



**INSTITUTO POLITÉCNICO NACIONAL  
CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN**

No. 7    Serie: AZUL    Fecha: Marzo 98

**Hallando los Temas Principales  
en un Artículo en Español**

Dr. Adolfo Guzmán Arenas<sup>1</sup>

**RESUMEN**

Muchas operaciones sobre colecciones sistemáticas de datos pueden realizarse fácilmente por la computadora cuando los datos son numéricos:

- ¿A qué se refiere esta información? ¿De qué temas se trata?
- Resumir, compendiar. Obtener un resumen de las ventas de la tienda El Sol en mayo
- Comparar. Comparar las ventas de mayo de las tiendas El Sol y La Hormiga
- Hallar diferencias y parecidos. ¿En qué se parecen las ventas de las tiendas A y B?
- Hacer promedios. Hallar las ventas en el Sudeste de México, o en el otoño de 1997.
- Establecer tendencias. Extrapolar

En cambio, cuando los datos están en documentos en español, organizados en secciones, párrafos y oraciones, no es posible que la computadora realice las operaciones anteriores. Puesto que gran parte del conocimiento humano está en textos en lenguaje natural, conviene descubrir métodos para realizar tales operaciones. Para ello, la máquina debe **entender** el texto, comprenderlo.

Este artículo expone cómo analizar un documento escrito en lenguaje natural, y reconocer los **temas** principales de los que habla.

**Palabras clave:** Análisis inteligente, árbol de conceptos, clasitex, análisis de textos, lenguaje natural, análisis semántico.

<sup>1</sup> Director del Centro de Investigación en Computación - I.P.N.

## ADVERTENCIA

*“Este reporte contiene información desarrollada por el Centro de Investigación en Computación del Instituto Politécnico Nacional a partir de datos y documentos con derechos de propiedad y por lo tanto su uso queda restringido a las aplicaciones que explícitamente se convenga.*

*La aplicación no convenida exime al Centro de su responsabilidad técnica y da lugar a las consecuencias legales que para tal efecto se determinen.*

*Información adicional sobre este reporte podrá obtenerse recurriendo a la Unidad de Publicaciones y Reportes Técnicos del centro de Investigación en Computación del I.P.N. Av. Juan de Dios Bátiz s/n, teléfono 729-60-00 ext. 56500, 56608 y 56610”.*

## Hallando los temas principales en un artículo en español

Adolfo Guzmán Arenas<sup>1</sup>  
Centro de Investigación en Computación del I. P. N.  
aguzman@pollux.cenac.ipn.mx

### 1. INTRODUCCIÓN Y OBJETIVOS

En el Centro de Investigación en Computación del Instituto Politécnico Nacional, el *Laboratorio de Lenguaje Natural y Procesamiento de Texto* está investigando el procesamiento inteligente de texto, necesario para llevar a cabo las labores ejemplificadas en el resumen. Una de tales labores, el hallazgo de los temas principales que un artículo en español aborda, ha sido resuelta y es el tema de esta ponencia.

#### 1.1 La importancia de analizar español escrito

El análisis inteligente de texto permitirá que la computadora *entienda* los documentos escritos en lenguaje natural, por ejemplo, que haga resúmenes, encuentre tendencias, compare documentos (con respecto a un tema dado) y conteste preguntas no triviales.

##### Leído este texto

Las ranas viven en el agua

Benito Juárez está enterrado en el Panteón de San Fernando

Otra ventaja: se introduciría un nuevo método de darle órdenes a la computadora: en español, en vez de hacerlo en Fortran, C o Java.

##### Contestar esta pregunta

¿Se mojan las ranas?

¿En dónde está enterrado el dedo pulgar izquierdo de Benito Juárez?

Otra ventaja: constituiría para la máquina un nuevo método para aprender: solo tendría que leer libros, los mismos libros que las personas utilizan. Para enseñarle Física, ya no sería necesario fabricar procedimientos o métodos (y los objetos correspondientes) o estructuras de datos que guardarán conceptos y fórmulas de esa ciencia.

#### 1.2 Diferencia entre palabra y concepto

Una palabra o vocablo puede representar varios conceptos o acepciones. Los conceptos, en cambio (por definición) no son ambiguos:

---

<sup>1</sup> Este trabajo fue realizado por el autor en SoftwarePro International, quien posee todos los derechos sobre CLASITEX.

### ESTRELLA (palabra)

- estrella-astronómica (concepto) cuerpo sideral
- estrella-artista (concepto) persona famosa o destacada
- estrella-insignia (concepto) adorno
- estrella (concepto) rompe, destroza, quiebra (verbo)

Decía el filósofo Wittgenstein : Si las palabras son ambiguas y representan varios conceptos, descartemos las palabras y solo usemos los conceptos. Por supuesto, esto no puede hacerse en la práctica. Lo que sí podemos hacer es dilucidar o reemplazar en un texto, las palabras por los conceptos que denota. Para esto hay que *desambiguar*: escoger para cada palabra su acepción (concepto) correcta, de varias posibles. Para desambiguar hay que utilizar información del contexto (del resto del artículo, por ejemplo, o de los temas de los que el artículo o la oración habla). Esto se denomina *análisis semántico*, ya que atañe al significado de cada palabra.

### 1.3 Los conceptos forman árboles

Es interesante notar que los conceptos están organizados en familias o jerarquías, como puede constatarse al hojear un Thesaurus. La relación más usada es la de “subconjunto”, y los conceptos “hijos” denotan casos particulares o más específicos que sus padres. Tomemos un concepto que denota un objeto (sustantivo, sujeto) como ejemplo:

PRENDA DE VESTIR (concepto)

ZAPATO

huarache  
sandalia  
mocasín  
bota  
alpargata  
pantufla

CAMISA

camisa de manga corta  
camisa de manga larga

PANTALÓN

En este ejemplo, el concepto mocasín es un subconjunto del concepto zapato, lo que significa: todos los mocasines son zapatos, pero no todos los zapatos son mocasines. Hemos mostrado una pequeña parte del árbol de conceptos.

Ahora veamos la parte del árbol (de conceptos) donde se encuentran los procesos (acciones, verbos). En particular veremos cambios de posición:

## CAMBIAR

### MOVER

*CAMBIAR DE POSICIÓN (de coordenadas sobre la tierra)*

Flotar

Nadar

Andar

*caminar*

*correr*

*gatear*

*arrastrarse*

Volar

*GIRAR (cambiar de orientación)*

*VIBRAR*

### CAMBIAR DE TAMAÑO

*INFLAR-EXPANDIR*

*CONTRAER-EMPEQUEÑECER*

### CAMBIAR DE SALUD

*SANAR-RESTABLECER*

*ENFERMAR-AGRAVAR*

*MORIR*

### CAMBIAR DE ESTADO ECONÓMICO

*ENRIQUECER*

*EMPOBRECER*

## 1.4 El número de conceptos es finito

En particular, el número de conceptos del conocimiento común (ver próxima sección) es finito. Resulta, pues, posible formar los árboles de conceptos, o de una parte de ellos, y en esto se basa el algoritmo y programa, llamado CLASITEX, que halla los temas en un artículo.

## 2. EL CONOCIMIENTO COMÚN

Es el conocimiento que tiene un especialista cuando se le quitan los conocimientos de su especialidad. Por consiguiente, es el conocimiento que tiene un niño de, digamos, 9 años, o alguien que no ha aprendido o adquirido (los conocimientos de) una especialidad o profesión. [Ref. 1].

### 2.1 El árbol del conocimiento común

Como ya se vio en §1.3, es posible organizar los conocimientos, incluidos los conocimientos comunes, en un árbol. Este árbol se calcula debe tener entre uno y diez millones de conceptos. El Proyecto CYC [Ref. 1] se abocó a la enorme tarea de construir este árbol, a mano, es decir, introducir concepto tras concepto. Tardó diez años, pero no logró su objetivo. No obstante, creo

que la meta es válida, y que este fracaso no debe detener a otros investigadores, el árbol tarde o temprano se creará.

Este árbol contiene los conceptos (objetos, transformaciones, relaciones, etc.) que las personas (tal como un niño de 9 años) han adquirido a través de su experiencia sobre la faz de la tierra y su relación con otros seres humanos. No contiene conceptos especializados, de Química o Reología, digamos. *Estos conceptos en el árbol no los conoce de antemano la computadora.* Mas son necesarios para entender un texto en lenguaje natural. Es sensato, pues, tratar de introducirlos a la máquina.

El árbol contiene, además de “subconjunto”, otras relaciones (como “parte de”, “formado por”, y otras) entre un concepto y otro. Inclusive, las mismas relaciones forman árbol, y, como tal, son parte (excepto las relaciones especializadas) del mismo árbol del conocimiento común. Como ejemplo daremos la relación PARIENTE DE.

PARIENTE DE

*PADRE DE*

*HIJO DE*

Primer hijo de

Hijo preferido de

*HERMANO DE*

### 3. HALLANDO LOS TEMAS DE UN ARTÍCULO, INTENTO 1.

Como primer intento, podemos contar las palabras en el artículo. Aquéllas que más aparecen sin duda constituyen o denotan temas importantes. Por ejemplo, si aparece 17 veces la palabra “revolución”, el artículo habla sobre revolución.

#### 3.1 Problema: muchas palabras se refieren al mismo tema

Por ejemplo, un artículo solo menciona *revolución* dos veces, pero menciona a Zapata, a Francisco Villa, a Flores Magón, ...

Por ejemplo, portero, defensa izquierdo, Atlante, Necaxa, Luis García, todos hablan (sugieren el tema) fútbol soccer.

Esto sugiere que fabriquemos el árbol de conceptos, y que reemplacemos las palabras en el artículo por los conceptos respectivos. Hecho esto, hallemos los conceptos más populares: ellos son los temas principales buscados. Al contar conceptos, debemos contar no solo los conceptos “inmediatos” que una palabra denota (por ejemplo, la palabra “huarache” denota el concepto “huarache”), sino los conceptos hacia la raíz del árbol, que en el ejemplo serían zapato, prenda de vestir, etc. El método busca marcar o contar sobre el árbol de conceptos. Al finalizar el análisis del artículo, ciertas partes del árbol quedarán más marcadas (habrá contadores con mayores cantidades), y esas partes del árbol son los temas del artículo.

Problemas en el enfoque propuesto:

- Hay que hacer el árbol, manualmente. (Esto lo resolvemos haciendo en efecto una introducción manual).

- Mapear una palabra al concepto que denota no es fácil, pues es necesario desambiguar (Cf. §1.2). Si no lo hiciéramos, una palabra incrementaría el contador de cada uno de los conceptos que denota. (Esto se resuelve en Clasitex omitiendo la desambiguación, cosa que no introduce un incremento apreciable en la inexactitud, como veremos abajo).

### 3.2 Resumen del procedimiento propuesto

Halle los conceptos que más aparecen – esos son los temas. Para esto, observemos lo que sigue:

1. Contemos o miremos los conceptos, no las palabras.  
Refinamiento: Cuento los conceptos sobre el árbol de conceptos.  
Cuento los conceptos y sus generalizaciones (conceptos inmediatamente más generales)
  - Cuando halle “huarache”, cuente también zapato, prenda de vestir, ... En la versión actual de Clasitex, no se cuentan las generalizaciones, y no obstante los resultados obtenidos son aceptables. ♣
  - Algunas partes del árbol se ennegrecerán más que otras. Las regiones negras del árbol señalan los **temas** del artículo.
 Mientras más extenso es el artículo, mejor resulta este método.
2. Cada palabra “vota” por (aumenta el contador de) los conceptos que denota.  
cuando hallamos **canal** aumentamos 1 al concepto canal-de-agua, y al concepto canal-de-comunicaciones.
3. *Habrán votos equivocados*. Empero, a la larga, los votos correctos prevalecen.
  - C) También es posible desambiguar en un segundo paso: las palabras que votaron más de una vez retiran todos excepto uno de sus votos (conservan el voto “correcto”), una vez encontrados los temas principales al final del paso 1. ♣<sup>2</sup> Estos votos que se retiran no deberían (salvo en contadas ocasiones) cambiar los temas principales encontrados.
  - D) En la práctica, hemos visto que este segundo paso no es necesario: la sola acumulación de los temas, tanto correctos como equivocados, hace que los correctos destaquen. Es como si, en vez de eliminar el ruido de una señal, acumulamos muchas de ellas, de modo que el ruido resulta reducido en raíz cuadrada de  $n$ , el número de señales agregadas.

### 3.3 Hay conceptos denotados por una secuencia de más de una palabra en español

Por ejemplo,

COCINA (una palabra) – denota el concepto “cocina”, habitación donde se confeccionan alimentos.

NUEVO MÉXICO (dos palabras). Denota el concepto “Nuevo-México”, un estado de EE. UU.

Benito Juárez (dos palabras). Denota al Benemérito de las Américas.

Los Tres Mosqueteros (tres palabras). Denota un libro de Alejandro Dumas.

<sup>2</sup> Con este símbolo marcamos sugerencias y extensiones para trabajos futuros, o mejoras posibles a CLASITEX.



Instituto Politécnico Nacional (tres palabras). Denota un famoso Instituto Tecnológico.

Adolfo Guzmán Arenas (tres palabras). Denota al autor de este artículo.

Universidad Nacional Autónoma de México (cuatro palabras). Denota una universidad mexicana, la UNAM.

Por mi raza hablará el espíritu (cuatro palabras). Denota también a la UNAM.

La técnica al servicio de la patria (seis palabras). Denota al Instituto Politécnico Nacional.

Lo que esto implica es que, al barrer el texto, debemos buscar no solo palabras, sino pares, tríos, cuartetos, ..., de palabras, en nuestro recorrido del texto, averiguando para cada una de estas palabras, pares, tríos, etc., si denotan algún concepto o no.

### 3.4 El árbol usado en Clasitex

Este último ejemplo nos permite emplear en Clasitex un árbol que usa, en vez de la relación "subconjunto" entre sus nodos (ver §1.2), la relación "sugiere el tema de", o "vota por el tema ...". Por ejemplo, "Benito Juárez" sugiere el tema "México", el tema "Oaxaca", y el tema "Historia de México", aunque Benito Juárez es una persona, México es un país, etc.

#### 3.4.1 Palabras sin sentido

Esta simplificación tiene otra consecuencia. Una palabra como "tres" (que representa al concepto "3", un número impar) o como "entonces" no sugiere ningún tema, no vota por ningún tema. A estas palabras las hemos llamado, un tanto inapropiadamente, de "sentido débil" o "sin sentido", dando a entender que no contribuyen a dilucidar el tema del artículo.

## 4. PROCEDIMIENTO USADO EN CLASITEX

Describimos el algoritmo utilizado por Clasitex.

Recorremos con un apuntador el artículo en español (que está en un archivo), de izquierda a derecha.

1. (COMIENZO) Observa la secuencia o cadena de cuatro palabras señaladas por el apuntador. ¿Denotan algún o algunos conceptos?  
Sí : aumenta en 1 el contador de cada uno de esos conceptos. Ve al paso (5).  
No : Ve al paso siguiente (2).
2. Observa la secuencia o cadena de tres palabras señaladas por el apuntador. ¿Denotan algún o algunos conceptos?  
Sí: aumenta en 1 el contador de cada uno de esos conceptos. Ve al paso (5).  
No: Ve al paso siguiente (3).
3. Ídem la secuencia de dos palabras.
4. Observa la palabra señalada por el apuntador. ¿Denota algún o algunos conceptos?  
Sí: aumenta en 1 el contador de cada uno de esos conceptos. Ve al paso (5).  
No: ¿Es una palabra que no denota concepto alguno? Es decir, que sea palabra "sin sentido" según el §3.4.1.



Sí: ignórala. Ve al paso (5) siguiente.

No: Imprímela como “no sé qué significa esta palabra, o si no significa algo” (Hay que agregar estas palabras en alguna ocasión posterior al árbol de Clasi-  
tex, o al conjunto de palabras “sin sentido”). Ve al paso (5) siguiente.

5. Mueve el apuntador a la derecha de la(s) palabra(s) ya analizada(s).

Y repite la operación o iteración: ve al paso 1 (COMIENZO). Cuando ya se haya analizado todo el texto (ya no podrá moverse el apuntador a la derecha), cuéntense los conceptos y repórtense los más numerosos como los temas del artículo.

En la práctica, se ha visto que el histograma de conceptos contiene una separación bien diferenciada entre los temas numerosos y los temas poco numerosos. Abajo veremos ejemplos.

#### 4.1 Ejemplo de conceptos representados por una palabra

Clasitex utiliza el directorio ARBOL1 para guardar los conceptos representados por una palabra (para ser más exactos, los conceptos “sugeridos” o “votados” por una palabra). Cada concepto es un archivo, y dentro del archivo están las palabras (palabras sencillas, en el caso de ARBOL1; dos palabras, en el caso de ARBOL2; tres palabras, en el caso de ARBOL3, etc.). Un listado parcial del directorio ARBOL1 es el siguiente:

accidente	encuesta	musica
acertijo	enfermedad	narcotrafico
acto-juridico	enloquecer	nayarit
africa	entregar	nino
agencia-noticia	envase-recipiente	noche
agricultura	error	noruega
aguascalientes	escandalo	noticia-internacional
aislado	escoger	oaxaca
alemania	escritura	oficina
amenaza	escuela-colegio	pacto-arreglo
amor-pasion	espana	pais
animal	estado-de-mexico	pan-bizcocho
anunciar-promulgar	excelsior	pan-partido
apicultura	experto	papa-catolico
arma-de-fuego	exponer-exhibir	pasado
asesorar	familia	paz
astronomia	famoso	pedir
atraso	ferrocarril	pelear
auto-parte	fiesta	periodico
autom-marca	filosofia	perro
autor-libro	finanzas	persona-vieja
autoritarismo	flor	platillo-comida
bahamas	francia	poder-judicial
bajacalifornia	fruta	poder-legislativo
banarse	futbol-soccer	poeta
banco-de-mexico	geografia	policia
bebida-alcoholica	gerencia	politecnico
beisbol	gobie-actual	politica
belgica	gobie-municipio	postre

bolsa-valores-mex	gobierno	pps-partido
buceo	guanajuato	prd-partido
budismo	guerra	premio
cambio	guerrero	prenada
campeche	guerrero-edo	preocupacion
campo-rural	habitacion	pri
cantar	hacer-fabricar	pri-partido
carretera	hecho-violento	problema-crisis
casa-parte	herr-carpintero	profesion
casamiento	herr-fierro	protestantismo
cereal	herr-fontanero	psicologia
chiapas	herr-jardinero	pt-partido
chihuahua	hidalgo	queretaro
chile	hipoteca	quintanaroo
ciencia	hombre	radiodifusion
cirugia	honestidad	reino
ciudad	hospital	relig-catolica
ciudad-mexico	hueso-humano	relig-cristiana
cns	iglesia	religion
colima	imss	renunciar
comer	indus-petrolera	reproducir
comercio	industria	restaurante
compositor-musica	inglaterra	ritmo-cadencia
compositor-musica-mex	inmigrante	roma
computadora	institucion	rusia
concept	inteligencia-seguridad	salud
congreso	intervencion-meter	sanluispososi
constitucion	inver-dinero	secr-comercio-mex
consumo	inver-extranjera	secr-hacienda-mex
convalescencia	iraq	secr-tesoro-ccuu
corea	irlanda	secuencia
correr	italia	sexual
corrupcion	jalisco	sinaloa
ctm	jardin	sobrevivir
cuerpo-humano	leguminosa	sociedad
cultura	lenguaje-calo	softwarepro
delincuencia	lentejuela	soga-hilo
democracia	ley	sonreir
deporte	libro-revista	substancia-quimica
desfile	llorar	suburbio
desgarrar-romper	lluvia	sufriendo
deuda-prestamo	lugar-en-ciudad	tabasco
devaluacion	magisterio	tamaulipas
dinero	maletin-veliz	taurino
dirigente	mar	tecnologia-ingenieria
discriminacion	marcha-protesta	telefono
discurso	marxismo	television
discutir	matematicas	tienda
disgusto	medicamento	tlaxcala
divorcio	mexico	trabajo
dormir	michoacan	trafico

Hallando los temas de un artículo en español

dulce-caramelo	miedo	triunfar
durango	militar	unidad-medida
ecologia	miseria-pobreza	utensilio-cocina
economia	misterio	venezuela
ecuu	mito-creencia	veracruz
el-financiero	morculos	vestimenta
eleccion-votar	mueble	viaje
empleado	muerte	viejo-antiguo
empleo-puesto	mujer	yucatan
empresa	mundo	zacatecas

Se podrá observar que (por haber utilizado Unix de Santa Cruz Operation) no se usaron acentos ni la letra ñ. ♣

#### 4.1.1 Ejemplos de archivos representando un concepto

En Clasitex, un nodo del árbol (un concepto) está representado por un archivo, en cuyo interior aparecen palabras sencillas, dobles, triples o cuádruples (no se están manejando frases de más de cuatro palabras, por el momento ♣<sup>3</sup>), que son aquellas palabras (o secuencias) que “votan por” o “se refieren al tema de” el concepto que el archivo denota. Por ejemplo, el archivo **acto-jurídico** (un concepto), contiene las siguientes palabras sencillas (palabras):

apelacion. audiencia. condena. condenada. condenadas. condenado. condenados. culpable. delictuosa. delictuosas. delictuoso. delictuosos. delito. embargo. juicio. procesado. procesados. procesamiento. procesar. proceso. sentencia. sentenciado. sobresellar.

Otro nodo del árbol del conocimiento común, versión Clasitex, es el concepto **platillo-comida**, representado por un archivo que contiene las siguientes palabras sencillas:

albondiga. albondigas. alimentar. alimenticia. alimenticias. alimenticio. alimenticios. alimento. arroz. birria. caldo. ceviche. comer. comida. degustar. enchiladas. flautas. freido. frito. guisado. guiso. menudo. mole. mondongo. paladear. pambazo. pambazos. pancita. picante. picosa. picoso. platillo. platillos. plato. pozole. pulpo. quesadilla. quesadillas. saborear. sabrosa. sabroso. servilleta. servilletas. sope. sopes. taco. tacos. tampiquena. taqueando. taquear. taquiza. tortilla. tortillas. tortillera. tortillero. tostada. tostadas. vianda.

Aquí se podrá notar que la versión actual de Clasitex carece de un desmembrador que descomponga una palabra en su raíz y su sufijo (y prefijos), y por esto mismo tenemos en el archivo palabras como alimentar, alimenticias, alimentos, etc., siendo que con la sola raíz aliment- sería suficiente. ♣

El ejemplo siguiente denota el concepto **platillo-comida**, y se muestra su contenido, que son pares de palabras que “votan” por el concepto. Este archivo se encuentra en el directorio arbol2, que contiene los conceptos (archivos) o nodos que a su vez contienen (son representados o votados por) pares de palabras.

agua fresca. aguas frescas. carne asada. pollo rostizado. pollos rostizados.

El concepto **casamiento**, y sus palabras dobles que votan por él:

media naranja. otra mitad. recién casada. recién casado. recién casados.

El concepto **utensilio-de-cocina**, y sus palabras dobles que votan por él:

bano maria. plato hondo. plato sopero. plato tendido. vaso refresquero.

<sup>3</sup> Con este símbolo marcamos sugerencias y extensiones para trabajos futuros, o mejoras posibles a CLASITEX.

El concepto **platillo-comida**, y sus palabras triples que votan por él:  
 chilpachole de jaiba. sopa de medula. sopa de pescado. sopa de tortilla.

El concepto **casamiento**, y las secuencias de tres palabras que lo sugieren como tema:  
 unir sus destinos. vestida de blanco.

El concepto **computadora**, y las secuencias de tres palabras que votan por él:  
 arquitectura de computadoras. base de datos. bases de datos. convertidor analogico digital. converti-  
 dor digital analogico. convertidores analogico digital. convertidores analogico digitales.  
 convertidores digital analogico. convertidores digital analogicos. editor de texto. editores de texto.  
 estructura de computadoras. hoja de calculo. hojas de calculo. programa de computo.  
 programa de graficacion. programas de computo. sistema de informacion.  
 sistemas de informacion.

El concepto **miedo-pavor**, y las secuencias de 3 palabras que lo sugieren:  
 lleno de pavor. lleno de temor. presa del panico. presas del panico. preso de  
 panico. preso de temor. preso del panico. presos del panico.

## 4.2 Palabras con sentido o significado débil, que no votan por concepto

Estas palabras se ignoran, y no contribuyen al contador de algún concepto. Se encuentran en un archivo que significa "las siguientes palabras serán ignoradas, pues no denotan concepto alguno. Son débiles de significado."

a.	abajo.	abalace.	abalanzaban.	abalanzado.	abalanzados.	abalanzamos.
	abalanzando.	abalanzar.	abalanzaran.	abalanzaras.	abalanzare.	abalanzariamos.
	abalanzarse.	abalanzas.	abalanzaste.	abalanzo.	abaldonado.	abaldonar.
	abalizada.	abalizadas.	abalizado.	abalizados.	abalizar.	abandona.
	abandonaba.	abandonabamos.	abandonaban.	abandonada.	abandonadas.	abandonado.
	abandonados.	abandonamiento.		abandonamos.	abandonando.	abandonar.
	abandonaras.	abandonare.	abandonaria.	abandonas.	abandonaste.	abandono.
	abate.	abatia.	abatiamos.	abatida.	abatidas.	abatido.
	abatidos.	abatimiento.	abatimientos.	abatir.	abatiras.	abatire.
	abatiremos.	abierta.	abiertas.	abierto.	abiertos.	abria.
	abriendo.	abril.	abriles.	abrimos.	abrio.	abrir.
	abriremos.	abriria.	absurda.	absurdamente.	absurdas.	absurdo.
	absurdos.	acababan.	acabada.	acabadas.	acabado.	acabados.
	acabamos.	acaban.	acabando.	acabar.	acabara.	acabaria.
	acabarian.	acabarias.	acabaron.	acabaste.	acabe.	acabemos.
	acabes.	acabo.	acatada.	acatadas.	acatado.	acatados.
	acatamiento.	acatamientos.	acatar.	acatada.	accesible.	acceso.
	acepta.	aceptaban.	aceptacion.	aceptada.	aceptamos.	aceptan.
	aceptar.	aceptaran.	aceptasen.	acepto.	acerca.	acercaba.
	acercada.	acercadas.	acercado.	acercados.	acercamiento.	acercamos.
	acercando.	acercandose.	acercar.	acercara.	acercarse.	acercuemos.

(Siguen muchas palabras más que empiezan con a.... el archivo continúa hasta la z).

#### 4.2.1 Árbol de conceptos sin sentido (con sentido débil), representados por una palabra sencilla

Este árbol contiene archivos, que a su vez contienen palabras sencillas sin sentido. A diferencia del archivo anterior del §4.2, aquí se han hecho árboles, para permitir posteriormente transferir algunos de estos conceptos a "conceptos con sentido", es decir, a árbol.

abajo	colores	feo-horror	llevar-a-cabo	rechazo		
abandonado		constituir		frio-caliente	lograr	renovar
afirmar		dia-mes		futuro	marcar	repeticion
afrenta		diferente		hablar	narrar	total
ahora		dificultad		igualar	nunca-siempre	unir
arbolsi3		eliminar		ilimitado	obedecer	
brillante		enunciar		impedir	obtener	
caminar		felicidad		ir-acudir	precipitar	

Uno de los archivos anteriores, **caminar** (un concepto "débil"), contiene las siguientes palabras sencillas:

andaban.	andado.	andamos.	andando.	andanza.	andanzas.	andar.	anduve.
anduvieron.	anduviste.	caminaba.	caminamos.	caminan.			
caminando.		caminante.	caminar.	caminas.	camine.		
camínemos.	camino.	recorrer.	recorreran.	recorri.	recorria.	recorrido.	reco-
rriendo.	recorrimos.	recorríste.	trasladaba.	trasladaban.	trasladada.	trasladado.	
trasladarse.	traslade.	trasladando.	trasladar.	trasladare.	trasladaren.	trasladarian.	
		traslado.					

El archivo **felicidad** (un concepto con sentido débil) contiene las siguientes palabras sencillas:

agradable.	agradablemente.	agradables.	agrado.	agradados.	alegría.	alegrías.
	felices.	felicidad.	felicitación.	felicitar.	felicito.	feliz.
satisfacción.	satisfacciones.	sentimiento.				

#### 4.2.2 Pares de palabras que no representan concepto alguno

desde abajo.	hasta abajo.	luego entonces.	no obstante.	por abajo.	por consiguiente.	por debajo.
jo.	sin barreras.	sin embargo.	sin límite.	sin límites.	sin obstáculo.	
sin obstáculos.						

## 5. RESULTADOS

### 5.1 OPERAN AL PAPA (Texto analizado por Clasitex)

Operaron al papa; largo, el periodo de rehabilitación.  
*Recibió una prótesis que reemplaza la cabeza del fémur.*

Se resbaló, insisten.

Ciudad del Vaticano, 29 de abril. (AP, AFP, EFE y ANSA). El papa Juan Pablo II fue sometido hoy a una intervención quirúrgica de "artroprótesis" tras haberse fracturado una pierna al caerse en el baño, y los médicos dijeron que tendrá varias semanas de recuperación y luego un largo periodo de rehabilitación.

Hallando los temas de un artículo en español



Una declaración de la Policlínica Gemelli en Roma dijo que la intervención quirúrgica, que según las técnicas modernas se considera de rutina, "estuvo muy bien." El Papa estaba nuevamente consciente y fue llevado de regreso a su habitación del décimo piso.

El Pontífice, que tiene 73 años, recibió una prótesis para reemplazar el hueso en la zona en que el fémur se articula con la cadera. La operación se inició alrededor de las 13:30 horas (11:30 GMT).

El papa se fracturó la parte superior del fémur derecho y se dislocó la cadera al salir de la ducha, mojado y descalzo, y resbalar en el piso de su habitación, el jueves por la noche. Fue atendido primero por su médico privado, Renato Buzzonetti, y durmió algunas horas con analgésicos.

Pero esta mañana fue trasladado al hospital Gemelli, la sexta vez que lo visita como paciente.

Ante la polémica sobre por qué fue ingresado el Papa 11 horas después de caer, el Vaticano señaló que el médico personal de Juan Pablo II estimó que no era necesario trasladarlo de manera urgente al hospital romano.

El vocero del Vaticano, Joaquín Navarro, dijo que la caída no fue provocada por pérdida del conocimiento ni por ninguna enfermedad.

Agrego que a Juan Pablo II se le ha colocado una prótesis de cierta aleación metálica que se ha introducido en el resto del fémur y cuya implantación ha sido perfecta gracias a la robustez del hueso propio de un hombre habituado a practicar deportes.

Asimismo, se le ha aplicado inmovilización interna mediante una pieza metálica, que ha hecho innecesaria la colocación de una escayola externa, precisó Navarro.

Antes de la intervención y cuando ya acusaba el dolor de la fractura, el Papa exclamó, según dijo Navarro: "Tal vez faltaba esto para el Año de la Familia".

Hace dos días, durante la audiencia pública de los miércoles, el Papa habló precisamente del valor redentor del sufrimiento humano.

El Papa dijo a los médicos que lo recibieron en el hospital: "Tienen que admirar mi fidelidad a la Católica", es decir, la Universidad, de la cual depende la Policlínica Gemelli.

El director del hospital, Emilio Tresalti, describió al Papa como "un hombre muy fuerte" cuyo estado general es "excelente".

Tresalti agregó: les digo solamente que mi madre, que tiene la misma edad de Juan Pablo II, fue operada de la misma cuestión, por el mismo equipo, el año pasado, y camina normalmente.

El boletín médico entregado a la prensa dijo que la intervención quirúrgica a la que fue sometido durante dos horas el Papa Juan Pablo II fue normal, habiéndose mantenido "los parámetros bioquímicos y funcionales dentro de los límites normales".

Después de la anestesia, el Papa "volvió en sí de modo tranquilo y normal. Sus condiciones son óptimas", agregó el boletín, que indicó además que Juan Pablo II permanecerá internado "dos o tres semanas", salvo complicaciones.

Durante la intervención perdió sangre pero esta misma sangre le fue introducida otra vez en el sistema circulatorio, no siendo necesaria ninguna transfusión externa, precisó el portavoz vaticano.

A los pocos meses del atentado de 1981, ocasión en la que le fueron practicadas varias transfusiones, el Papa padeció una enfermedad provocada por un virus probablemente transmitido por la sangre externa recibida.

El profesor Gianfranco Fineschi, jefe del equipo que operó al Pontífice, aseguró que el Papa se recuperará para hacer una vida normal, aunque no podrá volver a esquiar "si podrá realizar paseos por la montaña y nadar, que es la mejor terapia", manifestó.

La prótesis "funcionará" aunque "la cadera del Santo Padre no será ya como Dios la ha hecho, sino como la ha hecho un bioingeniero", comentó Fineschi.

La prótesis que se ha colocado al Papa, indicó, es de larguísima duración y muy costosa.

En ese momento, el Pontífice se encontraba acompañado del secretario del Estado Vaticano, cardenal Ángel Sonado; el sustituto de la Secretaría, Giambattista Re; su secretario privado, el sacerdote polaco Stanisław Dziwisz, los doctores Fineschi y Buzzonetti, y dos religiosas que otras veces lo han atendido.

Los médicos dijeron que, luego de las dos o tres semanas en cama, el Papa tendrá una rehabilitación de algunos meses, tal vez con ayuda de un bastón.

Pero luego podrá dejar el bastón, agregó un cirujano.



El Director del hospital dijo que había conversado con el Papa y no mostro signos de tension. Estaba solo preocupado por haber debido renunciar a su programado viaje a Sicilia durante este fin de semana.

El papa celebro misa -sentado-, esta mañana, en su apartamento privado, antes de abandonar el Vaticano para ser intervenido en la Policlinica Gemelli.

Navarro dijo que el Papa cancelaria los viajes y reuniones previstos para las proximas semanas.

El Papa se estaba preparando para un viaje a Sicilia, donde pronunciaria una alocucion contra la mafia y tenia una nutrida agenda para el mes proximo.

Estaba previsto que presidiera la clausura de un sinodo de obispos africanos, encabezara una conferencia de cardenales sobre los preparativos para la celebracion del ano 2000 y partiera el 13 de mayo para una visita de tres dias a Belgica. Tambien tenia programada una reunion con el canciller aleman Helmut Kohl.

El 18 de mayo el Papa cumplira 74 anos.

En junio, segun su agenda, tenia previsto entrevistarse con el presidente Bill Clinton, en el Vaticano.

Desde su cama, Juan Pablo continuara guiando a la Iglesia catolica, sin delegar la herencia de Pedro, asistido por el cardenal Sodano.

Aun convaleciente, podra firmar, como lo hizo en otras ocasiones, decretos urgentes.

Es perfectamente consciente y capaz de hacerlo, subrayo Navarro, quien sonriendo, anadio que no hay necesidad, ni estan previstas "regencias".

"Aqui, entre nosotros, no existe el problema de algun maletin negro con los codigos secretos", anadio Navarro.

El papa recibio los buenos augurios de personalidades de todas partes, el nuevo primer ministro de Italia, Silvio Berlusconi; el presidente italiano, Oscar Luigi Scalfaro; el rey de Espana Juan Carlos y los monjes budistas del Dalai Lama, entre otros.

Scalfaro senalo en su mensaje al Papa que "mientras, en el misterio amoroso de la Providencia, otro sufrimiento le es solicitado como oferta para la humanidad tan confusa y doliente, le envio mi augurio de rapida recuperacion, que es el augurio de todo el pueblo italiano."

El practicamente ex premier italiano, Carlo Azeglio Ciampi, los presidentes del Senado, Carlo Scognamiglio de la Camara de Diputados, Irene Pivetti, enviaron tambien telegramas en los cuales manifiestan su dolor por el accidente de Juan Pablo II y la esperanza de un restablecimiento rapido.

Al hospital empezaron a llegar numerosas monjas para rezar por la salud del jefe maximo de la Iglesia catolica.

Una de las primeras personas que llego a la policlinica fue la madre Elizabeth Patrizi, de la orden fundada por el martir de Auschwitz, con un gran ramo de flores.

Desde que fue elegido pontifice en 1978, Juan Pablo II ha pasado 106 dias en la Policlinica Gemelli, adonde ha sido intervenido seis veces.

La primera fue el 13 de mayo de 1981, a causa del atentado en la plaza de San Pedro.

La segunda, pocas semanas despues de salir por el atentado, a causa de una enfermedad viral.

En julio de 1992, volvio al hospital para ser operado de un tumor benigno del intestino.

En los meses sucesivos volvio para varios controles.

En diciembre de 1993 debio ser internado por una luxacion del hombro derecho causada por una caida durante una audiencia.

### 5.1.1 Resultados para "Operan al Papa"

En el archivo "Cuen-xxx", Clasitex deja las cuentas de los conceptos.

papa-catolico: 49	cirugia: 28	familia: 26	enfermedad: 17	accidente: 16
relig-catolica: 15	convalecencia: 11		hospital: 10	
anunciar-promulgar: 8	medicamento: 8	cuerpo-humano: 7	dirigente: 7	lugar-en-ciudad: 7
violento: 6	hueso-humano: 6	roma: 6	habitacion: 5	hecho-sufrimiento: 5
banarse: 4	casa-parte: 3	deporte: 3	economia: 3	empleo-puesto: 3
imss: 3	intervencion-meter: 3		acto-juridico: 2	alemania: 2
bebida-alcoholica: 2	budismo: 2		ciudad-del-vaticano: 2	dormir: 2
empresa: 2	escuela-colegio: 2		espana: 2	exponer-exhibir: 2

Hallando los temas de un artículo en español

fiesta: 2	gobie-actual: 2	hombre: 2	ley: 2	mueble: 2
noche: 2	religion: 2	sexual: 2	viaje: 2	africa: 1
agencia-noticia: 1	aleacion-metalica: 1		amor-pasion: 1	belgica: 1
computadora: 1	discutir: 1	eeuu: 1	eleccion-votar: 1	escritura: 1
flor: 1	gobierno: 1	jardin: 1	libro-revista: 1	maletin-veliz: 1
misterio: 1	muerte: 1	noticia-internacional: 1		paz: 1
periodico: 1	pierna-humana: 1		poder-legislativo: 1	renunciar: 1
salud: 1	secuencia: 1	sonreir: 1	tecnologia-ingenieria: 1	

Como se ve, los dos o tres primeros temas tienen una cuenta sobresaliente, y son los temas del artículo. El resto tiene una cuenta menor; esos contadores pueden estar correctos, o pueden deberse a votos incorrectos al votar una palabra por los varios conceptos que denota. Obsérvese asimismo que ciudad-del-vaticano solo recibió dos votos, debido a que la versión actual de Clasitex no tiene el algoritmo de propagación de votos (§3.2, observ. 1).

En el archivo "Res-xxx", Clasitex escribe en cada línea, primero el concepto escogido, y luego la palabra (o palabras) que lo seleccionaron. Este archivo se utiliza para entender los votos de palabras hacia conceptos, permitiendo afinar a Clasitex.

accidente: accidente.	accidente: disloco.	accidente: fractura.
accidente: fracturado.	accidente: fracturo.	accidente: hospital.
accidente: hospital.	accidente: hospital.	accidente: hospital.
accidente: hospital.	accidente: hospital.	accidente: hospital.
accidente: luxacion.	accidente: resbalar.	accidente: resbalo.
accidente: terapia.	acto-juridico: audiencia.	acto-juridico: audiencia.
africa: africanos.	agencia-noticia: afp.	
aleacion-metalica: aleacion metalica.		alemania: canciller aleman.
alemania: helmut kohl.	amor-pasion: amoroso.	anunciar-promulgar: alocucion.
anunciar-promulgar: boletin	.anunciar-promulgar: comento.	
anunciar-promulgar: conferencia.	anunciar-promulgar: declaracion.	
anunciar-promulgar: manifesto.	anunciar-promulgar: portavoz.	
anunciar-promulgar: vocero.	banarse: bano.	banarse: descalzo.
banarse: ducha.	banarse: mojado.	
bebida-alcoholica: presidente.	bebida-alcoholica: presidente.	
belgica: belgica.	budismo: dalai lama.	budismo: monjes budistas.
casa-parte: bano.	casa-parte: piso.	casa-parte: piso.
cirugia: anestesia.	cirugia: artroprotesis.	cirugia: cirujano.
cirugia: escayola.	cirugia: implantacion.	cirugia: inmovilizacion.
cirugia: intervencion quirurgica.	cirugia: intervencion quirurgica.	cirugia: intervencion quirurgica.
cirugia: intervencion.	cirugia: intervencion.	cirugia: operacion.
cirugia: operada.		

(Siguen otras palabras, he cortado la lista)

## 5.2 Una mujer española. Texto analizado por Clasitex

Selecciones del Reader's Digest. Junio 1994.

Una mujer española más guapa que fea, llamada María Antonia Pérez Blanco, ha sido la protagonista del más reciente escándalo del sexo y del poder en el Reino Unido. La historia de esta increíble aventura amorosa que ha ocupado las principales páginas de todos los medios informativos de Inglaterra, ha eclipsado las noticias de la propia corte de su graciosa majestad Isabel II.

No deja de ser apasionante y curioso el desarrollo de un relato que, como dice la letra de un aplaudido "bolero", "es la historia de un amor como no hay otro igual".

La verdadera historia de María Antonia Perez se origina cuando llega a Londres de niña, después del divorcio de sus padres; transcurrió el tiempo y la niña se convirtió en una joven dotada de grandes encantos físicos y decidió que era mejor la "buena vida" que tener que trabajar. Y de pronto se convirtió en la amante de un hombre de negocios iraquí, que la introdujo en los círculos de la alta sociedad inglesa. Al morir su amante, María Antonia adopta el nombre de "Bienvenida", nombre con el cual seguirá ya su trayectoria en la vida. Su amante no le dejó herencia alguna, pero Bienvenida Perez sabía ya el camino que había que seguir, y lo siguió.

Después de trabajar poco tiempo como secretaria en una compañía inmobiliaria, se convirtió en diseñadora de modas, inventándose un título en "diseño y confección" expedido supuestamente en Dallas, Texas, EU. Y dan comienzo los "milagros": como por arte de magia sus diseños y confecciones comienzan a ser seguidos por las damas de la aristocracia británica que la invitaban a fiestas, reuniones y cenas. Y fue en una de ellas donde conoció a sir Antony Buck, un alto personaje adinerado y diputado conservador, con quien se casó vestida de blanco a las tres semanas exactas de haberlo conocido.

Y es a partir de ese matrimonio, cuando se convertirá en lady Bienvenida Buck y cuando da comienzo el gran escándalo.

Lady Buck no se conforma con tener un esposo, sino que aspira a tener también un "amante" cosa que, por medio de sus encantos físicos, logra, cuando en otra reunión conoce a sir Peter Harding, un militar profesional, de muy alto rango y que es, nada menos, que ministro de la Defensa Nacional.

Pasa el tiempo, ya tenemos a Lady Buck convertida en esposa y con un "amante." Entonces a Bienvenida Perez se le ocurre otra "idea", que también le reportará cuantiosos beneficios económicos y hará llegar el escándalo a su punto máximo: discretamente se pone en contacto con un conocido hombre de "relaciones públicas", el cual a cambio de una comisión de 20%, negocia con un periódico sensacionalista y por una cantidad de 175.000 libras esterlinas, la historia de Bienvenida Perez, la española esposa de un diputado con escándalo en la Cámara de los Comunes y, al mismo tiempo, amante del jefe del ejército británico.

El periódico acepta el pago a cambio de la historia, pero exige lógicamente una condición: esta será un "beso en la boca" a su amante sir Peter Harding.

Así queda acordado. Y pocos días después, en un restaurante londinense, cuando Bienvenida aproxima sus rojos labios al amante, el cual le responde calurosamente, aparece un fotógrafo que recoge el momento estelar.

Momento estelar que apareciera en primera plana del periódico sensacionalista y cuyos efectos serán fulminantes: sir Peter Harding se ve obligado a dimitir, Bienvenida Perez, lady Buck, cobra las 175.000 libras esterlinas. La vida militar de sir Peter Harding totalmente arruinada; sir Antony Buck tramita el divorcio y lo obtiene; la esposa de sir Peter Harding, madre de tres hijos, hace lo mismo: los servicios secretos ingleses investigan ahora si Harding reveló a la española algunos "datos" de carácter militar y si el amante iraquí le comunicó algún secreto sobre la guerra del Golfo Pérsico. Y entre tanto, Bienvenida Perez cierra con broche de oro este escándalo contrayendo matrimonio con el ciudadano, de origen ruso, Nikola Sokolov, asegurando ella, que es "por amor", y recuerda la célebre canción que dice: "La española cuando besa, es que besa de verdad". Pero el beso que dio a sir Peter Harding fue el "beso de Judas".

### 5.2.1 Resultados para "Una mujer española"

El archivo Cuen-una-mujer-española contiene el siguiente histograma de conceptos:

inglaterra: 27	sexual: 27	amor-pasión: 21	escándalo: 10
periodicos: 6	españa: 4	iraq: 4	eeuu: 3
fiestas: 3	ahora: 2	congreso: 2	gobierno: 2
institución: 2	poder-legislativo: 2	reino: 2	casamiento: 1
cuerpo-humano: 1	días: 1	dirigente: 1	elecciones: 1
familia: 1	fútbol-soccer: 1	inteligencia-seguridad: 1	muerte: 1
musica: 1	rusia: 1	sociedad: 1	

Hallando los temas de un artículo en español

### 5.3 Perro de mar, México en el Mundial. Textos analizados por Clasitex

Un viejo perro de mar. Por Peter Muilenburg. Selecciones del Reader's Digest. Junio 1994  
Siempre nos había tenido con el alma en un hilo, hasta que de veras lo perdimos.

Aunque Santos ladraba furiosamente ante cualquier señal de que sus amos estuvieramos en peligro, ni siquiera lanzo un gemido cuando resbalo y cayo por la borda, a unos ocho kilometros de la costa venezolana. No lo echamos de menos hasta la mañana siguiente, despues de que anclamos nuestro queche, Breath, tras la escollera de Puerto Azul. De pronto, mi hijo Diego grito:

-Donde esta Santos?

Buscamos por todos los rincones. Diego incluso abrio la ultima bolsa de tortillas de maiz tostadas que nos quedaba, y la hizo crujir mientras llamaba a Santos, pero el animal habia desaparecido.

Mi esposa, Doroty, y yo fuimos a notificar la perdida al capital del puerto, que en esos momentos se hallaba ocupado con un torneo de pesca deportiva. No nos dio esperanzas, pero prometio indagar.

Caminamos hasta el bar de la playa, ordenamos dos cervezas y comenzamos a discutir lo que le diriamos a Diego.

En repetidas ocasiones habiamos temido por la vida de Santos. Siempre habia sido un pilluelo jugueton. Un amigo dijo que le daba la impresion de un hamster disfrazado de lobo, aunque mas bien parecia una zorra cruzada con un husky. Era el fiel companero de Diego.

Nos lo habian regalado algunos anos atras en un puerto de Florida, donde anclamos junto a una goleta con dos personas y siete perros a bordo. Jeanne y Vince eran criadores de schipperkes, perritos belgas adiestrados para vivir en las barcasas de los canales, y acababan de nacer cinco cachorros. Jeanne se ofrecio a escoger uno para nosotros.

Su barco necesita un schipperke -me aseguro.

Yo no era de la misma opinion, pero mis objeciones se desvanecieron ante las muestras de entusiasmo de la tripulacion. Recomende a los muchachos que eligieran un cachorrillo vivaz, y regresaron con una bola de pelo negro azabache que me miro a los ojos y me gruño.

Con el tiempo, Santos llego a ser un excelente perro de mar. Schipperke significa "capitancito" en flamenco. Entre las obligaciones de estos perros, en las barcasas donde los criaron originalmente, estaba la de nadar hasta la orilla de los frios canales para mordisquearles las patas a los caballos que tiraban de dichas embarcaciones, y acicatearlos de este modo. Tambien estaban adiestrados para ladrar si una persona caia al agua, por lo cual eran excelentes nineros. E incluso se esperaba que ayudaran al capitán a navegar.

Santos destacaba en todas estas tareas. Una noche, cuando amenazaba una tormenta, fuimos a buscar abrigo en la isla de Mayaguana, en las Bahamas. El sol no se habia asomado en 36 horas, y no estabamos seguros de nuestra posicion. Santos se levanto de repente y se puso a olfatear el viento y a gemir con ansiedad. Nos dejamos guiar por su nariz, y dos horas mas tarde avistamos tierra.

Este perro de mar era, ademas, un grato companero, especialmente durante las prolongadas vigiliass nocturnas. En tales ocasiones se deslizaba sin invitacion sobre el regazo del solitario timonel y, dando un suspiro de gusto, le apoyaba el hocico en la mano.

La inteligencia en un perro puede resultar peligrosa. Santos sabia bien lo que se esperaba de el, pero seguia sus propios impulsos, para divertirse, y siempre se metia en aprietos. Cuando cumpli tres meses ya habia estado a punto de ahogarse en dos ocasiones. En una de ellas andaba correteando por la cubierta y de pronto se fue por un imbornal. Oimos aranzos en el casco y, al asomarnos por la borda vimos a Santos chapoteando en el agua. Por fortuna el mar estaba en calma.

Un mes despues volvio a caerse, y la resaca lo arrastro mar adentro mientras nosotros almorzabamos en el interior del barco. Una persona que estaba buccando con un tubo respiratorio vio su cabeza negra y sus puntiagudas orejas, y lo rescato.

Era hora de que nos reunieramos con Diego. Ya no podiamos aplazar lo inevitable.

-Bueno -intente consolar a Doroty-, con la vivacidad de ese perro, fue una suerte haberlo tenido tanto tiempo.

Sus ojos se anegaron mientras decia:

-No hago mas que recordar los buenos momentos: con que suavidad tomaba la comida de mis manos, y como me lo agradecia siempre. Era un perro adorable.



Y se quedo solo en alta mar!

Una lagrima rodo por su mejilla.

Fuimos adonde estaba el esquire, y ya empezaba yo a soltar las amarras cuando oimos que nos llamaban. El capitán del puerto bajo corriendo la escalera de su oficina.

-No lo van a creer! -dijo-. Acabo de comunicarme por radio con los barcos de pesca para que me informaran de las cantidades que obtuvieron en la competencia.

Hizo una pausa para recobrar el aliento.

-Y...?

-Del ultimo me dieron que no habian pescado nada... excepto un perrito negro!

Aguardabamos en el muelle cuando la nave entro a puerto. Al estilo latino, toda la familia se hallaba a bordo, desde los niños hasta la abuela. Esta habia divisado a Santos entre las olas, e insistio en que su hijo se le acercara. Lo recogieron con una red de mano, y al final del día lamentaron tener que separarse de él.

De regreso en el Breath, Santos recibio una calurosa bienvenida, y un plato de los mismo que todos cenamos. Pero tenia la mirada perdida, como si ante ella acabara de pasar toda su existencia. Apenas terminada la cena se quedo dormido. Cuando, mas tarde, me fui a la cama, pense en lo bonita que era su vida, y me pregunte si por haber estado tan cerca de la muerte se habria vuelto mas sensato.

Mi fantasia se esfumo al amanecer, cuando nos despertaron unos ladridos: un pescador se habia atrevido a pasar a menos de 50 metros de nuestro barco. Horas despues, al ver a algunos niños jugando con un pastor alemán en la playa, Santos salto al mar sin vacilacion y se fue en pos de la aventura. Al poco rato, una persona golpeo en el caso de nuestro queche.

-Hola! Es este su perro?

-Si! -respondi-. Ese es nuestro perro!

#### CUEN-PERRO-DE-MAR

mar: 51	perro: 22	animal: 9	familia: 5
cuerpo-humano: 4		platillo-comida: 4	
comer: 3	deporte: 3	hombre: 3	nino: 3
noche: 3	religion: 3	unidad-medida: 3	
accidente: 2	amor-pasion: 2	bahamas: 2	bebida-alcoholica: 2
belgica: 2	dormir: 2	hecho-violento: 2	miedo: 2
aislado: 1	amenaza: 1	beisbol: 1	buceo: 1
carretera: 1			
casa-parte: 1	casamiento: 1	cereal: 1	convalescencia: 1
discutir: 1	eeuu: 1	enfermedad: 1	escandalo: 1
escoger: 1	escuela-colegio: 1	geografia: 1	habitacion: 1
inteligencia-seguridad: 1		italia: 1	llorar: 1
maletin-veliz: 1	militar: 1	mito-creencia: 1	lluvia: 1
oficina: 1	paz: 1	persona-vieja: 1	mueble: 1
muerte: 1		venezuela: 1	radiodifusion: 1
utensilio-cocina: 1			soga-hilo: 1
			vestimenta: 1
			viejo-antiguo: 1

#### MÉXICO EN EL MUNDIAL (ARTÍCULO ANALIZADO POR CLASITEX)

ARTICULO. Mexico: por su propia superacion. (Selecciones, junio 1994).

el Campeonato Mundial de los Estados Unidos 1994 significa la oportunidad de superacion para el seleccionado mexicano de futbol que dirige Miguel Mejia Baron. Un Mundial esperanzador para el futbol mexicano que así regresa al concierto del futbol internacional despues de aquel castigo de dos años que nos privo de asistir a la cita de Italia 1990.

En su paso por el Mundial, a Mexico le ha tocado en el Grupo E, el llamado de la muerte donde participara frente a los seleccionados de Italia, Eire y Noruega. La ventaja de esta situacion es que los tres europeos practican un mismo estilo de juego y bajo esa base, Mexico tiene la ventaja de no tener que entenderse por el momento con otros estilos diferentes. En su proceso, Mejia Baron observo a cerca de 60 jugadores y saco sus debidas conclusiones para afrontar el reto mundialista.

Hallando los temas de un artículo en español

El equipo mexicano cuenta con grandes individualidades que en conjunto pueden dar un extraordinario resultado. Todo comienza atrás con la presencia de Jorge Campos, el arquero-libero que juega al filo de la navaja, pero que con su inteligencia sale a flote. El Capi Ramirez Perales que con prestancia y buen futbol sale tocando de atrás y por el lateral izquierdo la presencia de Ramon Ramirez, convertido en figura importante del esquema nacional, en la media cancha es importante la contencion de Ignacio Ambriz y el trabajo creativo que puedan aportar Benjamin Galindo y Alberto Garcia Aspe, adelante opciones brillantes con Hugo Sanchez, maduro y en gran momento, la velocidad de Luis Garcia, el arranque de Zague y el olfato goleador de Carlos Hermosillo. Ellos forman la columna de un equipo que tiene ante sí \_tambien en cada posicion\_ elementos que han respondido a la confianza del tecnico nacional.

#### CUEN-MEXICO-EN-EL-MUNDIAL

futbol-soccer: 14	deportes: 8	mexico: 6	geografía: 3	mundo: 3
dias-meses: 2	escoger: 2	italia: 2	eeuu: 1	enfermedades: 1
hechos-violentos: 1		inteligencia-seguridad: 1	irlanda: 1	muerte: 1
musica: 1	noruega: 1	textiles-ropa: 1	trabajo: 1	trafico: 1

### 5.4 Conclusiones y comentarios

Hemos restringido nuestro análisis de textos en español a artículos que aparecen en revistas y periódicos de circulación general, no especializada. Esto con el fin de limitar el número de conceptos especializados.

En artículos cortos (México en el Mundial), Clasitex ofrece menos seguridad en sus temas: los histogramas contienen menos conceptos, y la diferencia entre los temas principales y los temas secundarios (y el ruido) es menos marcada.

El enfoque propuesto parece prometedor, pero hay que hacer muchos más experimentos para llegar a alguna conclusión definitiva. En particular, este enfoque requiere construir manualmente el árbol de conceptos.

La implementación actual está hecha usando el lenguaje de comandos (Shell) de Unix, y las utilerías awk, sed, etc. Esto permite modificaciones fáciles al programa, pero da una ejecución lenta (aproximadamente a la velocidad de lectura de una persona). Cuando se tenga una versión estable, reescribiremos los programas en C, guardando el árbol en estructuras en memoria, accedidas con índices. Emplearemos dispersión al azar ("hashing") si se usare memoria secundaria.

Algunas mejoras a Clasitex aparecen marcadas con \* en el texto. Lo principal es aumentar considerablemente el árbol de conceptos.

### 5.5 Bibliografía

1. Douglas B. Lenat and R. V. Guha. *Building large knowledge-based systems*. Addison Wesley, 1990
2. A. Guzmán. Finding the main themes in a Spanish document. To appear in *Expert Systems with Applications*, Elsevier, Jan. 1998.
3. A. Guzmán. Hallando los temas principales en un artículo en español. *Soluciones Avanzadas* 5, 47 (15 julio 97), págs. 58-63, y 5, 49 (15 sept. 97), págs. 66-71.