

Identificación de emociones a partir de texto usando desambiguación semántica

Emotion identification from text using semantic disambiguation

David García¹ y Francesc Alías²

¹GPMM - Grup de Processament Multimodal

²GTAM - Grup de Tecnologies Audiovisuals i Multimèdia

Enginyeria i Arquitectura La Salle

Universitat Ramon Llull

Quatre Camins 2, 08022 Barcelona

{dgarcia, falias}@salle.url.edu

Resumen: Este artículo presenta un sistema de identificación de emociones basado en texto con una arquitectura independiente del idioma. Éste usa distintas tareas de procesamiento del lenguaje natural además de un diccionario afectivo. Su principal novedad es la incorporación de un desambiguador semántico que permite considerar el significado de la palabra en la frase antes de categorizarla emocionalmente. Los experimentos muestran la mejora obtenida sobre un corpus de titulares en inglés.

Palabras clave: Etiquetado de emociones en textos, análisis gramatical y semántico de textos, WordNet, desambiguación semántica

Abstract: This paper presents a text-based emotion identification system based on text implemented by means of a language-independent architecture. The system includes several natural language processing tasks besides an affective keyword dictionary. The main novelty of the system is the incorporation of a semantic disambiguation module which focuses on the meaning of the word within the sentence before labelling it emotionally. The achieved results show the this purpose. The conducted experiments show the achieved improvement on a corpus of English headlines.

Keywords: Labelling of emotions in texts, grammatical and semantic analysis of texts, WordNet, word sense disambiguation

1. Introducción

Hoy en día, las tecnologías del habla son un elemento clave para conseguir interfaces hombre-máquina lo más naturales posible. En lo que se refiere a su etapa de salida, los Conversores de Texto en Habla (CTH) han alcanzado una buena inteligibilidad para sistemas de propósito general y una buena naturalidad para sistemas de dominio restringido. No obstante, todavía se continúa trabajando para lograr CTH expresivos, debido a la dificultad que supone, por un lado, generar el mensaje emocionado, y por otro, extraer la emoción del texto directamente. Para esto último, será de gran utilidad aplicar técnicas de Procesamiento del Lenguaje Natural (PLN) sobre los textos de entrada más allá del típico análisis de los sistemas de CTH (normalización, transcripción y estimación prosódica) (Alías et al., 2008).

Siguiendo con esta idea, en este trabajo se presenta un primer sistema que pretende extraer automáticamente la emoción intrínseca

del texto basándose en la aplicación de técnicas de PLN junto con un diccionario afectivo. El objetivo es que el sistema de CTH pueda generar el habla sintética con las características expresivas adecuadas al contenido emocional del texto. En la sección 2 se presentan trabajos relacionados con el sistema propuesto, descrito en la sección 3. Seguidamente, en las secciones 4 y 5, se analiza su funcionamiento y se discuten los resultados obtenidos. Finalmente, en la sección 6 se exponen las conclusiones y las líneas de futuro.

2. Trabajos relacionados

A continuación se presenta un breve resumen sobre la investigación en el ámbito de la identificación emocional de textos. Primero, se describen los modelos psicológicos emocionales que mejor se adaptan al mundo computacional. Seguidamente, se detallan las técnicas más utilizadas para la detección automática de emociones a partir de texto.

2.1. Modelo emocional

A continuación se revisan los modelos psicológico-computacionales más relevantes aplicables al ámbito de este trabajo.

El modelo más intuitivo para representar emociones es el basado en categorías emocionales, como son: alegría, tristeza, ira, etc. (Plutchik y Kellerman, 1980; Ekman, 1993). Un mejora de éste, es el modelo *Circumplex* (Schlosberg, 1952), que utiliza una circunferencia con dos ejes que representan sendas características emocionales, dando lugar a diferentes versiones según cuales sean: *i*) valencia (positivo/negativo) y activación (excitado/tranquilo) (Russell, 1980); o *ii*) positividad y negatividad (Watson y Tellegen, 1985).

Asimismo, existe un modelo similar al *Circumplex* llamado dimensiones emocionales (Schlosberg, 1954), que cuantifica las dimensiones de valencia, activación y control (dominado/dominante) mediante un vector de tres elementos. Finalmente, y en contraposición a los modelos anteriores, existe el modelo OCC (Ortony, Clore, y Collins, 1988), que presenta una jerarquía cognitiva de las emociones evitando el uso de categorías y dimensiones.

2.2. Detección de emociones

Este apartado resume las técnicas más usadas para la detección de emociones en textos en función del enfoque que siguen.

2.2.1. Basados en recuperación de información

Una primera técnica es hacer un análisis PMI-IR (*Pointwise-Mutual Information and Information Retrieval*) basado en dos iteraciones (Read, 2004). Una primera, usada inicialmente por Turney (2002) y denominada SO-PMI-IR (*Semantic Orientation using PMI-IR*), que indica la polaridad del texto, es decir, si es positivo o negativo. Y una segunda, conocida como AO-PMI-IR (*Affective Orientation using PMI-IR*), que extrae la dimensión afectiva predominante del texto.

Otro método es el etiquetado semántico + minería web (Lu, Hong, y Cruz-Lara, 2005), que reduce el texto a dos adjetivos que representan al núcleo del sujeto y del predicado, gracias a un análisis morfosintáctico y semántico. A continuación se buscan los hiperónimos de ambos adjetivos con una tabla de conversión. Por último, usando un conjunto de reglas cuyo patrón es verbo-hiperónimos, se determina la emoción del texto.

2.2.2. Basados en clasificación de textos

Una de las técnicas utilizadas con mayor éxito dentro del ámbito de la clasificación temática de grandes colecciones de texto es la basada en *Support Vector Machine* (SVM) (Joachims, 1998). En (Leshed y Kaye, 2006) se presenta un clasificador emocional de blogs que utiliza SVM.

En (Turney y Littman, 2003) se presenta un sistema de identificación de la polaridad del texto basado en *Latent Semantic Analysis* (LSA). Para saber la polaridad de cada palabra del texto, se calcula la diferencia entre su similitud con un conjunto de palabras positivas y otro de palabras negativas.

El problema fundamental de ambas técnicas radica en el elevado volumen de datos (entrenamiento y test) necesario para asegurar su buen funcionamiento. En el ámbito de los CTH lo más usual es disponer de textos de entrada reducidos (i.e. una frase) (Aliás et al., 2008). Asimismo, al trabajar con el texto como si fuera un único elemento, la información semántica que se puede obtener es poco precisa, no detectando, p.ej. la polisemia o la ironía (Liu, Lieberman, y Selker, 2003).

2.2.3. Basadas en diccionario afectivo

Estas técnicas se basan en buscar las palabras afectivas que contiene el texto en un diccionario de vocablos afectivos construido previamente. Destaca *Emotional Keyword Spotting* (EKS), debido a su sencillez de implementación. La emoción global del texto se determina a partir de la media de los valores emocionales de cada una de las palabras clave detectadas. Un ejemplo de esta técnica se presenta en (Francisco, Hervás, y Gervás, 2005), donde se aplica a la lectura expresiva de cuentos. Una extensión de EKS es la denominada afinidad léxica, que exporta la emoción de las palabras clave a sus palabras cercanas (Liu, Lieberman, y Selker, 2003). Ambas son incapaces de detectar cambios de polaridad de la emoción debido a elementos del texto, p.ej. negaciones (Francisco y Gervás, 2006).

Estas técnicas han sido aplicadas en el ámbito de los CTH, especialmente EKS (Sugimoto et al., 2004; Hofer, Richmond, y Clark, 2005), debido principalmente a su bajo coste computacional, esencial para lograr síntesis en tiempo real.

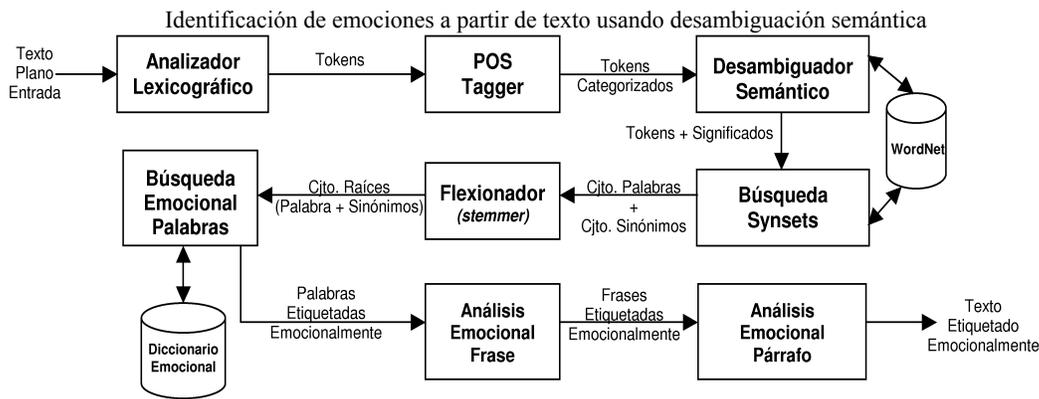


Figura 1: Diagrama de bloques del sistema

2.2.4. Otros

Caso aparte es el trabajo de Liu, Lieberman, y Selker (2003), que extraen conceptos de una voluminosa base de conocimiento del sentido común. La ventaja que aporta este sistema es la capacidad de detectar emociones en frases donde a priori no hay una emoción definida explícitamente. Se trata de una técnica compleja debido al tratamiento semántico que se debe hacer de los elementos de la base de conocimiento.

Asimismo, Ovesdotter, Roth, y Sproat (2005) presentan un sistema complejo que incorpora técnicas de inteligencia artificial para predecir la emoción del texto en el ámbito de la lectura de cuentos. Éste utiliza, además de palabras afectivas, parámetros del texto como la temática, la longitud de las frases, etc.

3. Sistema propuesto

El sistema que presentamos utiliza EKS y sigue una arquitectura base afín a la de Francisco, Hervás, y Gervás (2005), incorporándole un módulo de desambiguación semántica que permite extender dinámicamente el conjunto de palabras del texto a buscar en el diccionario afectivo (véase la figura 1).

A partir de un texto de entrada, el sistema obtiene como salida el texto etiquetado emocionalmente mediante marcas XML. Hasta el momento, el sistema se ha diseñado para detectar las siguientes emociones: alegría, sorpresa, tristeza, enfado y miedo, así como el estado neutro (sin emoción). Asimismo, la arquitectura es independiente del idioma, presentando en este trabajo su aplicación para el inglés, junto a las primeras adaptaciones para el español (con menores prestaciones¹).

¹Esta menor prestación es debido a la imposibilidad de disponer de WordNet 2.0 en español.

En cuanto al modelo emocional, el sistema utiliza el basado en dimensiones emocionales debido: *i*) al uso de los diccionarios afectivos basados en este modelo —ANEW (*Affective Norms for English Words*) (Bradley y Lang, 1999) y ANSW (*Affective Norms for Spanish Words*) (Redondo et al., 2007)—; y *ii*) a su futura integración dentro de un CTH expresivo, donde trabajar con datos continuos permite generar voz sintética con mayor variedad emocional (Hofer, Richmond, y Clark, 2005; Schröder, 2004).

A continuación, se describe cada uno de los módulos del sistema.

3.1. Analizador lexicográfico y PoS tagger

El primer paso consiste en analizar el texto de entrada mediante un conjunto de expresiones regulares (*tokens*) que determinan la gramática del idioma de análisis. Como resultado, se obtiene la categorización gramatical básica de las palabras del texto, permitiendo filtrar aquellas sin significado afectivo potencial (*stop words*), como por ejemplo los artículos. Después, se procede a desambiguar las palabras restantes, que comparten la misma expresión regular, para determinar su función gramatical (sustantivos, adjetivos y verbos). Para ello se utiliza el *PoS tagger* proporcionado por la Universidad de Stanford (Toutanova y Manning, 2000) para el inglés, y Qtag² para el español, debido a la implementación en Java de todo el sistema.

3.2. Desambiguador semántico

Cuando ya se tienen las palabras potencialmente emocionales, junto a sus correspondientes etiquetas gramaticales, se procede a

²www.english.bham.ac.uk/staff/omason/software/qtag.html

desambiguar el significado de los sustantivos del texto. El objetivo es que el siguiente módulo (búsqueda de *synsets*) sólo busque los sinónimos, antónimos e hiperónimos del significado del sustantivo en el texto, reduciendo así la probabilidad de falsa asignación (p.ej. se asigna a *ratón* los valores del diccionario correspondientes al significado de *animal* cuando en el texto se refiere a *dispositivo electrónico*). En este trabajo se utiliza una modificación del algoritmo *maximum relatedness disambiguation* propuesto por Pedersen, Banejee, y Patwardhan (2005), aplicándolo a todos los sustantivos del texto y con una ventana de desambiguación que abarca toda la frase. En cuanto a la medida de similitud semántica entre sustantivos, se utiliza la definida por Jiang y Conrath (1997), implementada en la librería *Java Similarity Library*³.

3.3. Búsqueda de *synsets*

Este módulo se encarga de extraer los sinónimos, antónimos e hiperónimo más directo de los adjetivos, verbos y sustantivos (con el significado indicado por el módulo anterior). De este modo, se amplía el espacio en la búsqueda de palabras. Así, aunque una palabra no esté en el diccionario afectivo, se puede buscar si lo está su sinónimo, antónimo o hiperónimo. Para encontrar este conjunto semántico se utiliza WordNet 2.0⁴. Al no disponer de WordNet para el español, sólo se usarán el desambiguador semántico y la búsqueda de *synsets* para el inglés.

3.4. Flexionador

Seguidamente se elimina la flexión de las palabras potencialmente afectivas y de sus sinónimos, antónimos e hiperónimo. Al trabajar con la raíz de la palabra se consigue tener una mejor cobertura del idioma. De este modo, no se necesitará tener *perro* y *perros* en el diccionario, sino que *perr* será la representación única para ambos vocablos. Esto resulta fundamental para los idiomas fuertemente flexionadores como el español. En consecuencia, el flexionador se aplicará tanto a la hora de construir el corpus como en el momento previo a la búsqueda de palabras. En este trabajo se usa el flexionador de Porter (1980) para el inglés, y el que ofrece el paquete Snowball⁵ para el español.

³eden.dei.uc.pt/~nseco/javasimlib.tar.gz

⁴wordnet.princeton.edu

⁵snowball.tartarus.org

3.5. Diccionario afectivo

Éste es el elemento clave de cualquier identificador de emociones basado en EKS. En este trabajo se utiliza ANEW para el inglés y su adaptación ANSW para el español. Ambos contienen 1034 palabras puntuadas en las tres dimensiones (valencia, activación y control) y normalizadas en este trabajo entre 0 y 10. En el diccionario, las palabras, además de estar etiquetadas emocionalmente, también tienen asignada la etiqueta gramatical para la cual fueron evaluadas (p.ej. la palabra *kiss* sólo aparece como nombre). Este campo será muy útil para distinguir casos como por ejemplo *afectar* y *afecto*, que al eliminar la flexión (se quedan en *afect*) crearían ambigüedad.

3.6. Búsqueda de palabras

Este módulo se encarga de buscar en el diccionario afectivo las palabras junto con sus sinónimos, antónimos e hiperónimos, una vez filtradas y lematizadas. Se trata de un proceso iterativo que finaliza cuando se encuentra un emparejamiento (coincidencia de raíz y categoría gramatical), o el conjunto de datos queda vacío. En caso de darse un emparejamiento, se le asigna a la palabra del texto los valores de valencia, activación y control indicados en el diccionario. En el caso de que el emparejamiento sea con el antónimo, se le asigna los valores complementarios. Si no hay éxito en la búsqueda, se le asigna el valor -1. Con este proceso se etiquetarán emocionalmente las palabras *afectivas* del texto.

3.7. Análisis a nivel de frase

Los valores de valencia, activación y control de cada frase se obtienen del promedio de los valores de las palabras emocionadas que forman la frase. Asimismo, se asigna a cada frase una etiqueta emocional en función de los valores promedio obtenidos. Estas etiquetas son: alegría, sorpresa, tristeza, enfado, miedo y estado neutro. Los valores asignados a estas etiquetas han sido establecidos siguiendo (Albrecht et al., 2005), situando como valores centroide de cada etiqueta los correspondientes a las palabras más comunes de cada emoción, y ajustándolos mediante la realización de pequeñas pruebas (véase el cuadro 1).

El sistema debe tomar en consideración ciertos fenómenos lingüísticos que pueden modificar la emoción de la frase, como los *modificadores emocionales* y las *negaciones*.

Categoría	Val.	Activ.	Control
Sorpresa	≥ 8.5	≥ 6.35	≥ 6.5
Alegría	≥ 6.445	≥ 5.86	≥ 5
Tristeza	≤ 3	≤ 4.575	> 1.5 y ≤ 3.75
Enfado	≤ 3.25	≥ 6.25	> 3.5 y ≤ 4.5
Miedo	≤ 3	< 7.5	≤ 3.75
Neutro	Resto o -1	Resto o -1	Resto o -1

Cuadro 1: Relación entre dimensiones y categorías emocionales del sistema propuesto.

En este trabajo, para el primero de ellos, sólo se toma en consideración la presencia de adverbios de cantidad detectados mediante el analizador lexicográfico. Resaltar que se consideran seis intensidades distintas (tres positivas y tres negativas). Según la intensidad asociada, se le asigna un valor fijado experimentalmente (± 0.25 , ± 0.40 ó ± 0.66) que incrementará/decrementará la valencia de la palabra emocional más cercana en el sentido adecuado. Para saber más sobre modificadores de valencia, se recomienda consultar (Polanyi y Zaenen, 2004).

Por su parte, el efecto causado por las negaciones se trata con un proceso que asigna a los *emotional keywords* valores complementarios cuando se encuentran después de un adverbio de negación. Para ello, se ha considerado, de un modo relajado, que la presencia de una negación no tiene por qué negar todas las palabras que la siguen. Así pues, el sistema niega todo lo que sigue a la negación hasta encontrar una conjunción adversativa (p.ej. sino, pero, etc.), manteniendo los valores que las siguientes palabras tienen en el diccionario afectivo. En (Francisco y Gervás, 2006) se presenta otro método más completo basado en el análisis de dependencias sintácticas.

3.8. Análisis a nivel de párrafo

Tras tener todas las frases categorizadas emocionalmente, se analiza el texto a nivel de párrafo asignándole la media, por separado, de las tres dimensiones de todas las frases que forman dicha estructura lingüística. Asimismo, se le asigna una etiqueta emocional. El análisis más exhaustivo de esta estructura queda abierto para trabajos futuros.

4. Análisis del sistema

En el estudio que se presenta a continuación se utiliza un corpus formado por 250 titulares de diarios en inglés, cuya emoción ha sido etiquetada manualmente estableciendo su grado de pertenencia (dentro del rango

[0,100]) a cada una de las categorías consideradas: enfado, disgusto, miedo, alegría, tristeza y sorpresa (Strapparava y Mihalcea, 2007). Para poder disponer de unos datos afines con los de la salida del sistema propuesto, resulta necesario reetiquetar manualmente las frases del corpus con la categoría emocional más representativa. Mencionar que se asigna al titular la etiqueta *enfado* si la emoción que destaca es *disgusto*, al ser ésta la emoción más afín de las consideradas por nuestro sistema.

El objetivo de este experimento es demostrar la mejora sucesiva del sistema a medida que se incorporan nuevos módulos. Para ello, se estudia la eficiencia de cuatro configuraciones distintas de nuestro sistema:

- **Simple:** sin WordNet ni desambiguación semántica. Sólo se hará una búsqueda de las raíces de las palabras potencialmente emotivas dentro del diccionario afectivo.
- **WN-Hiperónimos:** incorpora la ontología WordNet para incluir los hiperónimos en la búsqueda dentro del diccionario afectivo. Esta configuración intenta emular el sistema descrito en (Francisco, Hervás, y Gervás, 2005), pero sin la presencia de un diccionario adicional construido por los autores.
- **WN-Full:** es idéntica a la configuración anterior, pero añadiendo en la búsqueda los sinónimos y antónimos.
- **WN-Sem:** incorpora el módulo de desambiguación semántica de sustantivos al *WN-Full*. De este modo, se buscará en el diccionario únicamente los sinónimos, antónimos e hiperónimos que sólo hacen referencia al significado del sustantivo dentro del texto de entrada.

En la figura 2 se presenta la tasa de acierto de clasificación de las distintas configuraciones del sistema. De la gráfica, se puede observar una mejora absoluta de un 2% entre el sistema *Simple* y el *WN-Sem*, con una tasa de acierto global significativamente superior a la de un clasificador aleatorio (16,67% dadas las 6 categorías). Asimismo, se puede ver una tendencia creciente (con pequeñas mejoras sucesivas) a medida que se incorporan nuevas prestaciones al sistema. Una vez constatada esta tendencia global, se estudia el comportamiento de las cuatro configuraciones para cada una de las emociones de manera individual (véase la figura 3). Se puede

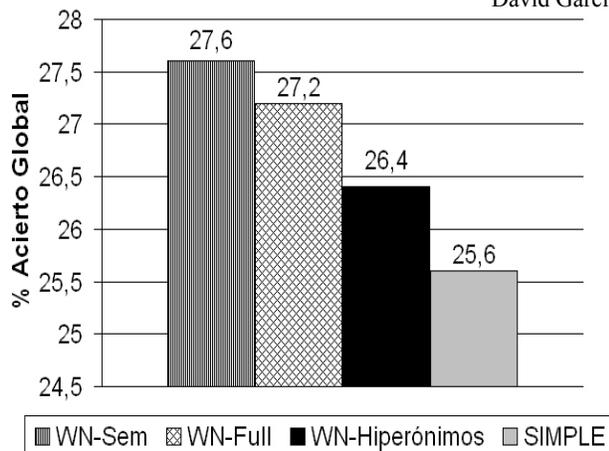


Figura 2: Tasa de acierto global de las diferentes configuraciones del sistema propuesto.

apreciar cómo la configuración *WN-Sem* presenta, como mínimo, la misma tasa de acierto que *Simple*. Destaca también la falta de acierto por parte de *todas* las configuraciones sobre las emociones *sorpresa* y *tristeza* (a excepción de un acierto de una frase triste por parte de *WN-Hiperónimos*, siendo la *tristeza* la única emoción en la que gana a *WN-Sem*). En el cuadro 2 se presenta la matriz de confusión de la mejor configuración del sistema: *WN-Sem*.

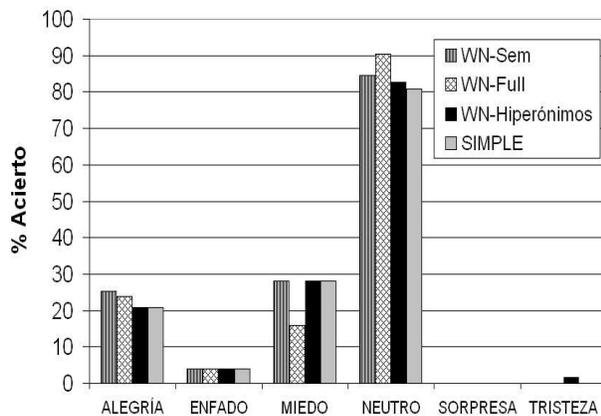


Figura 3: Porcentaje de acierto de las cuatro configuraciones para cada emoción.

5. Discusión

De los resultados obtenidos, se ha observado la mejora paulatina del sistema a medida que se incorporan nuevos módulos. No obstante, las reducidas mejoras porcentuales de la tasa de acierto llevan a analizar cualitativamente el funcionamiento de las distintas configuraciones del sistema mediante titulares concretos o ejemplos propios. De es-

Eti/Clas	Sorp.	Ale.	Tris.	Enf.	Mie.	Neu.
Sorp.	0 %	30 %	0 %	5 %	0 %	65 %
Ale.	0 %	25.37 %	0 %	0 %	2.98 %	71.64 %
Tris.	0 %	8.19 %	0 %	4.91 %	22.95 %	63.93 %
Enf.	0 %	8 %	0 %	4 %	12 %	76 %
Mie.	0 %	4 %	8 %	0 %	28 %	60 %
Neu.	0 %	9.61 %	0 %	1.92 %	3.84 %	84.61 %

Cuadro 2: Matriz de confusión del sistema para la configuración *WN-Sem*.

te análisis se concluye que, en ocasiones, las distintas versiones del sistema identifican la emoción de la frase correctamente pero por diferentes motivos. Éste es el caso de la frase “*This man is glad*” (Este hombre está alegre), categorizada por las cuatro configuraciones como *alegre*. Las versiones *WN-Sem* y *WN-Full* tienen en cuenta tanto la palabra *man* como *glad* (ya que encuentran su sinónimo *happy* en ANEW), mientras que las otras dos sólo utilizan *man*. Las dos últimas configuraciones aciertan debido a que *man* por sí solo ya está categorizada como *alegría* como consecuencia de los umbrales establecidos, cuestión totalmente subjetiva y discutible. Así pues, si en esta frase se cambiara *man* por otro vocablo que no estuviera directamente en ANEW, la clasificación fallaría para *Simple*, mientras que *WN-Hiperónimos* fallaría si no estuviera el hiperónimo de la nueva palabra. La diferencia principal entre *WN-Sem* y *WN-Full* está en que el primero identifica la emoción del texto siendo consecuente con su semántica, puesto que utiliza el sinónimo más adecuado para sustituir la palabra en función del contexto. Esto hace que *WN-Sem* sea más rico semántica y cualitativamente, aunque a veces *WN-Full* acierta igual que *WN-Sem* debido a los umbrales del sinónimo elegido (puede ser el mismo que en *WN-Sem*).

Para justificar por qué la tasa de acierto global, aunque significativamente superior a un sistema aleatorio, sobrepasa escasamente el 27 %, se debe decir que existen muchos titulares cuya semántica es imposible de extraer por cualquiera de los cuatro sistemas. Esto sucede, sobre todo, en los titulares en los que aparecen eventos concretos o personajes conocidos. Así, la frase “*Bush insists troops stay in Iraq, predicts midterm victory*” (Bush insiste en que las tropas permanezcan en Iraq, predice una victoria a medio plazo) es identificada por el sistema como po-

sitiva (debido a la idea de victoria), cuando las personas, conocedoras de una mayor información semántica, la podemos interpretar negativamente (como oposición a dicha decisión política). Así pues, tanto la subjetividad como el volumen de conocimiento semántico de las personas, son factores clave que afectan a los resultados.

Por último, en el cuadro 2 se puede ver al detalle qué sucede con la falta de acierto para *sorpresa* y *tristeza*. En el primer caso, ésta es confundida con *alegría* y *neutro*. Este hecho parece lógico, puesto que existen dos tipos de sorpresa: positiva y negativa. Para el caso positivo, que es el que se considera en este trabajo, una sorpresa puede ser perfectamente entendida como una alegría, lo cual no es un error demasiado crítico y más, si en nuestro caso, la futura salida será una voz emocionada. Suponemos que el 65 % asignado a neutro sería mucho menor si el sistema clasificara mejor la emoción *tristeza*, puesto que este porcentaje se repartiría entre ambas categorías. Por tanto, el caso más crítico es la *tristeza*, puesto que el hecho de que un 63.93 % de las frases se clasifiquen como *neutras* nos indica que hay que mejorar la clasificación de oraciones tristes. Destacar también que si se hicieran dos grandes grupos entre las categorías consideradas (sin contar *neutro*), podríamos obtener *positivas* (*sorpresa* y *alegría*) y *negativas* (el resto). Es decir, el grupo *negativas* estaría compuesto por tres emociones las cuales son muy cercanas entre sí a nivel de valores de las dimensiones (véase el cuadro 1), pero muy dispares en cuanto a semántica; mientras que en el primer grupo, aunque también cercanas en dimensiones, se pueden entender como una sola emoción tal y como se ha comentado anteriormente.

6. Conclusiones y trabajo futuro

En este trabajo se presenta un primer paso para la implementación de un sistema de detección automática de emociones a partir de un texto cualquiera. A diferencia de otros trabajos similares, el presente artículo presenta un sistema capaz de detectar emociones para el inglés y el español, considerando la sinonimia, antonimia e hiperonimia en función del significado de las palabras (desambiguación) dentro del texto. Aunque la incorporación de sucesivos módulos de procesamiento del lenguaje natural ha permitido una mejora sucesiva de la tasa de acierto global del sistema,

del análisis de los resultados obtenidos por categorías emocionales se puede observar la necesidad de continuar trabajando para mejorar las prestaciones del sistema.

Además de esto, se desea disponer del sistema completo para el español, al no tener la misma profundidad de análisis que en inglés. También se quiere incorporar un análisis de dependencia sintáctica que reconozca los elementos afectados por un cuantificador o un negador. Asimismo se desea un mejor análisis de estructuras lingüísticas, como: condicionales, adversativas, párrafos, etc. También se quiere, para ambos idiomas, reemplazar el actual módulo flexionador por un lematizador, mejorando así el acierto de búsqueda dentro del diccionario afectivo. A su vez se pretende trabajar con distintos diccionarios afectivos (p.ej. WordNet Affect, etc.)

Por otra parte, para comparar mejor el sistema con otros existentes, se desea crear un corpus de test multilingüe (inicialmente en inglés y español) para poder evaluar los sistemas de igual manera. Este corpus también permitirá evaluar el comportamiento del sistema para textos en español. Finalmente, respecto a su incorporación en un sistema CTH, se pretende encontrar las primeras relaciones entre parámetros de texto y de voz.

7. Agradecimientos

Los autores desean agradecer a J. Redondo la cesión del corpus ANSW.

Bibliografía

- Alías, F., X. Sevillano, J.C. Socoró, y X. Gonzalvo. 2008. Towards high quality next-generation Text-to-Speech synthesis: a Multi-domain approach by Automatic Domain Classification. *IEEE Transactions on Audio, Speech and Language Processing*. (aceptado).
- Albrecht, I., M. Schröder, J. Haber, y H. Seidel. 2005. Mixed feelings: expression of non-basic emotions in a muscle-based talking head. *Virtual Real*, 8(4):201–212.
- Bradley, M. y P. Lang. 1999. Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings. Informe Técnico C-1, Gainesville, FL, The Center for Research in Psychophysiology, Universidad de Florida.
- Ekman, P. 1993. Facial expression of emotion. *American Psychologist*, 48:384–392.
- Francisco, V. y P. Gervás. 2006. Análisis de dependencias para la marcación de cuentos con emociones. *Procesamiento de Lenguaje Natural*, (37):1135–5948, Septiembre.

- Francisco, V., R. Hervás, y P. Gervás. 2005. Expresión de emociones en la síntesis de voz en contextos narrativos. *Simpósio de Computación Ubicua e Inteligencia Ambiental*, Septiembre.
- Hofer, G., K. Richmond, y R. Clark. 2005. Informed blending of databases for emotional speech synthesis. En *Proc. Interspeech*.
- Jiang, J. y D. Conrath. 1997. Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. páginas 19–33.
- Joachims, T. 1998. Text categorization with support vector machines: learning with many relevant features. En Claire Nédellec y Céline Rouveilol, editores, *Proceedings of ECML-98*, numero 1398, páginas 137–142, Chemnitz, DE. Springer Verlag, Heidelberg, DE.
- Leshed, G. y J. Kaye. 2006. Understanding how bloggers feel: recognizing affect in blog posts. En Gary M. Olson y Robin Jeffries, editores, *CHI Extended Abstracts*, páginas 1019–1024. ACM.
- Liu, H., H. Lieberman, y T. Selker. 2003. A model of textual affect sensing using real-world knowledge. En *IUI '03: Proceedings of the 8th international conference on Intelligent user interfaces*, páginas 125–132, New York, NY, USA. ACM.
- Lu, C., J. Hong, y S. Cruz-Lara. 2005. Emotion detection in textual information by semantic role labeling and web mining techniques. *National ChiNan University and Universities of Nancy*.
- Ortony, A., G. Clore, y A. Collins. 1988. *The cognitive structure of emotions*. Cambridge University Press.
- Ovesdotter, C., D. Roth, y R. Sproat. 2005. Emotions from text: machine learning for text-based emotion prediction. En *Proceedings of HLT/EMNLP*, páginas 579–586, Vancouver, Canadá.
- Pedersen, T., S. Banejee, y S. Patwardhan. 2005. Maximizing semantic relatedness to perform word sense disambiguation. Marzo.
- Plutchik, R. y H. Kellerman. 1980. A general psychoevolutionary theory of emotion. *Emotion: Theory, research, and experience*, 1:3–33.
- Polanyi, L. y A. Zaenen. 2004. Contextual valence shifters. En *Exploring Attitude and Affect in Text: Theories and Applications*. AAAI Spring Symposium Series.
- Porter, M.F. 1980. An algorithm for suffix stripping. *Program*, 14(3):130–137, Julio.
- Read, J. 2004. Recognising affect in text using pointwise-mutual information. *University of Sussex*.
- Redondo, J., I. Fraga, I. Padrón, y M. Comesaña. 2007. The Spanish adaptation of ANEW (Affective Norms for English Words). *Behavior Research Methods*, 39(3):600–605, Agosto.
- Russell, J. 1980. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161–1178.
- Schlosberg, H. 1952. The Description of Facial Expressions in Terms of Two Dimensions. *Journal of Experimental Psychology*, 44:229–237.
- Schlosberg, H. 1954. Three dimensions of emotion. *Psychol. Rev.*, 61(2):81–88.
- Schröder, M. 2004. Dimensional emotion representation as a basis for speech synthesis with non-extreme emotions. En *Proc. Workshop on Affective Dialogue Systems*, páginas 209–220, Kloster Irsee, Germany. Springer Verlag.
- Strapparava, C. y R. Mihalcea. 2007. SemEval-2007 Task 14: Affective Text. *Proceedings of the 4th International Workshop on the Semantic Evaluations (SemEval)*, Junio.
- Sugimoto, F., K. Yazu, M. Murakami, y M. Yoneyama. 2004. Method to Classify Emotional Expressions of Text and Synthesize Speech. *Proceedings of the IEEE-EURASIP International Symposium on Control, Communications, and Signal Processing*, páginas 611–614.
- Toutanova, Kristina y Christopher D. Manning. 2000. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. En *Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in NLP and very large corpora*, páginas 63–70, Morristown, NJ, USA. Association for Computational Linguistics.
- Turney, P. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. páginas 417–424.
- Turney, P. y M. Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Trans. Inf. Syst.*, 21(4):315–346.
- Watson, J. y A. Tellegen. 1985. Toward a consensual structure of mood. *Psychological bulletin*, 98:219–235.