, 1972; Givón, 1983; Matthiessen y Thompson, 1988; Hatim y Mason, 1990). Por lo que hace referencia a la tercera y última de las cuestiones, la aparición y la disposición en la superficie del modelo textual de señales o marcas que apelen a la recuperabilidad de la información, se ha parametrizado, siguiendo los postulados de la *Rhetorical Structure Theory* y de la *Relational*

*Propositions Theory*, en los siguientes valores: **informativa**, **evaluativa**, **social**, **directiva** y **discursiva** (Habermas, 1971; Mann y Thompson, 1988; Mochizuki, Honda y Okumura, 1998; Nagao y Hasida, 1998).

Reconocido, por tanto, el valor de la semántica en la determinación del valor funcional del texto, el siguiente elemento a considerar en la representación del texto como muestra lingüística de orden superior ha sido la formalización de las relaciones pragmáticas y semánticas en la superficie del texto. Hemos afirmado que la deíxis del texto sería la que llevase a cabo la función de relacionar su carácter pragmático con el del contenido; sin embargo, esta deíxis textual ha de presentar marcas que indiquen en qué sentido el texto es referencial. De otro modo, el procesamiento de los textos no podría llevarse a cabo más que de un modo unívoco, desatendiendo las variaciones del texto en cuanto actualización real del lenguaje en una instancia comunicativa.

Para ello, hemos optado por encontrar en el texto los indicios que desde el plano puramente formal apelan al valor social del significado textual. Si tomamos de nuevo el concepto de argumentación para relacionar escalarmente los grupos de variables, la formalización nos permitirá descubrir formas prototípicas de actualizar las relaciones semánticas que son propias de determinados modos de deíxis social. Por tanto, la unidad del texto como convención se convierte en la convención de su forma, sin que esto implique que la normalización de la forma es el único modo de asignar significado.

Así, y siguiendo la metodología que hemos llevado a cabo sobre los grupos de variables anteriores, en este momento hemos buscado respuesta a preguntas como: ¿qué tipo de formalizaciones son propias de las diferentes estructuras semánticas? ¿cómo podemos normalizar la deíxis social por medio de los marcadores de superficie? En conclusión, ¿qué tipo de referencia léxica cubre las necesidades funcionales del texto? La incorporación de la argumentación entre los tres grupos de variables nos ha facilitado el camino hacia la multidimensionalidad en nuestra metodología.

La caracterización paramétrica de las variables que dan respuesta a estas preguntas se ha llevado a cabo del siguiente modo: tomando como base la deíxis de los componentes lingüísticos del texto, ésta se ha caracterizado a partir de dos variables cuya naturaleza formal se muestre preeminente respecto a consideraciones pragmáticas o semánticas; en otras palabras, dado que el lingüista o el traductor ya ha llevado a cabo una labor de desglose pragmático y semántico del componente formal del modelo textual, será dicho componente el que actúe como portador de la iconicidad del modelo textual.

El primer grupo de parámetros responde a la cuestión de las formalizaciones propias de las estructuras semánticas para cada uno de los tipos –modelos– textuales; para ello, nos hemos centrado en el concepto de conjunción (Halliday y Hasan, 1976; Halliday y Martin, 1993), por un lado, y en el de etiquetas y marcadores por otro (Fraser, 1997). La suma de estas dos perspectivas ha posibilitado que nuestra metodología señale no sólo las transiciones entre unidades sintácticas o textuales de orden menor –de lo cual se ocupa la conjunción– sino también las transiciones entre focos temáticos contextuales dentro de cada una de dichas unidades. Dada la prevalencia del análisis formal, el conjunto de parámetros para el procesamiento analítico del texto se ha dividido, a su vez, en torno a tres bloques marcados por la conjunción, a saber: **aditiva**, **causal** y **temporal**, sobre los que se ha operado el concepto de etiquetas y marcadores. Este proceso de delimitación

conjuntiva en nuestra metodología, ha tenido el siguiente resultado: **aditiva simple, aditiva enfática, aditiva de similaridad y aditiva marginal; causalidad per se, causalidad consecutiva, causalidad final y causalidad inversa; temporalidad conclusiva y temporalidad terminativa.**

La segunda pregunta que formulábamos en cuanto a la determinación de variables formales se refiere está estrechamente ligada con el concepto de marcadores de superficie al que habíamos aludido dentro del ámbito semántico; al contrario de lo que ocurría en el caso previo de la conjunción, donde el concepto semántico es secundario a lo que la formalización conjuntiva representa en el plano textual, en esta ocasión el análisis debe nutrirse de las consideraciones acerca de dichos marcadores de superficie. De otro modo, la aplicación de los presupuestos teóricos de la *Veins Theory* (Cristea, Ide y Romary, 1998) o de la cohesión léxica (Halliday y Hasan, 1976; Gallo y Rink, 1979; Hoey, 1991; Fox, 1992; Stede y Umbach, 1998; Sager et alii, 1980) resultaría hasta cierto punto falaz, ya que asumiría una separación irreal entre significante y significado textual. Teniendo en cuenta esta premisa para la necesaria retroalimentación y la constatación del grado de verdad en la marcación de esta variable, los parámetros resultantes han sido los siguientes: las relaciones formales se establecen tomando como foco contextual dominante un hiperónimo textual no especializado –o **nombre general**–, un **hiperónimo genérico** no especializado, un **término** –asumiendo por término el concepto de “especializado”– y, por último, un **sinónimo** –considerando sinónimo cualquier hipónimo “no general” y “no especializado” que encuentre a su hiperónimo dentro del texto.

## 5. APLICACIÓN SOBRE UN CORPUS LINGÜÍSTICO

Una vez establecidos los parámetros que de nuestra aplicación metodológica, procedimos a aplicarlo sobre un corpus de textos. La naturaleza traductológica del estudio nos ofrecía dos posibilidades: bien emplear para la aplicación un corpus de traducciones, o bien utilizar un corpus de textos comparables; entendemos que los corpus comparables son aquellos formados por muestras redactadas originariamente en dos o más lenguas y que presentan una serie de semejanzas (Bravo Gozalo y Fernández Nistal 1998, 226), mientras que corpus de traducciones son los corpus compuestos por textos traducidos o meta de una o de varias lenguas. El carácter funcional de nuestra propuesta nos ha inducido a emplear para la aplicación un corpus de textos comparables, decisión que se explica si consideramos que, por un lado, el corpus paralelo<sup>1</sup> iba a determinar de algún modo las relaciones pragmáticas que se habrían de mostrar en los textos traducidos, lo cual resultaría poco satisfactorio para nuestro propósito. Por otro lado, las relaciones que hemos establecido desde las variables semánticas y lingüísticas podrían mostrar una tendencia a la subordinación respecto a la lengua origen.

Al igual que la primera caracterización de nuestro corpus está determinada por el objetivo de nuestra investigación, las características más específicas que presenta también encuentran su razón de ser en la constatación funcional de los textos a partir de los grupos de variables señaladas. Así, el hecho de establecer una tipología de carácter general nos ha llevado a recopilar textos cuyo grado de especialización no fuera un argumento para la selección. En segundo lugar, si estimamos que los resultados de un corpus serán tanto más significativos cuanto mayor sea su tamaño, hemos

---

1 Bravo Gozalo y Fernández Nistal definen los corpus paralelos de la siguiente forma: “los corpus que están formados por un conjunto de textos fuente (en una lengua de partida) y sus traducciones” (1998, 225).

construido un corpus al que no podemos considerar cerrado sino que definiremos como de tipo monitor, es decir:

*“dynamic rather than a static phenomenon, consisting of very large amounts of electronically-held text which will pass through the computer”* (Renouf 1987, 21).

Esta característica, a su vez, nos lleva a presentar la última instancia de definición del mismo, que tiene que ver con la representatividad del corpus con respecto al conjunto global de textos que potencialmente podrían constituirlo. Desde este punto de vista nuestro corpus podría denominarse como un *sample corpus*, puesto que las selecciones de textos que lo integran son una muestra de las potenciales actualizaciones textuales que pudiéramos encontrar. Teniendo en cuenta estas características, la selección de textos se realizó de manera semi-dirigida; el sesgo al que hacemos referencia al denominar la selección como semi-dirigida tiene que ver únicamente con la intención de garantizar la comparabilidad real de los textos por fecha de publicación, tamaño de las muestras y difusión. En cualquier otro sentido la selección textual se ha desarrollado de manera aleatoria.

Tras la confección del corpus, procedimos a la marcación de los grupos de variables sobre los textos con la intención de demostrar nuestra hipótesis de un modo empírico: la interacción de los parámetros marcados debería dar como resultado el señalamiento de unas distancias absolutas entre los tipos –modelos– textuales pero, al mismo tiempo, daría respuesta también a las desviaciones de los textos como actuaciones reales en cada una de las variables diseñadas en nuestra metodología. De este modo, podríamos apuntar a que el ulterior procesamiento del texto para aplicaciones lingüísticas –en lingüística computacional, por citar sólo un ejemplo– o traductológicas, quedaría por fin satisfecho. La puesta en funcionamiento de nuestra metodología puede verse ejemplificada en el Apéndice que acompaña a este artículo; en los gráficos se muestra el procedimiento para nuestra aplicación tomando siempre dos variables o conjunto de variables de ámbitos diferentes. Se señala en los mismos cuáles son las variables que se han considerado y las distancias absolutas y relativas de los parámetros respecto al núcleo señalado por el tipo textual.

En consonancia con nuestro modelo empírico, la caracterización de los núcleos estadísticos señalados como tipos textuales y su posterior nominalización como **divulgativos, especializados, jurídicos, literarios, periodísticos, publicitarios** u **obras de referencia** se llevó a cabo tras aplicar la misma metodología de análisis a un elemento textual que pudiera considerarse canónicamente típico para cada uno de los casos; consecuentemente, los valores obtenidos para los parámetros de dicho elemento textual sirvieron como invariante de comparación en nuestra aplicación.

Tras la aplicación, y dado que la interpretación de los resultados teniendo como herramienta de trabajo corpus muy poco numerosos, si bien representativos, hemos procedido a la implementación de la metodología de corpus con otra de corte estadístico. En este sentido, hemos optado por el análisis factorial de correspondencias porque ofrecía una explicación más satisfactoria a la amplia gama de textos y de variables de las que disponíamos. La selección de este modelo de análisis nos ha permitido llevar a cabo, a su vez, el concepto algorítmico de función que habíamos postulado como básico para nuestra propuesta tipológica: el resultado de cada una de las variables para cada uno de los textos tendrá el mismo valor, lo que redundará en la separación entre los tipos textuales y la función. Como se ha podido observar, las diferencias que surgen entre los diferentes textos según las variables pueden constatarse incluso de un modo visual.

## 6. CONCLUSIONES

A la vista de los resultados de nuestra aplicación metodológica hemos alcanzado las siguientes conclusiones: En primer lugar, el procesamiento de los textos que se encierra en nuestra tipología se ha mostrado eficaz para dar respuesta no sólo a la normalización del texto como déictico social sino también a los comportamientos que conducen a dicha normalización. Partiendo de este planteamiento encontramos explicación a la dimensionalidad que manifiestan ciertos parámetros o grupos de parámetros integrados en los textos respecto a las ocurrencias normativas del texto virtual correspondiente.

Íntimamente relacionada con ésta se encuentra la cuestión funcional. Frente a modelos taxonómicos previos, los datos nos han permitido señalar a la noción de función como suma de los valores pragmáticos, semánticos y formales que integran el texto. La perspectiva lingüística que hemos añadido al concepto funcional nos ha ofrecido la posibilidad de señalar la naturaleza que posee el texto como déictico social y textual. En consecuencia, la convencionalización del texto, o de los tipos textuales, como elemento social adquiere por un lado el carácter normativo necesario para desempeñar su labor comunicativa y, por tanto, funcional. Por otro lado, permite la variación del texto como déictico textual, lo que integrará a los elementos pragmáticos, semánticos y lingüísticos, las herramientas de la equivalencia en los Estudios de Traducción, con el carácter funcional que en realidad encierran.

En resumen, si retomamos la cita de Sager, podremos concluir que en nuestra propuesta hemos intentado aunar la cuestión del significado social a través de la forma –en nuestro caso los textos o los tipos textuales– con los códigos que los integran sin perder de vista que la normalización del significado siempre la realizarán los individuos –las desviaciones y nunca las sociedades.

## 7. BIBLIOGRAFÍA:

- BASSNETT MC GUIRE, S. (1997). "Text types and Power Relations". En TROSBORG, A (ed), *Text Typology and Translation*. Amsterdam: John Benjamins, pp 87-99.
- BHATIA, V. (1993). *Analysing Genre: Language Use in Professional Settings*. Londres: Longman.
- BEAUGRANDE, R. (1980). *Text, Discourse and Process: Towards a Multidisciplinary Science of Texts*. Norwood, N.J: Ablex.
- BIBER, D. (1989). "A Typology of English Texts", en *Linguistics* 27, pp 3-43.
- : (1994). "Corpus-Based Approaches to Issues in Applied Linguistics" en *Applied Linguistics* 15:2, pp 169-189.
- BIBER, D & FINEGAN, E. (1994). *Sociolinguistic Perspectives on Register*. Nueva York: O.U.P.

- BRAVO GOZALO, J.M y FERNÁNDEZ NISTAL, P (1998). “La lingüística del corpus, las nuevas tecnologías de la información y los Estudios de Traducción en la década de 1990”. En FERNÁNDEZ NISTAL, P y BRAVO GOZALO, J. M. (eds). *Traducción: orientaciones lingüísticas y culturales*. Valladolid: Universidad de Valladolid, pp 205-257.
- BÜHLER, K. (1965). *Sprachtheorie*. Stuttgart: Fischer.
- CHAFE, W.L (1976). “Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View”. En LI, CH (1976), *Subject & Topic*. Nueva York: Academic Press, pp 20-55.
- CHESTERMAN, A. (1996). “On similarity”, *TARGET* 8 (1), PP 159-164.
- : (1998). *Contrastive Functional Analysis*. Amsterdam: John Benjamins.
- CRISMORE, A & FARNSWORTH, R. (1990). “Metadiscourse in Popular and Professional Science Discourse”. En NASH, W (ed), *The Writing Scholar: Studies in Academic Discourse*. Newbury Park: Sage Publications, pp 118-136.
- CRISTEA, D, IDE, N & ROMARY, L (1998). “Veins Theory: A Model of Global Discourse Cohesion & Coherence”, *COLING-ACL* 17, págs. 281-285.
- DANEŠ, F (1970). “Thur Linguistischen Analyse der Textstruktur”, *FOLIA LINGUISTICA*: 4, pp 72-78.
- : (1970/1976). “Functional Sentence Perspective and the Organization of the Text”, en DANEŠ, F (ed), *Papers on Functional Sentence Perspective*, Praga: Academia.
- : (1987). “On Prague School Functionalism in Linguistics”, en DIRVEN, R & FRIED, V (eds.), *Functionalism in Linguistics*, Amsterdam: John Benjamins, pp 3-38.
- DANEŠ, F & VIEHWEGER, D (eds.) (1977). *Probleme der Textgrammatik*, Berlín, Akademie.
- DELISLE, J (1980). *L'analyse du Discours comme Méthode de Traduction*. Ottawa: University of Ottawa.
- EVEN-ZOHAR, I (1979). “Polysystem Theory”, *POETICS TODAY*, 1 (1)-1 (2), pp 287-310.
- : (1990). “Polysystem Studies” Special issue of *POETICS TODAY*, XI(1).
- FERNÁNDEZ ANTOLÍN, M.J (2002). *La problemática de los tipos textuales inglés / español en los Estudios de Traducción*. Chadwyck.
- FILLMORE, Ch. (1968). “The Case for Case”. En BACH, E & HARMS, R (eds.). *Universals in Linguistic Theory*, Nueva York: Holt, pp 1-88.

- FILLMORE, Ch. (1975). "Santa Cruz Lectures on Deixis". Indiana University: inédito.
- FINEGAN E. & BIBER, D. (1994). "Introduction, Situating Register in Sociolinguistics". En BIBER, D & FINEGAN, E (eds.). *Sociolinguistic Perspectives on Register*. Nueva York: O.U.P, pp 3-10.
- FOX, C (1992). "Lexical Analysis and Stoplists". En FRAKES, W & BAEZAYATES, R (eds.), *Information Retrieval: Data Structures and Algorithms*. Upper Saddle River: Prentice Hall, pp 102.-130.
- FRASER, B (1997). "Commentary Pragmatic Markers in English", en *Estudios Ingleses de la Universidad Complutense* 5, pp 115-127.
- GALLO, J & RINK, H (1979). *Shaping College Writing*, Nueva York: Harcourt Brace.
- GAZDAR, G (1979). *Pragmatics: Implicature, Presupposition and Logical Form*. Nueva York: Academic Press.
- GIVÓN, T (1983). *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins.
- GÖMMLICH, K. (1993). "Text Typology and Translation-Oriented Text Analysis". En WRIGHT, S.E & WRIGHT, L.D (eds.), *Scientific and Technical Translation*. Amsterdam: John Benjamins, pp 175-183.
- GRICE, H.P (1975). "Logic and Conversation". En COLE, P & MORGAN, J.L (eds.). *Syntax and Semantics, vol. iii: Speech Acts*. Nueva York: Academic Press, pp 41-58.
- GROSZ, B & SIDNER, C. (1986). "Attention, Intention and the Structure of Discourse", en *Computational Linguistics* 12, pp 175-204.
- GROSZ, B. JOSHI, A & WEINSTEIN, S. (1995). "Centring a Framework for Modelling the Local Coherence of Discourse", en *Computational Linguistics* 12 (2), pp 203-225.
- HABERMAS, J (1971). "Vorbereitende Bemerkungen zu einer Theorie der Kommunikativen Kompetenz". En HABERMAS, J & LUHMANN, N (eds.), *Theorie der Gesellschaft oder Sozialtechnologie*. Frankfurt: Surhkamp, pp. 101-141.
- HALLIDAY, M.A.K. (1978). *Language as Social Semiotics*, Londres: Arnold.
- : (1985). *Language, Context and Text: aspects of language in a social-semiotic perspective*, Geelong: Deakin University Press.
- : (1993). "The Discursive Technology of Science". En HALLIDAY, M.A.K. & MARTIN J.R (eds.) (1993), *Writing Science. Literacy and Discursive power*. Londres: The Falmer Press, pp. 1-50.

- HALLIDAY, M.A.K. & HASAN, R (1976). *Cohesion in English*, Londres: Longman.
- HATIM, B & MASON, I (1990). *Discourse and the Translator*, Londres: Longman.
- HOEY, M (1991). *Patterns of Lexis in Text*, Oxford: O.U.P.
- HOLUB, R.C (1984). *The Reception Theory*, Londres: Routledge.
- HORN, L.R (1972). *On the Semantic Properties of the Logical Operators in English*, Bloomington: Indiana University Linguistics Club.
- HOSNER, D.W & LEMESHOW, S. (1989). *Applied Logistic Regression*. Nueva York: John Wiley and Sons.
- HOUSE, J (1981). *A Model for Translation Quality Assessment*, Tubinga: GNV.
- ILYAS, A. (1994). "A Typology for Translation". En BEAUGRANDE, R ET ALII (eds.). *Language, Discourse and Translation*. Amsterdam: John Benjamins, pp 45-53.
- KARTUNNEN, L & PETERS, S (1975), "Conventional Implicature in Montague Grammar", *Proceedings of the first annual meeting of the Berkley Linguistic Society*, pp 266- 278.
- KOLLER, W. (1995). "The Concept of Equivalence and the Object of Translation Studies", *TARGET* 7(2), pp 191-222.
- LEVINSON, S.C (1979). "Activity Types and Language", *LINGUISTICS*, 17, pp 367-399.
- : (1983). *Pragmatics*, Cambridge: C.U.P.
- LYONS, J (1977). *Semantics*, Cambridge: C.U.P.
- MCCULLAGH, P & NELDER, J.A. (1983). *Generalized Linear Models*. Londres: Chapman & Hall.
- MANN, W.C & THOMPSON, S. (1988). "Rhetorical Structure Theory: Toward a Functional Theory of Text Organization", *TEXT* (3), pp 243-281.
- MARCUS, M ET ALII (1994). "Building a Large Annotated Corpus of English: the Penn Treebank", *COMPUTATIONAL LINGUISTICS*, 19, pp 313-330.
- MATTHIESSEN, C & THOMPSON, S (1988), "The Structure of Discourse and Subordination". En HAIMAN, J & THOMPSON, S (eds.), *Close Combining in Discourse and Grammar*. Amsterdam: John Benjamins, pp 78-95.
- MCROY, W. (1992). "Using Multiple Knowledge Sources for Word Sense Discrimination", en *Computational Linguistics* 18(1), pp 1-30.

- MOCHIZUKI, H, HONDA, T & OKUMURA, M. (1998). "Text Segmentation with Multiple Surface Linguistic Cues", en *Coling-Acl* 17, pp 881-885.
- NAGAO, K & HASIDA, K. (1998). "Automatic Text Summarization Based on the Global Document Annotation", en *Coling-Acl* 17, pp 917-921.
- NORD, C. (1991). *Text Analysis in Translation. Theory, Methodology and Didactic Application of a Model for Translation Oriented Text Analysis*. Amsterdam: Rodopi.
- OKUMURA, M & HONDA, T. (1998). "Word Sense Disambiguation and Text Segmentation Based on Lexical Cohesion", en *Proceedings of the Fifteenth International Conference on Computational Linguistics*, pp 755-761.
- OSGOOD, C. ET ALII (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.
- RABADÁN, R. (1991). *Equivalencia y Traducción. Problemática de la equivalencia transléctica inglés-español*. León: Universidad de León.
- REISS, K. (1976). *Texttyp und Übersetzungsmethode*. Kronberg: Scriptor.
- REISS, K. & VERMEER, H. (1984). *Grundlegung einer allgemeinen Translationstheorie*. Tubinga: Max Niemeyer.
- RENOUF, A. (1987). "Corpus Development". En SINCLAIR, J (Ed.) (1987), *Looking up*. Londres: Harper Collins Publishers, pp 1-40.
- ROBERTS, R. (1995). "Towards a Typology of Translations". En *Hieronimus Complutensis*, 1, pp 69 -78.
- RUSSEL, B (1905). "On Denoting", *MIND* 14, pp 479-493.
- SAGER, J.C ET ALII (1980). *English Special Languages*. Wiesbaden: Branstetter Verlag.
- SAGER, J.C (1993). *Language Engineering and Translation. Consequences of Automation*. Amsterdam: John Benjamins.
- SNELL B & CRAMPTON, P. (1983). "Types of Translations". En PICKEN, C. (eds.). *The Translator's Handbook*. Londres: Aslib, pp 129-139.
- SOVRAN, T (1992). "Between Similarity and Sameness", *JOURNAL OF PRAGMATICS*, 18.4, pp 329-344.
- STEDE, M. & UMBACH, C. (1998). "Dimlex: a Lexicon of Discourse Markers for Text Generation and Understanding, *Coling-Acl*, 17, pp 1238-1242.

- TOURY, G (1980). *In Search of a Theory of Translation*, Tel Aviv: The Porter Institute for Poetics and Semiotics.
- : (1981). “Translated Literature: System, Norm, Performance”, *POETICS TODAY* 2 (4), pp 9-27.
- : (1985). “A Rationale for Descriptive Translation”. En HERMANS, T (ed) *The Manipulation of Literature. Studies in Literary Translation*. Londres: Croom Helm, pp 16-41.
- : (1995). *Descriptive Translations Studies and Beyond*, Amsterdam: John Benjamins.
- TROSBORG, A. (ed) (1997). *Text Typology and Translation*. Amsterdam: John Benjamins.
- TRUJILLO, R. (1996). *Principios de semántica textual*. Madrid: Arco.
- WATANABE, H. (1996). “A Method for Abstracting Newspapers Articles by Using Surface Clues”, en *Coling-Acl*, pp 974-979.
- WERLICH, E (1976). *A Text Grammar of English*, Heidelberg: Quelle & Meyer.
- WIERZBICKA, A (1991). *Cross-Cultural Pragmatics. The Semantics of Human Interaction*, Berlín: Mouton de Gruyter.
- : (1996). *Semantics. Primes and Universals*, Oxford: O.U.P.
- : (1975). *Presupposition & Non-Truth Conditional Semantics*, Londres: Academic Press.

## 8. APÉNDICE

### *Abreviaturas, variables y dimensiones*

Antes de la exposición de resultados, creemos necesario incluir una tabla con las abreviaturas de las variables y de los parámetros:

Intención='Int'	Implicaturas='Impl'
Estatus TO='EsTO'	Tematización='Tema'
Referencia='Ref'	Conjunción aditiva='Adit'
Conjunción temporal='Temp'	Cohesión léxica='Cohlex'
Receptor='Rec'	
Marcación semántica='Msem'	
Conjunción causal='Caus'	

Int

1= "Refer" 2= "Conat" 3= "Expres" 4= "Textual"

Impl

1= "Basica" 2= "Paralela" 3= "Comentar" 4= "Dicursi"

Rece

1= "Interp" 2= "Textual" 3= "Metacom"

EsTO

1= "Convenc" 2= "Particul" 3= "General"

Tema

1= "Intbas" 2= "Textbas" 3= "Textpara" 4= "Textcom" 5= "Metadisc" 6= "Metacom2"

Msem

1= "Infor" 2= "Evaluat" 3= "Social" 4= "Directiv" 5= "Discursiv"

Ref

1= "visible" 2= "invisib" 3= "semivis"

Adit

1= "simple" 2= "enfátic" 3= "similar" 4= "margin"

Caus

1= "perse" 2= "consecut" 3= "final" 4= "inversa"

Temp

1= "conclusi" 2= "terminat"

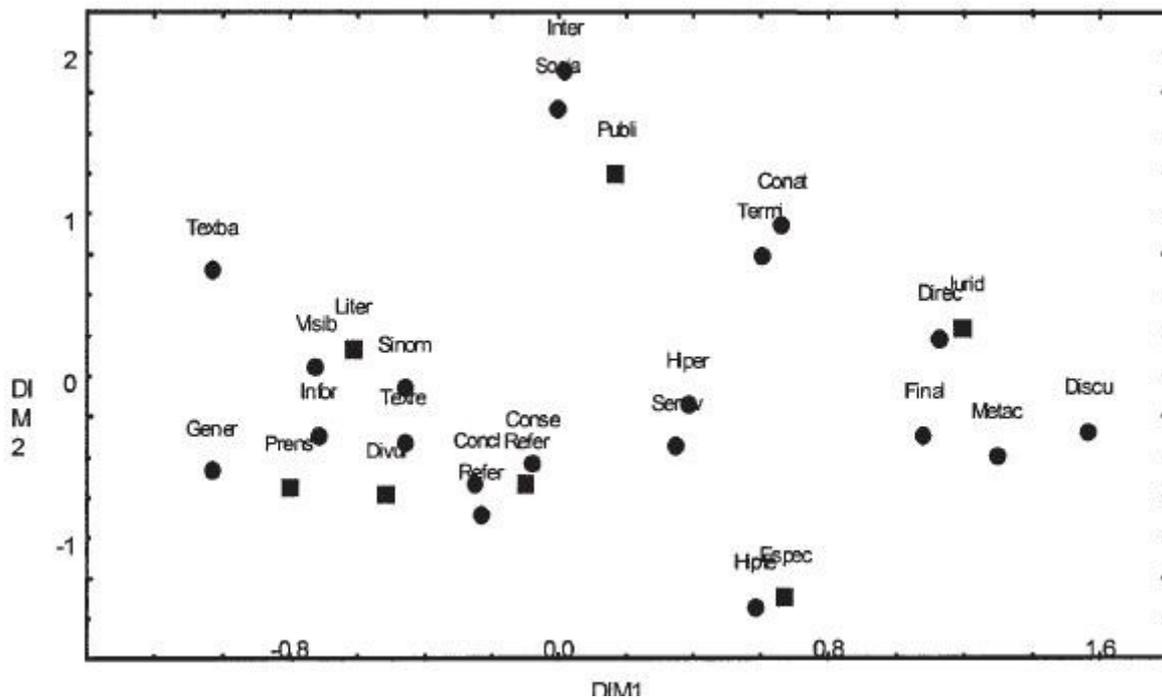
Cohlex

1= "nomgen" 2= "hiperon" 3= "Hiperter" 4= "Sinonim".

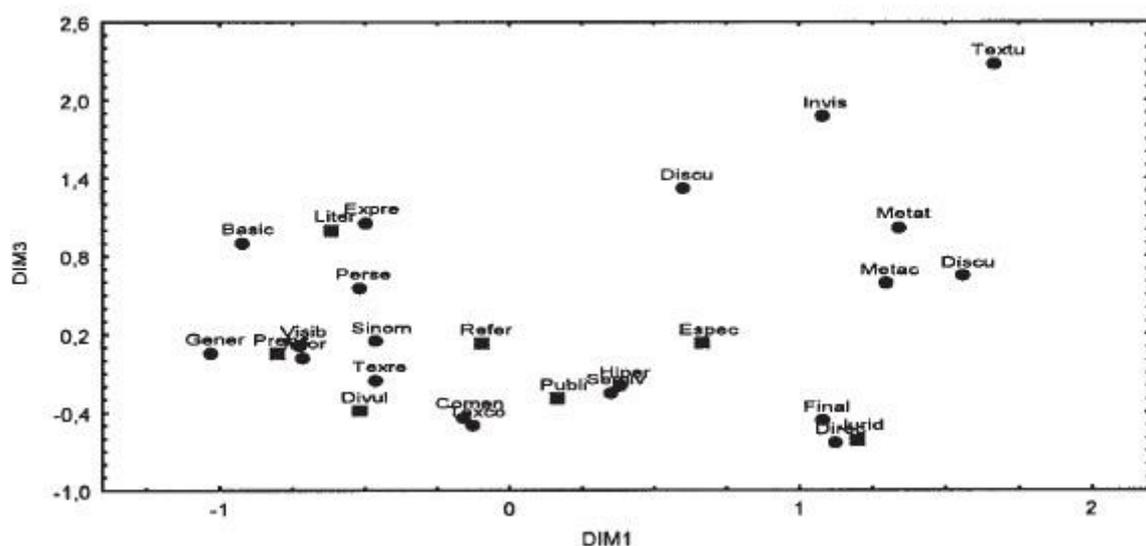
*Calidad de la representación para cada uno de los tipos modelos textuales*

Divu	0.108923
Espe	0.425699
Jur	0.571911
Litl	0.257889
Pres	0.151449
Publ	0.509800
Refr	0.039673
esp	0.010981
ingl	0.010981

**Gráfico 1: Caracterización de la Tematización y de la Intención.**  
*Los parámetros sobre las posiciones de los textos modelos- textuales.*



**Gráfico 2: Caracterización de la Tematización, la Intención, la Referencia y la Conjunción.**  
*Los parámetros sobre las posiciones de los modelos textuales.*



*Gráfico 3: Caracterización del Estatus del Texto Origen y de la Marcação Semántica. Dimensionalidad de estos parámetros sobre las posiciones de los tipos modelos- textuales*

