

Zirano Cúpula, una aplicación de las variables conceptuales de la semántica conceptual multilenguaje.

‘Zirano Cúpula’, an application of the conceptual variables in the conceptual multilanguage semantics.

Miguel Santolaria García

e-Lexis Ziento Volando
Pirineos 1, B. 22150 Arguis. Huesca
internacional@zirano.com

Carmen Santolaria Gracia

e-Lexis Ziento Volando
Mosen Femenia, 3, 6. 46006
Valencia

Manuela Giral Zanuy

e-Lexis Ziento Volando
Jota Aragonesa, 2, 4º, D. 22004
Huesca

Ana Santolaria Gracia

e-Lexis Ziento Volando
Zarandia 22, 2º. 22002 Huesca
anasantolaria@zirano.com

Marco Calvo Lafarga

e-Lexis Ziento Volando
Mosen Femenia, 3, 6. 46006 Valencia
evaluadordenodos@zirano.com

José Ángel García Landa

Universidad de Zaragoza

garciala@unizar.es

Resumen: Los Diccionarios Analógico Conceptuales que contiene la Plataforma Intercultural Multilenguaje Zirano (Zirano Cúpula) se basan en la Semántica Conceptual Multilenguaje, donde el átomo inicial es **el concepto**, con ello desaparecen algunas limitaciones de los sistemas que parten del inglés, como WordNet o Universal Networking Language.

La *doble representación* en el signo lingüístico permite que cada signo actúe como una *variable conceptual*: representamos con él todos los significantes que puede llegar a tener (a veces tantos como lenguas hay en el mundo). Esto permite búsquedas inclusivas en corpora desambiguados. Así, por ejemplo, en nuestro Diccionario Analógico Conceptual Castellano, ‘noche’ no sólo está asociado con ‘nocturno’ o ‘anochecer’ sino también con ‘aquejarre’, ‘cena’, ‘estrigiforme’, ‘asmático’, ‘vivaquear’, etc., conceptos/palabras que contienen la *variable conceptual* ‘noche’.

Palabras clave: Semántica. Lexicografía. Lexicón. Ontología. Procesamiento del lenguaje natural. Tradumática. Desambiguador. Lingüística de corpus. Recuperación de información multilingüe. Lenguas con recursos limitados.

Abstract: The Analogical Conceptual Dictionaries present in the ‘Plataforma Intercultural Multilenguaje Zirano (Zirano Cúpula)’ are based on Multilanguage Conceptual Semantics, where the initial atom is **the concept**. Therefore several limitations of the systems based on English (like WordNet or Universal Networking Language) disappear.

The *double representation* in the linguistic sign makes possible that a sign acts as a *conceptual variable*. That sign represents all the signifiers that it can have (sometimes, as many as languages in the world). This makes possible inclusive searches in disambiguated corpora. So, in our Spanish Analogical Conceptual Dictionary ‘night’ is associated not only with ‘nocturne’ or ‘night fall’ but also with ‘aquejarre’, ‘dinner’, ‘estrigiforme’, ‘quiropter’, ‘asthmatic’, ‘to bivouac’, etc., concepts/words which contain the *conceptual variable* ‘night’.

Keywords: Semantics. Lexicography. Lexicon. Ontology. Natural language processing. Translation. Disambiguator. Corpus linguistics. Multilingual information retrieval. Limited resources languages.

Antecedentes.

John Wilkins (1668) analizó la posibilidad de crear un idioma universal y planteó sus bases teóricas.

Partiendo de las mismas premisas y, sorprendentemente para el autor (Santolaria, 2004) utilizando casi las mismas estrategias, el ‘idioma universal’ que hemos desarrollado cumple las mismas funciones con las que teorizó John Wilkins. La diferencia principal es que el polígrafo inglés planteó un lenguaje para seres humanos y, en nuestro caso, se trata de un lenguaje para máquinas al servicio de los seres humanos.

En Semántica Conceptual Multilenguaje el átomo básico es el **concepto**. La *doble representación en el signo lingüístico* permite que cada signo actúe como una *variable conceptual* multilenguaje: representamos con él todos los significantes que puede llegar a tener (en teoría tantos como lenguas hay en el mundo). Esto permite búsquedas inclusivas en corpora desambiguados.

Los Diccionarios Analógico Conceptuales Zirano.

Una de las aplicaciones desarrolladas aplicando *variables conceptuales* de semántica conceptual son los diccionarios analógico conceptuales Zirano presentes en la Plataforma Intercultural Multilenguaje Zirano, situada en la página web www.zirano.com.

Parten de que el léxico de cada idioma es reflejo de la cultura en el que se desarrolla y a la que sirve. En un diccionario alfabético monolingüe de un idioma determinado están presentes muchos de los rasgos culturales del grupo humano que habla ese idioma.

Un Diccionario Analógico Conceptual, partiendo de un diccionario alfabético monolingüe de determinado idioma, por medio de la aplicación de variables conceptuales, permite extraer sistemáticamente de cada definición (fase

de creación) y recupera automáticamente (fase de consulta) sus semas, los rasgos culturales presentes en ella, los ‘ecos’ o sugerencias que las vivencias del investigador hagan surgir en su mente más los hallazgos estéticos de algunos de los literatos y creadores que han escrito en dicho idioma.

Esta tabla contiene parte de lo recuperado a partir de la variable conceptual ‘noche’.

Cada una de las palabras o expresiones que aparecen en esta tabla contienen la *variable conceptual* ‘noche’.

SUSTANT.	balancín	cena
acampada	banco	día
ahogo	banco	disc-jockey
alameda	bandolina	dondiego
alcaraván	bandurria	equinoccio
ambigú	basura	erizo
ánimas	brisa	escarcha
anochece	brugo	estantigua
aquejarre	búho	estrella
aroma	cabaré	estrigiforme
asma	cabareter@	fanal
autillo	cabezada	faro
bache	cabecal	garduña
bacín	cabrería	hiena
balada	cama	hipopótamo
balancín	claridad	imaginaria

ADJETIVOS

- altas(horas)
- antemeridian@
- asmatic@
- asmatic@
- bullanguer@
- cabareter@
- clandestin@
- estrigiformes
- nictálope
- noctámbul@
- nocturn@
- nocturn@
- otr@
- quirópter@
- tisanuro
- trasnochad@

VERBOS

- acamarse
- anochece
- anochece
- arropar
- atardecer
- correr
- crujir
- dormir
- escarchar
- ocultar
- oscurecer
- pernoctar
- rondar
- trasnochar
- velar
- vivaquear

ADVERB.

- anochece
- anteanoche
- sereno
- tarde
- temprana

OTROS

- noche
- des-noche
- Tantas
- Arma
- clar@

El acceso a un diccionario analógico conceptual en lengua desconocida, desde una lengua conocida. Traducción (semi)automática.

La Plataforma Intercultural Multilinguaje Zirano (Zirano Cúpula) contiene varios diccionarios analógico conceptuales que se construyen y se consultan por medio de una página web. La Plataforma permite acceder a los rasgos de las lenguas / culturas que la constituyen de una forma cómoda, directa, sistemática, contrastiva (varias culturas al mismo tiempo).

Los resultados de la consulta se leen simultáneamente en el idioma consultado y en el de consulta.

Esto es parte de lo recuperado del Diccionario Analógico Conceptual Aragonés haciendo consulta en Castellano de la variable conceptual ‘herramienta’.

alicates	alicates
arambre	alambre
artulario	instrumento
atornillador	desatornillador
batán	batán
billabarquín	taladro manual
brazo	brazo (anatomía)
burra	callosidad
chuego	juego (desajuste)
clau (mecanica)	llave (mecánica)
clau (punta)	clavo (carpint.)
crepaza	grieta en la piel
dalla	guadaña
enclusa	yunque
estenazas	tenazas
estral	hacha
estrapaluzio	ruido grande
falz	hoz
ferramienta	herramienta
man	mano
martiello	martillo
metal	metal
sarrucho	serrucho
tacha	tachuela

En la aplicación, al colocar el cursor del ratón encima de una palabra, surgen las definiciones de las parejas de palabras (falz /hoz) en aragonés y en castellano.

Traductor multilinguaje (semi)automático retroactivo.

Para estar completa, la aplicación necesita el desarrollo de un traductor multilinguaje (semi)automático retroactivo porque

a.- en la definición original puede haber detalles culturales interesantes que no son accesibles al consultante al no conocer la lengua que consulta (al no poder leer la definición).

b.- existen conceptos (definiciones) en la lengua / cultura investigada que no existen en la lengua de la persona consultante, no son pertinentes culturalmente, por lo tanto no es posible hallar su definición en la lengua de la persona consultante.

El traductor que necesitamos ha de tener las siguientes características:

Multilinguaje = Traducir simultáneamente todos los idiomas presentes en la Plataforma.

(Semi)automático =Las palabras y expresiones de las definiciones que ha de traducir van a estar desambiguadas, ya que este es un proceso que se realiza durante la construcción de los Diccionarios Analógico Conceptuales.

Retroactivo = Actualmente hay dos idiomas plenamente activos en la Plataforma, en el futuro van a incorporarse más. Los textos deben quedar preparados para ser traducidos a los idiomas actuales y a los futuros, sin ninguna modificación posterior.

Bibliografía

Wilkins, J. 1668. An Essay Towards a Real Character and a Philosophical Language.
 Borges, J. 1952. El idioma analítico de John Wilkins. Ensayo, en Otras inquisiciones. Emecé.
 Santolaria García, Miguel. 2004. Diccionarios Analógico Conceptuales Zirano. Aplicación informática, en www.zirano.com
 WordNet en <http://wordnet.princeton.edu/>
 Universal Networking Language en <http://opera.dia.fi.upm.es/spanish/unlprogram.htm>