

La Audiodescripción desde la representación del conocimiento general. Configuración semántica de una gramática local del texto audiodescrito

Catalina Jiménez Hurtado

Universidad de Granada

Audio Description is a new text type which offers the prototypical receiver – blind or partially sighted people – a narrated representation of what is occurring at specific moments in the other audiovisual text to which it is subordinated. An AD script is subject to multiple subordination: to what is happening on the screen, to the distribution of silent spaces in the other text, and to the amount of time provided in each space. This fact, however, does not prevent the AD script from offering a representation of the general knowledge encoded by specific fragments of the film. Such a representation of knowledge, understood as the description in words or other signs of a specific reality, may serve as a semantic basis for the creation of a local grammar of AD scripts.

1. Introducción: metodologías y análisis

La aparición de un nuevo objeto de estudio supone la aplicación y validación de antiguas metodologías de análisis científico, así como la posible creación de nuevas estructuras y herramientas de análisis que evidencien la adaptabilidad y capacidad de implementación de una teoría. De hecho, este tipo de procesos implica una renovación en la creatividad científica y la puesta en marcha de una maquinaria de análisis que forzosamente ha de sufrir los oportunos y convenientes cambios y adaptaciones al nuevo medio sin que por ello sufran sus fundamentos teóricos y empíricos.

Uno de los objetos de estudio que ha surgido en esta nueva etapa de la pretendida accesibilidad universal a los medios de comunicación es el fenómeno de la audiodescripción plasmada en un texto que llamaremos *guión audiodescrito* (en adelante GAD) y que se configura lingüísticamente como un texto que se construye con el fin de describir para las personas ciegas o de baja visión lo que ocurre en la pantalla durante los silencios de la película, serie o documental (Ballester en prensa).

Este fenómeno obliga a los estudiosos de la lengua al menos a dos reflexiones: por un lado, es necesaria la aplicación de metodologías clásicas de análisis, así como elucubrar en torno a la validez de las mismas o en caso contrario, obligaría a desarrollarlas y ampliarlas. Por otro lado, con los resultados de un primer acercamiento de análisis, es responsabilidad de las instituciones crear un protocolo de formación para los nuevos profesionales y su ubicación en un ámbito académico y curricular. Sólo así se evitaría uno

de los problemas sociales universalmente heredados, esto es, la invisibilidad social e institucional de los nuevos profesionales, los audiodescriptores.

En lo que sigue, plantaremos algunas perspectivas de análisis de un nuevo tipo textual, el GAD, sus parámetros y características desde el análisis lingüístico y traductológico.

2. Describir lo que se ve: una nueva forma de representar el conocimiento

La lingüística en todas sus vertientes, enfoques y variantes lleva algunos años de enhorabuena, gracias a la aparición de nuevos casos de estudio debido a la proliferación de tipos textuales que se presentan en formatos también novedosos como es, por ejemplo, el caso de los textos que aparecen en Internet. El hecho de que los contenidos hayan de presentarse de una forma determinada condiciona la selección de los mismos, así como el estilo lingüístico.

Con la aparición en el mercado de un nuevo tipo textual como es el GAD hemos de indagar en las formas de abordarlo desde diferentes enfoques. Muchos de estos enfoques lingüísticos y traductológicos se basan en la observación previa de las características semánticas y gramaticales de los nuevos textos (Bourne y Jiménez 2007).

2.1. Metodología de análisis: el corpus

La idea de representar el conocimiento está relacionada con la necesidad de recopilar de forma estructurada datos sobre objetos, eventos o relaciones entre ellos, de modo que se puedan almacenar y recuperar con facilidad. Representar el conocimiento, en palabras de Faber (2002: 3)

Significa escribir mediante algún tipo de lenguaje u otro medio de comunicación descripciones, símbolos o imágenes que corresponden a un estado de cosas en el mundo.

A pesar de la aparente sencillez del proceso, este tipo de descripciones se está convirtiendo en algo muy complejo y laborioso y, sobre todo, solo podrá ser creíble si está basado en el análisis semiautomático de un corpus realmente representativo.

Los métodos empíricos y estadísticos de análisis lingüístico con un razonable componente de intervención humana han resultado ser muy fructíferos a lo largo de la historia de la lingüística. Como bien afirma Pérez Hernández (2002: 129): “un corpus se puede usar de formas y con fines muy diferentes: para *validar*, para *ejemplificar* o para *construir* una teoría de la lengua y los diferentes aspectos que ésta implica”.

La consecuencia más interesante y valiente de utilizar un corpus y su metodología de análisis correspondiente es el hecho de que con ello estu-

diamos el uso de una lengua en un contexto textual y social determinado, confiando sólo en parte en lo que supone la intuición lingüística y afianzando la idea de una metodología descriptiva de los recursos lingüísticos utilizados.

Sin embargo, este proceso no está exento de dificultades; de hecho, la recopilación de un corpus debe hacerse teniendo siempre en cuenta unos criterios específicos entre los que destaca el hecho de que ha de estar compuesto por textos reales y actualizados en un contexto dado. En el caso de los GAD hemos de clasificarlo además por géneros fílmicos que son, en última instancia, formas de eventos comunicativos.

Para este estudio hemos recopilado un total de 149 películas de diferentes nacionalidades audiodescritas en español. Este corpus consta de un millón de palabras aproximadamente.

2.2. Concepto y finalidad de la representación del conocimiento

A pesar del avance que han experimentado estos estudios en la representación del conocimiento experto (RCE), resulta sorprendente su falta de aplicación a la representación del conocimiento *general*.

La RCE ofrece las bases gramaticales de los textos especializados de diferentes ámbitos. Describe en un lenguaje natural la forma en que los expertos producen los textos de su ámbito de saber, lo que a su vez permite crear una imagen de la organización de ese conocimiento por subáreas o categorías (Jiménez Hurtado 2007, Jiménez y Seibel 2004).

Si con una estructura de representación del conocimiento se describe una imagen conceptual de la organización de una realidad, de algo que está ocurriendo, esto es, de un evento, no nos puede extrañar que este fenómeno evite eventos con carácter general. Habría que buscar textos auténticos que describan eventos cotidianos con una función determinada para un tipo de receptor prototípico. Con estos requisitos, hasta ahora sólo se podía pensar en textos pertenecientes al ámbito de la ficción o los periodísticos cuando describen hechos generales acaecidos en un momento y lugar dados. Este tipo de estudios no se ha abordado precisamente por la magnitud del trabajo y quizá por la falta de funcionalidad del mismo.

Traboulsi *et al.* (2004) están trabajando precisamente en la representación de determinado tipo de eventos, en parte generales, que han sido aislados de textos periodísticos. A pesar de ser textos de divulgación o generales, éstos se encuentran siempre en secciones temáticas especializadas de los periódicos como la sección economía, deportes, etc. Los propios autores hablan de *dominio específico*:

We present a corpus-based method for identifying and learning patterns describing events in a specific domain by examining the manner in which: (a) a small number of keywords in the domain are distributed throughout the corpus; and, (b) a local grammar that is idiosyn-

cratic of a class of events in the domain, governs the usage of the keywords. (2004: 749)

El método que presentan, muy cercano al que se utiliza en la representación del conocimiento experto, se basa en el análisis de frecuencias, de las colocaciones y de las concordancias resultantes de la búsqueda semiautomática de determinados conceptos.

En primer lugar, se seleccionan los lexemas de uso más frecuente, se identifican las colocaciones y se crean concordancias que posteriormente se vuelven a examinar. De este modo se pueden describir los patrones semánticos, sintácticos y/o morfosintácticos que registra el tipo de textos representado en el corpus. Es evidente que las palabras más frecuentes son aquellas que pertenecen a clases cerradas como artículos, determinantes, etc. Sin embargo, la búsqueda de colocaciones y su análisis se inicia con los lexemas de clases abiertas.

La frecuencia de uso simultáneo de un lexema de clase abierta con otra serie de lexemas, es decir, la recurrencia estadísticamente relevante de ciertos patrones semánticos y sintácticos es considerada funcional y, por tanto, significativa para incluirla en un esquema gramatical que represente a ese tipo textual. Por ejemplo, los autores descubrieron que el uso simultáneo del lexema *rose* junto a *fell*, *up* y *down* más un número cardinal y la palabra *percent*, es la fórmula más utilizada para describir los movimientos de caída y subida de los precios de las acciones de determinadas empresas. El corpus de estos análisis procede de textos periodísticos y, por tanto, divulgativos no especializados; sin embargo, el hecho de que se haya utilizado la sección de economía implica que no se pueda denominar en rigor representación del conocimiento general.

En este sentido, el GAD se puede considerar, desde muchas perspectivas, un tipo textual que representa el conocimiento general. En primer lugar, describe, excepto cuando lo hace con documentales, escenas cotidianas que fundamentalmente incluyen acciones, descripción de personajes y ambientes como paisajes o interiores. En segundo lugar, esos eventos describen escenas concretas de una realidad-ficción incluida en otro texto más amplio que tiene una función comunicativa determinada. La función comunicativa del texto fílmico (texto origen) antes de ser audiodescrito impone en el audiodescriptor una selección funcionalmente relevante de lo que va a recoger en su GAD (texto meta A). Este texto está, por un lado, subordinado a la función comunicativa del texto origen y, por otro, a la imagen concreta, incluida la pista sonora, descrita en el GAD. Ambas restricciones hacen que el GAD represente un conocimiento general que podemos clasificar en tipos de estructuras conceptuales o categorías cognitivas subordinadas a una imagen que siempre podemos recuperar y analizar, relacionando así imagen visual con imagen escrita. En palabras de Salway (2007):

The relationship between the visual and the verbal, between vision and language and between image and text has fascinated scholars in

many disciplines for many centuries. Audio description provides a novel, tangible and important scenario for exploring these relationships, and it presents immediate and stimulating challenges for research and audiovisual translation.

La tercera restricción que sufre este texto la impone su propia función comunicativa: ofrecer a las personas ciegas y de baja visión una descripción adecuada de las imágenes de la película sólo en los momentos en que no exista emisión acústica. *Adecuada* en este caso significa informativa y relevante: no ofrecer ni demasiada información, ya que la persona ciega no podría procesarla, ni demasiado poca de modo que la persona ciega no acceda a la función comunicativa del texto original.

Este nuevo tipo textual permite, por tanto, acotar realidades generales con una función comunicativa doble: adecuación al receptor prototípico o persona ciega y adecuación a la función comunicativa del texto original.

3. Construyendo una gramática local del guión audiodescrito

3.1. La base metodológica

Tanto la guía de la AD en Gran Bretaña (*ITC Guidance on Standards for Audio Description*), como la Norma AENOR (Norma UNE 153020) española ofrecen unas pautas concretas de los objetos y eventos que han de describirse (personajes, acciones cada vez que se produce un cambio de las mismas, ambientación en general y localizada en un espacio y un tiempo dados, etc.) así como de sus cualidades (color, tamaño, propiedad, etc.).

La existencia de estas indicaciones simplifica en un principio la búsqueda y facilita la creación de primeras pautas de comportamiento textual. De hecho, parte del trabajo del grupo de investigación Tracce¹ se dedica a etiquetar los GAD teniendo en cuenta estas directrices de modo que se puedan recuperar con facilidad formas de audiodescribir la misma función comunicativa, en diferentes tipos de guiones y géneros fílmicos.

Por lo tanto, etiquetar en los GAD los elementos que, según estas normas, son susceptibles de ser descritos y analizar con un corpus las concordancias resultantes o la recuperación de colocaciones por etiquetas mostraría los recursos lingüísticos que utilizan los audiodescriptores para tal fin.²

Sin embargo, en estos momentos, la idea de trabajar con corpora también se debe al hecho de que queramos analizar los patrones lingüísticos recurrentes que se utilizan en la creación de los GAD, para de esta manera descubrir los primeros pasos metodológicos que nos ayuden a establecer una gramática de este tipo textual.

Para ello, es necesario buscar patrones en la utilización de la semántica, la morfosintaxis, las estructuras sintácticas más utilizadas, así como los elementos pragmático-discursivos que se suelen activar. Se trata de

seleccionar precisamente una serie de patrones lingüísticos recurrentes para expresar cualquier tipo de evento (suceso, estado, proceso, cualidad o relación) del que se predique algo.

3.2. La estructura semántica de la gramática local del GAD

Una gramática es cualquier conjunto de reglas que dan cuenta de las regularidades de estructura que manifiesta un determinado lenguaje. Las unidades de análisis que utiliza pueden ser estructurales (las partes de la oración), de significado (los casos) o centrarse en la estructura semántica del lexicón primario de la lengua en cuestión que es la forma en que nosotros iniciaremos el análisis.

Una gramática local, por su parte, es la descripción del comportamiento semántico y sintáctico de una categoría léxica (o parte de ella) que presenta cierta homogeneidad. Las gramáticas locales se crean para describir un uso determinado de una lengua de especialidad o simplemente para reconocer patrones de uso lingüístico en determinados tipos textuales.

El giro cognitivista de la semántica léxica y sus aplicaciones textuales indagan sobre cómo se organizan los conceptos en la mente humana, esto es, las formas de categorizar la realidad, para trasladar sus resultados al ámbito de la ingeniería del conocimiento. Los trabajos más recientes revelan que las relaciones léxicas del lexicón de una lengua natural codifican a su vez un mapa o red de relaciones conceptuales. De hecho, el léxico de una lengua es un almacén dinámico de conocimiento cuya configuración codifica la información que compartimos sobre el mundo, por lo que se puede considerar que un lexicón organizado y jerárquicamente estructurado de una lengua natural es un tipo de gramática que simboliza nuestro conocimiento del mundo a través de las expresiones lingüísticas. En palabras de Faber y Mairal (1997: 15): “In this respect lexical units are to some extent ‘translations’ of our interpretation of perceptual data because they and their combinatorial possibilities are the semantic representation of or model of the world”.

El significado se considera una representación interna del conocimiento que se reparte en categorías cognitivas que son las encargadas de elaborar y procesar la información. Esta especie de lenguaje del pensamiento se descubre analizando los lexemas estructurados en categorías y sus respectivos predicados. Cada lexema del lexicón pertenece a una categoría semántica o esquema cognitivo que codifica la información de nuestra experiencia mental y física sobre el mundo y nuestra forma de interactuar en él (Talmy 1996). Después de analizar unos 15.000 verbos del inglés del lexicón primario Faber y Mairal (1999: 88) llegan a la conclusión de que éstos se pueden catalogar en las siguientes categorías: existencia (ser/estar), cambio (cambiar), posesión (tener/poseer), habla (decir/hablar), sentimiento (sentir), acción general (hacer), cognición (saber/creer), movimiento (mover), percepción general (notar/percibir), percepción física (ver/oír/gustar/oler/tocar) y posición (estar/poner). Cada una de estas diez

categorías incluye verbos cuya relación interna básica es la jerárquica. Cuanto más arriba se encuentre un verbo en la jerarquía, más posibilidades tendrá de ser utilizado en textos generales y de ocupar un lugar preponderante en la creación de extensiones metafóricas, lo que a su vez significa que se trata de un verbo con un significado muy general y de una gran extensión.

4. La gramática de una categoría en los GAD: el MOVIMIENTO

4.1. El MOVIMIENTO desde la semántica léxica

Una de las cuestiones que señala la Norma española UNE de la audiodescripción es la necesidad de describir los cambios que ocurren en pantalla, así como las acciones que realizan los personajes en la misma. Desde un punto de vista de la semántica cognitiva, tanto el concepto de *cambio* como el de *acción general* implican un patrón subyacente de MOVIMIENTO. La relevancia de una categoría dentro del sistema semántico de una lengua está directamente relacionada con la cantidad de conceptos que se categoricen dentro de la misma. La categoría MOVIMIENTO es de las más importantes, tanto por su complejidad estructural, como por la cantidad de lexemas que se incluyen en ella (Cifuentes Honrubia 1999). La imagen que sigue es la estructuración de la categoría de MOVIMIENTO según la distribución de los verbos que incluye.

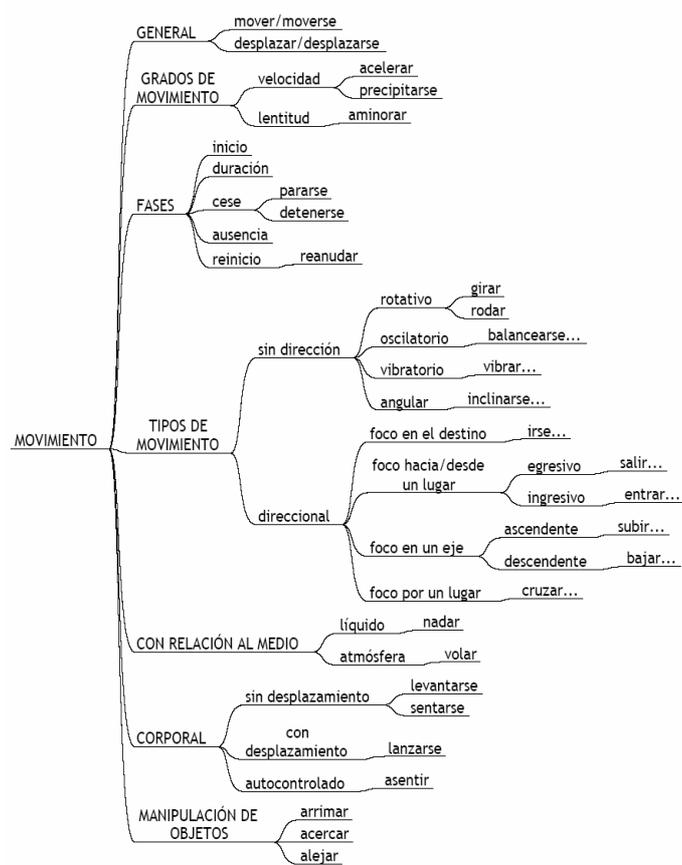


Figura 1: el movimiento desde la semántica léxica

4.2. El MOVIMIENTO en los GAD

Si observamos la lista de frecuencias del corpus de guiones audiodescritos, el primer lexema semánticamente relevante es *mira*³ en el lugar 19, aunque le siguen muy de cerca *sale*, *entra* y *acerca*, en los lugares 28, 29 y 30 respectivamente, verbos pertenecientes a la categoría movimiento.

17	POR	6.857	0,64
18	LAS	6.804	0,63
19	MIRA	5.020	0,47
20	SOBRE	3.685	0,34
21	LO	3.304	0,31
22	HACIA	3.191	0,30
23	ELLA	3.102	0,29

24	ESTÁ	3.032	0,28	
25	SUS	2.999	0,28	
26	ÉL	2.803	0,26	
27	PUERTA	2.776	0,26	
28	SALE	2.503	0,23	
29	ENTRA	2.372	0,22	
30	ACERCA	2.336	0,22	
31	MANO	2.247	0,21	
32	CASA	2.093	0,19	

Analizar los eventos de textos especializados entraña relativamente poca dificultad ya que éstos describen una realidad delimitada en una serie de categorías recurrentes. Salway (2007) explica que ha descubierto cuatro tipos de eventos en los GAD. El evento de *cambio de escena*, el evento del *foco de atención de los personajes*, el evento de la *comunicación no verbal* y el evento de *cambio de situación*.

Analizados con detenimiento, los cuatro se ven cognitivamente afectados por el concepto de MOVIMIENTO. El primero, porque un cambio de escena necesita del movimiento para realizarse; el segundo, porque el foco de atención implica igualmente un movimiento continuo de la mirada (ya sea de la cámara o del personaje); el tercero, porque la comunicación no verbal suele conllevar un movimiento corporal y, el último, porque un cambio de situación sería imposible sin un tipo de movimiento. Podemos por tanto afirmar que después de la categoría cognitiva de la *percepción física* por la vista (mirar/ver), la siguiente categoría posiblemente sea la de movimiento.

Tras haber extraído concordancias de los verbos que lideran las diferentes categorías semánticas, siempre conjugados en tercera persona del singular del presente de indicativo, hemos obtenido los siguientes resultados: la *percepción física* es evidentemente la más sobresaliente con 7394 ocurrencias (mira*), le sigue la de *existencia* con 3205 (está*/permanece*), aunque estos resultados, precisamente por la naturaleza de la combinatoria del verbo *estar* y por no haber sido depurado el conjunto de concordancias, no se pueden considerar estadísticamente relevantes. La siguiente es la categoría *posición* con 1415 ocurrencias en pone* y 882 con coloca*, la categoría de los verbos de *habla* (habla* 930 y dice* 108) y el *movimiento* con 295 ocurrencias exclusivamente con mueve*. El resto de las categorías, representadas como las anteriores con sus lexemas más generales (posesión con tiene*; cambio con cambia*, acción con hace*, los verbos de sentimiento con siente* y los de cognición con sabe*) distaban mucho de acercarse al porcentaje de las anteriores. De los resultados anteriores se desprende por tanto que con análisis más profundos, en los que estén presentes todos los verbos de movimiento, éste podría convertirse, después de *mirar*, en una de las áreas más destacadas en la gramática del GAD.

Para iniciar la configuración de la gramática del texto audiodescrito, deberíamos extraer de las concordancias las proyecciones sintácticas y

las relaciones semánticas. Del mismo modo, si deseamos analizar la forma en que los audiodescriptores representan el movimiento en sus textos, hemos de analizar el comportamiento de todos los verbos que se incluyen en esta categoría, identificar sus rasgos gramaticales más recurrentes, crear una estructura que resuma su comportamiento sintáctico y asociar los resultados a los diferentes roles semánticos.

En lo que sigue y solamente a modo de ejemplo, vamos a ofrecer los resultados de los análisis de las concordancias del verbo más representativo de esta categoría *mover(se)* en la tercera persona del singular del presente de indicativo. Al introducir *mueve** en la herramienta *concord* obtuvimos un total de 295 ocurrencias. Una vez analizadas, el mapa conceptual del movimiento en los GAD quedó como sigue en la figura 2.

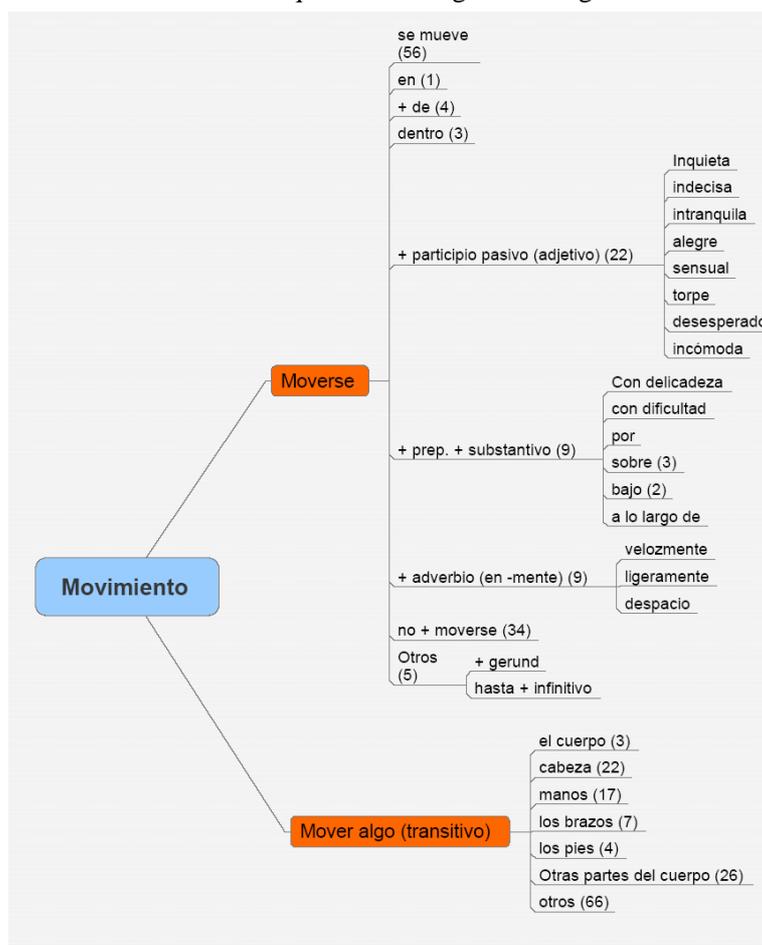


Figura 2: El movimiento en los GAD

La figura no es más que una muestra que ilustra la metodología, sin embargo, podemos extraer algunas conclusiones que nos permitan seguir trabajando en la gramática. La primera curiosidad que sobresale del esquema es el hecho de que el movimiento transitivo (mover algo) esté mayoritariamente relacionado con el cuerpo humano. La audiodescripción del 90% de los sujetos animados, ya sean personajes humanos o animales, muestra el movimiento de una parte del cuerpo. Como se aprecia en el esquema, la cabeza y las manos suelen ser el foco de atención principal. En la clasificación *otras partes del cuerpo* se incluyen ojos, pelos y boca.

A la vista de estos resultados mínimos y con la cautela que requiere este primer análisis de los guiones en español, nos atrevemos a afirmar que, en efecto, como explica Salway (2007), los GAD describen mayoritariamente también en español los eventos de foco de atención, codificado en *mira** y en todo lo relacionado con el movimiento de la cabeza y los ojos. Del mismo modo, además del canal de la comunicación acústica que las personas ciegas reciben igual que el resto de espectadores, es a través del movimiento corporal donde se codifican, en las diferentes culturas, la interacción entre los personajes. Moviendo los brazos, los pies y las manos sobre todo nos comunicamos sin palabras, una vez que ya hemos detectado estados de ánimo y sentimientos en el otro, precisamente también gracias a la actitud corporal.

En la primera parte de la figura se pueden observar los patrones de *move*. Nuevamente destaca el hecho de que cuando el sujeto es animado, la estructura más representada es la de *move + participio pasivo o adjetivo: se mueve inquieta, indecisa, intranquila*. Una vez más, los sentimientos, la actitud emocional y el estado de ánimo en general que denota el movimiento corporal es muy relevante en las audiodescripciones y al parecer algo en lo que tenemos que seguir investigando.

5. Conclusiones

No deja de ser sorprendente que esta primera aproximación al movimiento en los GAD en español revele pautas semánticas y discursivas muy parecidas a las que se realizan en lengua inglesa.

El principio básico para la creación de una gramática local de base semántica (Harris 1991) es ofrecer los patrones semánticos y sintácticos más representativos de una categoría o esquema semántico-cognitivo determinado. En los GAD en español está todo por hacer.

Bibliografía

- Ballester Casado, Ana (en prensa). "La audiodescripción: apuntes sobre el estado de la cuestión y las perspectivas de investigación". *Tradterm*. Sao Paulo.
- Bednarek, Monika (2007). "Local Grammar and Register Variation: Explorations in Broadsheet and Tabloid Newspaper Discourse", *ELR Journal*, 1 (1), <http://ejournals.org.uk/ELR/article/2007/1>. (last visited on 10 December 2007).

- Bourne, Julian; Jiménez Hurtado, Catalina (2007). "From the visual to the verbal in two languages: a contrastive analysis of the audiodescription of *The Hours* in English and Spanish". Díaz Cintas, Jorge; Orero, Pilar; Rемаel, Aline (eds.): *Media for All. Subtitling for the Deaf, Audio Description, and Sign Language*. Amsterdam: Rodopi., 134-142.
- Cifuentes Honrubia, José Luis (1999). *Sintaxis y semántica del movimiento*. Aspectos de gramática cognitiva. Alicante: Instituto de Cultura Juan Gil-Albert.
- Faber, Pamela (2002). "Investigar en terminología". Faber, Pamela; Catalina Jiménez (eds). *Investigar en Terminología*. Granada: Comares, 3-23.
- Faber, Pamela, Ricardo Mairal (1997). "Towards a typology of predicate schemata in a funcional-Lexematic Model". Wotjak, Gerd (ed.). *Toward a functional lexicology*. Frankfurt: Peter Lang, 11-36.
- Faber, Pamela, Ricardo Mairal (1999). *Constructing a Lexicon of English Verbs*. Berlin: Mouton de Gruyter.
- Harris, Zellig (1991). *A Theory of Language and Information: A Mathematical Approach*. Clarendon Press.
- Jiménez Hurtado, Catalina (2007). "De imágenes a palabras: la audiodescripción como una nueva modalidad de traducción y de representación del conocimiento", Wotjak, Gerd *Quo vadis Translatologie?* Leipzig: Frank und Timme, 143-160.
- Jiménez, Catalina / Seibel, Claudia (2004). "El lenguaje controlado para una definición variable: 'no ambiguity through homonyms, no redundancy through synonyms'". Faber, Pamela / Jiménez, Catalina / Wotjak, Gert (Hrsg.). *Léxico especializado y comunicación interlingüística*. Granada: Granada Lingvistica, 117-130.
- Pérez Hernández, Chantal (2002). "Terminología basada en corpus: principios teóricos y metodológicos". Faber, Pamela; Catalina Jiménez (eds). *Investigar en Terminología*. Granada: Comares, 127-166.
- Salway, Andrew (2007). Seminario ofrecido en la Universidad de Granada. Marzo 2007.
- Talmy, Leonard (1996). "The windowing of attention in language". Masayoshi Shibatani; Sandra Thompson (eds). *Grammatical Constructions: Their Form and Meaning*. Oxford: Oxford University Press, 235-287.
- Traboulsi Hayssam, David Cheng, Khurshid Ahmad (2004). "Text Corpora, Local Grammars and Prediction" Proc. of the 4th International Language Resources and Evaluation Conference, vol. 3, Lisbon, 2004, 749-752.

¹ TRACCE: Evaluación y gestión de los recursos de accesibilidad para discapacitados sensoriales a través de la traducción audiovisual: La audiodescripción para ciegos. Protocolo para formar a formadores. Código: SEJ2006-01829/PSIC.

² Para la realización de este estudio, se optó por la herramienta informática de análisis *Wordsmith Tools*, creado por Mike Scott de la Universidad de Liverpool para *Oxford University Press*, por tratarse de un programa que se utiliza cada vez más en estudios con *corpora* especializados. Además, ofrece una serie de herramientas muy útiles y de relativamente fácil manejo. La extracción de líneas de concordancia (*Key Word in Context - KWIC*) con la aplicación llamada *Concord* es de gran utilidad ya que permite seleccionar un término que posteriormente aparece en el centro de un co-texto que suele tener la longitud de una línea con lo que cada línea nos ofrece información contextual acerca de las colocaciones semánticas y sintácticas con las que aparece dicho término

³ Coincidiendo con los resultados de las investigaciones llevadas a cabo por Salway para el verbo *to look* y su variante *looks*.