

Creación de bases de conocimiento lingüístico para el aprendizaje de idiomas

A. Vaquero y F. Sáenz

{vaquero,fernan}@sip.ucm.es

Resumen— En primer lugar centramos nuestra atención en el aprendizaje de idiomas mediante el uso y construcción de diccionarios multilingües para facilitar la asimilación de conceptos lingüísticos fundamentales: léxico, acepciones, categorías semánticas, ciertas relaciones léxico-semánticas y taxonomía. Se subrayan diversos conceptos lingüísticos como objetivos pedagógicos, así como el modelo de aprendizaje seleccionado, con un enfoque constructivista, que constituyen los requisitos para el desarrollo de las herramientas de uso y administración. Hemos desarrollado herramientas informáticas para encauzar las actividades que permiten construir y usar diccionarios bilingües. También hemos desarrollado una herramienta de administración que soporta la consistencia semántica del léxico, lo que permite detectar diversas omisiones y posibles inconsistencias. Se describen las interfaces de uso de dichas herramientas. Como continuación de esta línea de trabajo, para seguir avanzando en el dominio de las lenguas con objetivos más ambiciosos, se proponen nuevas herramientas para crear y manejar bases de conocimiento lingüístico, parte de las cuales ya están construidas y se mencionan aquí. El uso de estas herramientas suponen indudables ventajas para alcanzar el dominio de la lengua y aprender idiomas.

Index Terms—Computer Networks and Education, Design of educational software, E-learning, Software architectures for E-learning, Standardization.

I. INTRODUCCIÓN

El uso de las computadoras en la escuela debe de ser controlado por el profesor [4] y, por tanto, éste tiene que prepararse para estar dispuesto a asumir este control. Además de la cultura informática que cada profesor debe de tener en general, el "profesor usuario" debería poseer una formación específica en Informática Educativa [Erickson, 1994]. Particularmente importante es el aspecto de la Lengua, no sólo por la comunicación directa entre profesor y alumno, sino porque no se puede enseñar lo que no se conoce aunque se cuente para ello con buenos medios.

Pero existe un escaso soporte para el profesorado y una ausencia de planes para su adecuada formación. Además de estas dificultades, cuando se encomienda a los profesores

aplicar las computadoras en el aula, se detectan diversos problemas, además de la escasez de hardware, que impiden el uso de las computadoras a lo largo del curriculum [8], entre los que cabe destacar: calidad del software, accesibilidad al software, selección del software apropiado para la enseñanza en las propias clases, escasa fiabilidad de las revisiones y evaluaciones del software educativo, dificultad de integrar el software existente en la enseñanza, tiempo escaso del profesor antes de poder utilizar la informática en el aula, etc. Más aún, a medida que cambia el método de aprendizaje aparecen nuevos problemas. Los cambios se centran en la propia dinámica del aprendizaje, o sea, en el proceso de comprensión del estudiante. Este nuevo punto de vista en la naturaleza del aprendizaje está basado en las teorías cognitivas [1] [16] e induce nuevas formas de enseñanza [22]. La adquisición del conocimiento es un proceso de construcción activo del estudiante y es ese tipo de actividad la que hay que promover.

Con esta nueva visión del problema, los materiales necesarios para la enseñanza han de construirse pensando más en el estudiante que en el curriculum [10]. Consecuentemente se deberían construir nuevos materiales y herramientas informáticas como ayuda al estudiante. Unas para liberarlo de sobrecargas pesadas y actividades rutinarias, como calculadoras, manipuladores simbólicos, herramientas gráficas y estadísticas, procesadores de texto, hojas de calculo, bases de datos, etc. Otras para realizar actividades que motiven a aprender determinadas materias. Las herramientas utilizadas aquí son de este tipo.

Cuando se detecta un dominio débil de conocimientos y habilidades en la población estudiantil, se debe intentar subsanar esta deficiencia poniendo a su disposición entornos informáticos apropiados. Uno de los dominios especialmente débil resulta ser la Lengua. La falta de dominio de la lengua nativa resulta cada vez más preocupante. La clave de la falta de comprensión es el léxico. Existe una evidencia experimental de la dependencia existente entre la comprensión de la lectura y el vocabulario [9] [21].

Para mejorar el nivel de dominio de la lengua, cada alumno debería manejar herramientas específicas con funcionalidad para construir, consultar y modificar piezas del lenguaje. Los recursos convencionales tales como diccionarios, glosarios, tesauros, enciclopedias, etc. , no son suficientes, como muestra el hecho generalizado de su escasa utilización. Los diccionarios electrónicos [24] deben motivar al estudiante más que los impresos en papel, siguiendo el modelo de aprendizaje basado en el constructivismo [2]. Además de la rapidez y

Manuscript received June 14, 2002.

A. R. Vaquero Sánchez, Dept. Sistemas Informáticos y Programación, Universidad Complutense de Madrid, España (e-mail: avaquero@sip.ucm.es).

F. Sáenz Pérez, Dept. Sistemas Informáticos y Programación, Universidad Complutense de Madrid, España (tel: +34913945186, e-mail: fernan@sip.ucm.es).

simplicidad de la consulta de términos, la computadora permite al estudiante desarrollar una serie de nuevas tareas con fines claramente pedagógicos. El objetivo pedagógico global que se persigue es el significado de las palabras [18]. La definición constituye la tarea para llegar a aprender la dependencia del significado de una palabra de otras palabras, así como ocurre con la clasificación de las palabras en categorías semánticas. Objetivos específicos son algunas relaciones como polisemia y sinonimia y sus implicaciones en la clasificación. También puede ser interesante relacionar palabras de diferentes lenguas. Todos estos objetivos se pueden alcanzar por un procedimiento constructivo y en colaboración entre estudiantes y profesores en el aula. Con herramientas reactivas adecuadas provistas de interfaces amigables podría llevarse a cabo esta misión [25].

En este artículo se proponen herramientas informáticas para construir y consultar diccionarios electrónicos bilingües y multilingües, y en general bases de conocimiento lingüístico, que sirvan como medio de aprendizaje para ir mejorando continuamente el dominio de idiomas. Además, también se presentan los modelos conceptuales que dan soporte a la información que integra los conceptos lingüísticos en los que estamos interesados para el aprendizaje de idiomas. Se trata de rellenar un hueco existente en el conjunto de herramientas constructivas para mejorar el dominio de la lengua, en cada materia y en un amplio rango de niveles educativos. Para situar nuestras herramientas en su justo lugar hay que distinguir entre aprendizaje constructivista en ambientes controlados por el usuario (entornos completamente libres) y los de navegación con multimedia [15]. Nuestras herramientas pertenecen al primero de estos modelos de aprendizaje. El segundo resulta más apropiado para aprender otras partes distintas del léxico [7]. Sin embargo, ambos tipos de herramientas resultan complementarios y no separables de forma absoluta [20] [6]. Por lo que conocemos no existen descritas en la literatura herramientas similares.

El artículo se organiza de la forma siguiente. En la sección 2 se resaltan los conceptos lingüísticos que consideramos objetivos del aprendizaje. Las herramientas informáticas de aprendizaje se presentan en la sección 3. En la sección 4 se presenta el modelo conceptual del diccionario bilingüe. En la sección 5 se presenta el modelo conceptual del diccionario multilingüe. Finalmente, la sección 6 resume algunas conclusiones y proporciona líneas para el desarrollo de futuros trabajos.

II. CONCEPTOS LINGÜÍSTICOS INCORPORADOS COMO OBJETIVOS PEDAGÓGICOS

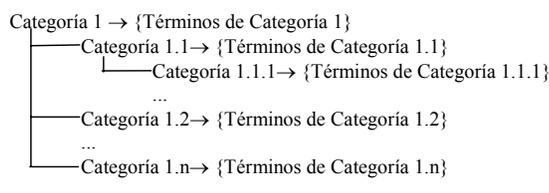


Fig. 1. Taxonomía.

En esta sección se presentan los conceptos lingüísticos tenidos en cuenta en nuestra propuesta que pueden ser explotados ventajosamente en el aprendizaje de lenguas. Las siguientes secciones muestran la forma en que hemos incorporado estos conceptos a nuestras herramientas para ser tenidos en cuenta durante las actividades que podrán desarrollar los estudiantes.

A. Orden, clasificación y ontología

Normalmente los diccionarios monolingües presentan un orden alfabético que obedece a una clasificación simple de términos en conjuntos unitarios según su forma lexicográfica. Otras clasificaciones posibles y menos ingenuas son de tipo derivativo (según la raíz de la palabra), gramaticales y semánticas. Las clasificaciones derivativas [11] no son corrientes y las clasificaciones gramaticales no son adecuadas para los diccionarios. Finalmente, las clasificaciones semánticas agrupan términos mediante categorías semánticas (por ejemplo, los diccionarios de sinónimos y antónimos o los diccionarios ideológicos [3]). Las categorías semánticas no sólo permiten la clasificación de acepciones sino la más significativa de taxonomía de acepciones. Las bases de datos léxicas convencionales, tales como WordNet [13], presentan clasificación de términos tales como sinonimia (agrupados en los denominados *synsets*). Las ontologías van más allá del papel de la taxonomía de acepciones [14]. Nuestras herramientas soportan gran parte de las propiedades y relaciones que implica este importante concepto, según se explicará a lo largo de este artículo.

Las categorías semánticas carecen de utilidad en las búsquedas de términos, puesto que las acepciones se corresponden en general con un conjunto de términos (sinónimos). Sin embargo, juegan un papel importante en el aprendizaje mediante el uso y construcción de diccionarios porque cada acepción de un término dado (polisemia y/o homonimia) está precisamente identificada por su categoría semántica (por brevedad, de ahora en adelante nos referiremos a ellas únicamente como categorías) en lugar de un número secuencial poco orientativo. Por tanto, se tiene una taxonomía o clasificación para las acepciones, pero no una ordenación término a término, puesto que las acepciones son ideas abstractas que no pueden ser expresadas en general sólo por una palabra. Es reconocido comúnmente que el mejor orden para las consultas es el lexicográfico (una clasificación derivativa es un contraejemplo de esto, pero incluso mantiene un orden lexicográfico repitiendo entradas y añadiendo enlaces). La figura 1 ilustra el orden de una taxonomía como una jerarquía, presentando una taxonomía de categorías junto al conjunto de términos pertenecientes a cada categoría. Desde

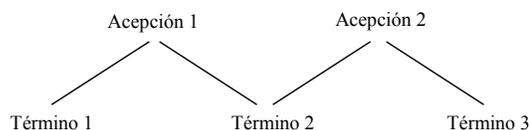


Fig. 2. Polisemia y sinonimia.

este punto de vista existe un orden completo lexicográfico (siempre que las categorías estén identificadas por términos o frases). La jerarquía es una estructura natural para la clasificación de acepciones. Cada nodo de la jerarquía corresponde a una categoría. En principio, cada categoría de la jerarquía se puede usar sin importar su nivel en ella. Nótese que cada categoría en la jerarquía contiene al menos el término que da nombre a la categoría, de forma que ninguna categoría está vacía. Por otra parte, la creación de nuevas categorías como intersección de otras predefinidas debería ser evitada con el fin de conseguir compactación. Esta limitación desaparece cuando pasamos a una ontología como conocimiento conceptual independiente de idiomas.

Desde un punto de vista educativo, el fin no es desarrollar un diccionario general (de hecho, es un trabajo de enorme envergadura que todavía hoy se lleva a cabo por investigadores lingüísticos), sino diccionarios especializados restringidos a un dominio semántico o lingüístico, con el fin de hacer más sencilla la categorización de acepciones así como la definición de categorías. Existen numerosas ventajas en clasificar acepciones como taxonomías. Primero, la taxonomía de acepciones es una característica útil para un diccionario electrónico, porque las acepciones expresan una semántica adicional que proporciona al lector una mayor información (mayor que la de los números secuenciales indicada anteriormente). Segundo, el sistema puede ganar también una nueva dimensión, porque es posible generar automáticamente diccionarios especializados bajo diferentes categorías (un diccionario deportivo puede tratar diccionarios de fútbol, baloncesto o tenis). Tercero, ayuda a desarrollar diccionarios equilibrados incorporando un número suficiente de términos de todos los campos. Una vez clasificados los términos, es fácil comprobar el número de términos que se encuentran en una categoría determinada. Cuarto, ayuda a distribuir el trabajo entre varios autores asignando categorías a los mismos. Un equipo de autores puede desarrollar un diccionario completo especializado, dividiendo el trabajo por categorías, con el fin de promover el trabajo colaborativo entre los estudiantes. Esto último significa finalmente que las categorías deben ser definidas, lo que implica una ventaja educativa añadida, puesto que los estudiantes tendrán que organizar ideas de un modo formal, proceso soportado por la herramienta de autor, tratada en la siguiente sección.

B. Polisemia y sinonimia

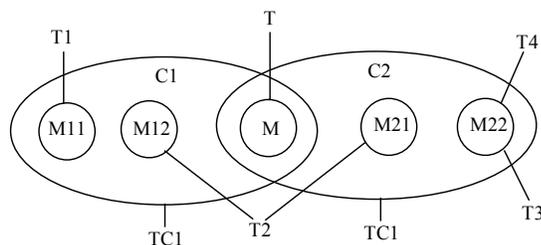


Fig. 3. Relaciones entre categorías, acepciones y términos. Definición por extensión.

En todos los idiomas existe el conocido problema de la denominación ::[Katzenberg, 1993], que presenta dos aspectos: uno es la polisemia (bajo un punto de vista sincrónico, es decir, incorpora tanto la propia polisemia como la homonimia), mediante la cual un término puede tener distintos significados, y el otro es la sinonimia, por la cual a un solo significado se le pueden asignar distintos términos, según se puede observar en la figura 2. En ella, Término 1 y Término 2 son sinónimos y comparten un mismo significado (Acepción 1), y lo mismo ocurre con Término 2 y Término 3 con respecto a otro significado diferente (Acepción 2). Además, Término2 es polisémico puesto que tiene dos posibles significados.

C. Relaciones

1) Relaciones básicas

En esta sección destacamos algunas observaciones acerca de las relaciones entre categorías, acepciones y términos. Por una parte, un término dado puede pertenecer a varias categorías con significados distintos. Por otra parte, ese término puede pertenecer a distintas categorías con el mismo significado. La figura 3 muestra dos categorías (C1 Y C2) que contienen respectivamente los significados {M11, M12, M} y {M, M21, M22}. Cada significado tiene asociado uno o más términos. El término T2 está asociado a los significados M12 y M21, que pertenecen respectivamente a las categorías C1 y C2. También aparece el término T asignado al significado M, que pertenece a ambas categorías C1 y C2. La polisemia, como se puede ver, está presente en T2, y la sinonimia está también presente en T3 y T4. T1 no tiene ni polisemia ni sinonimia, TC1 y TC2 son los términos usados para denotar las categorías C1 y C2 respectivamente.

En la figura 3, el conjunto de significados {M11, M12, M} en C1 es la definición por extensión de la categoría C1. Se observa también que una categoría tiene el significado descrito por su definición. La figura no expresa este hecho. Para expresar que las acepciones están relacionadas con las categorías, se transforma el esquema de la figura 3 en el ilustrado en la figura 4. Ahora, C1 es el significado de la categoría C1 y TC1 es el término asociado a dicho significado, lo que se aplica igualmente con respecto a C2 y TC2. Luego en cada categoría se dispone de un significado más. Este significado proporciona la definición por intensión de la categoría.

Para una lengua dada, se tiene un conjunto de términos que

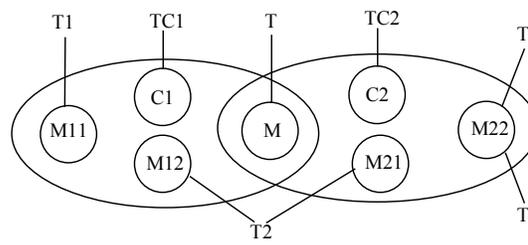


Fig. 4. Relaciones entre categorías, acepciones y términos. Definición por intensión.

mantienen las relaciones en las categorías y acepciones mostradas en la figura 4. Si pensamos en varias lenguas, a cada una se le aplica lo mismo. Por tanto, las relaciones entre los términos de diferentes lenguas provienen de considerar al mismo tiempo los esquemas involucrados en cada lengua.

2) Otras relaciones léxicas y semánticas

En cualquier idioma el conocimiento contenido en el discurso es de dos clases: conceptual y lingüístico. Los términos y las frases se refieren a conceptos, pero tienen en cada lengua unas características estructurales y morfológicas propias. El dominio de una lengua pasa por distinguir entre ambas clases de conocimiento. En realidad el dominio de las lenguas pasa por sucesivas fases hasta llegar a aprender a distinguir entre los conceptos y la forma lingüística de expresarlos en un idioma. Hay que aprender los conceptos y las relaciones entre los mismos, el léxico y las propiedades lingüísticas de los términos, la composicionalidad definida por la estructura sintáctica y las vinculaciones entre términos y conceptos. Todos estos objetivos pedagógicos han sido tenidos en cuenta para ser potencialmente alcanzados realizando las tareas soportadas por nuestras herramientas.

Aunque no es exactamente igual ontología que conocimiento conceptual del discurso, no hay medio computacional más adecuado para representarlo. Todas las relaciones (meronimia, holonimia, hipernimia, hiponimia,...) representadas en las bases de datos léxicas más completas, como WordNet, están presentes en las bases de datos basadas en ontologías, como el sistema MikroKosmos [12] basado en la ontología Ontos; pero en éstas están presentes de una forma estructurada por niveles. En la ontología se representan los conceptos y las relaciones entre los mismos, mientras que en

cada léxico están los términos de un idioma y sus propiedades lingüísticas, así como su asociación con los conceptos de la ontología. La vinculación entre ontología y léxicos es la clave para coordinar adecuadamente todas las relaciones léxicas y semánticas.

III. HERRAMIENTAS DE APRENDIZAJE

En este apartado se explica hasta dónde han evolucionado las herramientas desarrolladas originalmente [23] con el resultado de la experiencia obtenida en su aplicación.

A. Consulta del diccionario

La herramienta de usuario es una interfaz de consulta que permite recuperar fácilmente la información de términos tanto en inglés como en español así como las relaciones existentes entre ambos, que están contenidas en la denominada base de datos terminológica. Esta interfaz permite al usuario navegar por las categorías semánticas, permitiendo asimismo la búsqueda de información relevante sobre cualquier término (definición, términos relacionados, traducción, sinónimos, ...)

La figura 5 muestra la configuración inicial de la interfaz de esta herramienta. El cuadro de texto Términos contiene un término seleccionado de la lista de términos almacenados en la base de datos. El usuario puede escribir en este cuadro de texto el término del que desea información y, según escribe se va seleccionando el término más parecido en la lista de abajo. La selección del término también se puede realizar con el ratón sobre esta lista. Más abajo aparece el cuadro Categorías, que muestra la lista de categorías a las que pertenece el término seleccionado en el cuadro Términos, cada una correspondiente a una acepción diferente del término. El

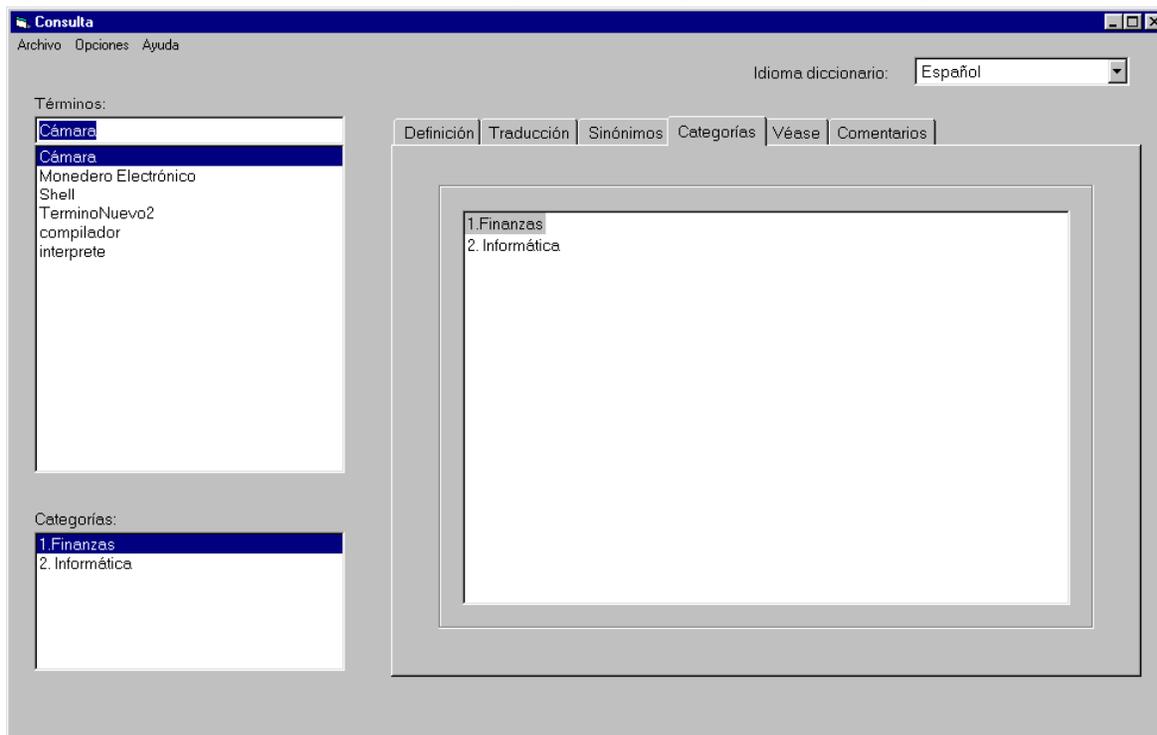


Fig. 5. Vista inicial de la interfaz de la aplicación de consulta .

cuadro desplegable Idioma diccionario permite seleccionar el idioma origen entre los dos disponibles para las traducciones. La zona derecha de la interfaz contiene varias fichas que se explican a continuación.

La ficha Categorías presenta el árbol de categorías definido para el diccionario sin expandir sus nodos. Pulsando sobre los símbolos + se accede a las categorías hijo en esta estructura jerárquica. Si se selecciona cualquiera de estas categorías, la lista de términos se actualiza con los términos que pertenezcan a la categoría seleccionada. Además se selecciona automáticamente el primero de ellos en el cuadro de texto Términos y se muestran sus categorías asociadas en el cuadro Categorías.

La ficha Definición muestra la definición textual de la acepción del término seleccionado en el cuadro Términos correspondiente a la categoría seleccionada en el cuadro Categorías. Cualquier término de la definición textual que se encuentre en la base de datos aparece resaltado como

hipervínculo de forma que se puede pulsar para acceder a la información de ese término.

Las fichas Traducción, Sinónimos, Véase y Comentarios funcionan de forma similar a la anterior pero con referencia respectivamente a la traducción del término, sus sinónimos, los términos relacionados a los que se remite para su consulta bajo cada acepción y los comentarios. En todos estos casos también están disponibles los hipervínculos.

Finalmente, la barra de menú de la aplicación consta de tres elementos: Archivo, Opciones y Ayuda. Con el primero se puede seleccionar la base de datos del diccionario y salir de la aplicación. Con el segundo se puede seleccionar el idioma con que se muestra la interfaz y el tipo de actualización de información (bajo demanda, con una doble pulsación de ratón o pulsando la tecla Intro, o bien automáticamente; la primera opción es útil cuando las consultas en la base de datos son muy costosas en grandes diccionarios, permitiendo una navegación más fluida). Finalmente, con Ayuda se accede a la

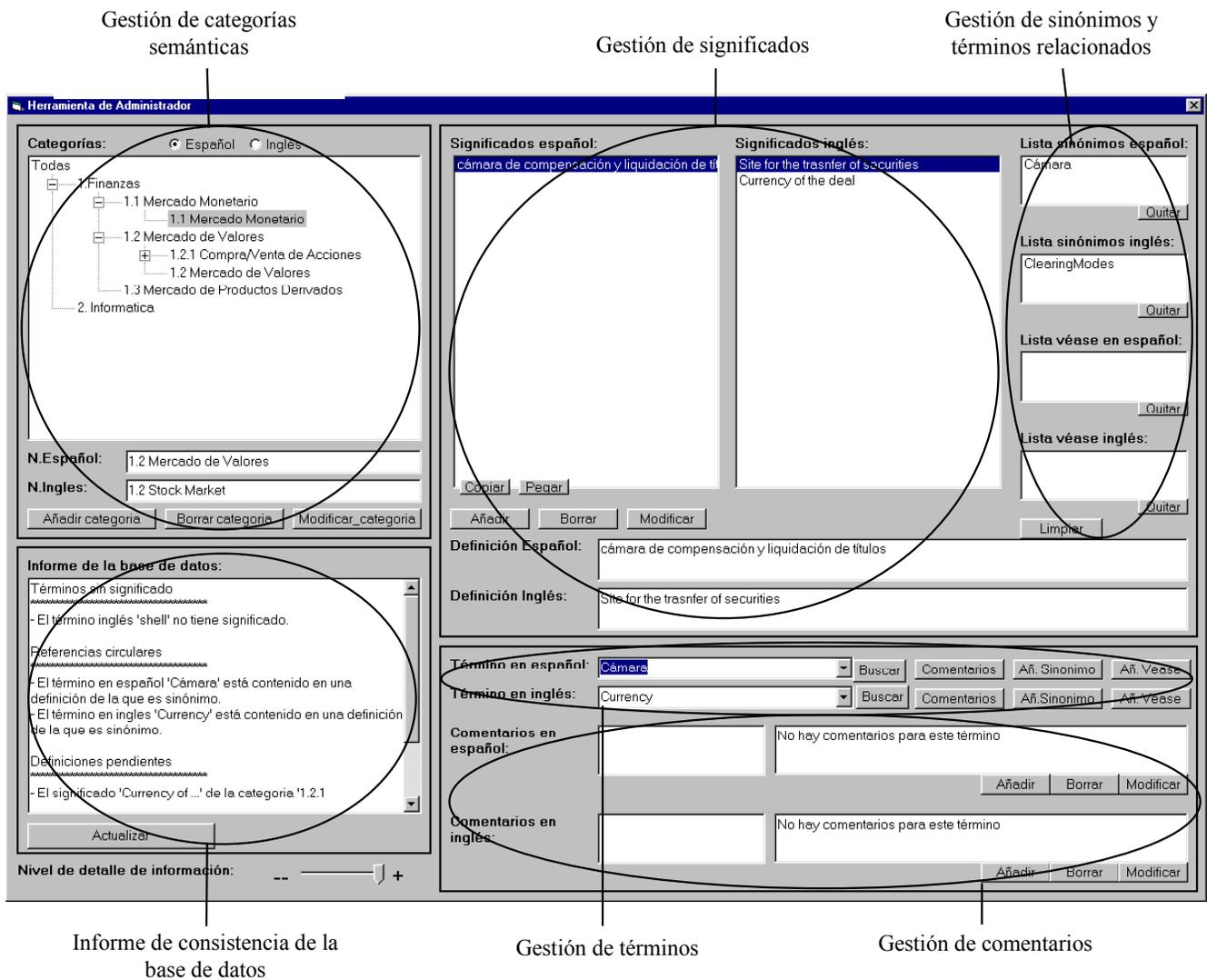


Fig. 6. Aplicación para la creación de diccionarios.

ayuda de la aplicación.

B. Creación de diccionarios

La herramienta de creación de diccionarios permite al diseñador añadir nuevos términos a la base de datos terminológica junto con la información relevante, tal como la definición, categorías semánticas, acepciones, conjuntos de sinónimos y términos relacionados. Se ha desarrollado una interfaz de usuario para los idiomas inglés y español (seleccionable con los botones de radio ubicados en la parte superior izquierda) que se muestra en la figura 6. Contiene varias áreas de gestión (señaladas separadamente en la figura) que se explican a continuación.

En el área de gestión de categorías semánticas se gestionan las operaciones referidas al mantenimiento de categorías semánticas. Tiene varios controles: una vista jerárquica de las categorías semánticas (que puede comprimirse o expandirse pulsando en los símbolos +), campos de texto para los nombres de la categoría semántica (en ambos idiomas), y los botones Añadir categoría, Borrar categoría y Modificar categoría. El punto de inserción cuando se añade una nueva categoría semántica es la categoría semántica resaltada.

El área para la gestión de significados consiste en dos listas de acepciones en los dos lenguajes y los botones Añadir, Borrar y Modificar para la inserción, el borrado y la modificación de las acepciones, así como los botones de edición (Cortar y Pegar). Ambas listas muestran las acepciones mediante sus definiciones textuales y cada línea seleccionada resalta tanto la definición inglesa como la española. Además, la información correspondiente a esa acepción (conjunto de sinónimos, ...) se actualiza, poniéndose de manifiesto que existe una correspondencia uno a uno entre las acepciones de todos los lenguajes. Obsérvese que las acepciones se muestran bajo la mejor representación posible en un lenguaje dado, es decir, su definición textual.

Con el área de gestión de términos se permiten añadir, borrar y modificar términos, tanto para asignarlos al conjunto de sinónimos de una acepción, como para incluirlos en las listas de términos relacionados. También se pueden asociar comentarios a los términos desde esta área.

El área de gestión de sinónimos y términos relacionados

contiene las listas de sinónimos (conjunto de sinónimos, *synset*) y las listas de términos relacionados en los dos lenguajes del diccionario, que corresponden a la acepción resaltada en el área de gestión de significados.

El área de informe de la consistencia de la base de datos presenta un informe (cuadro de texto Informe de la base de datos) sobre la consistencia de la base de datos. Este informe tiene cuatro apartados. En el primero, Términos sin significado, se informa de los términos que no se han asignado aún a un significado. El apartado referencias circulares informa de las definiciones textuales circulares (existen diversos diccionarios comerciales con este error). El apartado Definiciones pendientes lista las acepciones para las que no se ha aportado una definición textual. Finalmente, Definiciones sin sinónimos, lista las definiciones a las que no se ha asignado ningún término. Este informe resulta muy importante en la construcción de diccionarios, ya que los diccionarios no se pueden construir de forma consistente en un paso sino que se han de realizar de forma constructiva en base a los términos y las relaciones entre ellos (polisemia y sinonimia).

IV. MODELO CONCEPTUAL DEL DICCIONARIO BILINGÜE

El desarrollo de las herramientas introducidas anteriormente está basado en el ciclo de diseño clásico de bases de datos. Su primera y fundamental etapa es el diseño del modelo conceptual, para el que hemos usado el modelo entidad-relación. En la figura 7 se presenta el modelo conceptual que hemos desarrollado para estas herramientas. En esta figura, y de acuerdo con convenios de [17][19], las entidades se representan por rectángulos, los atributos de las entidades por elipses, y las relaciones entre entidades por líneas que pueden ser direccionales o no. Si a la entidad B le llega una línea con flecha que viene de A, significa que la relación tiene una cardinalidad de uno (A) a varios (B). Las flechas dobles entre las entidades A y B denotan que la cardinalidad es de varios a varios. Las líneas no direccionales denotan que la cardinalidad es de uno a uno y también se usan para unir atributos con sus entidades o relaciones. Los nombres de las relaciones se usan para etiquetar cada una de las líneas (no se muestran en la figura).

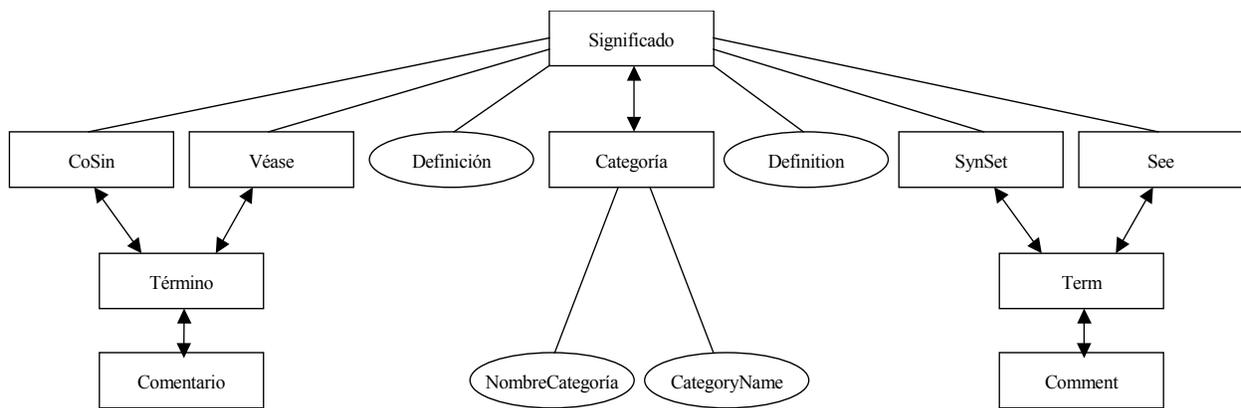


Fig. 7. Modelo entidad-relación para un diccionario inglés-español.

En este modelo conceptual se muestra la entidad Significado, de la que dependen fundamentalmente el resto de entidades y que es independiente del idioma. La entidad CoSin denota el conjunto de sinónimos en español (CoSin, conjunto de sinónimos) y el tipo de relación entre ambas entidades es uno a uno. La entidad Término representa todos los términos de la base de datos terminológica y su relación con CoSin es varios a varios porque un conjunto de sinónimos puede contener varios términos (sinónimos) y un mismo término puede estar en diferentes conjuntos de sinónimos (polisemia). La entidad Véase denota el conjunto de términos españoles relacionados bajo un determinado significado y su relación con Significado es uno a uno. La entidad Comentario representa todos los posibles comentarios que se pueden asociar a los términos. La relación entre Véase y Término es varios a varios porque un término puede referirse a otros y el mismo término puede aparecer referenciado por varios términos. La relación entre Categoría y Significado es varios a varios porque hay varios significados correspondientes a una categoría y un mismo significado puede estar en varias categorías (esta situación se pretende que se reduzca al máximo para mantener la clasificación disjunta). Esta relación implica que nuestra clasificación no es léxica (no hay una relación directa entre Categoría y Término) sino semántica (se relacionan significados con categorías, es decir, se catalogan los significados). La relación entre Comentario y Término es varios a varios porque un mismo término puede tener varios comentarios y el mismo comentario se puede referir a varios términos. La entidad Categoría tiene dos atributos: NombreCategoría y CategoryName, que corresponden con los nombres textuales de la categoría en cada idioma. La entidad Meaning tiene dos atributos: Definición y Definition, que

corresponden a las definiciones textuales en cada idioma. El resto de entidades, SynSet, See, Term y Comment, son homólogas a las entidades respectivas CoSin, Véase, Término y Comentario.

V. MODELO CONCEPTUAL DEL DICCIONARIO MULTILINGÜE

El siguiente paso que pretendemos abordar es la creación de herramientas de consulta y creación de diccionarios multilingües. Como primera etapa es necesario construir una base sólida de representación conceptual como se ha indicado en el anterior apartado y, para ello, hemos creado un modelo entidad-relación para modelar un diccionario multilingüe dinámico. Por dinámico nos referimos a la capacidad del usuario de las herramientas para modificar el número de idiomas sin que este paso altere la estructura del modelo entidad-relación. La figura 8 muestra el modelo entidad-relación que proponemos.

Este modelo es necesariamente más elaborado que el anterior para representar su característica dinámica. Aparece una nueva entidad, Idioma, que representa todos los idiomas representables en el diccionario multilingüe. Además se indican con rombos las relaciones que aparecen. Las entidades CoSin y Véase se han convertido en relaciones (representación más adecuada para ellas, ahora que también ha sido necesario hacer explícitas otras relaciones). La estructura entidad-relación entre significados, términos y comentarios es similar a la del modelo conceptual anterior. Sin embargo, las entidades Término y Comentario se refieren a los conjuntos de todos los términos y comentarios, sin importar el idioma. La clave que indica el idioma al que pertenece un término es la relación IdiomaTérmino. Los comentarios, en cambio, se asocian directamente a los términos mediante la relación

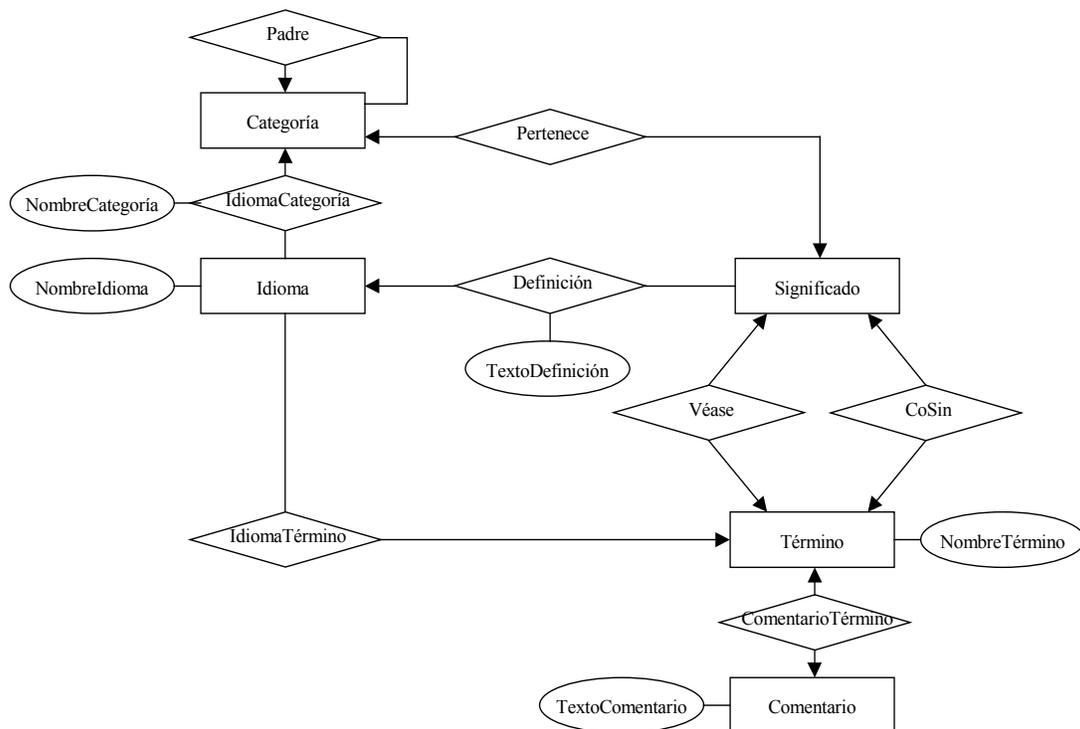


Fig. 8. Modelo entidad-relación para un diccionario multilingüe.

ComentarioTérmino. En este caso no es necesario hacer explícito el idioma del comentario puesto que un comentario se asocia a un término cuyo idioma se conoce por la relación IdiomaTérmino. El comentario en si se representa por el atributo TextoComentario de la entidad Comentario. Por otra parte, la entidad Categoría se asocia a un idioma con la relación IdiomaCategoría. En este caso sí es necesario hacer explícito el idioma de la categoría dado que el significado es independiente del idioma. Nótese que ahora el atributo NombreCategoría va asociado a la relación IdiomaCategoría, es decir, el concepto Categoría es independiente del idioma, no así el texto que la describe. Además aparece la relación Padre que no se había hecho explícita en el modelo anterior. Con ello se consigue especificar explícitamente la relación jerárquica entre las categorías (aunque en esta representación conceptual no se restringen implícitamente las jerarquías a árboles; es necesario incluir esta información en el esquema conceptual como restricción añadida). La relación Pertenece hace explícita la relación que ya se había identificado en el anterior modelo. Las definiciones no son ahora atributos de la entidad Significado, sino que se modelan con el atributo TextoDefinición de la relación Definición entre Significado e Idioma (es decir, un significado tiene una definición en un determinado idioma).

VI. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se han presentado herramientas para ser usadas en el aprendizaje de lenguas con un enfoque constructivista. El lenguaje es una materia técnica compleja. Hay que aprender desde unos conceptos elementales hasta otros muy complejos. En este camino las herramientas han de ceñirse a los objetivos pedagógicos marcados, determinados por los conceptos lingüísticos a aprehender. Aquí se explica el estado actual de esas herramientas, el modelo conceptual en el que están basadas y las interfaces para su uso. Con ello se tiene una visión de los ámbitos en que pueden beneficiarse de su aplicación y de los modos de utilización de las mismas.

Hay que continuar en esta prometedora línea de investigación y desarrollo por varios motivos. Uno es la plasmación de una metodología completa y bien fundada para integrar recursos lingüísticos, que hoy se hacen en múltiples sitios heterogéneamente. Otro es incluir conceptos lingüísticos más avanzados.

7. REFERENCIAS BIBLIOGRÁFICAS

- [1] D.P. Ausubel, "Educational Psychology: A cognitive View", New York, Holt, Reinhart and Winston, 1968.
- [2] A. Cabrera, "Informática Educativa: La revolución constructivista". Informática y Automática, Vol. 28, n. 1, marzo 1995.
- [3] Casares. Ideological Spanish Dictionary.
- [4] L. Cuban, "Teachers and Machines: The classroom use of technology since 1920". Teachers College Press, Columbia University, New York, 1987.
- [5] F.J. Erickson, and J.A. Yonk, "Computer Essentials in Education. The teaching tools". McGraw-Hill Book Co., 1994.
- [6] A. Fernández-Valmayor, C. López-Alonso, S. Arlette, and B. Fernández-Manjón, "The Design of a Flexible Hypermedia System: Integrating an Interactive Learning Paradigm for Foreign Language Text Comprehension", International Working Conference on Building Electronic Educational Environments, IFIP, Irvine, California, pp. 51-65, 1999.
- [7] S.R. Goldman, "Reading, Writing, and Learning in Hypermedia Environments", Cognitive Aspects of Electronic Text Processing (Ed. H. Van Oostendorp and S. Mul), Norwood, NJ. Ablex Publications, 1996.
- [8] B. Hodgson, "The roles and the needs of the teacher". Proceedings of the Working Conference "Integrating Information Technology into Education", IFIP, Barcelona, pp. 25-34, October 17-21, 1994.
- [9] D.D. Johnson, and P.D. Pearson, "Teaching Reading Vocabulary", Ed. Holt, Reinhard & Winston, New York, 1978.
- [10] J. Karat, "Evolving the scope of user-centered design" Communications of the ACM, Vol. 40, N. 7, July, 1997.
- [11] María Moliner. Derivative Spanish Dictionary.
- [12] <http://crl.nmsu.edu/Research/Projects/mikro/index.html>
- [13] G. Miller, "WordNet: A Lexical Data Base for English", Communications of the ACM, Vol. 38, 11, 1995.
- [14] S. Nirenburg, V. Raskin, and B. Onyshkevich, "Apologiae Ontologiae", Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation, Center for Computational Linguistics, Catholic University, Leuven, Belgium, pp. 106-114, 1995.
- [15] K. Norman, "Navigating the educational space with HyperCourseware". Hypermedia, Vol. 6, enero 1994.
- [16] M.I. Posner (Ed.), "Foundations of Cognitive Science". Cambridge, Mass., MIT press, 1989.
- [17] R.S. Pressman, "Software Engineering. A Practitioner's Approach", McGraw-Hill, 1997.
- [18] M. R. Quillian, "Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities". Brachmen, R. J. y Levesque, H. J., Eds., Reading in Knowledge Representation. Morgan Kaufman, 1967.
- [19] A. Silberschatz, H.F. Korth, S. Sudarshan, "Data Base System Concepts", WCB/McGraw-Hill, 2001.
- [20] P. Teusch, T. Chanier, Y. Chevalier, D. Perrin, F. Mangenot, J.P. Narcy, and J.de Saint Ferjeux, "Environnements interactives pour l'apprentissage en langue étrangère". Hypermedias et Apprentissage, 3 (Ed. E. Brouillard), 1996, pp. 247-256.
- [21] R.L. Thorndike, "Reading Comprehension Education in Fifteen Countries", Ed. Wiley, 1973.
- [22] K. Tobin, and D. Tippings, "Constructivism as a referent for teaching and learning". K. Tobin (Ed.), "The practice of Constructivism in Science Education". AAAS Press. Washington, DC, 1993, pp. 3-21.
- [23] A. Vaquero, F. Sáenz y A. Barco, "Improving the Language Mastery through Responsive Environments" en "Computers and Education. Towards an Interconnected Society", Kluwer.
- [24] Y.A. Wilks, D.C. Fass, C.M. Guo, J.E. McDonald, T. Plate, and B.M. Slator, "Providing machine tractable dictionary tools". Machine Translation, 5, 1990, pp. 99-151.
- [25] D. Zeltzen and R. K. Addison, "Responsive virtual environments", Communications of the ACM, Vol. 40. N. 8, August, 1997.