

---

**EmoTraductor 2.0. Análisis emocional de textos en  
español**  
**EmoTraductor 2.0. Emotional analysis of texts in  
Spanish**

---



**Trabajo de Fin de Máster**  
**Curso 2019–2020**

**Autora**  
**Paloma Gutiérrez Merino**

**Directoras**  
**Virginia Francisco Gilmartín**  
**Raquel Hervás Ballesteros**

**Máster en Ingeniería Informática**  
**Facultad de Informática**  
**Universidad Complutense de Madrid**



EmoTraductor 2.0. Análisis emocional de  
textos en español  
EmoTraductor 2.0. Emotional analysis of  
texts in Spanish

Trabajo de Fin de Máster en Ingeniería Informática  
Departamento de Ingeniería del Software e Inteligencia Artificial

**Autora**  
Paloma Gutiérrez Merino

**Directoras**  
Virginia Francisco Gilmartín  
Raquel Hervás Ballesteros

**Convocatoria:** *Julio 2020*  
**Calificación:** *9*

Máster en Ingeniería Informática  
Facultad de Informática  
Universidad Complutense de Madrid



# Autorización de difusión

El abajo firmante, matriculado en el Máster en Ingeniería en Informática de la Facultad de Informática, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo Fin de Máster: “EmoTraductor 2.0. Análisis emocional de textos en español”, realizado durante el curso académico 2019-2020 bajo la dirección de Virginia Francisco Gilmartín y Raquel Hervás Ballesteros en el Departamento de Ingeniería del Software e Inteligencia Artificial, y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en Internet y garantizar su preservación y acceso a largo plazo.

Paloma Gutiérrez Merino

24 de julio de 2020



# Agradecimientos

En primer lugar quiero dar las gracias a mis directoras en este trabajo, Virginia y Raquel, que a pesar de todas las dificultades con las que hemos tenido que lidiar este curso han dedicado mucho esfuerzo y tiempo a este trabajo y a mi formación. Siempre es un placer trabajar con vosotras y conseguís que quiera conseguir y mejorar en todo lo que me proponga.

En segundo lugar, quiero dar las gracias a mi familia y amigos por el apoyo que me han dado durante todo el máster y concretamente en este último año y empujón con este trabajo. Sobre todo quiero agradecerse a mi futuro compañero de vida que es capaz de conseguir que trabaje muchas horas con una sonrisa.



# Resumen

## **EmoTraductor 2.0. Análisis emocional de textos en español**

Existen personas para las que resulta difícil interpretar y reconocer el contenido afectivo de un texto. En personas con Trastornos de Espectro Autista (TEA) estas dificultades se incrementan, surgiendo la necesidad de crear herramientas para conseguir la integración digital de personas con este tipo de dificultad.

El EmoTraductor 1.0 es una primera aproximación a este tipo de herramientas para el castellano. Este trabajo tiene como objetivo evaluar dicha herramienta e intentar mejorar su marcado emocional añadiendo distintas mejoras como el tratamiento de la negación o los modificadores. Este TFM realiza un análisis de los resultados obtenidos con el EmoTraductor 1.0 para identificar sus deficiencias y mejorarlo. Entre las mejoras propuestas caben destacar: el tratamiento de los modificadores y el tratamiento de la negación. Cada una de las mejoras propuestas ha sido implementada y posteriormente evaluada para comprobar su validez y la mejora en el marcado conseguida. Todas aquellas mejoras que se tradujeron en un avance en la detección de la carga emocional de un texto han sido incorporadas a una nueva versión de la herramienta, el EmoTraductor 2.0.

Finalmente, tras una evaluación final de los resultados, la nueva versión del EmoTraductor obtiene mejores resultados que la primera versión. No obstante, también se identifica aún trabajo por realizar para continuar mejorando el marcado emocional como mejorar el tratamiento de la negación o la ampliación del diccionario emocional que usa el sistema como base.

## **Palabras clave**

Análisis emocional, Computación Afectiva, Emociones, categorías emocionales, Negación, Modificadores, Accesibilidad



# Abstract

## **EmoTraductor 2.0. Emotional analysis of texts in Spanish**

Sometimes, it is difficult to interpret the information we find on the internet and recognize the affective content of a text, and the difficulty increases as a result of the growth of the Internet. In people with Autism Spectrum Disorders (ASD) these difficulties increase even more. As a consequence, the need to create tools to achieve digital integration for people with different abilities appears.

Thanks to affective computing, great advances are being made for its digital integration by creating tools capable of emotionally labeling texts using emotional marking methods.

Although most of the work and advances in affective computing have been done in English, we know some tools in Spanish that are doing the same task or we expect them to do it in the future.

One of these tools is the EmoTraductor 1.0 developed in 2018 at the Faculty of Computing at the Complutense University of Madrid.

This project aims to evaluate this tool and try to improve its emotional marking of texts by adding different improvements such as negations or modifiers.

A study of EmoTraductor 1.0 is carried out to improve its deficiencies, the treatment of modifiers is studied and a treatment method is applied. The treatment of negation is also studied, presenting different methods. Each of the proposed improvements is associated with a serie of evaluations to check their validity. Finally, it is possible to improve the emotional marking of texts by correcting the malfunctions of the EmoTraductor 1.0 application, adding the treatment of modifiers and the treatment of negation.

## **Keywords**

Emotional Analysis, Affective Computing, Emotions, Emotional Categories, Denial, Modifiers, Accessibility



# Índice

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación	1
1.2. Objetivos	2
1.3. Estructura de la memoria	3
<b>2. Estado del arte</b>	<b>5</b>
2.1. Computación Afectiva	5
2.2. Representación de las emociones	7
2.3. Marcado emocional en textos	8
2.4. Diccionarios afectivos	11
2.4.1. ANEW	11
2.4.2. S-ANEW	11
2.4.3. WordNet-Affect	13
2.4.4. Diccionario de Díaz et al.	13
2.4.5. Corpus de Ferré	14
2.4.6. MADS (Madrid Affecive Database for Spanish)	15
2.4.7. Diccionario de Ferré	16
2.4.8. Diccionario de Hans Stadthagen-González	16
2.5. El papel de los emoticonos en el marcado emocional de textos	18
2.6. El papel de los modificadores en el marcado emocional de textos	19
2.7. Papel de la negación en el marcado emocional de textos	23
2.8. Herramientas existentes para detectar la emoción de un texto	26
2.8.1. EmoTag	26
2.8.2. Sentitex	29
2.8.3. Modelo de clasificación de textos según su polaridad e intensidad	30
2.8.4. Lingmotif	33
<b>3. EmoTraductor 1.0: una aplicación web para el marcado emocional</b>	<b>37</b>
3.1. Aplicación web del EmoTraductor 1.0	37
3.2. API de servicios web del EmoTraductor 1.0	39
3.2.1. Servicios para el análisis emocional de una palabra	40
3.2.2. Servicios para el análisis emocional de una frase	42
3.2.3. Servicios para el análisis emocional de un texto	44
3.3. Resultado de la evaluación de la herramienta	47

3.3.1.	Evaluación del mercado emocional . . . . .	47
3.3.2.	Evaluación de la aplicación web . . . . .	47
3.3.3.	Posibles mejoras sobre el EmoTraductor 1.0 . . . . .	48
<b>4.</b>	<b>Primeros pasos hacia el EmoTraductor 2.0</b>	<b>49</b>
4.1.	Detección de mejoras preliminares . . . . .	49
4.2.	Normalización de los valores emocionales del diccionario . . . . .	50
4.3.	Ampliación del diccionario emocional . . . . .	53
4.4.	Cambio en el criterio de selección de palabras emocionales . . . . .	53
4.5.	Corrección en el servicio de grados emocionales de un texto . . . . .	55
4.6.	Mejora en el servicio de obtención de grados emocionales de una palabra . . . . .	55
4.7.	Peso de las palabras no emocionales en la emoción final de la frase . . . . .	57
4.8.	Evaluación . . . . .	58
4.8.1.	Diseño de la evaluación . . . . .	58
4.8.2.	Resultados . . . . .	59
4.8.3.	Discusión de resultados . . . . .	67
<b>5.</b>	<b>Detección de modificadores y su papel en la marcación emocional</b>	<b>69</b>
5.1.	Listado de modificadores . . . . .	70
5.2.	Detección del ámbito de los modificadores . . . . .	70
5.3.	Marcación emocional de una frase teniendo en cuenta los modificadores . . . . .	74
5.4.	Evaluación del impacto del tratamiento de los modificadores . . . . .	76
5.4.1.	Diseño de la evaluación . . . . .	77
5.4.2.	Resultados . . . . .	77
5.4.3.	Discusión de resultados . . . . .	78
<b>6.</b>	<b>Detección de la negación y su papel en la marcación emocional</b>	<b>81</b>
6.1.	Tipos de negación . . . . .	81
6.2.	Tratamiento de la negación nominal . . . . .	82
6.3.	Tratamiento de la negación sintáctica . . . . .	83
6.3.1.	Anulación de la aportación emocional de palabras negadas . . . . .	85
6.3.2.	Negación basada en antónimos de palabras . . . . .	87
6.3.3.	Inversión de los grados emocionales de las palabras en el ámbito de la negación . . . . .	89
6.3.4.	Inversión de los grados emocionales significativos de las palabras en el ámbito de la negación . . . . .	90
6.3.5.	Uso de la antonimia de las categorías emocionales . . . . .	93
6.4.	Conclusiones . . . . .	97
<b>7.</b>	<b>EmoTraductor 2.0</b>	<b>99</b>
7.1.	Mejoras añadidas al sistema . . . . .	99
7.2.	Cambios en la interfaz de la aplicación web . . . . .	106
7.3.	Cambios en los servicios web . . . . .	107
7.4.	Evaluación del mercado emocional de texto del EmoTraductor 2.0 . . . . .	107
<b>8.</b>	<b>Conclusiones y Trabajo Futuro</b>	<b>111</b>
8.1.	Conclusiones . . . . .	111

8.2. Trabajo Futuro . . . . .	115
<b>9. Introduction</b>	<b>117</b>
9.1. Motivation . . . . .	117
9.2. Objectives . . . . .	118
9.3. Memory structure . . . . .	118
<b>10. Conclusions and future work</b>	<b>121</b>
10.1. Conclusions . . . . .	121
10.2. Future work . . . . .	125
<b>Bibliografía</b>	<b>127</b>
<b>A. Textos para evaluar el mercado emocional</b>	<b>131</b>
A.1. Textos usados en la evaluación del EmoTraductor 1.0 . . . . .	131
A.1.1. Aladdín . . . . .	131
A.1.2. Blancanieves . . . . .	131
A.1.3. Blancanieves y sus amigos . . . . .	131
A.1.4. Blog agradecimiento . . . . .	131
A.1.5. Blog enfado . . . . .	132
A.1.6. Blog moda . . . . .	132
A.1.7. Blog muerte . . . . .	132
A.1.8. Blog reflexión . . . . .	132
A.1.9. Cenicienta . . . . .	133
A.1.10. El castillo de bestia . . . . .	133
A.1.11. El ratoncito Pérez . . . . .	133
A.1.12. La bella durmiente . . . . .	134
A.1.13. La bella y la bestia . . . . .	134
A.1.14. La margarita . . . . .	134
A.1.15. La sirenita . . . . .	134
A.1.16. Mulán . . . . .	135
A.1.17. Noticia OMS . . . . .	135
A.1.18. Noticia muerte . . . . .	135
A.1.19. Noticia nacimiento . . . . .	136
A.1.20. Noticia política . . . . .	136
A.1.21. Pinocho . . . . .	136
A.1.22. Pocahontas . . . . .	137
A.1.23. Tipos . . . . .	137
A.2. Textos añadidos en la evaluación del nuevo EmoTraductor . . . . .	137
A.2.1. Centro comercial . . . . .	137
A.2.2. Conversación1 . . . . .	138
A.2.3. Conversación2 . . . . .	138
A.2.4. DobleNegación1 . . . . .	138
A.2.5. DobleNegación2 . . . . .	138
A.2.6. DobleNegación3 . . . . .	138
A.2.7. FraseCorta . . . . .	138
A.2.8. FraseCorta2 . . . . .	138

A.2.9. Negación . . . . .	138
A.2.10. Negación2 . . . . .	138
A.2.11. Negaciones . . . . .	139
A.2.12. Negaciones coordinadas . . . . .	139
A.2.13. Noticia deportes . . . . .	139
A.2.14. Peli aburrida . . . . .	139
A.2.15. Preguntas tristes (oración) . . . . .	139
A.2.16. Texto1 . . . . .	140
A.2.17. Texto2 . . . . .	141
A.2.18. Texto miedo a perder el amor . . . . .	141
A.2.19. Tristeza . . . . .	142
A.2.20. Tristeza2 . . . . .	142

# Índice de figuras

2.1. Clasificación de 16 palabras emocionales en el modelo circunflejo de Russell	8
2.2. Porcentajes de probabilidad de uso afectivo en palabras del diccionario “alegría” según Díaz et al. (2014)	14
2.3. Emoticonos y sentimiento/palabra asociada de Vashisht y Thakur (2014)	20
2.4. Léxico de emoticonos de Wegrzyn-Wolska et al. (2016)	21
2.5. Lista de modificadores de Carrillo (2011)	22
2.6. Árbol de dependencias obtenido para la frase “Ese ordenador es muy rápido pero no recomiendo comprarlo” según Vilares et al. (2013)	24
2.7. Reglas para definir el ámbito de la negación por Jiménez et al. (2015)	24
2.8. Árbol de dependencias obtenido para la frase “La película no empieza mal con algunas escenas interesantes para introducirte en el drama vivido por el personaje principal” según Jiménez et al. (2015)	25
2.9. Términos incluidos en la lista de negaciones de Carrillo (2011)	26
2.10. Lista de delimitadores del ámbito de la negación de Carrillo (2011)	26
2.11. Funcionamiento de la inversión de la categoría emocional en el EmoTag.	28
2.12. Valores emocionales para una frase de ejemplo en EmoTag.	29
2.13. Niveles de etiquetas emocionales en OntoEmotion.	29
2.14. Ejemplo de datos proporcionados por Lingmotif	34
2.15. Ejemplo de un texto marcado por Lingmotif	34
2.16. Ejemplo de reglas de contexto que emplea Lingmotif	35
3.1. Esquema de la arquitectura del EmoTraductor	38
3.2. Página principal EmoTraductor 1.0	38
3.3. Resultados del traducción	39
3.4. Resultado de colocar el ratón sobre la palabra «tragedia»	39
3.5. Resultado servicio grados de la palabra de la palabra abejas	40
3.6. Resultado servicio emoción mayoritaria de la palabra “abejas”	42
3.7. Resultado servicio emoción mayoritaria en caso de que tenga más de una emoción mayoritaria	42
3.8. Resultado del servicio de grados emocionales de la frase para la frase “Me gustan las naranjas”	43
3.9. Resultado del servicio de emoción mayoritaria para la frase “Me gustan las naranjas”	44
3.10. Resultado del servicio de grados emocionales de textos para “Salir a la calle es divertido. Ir al circo es divertido”	45

3.11. Resultado del servicio de grados emocionales de textos para “El pequeño cocodrilo. Para aprender sus cantares usa las aguas del Nilo con sus notas musicales.” . . . . .	46
3.12. Resultado servicio emoción mayoritaria del texto para el texto “Me gusta ir al parque. Es muy divertido” . . . . .	47
5.1. Modificadores usados en el trabajo de Carrillo (2011) . . . . .	72
5.2. Árbol de dependencias generado por Spacy para la frase “La niña está muy feliz y contenta” . . . . .	73
5.3. Árbol de dependencias creado por Spacy de la frase “La niña está muy feliz y contenta” . . . . .	73
5.4. Árbol de dependencias creado por Spacy de la frase “La mujer es súper calmada pero le gustan mucho las fiestas” . . . . .	74
6.1. Árbol de dependencias creado por Spacy de la frase “Mi madre estaba casi enfadada cuando salí de casa” . . . . .	83
6.2. Árbol de dependencias creado para la frase “No tenemos dinero” . . . . .	85
6.3. Árbol de dependencias creado para la frase “No vivimos en esa casa fea” . . . . .	85
6.4. Árbol de dependencias creado para la frase “La niña no es fea” . . . . .	85
6.5. Árbol de dependencias creado para la frase “Ella no es una niña fea” . . . . .	86
7.1. Árbol de dependencias creado por Spacy de la frase “No estoy enfadada” . . . . .	100
7.2. Árbol de dependencias creado por Spacy de la frase “¡Estoy muy alegre!” . . . . .	101
7.3. Interfaz de la aplicación web EmoTraductor 1.0 . . . . .	106
7.4. Interfaz de la aplicación web EmoTraductor 2.0 . . . . .	106

# Índice de tablas

2.1.	Primera y última palabra del diccionario ANEW y sus campos . . . . .	11
2.2.	Primera y última palabra del SANEW y sus campos . . . . .	12
2.3.	Extracto de cinco palabras del Corpus de Ferré. . . . .	15
2.4.	Extracto del diccionario de Hinojosa . . . . .	16
2.5.	Fragmento del diccionario de Ferré . . . . .	17
2.6.	Resultado de la agrupación de etiquetas emocionales por niveles según On- toEmotion. . . . .	29
3.1.	Resultado servicio grados palabra Emotraductor 1.0 con la palabra “alegría”	41
3.2.	Resultado del servicio de grados emocionales de un frase con la frase “Estoy alegre y feliz” . . . . .	43
3.3.	Resultado del servicio de grados emocionales para textos del Emotraductor 1.0 . . . . .	45
4.1.	Grados emocionales de las palabras y de la frase “Me gusta abrazar a la chica”	51
4.2.	Comparativa de los resultados grados emocionales de “Me gusta abrazar a la chica” . . . . .	51
4.3.	Comparativa de los resultados de los grados emocionales de “¡Me gusta abra- zar a la chica! Me gusta mucho.” . . . . .	51
4.4.	Grados emocionales de las palabras de “Cuando habla el tío odioso sola- mente pienso en abrazar perros” y su resultado total como frase según el EmoTraductor 1.0 . . . . .	52
4.5.	Grados emocionales de las palabras de “Cuando habla el tío odioso sola- mente pienso en abrazar perros” y su resultado total como frase usando el diccionario normalizado . . . . .	52
4.6.	Comparativa de los resultados del servicio de grados emocionales de “¡Cuan- do habla el tío odioso solo pienso en abrazar perros! Eso me relaja.” . . . . .	52
4.7.	Grados emocionales de las palabra “nubes” normalizada . . . . .	54
4.8.	Grados emocionales de las palabras candidatas a ser emocionales de “Estaba preocupado con la situación” normalizados . . . . .	54
4.9.	Grados emocionales de la frase “Estaba preocupado con la situación” . . . . .	55
4.10.	Resultado de las palabras “abadía” y “abad” en el diccionario . . . . .	57
4.11.	Resultado de las palabras “abuelo” y “abuela” en el diccionario . . . . .	57
4.12.	Resultado de los grados de las palabras candidatas a ser emocionales en la frase: “Si achuchas a los bebés afligidos ayudas” . . . . .	58

4.13. Resultado de los grados finales de la frase “Si achuchas a los bebés afligidos ayudas” teniendo en cuenta distintos pesos para las palabras no emocionales.	58
4.14. Resultados de los textos de prueba del EmoTraductor 1.0 . . . . .	60
4.15. Resultados de los textos de prueba después de aplicar la normalización del diccionario. . . . .	61
4.16. Resultados de los textos tras las primeras mejoras . . . . .	62
4.17. Tabla I de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.	63
4.18. Tabla II de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.	64
4.19. Tabla III de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.	65
4.20. Tabla IV de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.	66
5.1. Grados emocionales devueltos por el EmoTraductor 1.0 para tres frases que difieren únicamente en los modificadores. . . . .	69
5.2. Modificadores incluidos en el EmoTraductor con sus porcentajes asociados .	71
5.3. Grados emocionales considerando modificadores de “La niña está muy feliz y contenta” . . . . .	74
5.4. Categorías gramaticales de la frase “La mujer es súper calmada pero le gustan mucho las fiestas.” . . . . .	75
5.5. Grados emocionales sin y con modificadores de la frase “La mujer es súper calmada pero le gustan mucho las fiestas.” . . . . .	76
5.6. Grados emocionales finales para la frase “La mujer es súper calmada pero le gustan mucho las fiestas.” . . . . .	77
5.7. Grados emocionales considerando modificadores . . . . .	77
5.8. Grados emocionales considerando modificadores para distintas frases con y sin modificadores . . . . .	78
5.9. Comparación de los resultados de los textos de prueba después las primeras mejoras y al añadir los modificadores. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco). . . . .	79
6.1. Listado de modificadores considerados en la negación nominal de nuestro trabajo. . . . .	82
6.2. Grados emocionales de las palabras emocionales de la frase “Mi madre estaba casi enfadada cuando salí de casa”. . . . .	83
6.3. Grados emocionales de la frase “Mi madre estaba casi enfadada cuando salí de casa”. . . . .	84
6.4. Grados emocionales de la frase “Me gustó ir al cine” y “No me gustó ir al cine” con el EmoTraductor 1.0 . . . . .	86
6.5. Grados emocionales de la frase “No me gustó ir al cine” con la primera aproximación implementada para tratar la negación. . . . .	87
6.6. Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir la negación eliminando los grados de las palabras en el ámbito de la negación. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco). . . . .	88

6.7.	Grados emocionales en el diccionario afectivo para las palabras “ir” y “venir”	89
6.8.	Grados emocionales del diccionario e invertidos de las palabras “ir” y “cine”	90
6.9.	Grados emocionales sin considerar y considerando la negación para la frase “No iremos al cine”	90
6.10.	Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir la negación por inversión de todas las emociones. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).	91
6.11.	Grados emocionales de la palabras “baila” y “fiesta” según el diccionario y después de aplicar el método de inversión para las emociones que tienen un grado mayor que 1,5	92
6.12.	Evaluación de la frase “Ella no baila en la fiesta” sin tener en cuenta la negación, teniendo en cuenta la negación usando el primer método de inversión y con el nuevo método de inversión de emociones superiores a 1,5	93
6.13.	Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir el segundo método de la negación por inversión en solo las categorías emocionales que superan el 1,5. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).	94
6.14.	Grados emocionales del diccionario para “adicto” y “drogas”	96
6.15.	Grados emocionales de la frase “No es adicto a las drogas”	96
6.16.	Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir el tercer método de la negación , usando la división de tipos de emociones en fuertes y débiles y las relaciones de antonimia. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).	98
7.1.	Palabras de la frase “No estoy enfadada” con su posición, categoría gramatical y si son candidatas a ser emocionales o no.	101
7.2.	Palabras de la frase “¡Estoy muy alegre!” con su posición, categoría gramatical y si son candidatas a ser emocionales o no.	101
7.3.	Palabras del texto introducido y su respuesta ante la primera evaluación que se les realiza con Spacy y PyStemmer	102
7.4.	Palabras candidatas a ser emocionales en nuestro texto “No estoy enfadada. ¡Estoy muy alegre!” junto con si son palabras emocionales y sus grados emocionales.	103
7.5.	Palabras emocionales en nuestro texto “No estoy enfadada. ¡Estoy muy alegre!” con los efectos de la negación y los modificadores aplicados	105
7.6.	Grados emocionales de la frase “No estoy enfadada” con el EmoTraductor 2.0	105
7.7.	Grados emocionales de la frase “¡Estoy muy alegre!” con el EmoTraductor 2.0	105
7.8.	Grados emocionales del texto “No estoy enfadada. ¡Estoy muy alegre!” con el EmoTraductor 2.0	106
7.9.	Comparación de los resultados de los textos de prueba entre lo que devolvía el EmoTraductor 1.0 y el EmoTraductor 2.0. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).	109



# Introducción

**RESUMEN:** En este capítulo se va a realizar una introducción al trabajo que vamos a presentar. En primer lugar, en la Sección 1.1 se explicará la motivación que hay detrás de este trabajo. A continuación en la Sección 1.2 se presentarán los objetivos propuestos para el trabajo. Por último, en la Sección 1.3 se muestra la estructura que tendrá la memoria.

## 1.1. Motivación

Gracias al rápido crecimiento de Internet y la rápida evolución de los sistemas de comunicación, cada vez tenemos disponible más información al alcance de nuestra mano. Esto supone muchas ventajas en nuestro día a día, pero también incrementa la complicación que conlleva entender e interpretar toda la información que obtenemos y saber cómo manejarla. Las dificultades que comentamos en algunas ocasiones aumentan debido a la falta de empatía emocional que poseen las personas con Trastornos del Espectro Autista (TEA), especialmente con Síndrome de Asperger. Esta falta de empatía emocional les impide reconocer el contenido afectivo de un texto pudiendo malinterpretar la carga emocional del mismo. También les impide en ocasiones poder comunicarse adecuadamente online con otras personas ya que no saben qué emociones pueden ser interpretadas en los textos que ellos mismos escriben.

Para solucionar estos problemas surge la necesidad de crear herramientas capaces de interpretar la carga afectiva de los textos para ayudar a la integración digital de las personas con este tipo de dificultades. Gracias a la computación afectiva estas herramientas pueden desarrollarse utilizando técnicas de Procesamiento de Lenguaje Natural (PLN). Sin embargo, a pesar de que el Procesamiento de Lenguaje Natural se emplee cada vez en más situaciones y de manera exitosa, sigue siendo complicado encontrar el mejor método para interpretar el contenido afectivo de un texto debido al grado de subjetividad no solo de las emociones, sino también del propio lenguaje.

A toda la complicación que supone interpretar los textos mediante PLN se suma que la mayor parte de los estudios y trabajos que se realizan en este campo se hacen en lengua inglesa. Así, encontramos pocos trabajos y herramientas que se utilicen de manera estándar o con mayor relevancia en castellano y que puedan servirnos de guía para cualquiera de los aspectos que necesitamos. No obstante, sí existen algunas herramientas que utilizan

el castellano para realizar este tipo de interpretaciones. Por ejemplo, una de estas herramientas es el EmoTraductor 1.0 desarrollado en 2018 en la Facultad de Informática de la Universidad Complutense de Madrid, como parte del TFG de la autora de este TFM junto a otras dos compañeras (Eugercios et al., 2018).

El EmoTraductor 1.0 es una aplicación web desarrollada con el objetivo de marcar emocionalmente textos introducidos por los usuarios mediante una interfaz web. Esta aplicación utiliza categorías emocionales para marcar los textos, concretamente las cinco emociones básicas (tristeza, miedo, alegría, enfado y asco). La presencia de cada una de las categorías en el texto se muestra mediante lo que llamamos grados emocionales, que son números que indican la presencia de la emoción en el texto (siendo el 0 el valor que indica su ausencia). Aunque los resultados obtenidos en el EmoTraductor 1.0 fueron prometedores, el marcado emocional presentaba deficiencias debido a que no se consideraban algunos fenómenos complejos del lenguaje natural.

El principal objetivo de este trabajo será mejorar el EmoTraductor 1.0 para mejorar el marcado emocional de texto incorporando un procesamiento de lenguaje natural más avanzado que el realizado en la versión anterior. En este TFM analizaremos cómo distintos fenómenos que se presentan en el lenguaje natural (negaciones, modificadores...) afectan a la carga emocional de un texto e incorporaremos su tratamiento al traductor emocional con el fin de mejorar los resultados que este obtiene.

## 1.2. Objetivos

El objetivo principal de este trabajo es mejorar los mecanismos de marcado emocional que ofrece la herramienta EmoTraductor 1.0. Esta herramienta es una buena primera aproximación a un marcador de emociones en texto, aunque tiene algunas limitaciones como un diccionario emocional pequeño o un procesamiento muy simple del texto que deja fuera aspectos importantes del lenguaje que influyen de manera directa en la carga emocional como la negación.

Para poder cumplir con nuestro objetivo principal, mejorar el marcado emocional de textos, estudiaremos el funcionamiento del EmoTraductor 1.0 detectando los puntos débiles del marcado e intentando mejorarlos, y después ampliaremos el sistema para añadir algunas mejoras.

Con esto en mente pasamos a enumerar nuestros objetivos específicos:

- Solucionar los malos funcionamientos encontrados tras la evaluación del EmoTraductor 1.0.
- Aumentar el diccionario emocional para ampliar la cobertura del EmoTraductor.
- Añadir el tratamiento de los modificadores que pueden alterar la carga emocional de las palabras a las que afectan.
- Añadir el tratamiento de la negación en los textos.
- Evaluar todas las mejoras introducidas para decidir cuáles de ellas son finalmente incorporadas a la nueva versión de nuestra herramienta.

Todo el código y pruebas realizadas se encuentran en el siguiente repositorio público de GitHub:

<https://github.com/NILGroup/NIL-API-Emociones/tree/Paloma>

## 1.3. Estructura de la memoria

La estructura de este documento consta de los siguientes capítulos:

- En el capítulo 2 se hace un repaso del estado del arte sobre el tema a tratar partiendo de una primera visión sobre qué es la computación afectiva, la representación de las emociones y el marcado emocional de textos. Después se exponen algunos de los diccionarios emocionales más relevantes para nuestro trabajo y se explica la importancia de los emoticonos, los modificadores y las negaciones en el marcado emocional de textos. Por último, se examinan algunas de las herramientas de marcado emocional existentes.
- En el capítulo 3 se presenta la aplicación web EmoTraductor 1.0, junto con sus servicios web, desarrollada anteriormente a este trabajo. Será nuestro punto de partida y se estudiará para mejorar su funcionamiento.
- En el capítulo 4 aparecen los resultados de un primer análisis de la aplicación web EmoTraductor 1.0 y se explican una serie de mejoras que se realizan y evalúan sobre la aplicación. Estas mejoras van desde cambios en el diccionario emocional que usa de base hasta el cambio completo de un servicio web.
- En el capítulo 5 se explica cómo se van a tratar los modificadores en el marcado emocional de textos que utiliza el EmoTraductor, y se evalúan los resultados obtenidos.
- En el capítulo 6 se explica cómo se van a tratar las negaciones en el marcado emocional de textos que utiliza el EmoTraductor, incluyendo la evaluación del impacto de las aproximaciones planteadas.
- En el capítulo 7 se discutirá la validez de las nuevas mejoras y su aplicación, así como los cambios necesarios que se han debido hacer en la aplicación.
- En el capítulo 8 se presentan las conclusiones del TFM junto con el trabajo futuro.
- Los capítulos 9 y 10 son los capítulos 1 y 8 traducidos al inglés.



## Estado del arte

**RESUMEN:** En este capítulo se van a tratar los aspectos más importantes tanto de la computación afectiva como de las diferentes tecnologías y metodologías que se van a utilizar en este TFM. En la Sección 2.1 se hace un breve recorrido por la computación afectiva. En la Sección 2.2 veremos distintos métodos de representación de las emociones. En la Sección 2.3, entraremos en profundidad en la detección de emociones en textos. En la Sección 2.4 conoceremos los diccionarios afectivos más relevantes para nuestro trabajo. En las Secciones 2.5, 2.6 y 2.7 veremos la importancia de los emoticonos, modificadores y negaciones en el marcado emocional respectivamente. Finalmente, en la Sección 2.8 veremos algunas herramientas de marcado emocional de textos ya existentes.

### 2.1. Computación Afectiva

La computación afectiva es el estudio y el desarrollo de sistemas capaces de reconocer, procesar, interpretar y/o simular emociones humanas (Baldassarri, 2016). La computación afectiva se ha convertido en una rama de investigación importante dentro de la Inteligencia Artificial debido a que cada día las personas están rodeadas por multitud de emociones, y cualquier sistema inteligente debe poder interpretarlas y reproducirlas. Algunas teorías, tanto antiguas como más modernas, han demostrado que las emociones influyen en muchos aspectos del conocimiento humano; no solo en la creatividad o en la inteligencia sino también en la percepción, el aprendizaje, el procesamiento racional y la toma de decisiones (Baldassarri, 2016; Picard, 1997).

Actualmente la computación afectiva está extendida por distintos campos. En el libro de la doctora Picard (1997), ya se explicaban campos de aplicación en los que esta rama podría ir avanzando. En la actualidad algunos de estos campos de aplicación en los que se está desarrollando la computación afectiva podrían ser (Baldassarri, 2016; Picard, 1997; el Kaliouby et al., 2006; Cuevas, 2015) :

- Seguridad: Muchas veces, la falta de atención o aburrimiento puede provocar en las personas despistes y accidentes. A través de la medición del aburrimiento o el estrés podemos saber qué atención está poniendo una persona en la realización de las tareas y que la máquina que esté usando actúe de una manera o de otra. Por ejemplo, podremos mejorar la seguridad en las carreteras controlando el estrés o el enfado de las personas a través de música o conversaciones cortas con el propio coche.

- **Salud:** Se ha demostrado que existen numerosos tratamientos para ayudar a personas con problemas o trastornos como el autismo o síndrome de Asperger en los que intervienen las emociones. Podríamos mostrar a una persona con este tipo de problemas a diferenciar qué está sintiendo en cada momento ayudado simplemente de un ordenador. También el análisis de emociones o la computación afectiva puede ayudar a personas sin voz, ya que a través de esta rama de investigación podríamos conseguir que una máquina exprese las palabras de una persona poniendo intensidad o haciendo entender las emociones. Otro aspecto relacionado con la salud es la compañía que puede dar un sistema capaz de entender y demostrar las emociones. En Japón esta utilización de la computación afectiva es más habitual y han creado robots y mascotas capaces de hacer compañía a las personas que están solas.
- **Ocio/Videojuegos:** Gracias a la computación afectiva, utilizando algunos de los elementos que se explicarán más adelante, podríamos analizar la experiencia de usuario de los juegos en directo y hacer que el juego continúe de una manera u otra dependiendo de lo que produce en el jugador.
- **Marketing:** Este ámbito se parece mucho al anterior. A través de la computación afectiva seríamos capaces de tener feedback inmediato de las emociones que surgen a los compradores o usuarios al utilizar, probar o comprar un producto. Se podrían, así, mejorar los productos adaptándolos a las necesidades, gustos y emociones de los compradores.
- **Aprendizaje y formación:** Actualmente, gracias al e-learning y al m-learning, ya estamos viviendo la educación a través de la tecnología. Si a este tipo de educación le añadimos la medición de las emociones podemos mejorar y hacer que los alumnos aprendan más rápido (evitando la frustración) y que disfruten aprendiendo. Teniendo en cuenta las emociones que sienten los alumnos, las máquinas podrán actuar en consecuencia.
- **Toma de decisiones:** Como decíamos antes, se ha demostrado que las emociones son necesarias para la toma de decisiones. Algunos estudios demuestran que la toma de decisiones sin tener en cuenta las emociones se ve tan afectada como la toma de decisiones teniendo en cuenta demasiado las emociones. Por lo tanto, si queremos tomar decisiones inteligentes, debemos analizar y conocer las emociones. Lo mismo ocurre con las máquinas; si queremos que sean inteligentes, debemos enseñarles a distinguir las emociones y esto mejorará la comunicación persona-máquina, ayudándonos en el proceso complejo ante el que nos encontramos.

Todos estos campos de aplicación nos llevan a intentar crear computadores o programas capaces de detectar el estado emocional del usuario. Para ello existen diferentes métodos para detectar emociones:

- **Reconocimiento textual:** Se reconocen emociones a través de las palabras utilizadas.
- **Reconocimiento facial:** Se reconocen emociones a través de expresiones con la cara, por ejemplo las sonrisas o los pucheros de un bebé.
- **Reconocimiento vocal:** Se reconocen emociones analizando la modulación y el tono de la voz.

- Reconocimiento mediante señales fisiológicas: Se reconocen emociones por cambios que sufre nuestro organismo ante ciertas emociones por ejemplo una aceleración del pulso o sudor sería indicativo del enfado o miedo.
- Reconocimiento multimodal: Es una mezcla de los reconocimientos anteriores (mezclando visual, auditivo, fisiológico y otros).

Todos los reconocimientos anteriores se basan en distintas tecnologías o herramientas como pueden ser (Baldassarri, 2016; Picard, 1997; Bosquez et al., 2018):

- Cámaras para el reconocimiento facial. Mediante estas capturamos las emociones por propiedades visibles como el color de la piel detectando posibles variables en el color indicando enfado, vergüenza o miedo, gestos o detección facial.
- Eye-trackers para el reconocimiento mediante señales fisiológicas. Esta tecnología mide el punto de fijación de la mirada o el movimiento de los ojos respecto a la cabeza. Con tecnologías como esta podemos capturar emociones por el movimiento y seguimiento de las pupilas.
- Micrófonos para el reconocimiento vocal. Capturamos emociones a través de la entonación, el tono o el volumen de un discurso o conversación.
- Lenguaje para el reconocimiento vocal, textual y multimodal. A través del entendimiento del lenguaje, captaremos las emociones en textos o voz.
- Otros como la respiración, el pulso o la temperatura corporal valen como medios para el reconocimiento mediante señales fisiológicas.

A continuación, en la Sección 2.2, presentamos los métodos de representación de las emociones y en la Sección 2.3, entraremos en profundidad en el reconocimiento de emociones en texto, ya que será este tipo de reconocimiento el que emplearemos en nuestro trabajo.

## 2.2. Representación de las emociones

Una emoción, según la RAE, es la “alteración del ánimo intensa y pasajera, agradable o penosa, que va acompañada de cierta conmoción somática”, es decir, es un alteración que surge en las personas como una reacción subjetiva ante la situación que vive, y que puede ir acompañada de cambios fisiológicos como puede ser sudoración o aceleramiento del ritmo cardíaco.

Actualmente existen varios métodos para representar emociones, esto se debe a que existe un problema para decidir o crear un estándar de cómo representarlas por el grado de subjetividad que supone una emoción. Los principales métodos usados para representar emociones son (Cowie y Cornelius, 2003):

- Categorías emocionales. Consiste en representar las emociones usando palabras emocionales o etiquetas afectivas como por ejemplo: alegría o tristeza.
- Dimensiones emocionales. Se usan los aspectos esenciales de las emociones (dimensiones) para representarlas (Wundt, 1913). Se suelen utilizar tres dimensiones para representar las emociones (Lang et al., 1997):

1. Evaluación o valencia: La evaluación puede ser positiva o negativa. Por ejemplo, la alegría tendría una evaluación positiva mientras que el enfado tiene una evaluación negativa.
  2. Activación: Indica cómo de activa es la emoción. Por ejemplo, la sorpresa o la alegría son emociones muy activas, con mucha intensidad, mientras que la tristeza es menos activa.
  3. Control: Representa el control de la emoción sobre la persona. Por ejemplo, el enfado o la ira tendrían un control mayor que el de la tristeza.
- Modelos circunflejos: Russell (1980), dice que las dimensiones emocionales están interrelacionadas y se puede crear un modelo espacial en círculo con un orden donde el eje X representa la evaluación (positivo-negativo) y el eje Y la activación (activado-desactivado). En la Figura 2.1 se puede ver el modelo circunflejo de Russell donde están colocadas algunas emociones. En dicha figura vemos que, por ejemplo, la palabra “tenso” está en la zona superior izquierda indicando que es una emoción negativa y activa.



Figura 2.1: Clasificación de 16 palabras emocionales en el modelo circunflejo de Russell

### 2.3. Marcado emocional en textos

El marcado emocional de textos consiste principalmente en etiquetar o marcar distintas palabras, frases o textos con sus emociones asociadas. Existen distintos métodos de marcado emocional (Francisco, 2008; Francisco y Gervás, 2006):

- *Keyword spotting* o marcado de palabras clave. Consiste en clasificar un texto conforme a las palabras emocionales que contiene el mismo. Por ejemplo, la frase “La Cenicienta está feliz” se marcaría como alegre por la aparición de la palabra “feliz”, que tiene asociada la emoción alegre. El principal problema de este método es que no tiene en cuenta los distintos modificadores que pueden aparecer como, por ejemplo, la negación, u otros fenómenos del lenguaje que pueden modificar la carga emocional del texto. En el caso de la frase “La Cenicienta no está feliz”, este método seguiría encontrando la palabra «feliz» y marcaría el texto erróneamente como alegre.
- Afinidad léxica. No solo detecta las palabras claramente emocionales como “feliz”, “triste”, “contento” o “enfadado”, sino otras palabras que denotan emoción. Para ello, asocia a las palabras una probabilidad de representar una emoción u otra. Por ejemplo, la palabra “accidente” podrá tener asociada en un 75 % la emoción tristeza y en un 25 % la emoción alegre porque podríamos encontrarnos las frases “Murió en un accidente de tráfico”, donde claramente la palabra es triste, o podríamos ver la frase “Nos enamoramos por accidente” donde accidente es un acontecimiento bonito o alegre. Para asignar las probabilidades a las distintas palabras será necesario contar con un corpus de textos con emoción asociada. Al igual que el método anterior, no tiene en cuenta el contexto de la frase, y, por ello, ante la aparición de, por ejemplo, una negación, los resultados no serían correctos.
- Procesamiento estadístico del lenguaje natural. Consiste en entrenar un algoritmo de aprendizaje automático con un corpus de textos emocionales marcados. Este método solo funciona con textos extensos, y queda muy restringido al tipo de textos y vocabulario con el que ha sido entrenado el algoritmo.
- Métodos basados en el conocimiento del mundo real. Emplea el conocimiento obtenido de las actitudes emocionales de las personas en distintas situaciones, objetos y acciones, para marcar los textos. Miran la calidad afectiva de la semántica que contiene un texto y por ello no necesitan que los textos tengan palabras claves emocionales. Para estos métodos es necesario tener un amplio conocimiento de las actitudes emocionales de las personas.

Estos métodos de marcado necesitan un fenómeno lingüístico sobre el que trabajar (texto, frase o palabra). Nosotros trabajaremos sobre la palabra. Según la RAE, una palabra es una “unidad lingüística, dotada generalmente de significado, que se separa de las demás mediante pausas potenciales en la pronunciación y blancos en la escritura”. Basándonos en esta definición, podríamos decir que en caso de que una palabra esté dotada de significado llevará asociada una connotación emocional. Usando esta unidad mínima, la palabra, podríamos crear infinitas frases distintas. Así podemos considerar, la frase como una unidad lingüística con sentido completo. Podríamos decir entonces que la carga emocional de una frase (o el conjunto de estas, el texto) es el efecto combinado de palabras relevantes que se encuentran en ella en función de la clasificación emocional de estas palabras (Taboada et al., 2011).

A pesar de utilizar la palabra como base, porque ya es una unidad con significado y con carga emocional, nos encontramos con algunos textos o frases donde algunas palabras no aportan carga emocional al conjunto. Para decidir qué palabras sí aportan o podrían aportar emociones a una frase o texto debemos analizar las palabras. Por ejemplo en la frase “Vamos a ver la casa de mi querida abuelita” vemos claramente que las palabras “a”, “la”, “mi” y “de” no aportan significado ni emociones a la frase y que sirven únicamente como nexos. En la frase “Voy a por el regalo de cumpleaños para la fiesta y ya”, volvemos

a encontrar palabras que no aportan emociones “a”, “por”, “el”, “para”, “la”, “y” y “ya” que no sabríamos indicar si aportan carga emocional o no.

En algunos trabajos (Taboada et al., 2011; Polanyi y Zaenen, 2006), vemos que se utiliza el tipo de palabra, morfológicamente hablando, para reconocer qué palabras podrían tener carga emocional o no, encontrando como categorías emocionales válidas los adjetivos, adverbios, sustantivos y verbos.

En la frase “Cada fracaso suponía un nuevo golpe” vemos que hay ciertas palabras que nos transmiten la emoción de tristeza, enfado o miedo como pueden ser las palabras “fracaso” y “golpe”, mientras que encontramos otras como “cada” o “un” que no aportan ningún valor al contenido emocional de la frase. En la frase “Ayer estaba muy feliz”, encontramos palabras como “feliz” que nos transmiten la emoción de alegría y palabras como “ayer” que no nos aportan contenido emocional. Pero en este ejemplo podemos ver también el modificador “muy” que aunque no aporta ninguna carga emocional si la evaluamos de manera aislada, en el contexto de la frase vemos que intensifica las emociones asociadas a la palabra “feliz”. En la frase “Los fantasmas no me dan miedo”, encontramos palabras como “fantasmas” y “miedo” que nos transmiten emociones y palabras que no nos aportan nuevo contenido emocional, no son palabras emocionales, como “los” y “me” y la palabra “no” que aunque de manera aislada no aporta carga emocional, vemos que modifica la intensidad de la emoción que nos transmite la palabra “miedo” llegando incluso a hacer desaparecer la emoción. En las Secciones 2.6 y 2.7 mostraremos como se pueden gestionar los modificadores y la negación a la hora de marcar emocionalmente un texto.

Teniendo en cuenta que la palabra es la unidad mínima de carga emocional y que a través de ella iremos descubriendo las emociones de frases y textos, necesitamos un método para saber qué connotación emocional posee cada palabra. Con este fin se han creado distintos diccionarios afectivos, presentaremos los más relevantes para nuestro trabajo en la Sección 2.4.

Si aumentamos la unidad de lenguaje que evaluamos, llegamos finalmente a los textos. Un texto está compuesto por frases y estas pueden ser de distintos tipos: enunciativas, exclamativas o interrogativas. El tipo de una frase también pueden aportar carga emocional al texto, ya que la emoción asociada a una frase con las mismas palabras pero con distinto tipo puede ser muy diferente (por ejemplo “Te odio”, “¡Te odio!” o “¿Te odio?”). En el ejemplo anterior, “Te odio” transmite enfado y tristeza; “¡Te odio!” también nos transmite enfado y tristeza pero con mayor intensidad y, por último, “¿Te odio?” nos transmite menos enfado y tristeza que las otras, parece que el hablante no sabe del todo lo que acaba de sentir, lo que hace que disminuya la intensidad de la emoción.

Esta es solo una clasificación lingüística de las frases en textos, pero existen otras clasificaciones sintácticas que también será importante tener en cuenta a la hora de determinar la carga afectiva del texto. Las oraciones coordinadas son un ejemplo: dependiendo del tipo de coordinada aumentará o disminuirá la intensidad de las emociones. Por ejemplo, si decimos “Te quiero pero te quiero poco”, a pesar de que la frase parezca muy alegre porque queremos a alguien el “pero” resta importancia a la oración que viene a continuación y limita la cantidad de alegría que puede transmitir la frase.

Más allá de la estructura lingüística de los textos y frases, por la evolución del lenguaje, aparecen los emoticonos, que aportan nueva carga emocional a la frase o texto donde están incluidos. Si evaluamos por separado los emoticonos como palabras, vemos que cada uno posee una carga emocional diferente, ya que su intención es, precisamente, mostrar la emoción que el hablante está sintiendo en el momento de hablar o escribir. Por ejemplo, el emoticono “:O” nos indica sorpresa, el emoticono “:)” nos va a denotar alegría, o el emoticono “:(” nos transmite tristeza. Si los vemos en distintas frases, observamos como

cambia también la carga emocional de las mismas. Por ejemplo, en la frase “Estoy feliz :)” vemos una mayor carga de alegría que en la frase “Estoy feliz”. La frase “No me gusta que hagas eso” podría denotar enfado y tristeza, mientras que si la frase es “No me gusta que hagas eso :(” percibimos una mayor carga en cuanto a la tristeza y en menor medida el enfado. En la Sección 2.6 mostraremos cómo se pueden tratar los emoticonos desde el punto de vista de la detección emocional.

## 2.4. Diccionarios afectivos

Un diccionario afectivo es un conjunto de palabras que tienen asociadas emociones. Su objetivo es servir de base para el marcado emocional de textos. En esta Sección presentaremos los diccionarios afectivos más relevantes para nuestro trabajo.

### 2.4.1. ANEW

El Affective Norms for English Words (Bradley y Lang, 1999), por sus siglas ANEW, es un diccionario que posee 1.034 palabras inglesas y que utiliza las dimensiones emocionales (evaluación, activación, control) para etiquetar emocionalmente cada palabra. El objetivo de este diccionario, al igual que el de muchos otros, es crear material estándar para los estudios sobre las emociones. El diccionario se creó mediante una serie de experimentos que fueron realizados por alumnos de una clase introductoria a la Psicología. Cada experimento presentaba entre 100 y 150 palabras diferentes y para cada conjunto de palabras se preparó una hoja de papel con las palabras individuales en 4 columnas y 14 filas (56 palabras por hoja). Los sujetos debían calificar en estas hojas las dimensiones de evaluación, activación y control en una escala Likert del 0 al 9 con las palabras fuera de contexto.

En la Tabla 2.1 podemos ver la primera palabra del diccionario (“*abduction*”) y la última palabra de este (“*zest*”) con el número de palabra que ocupa en el diccionario y los valores asociados a cada una de las dimensiones emocionales. La palabra “*abduction*” tiene una connotación negativa, con una activación media y un control bajo. La palabra “*zest*”, sin embargo, tiene una connotación positiva, una activación media y un control mayor que “*abduction*”.

Descripción	Nº palabra	Evaluación	Activación	Control
abduction	621	2,76	5,53	3,49
zest	1040	6,79	5,59	6,00

Tabla 2.1: Primera y última palabra del diccionario ANEW y sus campos

Este diccionario es uno de los más importantes y por ello ha sido traducido a varios idiomas, entre ellos español (S-ANEW).

### 2.4.2. S-ANEW

S-ANEW (Redondo et al., 2007) es una traducción del diccionario ANEW al castellano. Para crear el diccionario se tradujeron las 1034 palabras del diccionario ANEW (Bradley y Lang, 1999) y gracias a estudiantes de Psicología fueron marcadas en las mismas dimensiones emocionales que en ANEW (evaluación, activación y control) usando la misma metodología (Bradley y Lang, 1999).

Además de la palabra con sus dimensiones emocionales, el diccionario contiene los siguientes campos por cada palabra:

- Número de palabra: es una numeración que coincide con la de cada palabra en el diccionario ANEW para relacionar las palabras de ANEW con S-ANEW, ya que el orden no coincidirá pues en ambos diccionarios las palabras se encuentran en orden alfabético.
- E-word: es la palabra inglesa relacionada.
- S-word: es la palabra en castellano.
- Evaluaciones de las palabras: se incluyen los valores medios y la desviación estándar para la tres dimensiones emocionales. Se muestran los datos de la muestra global seguidos de los datos correspondientes a las mujeres y los hombres.
- Índices psicolinguísticos: son índices para marcar algunas características que pueden resultar útiles de los términos evaluados. Existen índices objetivos y subjetivos. Los objetivos son el número de letras, el número de sílabas, la clase gramatical, la frecuencia de aparición y el número de palabras vecinas (estas son las palabras que son iguales excepto por una letra). Los subjetivos son la familiaridad (cuánto de conocida es la palabra para el evaluador), la concreción (cuánto de concreto es el concepto de la palabra evaluada) y la imaginabilidad (cuánto de abstracto o imaginario es el concepto que representa la palabra).

En la Tabla 2.2 podemos ver también la primera palabra del diccionario ANEW “*abduction*” y la última “*zest*” en su versión española. Se puede observar que los resultados para las mismas palabras en los diccionarios no coinciden, dado que en el S-ANEW los términos fueron re-evaluados una vez traducidos por hispanohablantes.

Campos del diccionario	Palabras	
	rapto	brío
S-word	rapto	brío
E-word	abduction	zest
Nº palabra	621	1040
Evaluación	1,67	5,78
Desviación normal evaluación	1,37	2,14
Activación	7,00	5,77
Desviación normal activación	1,97	2,24
Control	2,36	5,32
Desviación normal control	1,69	2,08
Nº Letras	5	4
Nº Sílabas	2	2
Clase gramatical	Nombre	Nombre
Frecuencia de aparición	4,29	4,46
Palabras vecinas	0	1
Familiaridad	4,72	No disponible
Concreción	3,7	No disponible
Imaginabilidad	4,98	No disponible

Tabla 2.2: Primera y última palabra del SANEW y sus campos

### 2.4.3. WordNet-Affect

WordNet-Affect (Strapparava y Valitutti, 2004) es un recurso lingüístico desarrollado a partir de WordNet (Miller et al., 1990) y WordNet Domains (Magnini y Cavaglià, 2000). WordNet es un diccionario online inglés donde hay más de 118.000 palabras distintas con todos sus conceptos o significados en inglés. WordNet Domains es el mismo diccionario añadiendo los llamados Subject Field Codes (SFC), que son campos que indican en qué dominios es relevante el concepto (por ejemplo Medicina, Arquitectura, Deporte o Política). Hay que tener en cuenta que una misma palabra puede tener varios conceptos asociados, como por ejemplo la palabra “operate” que puede significar cosas distintas dependiendo del ámbito, si es Medicina o Matemáticas. Partiendo por tanto de WordNet Domains, WordNet-Affect añade lo que podríamos llamar «etiquetas de dominio afectivo». Estas etiquetas traducidas al español son: *estados de ánimo*, *situaciones que provocan emociones* o *respuestas emocionales*. Dentro de cada una de estas etiquetas, podemos encontrar otras más específicas que están vinculadas con rasgos de personalidad, comportamientos, actitudes, estados físicos y corporales y sentimientos. Para obtener estas etiquetas se mapearon los términos escogidos para el WordNet Affect con los synsets que se encuentran en WordNet. Se eliminaron todos los synsets de las palabras que no reconocían a estas como afectivas. Después, se verificó la coherencia de información afectiva de los synsets fijándose sobretodo en si había información de sinónimos con valores incompatibles y para etiquetar manualmente aquellas entradas que no se hubieran etiquetado automáticamente. Más tarde se añadieron etiquetas nuevas como *positive*, *negative*, *ambiguous* y *neutral*. WordNet-Affect contiene 4.787 palabras. Este diccionario se ha ido traduciendo a distintos idiomas en distintos proyectos.

### 2.4.4. Diccionario de Díaz et al.

Existe un diccionario en español más reciente (Díaz et al., 2014) basado en WordNet-Affect (Strapparava y Valitutti, 2004) que propone un diccionario donde las palabras se etiquetan mediante categorías emocionales. Usa como etiquetas emocionales las emociones básicas de Ekman (1972): *«alegría»*, *«enfado»*, *«miedo»*, *«repulsión»*, *«sorpresa»* y *«tristeza»*.

Para crear este diccionario se toman las palabras de WordNet-Affect (Strapparava y Valitutti, 2004), en inglés, categorizadas con las etiquetas de “alegría”, “repulsión”, “enfado”, “miedo”, “sorpresa” y “tristeza”, obteniendo 1.110 palabras. El siguiente paso fue traducir al castellano las 1.110 palabras utilizando tres traductores automáticos distintos. Cada palabra fue por lo tanto traducida tres veces, una vez por cada herramienta, y junto con cada traducción se guardó la categoría emocional asociada a la palabra inglesa. Después, se unieron todas las traducciones de las palabras y eliminando las palabras castellanas repetidas, quedaron 3.591 palabras en el diccionario. Para garantizar que los resultados eran coherentes y que cumplían con la calidad requerida se revisó manualmente el diccionario y finalmente quedaron en el diccionario 2.036 palabras.

Definidas las palabras que el diccionario alberga se evaluó por voluntarios si cada palabra tenía relación con la emoción con la que fueron asociadas (las que tenían asociadas en el WordNet-Affect (Strapparava y Valitutti, 2004)). En esta evaluación se pidió a los evaluadores que hicieran una estimación de con qué frecuencia la palabra se utilizaba para denotar la etiqueta emocional que tenían asociada. Para ello utilizaron las opciones: “nula”, “baja”, “media” y “alta”, donde “nula” indicaba que en ningún contexto la palabra denotaba la etiqueta asociada, “baja” en algunos contextos la palabra denotaba algo relacionado con la emoción de la etiqueta, “medio” indicaba que se usa con igual frecuencia la palabra

denotando la emoción de la etiqueta que sin denotar la emoción y “alta” indica que la palabra casi siempre o siempre denota la emoción asociada. Para facilitar a los evaluadores la decisión se incluyó la definición de cada palabra. La finalidad de la evaluación fue saber hasta qué punto un término se puede asociar con una emoción en particular. Por lo tanto, este diccionario nos indica la estimación de la frecuencia con que una palabra se usa con un sentido que lo relaciona con una emoción. Cada palabra por tanto tendrá asociada no sólo una emoción sino que además tendrá la probabilidad de que se esté usando la emoción o no. En la Figura 2.2, vemos un fragmento de los valores de probabilidad de uso afectivo para la categoría “alegría”. Por ejemplo, en la palabra “abundancia” el 50 % de los evaluadores eligieron la opción “media” lo que indica que esta palabra se utiliza en un contexto que denota alegría casi con igual frecuencia que cuando se usa en un contexto donde no se denota emoción y además el otro 50 % eligió el valor “alta” por lo que desde el punto de vista de los evaluadores la palabra “abundancia” siempre o casi siempre denota alegría.

PROBABILIDAD DE USO AFECTIVO					
#	PALABRA	NULA [%]	BAJA [%]	MEDIA [%]	ALTA [%]
1	<i>abundancia</i>	0	0	50	50
2	<i>acabalar</i>	40	0	60	0
3	<i>acallar</i>	50	40	10	0
4	<i>acatar</i>	50	40	10	0
5	<i>acción</i>	30	30	30	10
6	<i>aceptable</i>	0	20	80	0
7	<i>aceptación</i>	0	20	50	30
8	<i>acicate</i>	10	50	40	0
9	<i>aclamación</i>	10	10	10	70

Figura 2.2: Porcentajes de probabilidad de uso afectivo en palabras del diccionario “alegría” según Díaz et al. (2014)

Por tanto, lo más relevante de este diccionario es que etiqueta las palabras en una categoría emocional pero también incluye la probabilidad de uso afectivo para cada una. Es decir, indica la frecuencia con que se usan con un sentido emocional pudiendo así considerar diferentes contextos de aplicación de la palabra. Este punto es muy interesante para el estudio de una emoción en un texto o una oración.

#### 2.4.5. Corpus de Ferré

El corpus de Ferré et al. (2012), contiene 380 palabras en castellano pertenecientes a tres categorías semánticas diferentes (animales, personas y objetos) con valores para la valencia y la activación. También evalúa otros aspectos interesantes como la concreción, que indica cuanto de concreta o abstracta es una palabra, y la familiaridad que indica cómo de familiar resulta una palabra. Por ejemplo, la palabra “abeja” tiene un valor de un 8,52 en concreción sobre 9, lo que indica que es una palabra que representa algo concreto y de familiaridad tiene un 6,32 sobre 9, lo que indica que los evaluadores consideraban la palabra bastante familiar.

Para asignar los distintos valores a las palabras se contó con 504 estudiantes de Psicología, Ciencias de la Comunicación y Educación de las universidades Rovira y Virgil en Tarragona, Universidad Autónoma de Madrid y Universidad de Santiago de Compostela. Se hicieron dos cuestionarios diferentes, uno donde se clasificaban la valencia y activación y otro donde se evaluaban la familiaridad y la concreción. El primer cuestionario fue realizado

por 304 participantes, mientras que el segundo solo por 200. Las palabras fueron marcadas en cada aspecto (valencia, activación, concreción y familiaridad) usando una escala Likert del 1 al 9. En base a los resultados obtenidos, decidieron clasificar las palabras en negativas, positivas y neutras, siendo las que tienen valencia entre 1 y 4 palabras negativas, las que se encuentran entre 4 y 6 neutras y las que tienen valencia entre el 6 y 9 positivas. Conforme a esta clasificación el corpus contiene 113 palabras negativas, 194 neutras y 73 positivas. En base a los resultados obtenidos, decidieron dividir el conjunto en dos, palabras con baja activación (220 palabras) y palabras con alta activación (160 palabras). Para esta división se utiliza el valor 5 como punto de división.

Como resultado de las marcaciones realizadas por los estudiantes se generó un corpus con las calificaciones de valencia, activación, concreción y familiaridad para 380 palabras. El diccionario también contiene los siguientes índices psicolingüísticos: frecuencia de aparición y número de letras. En el corpus aparecen primero los nombres de la categoría semántica «animales», seguidos de las que se refieren a «personas» y la última categoría es «objetos».

En la Tabla 2.3, podemos ver algunas palabras del corpus centrándonos en las columnas de activación y valencia, donde podemos ver que las palabras “Delfín” e “Hijo” son consideradas positivas, mientras que las palabras “Chulo” y “Abeja” están marcadas como negativas y la palabra “Alfombra” es considerada neutra en este corpus. También vemos que son consideradas palabras con alta activación “Delfín”, “Chulo”, “Hijo” y “Abeja” mientras que “Alfombra” está considerada como palabra con baja activación.

Categoría	Palabra	Valencia	Clasificación por valencia	Activación	Clasificación por activación
Animal	Delfín	7,74	Positiva	5,21	Alta
People	Chulo	2,89	Negativa	5,95	Alta
Object	Alfombra	5,03	Neutral	2,92	Baja
People	Hijo	7,22	Positiva	6,16	Alta
Animal	Abeja	3,34	Negativa	6,79	Alta

Tabla 2.3: Extracto de cinco palabras del Corpus de Ferré.

#### 2.4.6. MADS (Madrid Affective Database for Spanish)

MADS (Hinojosa et al., 2016) contiene 875 términos sacados de LEXESP (Sebastián-Gallés et al., 2000) y de trabajos previos del mismo grupo de investigadores y tiene como fin complementar a diccionarios importantes ya existentes como el S-ANEW (Redondo et al., 2007) o el Corpus de Ferré (Ferré et al., 2012), y es por esto que no incluye palabras que existan en alguno de los dos, exceptuando 16 palabras de ANEW y 11 del Corpus de Ferré que servirán como control de las evaluaciones. Para crear este diccionario se marcaron las 875 palabras con cinco categorías emocionales (*alegría, enfado, tristeza, miedo y asco*) y con dos dimensiones emocionales (*valencia y activación*). Para el marcado participaron 660 personas cuya lengua materna era el español y cada palabra fue marcada por 30 personas. Las emociones básicas se marcaron usando una escala del 1 al 5 (siendo 1 la ausencia de la emoción y el 5 la presencia plena de la emoción). Cabe destacar que una palabra denota una determinada emoción si tiene un valor mayor al 2,5 para dicha emoción. Para evaluar la valencia, los participantes eligieron para cada término un número del 1 al 9 (siendo 1 muy negativo y 9 muy positivo). Una vez marcadas las palabras se clasificaron como

positivas, negativas o neutrales según la valencia de cada palabra: del 1 a 4 son palabras negativas, de 4 a 6 neutrales y de 6 a 9 positivas. La activación se marca usando una escala igual a la de la valencia y los evaluadores eligieron un número del 1 al 9 (siendo 1 muy poca activación y 9 mucha activación).

El diccionario, además de las etiquetas ya vistas, contiene también: la palabra en inglés, la frecuencia de aparición de la palabra, el número de letras y sílabas y una serie de variables psicolingüísticas similares a las expuestas en el diccionario S-ANEW (Redondo et al., 2007).

En la Tabla 2.4 vemos dos palabras de las 11 que tiene este diccionario en común con el Corpus de Ferré (Ferré et al., 2012) para demostrar que los resultados no deben ser los mismos, y añadiremos la primera palabra y la última del diccionario. Vemos que la palabra “Delfín” en este diccionario es más o menos semejante en cuanto a valencia y activación respecto al Corpus de Ferré; mientras que la palabra “Chulo” en este diccionario es más positiva que en el Corpus de Ferré. Gracias a las categorías emocionales que añade el diccionario de Hinojosa, podemos decir que la palabra “Delfín”, siguiendo sus valores en cada una de las categorías, denota alegría; la palabra “Chulo” enfado, la palabra “Abandono” denota enfado, tristeza, miedo y asco y la palabra “Zopenco” no denota ninguna emoción.

Palabra	Palabra en inglés	Valencia	Activación	Alegría	Enfado	Tristeza	Miedo	Asco
Delfín	dolphin	7,33	5,17	3,90	1,07	1,27	1,50	1,20
Chulo	pimp	3,73	5,00	2,17	2,62	1,76	1,52	2,62
Abandono	abandonment	1,47	6,47	1,03	3,30	4,43	3,33	2,63
Zopenco	stupid	3,33	5,07	1,43	1,83	2,37	1,57	1,50

Tabla 2.4: Extracto del diccionario de Hinojosa

#### 2.4.7. Diccionario de Ferré

Este diccionario (Ferré et al., 2017) está compuesto por 2.266 palabras en castellano marcadas con cinco categorías emocionales: *alegría*, *enfado*, *miedo*, *asco* y *tristeza*. Las palabras con las que cuenta son las incluidas en los diccionarios S-ANEW (Redondo et al., 2007), el Corpus de Ferré (Ferré et al., 2012) y el de Guasch et al. (2016). De las 2.814 palabras obtenidas se eliminan las repetidas y las que apareciesen en el diccionario de Hinojosa et al. (2016) obteniendo las 2.266 que finalmente incluye el diccionario. El proceso de marcado consistió en que 1.380 personas asignaron un valor del 1 a 5 para cada categoría emocional, siendo 1 la ausencia de la emoción y el 5 la máxima aparición de la emoción en la palabra. Las palabras con un valor inferior al 2,5 se consideran neutrales mientras que las que tienen un valor superior al 2,5 en más de una categoría pertenecerán a aquella con mayor valor.

En la Tabla 2.5 podemos ver las palabras que hemos visto en el Corpus de Ferré pero que no están en el diccionario de Hinojosa dado que ninguna palabra del diccionario de Hinojosa se encuentran en este diccionario. No podemos compararlas con los valores del Corpus de Ferré porque no se utiliza el mismo método de marcado para las palabras, pero podemos ver que la única palabra que denota emoción es “Abeja” denotando miedo.

#### 2.4.8. Diccionario de Hans Stadthagen-González

Este diccionario (Stadthagen-González et al., 2018) contiene 14.031 palabras en castellano obtenidas de trabajos anteriores del autor (Stadthagen-González et al., 2017), con

Palabra	Palabra en inglés	Alegría	Enfado	Tristeza	Miedo	Asco
Abeja	Bee	1,47	1,97	1,47	3,53	2,07
Alfombra	Carpet	1,60	1,20	1,30	1,13	2,03
Hijo	Son	4,43	1,23	1,43	2,00	1,33

Tabla 2.5: Fragmento del diccionario de Ferré

valores para las categorías emocionales alegría, tristeza, miedo, enfado y asco. Para evaluar las 14.031 palabras con sus categorías emocionales, se involucró a 2.010 españoles nativos, la mayoría de ellos estudiantes de distintas universidades de España. El conjunto de palabras se dividió en diferentes bloques, cada bloque contenía aproximadamente 253 palabras, y cada uno de ellos fue calificado por 20 participantes. La evaluación se hizo como en otros diccionarios, utilizando una escala Likert para cada emoción del 1 al 5, donde el 1 indica que no contiene la emoción calificada y el 5 significa que la palabra contiene completamente la emoción clasificada. En este caso, también se proporcionó a los evaluadores una opción para indicar que desconocían la palabra y por tanto no podían evaluarla. Después de analizar los resultados se decidió eliminar algunos de ellos por falta de información (evaluaciones incompletas) o por repetición de la misma clasificación en más del 95 % de las distintas palabras, es decir que para el 95 % de las distintas palabras clasificadas por el mismo evaluador se puso la misma evaluación sin tener en cuenta la palabra (se eliminaron un 4,2 % de los bloques recopilados), hubo 3.940 palabras que se tuvieron que descartar del diccionario porque más del 45 % de los evaluadores habían marcado la opción que indicaba que no conocían la palabra y por tanto no valía la calificación hecha. Para añadir la valencia y activación se utilizó el diccionario original de donde se sacaron las palabras (Stadthagen-González et al., 2017) que toma valores para ambas entre 1 y 9, siendo en caso de la valencia 1 negativo o infeliz y el 9 positivo o feliz y en caso de la activación siendo el 1 pasivo o tranquilo y el 9 activo o excitado. Después, utilizando EsPal (Duchon et al., 2013), que es un repositorio que contiene distinta información sobre la frecuencia, la ortografía, la fonología e información subjetiva (emocional) de palabras en español; añaden la dominancia de las palabras. Se añade la dominancia todas las palabras excepto a 95 que se hace manualmente porque no existen en la plataforma. Como resultado se obtiene un diccionario con 10.491 palabras en castellano, cada una con los siguientes campos:

- Word: muestra la palabra en español.
- Few\_Raters: tiene una «X» en caso de que la palabra tenga palabras con menos de diez calificaciones para al menos una emoción.
- Happiness\_Mean: es el valor promedio para todas las respuestas válidas en la categoría alegría.
- Happiness\_SD: es la desviación estándar de las respuestas válidas en la categoría alegría.
- Happiness\_%Raters: es el porcentaje de participantes (de un total de 20) que conocían y calificaron la palabra en la categoría alegría.
- Disgust\_Mean: es el valor promedio para todas las respuestas válidas en la categoría asco.
- Disgust\_SD: es la desviación estándar de las respuestas válidas en la categoría asco.

- Disgust\_ %Raters: es el porcentaje de participantes (de un total de 20) que conocían y calificaron la palabra en la categoría asco.
- Anger\_Mean: es el valor promedio para todas las respuestas válidas en la categoría enfado.
- Anger\_SD: es la desviación estándar de las respuestas válidas en la categoría enfado.
- Anger\_ %Raters: es el porcentaje de participantes (de un total de 20) que conocían y calificaron la palabra en la categoría enfado.
- Fear\_Mean: es el valor promedio para todas las respuestas válidas en la categoría miedo.
- Fear\_SD: es la desviación estándar de las respuestas válidas en la categoría miedo.
- Fear\_ %Raters: es el porcentaje de participantes (de un total de 20) que conocían y calificaron la palabra en la categoría miedo.
- Sadness\_Mean: es el valor promedio para todas las respuestas válidas en la categoría tristeza.
- Sadness\_SD: es la desviación estándar de las respuestas válidas en la categoría tristeza.
- Sadness\_ %Raters: es el porcentaje de participantes (de un total de 20) que conocían y calificaron la palabra en la categoría tristeza.
- Valence\_Mean: valor promedio de las evaluaciones para el aspecto de valencia en el diccionario base.
- Valence\_SD: valor de la desviación estándar del diccionario base para el aspecto de valencia.
- Valence\_ %Raters: porcentaje de evaluadores que conocían y calificaron la palabra para el aspecto de valencia en el diccionario.
- Arousal\_Mean: valor promedio de las calificaciones de los evaluadores para el aspecto de activación en el diccionario base.
- Arousal\_SD: desviación estándar del diccionario base para el aspecto de activación.
- Arousal\_ %Raters: porcentaje de evaluadores que conocían y calificaron la palabra para el aspecto de activación en el diccionario.
- Dominant\_POS: valor de dominancia sacado de EsPal.
- %Dominant\_POS: porcentaje de la dominancia obtenido de EsPal.

## 2.5. El papel de los emoticonos en el mercado emocional de textos

Un emoticono según la RAE es la «representación de una expresión facial que se utiliza en mensajes electrónicos para aludir al estado de ánimo del remitente», por lo que no es otra cosa que una secuencia de caracteres ASCII que representa una cara humana y expresa

una emoción. Los orígenes de los emoticonos son diversos, pero parece que su creación surge por la necesidad de los humanos de unir palabras y emociones en sus interacciones (Cuadrado et al., 2014). Como hemos visto en la Sección 2.3 es importante tenerlos en cuenta a la hora de analizar emocionalmente los textos ya que pueden ayudar a interpretar el contenido emocional del mismo. A continuación, vamos a presentar algunos trabajos que han tratado los emoticonos a la hora de marcar emocionalmente un texto.

Existe un estudio sobre el uso de los emoticonos junto con las jergas que mejoran la exactitud de la aplicación PosNeg Opinion (Amores et al., 2015), una aplicación para polarizar opiniones en la web. Amores et al. (2016) recogían un conjunto de emoticonos publicados en una serie de listas<sup>1</sup>, y con todo el conjunto de emoticonos hacen una relación entre el emoticono y la palabra a la que sustituyen. Por ejemplo, el emoticono “=)” según las fuentes significa “Happy face”, por lo que dicho emoticono quedaría asociado a la palabra “Happy” o “Feliz” en español. Una vez han obtenido todas las palabras que representan los emoticonos utilizan recursos léxicos como SentiWordNet (Baccianella et al., 2010) para obtener la polaridad y la emoción de cada una de las palabras asociadas a los emoticonos para asociar a cada uno de ellos la polaridad y las emociones correspondientes, creando así el diccionario. En el estudio también se explica que un emoticono no cumple la misma función en un texto o frase que una palabra y la mayoría de las veces no afectan solo a un término sino a todo el contexto. Es por esto que se considera y defiende que el valor del emoticono debe tenerse en cuenta a la hora de tomar la decisión final de la polaridad del texto añadiendo valor y no tenerlo en cuenta como un término.

Otro estudio realizado sobre los estados de Facebook y algunos de sus comentarios (Vashisht y Thakur, 2014) también clasifica algunos de los emoticonos más utilizados según su polaridad. En este trabajo se analizan 1.250 estados y 2.050 comentarios de Facebook y se creó un registro con todos los símbolos tipográficos que se utilizaban para expresar cada emoción en particular, generando el resultado que se muestra en la Figura 2.3 donde se relaciona cada emoticono con su emoción o palabra asociada y su polaridad (positiva o negativa).

También existe un estudio similar al anterior (Wegrzyn-Wolska et al., 2016) basado en Twitter en lugar de Facebook. En este trabajo tras el análisis de 96.269.892 tweets<sup>2</sup> eligen los 50 emoticonos principales según su porcentaje de aparición y a cada emoticono le asocian una polaridad (negativo o positivo), una emoción específica y un peso que representa de nuevo la polaridad de la palabra. Este léxico se muestra en la Figura 2.4. Después de crear este léxico, en el mismo estudio intentan saber qué peso o importancia tienen los emoticonos en la carga emocional de los distintos tweets, los autores llegan a la conclusión de que algunos emoticonos dominan el sentimiento de un tweet y conquistan la emoción, descartando muchos valores de otros términos, como por ejemplo “Estoy muy feliz :(” donde la “:(” anulará la emoción de alegría que da la palabra “feliz”.

## 2.6. El papel de los modificadores en el mercado emocional de textos

Un modificador es una palabra que modifica la emoción asociada a una palabra o a un conjunto de palabras (Andreevskaia y Bergler, 2007). Algunos ejemplos de modificadores son las palabras “muy”, “poco” o “casi” en las siguientes frases “Estoy muy triste”, “Me gusta poco eso que haces” o “Ella está casi llorando”. Existen distintos trabajos de marcado

<sup>1</sup>[www.netlingo.com](http://www.netlingo.com), [www.noslang.com](http://www.noslang.com), [es.wikipedia.org/wiki/Anexo:Emoticonos](http://es.wikipedia.org/wiki/Anexo:Emoticonos)

<sup>2</sup>El estudio se puede ver en <http://www.datagenetics.com/blog/october52012/index.html>

SNo.	Classification of Emoticons	Typographical symbols used				Sentiment
1	Happiness/ Smile	:-)	=)	:)	:]	Positive
2	Sadness	:(	=(	:[	:(	Negative
3	Wink	;-)	;) )			Positive
4	Teasing/ kidding	;P	;P	;p	;p	Positive
		:P	:P	:p	:p	
5	Amused	:D	:D		=D	Positive
6	Anger	>:(	>:-(			Negative
7	Kiss	:-*	:*			Positive
8	Confused	o.O	O.o			Negative
9	Embarrassed	:-\$	=\$	>///<		Negative
10	Devil	3:)	3:-)			Negative
11	Cool	8-)	B-)	8)	B)	Positive
12	Unsure	:/	:/	\:	:\	Negative
13	Cry	:'(	:'-()			Negative
14	Love/Heart	<3				Positive
15	Shy	^_^				Positive
16	Blessed	O:-)	O:)			Positive
17	Hug	>(^_^)<	<(^_^)>			Positive
18	Squint	-_-				Negative
19	Surprised	:-o	:-O	:o	:-o	Positive
20	Thumbs up	(y)				Positive
21	Thumbs down	(n)				Negative

Figura 2.3: Emoticonos y sentimiento/palabra asociada de Vashisht y Thakur (2014)

emocional que tienen en cuenta los modificadores a la hora de determinar la emoción del texto y que vamos a presentar a continuación.

Andreevskaja y Bergler (2007) tienen en cuenta los modificadores en su sistema de clasificación emocional de titulares de noticias. La razón por la que decidieron tener en

Emoticon	Value	Translation	Weight
:) :D :-) ;) XD :] =) (: ;-) =D =] :-D ^ _ ^ (8 :o) (;=o 8) ;o) (= [ : 8D :]	POSITIVE	happy	1
:o ;O o:	POSITIVE	surprise	1
=P :-P ;P=P	POSITIVE	playful	1
;D ;]	POSITIVE	wink	1
\m/	POSITIVE	salute	1
:( D: =( : ) : [ : ( =[	NEGATIVE	sad	-1
=/ :-/ \ / :-/ =\	NEGATIVE	annoyed	-1
:'(	NEGATIVE	crying	-1
:@	NEGATIVE	angry	-1
:!	NEGATIVE	indifferent	-1

Figura 2.4: Léxico de emoticonos de Wegrzyn-Wolska et al. (2016)

cuenta los modificadores en su análisis es porque los titulares de noticias habitualmente no cuentan con muchas pistas emocionales por la falta de palabras y por tanto cada palabra que apareciera tenía que ser aprovechada al máximo para tener resultados correctos. Para tratar de tener en cuenta todos los elementos relevantes de la estructura sintáctica de las oraciones agregaron al sistema la elaboración de árboles de análisis de dependencias y una serie de reglas para cambiar la valencia de las palabras que aparecen en los titulares. También crearon automáticamente un listado en inglés de aumentativos y diminutivos de palabras. Al listado también añadieron las negaciones más comunes del inglés. Después de todas la incorporaciones el listado contaba con 490 palabras y expresiones en total. Cada entrada de este listado contaba con la acción que realizaba la palabra o expresión sobre la emoción que poseía la palabra o grupo de palabras a la que acompañaba y un alcance para determinar qué palabras o grupo de palabras se veían afectadas. Es decir, cada entrada del listado contará con una regla para delimitar el ámbito de actuación y la actuación que provocará en la frase. Para delimitar el ámbito de cada entrada se utiliza el sistema de análisis de árboles producido por MiniPar (Lin, 2003).

Otros trabajos, dividen los modificadores en intensificadores y reductores (Kennedy y Inkpen, 2006; Carrillo, 2011), siendo los intensificadores aquellos que aumentan las emociones de las palabras afectadas por ellos (por ejemplo “muy”, “tan” o “extremadamente”), y los reductores los que las disminuyen (por ejemplo “poco”, “menos” y “débilmente”).

Kennedy y Inkpen (2006) tratan los modificadores para variar la valencia de las distintas palabras, cambiando así la polaridad positiva-negativa dependiendo del modificador asociado a la palabra. En el caso de que el modificador esté asociado a una palabra con valencia positiva, si el modificador es un intensificador se le sumará 1 al valor de la valencia de la palabra afectada por el intensificador y si el modificador es un reductor se le restará 1, disminuyendo el valor de la valencia. En caso de que el modificador afecte a una palabra con valencia negativa los valores de intensificador y reductor cambian, si va acompañada por un intensificador se restará 1 al valor de valencia haciendo más negativa la palabra y si va acompañada de un reductor se sumará 1 haciendo más positiva la palabra. Por ejemplo, en la frase “This movie is very good”, si partimos de que “good” tiene una valencia de 2, al añadir la palabra “very” debemos sumar 1 a su valencia convirtiéndolo en un 3 porque las palabras “very good” indican algo más positivo que la palabra “good” en solitario. En la frase “This movie is barely any good”, el término “barely”, es un reductor, y nuevamente si “good” tiene una valencia 2, al añadir el reductor ahora tendría un valor de 1, haciendo

la frase menos positiva. En caso de que expresemos algo negativo como en “This movie is very bad”, suponiendo que la palabra “bad” tiene como valor de valencia un -2, al añadir la palabra “very”, un intensificador, hará la frase más negativa y por ello le restamos 1 y obtenemos una valencia de -3. Si por el contrario la frase es “This movie is barely any bad” la frase ahora es más positiva que antes por lo que si “bad” tenía una valencia de -2, con el reductor “barely”, sumamos 1 quedando la valencia como -1.

Carrillo (2011) indica que no todos los cuantificadores tienen la misma intensidad como dicen Kennedy y Inkpen (2006). Para explicar esto compara las frases “The room was absolutly good” y “The room was quite good”. Podemos ver que aunque ambas indican que la habitación estaba bien, la primera sentencia indica mayor grado de satisfacción que la segunda cambiando únicamente el modificador. Carrillo (2011) crea un listado de términos que actúan como modificadores (ver Figura 2.5). Los elementos de esta lista, extraída de la literatura inglesa (Quirk et al., 1985), tienen asociado un valor en porcentaje indicando cómo debe variar la emoción de la palabra a la que acompañe el modificador.

<b>Términos</b>	<b>%</b>	<b>Términos</b>	<b>%</b>	<b>Términos</b>	<b>%</b>
Very	75	Small	-50	So	75
Great	60	Really	80	More	35
Much	35	Little	-75	Bit	-35
Some	40	Less	-55	Barely	-75
Hardly	-85	Almost	-20	Slightly	-75
Relatively	-50	Somewhat	-60	Fairly	65
Pretty	75	Rather	75	Quite	35
Perfectly	75	Obviously	75	Certainly	75
Completely	85	Definitely	95	Absolutely	95
Highly	75	Truly	75	Especialy	70
Particularly	45	Significantly	45	Awfully	75
Totally	85	Fully	85	Entirely	75
Strongly	55	Extremely	95	Incredibly	85
Terribly	75	Immensely	75	Such	55
Exceptionally	75	Exceedingly	85	Vastly	75
Super	75	Thoroughly	65	Tremendously	85
Ridiculously	-65	Extraordinarily	75	Total	75
Huge	75	Tremendous	85	Complete	75
Absolute	75	Incredible	75	Utter	75
Big	65	Bigger	35	Biggest	75
Extra	20	Major	35	High	55
Higher	45	Highest	65	Real	20
Minor	-75	Low	-75	Lower	-65
Lowest	-85	Few	-75	Fewer	-65
Fewest	-85	Lot	75	Difficult	-65
Only	-25				

Figura 2.5: Lista de modificadores de Carrillo (2011)

## 2.7. Papel de la negación en el mercado emocional de textos

En las primeras investigaciones sobre el análisis emocional de textos, se empleó un método sencillo para tratar la negación. Este método consistía en marcar con distintas etiquetas las palabras que se veían afectadas por la negación para luego cambiar el valor emocional de dichas palabras (Wiegand et al., 2010; Das y Chen, 2007; Polanyi y Zaenen, 2006). Por ejemplo, en algunos estudios la etiqueta utilizada para marcar las palabras afectadas por la negación era “NOT\_” añadiendo este prefijo al comienzo de todas las palabras afectadas. En el caso de la frase “I am not happy” se marcaría como “I am not NOT\_happy” para cuando se evalúe la palabra “happy” tener en cuenta que está afectada por la negación.

Más adelante otros autores, decidieron asumir que la negación no solo afecta a la palabra que se encuentra a continuación sino que también puede afectar a otras palabras. Para decidir a qué palabras afecta, Pang et al. (2002) deciden asumir que la negación solo afecta desde que aparece hasta el siguiente signo de puntuación, aunque en estos casos no aumentaba mucho el porcentaje de acierto general del clasificador de emociones. Otros en cambio, como Jia et al. (2009), hicieron un análisis de los árboles de dependencia de las frases o textos para identificar el ámbito en el que la negación tiene efecto.

Basándose en los estudios de análisis sintáctico de dependencias, Vilares et al. (2013) crearon un clasificador de opiniones en español, en el que se contemplan como negación únicamente tres palabras españolas: *no*, *nunca* y *sin*. En este trabajo para tratar la negación se realiza primero una identificación del alcance de la misma usando árboles de dependencias, donde cada palabra contemplada como negación tendrá su propio alcance y su propia manera de determinarlo viendo el árbol de dependencias. Una vez se ha definido el alcance de la negación, se obtiene la polaridad para el conjunto afectado y ese valor obtenido es modificado por una cantidad preestablecida para cada negación, para “no” y “nunca” el valor es 4 mientras que para “sin” el valor es 3,5. Por ejemplo para la frase “Ese ordenador es muy rápido pero no recomiendo comprarlo” se identificaría el “no” como negación, y basándonos en su árbol de dependencias (que podemos ver en la Figura 2.6), el ámbito de esta negación abarca solo a la palabra “recomiendo”. La palabra “recomiendo” tiene una polaridad de 2 a la que si restamos el valor preestablecido 4 de la negación “no” nos quedaría una polaridad de -2 (2-4).

Con ánimo de mejorar el trabajo realizado, Jiménez et al. (2015) amplían el corpus de palabras analizando críticas de cine, introduciendo “tampoco”, “nadie”, “jamás”, “ninguno”, “ni” y “nada” al listado de palabras reconocidas como negaciones. Para cada una de las palabras del corpus de negaciones establecieron reglas que indican a qué ámbito o conjunto de palabras afecta en un análisis de dependencias las cuales podemos ver en la Figura 2.7 y así cada palabra del corpus tendrá por lo tanto, un ámbito de aplicación determinado por distintas reglas. Después de tener definido el ámbito de la negación se procede a modificar la polaridad del fragmento que se ve afectado invirtiendo la polaridad. Por ejemplo en la frase: “La película no empieza mal con algunas escenas interesantes para introducirte en el drama vivido por el personaje principal”, la partícula “no” es la que marca la negación y afectará, según el árbol de dependencias. Podemos ver este listado en la Figura 2.8. Si observamos el listado de las palabras de la frase “La película no empieza mal con algunas escenas interesantes para introducirte en el drama vivido por el personaje principal”, las palabras “empieza” y “mal” aparecen en dicho listado y sólo la palabra “mal” es considerada como poseedora de carga afectiva con una polaridad de -1 (negativa). Por lo tanto, la polaridad de la palabra “mal” se verá invertida convirtiéndose en una palabra con polaridad 1 (positiva).

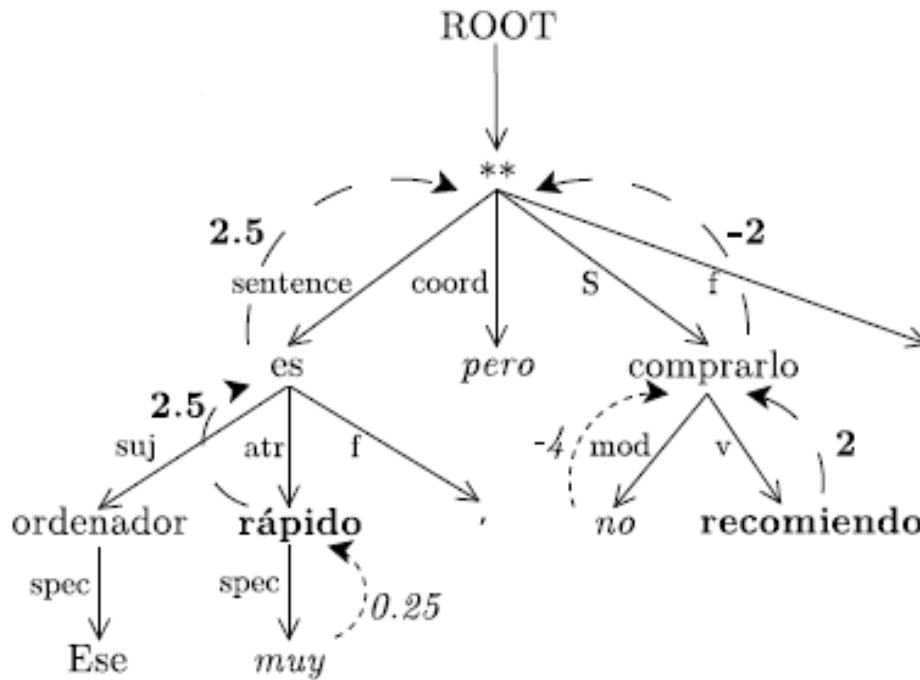


Figura 2.6: Árbol de dependencias obtenido para la frase “Ese ordenador es muy rápido pero no recomiendo comprarlo” según Vilares et al. (2013)

Partícula	Regla
no, tampoco, nadie, jamás, ninguno	Afecta al nodo padre y al árbol formado por el hermano de la derecha (incluido).
ni, sin	Afecta a todos los hijos y a todos los árboles formados por ellos hasta llegar a nodos hoja.
nada, nunca	Afecta al nodo padre.

Figura 2.7: Reglas para definir el ámbito de la negación por Jiménez et al. (2015)

Otros autores consideran que la negación es únicamente un modificador, como los que hemos visto en la Sección 2.6, que modifica la intensidad de la emoción (Whitelaw et al., 2005; Andreevskaia y Bergler, 2007; Kennedy y Inkpen, 2006; Carrillo, 2011). Carrillo (2011) considera a las llamadas negaciones suaves o implícitas como modificadores, este tipo de negación son las generadas con palabras como “less” o “rarely” y solo considera negación a las explícitas, las que aparecen con palabras como “not” o “never”. Para identificar las negaciones compara cada término del texto u oración con un listado de estas palabras,

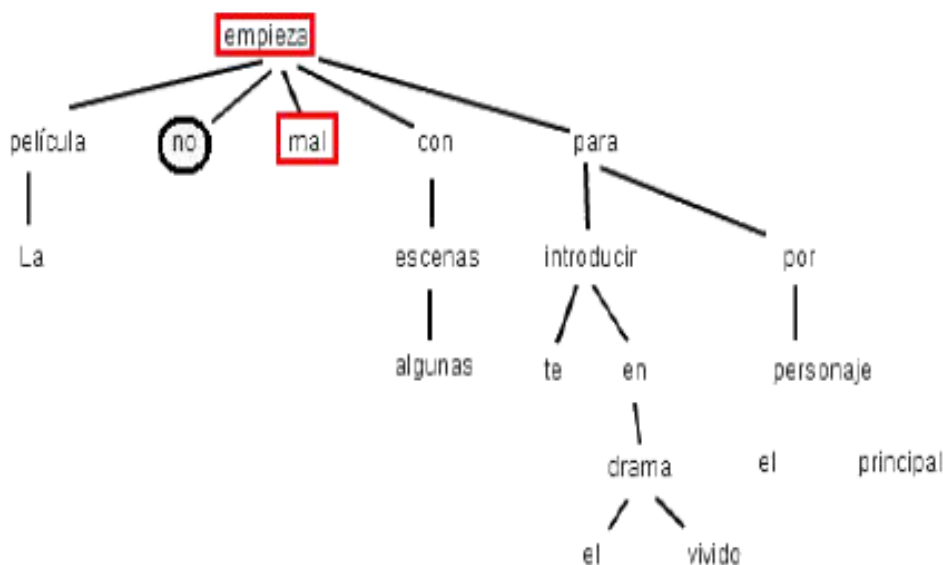


Figura 2.8: Árbol de dependencias obtenido para la frase “La película no empieza mal con algunas escenas interesantes para introducirte en el drama vivido por el personaje principal” según Jiménez et al. (2015)

que podemos ver en la Figura 2.9. También tiene un listado con expresiones que utilizan términos incluidos en la lista de negación, pero que en el contexto del texto u oración no indican este cambio de polaridad como pueden ser: “Not only”, “Not just” o “Not even”. Una vez identificados los términos de la negación, para definir el ámbito, utiliza árboles de dependencia sintáctica para que el cambio de polaridad solo afecte a la parte de la oración o texto afectado por la negación. Sin embargo, utilizar únicamente los árboles de dependencia, en ocasiones, es incorrecto por las oraciones subordinadas u otros casos donde varían los signos de puntuación; para mejorar este aspecto, Carrillo (2011) introduce una serie de reglas que podemos ver en la Figura 2.10 donde se indican las palabras que toma como delimitadores del ámbito de la negación, para que en los casos especiales se haga una correcta identificación del ámbito de la negación. Por ejemplo en la frase “I didn’t need any transportation because the hotel place is good” el sistema detecta la negación “didn’t” y determina como posible ámbito “need any transportation because the hotel place is good”, sin embargo, como en este ámbito se encuentra una de las palabras que aparece en las reglas de la Figura 2.10 “because” que indica el inicio de una oración subordinada, la regla se ejecuta y el ámbito de la negación se reduce a “need any transportation”.

Respecto a la forma de tratar la negación, Polanyi y Zaenen (2006) hacían que las palabras marcadas cambiarán su polaridad para volverse negativas en caso de ser positivas y viceversa. Este mismo método se utilizaba en los trabajos donde no se trata a la negación como un modificador (Wiegand et al., 2010; Das y Chen, 2007). Pero en caso de que no tratemos la polaridad o la intensidad de la emoción, sino que clasifiquemos los textos con categorías emocionales este método no es válido. Para superar este obstáculo, autores como Carrillo (2011) deciden usar la antonimia como salida y no el cambio de polaridad, es decir, cuando encuentran una palabra que se ve afectada por una negación, si la palabra es alegre, se consideraría triste. Esta metodología tiene como principal problema que en ocasiones no podemos encontrar una categoría emocional antónima y entonces la palabra es etiquetada

No	None	Non	Nor	Nothing	Neither	Nobody
Never	Nowhere	Not	n't	Don't	Dont	Doesn't
Doesnt	Didn't	Didnt	Haven't	Havent	Can't	Cant
Cannot	Couldn't	Couldnt	Needn't	Neednt	Wasn't	Wasnt
Weren't	Werent	Daren't	Darent	Hadn't	Hadnt	Isn't
Isnt	Aren't	Arent	Oughtn't	Oughtnt	Wouldn't	Wouldnt
Hasn't	Hasnt	Mightn't	Mightnt	Mustn't	Mustnt	Shouldn't
Shouldnt	Won't	Wont	Shan't	Shant		

Figura 2.9: Términos incluidos en la lista de negaciones de Carrillo (2011)

<b>Término</b>	<b>POS</b>
so, as, because, that, if until, since, unless, before	<b>IN</b>
while, about, than, despite	<b>IN</b>
who, what, whose	<b>WP</b>
Why, when, where	<b>WRB</b>
which	<b>WDT</b>
however	<b>RB</b>
that	<b>RB</b>
“ ” , = , : , ( , ) , ! , ?	<b>=</b>

Figura 2.10: Lista de delimitadores del ámbito de la negación de Carrillo (2011)

como palabra sin categoría emocional. Por ejemplo, con la categoría emocional de “miedo” tendríamos problemas, ya que no tener miedo no implica estar tranquilo o calmado. Los autores que tratan la negación como un modificador (Whitelaw et al., 2005; Kennedy y Inkpen, 2006) únicamente aplicarían el porcentaje asignado como si se tratase de un reductor.

## 2.8. Herramientas existentes para detectar la emoción de un texto

En esta Sección vamos a hacer un repaso de algunas de las herramientas existentes para marcar textos emocionalmente.

### 2.8.1. EmoTag

EmoTag (Francisco, 2008) es un sistema de marcado de textos orientado al dominio de los cuentos. Se decide este dominio porque la finalidad de los cuentos es enseñar a los niños a entender sentimientos. Al igual que otros marcadores que veremos más adelante utiliza la frase como unidad de marcado, es decir, se identifican las emociones a nivel de frase. Esta herramienta representa la emociones usando dimensiones emocionales (activación, evaluación y control) a las que se le asignará un valor y usando categorías emocionales teniendo 119 donde elegir para cubrir cualquier texto.

El proceso que sigue EmoTag para marcar las emociones de las frases es el siguiente:

1. División en frases. Primero se divide el texto introducido en frase y se obtiene el árbol de dependencias de la frase y la categoría léxica de cada palabra junto con su lema. Después cada frase es dividida en palabras y se evalúan una a una. Primero se descartarán aquellas palabras cuya categoría léxica pertenezca a la lista de categorías de parada o las que no se encuentren marcadas en el diccionario emocional *The General Inquirer* (Stone et al., 1996), que ofrece una lista de palabras marcadas con distintas categorías. Para las palabras no descartadas se obtiene el lema de la palabra. Al finalizar esta primera fase tendremos un árbol de dependencias de la frase y las palabras “emocionales”, las no descartadas, con su categoría léxica y su lema.
2. Obtención del valor emocional de cada palabra. Las palabras obtenidas en la frase anterior son buscadas en los diccionarios LEW y ANEW, en estos diccionarios se busca por el par (lema, categoría léxica). LEW es un diccionario afectivo creado por la misma autora y ANEW (Bradley y Lang, 1999) es un diccionario emocional que hemos visto ya en la Sección 2.4.1. La búsqueda en estos diccionarios devolverá para cada palabra el valor de cada dimensión (evaluación, activación y control) y para las categorías emocionales devolverá un factor de correlación entre la palabra y cada una de las emociones. En caso de que la palabra no se encuentre en LEW ni ANEW se buscarán cada uno de los hiperónimos de la palabra y se volverá a realizar la búsqueda. Tanto si se encuentra la palabra sin necesidad de los hiperónimos como si se han necesitado el contenido emocional será asociado a la palabra original evaluada. Si ninguno de los hiperónimos tiene información emocional la palabra se descartará del proceso. Al final de esta fase tendremos cada una de las palabras emocionales con información de sus dimensiones emocionales y las categorías emocionales.
3. Tratamiento de la negación. El tratamiento de la negación consiste en identificar todas las frases donde aparezca una negación e invertir el valor emocional de las palabras que se vean influenciadas por la negación. Para analizar el ámbito de actuación de la negación se utiliza el árbol de dependencias que se obtuvo en la primera fase. La inversión del contenido emocional se hace de dos maneras distintas, una para las dimensiones emocionales y otra para las categorías emocionales. En el caso de las dimensiones emocionales se obtiene el valor contrario haciendo una resta donde el sustraendo es 10 porque es el máximo valor para cada dimensión, el minuendo es el valor que tenía la dimensión obtenida de los listados disponibles y el resultado es el valor nuevo para la dimensión. En el caso de las categorías emocionales se obtiene la emoción contraria. Para encontrar la emoción contraria hay que recorrer cada una de las categorías que tienen un factor asignado para buscar su contrario, en caso de que se encuentre se asigna a la categoría inicial el factor 0 y a la contraria se le asigna el factor original. En la Figura 2.11 se puede ver cómo funciona este proceso de manera muy clara. En esta Figura vemos que lo evaluado tenía como emociones “surprise” con 5,079, “happiness” con 7,408, “admiration” con 2,088 y “neutral” con 2,648, para invertir los valores se buscan sus emociones antónimas, excepto para “neutral” que seguirá teniendo el mismo valor, encontramos en la lista de antónimos las emociones antónimas de “happiness” que es “sadness” y la de “admiration” que es “contempt”, por lo tanto a estas categorías antónimas les ponemos los grados 7,408 y 2,088 respectivamente. En cuanto a la emoción “surprise” que no hemos encontrado su antónimo en el listado para la inversión ponemos como valor un 0,000.
4. Emoción asignada a una frase. Después de todas las fases anteriores donde hemos evaluado cada palabra se determina el valor final de la frase. Lo primero que se hace es tener en consideración las oraciones subordinadas que hay dentro de la frase

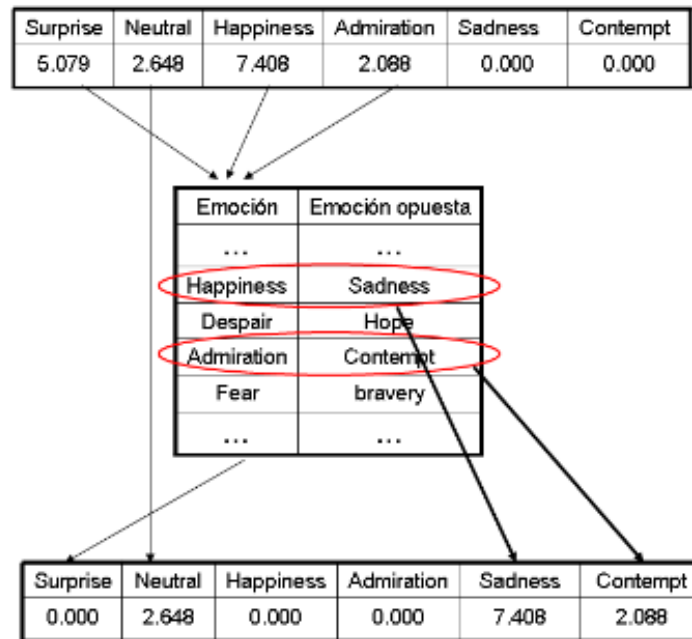


Figura 2.11: Funcionamiento de la inversión de la categoría emocional en el EmoTag.

calculando la emoción de cada una de ellas para considerarlas como una palabra más dentro de la frase principal con la emoción que se ha calculado para ella. Después, habiendo convertido las subordinadas en una palabra más, se consigue el valor final de oración. Por ejemplo en la frase en español “Me gusta bailar flamenco” donde la oración subordinada es “bailar flamenco”, primero se hallaría la carga emocional de esta frase subordinada y después se evaluaría la carga emocional de la frase completa como si la subordinada se hubiera sustituido por una palabra con la carga emocional de la subordinada. Para obtener el valor de las dimensiones emocionales se suman los valores de cada dimensión de las distintas palabras y se asigna el valor medio a cada una a las dimensiones de la frase (activación, evaluación y control), este valor se obtiene sumando los valores y dividiendo entre el número de palabras que han aportado valores. Para obtener la categoría emocional de la frase se emplea la ontología emocional OntoEmotion desarrollada en el mismo trabajo por la misma autora. Para elegir la categoría emocional de la frase después de haber evaluado cada palabra se suman los factores de cada emoción, después se agrupan las emociones según el nivel que tengan dentro de la ontología y con la ayuda de esta se procesan las emociones para obtener los padres de cada emoción en la ontología y estos se añaden a los que teníamos anteriormente con el factor de correlación asociado al concepto más específico y después la emoción con mayor factor es asignada a la frase. Por ejemplo, para una frase con 3 palabras a evaluar emocionalmente con los valores que tenemos en la Figura 2.12, vamos a calcular la emoción final asociada a la frase. Primero, siguiendo la categorización por niveles que proporciona OntoEmotion (ver Figura 2.13) agruparemos en niveles las emociones, quedando los resultados que vemos en la Tabla 2.6. A continuación, obtendremos las etiquetas padres de las emociones del Nivel 2 según la ontología. Siguiendo la Figura 2.13 vemos que el padre de las etiquetas “Indignantion”, “Sulking”, “Dspleasure” es “Anger” y que

Emotion	Anger	Fear	Surprise	Sadness	Indignation	Sulking	Displeasure	Amazement
Palabra 1	3.03	4.04	0	3.03	0	0	0	0
Palabra 2	0	2.02	5.05	1.01	2.02	0	0	0
Palabra 3	0	2.02	1.55	1.55	0	1.55	3.03	0.5
Total	3.03	8.08	6.60	5.59	2.02	1.55	3.03	0.5

Figura 2.12: Valores emocionales para una frase de ejemplo en EmoTag.

el de la etiqueta “Amazement” es “Surprise”. Sabiendo esto, sumamos el factor de correlación que posee cada etiqueta del Nivel 2 al valor que poseen sus etiquetas padres en esta frase quedando los resultados: “Fear” (8,08), “Anger” (9,63), “Sadness” (5,59) y “Surprise” (7,10). De estos valores, elegiremos el mayor y será la emoción final asociada la frase, en este caso será “Anger”.



Figura 2.13: Niveles de etiquetas emocionales en OntoEmotion.

Nivel 1	Fear (8,08), Anger (3,03), Sadness (5,59), Surprise (6,60)
Nivel 2	Indignation (2,02), Sulking (1,55), Displeasure (3,03), Amazement (0,5)

Tabla 2.6: Resultado de la agrupación de etiquetas emocionales por niveles según OntoEmotion.

### 2.8.2. Sentitex

Sentitex es un sistema de análisis de sentimientos para el español (Moreno et al., 2010a) que muestra los resultados usando gráficos y textos según la polaridad de los fragmentos de texto analizado, y proporciona una valoración global del texto.

Su característica más importante es que está basado en el conocimiento y se apoya en tres bases de datos: Words, MWords y CRules. Words es una base de datos, creada en este mismo trabajo, con más de 10.000 palabras, donde cada palabra tiene asignada una valencia que indica la carga afectiva. El rango de valores asociados a la valencia varía entre -2 y 2, siendo una palabra que tiene asociado un -2 una palabra muy negativa y una palabra que tiene asociado un 2 una palabra muy positiva. Para obtener las palabras se recurrió al diccionario de sinónimos de OpenOffice<sup>3</sup>. Para obtener las valencias asignadas a cada palabra, primero se asignaron valencias a las palabras en base a la semántica de la misma con un corpus de textos de lengua general y después lexicógrafos que trabajan con la herramienta deciden cual es el valor afectivo más adecuado para la palabra. MWords es otra

<sup>3</sup><https://extensions.openoffice.org/en/project/diccionario-de-correccion-ortografica-separacion-silabica-y-sinonimos-en-espanol-espana>

base de datos, también creada en este trabajo, con unas 17.000 entradas con expresiones que mejora los resultados del marcador para expresiones como “ir sobre ruedas”. Esta base de datos es necesaria ya que algunas expresiones aún teniendo una palabra positiva son expresiones negativas o como la expresión “ir sobre ruedas” que a pesar de estar formada por palabras neutras forma una expresión positiva. Estas dos bases de datos (Words y MWOrds) están sometidas siempre a actualizaciones ya que se siguen analizando textos diariamente para mejorar la información que contienen. CRules, también creada en el mismo trabajo, sirve para tener en cuenta las unidades léxicas que modifican la carga de otras unidades cercanas como ocurre con la negación o los modificadores. Estas reglas están basadas en los estudios de Polanyi y Zaenen (2006). Para conseguir el marcado emocional de un texto la herramienta sigue cuatro pasos:

1. Lematización y etiquetado morfológico del texto.
2. Asignación de valencias. Para ello la herramienta recorre la lista de unidades léxicas obtenidas y usando el lema de la palabra se busca la valencia en las tres bases de datos.
3. Aplicación de reglas de contexto. La herramienta vuelve a recorrer la lista de unidades léxicas y en caso de encontrar un modificador que cumpla las restricciones indicadas para una regla (posición, cercanía, naturaleza del elemento a modificar) se transforma la valencia de la unidad a la que modifica. Las reglas de contexto están jerarquizadas por si se cumple más de una regla al ser evaluada una unidad léxica.
4. Extracción de datos. Se obtienen datos derivados de los análisis anteriores como el índice afectivo, que es la cantidad con carga afectiva en relación con el número total de palabras, o el índice global que intenta aproximar la positividad o negatividad del texto.

Sentitext fue evaluada usando valoraciones de hoteles realizadas por clientes en la plataforma *Tripadvisor* (Moreno et al., 2010b). La evaluación pone de manifiesto algunas de las limitaciones de la herramienta relacionadas con aspectos formales como el uso de mayúsculas para enfatizar sensaciones, la repetición de letras o el uso de signos de puntuación; la semántica del dominio evaluado ya que hay palabras que habitualmente no son positivas o negativas pero al estar evaluando un hotel cobran cierta importancia, como por ejemplo las palabras “limpieza” y “suciedad” o “cálido” y “frío”; y de aspectos discursivos como la ironía, el sarcasmo o la cortesía lingüística.

### 2.8.3. Modelo de clasificación de textos según su polaridad e intensidad

Este modelo fue creado por Carrillo (2011) utilizando aprendizaje supervisado, análisis semántico y reglas lingüísticas. El objetivo es, dado un texto traducirlo a un conjunto de categorías emocionales que representarán el significado afectivo del texto. Estas categorías emocionales irán acompañadas con un grado de intensidad, pudiendo cada texto, tener más de una categoría emocional asociada con distintos o iguales grados de intensidad. El sistema trabaja solo con textos en inglés ya que hay un mayor abanico de herramientas para realizar análisis de textos. Independientemente se podría adaptar de manera sencilla a otros idiomas ya que el algoritmo no depende del idioma. El proceso de marcado conlleva 7 fases:

1. Pre-procesado. La primera fase consiste en un análisis sintáctico del texto donde se etiqueta cada uno de los términos según su categoría gramatical, se generan árboles sintácticos de cada oración y se identifican las dependencias entre términos.

2. Identificación de conceptos. Esta segunda fase se encarga de convertir cada término etiquetado en la fase anterior como *sustantivo*, *verbo*, *adjetivo* o *adverbio* en conceptos para conseguir que la herramienta trabaje a nivel semántico, usando los conceptos y no los términos, eliminando posibles ambigüedades. En esta fase también se recuperan todos los hiperónimos.
3. Identificación de categorías emocionales. Esta tercera fase se encarga de obtener para cada concepto su categoría emocional en caso de tenerla. En caso de que para el concepto no exista una categoría emocional asociada se utilizan los hiperónimos obtenidos en la fase anterior, pudiendo recurrir hasta a 3 hiperónimos intentando perder el mínimo significado del concepto. Para obtener la categoría emocional se utiliza un léxico que asocia los conceptos con categorías emocionales desarrollado en el mismo trabajo llamado SentiSense (Carrillo, 2011). SentiSense es un diccionario afectivo que contiene conceptos o significados en lugar de palabras. Cada palabra está etiquetada con una de las siguientes categorías emocionales *like*, *compassion*, *joy*, *anger*, *hate*, *expectancy*, *love*, *sadness*, *calmness*, *ambiguous*, *fear*, *hope*, *surprise* y *disgust*. SentiSense en total cuenta con 2.214 conceptos que abarcan 5.551 términos. En este paso también se busca en el sistema la relación de antonimia que está definido entre categorías emocionales como puede ser *joy-sadness*.
4. Detección de negación. Esta cuarta fase se divide en tres pasos diferentes.
  - a) Identificación de la negación. Se buscan partículas o términos de negación teniendo en cuenta distintos tipos de negación: las explícitas que son las introducidas por palabras como “not” o “never” y las implícitas que usan términos como “less” o “rarely”. Las negaciones implícitas se tratarán como modificadores. Para detectar las negaciones explícitas se compara cada término de las oraciones con un listado de negaciones que posee el sistema. Una vez detectadas pasará al segundo paso.
  - b) Identificación del ámbito de actuación de la negación. Para detectar el ámbito de la negación se usa el árbol sintáctico generado en la primera fase de pre-procesado. En el árbol se buscará el primer ancestro en común entre el término de la negación y el que le sigue inmediatamente después en la oración. Una vez encontrado se asume que todos los nodos hoja descendientes de ese ancestro que están a la derecha de la negación en la oración están afectados por ella. Dado que este método fallaría con oraciones subordinadas, se crean una serie de normas para eliminar de este ámbito cualquier oración candidata a ser subordinada.
  - c) Aplicación del impacto de la negación. Dado que es una herramienta que utiliza categorías emocionales no es válido el método de otros autores donde se invierte la polaridad de las emociones y propone un método más sofisticado que consiste en cambiar la categoría emocional por una antónima que se encuentre en el léxico y una vez cambiada disminuir la fuerza de la categoría emocional antónima ya que la negación implica una pérdida de intensidad, esta pérdida de fuerza consistirá en que la emoción en vez de tener asociada la intensidad al 100 % tendrá únicamente un 60 %. En caso de que una categoría emocional no tenga una antónima se dejará el término sin etiqueta emocional.
5. Detección de modificadores. Esta quinta fase consiste en identificar los modificadores presentes en el texto y decidir su ámbito de actuación. Para ello se sigue un proceso similar al de la detección de la negación explicado en el paso anterior y para ello se

incluye un listado de palabras que se consideran modificadores junto con la categoría gramatical que deben tener para actuar como modificador en una frase. Considerando que cada modificador modifica el contenido emocional de la frase en distinta medida, tal y como hemos visto en la Sección 2.6, el sistema tiene relacionado cada modificador que se encuentre en la lista mencionada antes con un porcentaje de actuación con el que variará la intensidad de la emoción, pudiendo intensificarla o disminuirla. El ámbito de los modificadores se determina usando el árbol sintáctico elaborado en la primera fase de la misma manera que con la negación.

6. Detección de modales. En esta fase el sistema tiene como fin detectar la modalidad y su ámbito. La modalidad puede ser definida como la actitud o punto de vista del emisor hacia el estado actual del mundo y puede invertir, disminuir o incluso no afectar en nada a la polaridad e intensidad del texto.
7. Clasificación automática. Después de tener todas las oraciones analizadas, identificadas con sus categorías emocionales de todos sus conceptos y resueltos todos los modificadores, negaciones y modalidad, el texto se traduce en un *Vector de Intensidades Emocionales* que será la entrada de un algoritmo de aprendizaje máquina donde será clasificado de acuerdo a su polaridad e intensidad emocional. Este vector tiene 14 posiciones, correspondiente a las 14 categorías emocionales que contiene SentiSense.

Vamos a ver cada uno de los pasos de este sistema con un ejemplo. Tomaremos como entrada al sistema la frase “Cleaning starts very early even on Sunday, furnitures are moved, very noisy, not possible to relax and sleep”.

1. Pre-procesado. En esta primera fase se realiza el análisis sintáctico de la frase, se etiqueta cada término con su categoría gramatical, se genera el árbol sintáctico de la oración y se identifican las dependencias entre términos.
2. Identificación de conceptos. Se convierte cada término etiquetado como *sustantivo*, *verbo*, *adjetivo* o *adverbio* en concepto. Resultando los conceptos de la frase: “cleaning”, “relax” y “noisy”.
3. Identificación de categorías emocionales. Usando SentiSense obtenemos que la categoría emocional de “cleaning” y de “relax” es *like* y la del concepto “noisy” es *dislike*.
4. Detección de negación. En esta fase nos damos cuenta que el concepto “relax” se encuentra en el ámbito de la negación “not” por lo que el sistema cambia su categoría emocional por la antónima (*disgust*).
5. Detección de modificadores. El sistema encuentra el modificador “very” que amplifica la carga afectiva de su ámbito y en este se encuentra “noisy” amplificando un 75 % su emoción.
6. Detección de modales. No se encuentra ningún modal en esta oración, por lo que continúa con la siguiente fase.
7. Clasificación automática. Se genera el VEI (Vector de Intensidades Emocionales) que quedaría siguiendo el orden de las categorías emocionales *like*, *compassion*, *joy*, *anger*, *hate*, *expectancy*, *love*, *sadness*, *calmness*, *ambiguous*, *fear*, *hope*, *surprise*, *disgust* de la siguiente manera  $[1,0 - 0,0 - 0,0 - 0,0 - 0,0 - 0,0 - 0,0 - 0,0 - 0,0 - 0,0 - 0,0 - 0,0 - 2,35]$  donde vemos que para la categoría *like* tenemos un 1,0 por

el concepto “cleaning”, mientras que en *disgust* tenemos un 2,35 conseguido por: la palabra “noisy”, que sin afectar por el modificador aportaría un 1,0, pero al sumar el 75 % se convertirá en un 1,75 y por la palabra “relax” que fue afectada por la negación y suma un 0,6 en vez de un 1,00 por solo aportar un 60 % por haber sido afectada por la negación. Este vector será introducido en el algoritmo de aprendizaje donde será clasificado de acuerdo a su polaridad e intensidad emocional.

Esta herramienta tiene como ventaja el trabajar a nivel de concepto en vez de a nivel término, lo que ayuda a diferenciar entre matices de la palabra y por tanto, mejorará el marcado. Por ejemplo al encontrarnos con el término “cancer”, este podría tratarse de un tumor, de una constelación o de un signo del zodiaco; pero al tratarlo como concepto se asociarán las emociones adecuadas al ámbito de la frase gracias a la fase 2 de identificación de conceptos, donde reconocerá a través del contexto el hiperónimo adecuado para el término.

Otro punto destacable del sistema es que tiene en cuenta las negaciones tratándola no solo para la polaridad sino también para categorías emocionales y trata los modificadores dándoles distintos pesos conforme al uso de cada uno.

#### 2.8.4. Lingmotif

Lingmotif (Moreno-Ortiz, 2017a) es una aplicación de Análisis de Sentimiento multiplataforma con una interfaz enfocada al usuario. La primera versión de la aplicación está preparada para trabajar con textos en español o inglés, en versiones posteriores soportará también el francés, alemán e italiano. Esta aplicación analiza los textos a nivel oración, esto quiere decir que analiza frase a frase los textos, lo que puede producir pérdidas de matices que aparezcan a nivel texto. También tiene en cuenta los modificadores que intensifican, disminuyen o invierten por completo las emociones que transmiten las frases. Únicamente trabaja con la valencia que indica si una frase o palabra es positiva o negativa.

La aplicación devuelve las siguientes métricas:

- TSS - Text Sentiment Score. Este número indica la puntuación general del texto. Utiliza un intervalo del 0 al 100 y en cuanto más pequeño es el valor asociado más negativo será el texto. Por ejemplo si evaluamos un texto y nos da como resultado un 12 significa que es un texto extremadamente negativo.
- TSI - Text Sentiment Intensity. Es la proporción de elementos que conllevan un sentimiento versus los elementos que no tienen. Utiliza también un intervalo del 0 al 100, cuanto más pequeño es el valor asociado más baja es la intensidad de emociones presentes en el texto. Si tuviéramos un 92 en este marcador significaría que el texto es extremadamente intenso, tiene muchas palabras que aportan carga emocional.
- Sentiment Profile. Es una representación gráfica del “flujo” de la opinión del texto, cómo varía la emoción a lo largo del discurso. En esta gráfica se va viendo como varía el TSS (eje Y) a lo largo del número de palabras que contiene el texto (eje X). De esta manera se puede ver si alguna parte es más positiva o negativa que otra.
- Text Analysis. Es un conjunto de métricas del texto: número de tokens, tipos, oraciones, palabras léxicas y palabras funcionales. La información que se obtiene se puede ver en la Figura 2.14 para un texto de muestra que se proporciona en el estudio, en este caso podemos ver que el texto poseía por ejemplo 6.898 palabras en 637 frases donde 2.916 cumplían una función léxica, que aportaban significado, mientras que

3.982 solo poseían función gramatical. También vemos que se han analizado 6.409 palabras simples y 489 multipalabras o expresiones formadas por más de una palabra.

- **Quantitative Sentiment Analysis.** Es el desglose de las cifras que se obtuvieron para calcular el TSS y el TSI. Se puede ver también en la Figura 2.14 los valores que se obtienen entre otros los elementos positivos, negativos y neutros que se encuentran y cuál es la puntuación positiva y negativa. Vemos que para el texto de ejemplo había 169 términos positivos, 298 negativos y 2.416 neutros. También aparece un valor llamado CVS (cambio de valencia de contexto) que indica cuántos fragmentos del texto se han visto afectados por alguna regla de contexto y, por tanto, se han visto cambiadas sus valencias por un modificador. En el ejemplo vemos que el CVS es 95, lo que significa que se han aplicado 95 modificadores distintos al texto.
- **Detailed Sentiment Analysis.** Se muestra el texto de entrada donde los sentimientos están codificados mediante colores de acuerdo con la polaridad que tengan. Vemos en la Figura 2.15 cómo se marca en el texto con el color verde las partes positivas del texto y con rojo las negativas.

**Text Analysis**

Text Stats			Sentences	Words by Function		Words by Form	
Tokens	Types	T/T Ratio		Lexical	Grammatical	Single Words	Multiwords
6898	1293	18.74%	637	2916	3982	6409	489

**Sentiment Analysis**

TSS	TSI	CVS segments	Positive items	Negative items	Neutral items	Positive Score	Negative Score
38	88	95	169	298	2416	379	693

Figura 2.14: Ejemplo de datos proporcionados por Lingmotif

**Detailed Sentiment Analysis**

Well, I actually **agree\_with** that.

I **agree\_with** everything she said.

I began this campaign because I was **so\_tired\_of** seeing such **foolish** things happen to our country.

This **is\_a\_great** country.

This **is\_a\_great** land.

I have **gotten\_to\_know** the people of the country over the last year-and-a-half that I have been doing this as a politician.

Figura 2.15: Ejemplo de un texto marcado por Lingmotif

Para obtener los resultados de las emociones Lingmotif se basa en los recursos léxicos que posee y que describimos a continuación. Estos recursos los tiene para cada idioma que pueda reconocer la aplicación. Los recursos en español están basados en los que se usaron para Sentitext (Moreno et al., 2010a) y solo se han adaptado. Estos son la base de los resultados que proporciona la herramienta y su calidad es directamente proporcional a la calidad de los recursos. Los recursos con los que cuenta son:

- **Core lexicon:** Es un léxico que contiene palabras individuales y expresiones de varias palabras o multiplabras con su sentimiento asociado (la valencia) y su función en las frases. La valencia es un número entero de -5 a -2 para elementos negativos, de -2 a

2 los neutrales y de 5 a 2 para positivos. Todas las palabras individuales tienen una valencia que no es 0. Es el más importante de todos los elementos del sistema.

- Un conjunto de reglas de contexto donde se define el uso de los modificadores. Funcionan especificando palabras o frases que puedan aparecer a continuación de la palabra sentimental identificada. Usan sumas y restas simples de números enteros usando una escala de -5 a 5. Cada regla tiene especificado la ubicación (izquierda o derecha) con respecto a la palabra a la que afecta, la extensión de su efecto (en número de palabras) y el valor que suma o resta el modificador (entre -5 y 5). En la Figura 2.16 vemos un ejemplo de algunas reglas que se usan en el sistema. Si no fijamos por ejemplo en la segunda de los modificadores de inversión sabemos que las palabras que tengan como categoría gramatical “adjetivo” (JJ), tanto si es positivo como negativo (+-), cuando aparece “not” por la izquierda (L) se deberá invertir (INV0) y que su ámbito cubre 2 palabras.

Shift type	Example Context Rule
Inversion	NN,-,avoid*,LR,5,INV0
	JJ,+-,not,L,2,INV0
Intensification	JJ,+-,seriously,L,2,INT3
	VB,+-,may_well,L,1,INT1
Downtoning	NN,-,mild,L,2,DOW1
	NN,+-,a_bit,L,2,DOW1

Figura 2.16: Ejemplo de reglas de contexto que emplea Lingmotif

- Un etiquetador de partes del discurso. Es una parte muy necesaria para poder marcar los límites de las reglas de contexto de los modificadores que hemos visto en el punto anterior. Simplemente se encarga de etiquetar cada palabra con su categoría gramatical.
- Un lematizador.

El proceso de análisis es el siguiente:

1. Preprocesamiento. El texto es escaneado para encontrar abreviaturas, contracciones y errores ortográficos con la finalidad de corregirlos.
2. Tokenización. Se hace una tokenización a nivel frase y a nivel palabra.
3. Identificación de las “multipalabras”. Los conjuntos se comparan con las expresiones de múltiples palabras que se encuentran en el léxico de la aplicación, en caso de existir se marcan y se asigna su valencia.
4. Identificación de palabras con polaridad. Las palabras individuales se buscan en el léxico, utilizando su lema, y se les asigna la valencia si son encontradas.
5. Aplicación de reglas de contexto. Las palabras con polaridad identificadas y las multipalabras se comparan con la lista de reglas de contexto. Si una regla de contexto coincide con el contexto de una palabra, todo el segmento del texto (el que indicará la regla que se ve afectado) se marca y etiqueta como una unidad del tipo CVS (cambio

de valencia de contexto). Este proceso se repetirá dos veces para tener en cuenta las secuencias de modificadores como en la frase “very very good”.

En una versión posterior del Lingmotif (Moreno-Ortiz, 2017b), añaden el tratamiento de emoticonos. Para tratarlos los incluyen en el léxico que posee la herramienta como si de palabras simples se tratase.

## EmoTraductor 1.0: una aplicación web para el mercado emocional

**RESUMEN:** En este capítulo vamos a presentar la aplicación web EmoTraductor 1.0, y sus mecanismos de marcado emocional, que son los puntos de partida del TFM presentado en este documento.

El EmoTraductor 1.0 es una aplicación web, desarrollada en 2018 en la Facultad de Informática de la UCM (Eugercios et al., 2018), que surgió de la necesidad de ayudar a interpretar el contenido emocional de un texto a personas que pudieran tener problemas para hacerlo, por ejemplo a personas con Síndrome de Asperger. El objetivo de la herramienta era detectar automáticamente la carga afectiva que tiene un texto y representarla con las cinco emociones básicas: alegría, tristeza, miedo, asco y enfado. El resultado era mostrado en la interfaz de la aplicación mediante el uso de colores y emoticonos. La herramienta fue evaluada con la colaboración de la Asociación Asperger de Madrid. Como resultado de esta evaluación se vio que aún quedaba trabajo por realizar en cuanto a la interpretación del contenido emocional, pero el trabajo que se hizo era suficiente como para establecer una base sólida que se podría ir mejorando.

El EmoTraductor es una aplicación web creada sobre una API de servicios web para el mercado emocional. Podemos ver un esquema muy básico de su arquitectura en la Figura 3.1. En las próximas secciones daremos más detalle de cada una de las partes.

### 3.1. Aplicación web del EmoTraductor 1.0

La versión 1.0 del EmoTraductor está disponible en <http://sesat.fdi.ucm.es/traductor/>. En la Figura 3.2 se muestra el aspecto de la página inicial de la aplicación. Se puede observar que a la izquierda y debajo del nombre *EmoTraductor* aparece el cuadro donde el usuario debe introducir la palabra, frase o texto que quiere marcar emocionalmente. Una vez introducido el texto, debemos pulsar en el botón Traducir. El resultado del marcado emocional aparecerá abajo con una barra emocional, y a la derecha donde aparecen destacadas las palabras emocionales del texto (ver Figura 3.3). La barra emocional es muy intuitiva y marca las emociones predominantes en el texto introducido. La porción coloreada de la barra de cada emoción es equivalente al peso que tiene en el texto introducido dicha emoción. Por ejemplo, si nos fijamos en la Figura 3.3, vemos que el texto introducido tiene como emociones asociadas la tristeza en mayor medida y el miedo en menor medida.

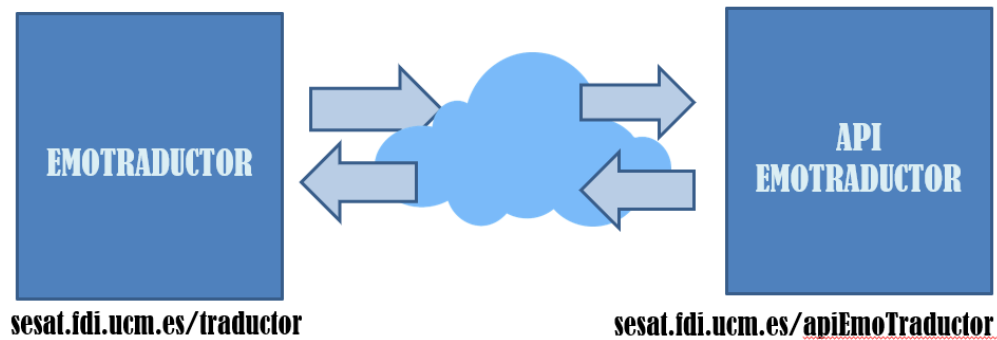


Figura 3.1: Esquema de la arquitectura del EmoTraductor

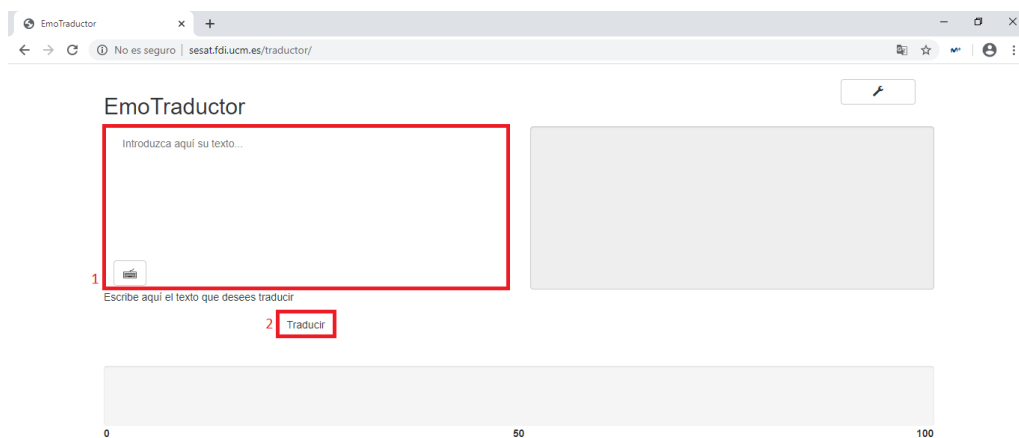


Figura 3.2: Página principal EmoTraductor 1.0

Una manera de conocer el porqué de las emociones resultantes del análisis realizado por la herramienta es a través de las pistas sobre las palabras emocionales del texto y su emoción asociada. Las palabras en negrita que aparecen en el cuadro de la derecha (sobre el mismo texto escrito por el usuario) indican las palabras que contribuyen a determinar la emoción del texto (son las llamadas emociones principales). Al situar el ratón sobre una de estas palabras en negrita, podremos ver qué emociones tiene esa palabra tal y como podemos ver en la Figura 3.4.

Como una medida alternativa y para mejorar la accesibilidad, existe la posibilidad de cambiar el color asociado a las emociones para que las personas con sinestesia<sup>1</sup> puedan adaptar los colores de las emociones a sus preferencias. Si pulsamos la llave inglesa de la parte superior derecha de la pantalla, aparecerá un menú donde solo se puede pulsar uno de los colores asociados a las emociones y cambiarlo.

<sup>1</sup>Unión de dos imágenes o sensaciones procedentes de diferentes dominios sensoriales, como relacionar una emoción con un color

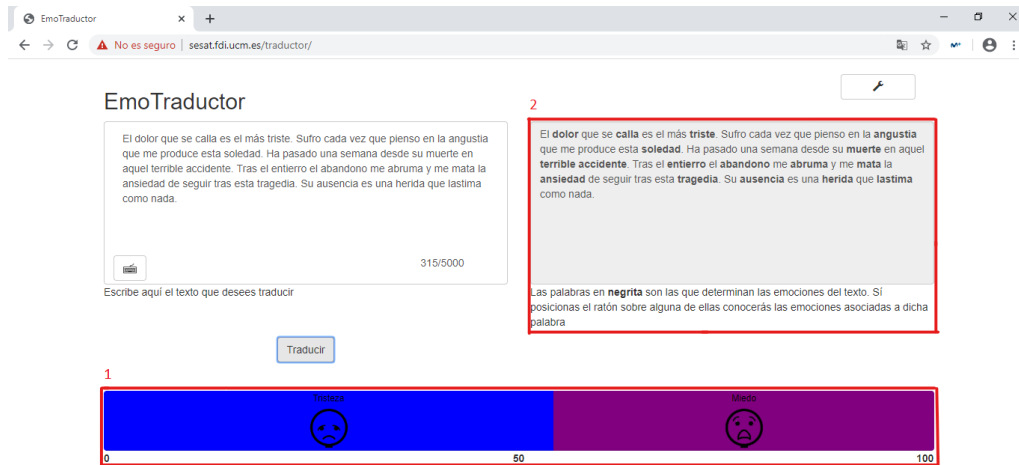


Figura 3.3: Resultados del traducción

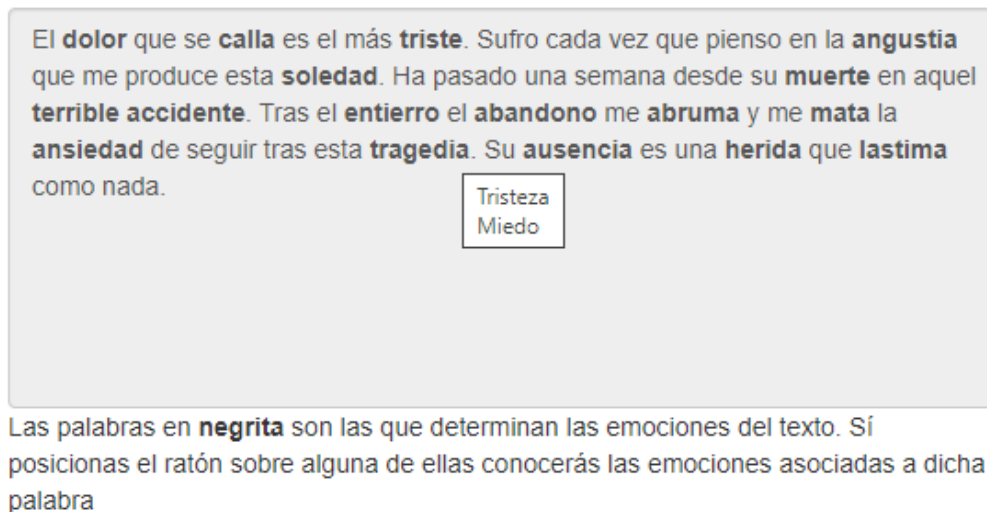


Figura 3.4: Resultado de colocar el ratón sobre la palabra «tragedia»

## 3.2. API de servicios web del EmoTraductor 1.0

Para que la aplicación pueda cumplir su objetivo, el marcado emocional de textos, se crearon distintos servicios web accesibles vía API<sup>2</sup> para su funcionamiento. En las siguientes secciones daremos más información de los servicios web implementados para el EmoTraductor.

<sup>2</sup><http://sesat.fdi.ucm.es/apiEmoTraductor/>

### 3.2.1. Servicios para el análisis emocional de una palabra

Estos servicios se basan en el diccionario emocional que se encuentra en la base de datos de la aplicación. El diccionario emocional, es el elemento fundamental para el análisis emocional de los textos, está compuesto por los diccionarios de Ferré (2017) y de Hinojosa (2016) que vimos en las Secciones 2.4.6 y 2.4.7 respectivamente. Fusionando ambos diccionarios, tenemos un diccionario con 3.141 palabras. Cada una de ellas tiene asociado su lexema y los grados de las categorías emocionales (tristeza, alegría, miedo, enfado y asco) con valores entre el 1 y el 5, donde 1 es la ausencia de la emoción y el 5 la presencia plena de esta. Los autores de los diccionarios originales consideran que una palabra es emocional si el grado de, al menos, una categoría emocional es mayor de 2,5.

Los servicios más importantes de esta parte de la API son:

- Servicio de grados emocionales: Este servicio devuelve, dada una palabra, los grados emocionales para cada una de las categorías emocionales. Para consultar los grados emocionales se debe usar la siguiente petición, por método GET:

```
http://sesat.fdi.ucm.es/emociones/palabra/gradosEmo?palabra=pBuscada
```

donde *pBuscada* será la palabra que queramos evaluar. La respuesta será enviada en JSON y contendrá cada una de las categorías emocionales con su grado asociado. Podemos ver como ejemplo, el resultado de la petición

```
http://sesat.fdi.ucm.es/emociones/palabra/gradosEmo?palabra=abejas
```

en la Figura 3.5, donde la palabra “abejas” tiene un grado de 1,17 en tristeza, 3,47 en miedo, 1,37 en alegría, 2,0 en enfado y 2,2 en asco. En caso de que la palabra no existiese en el diccionario, para cada una de las categorías emocionales devolvería un valor especial indicando la ausencia de la palabra en el diccionario.

```
{
  "Tristeza": "1.17",
  "Miedo": "3.47",
  "Alegría": "1.37",
  "Enfado": "2.0",
  "Asco": "2.2",
}
```

Figura 3.5: Resultado servicio grados de la palabra de la palabra abejas

Cuando la petición de grados emocionales llega al sistema, lo primero que se hace es obtener el lema de la palabra, utilizando la librería PyStemmer<sup>3</sup>, para acudir al diccionario emocional donde buscar los grados emocionales que le corresponden. Se utiliza este lema para tener en consideración cualquier palabra derivada de la que aparece en el diccionario, incrementando el número de palabras que podemos

<sup>3</sup><https://pypi.org/project/PyStemmer/>

clasificar. Por ejemplo, la palabra “cantó” no aparece en el diccionario y es una forma verbal de “cantar”, cuyo lema es “cant”. Al buscar usando el lema y no la palabra podemos asociar los valores de la palabra “cantar” a “cantó”. Lo mismo sucede con el verbo “cenar” que no aparece en el diccionario y cuyo lema es “cen”. Como sí aparece la palabra “cena” cuyo lema también es “cen” podemos asociar la información emocional del diccionario de la “cena” al verbo “cenar”.

Como ejemplo, con la palabra “alegría”, que tiene como lema “alegr”, se consulta el diccionario resultando lo que podemos ver en la Tabla 3.1.

Tristeza	Miedo	Alegría	Enfado	Asco
1,1	1,23	4,83	1,1	1,03

Tabla 3.1: Resultado servicio grados palabra Emotraductor 1.0 con la palabra “alegría”

Este servicio tiene algunos problemas en caso de que los lemas de las palabras estén repetidos, haciendo que el EmoTraductor devuelva un error y nunca se pueda analizar la palabra enviada. Esto ocurriría en palabras como “tímido” y “timado” donde el lema de cada palabra es “tim” y por tanto el servicio no sabría cuál de los dos debe devolver para obtener los grados correctos. También pasaría por ejemplo con las palabras “tontería” y “tonto” cuyo lema es “tont”. Esto puede ocurrir porque existen en el diccionario varias palabras que se forman a partir del mismo lema y, como siempre se busca primero el lema de la palabra sin tener en cuenta la palabra original, encontraremos lemas repetidos.

- Servicio de emoción mayoritaria: Este servicio tiene como finalidad devolver la categoría emocional con mayor grado de la palabra enviada, junto con el grado con el que lo consigue. Esta petición se hace mediante método GET:

`http://sesat.fdi.ucm.es/emociones/palabra/mayoritariaEmo?palabra=pBuscada`

donde *pBuscada* será la palabra que queramos evaluar. El resultado será un JSON que contiene un listado de emociones, con entre 1 y 5 emociones, que serán las mayoritarias y el grado con el que esta emoción o emociones son consideradas mayoritarias. Por ejemplo, si realizamos la petición

`http://sesat.fdi.ucm.es/emociones/palabra/mayoritariaEmo?palabra=abejas`

para la palabra “abejas” obtendremos que la emoción mayoritaria es el miedo con un 3,47 como grado emocional (ver Figura 3.6). En caso de que una palabra tuviera más de una emoción con el grado más alto, el servicio devolvería el grado junto con un listado de las emociones que poseen este grado. Podemos ver un ejemplo en la Figura 3.7 que indica que las emociones mayoritarias son el miedo y la alegría con un 3,47 como grados emocionales (este caso no existe en el diccionario del EmoTraductor 1.0).

Para obtener el resultado este servicio, una vez recibida la palabra, se obtiene el lema de la palabra usando PyStemmer y será este lema el que se busque en el diccionario emocional para obtener el valor de los grados. Estos valores se evaluarán para encontrar el mayor y se formará la respuesta JSON esperada.

```
{
  "grado": 3.47,
  "emociones": ["Miedo"]
}
```

Figura 3.6: Resultado servicio emoción mayoritaria de la palabra “abejas”

```
{
  "grado": 3.47,
  "emociones": [ "Miedo", "Alegria" ]
}
```

Figura 3.7: Resultado servicio emoción mayoritaria en caso de que tenga más de una emoción mayoritaria

### 3.2.2. Servicios para el análisis emocional de una frase

Estos servicios serán la base para los que analizan emocionalmente texto completos y a su vez se basan en los servicios de análisis emocional de palabras.

- Servicio de grados emocionales: Este servicio devuelve los grados emocionales para cada categoría emocional de la frase dada. Para consultar los grados emocionales se debe usar la siguiente petición, por método POST:

`http://sesat.fdi.ucm.es/emociones/frase/gradosEmo`

Al tratarse de un método POST, se debe incluir como parámetro de envío en la petición la variable obligatoria *frase* que será la que queramos evaluar (la variable obligatoria no tiene nada que ver con la palabra “frase” que aparece en la URL de la petición). La respuesta será enviada en JSON y contendrá cada una de las categorías emocionales con su grado asociado. En la Figura 3.8, podemos ver los resultados para la frase “Me gustan las naranjas”: 1,17 grados en tristeza, 1,16 en miedo, 4,43 en alegría, 1,2 en enfado y 1,17 en asco. En caso de que la frase no tenga emociones en la respuesta JSON cada categoría aparecerá con un valor especial reservado.

Como se indica al principio de esta misma sección, este servicio utiliza como base el de grados emocionales de una palabra ya explicado en la sección 3.2.1. Nada más llegar la frase al servicio, se usa el analizador sintáctico de Spacy<sup>4</sup>, que la evalúa para detectar la categoría gramatical de cada palabra. Según la categoría detectada se decide si es una palabra *candidata* a ser emocional. Consideramos una palabra *candidata* a ser emocional a la que tiene como categoría morfosintáctica verbo, adjetivo o sustantivo. Una vez localizadas las *candidatas* se envían una a una al servicio de **grados emocionales de una palabra** explicado en la Sección 3.2.1. Una vez este servicio devuelve el resultado, consideraremos que una palabra es emocional si alguno de sus grados supera el 2,5 en alguna de sus categorías y dejarán de ser *candidatas* para ser consideradas emocionales, que quiere decir que aportarán su carga emocional a la frase. Con los valores obtenidos para las palabras emocionales, se sumarán los

<sup>4</sup><https://spacy.io/>

```

{
  "Tristeza": "1.17",
  "Miedo": "1.6",
  "Alegria": "4.43",
  "Enfado": "1.2",
  "Asco": "1.17"
}

```

Figura 3.8: Resultado del servicio de grados emocionales de la frase para la frase “Me gustan las naranjas”

grados y se dividirá entre el número de palabras consideradas emocionales, haciendo la media. El resultado de esta media será el resultado que devolverá el servicio.

Veamos con un ejemplo la funcionalidad del servicio. Enviamos la frase “Estoy alegre y feliz” y Spacy analizará la frase y nos catalogará las palabras “alegre” y “feliz” como candidatas para ser emocionales por ser adjetivos, mientras que las palabras “estoy” (los verbos copulativos son considerados auxiliares y no verbos) e “y” las descartará. Se enviarán las palabras candidatas al servicio de grados emocionales, que nos devolverá los valores de los grados de ambas palabras que podemos ver en la Tabla 3.2 en las dos primeras filas. Con estos valores, vemos que ambas palabras son emocionales, porque superan el 2,5 en la categoría alegría y por tanto ambas colaborarán para decidir los grados emocionales de la frase. A continuación, se sumarán los valores, como vemos en la Tabla 3.2 en la tercera fila, y se dividirá entre el número de palabras emocionales, en este caso 2, dando como resultados los que vemos en la última fila de la Tabla 3.2.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
alegre	1,1	1,23	4,83	1,1	1,03
feliz	1,1	1,2	4,63	1,0	1,0
Suma	2,2	2,43	9,46	2,1	2,03
NºPalabras	2				
Valor final frase	1,1	1,21	4,73	1,05	1,02

Tabla 3.2: Resultado del servicio de grados emocionales de un frase con la frase “Estoy alegre y feliz”

Este servicio tiene como limitación que no se consideran la negación ni los modificadores. Al no tener en cuenta este tipo de elementos, en muchas frases el resultado es incorrecto. Por ejemplo, en la frase “No tengo miedo”, el EmoTraductor devolverá como emoción mayoritaria el miedo, cuando en realidad la frase está indicando todo lo contrario. En el caso de los modificadores pasará algo similar. Si evaluásemos la frase “Tengo poco miedo” y la comparásemos con la frase “Tengo mucho miedo” deberíamos notar en los resultados que la primera tiene menos grados en la categoría emocional miedo y la segunda debe tener resultados mayores; sin embargo, los resultados serán iguales en este caso.

- Servicio de emoción mayoritaria: Esta petición devuelve la categoría emocional con mayor grado emocional en la frase. Para consultar la emoción mayoritaria de una frase se debe usar la siguiente petición, por método POST:

`http://sesat.fdi.ucm.es/emociones/frase/mayoritariaEmo`

Al tratarse de un método POST, se debe incluir como parámetro de envío en la petición la variable obligatoria *frase* que será la que queramos evaluar (la variable obligatoria no tiene nada que ver con la palabra “frase” que aparece en la URL de la petición). La respuesta será enviada en JSON y contendrá, como sucedía en el servicio de emoción mayoritaria para palabras, el grado con el que es considerada la emoción como mayoritaria y un listado con la emoción o emociones que poseen ese grado asociado. Podemos ver en la Figura 3.9 el resultado de la frase “Me gustan las naranjas”, donde la alegría es la emoción mayoritaria con un 4,43. En caso de que la frase no tenga emociones en la respuesta JSON aparecerá la lista e categorías emocionales vacía.

```
{
  "grado": 4.43,
  "emocion": ["Alegría"]
}
```

Figura 3.9: Resultado del servicio de emoción mayoritaria para la frase “Me gustan las naranjas”

Este servicio se basa en el explicado anteriormente (servicio de grados emocionales de una frase), dado que para decidir qué emoción es mayoritaria debemos obtener primero los grados emocionales de la frase completa. Una vez obtenida la respuesta, se evalúa y se decide qué categoría emocional tiene mayor grado asociado.

### 3.2.3. Servicios para el análisis emocional de un texto

Los servicios del bloque texto están creados para tratar un texto compuesto por una o más frases, y usan como base los servicios de tratamiento de una frase explicados en la Sección 3.2.2.

- Servicio de grados emocionales: Este servicio devuelve los grados emocionales para cada categoría emocional del texto. Para consultar los grados emocionales de un texto se debe usar la siguiente petición, por método POST:

`http://sesat.fdi.ucm.es/emociones/texto/gradosEmo`

Se debe incluir como parámetro de envío en la petición POST la variable obligatoria *texto* que será el que queramos evaluar (la variable obligatoria no tiene nada que ver con la palabra “texto” que aparece en la URL de la petición). La respuesta será enviada en JSON y contendrá, para cada categoría emocional su grado asociado. Podemos ver en la Figura 3.10 el resultado del texto “Salir a la calle es divertido. Ir al circo es divertido”, donde podemos ver que tiene como grados: 1,34 de tristeza, 1,41

de miedo, 4,05 de alegría, 1,11 de enfado y 1,18 de asco. En caso de que el texto no tenga emociones en la respuesta JSON cada categoría aparecerá con un valor especial reservado para ello.

```
{
  "Tristeza": "1.34",
  "Miedo": "1.41",
  "Alegría": "4.05",
  "Enfado": "1.11"
  "Asco": "1.18",
}
```

Figura 3.10: Resultado del servicio de grados emocionales de textos para “Salir a la calle es divertido. Ir al circo es divertido”

Para obtener el resultado, se comienza con la separación del texto en distintas frases utilizando los signos de puntuación. Una vez divididas se analiza el tipo de frase: enunciativa, exclamativa o interrogativa. Dependiendo del tipo de frase las emociones de estas tendrán más peso o menos en el resultado final: las enunciativas tendrán el valor resultado obtenido por el servicio de grados emocionales de una frase, las exclamativas tendrán el doble de ese valor y las interrogativas la mitad. Una vez analizado el tipo de frase se envían al servicio de grados emocionales de frase. Cuando se obtienen las respuestas del servicio de frases para todas las frases que contenga el texto, se procede a la ponderación por tipo de frase, a la suma de los valores y a hacer la media con el número de frases evaluadas. Veamos un ejemplo utilizando el texto: “Ayer vi a tu hermano. ¡Qué guapo está! ¿Cuándo ha vuelto de Londres?”. El servicio dividiría este texto en tres frases distintas: “Ayer vi a tu hermano”, “¡Qué guapo está!” y “¿Cuándo ha vuelto de Londres?”, las clasificaría como enunciativa, exclamativa e interrogativa respectivamente, y las enviaría al servicio obteniendo el resultado para cada una de ellas que podemos ver en la Tabla 3.3 en las tres primeras filas. A continuación, sumaría los valores teniendo en cuenta las ponderaciones explicadas, que podemos ver en la cuarta fila de la Tabla 3.3 y después haría la media utilizando el número de frases evaluadas (en este caso 3), obteniendo el resultado que vemos en la última fila de la tabla.

	Tristeza	Miedo	Alegría	Enfado	Asco	Peso
Frase 1 (Enunciativa)	1,47	1,1	4,3	1,47	1,1	x1
Frase 2 (Exclamativa)	1,12	1,27	3,77	1,17	1,07	x2
Frase 3 (Interrogativa)	1,0	1,0	1,0	1,0	1,0	/2
Suma	4,21	4,14	11,84	4,31	3,74	
NºFrases	3					
Valor final frase	1,41	1,38	3,95	1,44	1,25	

Tabla 3.3: Resultado del servicio de grados emocionales para textos del Emotraductor 1.0

- Servicio de palabras emocionales: Este servicio devuelve la lista de las palabras emocionales de un texto junto con los grados emocionales de cada una. Para encontrar las palabras emocionales de un texto se debe usar la siguiente petición, por método POST:

`http://sesat.fdi.ucm.es/textosGuay/`

Se debe incluir como parámetro de envío en la petición POST la variable obligatoria *texto* que será el que queramos evaluar.

La respuesta será enviada en JSON y contendrá dos atributos. El primero será la evaluación total de las emociones en el texto, donde en una lista se mostrarán los mismos resultados que se muestran para el servicio de grados emocionales de un texto. El segundo será un listado de todas las palabras que intervienen para conseguir este resultado. En la Figura 3.11 podemos ver el resultado de analizar el texto “El pequeño cocodrilo. Para aprender sus cantares usa las aguas del Nilo con sus notas musicales.” donde vemos la lista de valores de cada emoción en orden tristeza, miedo, alegría, enfado y asco, precedida por la palabra “emociones”, y el listado de palabras emocionales en “palabras”.

```
{
  "emociones": ["1.42", "2.38", "2.74", "1.28", "1.41"]
  "palabras": ["pequeño", "cocodrilo", "entonar",
  "cantares", "usa", "aguas", "nilo", "notas", "musicales"]
}
```

Figura 3.11: Resultado del servicio de grados emocionales de textos para “El pequeño cocodrilo. Para aprender sus cantares usa las aguas del Nilo con sus notas musicales.”

Para obtener estos resultados se ejecuta el servicio de grados emocionales pero se almacena en una lista cada una de las palabras analizadas en cada una de las frases del texto resultando el listado de palabras que se devuelve.

- Servicio de emoción mayoritaria: Este servicio devuelve la categoría emocional y el grado, de la categoría con mayor grado asignada en el texto. Para consultar la emoción mayoritaria de un texto se debe usar la siguiente petición, por método POST:

`http://sesat.fdi.ucm.es/emociones/texto/mayoritariaEmo`

Se debe incluir como parámetro de envío en la petición POST la variable obligatoria *texto* que será el que queramos evaluar (la variable obligatoria no tiene nada que ver con la palabra “texto” que aparece en la URL de la petición).

La respuesta será enviada en JSON y contendrá, como sucedía en el resto de servicios de este tipo, el grado con el que es considerada la emoción como mayoritaria y un listado con la emoción o emociones que poseen ese grado asociado. Podemos ver en la Figura 3.12 el resultado del texto “Me gusta ir al parque. Es muy divertido”, donde la alegría es la emoción mayoritaria con un 4,39.

Este servicio se basa en el explicado anteriormente (grados emocionales de un texto) dado que para decidir qué emoción es mayoritaria debemos obtener primero los grados emocionales del texto completo. Una vez obtenida la respuesta, se evalúa y decide qué categoría emocional tiene mayor grado asociado.

```
{  
  "grado": "4.39",  
  "emociones": [ "Alegría"]  
}
```

Figura 3.12: Resultado servicio emoción mayoritaria del texto para el texto “Me gusta ir al parque. Es muy divertido”

### 3.3. Resultado de la evaluación de la herramienta

Esta primera versión del EmoTraductor fue evaluada con la ayuda de la Asociación Asperger de Madrid en un encuentro en la Facultad de Informática, donde se evaluó con usuarios reales el marcado emocional de textos y la usabilidad de la interfaz diseñada.

#### 3.3.1. Evaluación del marcado emocional

Esta evaluación se realizó con la finalidad de ver la corrección del etiquetado emocional del sistema. Para ello se utilizó un conjunto de noticias, blogs y cuentos que se pueden encontrar en el Apéndice A y que usaremos posteriormente para seguir evaluando la herramienta con el fin de poder comparar resultados.

El resultado de la evaluación fue que la herramienta era capaz de marcar todos los textos pero que era esencial tener en cuenta las negaciones para poder mejorar la clasificación que hace el sistema, ya que en muchos textos las emociones que se consideraban mayoritarias no eran las correctas sino más bien la contraria. También se vio que el límite de 2,5 que se usaba para decidir cuándo una emoción era considerada como significativa era bastante interesante ya que permitía descartar rápidamente cuáles de las emociones no lo eran. Por último, se descubrió que casi todos los resultados obtenidos con estos textos tendían a la alegría y que quizá hubiera sido mejor otro tipo de textos para probar el marcado ya que la mayoría de los textos elegidos eran cuentos infantiles con un final feliz.

#### 3.3.2. Evaluación de la aplicación web

Como se decía en la introducción, se pudo realizar gracias a la Asociación Asperger Madrid una evaluación con usuarios finales. Esta evaluación tenía como objetivo conocer la opinión en cuanto a utilidad y usabilidad de usuarios finales, y no evaluar el método de marcado ya que se había evaluado de otra manera y porque los usuarios finales de la aplicación no podían decir si los resultados eran los correctos o no (al tener todos Síndrome de Asperger).

Esta evaluación consistió en tres partes: una primera con datos demográficos para conocer a los usuarios, una segunda que consistía en evaluar distintos textos usando la aplicación y contestar a una serie de preguntas sencillas sobre la interpretación de los resultados y una tercera para indicar cómo de usable les resultaba la aplicación.

Como resultado de esta evaluación podríamos decir que:

- Los colores y los emoticonos usados por defecto en la aplicación son correctos. Y es útil la opción de personalizar los colores.

- El diseño de la barra emocional que muestra el resultado de la traducción emocional es intuitiva.
- Mejorando el método de marcado el Emotraductor puede ser una herramienta útil en el día a día del colectivo objetivo, pero se debería aumentar el corpus de palabras identificadas y tener en cuenta expresiones coloquiales, modificadores y negaciones.

### **3.3.3. Posibles mejoras sobre el EmoTraductor 1.0**

Se presentaron las siguientes tareas como posibles para mejorar la aplicación en su conjunto. Las vamos a dividir entre las que afectarían al mercado emocional y las que afectan a la aplicación web.

#### **3.3.3.1. Mercado emocional**

- Añadir el reconocimiento de más tipos de frases como negativas o subordinadas y no únicamente reconocer las exclamativas, interrogativas y afirmativas.
- Insertar más vocabulario para el marcado, incluyendo frases hechas y expresiones típicas, al diccionario.
- Creación de un log para poder aprender de este y añadir nuevas palabras de los textos más usados y poder depurar errores y mejorar el marcado.

#### **3.3.3.2. Aplicación web**

- Añadir la posibilidad de cambiar las imágenes asociadas a las emociones.
- Detectar automáticamente los colores oscuros en la barra emocional para evitar problemas con las imágenes y el texto.
- Añadir la introducción de texto por voz para no escribir los textos y que se pueda usar la aplicación en cualquier ámbito.
- Implementar una aplicación móvil para que el uso del EmoTraductor se extienda.
- Actualizar la interfaz para que cumpla un nivel de conformidad AAA.

## Primeros pasos hacia el EmoTraductor 2.0

**RESUMEN:** En este capítulo vamos a explicar cada una de las mejoras preliminares que se han realizado en los servicios de marcado emocional que usa el EmoTraductor 1.0 y realizaremos una evaluación para saber si los cambios realizados se traducen en una mejora de la cobertura de este sistema.

En la primera Sección de este capítulo vamos hacer una detección de las mejoras que necesita el EmoTraductor 1.0. A continuación, cada una de las siguientes secciones tratará una mejora preliminar de los servicios de marcado emocional que usa el EmoTraductor 1.0. Después de las mejoras se han evaluado los resultados usando los textos que se emplearon en el trabajo original (Eugercios et al., 2018) junto con una nueva colección de textos seleccionados para poder evaluar los cambios con textos que presentan características que fomentan que la aplicación de los cambios tengan efectos en sus resultados. Por ejemplo se han añadido textos que utilizan negaciones, modificadores, frases de distintos tipos y léxico más rico.

### 4.1. Detección de mejoras preliminares

Tras la evaluación del EmoTraductor 1.0 realizada, al comienzo de este trabajo se ha realizado otra encontrando algunos puntos que se pueden mejorar. A continuación se explican las mejoras identificadas durante las dos evaluaciones realizadas:

- En los servicios de grados emocionales, cuando lo evaluado no tiene resultados de presencia de una emoción se utiliza el 1 ya que en el diccionario afectivo que utilizamos el 1 es la ausencia de la emoción. Esto hace que en algunas ocasiones los resultados no sean adecuados, ya que cuando una palabra no tiene una emoción incrementamos en 1 el valor de dicha emoción en la frase que estamos evaluando cuando en realidad no deberíamos de aportar ningún valor. Sin embargo, cuando nuestro sistema está evaluando una frase que no tiene ninguna palabra emocional devuelve un 0 indicando la ausencia de emoción en la frase, lo cual no es coherente con lo anterior. En la Sección 4.2 explicaremos cómo hemos solventado esta falta de coherencia para obtener mejores resultados.
- Uno de los puntos débiles detectado en la evaluación con usuarios finales del EmoTraductor 1.0 era que el diccionario emocional que se usaba se quedaba corto y había

muchas palabras cotidianas que no se tenían en cuenta para determinar la carga afectiva de un texto porque no se encontraban en nuestro diccionario. Esta mejora se tratará en la Sección 4.3.

- En el servicio de grados emocionales de una frase, se considera que una palabra es emocional si alguno de los grados de sus categorías emocionales supera un valor límite establecido (2,5). Al poner un límite no se tienen en cuenta todas las palabras que obtenemos del diccionario cuyos valores no superan el límite, perdiendo la contribución de estas palabras en el resultado final de la carga afectiva del texto. Esta mejora se tratará en la Sección 4.4.
- En el servicio de grados emocionales de un texto, que vimos en la Sección 3.2.3, había un error al determinar el número de frases contenidas en un texto (denominador empleado para obtener la carga emocional final de un texto) ya que en lugar de sumar uno para cada frase se sumaba el peso de la frase en la carga emocional. Esta forma de proceder era incorrecta y hacía que los resultados no fueran buenos. En la Sección 4.5 explicamos cómo hemos corregido este error.

## 4.2. Normalización de los valores emocionales del diccionario

Como ya se explicó en la Sección 4.1, las palabras del diccionario emocional que utiliza el EmoTraductor 1.0 usan una escala del 1 al 5, donde el 1 muestra la carencia de la emoción en la palabra y el 5 la máxima presencia de esta. Sin embargo, al probar con cuidado la herramienta, nos damos cuenta de que en ocasiones el hecho de que la carencia de emoción sea un 1 hace que dicha emoción aumente en 1 el grado emocional de la emoción en la frase. Por ejemplo, en la frase “Me gusta abrazar a la chica”, las palabras candidatas a ser emocionales son “gusta”, “abrazar” y “chica” ya que todas aquellas tienen alguna de sus emociones con un grado superior a 2,5. Como podemos ver en la Tabla 4.1, las tres palabras candidatas superan el 2,5 en la categoría emocional alegría y por ello son palabras emocionales alegres. Como hemos visto en la Sección 3.2.2, el servicio de grados emocionales de una frase suma todos los grados de las palabras emocionales y divide entre el número de palabras. En este caso, donde “abrazar” tiene dos emociones (enfado y asco), con un 1 en el diccionario (que representa la ausencia de esta emoción) está sumando grados en estas categorías cuando en realidad no debería aportar nada por no tener presencia. En la Tabla 4.1 en la fila final vemos el resultado que se da a la frase según el EmoTraductor 1.0 donde vemos que al enfado y al asco se le ha sumado un grado emocional más por la palabra abrazar cuando en realidad el 1 significa la ausencia de la emoción.

La solución que proponemos a este problema es cambiar todo el diccionario para que los valores posibles de cada emoción vayan del 0 al 4 en lugar de hacerlo de 1 a 5. A esto lo llamaremos normalizar. En la Tabla 4.2 vemos la comparativa del marcado original del EmoTraductor 1.0 con el resultado obtenido tras normalizar el diccionario. Aparentemente parece que las emociones sufren un cambio significativo, pero teniendo en cuenta que variamos la escala, los cambios no son muy significativos ya que podríamos decir que hemos restado un 1 a los resultados. Sin embargo, si esta frase no fuese enunciativa sino que fuese exclamativa por ejemplo, el valor final de esta frase habría que multiplicarlo por 2 (ya que es el peso que damos a las frases exclamativas) haciendo que la penalización por tener 1 como límite inferior sea mucho mayor. Vamos a ver esto con un ejemplo. Supongamos que tenemos el texto: “¡Me gusta abrazar a la chica! Me gusta mucho.” Podemos ver en la Tabla 4.3, el resultado de la evaluación del texto. Vemos que si tenemos marcada la ausencia de la emoción con un 1 los resultados del enfado (1,76) y el asco (1,72) son bastante superiores

que si tenemos marcada la ausencia de la emoción con un 0 donde los grados de enfado (0,26) y asco (0,23) son menores. Esto se debe a que en este caso debíamos ponderar las emociones de la frase exclamativa “¡Me gusta abrazar a la chica!” multiplicando por dos las emociones que aparecen, tras haber normalizado el diccionario no será lo mismo multiplicar por dos un 1,16 que teníamos en la frase como enfado en el EmoTraductor 1.0 (ver en la Tabla 4.2) que multiplicar dos por 0,16 que tendríamos como resultado tras normalizar el diccionario (ver en la Tabla 4.2). Es por esto que es importante la realización de la normalización, no tanto para las oraciones enunciativas sino para las frases exclamativas o interrogativas que ponderamos.

Elemento	Tristeza	Miedo	Alegría	Enfado	Asco
gusta	1,17	1,6	<b>4,43</b>	1,2	1,17
abrazar	2	1,63	<b>4,73</b>	1	1
chica	1,38	1,34	<b>2,97</b>	1,28	1,24
Me gusta abrazar a la chica	1,51	1,52	4,04	1,16	1,14

Tabla 4.1: Grados emocionales de las palabras y de la frase “Me gusta abrazar a la chica”

Caso	Tristeza	Miedo	Alegría	Enfado	Asco
EmoTraductor 1.0	1,5	1,52	4,04	1,16	1,14
Diccionario normalizado	0,52	0,52	3,04	0,16	0,14

Tabla 4.2: Comparativa de los resultados grados emocionales de “Me gusta abrazar a la chica”

Caso	Tristeza	Miedo	Alegría	Enfado	Asco
EmoTraductor 1.0	2,09	4,86	6,25	1,76	1,72
Diccionario normalizado	1,21	0,82	4,75	0,26	0,23

Tabla 4.3: Comparativa de los resultados de los grados emocionales de “¡Me gusta abrazar a la chica! Me gusta mucho.”

Veamos otro ejemplo con la frase “Cuando habla el tío odioso solamente pienso en abrazar perros”. Las palabras candidatas a ser emocionales son: “habla”, “tío”, “odioso”, “pienso”, “abrazar” y “perros”. De estas, como podemos ver en la Tabla 4.4 con los valores para cada palabra según el EmoTraductor 1.0, solo son emocionales “tío”, “odioso”, “abrazar” y “perros”. Vemos también que la palabra “odioso” tiene ausencia de alegría y “abrazar” de enfado y asco. En la Tabla 4.4 podemos ver cómo calcula el EmoTraductor 1.0 el valor emocional de la frase teniendo en cuenta en la suma de grados emocionales los unos que indican ausencia de emoción (de alegría en la palabra “odioso” y enfado y asco en la palabra “abrazar”) estarán hinchando emociones a las que no les corresponden un grado tan alto. Siguiendo el mismo ejemplo en la Tabla 4.5 vemos cómo quedarían las palabras de la frase según el diccionario normalizado y en la última fila el resultado final de la frase en la versión que usa el diccionario normalizado. Si comparamos las últimas filas de las Tablas 4.4 y 4.5, una vez más vemos que los cambios no son muy significativos, pero si utilizamos esta frase en un texto con ponderación los resultados volverá a variar enormemente. En la

Tabla 4.6, vemos el resultado del texto “¡Cuando habla el tío odioso solo pienso en abrazar perros! Eso me relaja.” donde vemos un cambio significativo ya que pasaríamos tener como significativas las emociones de tristeza y alegría y muy cercanas a serlo el miedo y el enfado mientras que con los resultados usando el diccionario normalizado tenemos que la alegría destaca sobre las demás notablemente.

Elemento	Tristeza	Miedo	Alegría	Enfado	Asco
habla	1	1	1	1	1
tío	1,5	1,23	<b>2,97</b>	1,43	1,3
odioso	<b>2,79</b>	2,24	1	<b>3,86</b>	<b>3</b>
pienso	1	1	1	1	1
abrazar	2	1,63	<b>4,73</b>	1	1
perros	1,2	1,5	<b>3,67</b>	1,17	1,33
Resultado de frase	1,87	1,65	3,09	1,86	1,65

Tabla 4.4: Grados emocionales de las palabras de “Cuando habla el tío odioso solamente pienso en abrazar perros” y su resultado total como frase según el EmoTraductor 1.0

Elemento	Tristeza	Miedo	Alegría	Enfado	Asco)
habla	0	0	0	0	0
tío	0,5	0,23	<b>1,97</b>	0,43	0,3
odioso	<b>1,79</b>	1,24	1	<b>2,86</b>	<b>2</b>
pienso	0	0	0	0	0
abrazar	1	0,63	<b>3,73</b>	0	0
perros	0,2	0,5	<b>2,67</b>	0,17	0,33
Resultado de frase	0,87	0,65	2,09	0,86	0,65

Tabla 4.5: Grados emocionales de las palabras de “Cuando habla el tío odioso solamente pienso en abrazar perros” y su resultado total como frase usando el diccionario normalizado

Caso	Tristeza	Miedo	Alegría	Enfado	Asco)
EmoTraductor 1.0	2,56	2,24	4,94	2,37	2,18
Diccionario normalizado	0,88	0,73	3,43	0,87	0,68

Tabla 4.6: Comparativa de los resultados del servicio de grados emocionales de “¡Cuando habla el tío odioso solo pienso en abrazar perros! Eso me relaja.”

En el ejemplo anterior, también podemos ver que palabras como “habla” o “pienso” que tienen ausencia de emoción por no aparecer en el diccionario en el EmoTraductor 1.0 no se están teniendo en cuenta para realizar la suma total de las emociones de la frase. Si estas emociones no son incluidas, tampoco se deben incluir en la suma los unos que tienen las palabras emocionales, como “odioso” en alegría, porque indican la carencia de la emoción en la palabra. Debemos conseguir que estos grados considerados como ausencia no aporten valor a la decisión final de la emoción y que como hemos visto ahora mismo solo ayudan a aumentar emociones que no nos interesan en vez de ayudar a diluir la emoción en la frase.

Por todo esto, decidimos que la opción más sencilla para que la ausencia de emoción

no aumente de manera incorrecta el grado de dichas emociones ausentes, será normalizar los grados del diccionario, de manera que la nueva escala sea del 0-4 en lugar del 1-5, dado que necesitamos reducir el valor de la ausencia de la emoción a 0 para evitar todos los problemas que hemos visto anteriormente con el aumento de las emociones ausentes. Dado que el EmoTraductor utilizaba el límite 2,5 para determinar si una emoción era significativa o no y para considerar si una palabra era emocional o no, al reducir el rango de valores a números entre 0 y 4 es necesario reducir el límite a uno equivalente. El límite 2,5 estaba escogido por los autores del diccionario emocional (Hinojosa et al., 2016; Ferré et al., 2017) por lo que vamos a hacer una equivalencia a la nueva escala, resultando el 1,5 como nuevo valor límite. Por lo tanto, ahora consideraremos que una palabra es emocional si alguno de sus grados es superior a 1,5.

Los resultados de este cambio se analizarán, junto con el resto de mejoras en la Sección 4.8

### 4.3. Ampliación del diccionario emocional

La evaluación del EmoTraductor 1.0 con usuarios reales puso de manifiesto la necesidad de ampliar el diccionario emocional ya que descubrimos que había muchas palabras de nuestro día a día que no se encontraban en el diccionario que usaba la herramienta. Aumentando el número de palabras contenidas en el diccionario afectivo aumentaremos la cobertura del traductor al disminuir el número de palabras no evaluadas (o mal evaluadas) por considerar palabras no emocionales a palabras que sí tienen carga emocional pero que no aparecen en el diccionario original. Para aumentar el diccionario original hemos incluido en él las palabras del diccionario de Stadthagen-González et al. (2018), pasando de tener de 3.141 palabras a tener 14.031 palabras. Este cambio debería de tener un impacto positivo en los resultados obtenidos para los textos del corpus de evaluación ya que ahora nuestro diccionario posee palabras tan comunes como: “camino”, “pensar” o “trabajo” que no se encontraban en el diccionario original.

También, se ha ampliado el diccionario añadiendo nuevos campos a las palabras contenidas en él. En concreto se han añadido los siguientes campos a cada palabra: categoría morfológica (nombre, nombre propio, verbo, adverbio o adjetivo), número (singular o plural) y género (femenino o masculino). Veremos la utilidad de esta segunda ampliación en la Sección 4.6.

### 4.4. Cambio en el criterio de selección de palabras emocionales

Otro de los cambios explorados para la nueva versión del EmoTraductor ha sido la disminución del límite para considerar cuándo una palabra candidata es realmente una palabra emocional. Anteriormente, antes de la normalización como indicábamos en el comienzo de la sección 3.2.1, se consideraba una palabra emocional si alguna de sus emociones superaba el 2,5 y tras la normalización, como contábamos al final de la sección 4.2, este límite fue cambiado a 1,5. Analizando los resultados de la clasificación emocional de algunas frases nos dimos cuenta de que este límite era demasiado alto y hacía que algunas palabras claramente emocionales (aunque fuera de forma débil) se quedaban fuera y no eran tenidas en cuenta a la hora de determinar la carga afectiva de una frase. Por ejemplo, en la frase “Me gustaría abrazar nubes” la palabra “nubes” aparece en el diccionario, con los grados que se muestran en la Tabla 4.7, pero ninguno de sus grados supera el límite de 1,5 y por

tanto sería una palabra candidata a ser emocional pero no llegaría a serlo y por ello sus grados no aportarían nada a la frase. Si estos grados sí aportasen carga emocional a la frase o texto al que pertenece podrían variar el grado de cada categoría. En la Tabla 4.7 vemos que la palabra “nubes” puede aportar 1,37 de alegría que es casi el límite establecido (1,5). Si tenemos en cuenta esta restricción muchas de las palabras nuevas que aporta la mejora del aumento del corpus es insignificante ya que existen palabras que no superan el umbral marcado. Además, una palabra con un valor bajo en alegría en el diccionario emocional como “nubes” significa que normalmente esta palabra no indica mucha alegría pero si está acompañada de palabras muy alegres nos ayudará a determinar que ese es el caso en que la palabra indica alegría. Por tanto, si eliminamos este límite, las nuevas palabras que no superen el límite para alguna emoción podrán considerarse emocionales y enriquecerán los resultados en el marcado emocional de los textos.

Vamos a ver un ejemplo en el que se ve claramente cómo eliminando el límite de marcado emocional obtenido es mucho más preciso. Tomamos como punto de partida la frase “Estaba preocupado con la situación”. En la Tabla 4.8 vemos los grados emocionales de las palabras candidatas a ser emocionales. En el caso de mantener el límite en el 1,5 únicamente aportarían emoción, es decir, se convertirían en emocionales, las palabras “estaba” y “preocupado” y despreciaríamos los grados emocionales de la palabra “situación” a pesar de que tiene un grado bastante cercano al 1,5 (1,37) en alegría. En la Tabla 4.9 vemos los resultados de la frase en caso de que tengamos en cuenta el límite establecido en 1,5 en la primera fila, donde la emoción significativa (por superar el 1,5) de la frase es el miedo, cuando la aparición de esta emoción no parece ser del todo correcta. Sin embargo, si nos fijamos en los resultados de la segunda fila, tras haber eliminado el límite del 1,5 para considerar una palabra como emocional y por tanto la palabra “situación” aporte su carga emocional a la frase, podemos ver que únicamente añadiendo a la suma de los grados totales de la frase los grados de esta palabra, el resultado final de la frase mejora ya que eliminamos el miedo como emoción significativa (superior al 1,5) y dejamos una frase donde el miedo está a punto de ser significativo pero no llega a serlo.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
nubes	0,47	0,3	1,37	0	0,07

Tabla 4.7: Grados emocionales de las palabra “nubes” normalizada

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
estaba	0,05	0,05	2,1	0	0,1
preocupado	2,73	2,97	0,1	1,07	0,4
situación	1,2	1,27	1,13	0,9	0,77

Tabla 4.8: Grados emocionales de las palabras candidatas a ser emocionales de “Estaba preocupado con la situación” normalizados

Situación	Tristeza	Miedo	Alegría	Enfado	Asco
Con límite 1,5	1,39	1,51	1,1	0,53	0,5
Sin límite	1,33	1,43	1,11	0,66	0,42

Tabla 4.9: Grados emocionales de la frase “Estaba preocupado con la situación”

## 4.5. Corrección en el servicio de grados emocionales de un texto

Como contábamos en la Sección 4.1 tras evaluar el servicio de los grados de los textos, hemos descubierto que no estaba funcionando de la manera en la que se explicaba y estaba diseñado, y que podemos ver en la sección 3.2.3. El funcionamiento era el siguiente: cuando se encontraba una frase enunciativa se contaba con un peso de unidad, en caso de interrogativa la mitad de la enunciativa y en caso de exclamativa se duplicaba el peso de la enunciativa. Hasta aquí todo es correcto. Una vez sumadas y ponderadas todas las emociones, se debía dividir entre el número de frases, pero en realidad se estaba dividiendo por la suma de ponderaciones de las frases. Esto quiere decir que si teníamos una frase enunciativa, una interrogativa y una exclamativa se dividían los grados emocionales ponderados entre 3,5 (1 enunciativa + 2 de la exclamativa + 0.5 de la interrogativa) en lugar de entre 3 que es el número de frases que aparecen. En la nueva versión de este servicio el error ha sido solventado y ya si que se divide por el número de frases.

## 4.6. Mejora en el servicio de obtención de grados emocionales de una palabra

En el EmoTraductor 1.0 el servicio que obtenía los grados emocionales de una palabra usaba el lema de la palabra que le llegaba como entrada para buscar en el diccionario emocional. Esto funcionaba porque en el diccionario afectivo de esta primera versión solo existía una entrada por lema para las palabras de los textos evaluados. Sin embargo existían lemas repetidos, teníamos por ejemplo las palabras “tímido” y “timado” donde el lema para ambos es “tim”. Con el aumento del diccionario del que hablamos en la Sección 4.3 ahora tenemos varias entradas más en el diccionario con el mismo lema y por tanto no podremos realizar la búsqueda en el diccionario solo por lema, debemos buscar una forma de hacer una búsqueda que devuelva un único registro. Para solucionarlo hemos aprovechado el incremento de campos en el diccionario que contábamos en la Sección 4.3 y ahora el nuevo servicio no solo usa el lema de la palabra para realizar la búsqueda sino que para buscar una palabra si no la encontramos por lema podemos emplear otras características de la palabra como el género, el número o el tipo de palabra. Además, para garantizar que encontramos la palabra ahora obtenemos los lemas a través de dos herramientas, PyStemmer y SpaCy, que pueden darnos lemas diferentes. Por ejemplo, con la palabra “castigo” PyStemmer devolverá como lema “castig” porque considerará que viene del verbo “castigar” donde su lema es “castig”, y SpaCy devolverá “castigar”. Otro ejemplo de lemas distintos es la palabra “arboleda”, para esta palabra PyStemmer devuelve “arboled” y SpaCy “arboleda”.

La nueva versión de este servicio ahora funciona así:

1. Procesamiento de la palabra con Spacy y PyStemmer. En un primer instante, cuando la palabra es recibida por el servicio se obtiene toda la información posible de la palabra: categoría morfológica, género, número, lema con Spacy y lema con PyStemmer.

2. Búsqueda de la palabra en el diccionario afectivo. Se busca en el diccionario emocional la palabra tal cual está escrita en el texto. En caso de encontrarla el servicio devolvería los grados emocionales y no continuaría con los siguientes pasos.
3. Búsqueda de la palabra en singular. En este paso, con el dato del número de la palabra obtenido en el primer paso se evalúa si estamos ante una palabra en plural o singular. Si la palabra es singular o plural no acabada en -es o -s, pasaremos al siguiente paso. Si la palabra está en plural y acaba en -es o -s quitaremos la partícula del final y volveremos a buscar la palabra en el diccionario. Este paso se debe a que algunas palabras que se encuentran en el diccionario serían evaluadas incorrectamente si obtuviésemos directamente el lema de la palabra y lo buscásemos. Por ejemplo, en el diccionario encontramos las palabras “abadía” y “abad”. En la Figura 4.10 podemos ver que ambas palabras tienen como lema “abad” y también que “abadía” es considerado un nombre propio, morfológicamente hablando (PROPN) mientras que “abad” es un nombre común. Cuando evaluamos la palabra “abadías”, plural de “abadía”, Spacy detecta la palabra como un nombre común, femenino, plural con lexema “abad”. Si buscásemos el lexema con el tipo de palabra, buscaríamos “abad” “NOUN” y obtendríamos la información de la palabra “abad” del diccionario en vez de la información que realmente queremos “abadía”. En cambio, si quitamos primero la terminación que convierte en plural la palabra, -s, buscaríamos en el diccionario la palabra “abadía” y nos daría el resultado correcto.

Lo mismo ocurre con la palabra “abuelas”. En el diccionario contamos con las palabras “abuela” y “abuelo”. En la Figura 4.11 vemos que los lemas son iguales “abuel” para ambos casos y que los valores emocionales cambian lo suficiente para que “abuelo” sea considerada una palabra un poco triste y muy alegre y “abuela” solo muy alegre. Cuando evaluamos la palabra “abuelas” primero la buscaríamos en el diccionario y no la encontraríamos. Spacy reconoce correctamente que es un nombre común, femenino y plural y buscaría en el diccionario la raíz “abuel” con el tipo de palabra “NOUN” encontrando la información de “abuelo” y no la que necesitamos de “abuela”.

4. Búsqueda de la palabra por su lema en el diccionario. En caso de que la palabra no haya sido encontrada aún o se haya detectado como singular o plural no acabada en -es o -s, usaremos el lema de la palabra que nos han devuelto Spacy y PyStemmer. En caso de que estas herramientas no hayan obtenido el mismo lema pasaremos al paso 6. En cambio si las herramientas han obtenido el mismo lema realizaremos la búsqueda en el diccionario del lema con la categoría morfológica de la palabra evaluada. En caso de que no la encontrase porque haya dos palabras en el diccionario con la misma categoría morfológica y el mismo lema, añadiríamos el género de la palabra en la búsqueda.
5. Búsqueda del lema como palabra. Si aún así no hemos encontrado la palabra en el diccionario, usaremos el lema como palabra para buscar en el diccionario. De esta manera si la palabra introducida por el usuario que estamos analizando es una palabra derivada de otra y el lema obtenido es incorrecto, podremos encontrar, al menos, la palabra de la que deriva y asociar su carga emocional a la palabra original.
6. Búsqueda con dos lemas distintos. En caso de que PyStemmer y Spacy no hayan obtenido el mismo lema, haríamos los pasos 4 y 5 con cada uno de los lemas obtenidos, probando siempre primero el obtenido con PyStemmer y después con el lema de SpaCy. Esto se hace por este orden porque los lemas que aparecen en el diccionario emocional han sido obtenidos por PyStemmer.

Palabra	Lexema	Tristeza	Miedo	Alegría	Ira	Asco	Tipo de palabra	Género	Número
abad	abad	1,35	1,56	1,71	1,24	2,23	NOUN	M	S
abadía	abad	1,43	1,22	2,33	1,21	1,21	PROP	X	X

Tabla 4.10: Resultado de las palabras “abadía” y “abad” en el diccionario

Palabra	Lexema	Tristeza	Miedo	Alegría	Ira	Asco	Tipo de palabra	Género	Número
abuela	abuel	2,3	1,67	4,33	1,13	1,1	PROP	X	X
abuelo	abuel	2,6	1,2	4,6	1,07	1,1	NOUN	M	S

Tabla 4.11: Resultado de las palabras “abuelo” y “abuela” en el diccionario

## 4.7. Peso de las palabras no emocionales en la emoción final de la frase

En el EmoTraductor 1.0 cuando una palabra no aparece en el diccionario emocional no es considerada para nada a la hora de calcular la carga afectiva de un texto, es como si la palabra no existiera en la frase. Por ejemplo, en la frase “La casa estaba acabada y adecuada solo para nosotros” solo tenemos como palabra emocional “casa” por lo que para el cálculo de las emociones finales solo se tendrá en cuenta esta palabra descartando el resto.

En el EmoTraductor 1.0 si una palabra no era emocional no era contabilizada en el número de palabras que tiene la frase y que utilizamos para dividir la suma de grados emocionales de todas las palabras emocionales. Analizando los resultados obtenidos por el EmoTraductor 1.0 nos dimos cuenta que quizás esta no era una aproximación correcta y que las palabras candidatas que finalmente no son emocionales sí deberían ser contabilizadas como palabras en la división. Si hacemos esto, la frase que pusimos como ejemplo no obtendrían el mismo valor, siendo menor la carga emocional de la frase por tener más palabras.

Nuestra hipótesis es que cuando una frase tiene muchas palabras no emocionales (que no se encuentran en el diccionario) estas palabras deberían de suavizar la carga emocional de la frase. Pero, ¿deben tener estas palabras el mismo peso que las palabras emocionales a la hora de crear el divisor? Para decidir el peso que deberían tener estas palabras hemos decidido dar distintos pesos a las palabras candidatas a ser emocionales sin entrada en el diccionario: 0 como está actualmente, 0,25, 0,5, 0,75 o 1. Por ejemplo, si tenemos la frase “Si achuchas a los bebés afligidos ayudas” donde las palabras candidatas a ser emocionales serían “achuchas” por ser un verbo, “bebés” por ser sustantivo, “afligidos” por ser un adjetivo y “ayudas” por ser un verbo. En la Tabla 4.12 podemos ver los valores que obtendrían estas palabras teniendo en cuenta que si todas las categorías emocionales tienen un 0, como “achuchas” es porque no tienen resultado en el diccionario. Para obtener el resultado de esta frase tendríamos en cuenta todas las palabras de la Tabla 4.12 menos “achuchas” por no tener entrada en el diccionario y por tanto ser considerada no emocional. Con estos datos el servicio sumaría los grados para cada emoción y dividiría entre el número de palabras emocionales (que están en el diccionario), en este caso 3. Pero si tuviéramos en cuenta la palabra “achuchar”, es decir, las candidatas a ser emocionales pero que no aparecen en el diccionario, se podría reducir el resultado final de los grados de manera que suavizásemos las emociones. El método para suavizar las emociones consiste en aumentar el divisor con el que se obtiene el resultado final dando un valor a las palabras candidatas que no terminan

siendo emocionales por no aparecer en el diccionario con el fin de no descartarlas sino que aporten y establezcan los grados emocionales. Por lo tanto con cada una de las propuestas que planteamos en el ejemplo propuesto evaluaríamos la frase con denominador 3 (cuando las palabras candidatas que no acaban siendo emocionales valen 0), 3,25 (cuando estas tienen un peso de 0,25), 3,5 (cuando pesan 0,5), 3,75 (cuando pesan 0,75) y 4 (cuando pesan 1), dando por tanto diferentes resultados finales en la frase que podemos ver en la Tabla 4.13.

Al igual que el resto de mejoras, en la siguiente Sección 4.8, vamos a ver cómo se evalúa la hipótesis y a que conclusiones llegamos.

Palabra	Tristeza	Miedo	Alegría	Ira	Asco
achuchas	0	0	0	0	0
bebés	0,29	0,87	3,55	0,32	0,35
afligidos	3,1	1,55	0,1	0,97	0,48
ayudas	1,23	1,27	2,03	0,43	0,3

Tabla 4.12: Resultado de los grados de las palabras candidatas a ser emocionales en la frase: “Si achuchas a los bebés afligidos ayudas”

Peso de las palabras	Tristeza	Miedo	Alegría	Ira	Asco
0	1,54	1,23	1,89	0,57	0,38
0,25	1,14	0,87	1,75	0,53	0,35
0,5	1,32	1,05	1,62	0,49	0,32
0,75	1,23	0,98	1,51	0,46	0,3
1	1,16	0,92	1,42	0,43	0,28

Tabla 4.13: Resultado de los grados finales de la frase “Si achuchas a los bebés afligidos ayudas” teniendo en cuenta distintos pesos para las palabras no emocionales.

## 4.8. Evaluación

En esta sección vamos a presentar cómo se han realizado las pruebas de las posibles modificaciones que planteamos para mejorar el EmoTraductor 1.0. En la Sección 4.8.1 veremos cómo se han realizado las pruebas, en la Sección 4.8.2 veremos los resultados obtenidos y en la Sección 4.8.3 la discusión de los resultados.

### 4.8.1. Diseño de la evaluación

Para poder analizar el impacto de cada una de las modificaciones planteadas, hemos seguido el mismo diseño de evaluación que seguimos en la evaluación original del EmoTraductor 1.0. Para esta evaluación se han utilizado los textos que se usaron en la evaluación del EmoTraductor 1.0 junto con nuevos textos. La finalidad de usar los textos que se utilizaron en la primera evaluación del EmoTraductor 1.0 es poder comparar resultados y ver si estamos mejorando el marcado. El fin de buscar nuevos textos para evaluar el marcado se debe a la inclusión de las distintas mejoras, ya que los textos anteriores eran insuficientes y algunos de los cambios no seríamos capaces de evaluarlos con los textos iniciales, por

ejemplo, con los textos originales no se detectaría el problema de la existencia de dos entradas en el diccionario con el mismo lema. Para la elección de estos nuevos textos hemos tenido en cuenta lo siguiente:

- Se deben buscar textos más complejos y no solo blogs o cuentos infantiles, porque se ha aumentado el número de entradas en el diccionario y por lo tanto se deben buscar textos con vocabulario más complejo. Con este fin hemos añadido al corpus de evaluación noticias y textos con frases más complejas y más palabras.
- En próximos capítulos vamos a introducir otra serie de mejoras como son el tratamiento de la negación y de los modificadores y para evaluar el impacto de estas modificaciones necesitamos contar en el corpus de evaluación con textos que contengan negaciones y modificadores.

Siguiendo estas premisas se han añadido 20 textos nuevos para las evaluaciones. Todos los textos se pueden encontrar en el Apéndice A en dos secciones diferentes. En la primera, A.1, encontraremos los textos usados en el EmoTraductor 1.0, y en la segunda, A.2, encontraremos los añadidos en esta evaluación.

El primer paso de la evaluación fue procesar todos los textos del corpus de evaluación con el EmoTraductor 1.0. A continuación, se modificó la herramienta para cambiar los límites inferior y superior de las emociones (pasamos de tener un rango emocional que iba de 1 a 5 otro que iba de 0 a 4). Tras este cambio se volvieron a procesar todos los textos. Después, se incorporaron el resto de modificaciones al sistema (excepto la de los pesos de las palabras candidatas no emocionales) y se volvieron a procesar todos los textos. Por último, modificamos el sistema para cambiar los pesos de las palabras candidatas que finalmente no son emocionales y se procesaron los textos con los distintos pesos para comprobar cuál de ellos obtenía mejores resultados.

### 4.8.2. Resultados

En la Tabla 4.14 se muestran los resultados al evaluar los textos con el EmoTraductor 1.0, donde cabe destacar que algunos textos nuevos como “CentroComercial”, “Noticia deportes” y “Texto1” no han podido ser evaluados por errores en el sistema provocados por la duplicidad de palabras con el mismo lema en el diccionario.

En la Tabla 4.15 podemos ver los resultados de la evaluación de textos tras la aplicación de la normalización en el sistema. También es interesante destacar, al igual que en el caso anterior, que tras esta mejora los textos que en la evaluación anterior usando el EmoTraductor 1.0 daban problemas al ser evaluados siguen dando este problema.

A continuación, la siguiente prueba se realiza incluyendo el resto de mejoras (la ampliación del diccionario, el cambio en el criterio de selección de palabras emocionales, la corrección en el servicio de grados emocionales de un texto y la mejora en el servicio de obtención de grados emocionales de una palabra). Después de incluirlas se vuelven a evaluar todos los textos resultando las evaluaciones que aparecen en la Tabla 4.16 donde cabe destacar que hemos solucionado el problema que existía con los textos “CentroComercial”, “Noticia deportes” y “Texto2” y ahora ya sí se han podido analizar.

Los resultados de la evaluación del sistema configurado con distintos pesos para las palabras candidatas no emocionales se pueden ver en las Tablas 4.17, 4.18, 4.19 y 4.20

Texto	Tristeza	Miedo	Alegría	Enfado	Asco
Aladdin	1,44	1,67	3,79	1,26	1,14
Blancanieves	1,9	2,02	3,02	1,81	1,51
Blancanieves y sus amigos	1,29	1,4	4,18	1,17	1,1
Blog agradecimientos	1,29	1,2	2,77	1,04	0,95
Blog enfado	0,7	0,95	1,39	0,67	0,6
Blog moda	1,1	0,93	2,05	0,96	0,9
Blog muerte	2,24	1,91	2,76	1,62	1,28
Blog reflexion	2	1,99	3,12	1,8	1,56
Cenicienta	1,43	1,53	4,05	1,27	1,16
CentroComercial	X	X	X	X	X
Conversación1	0,17	0,18	0,63	0,18	0,17
Conversación2	0,7	0,68	1,93	0,62	0,56
DobleNegación1	1,97	1,54	2,62	1,7	1,31
DobleNegación2	1,75	1,77	4,14	1,34	1,26
DobleNegación3	1,2	1,2	3,23	1,12	1,4
El castillo de bestia	1,41	1,53	3,95	1,26	1,23
El ratoncito perez	1,31	1,47	2,25	1,16	1,06
FraseCorta	1,74	1,59	3,72	1,28	1,13
FraseCorta2	2,31	2,13	2,4	1,6	1,31
La bella durmiente	2,06	2,21	2,79	1,96	1,44
La bella y la bestia	1,62	2	3,55	1,49	1,24
La margarita	1,45	1,48	3,78	1,17	1,13
La sirenita	1,89	2,17	3,92	1,34	1,28
Mulan	1,34	1,25	2,07	1,18	0,92
Negación	2	2,06	3,33	1,37	1,24
Negación2	2,21	2,46	2,39	2,04	1,66
negaciones	3,02	2,19	2,31	2,09	1,5
Negaciones coordinadas	2,29	2,19	1,45	1,56	1,19
Noticia deportes	X	X	X	X	X
Noticia muerte	2,23	2,36	1,79	2,14	1,67
Noticia nacimiento	1,71	2,18	3,32	1,61	1,38
Noticia OMS	1,99	2,12	2,57	1,63	1,54
Noticia política	1,39	1,32	2,07	1,3	1,08
Peli aburrida	1,48	1,16	1,62	1,22	1,12
Pinocho	1,37	1,53	2,61	1,14	1,02
Pocahontas	1,55	1,54	2,07	1,16	0,99
Preguntas triste oración	1,26	1,5	1,83	1,22	0,92
Texto miedo perder amor	1,23	1,38	1,25	1,03	0,82
Texto1	X	X	X	X	X
Texto2	1,42	1,56	2,13	1,18	1,13
Tipos	0,47	0,6	1,33	0,43	0,41
Tristeza	1,58	1,6	2,77	1,27	1,08
Tristeza2	1,29	1,44	1,5	1,02	0,84

Tabla 4.14: Resultados de los textos de prueba del EmoTraductor 1.0

Texto	Tristeza	Miedo	Alegría	Enfado	Asco
Aladdin	0,44	0,67	2,79	0,26	0,14
Blancanieves	0,9	1,02	2,02	0,81	0,51
Blancanieves y sus amigos	0,29	0,4	3,18	0,17	0,1
Blog agradecimientos	0,54	0,45	2,02	0,28	0,2
Blog enfado	0,26	0,5	0,95	0,22	0,15
Blog moda	0,35	0,18	1,3	0,21	0,15
Blog muerte	1,24	0,91	1,76	0,62	0,28
Blog reflexion	1	0,99	2,12	0,8	0,56
Cenicienta	0,43	0,53	3,05	0,27	0,16
CentroComercial	X	X	X	X	X
Conversación1	0,01	0,02	0,47	0,02	0,01
Conversación2	0,18	0,16	1,41	0,1	0,04
DobleNegación1	0,97	0,54	1,62	0,7	0,31
DobleNegación2	0,75	0,77	3,14	0,34	0,26
DobleNegación3	0,2	0,2	2,23	0,12	0,4
El castillo de bestia	0,41	0,53	2,95	0,26	0,23
El ratoncito perez	0,56	0,72	1,49	0,41	0,31
FraseCorta	0,74	0,59	2,72	0,28	0,13
FraseCorta2	1,31	1,13	1,4	0,6	0,31
La bella durmiente	1,06	1,21	1,79	0,96	0,44
La bella y la bestia	0,62	1	2,55	0,5	0,24
La margarita	0,45	0,47	2,78	0,17	0,13
La sirenita	0,89	1,17	2,92	0,34	0,28
Mulan	0,71	0,62	1,44	0,56	0,29
Negación	1	1,06	2,33	0,37	0,24
Negación2	1,21	1,46	1,39	1,04	0,66
negaciones	2,02	1,9	1,31	1,09	0,51
Negaciones coordinadas	1,54	1,44	0,69	0,81	0,44
Noticia deportes	X	X	X	X	X
Noticia muerte	1,38	1,51	0,94	1,29	0,82
Noticia nacimiento	0,71	1,18	2,32	0,61	0,38
Noticia OMS	1,09	1,22	1,66	0,72	0,64
Noticia política	0,64	0,57	1,33	0,55	0,33
Peli aburrida	0,82	0,49	0,95	0,55	0,46
Pinocho	0,53	0,69	1,78	0,31	0,18
Pocahontas	0,88	0,88	1,4	0,5	0,33
Preguntas triste oración	0,63	0,87	1,2	0,58	0,28
Texto miedo perder amor	0,71	0,86	0,73	0,51	0,3
Texto1	X	X	X	X	X
Texto2	0,68	0,81	1,39	0,44	0,39
Tipos	0,13	0,27	1	0,1	0,08
Tristeza	0,73	0,76	1,93	0,43	0,23
Tristeza2	0,76	0,9	0,96	0,48	0,3

Tabla 4.15: Resultados de los textos de prueba después de aplicar la normalización del diccionario.

Texto	Tristeza	Miedo	Alegría	Enfado	Asco
Aladdin	0,53	0,81	2,54	0,31	0,22
Blancanieves	0,85	0,84	1,58	0,85	0,63
Blancanieves y sus amigos	0,37	0,46	2,82	0,23	0,18
Blog agradecimientos	0,52	0,45	2,14	0,43	0,28
Blog enfado	0,41	0,55	1,38	0,34	0,28
Blog moda	0,45	0,34	0,89	0,25	0,17
Blog muerte	0,87	0,64	1,32	0,57	0,29
Blog reflexion	0,72	0,7	1,65	0,58	0,39
Cenicienta	0,49	0,5	2,34	0,34	0,31
CentroComercial	0,57	0,58	1,26	0,47	0,37
Conversación1	0,19	0,22	0,99	0,21	0,16
Conversación2	0,38	0,32	1,52	0,27	0,24
DobleNegación1	0,97	0,54	1,62	0,7	0,31
DobleNegación2	0,61	0,66	2,65	0,4	0,29
DobleNegación3	1,62	1,48	0,71	2,36	3,79
El castillo de bestia	0,49	0,7	2,51	0,43	0,23
El ratoncito perez	0,43	0,61	1,51	0,4	0,33
FraseCorta	0,54	0,41	1,69	0,35	0,3
FraseCorta2	0,68	0,63	1,44	0,4	0,25
La bella durmiente	0,58	0,8	1,96	0,54	0,28
La bella y la bestia	0,66	0,95	2,36	0,42	0,21
La margarita	0,39	0,38	2,18	0,23	0,17
La sirenita	0,62	0,87	2,19	0,35	0,25
Mulan	0,78	0,88	1,86	0,84	0,49
Negación	0,56	0,59	2,08	0,28	0,19
Negación2	0,62	0,5	0,93	0,45	0,21
Negaciones	1	0,99	1,39	0,53	0,49
Negaciones coordinadas	1,18	0,99	1,29	0,76	0,47
Noticia deportes	0,53	0,67	1,4	0,55	0,37
Noticia muerte	1,14	1,22	1,13	1,09	0,73
Noticia nacimiento	0,46	0,59	1,76	0,35	0,29
Noticia OMS	0,82	0,83	1,19	0,64	0,58
Noticia política	0,69	0,63	1,42	0,78	0,56
Peli aburrida	0,96	0,68	1,38	0,62	0,44
Pinocho	0,6	0,59	1,41	0,41	0,3
Pocahontas	0,75	0,84	1,95	0,41	0,27
Preguntas triste oración	0,33	0,42	0,97	0,28	0,14
Texto miedo perder amor	0,74	0,94	1,4	0,6	0,38
Texto1	0,8	0,88	1,09	0,77	0,51
Texto2	0,56	0,61	1,52	0,46	0,31
Tipos	0,16	0,31	2,05	0,19	0,09
Tristeza	0,83	0,66	1,07	0,46	0,34
Tristeza2	0,7	0,76	0,84	0,47	0,36

Tabla 4.16: Resultados de los textos tras las primeras mejoras

Texto	Emoción	Peso 0	Peso 0,25	Peso 0,5	Peso 0,75	Peso 1
Aladdin	Tristeza	0,53	0,52	0,51	0,5	0,49
	Miedo	0,81	0,78	0,76	0,75	0,74
	Alegría	2,54	2,48	2,44	2,4	2,36
	Enfado	0,31	0,3	0,3	0,29	0,29
	Asco	0,22	0,21	0,2	0,2	0,19
Blancanieves	Tristeza	0,87	0,8	0,76	0,72	0,69
	Miedo	0,88	0,8	0,75	0,72	0,68
	Alegría	1,59	1,51	1,44	1,37	1,31
	Enfado	0,83	0,8	0,75	0,71	0,68
	Asco	0,57	0,6	0,57	0,54	0,51
Blancanieves y sus amigos	Tristeza	0,37	0,34	0,32	0,31	0,3
	Miedo	0,46	0,42	0,4	0,37	0,36
	Alegría	2,82	2,6	2,42	2,3	2,19
	Enfado	0,23	0,21	0,19	0,18	0,17
	Asco	0,18	0,16	0,15	0,14	0,14
Blog agradecimientos	Tristeza	0,5	0,49	0,46	0,44	0,42
	Miedo	0,44	0,43	0,41	0,4	0,38
	Alegría	2,18	2,04	1,97	1,91	1,86
	Enfado	0,42	0,4	0,38	0,37	0,35
	Asco	0,27	0,26	0,25	0,24	0,23
Blog enfado	Tristeza	0,4	0,39	0,37	0,36	0,35
	Miedo	0,49	0,5	0,47	0,44	0,42
	Alegría	1,33	1,3	1,24	1,19	1,16
	Enfado	0,35	0,32	0,3	0,29	0,27
	Asco	0,3	0,26	0,25	0,24	0,23
Blog moda	Tristeza	0,44	0,4	0,36	0,33	0,31
	Miedo	0,35	0,3	0,27	0,24	0,22
	Alegría	0,87	0,82	0,76	0,71	0,67
	Enfado	0,25	0,22	0,21	0,19	0,18
	Asco	0,17	0,16	0,15	0,14	0,13
Blog muerte	Tristeza	0,84	0,8	0,74	0,69	0,65
	Miedo	0,62	0,59	0,55	0,51	0,48
	Alegría	1,33	1,24	1,17	1,11	1,06
	Enfado	0,56	0,52	0,49	0,46	0,43
	Asco	0,28	0,27	0,25	0,23	0,22
Blog reflexion	Tristeza	0,71	0,68	0,65	0,63	0,61
	Miedo	0,71	0,66	0,63	0,61	0,59
	Alegría	1,62	1,55	1,48	1,43	1,39
	Enfado	0,58	0,54	0,52	0,5	0,48
	Asco	0,37	0,36	0,34	0,33	0,32
Cenicienta	Tristeza	0,48	0,48	0,48	0,48	0,47
	Miedo	0,5	0,5	0,5	0,49	0,49
	Alegría	2,34	2,31	2,28	2,26	2,24
	Enfado	0,35	0,33	0,33	0,33	0,32
	Asco	0,32	0,3	0,3	0,3	0,3
CentroComercial	Tristeza	0,56	0,55	0,54	0,52	0,51
	Miedo	0,56	0,56	0,54	0,53	0,52
	Alegría	1,25	1,22	1,19	1,15	1,13
	Enfado	0,47	0,46	0,44	0,43	0,42
	Asco	0,36	0,35	0,34	0,33	0,32
Conversación1	Tristeza	0,19	0,19	0,19	0,19	0,19
	Miedo	0,22	0,22	0,22	0,22	0,22
	Alegría	0,99	0,93	0,88	0,85	0,83
	Enfado	0,21	0,21	0,21	0,2	0,2
	Asco	0,16	0,15	0,15	0,15	0,15
Conversación2	Tristeza	0,42	0,35	0,32	0,31	0,3
	Miedo	0,37	0,3	0,28	0,27	0,26
	Alegría	1,4	1,33	1,19	1,09	1,01
	Enfado	0,34	0,25	0,24	0,23	0,23
	Asco	0,28	0,23	0,22	0,22	0,21

Tabla 4.17: Tabla I de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.

Texto	Emoción	Peso 0	Peso 0,25	Peso 0,5	Peso 0,75	Peso 1
DobleNegación1	Tristeza	0,16	0,97	0,97	0,97	0,97
	Miedo	0,27	0,54	0,54	0,54	0,54
	Alegría	2,37	1,62	1,62	1,62	1,62
	Enfado	0,3	0,7	0,7	0,7	0,7
	Asco	0,09	0,31	0,31	0,31	0,31
DobleNegación2	Tristeza	0,61	0,61	0,61	0,61	0,61
	Miedo	0,66	0,66	0,66	0,66	0,66
	Alegría	2,65	2,65	2,65	2,65	2,65
	Enfado	0,4	0,4	0,4	0,4	0,4
	Asco	0,29	0,29	0,29	0,29	0,29
DobleNegación3	Tristeza	1,62	1,62	1,62	1,62	1,62
	Miedo	1,48	1,48	1,48	1,48	1,48
	Alegría	0,71	0,71	0,71	0,71	0,71
	Enfado	2,36	2,36	2,36	2,36	2,36
	Asco	3,79	3,79	3,79	3,79	3,79
El castillo de bestia	Tristeza	0,47	0,49	0,49	0,49	0,49
	Miedo	0,68	0,7	0,7	0,7	0,7
	Alegría	2,53	2,51	2,51	2,51	2,51
	Enfado	0,43	0,43	0,43	0,43	0,43
	Asco	0,22	0,23	0,23	0,23	0,23
El ratoncito Pérez	Tristeza	0,43	0,42	0,41	0,4	0,39
	Miedo	0,64	0,59	0,58	0,57	0,56
	Alegría	1,5	1,47	1,44	1,41	1,39
	Enfado	0,43	0,39	0,39	0,38	0,38
	Asco	0,34	0,32	0,31	0,31	0,3
FraseCorta	Tristeza	0,54	0,53	0,51	0,5	0,49
	Miedo	0,41	0,4	0,39	0,38	0,37
	Alegría	1,72	1,65	1,6	1,56	1,52
	Enfado	0,32	0,34	0,34	0,33	0,33
	Asco	0,28	0,3	0,29	0,28	0,27
FraseCorta2	Tristeza	0,66	0,66	0,64	0,62	0,6
	Miedo	0,61	0,61	0,59	0,58	0,56
	Alegría	1,47	1,4	1,36	1,32	1,29
	Enfado	0,39	0,38	0,37	0,36	0,35
	Asco	0,24	0,25	0,24	0,23	0,23
La bella durmiente	Tristeza	0,58	0,57	0,56	0,55	0,54
	Miedo	0,8	0,79	0,77	0,76	0,75
	Alegría	1,96	1,92	1,89	1,86	1,84
	Enfado	0,54	0,53	0,51	0,5	0,49
	Asco	0,28	0,27	0,27	0,26	0,26
La Bella y la Bestia	Tristeza	0,64	0,66	0,66	0,66	0,66
	Miedo	0,88	0,95	0,95	0,95	0,95
	Alegría	2,22	2,36	2,36	2,36	2,36
	Enfado	0,4	0,42	0,42	0,42	0,42
	Asco	0,23	0,21	0,21	0,21	0,21
La margarita	Tristeza	0,4	0,38	0,38	0,37	0,37
	Miedo	0,38	0,37	0,37	0,37	0,36
	Alegría	2,19	2,15	2,12	2,1	2,07
	Enfado	0,23	0,23	0,22	0,22	0,22
	Asco	0,17	0,17	0,17	0,17	0,17
La sirenita	Tristeza	0,61	0,6	0,58	0,7	0,56
	Miedo	0,89	0,84	0,81	0,79	0,77
	Alegría	2,22	2,09	2,01	1,94	1,88
	Enfado	0,35	0,35	0,35	0,34	0,34
	Asco	0,25	0,25	0,24	0,24	0,23

Tabla 4.18: Tabla II de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.

Texto	Emoción	Peso 0	Peso 0,25	Peso 0,5	Peso 0,75	Peso 1
Mulan	Tristeza	0,68	0,74	0,71	0,69	0,68
	Miedo	0,76	0,83	0,79	0,77	0,75
	Alegría	1,86	1,82	1,79	1,77	1,75
	Enfado	0,72	0,79	0,76	0,73	0,71
	Asco	0,46	0,47	0,45	0,44	0,43
Negación	Tristeza	0,6	0,56	0,56	0,56	0,56
	Miedo	0,6	0,59	0,59	0,59	0,59
	Alegría	1,97	2,08	2,08	2,08	2,08
	Enfado	0,32	0,28	0,28	0,28	0,28
	Asco	0,23	0,19	0,19	0,19	0,19
Negación2	Tristeza	0,53	0,55	0,5	0,45	0,41
	Miedo	0,57	0,45	0,4	0,37	0,34
	Alegría	1,08	0,82	0,74	0,67	0,62
	Enfado	0,47	0,4	0,36	0,32	0,3
	Asco	0,31	0,19	0,17	0,16	0,14
Negaciones	Tristeza	0,69	0,84	0,73	0,64	0,57
	Miedo	0,8	0,84	0,72	0,64	0,57
	Alegría	1,5	1,17	1,01	0,89	0,79
	Enfado	0,37	0,44	0,38	0,34	0,3
	Asco	0,34	0,41	0,36	0,31	0,28
Negaciones coordinadas	Tristeza	1,18	1,18	1,18	1,18	1,18
	Miedo	0,99	0,99	0,99	0,99	0,99
	Alegría	1,29	1,29	1,29	1,29	1,29
	Enfado	0,76	0,76	0,76	0,76	0,76
	Asco	0,47	0,47	0,47	0,47	0,47
Noticia deportes	Tristeza	0,53	0,5	0,48	0,47	0,45
	Miedo	0,66	0,64	0,62	0,6	0,59
	Alegría	1,4	1,33	1,28	1,24	1,21
	Enfado	0,55	0,53	0,51	0,5	0,49
	Asco	0,4	0,35	0,34	0,33	0,32
Noticia muerte	Tristeza	1,08	1,1	1,06	1,03	1
	Miedo	1,17	1,17	1,13	1,09	1,06
	Alegría	1,11	1,08	1,03	1	0,96
	Enfado	1,04	1,04	1	0,97	0,94
	Asco	0,7	0,7	0,68	0,65	0,64
Noticia nacimiento	Tristeza	0,46	0,43	0,41	0,38	0,37
	Miedo	0,57	0,55	0,52	0,5	0,48
	Alegría	1,76	1,65	1,56	1,49	1,42
	Enfado	0,39	0,33	0,31	0,29	0,28
	Asco	0,29	0,28	0,26	0,25	0,24
Noticia OMS	Tristeza	0,81	0,79	0,76	0,74	0,71
	Miedo	0,82	0,8	0,77	0,75	0,72
	Alegría	1,22	1,15	1,11	1,08	1,05
	Enfado	0,63	0,62	0,59	0,57	0,56
	Asco	0,57	0,55	0,53	0,51	0,5
Noticia política	Tristeza	0,7	0,65	0,62	0,59	0,57
	Miedo	0,69	0,6	0,57	0,54	0,52
	Alegría	1,37	1,34	1,27	1,21	1,16
	Enfado	0,75	0,74	0,7	0,67	0,65
	Asco	0,55	0,52	0,5	0,48	0,46
Peli aburrida	Tristeza	0,92	0,95	0,94	0,93	0,93
	Miedo	0,81	0,67	0,6	0,66	0,65
	Alegría	1,4	1,37	1,35	1,33	1,32
	Enfado	0,82	0,62	0,61	0,6	0,59
	Asco	0,53	0,44	0,43	0,43	0,42

Tabla 4.19: Tabla III de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.

Texto	Emoción	Peso 0	Peso 0,25	Peso 0,5	Peso 0,75	Peso 1
Pinocho	Tristeza	0,59	0,59	0,58	0,58	0,57
	Miedo	0,57	0,59	0,58	0,57	0,57
	Alegría	1,37	1,39	1,38	1,36	1,35
	Enfado	0,43	0,41	0,4	0,4	0,39
	Asco	0,32	0,3	0,3	0,29	0,29
Pocahontas	Tristeza	0,75	0,7	0,66	0,63	0,61
	Miedo	0,84	0,8	0,77	0,74	0,72
	Alegría	1,95	1,89	1,84	1,8	1,76
	Enfado	0,41	0,39	0,38	0,37	0,35
	Asco	0,27	0,26	0,25	0,24	0,24
Preguntas triste oración	Tristeza	0,32	0,32	0,31	0,3	0,3
	Miedo	0,42	0,4	0,38	0,37	0,36
	Alegría	0,96	0,92	0,88	0,86	0,84
	Enfado	0,28	0,26	0,25	0,25	0,24
	Asco	0,14	0,13	0,13	0,12	0,12
Texto miedo perder amor	Tristeza	0,74	0,73	0,72	0,71	0,71
	Miedo	0,94	0,92	0,91	0,9	0,89
	Alegría	1,46	1,37	1,35	1,33	1,32
	Enfado	0,59	0,59	0,58	0,57	0,57
	Asco	0,38	0,37	0,37	0,37	0,36
Texto1	Tristeza	0,8	0,75	0,72	0,68	0,66
	Miedo	0,86	0,82	0,77	0,73	0,7
	Alegría	1,22	1,01	0,95	0,9	0,87
	Enfado	0,78	0,72	0,69	0,66	0,64
	Asco	0,52	0,48	0,46	0,44	0,42
Texto2	Tristeza	0,54	0,54	0,53	0,51	0,5
	Miedo	0,59	0,59	0,57	0,55	0,54
	Alegría	1,5	1,45	1,4	1,36	1,33
	Enfado	0,44	0,45	0,43	0,42	0,41
	Asco	0,33	0,3	0,29	0,28	0,28
Tipos	Tristeza	0,16	0,15	0,14	0,14	0,14
	Miedo	0,31	0,29	0,28	0,28	0,27
	Alegría	2,05	1,96	1,89	1,85	1,82
	Enfado	0,19	0,18	0,18	0,17	0,17
	Asco	0,09	0,08	0,08	0,08	0,08
Tristeza	Tristeza	0,84	0,82	0,81	0,8	0,79
	Miedo	0,67	0,65	0,64	0,63	0,63
	Alegría	1,01	1,05	1,02	1	0,99
	Enfado	0,49	0,45	0,45	0,44	0,44
	Asco	0,35	0,34	0,33	0,32	0,32
Tristeza2	Tristeza	0,7	0,7	0,7	0,69	0,69
	Miedo	0,77	0,76	0,76	0,75	0,75
	Alegría	0,88	0,84	0,84	0,83	0,83
	Enfado	0,47	0,47	0,46	0,46	0,46
	Asco	0,36	0,36	0,36	0,36	0,36

Tabla 4.20: Tabla IV de la comparación de los resultados de los textos de prueba para evaluar si debemos dar peso a las palabras no emocionales candidatas a serlo.

### 4.8.3. Discusión de resultados

Si comparamos los resultados de las Tablas 4.14 y 4.15 para analizar el impacto del cambio de límites inferior y superior de las emociones (la normalización) vemos, como ya explicábamos antes en la Sección 4.2, que los cambios no son muy significativos ya que los textos del corpus de evaluación no usan muchas frases exclamativas o interrogativas. Pero si utilizásemos ejemplos con estas casuística podríamos ver la mejora significativa. Por tanto, según el razonamiento que encontramos en la sección 4.2 y dado que la evaluación con los textos no indica lo contrario, consideramos que esta mejora es buena y la implementamos de manera definitiva en el sistema.

Si comparamos los resultados de la Tabla 4.15 (sistema con normalización) con los presentados en la Tabla 4.16 (sistema con la ampliación del diccionario, el cambio de criterio en la selección de palabras emocionales, la corrección en el servicio de grados emocionales de un texto y la mejora en el servicio de obtención de grados emocionales de una palabra) vemos las siguientes mejoras:

- Se ha solucionado el problema que hacía que los textos “CentroComercial”, “Noticia de deportes” y “Texto 1” no se pudieran evaluar. Anteriormente estos textos no se podían evaluar por problemas al buscar las palabras únicamente por lema porque en el diccionario había palabras con lemas duplicadas. En el caso del “CentroComercial”, fallaba porque la palabra “tímidas” se buscaba usando su lema “tim” que en el diccionario del EmoTraductor 1.0 sin incluir el nuevo podía pertenecer a las palabras “tímido” o “timado” e incluyendo el nuevo diccionario podían ser cuatro palabras distintas. Lo mismo ocurre con el texto “Noticia deportes” con la palabra “daño”, que con cualquiera de los dos diccionarios el servicio no sabía si tenía que devolver las emociones de la palabra “daño” o “dañar”. Y ocurre lo mismo con el “Texto1” con la palabra “firmados”, que en caso de usar el antiguo diccionario no sabría si elegir las emociones de “firma” o de “firme” y en el nuevo diccionario añadiríamos para elegir las emociones de “firmar”. Hasta que no añadimos la modificación del servicio web que obtiene los grados emocionales de una palabra, explicado en la sección 4.6, no se podían evaluar estos textos.
- En el texto “Blog muerte” vemos una pequeña mejoría en los grados finales del texto dado que perdemos la emoción de alegría en un texto claramente triste. Esto se debe a que ahora, gracias al nuevo diccionario, somos capaces de evaluar casi todas las palabras y se diluyen más las emociones que antes sobresalían.
- El texto “DobleNegación3” mejora porque hemos conseguido que la emoción mayoritaria sea asco que es la emoción principal del texto en vez de la alegría.
- En el texto “Conversación1” al evaluar más palabras conseguimos emociones más significativas que en los resultados anteriores. Ahora tenemos como valores 0,19 de tristeza, 0,22 de miedo, 0,99 de alegría, 0,21 de enfado y 0,16 de asco frente a 0,01 de tristeza, 0,02 de miedo, 0,47 de alegría, 0,02 de enfado y 0,01 de asco.
- En el texto “Tipos” vemos que el cambio de la ponderación en los textos con diferentes tipos de frases (enunciativas, exclamativas e interrogativas) también supone una mejora. Este texto fue elegido para ver cómo funcionan las ponderaciones según el tipo de frase, gracias a la modificación conseguimos que se considere la emoción alegría como sobresaliente en el texto.

Respecto a la evaluación del peso que deben tener las palabras candidatas no emocionales, si analizamos los resultados presentados en las Tablas 4.17, 4.18, 4.19 y 4.20 vemos que

los pesos inferiores al 1 no provocan variaciones significativas aunque van empeorando los resultados hasta llegar a los peores resultados con el peso 1. Con el peso 1 vemos que hay textos como “Conversación2” donde se han perdido todas las emociones y dejamos el texto como neutral, diluyendo todas las emociones cuando nuestro fin es mejorar el marcado y no empeorarlo. Por lo tanto, descartamos que un cambio en los pesos de las palabras no emocionales pueda mejorar de manera significativa como para tenerlo en cuenta. Decidimos entonces que las palabras candidatas a ser emocionales que no aparezcan en el diccionario, seguirán como en el EmoTraductor 1.0, y no aportarán nada al cálculo final de las emociones de una frase.

# Capítulo 5

## Detección de modificadores y su papel en la marcación emocional

**RESUMEN:** En este capítulo vamos a explicar la integración de los modificadores en la herramienta y su repercusión en los resultados de las marcaciones emocionales realizadas por nuestra herramienta.

Recordemos que un modificador es una palabra que modifica la emoción asociada a una palabra (Andreevskaia y Bergler, 2007). Tenemos dos tipos diferentes de modificadores, los aumentativos que aumentarán la carga emocional de la palabra o palabras sobre las que actúa y los reductores que disminuirán la carga emocional.

Como hemos visto en la Sección 2.6, que una herramienta de marcado emocional de textos reconozca ciertas palabras que actúan como modificadores ayuda a mejorar el marcado emocional que esta realiza. En las frases, “El muchacho está enfadado”, “El muchacho está poco enfadado” y “El muchacho está muy enfadado” podemos apreciar diferentes grados de enfado, en la segunda oración el muchacho está menos enfadado que en la primera y en la primera menos que en la tercera. Sin embargo, sin el uso de los modificadores el EmoTraductor 1.0, al no contemplar el tratamiento de los modificadores, daría el mismo resultado en la tres frases tal y como podemos ver en la Tabla 5.1.

Frase	Tristeza	Miedo	Alegría	Ira	Asco
El muchacho está enfadado	0,92	0,76	0,88	<b>1,6</b>	0,34
El muchacho está poco enfadado	0,92	0,76	0,88	<b>1,6</b>	0,34
El muchacho está muy enfadado	0,92	0,76	0,88	<b>1,6</b>	0,34

Tabla 5.1: Grados emocionales devueltos por el EmoTraductor 1.0 para tres frases que difieren únicamente en los modificadores.

Como acabamos de ver, tal y como explicábamos en la Sección 2.6, pueden existir modificadores que aumenten la carga emocional y otros que la disminuyan. Consideraremos en este apartado solo aquellos modificadores que aumenten la carga emocional ya que los que la disminuyen serán tratados como parte de la negación que se explicará en el próximo capítulo (Capítulo 6). En la nueva versión del EmoTraductor, con el fin de mejorar el marcado emocional, vamos a asociar a cada modificador un porcentaje que aumentará los grados de las palabras afectadas por dicho modificador.

En las próximas secciones veremos cómo se ha implementado el tratamiento de los modificadores en el sistema. En la Sección 5.1 veremos la base para el tratamiento de los modificadores, un listado de modificadores con un porcentaje de actuación asociado. En la Sección 5.2 veremos cómo identificamos el ámbito de un modificador. En la Sección 5.3 analizaremos el proceso paso a paso de cómo se tratan los modificadores en el sistema. Finalmente, en la Sección 5.4 veremos la evaluación que se ha realizado para estudiar la mejora que produce este cambio en el marcado de texto de la herramienta.

## 5.1. Listado de modificadores

Hemos decidido añadir al EmoTraductor 1.0 una lista basada en la lista de modificadores en inglés presentada en la tesis de Carrillo (2011) (ver Figura 5.1) traduciendo dichos modificadores al castellano y manteniendo los porcentajes asignados a cada uno de ellos.

La lista resultante se puede ver en la Tabla 5.2. En esta tabla podemos ver los modificadores originales, su traducción al castellano y el porcentaje de aumento que se aplicará a la carga emocional en de las palabras afectadas por el modificador. Todos los modificadores de la lista original que tenían un porcentaje negativo (“Hardly”, “Relatively”, “Ridiculously”, “Minor”, “Lowest”, “Fewest”, “Only”, “Small”, “Little”, “Less”, “Almost”, “Somewhat”, “Low”, “Few”, “Bit”, “Barely”, “Slightly”, “Lower”, “Fewer”, “Difficult”) se han eliminado del listado ya que son modificadores considerados reductores y los trataremos como parte de la negación. También hemos eliminado del listado original aquellos términos que se traducen al castellano con más de una palabra: “Higher” (más alto), “Rather” (más bien), “Highest” (mas alto), dado que el EmoTraductor 1.0 analiza palabra a palabra y aún no puede contemplar este tipo de composiciones. Sin embargo, si en el listado de modificadores existiese “alto”, cuando encontremos en el texto una expresión como “más alto que” no iremos al porcentaje que se planteó como “más alto” pero, sí al que se planteó para “alto”. Se han eliminado además, las palabras que al ser traducidas al castellano tienen el mismo significado que otras que ya poseíamos. En estos casos, para elegir qué palabra permanecerá en nuestro listado, hemos decidido quedarnos siempre con la primera palabra que tiene el significado. Por esta causa se han eliminado las palabras “Fully” ya que primero aparece “Completly”, “Exceedingly” ya que primero aparece “Extremely” y “Awfully” ya que primero aparece “Very”. Por último, también hemos eliminado el término “Pretty”, ya que al traducirlo obtendríamos “bonito” o “bonita” que en español son utilizados como adjetivos, que no son considerados modificadores, y serán palabras emocionales que ya aparecen en el diccionario afectivo usado por el EmoTraductor.

## 5.2. Detección del ámbito de los modificadores

Una vez integrada en el sistema la lista de modificadores y su porcentaje asociado vamos a considerar que una palabra está afectada por un modificador si éste está en su ámbito. Para reconocer el ámbito del modificador, es decir, a qué parte de la frase o a qué palabras está afectando el aumento que conlleva el modificador, hemos utilizado la herramienta Spacy. Esta herramienta nos permite crear un árbol de dependencias por cada frase. Por tanto, todas las palabras emocionales que tengan en alguno de sus nodos hijos en el árbol de dependencias el modificador, estarán afectadas por este y aumentarán sus grados emocionales.

Vamos a verlo con un ejemplo, para la frase “La niña está muy feliz y contenta” Spacy devolvería el árbol de dependencias que podemos ver, en la Figura 5.2. La palabra “muy”,

Modificador en inglés	Modificador traducido	% de actuación
Very	muy	+75 %
Great	gran	+60 %
Much	mucho	+35 %
Some	algún	+40 %
Perfectly	perfectamente	+75 %
Completely	completamente	+85 %
Highly	altamente	+75 %
Particularly	particularmente	+45 %
Totally	totalmente	+75 %
Strongly	fuertemente	+75 %
Terribly	terriblemente	+75 %
Exceptionally	excepcionalmente	+85 %
Super	super	+75 %
Huge	enorme	+75 %
Absolute	absoluto	+75 %
Big	grande	+65 %
Extra	extra	+20 %
Really	realmente	+80 %
Obviously	obviamente	+75 %
Definitely	definitivamente	+95 %
Truly	verdaderamente	+75 %
Significantly	significativamente	+45 %
Extremely	extremadamente	+95 %
Immensely	inmensamente	+75 %
Thoroughly	profundamente	+85 %
Extraordinarily	extraordinariamente	+75 %
Tremendous	tremendo/tremenda	+85 %
Incredible	increíble	+75 %
Major	mayor/mayormente	+35 %
Lot	mucha	+75 %
So	tan/tanto	+75 %
More	más	+35 %
Fairly	justamente	+65 %
Quite	bastante	+75 %
Certainly	ciertamente	+75 %
Absolutely	absolutamente	+95 %
Especially	especialmente	+70 %
Entirely	enteramente	+75 %
Incredibly	increíblemente	+85 %
Vastly	vastamente	+75 %
Tremendously	tremendamente	+85 %
Total	total	+75 %
Complete	completo/completa	+75 %
Biggest	grandísimo	+75 %
High	alto/alta	+55 %
Real	real	+20 %

Tabla 5.2: Modificadores incluidos en el EmoTraductor con sus porcentajes asociados

<b>Términos</b>	<b>%</b>	<b>Términos</b>	<b>%</b>	<b>Términos</b>	<b>%</b>
Very	75	Small	-50	So	75
Great	60	Really	80	More	35
Much	35	Little	-75	Bit	-35
Some	40	Less	-55	Barely	-75
Hardly	-85	Almost	-20	Slightly	-75
Relatively	-50	Somewhat	-60	Fairly	65
Pretty	75	Rather	75	Quite	35
Perfectly	75	Obviously	75	Certainly	75
Completely	85	Definitely	95	Absolutely	95
Highly	75	Truly	75	Especially	70
Particularly	45	Significantly	45	Awfully	75
Totally	85	Fully	85	Entirely	75
Strongly	55	Extremely	95	Incredibly	85
Terribly	75	Immensely	75	Such	55
Exceptionally	75	Exceedingly	85	Vastly	75
Super	75	Thoroughly	65	Tremendously	85
Ridiculously	-65	Extraordinarily	75	Total	75
Huge	75	Tremendous	85	Complete	75
Absolute	75	Incredible	75	Utter	75
Big	65	Bigger	35	Biggest	75
Extra	20	Major	35	High	55
Higher	45	Highest	65	Real	20
Minor	-75	Low	-75	Lower	-65
Lowest	-85	Few	-75	Fewer	-65
Fewest	-85	Lot	75	Difficult	-65
Only	-25				

Figura 5.1: Modificadores usados en el trabajo de Carrillo (2011)

es el modificador y queremos conocer su ámbito, por tanto nos fijamos en sus hijos y vemos que no tiene ninguno, con lo que podríamos pensar que no afecta a ninguna otra palabra. Sin embargo, si nos fijamos en los hijos de la palabra “feliz” podemos ver que aparece la palabra “muy” entendiendo así que esta palabra tiene relación con el modificador “muy”. Podemos entenderlo mejor usando la representación gráfica de este árbol, que encontramos en la Figura 5.3. Esta representación nos permite interpretar de una forma más visual el árbol de dependencias devuelto por Spacy, donde el comienzo de la flecha es el nodo origen y el final de esta indica que la palabra señalada afecta o se ve influenciada por la que tiene en su su origen. A partir del árbol de la Figura 5.3 vemos que de la palabra “muy” no sale ninguna flecha por lo que no se ve modificada por ninguna otra palabra pero sí le llega una flecha procedente de otra palabra (“feliz”), lo que implica que modifica a esa palabra. Por tanto, el modificador “muy” afectará únicamente en esta frase a la carga emocional de la palabra “feliz”.

Para poder gestionar estas dependencias en el sistema, recorreremos el árbol devuelto por Spacy como el que hemos visto (Figura 5.2) buscando los modificadores entre los hijos de las palabras emocionales.

```

Palabra: La
Hijos:
Palabra: niña
Hijos:
    La
Palabra: está
Hijos:
Palabra: muy
Hijos:
Palabra: feliz
Hijos:
    niña
    está
    muy
    contenta
Palabra: y
Hijos:
Palabra: contenta
Hijos:
    y

```

Figura 5.2: Árbol de dependencias generado por Spacy para la frase “La niña está muy feliz y contenta”

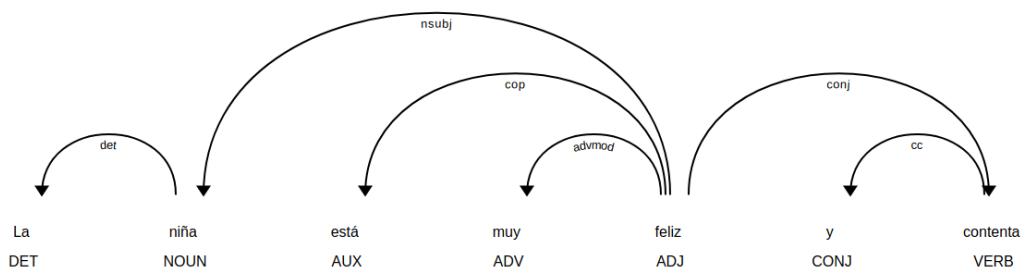


Figura 5.3: Árbol de dependencias creado por Spacy de la frase “La niña está muy feliz y contenta”

Por lo tanto, continuando el ejemplo, el sistema con este nuevo proceso descubriría que “feliz” tiene en su ámbito el modificador “muy” y por lo tanto buscará el porcentaje asociado a “muy” en el listado de modificadores (Tabla 5.2) y obtendrá como porcentaje 75%. A continuación, modificará las emociones asociadas a la palabra “feliz” con este porcentaje amplificador. En la Tabla 5.3, vemos los grados de cada una de las palabras emocionales que componen la frase que estamos usando de ejemplo, y si aplicamos algún modificador en ellas y en qué medida. En caso de verse afectada por modificadores se ve entre paréntesis el resultado de los grados después de la actuación de los modificadores, en la última fila vemos el resultado total de la frase, obtenida teniendo en cuenta que la palabra “feliz” usará los grados entre paréntesis.

Palabra	Tristeza	Miedo	Alegría	Ira	Asco	% modificadores
niña	0,43	0,47	2,37	0,3	0,07	0 %
feliz	0,1 (0,175)	0,2 (0,35)	3,63 (6,3525)	0 (0)	0 (0)	+75 %
contenta	0,19	0,19	3,52	0,13	0,16	0 %
Resultado total	0,27	0,34	4,08	0,14	0,08	

Tabla 5.3: Grados emocionales considerando modificadores de “La niña está muy feliz y contenta”

### 5.3. Marcación emocional de una frase teniendo en cuenta los modificadores

En esta sección vamos a ver cómo quedaría el servicio de obtención de los grados emocionales de una frase añadiendo el tratamiento de los modificadores. Vamos a usar como frase de ejemplo en esta sección la frase “La mujer es súper calmada pero le gustan mucho las fiestas”, cuyo árbol de dependencias se puede ver en la Figura 5.4.

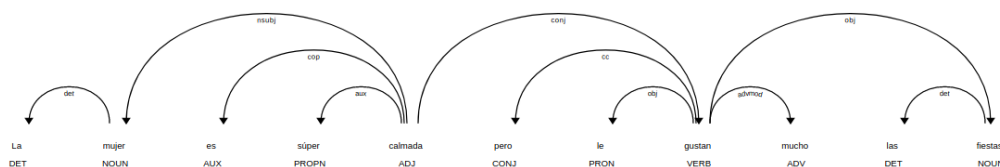


Figura 5.4: Árbol de dependencias creado por Spacy de la frase “La mujer es súper calmada pero le gustan mucho las fiestas”

Los pasos que sigue la nueva versión del servicio para obtener el contenido emocional de una frase son los siguientes (en **negrita** las modificaciones realizadas en el servicio para tratar los modificadores):

1. Obtención del árbol de dependencias de la frase y de las categorías gramaticales de cada palabra. Cuando la frase llega al servicio, utilizamos Spacy para detectar las categorías gramaticales de cada una de las palabras incluidas en la frase y para obtener el árbol de dependencias de la frase. Según la categoría gramatical de una palabra conoceremos si es una palabra candidata a ser emocional o no y si la palabra es candidata se almacenará en una lista. En la Tabla 5.4 se muestran todas las palabras de la frase de ejemplo junto con su categoría gramatical y si es candidata a ser emocional o no.
2. Detección de modificadores y de las palabras incluidas en el ámbito de los modificadores. Es importante que este paso se realice después de hallar las palabras candidatas a ser emocionales, ya que solo nos interesan aquellos modificadores que afecten a alguna de estas palabras. Para detectar las palabras afectadas por algún modificador recorreremos el árbol de dependencias (como se ha explicado en la Sección 5.2). Recorreremos las palabras de la frase y pararemos a evaluar los nodos hijos de las palabras candidatas a ser emocionales. Si en alguno de los nodos hijos de las candidatas aparece algún modificador que aparezca en la lista de modificadores, sabremos que

Palabra	Categoría gramatical	Candidata
La	DET	NO
mujer	NOUN	SI
es	AUX	NO
súper	PROPN	NO
tranquila	ADJ	SI
pero	CONJ	NO
le	PRON	NO
gustan	VERB	SI
mucho	ADV	NO
las	DET	NO
fiestas	NOUN	SI

Tabla 5.4: Categorías gramaticales de la frase “La mujer es súper calmada pero le gustan mucho las fiestas.”

esa palabra candidata a ser emocional está afectada por un modificador. Todas las palabras candidatas a ser emocionales, independientemente de si son afectadas por un modificador o no, serán almacenadas en una lista de tuplas. Estas tuplas consisten en la palabra candidata, su posición en la frase y un valor que indica el porcentaje total de actuación del conjunto de modificadores que le afectan y por tanto, modificarán la carga afectiva de la palabra. El conjunto puede estar formado por uno o varios modificadores.

En nuestro ejemplo la lista sería la siguiente:

$[[ \text{'mujer'}, 1, 0 ], [ \text{'calmada'}, 4, 75 ], [ \text{'gustan'}, 7, 35 ], [ \text{'fiestas'}, 10, 0 ]]$
--

Podemos ver que las palabras “calmada” y “gustan” están siendo afectadas por los modificadores “súper” y “mucho” con unos porcentajes de 75 % y 35 % respectivamente.

3. Obtención de la carga emocional de la frase. Este proceso se realiza palabra a palabra con la ayuda del diccionario afectivo del sistema. Si una palabra existe en el diccionario será emocional, se almacenarán los grados emocionales que tenga asignados en el diccionario y se sumarán al resultado total de la frase. En el proceso que aúna la carga emocional de cada palabra para calcular la frase, se evaluará el porcentaje de modificación que tiene asociado la palabra que va a añadir su carga emocional a la frase. Según el porcentaje que tenga se actuará de distinta manera:

- Si la palabra está en el listado y el valor de actuación asociado a la palabras es distinto de 0. Se usará el valor que aparece en la tupla que indica el porcentaje total de actuación del o los modificadores que afectan a la palabra, para así modificar el contenido afectivo de la palabra en el diccionario antes de ser sumado al total de la oración.
- Si la palabra está en el listado y el valor de actuación asociado a la palabra es 0. Se sumarán los grados emocionales que tiene asociados la palabra en el diccionario.

El porcentaje que aparece en el listado de tuplas asociado a una palabra afectará a todos los grados emocionales de cada categoría. En la Tabla 5.5 podemos ver para cada palabra de la frase de ejemplo los grados emocionales antes y después de aplicarles los porcentajes de los modificadores.

Palabra	Porcentaje a aplicar	Emoción	Grados sin modificador	Grados con modificador
mujer	0	Tristeza	0,45	0,45
		Miedo	0,3	0,3
		Alegría	2,23	2,23
		Enfado	0,3	0,3
		Asco	0,2	0,2
calmada	75	Tristeza	0,33	0,58
		Miedo	0,2	0,35
		Alegría	1,27	2,22
		Enfado	0,3	0,53
		Asco	0,03	0,05
gusten	35	Tristeza	0,17	0,23
		Miedo	0,6	0,81
		Alegría	3,43	4,63
		Enfado	0,2	0,27
		Asco	0,17	0,23
fiestas	0	Tristeza	0,13	0,13
		Miedo	0,28	0,28
		Alegría	3,69	3,69
		Enfado	0,07	0,07
		Asco	0,07	0,07

Tabla 5.5: Grados emocionales sin y con modificadores de la frase “La mujer es súper calmada pero le gustan mucho las fiestas.”

Una vez realizada la suma de las palabras, teniendo en cuenta en los casos necesarios los modificadores, y tal y como se hacía en el EmoTraductor 1.0, se dividirán los grados emocionales totales que se han obtenido entre el número de palabras emocionales que existían en la frase, haciendo así la media. En el caso de la frase de ejemplo el resultado obtenido será el que se muestra en la Tabla 5.6. En esta tabla vemos en la primera sección los valores que toman definitivamente las palabras candidatas según si están afectadas por un modificador cada uno de los grados de cada emoción, el número de palabras por el que dividir y el resultado final de la frase.

#### 5.4. Evaluación del impacto del tratamiento de los modificadores

En esta sección vamos a ver primero cómo evaluaremos el impacto de los modificadores tanto en los ejemplos que comenzamos poniendo al principio del capítulo como en los textos con los que evaluamos el marcado emocional obtenido por el EmoTraductor 1.0. Después observaremos los resultados y los discutiremos.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
mujer	0,45	0,3	2,23	0,3	0,2
calmada	0,58	0,35	2,22	0,53	0,05
gusten	0,23	0,81	4,63	0,27	0,23
fiestas	0,13	0,28	3,69	0,07	0,07
Suma de los grados	1,39	1,74	12,77	1,17	0,55
Número de palabras	4				
Resultado final	0,34	0,43	3,19	0,29	0,14

Tabla 5.6: Grados emocionales finales para la frase “La mujer es súper calmada pero le gustan mucho las fiestas.”

#### 5.4.1. Diseño de la evaluación

Esta evaluación se va dividir en dos partes. La primera será la evaluación del cambio del marcado obtenido en las frases “El muchacho está enfadado”, “El muchacho está poco enfadado” y “El muchacho está muy enfadado”. Usaremos estas frases porque son las que empleamos al principio del capítulo para ver la importancia de tener en cuenta los modificadores y veremos si se ha mejorado el marcado o no. También se evaluará usando otras frases donde podremos ver la actuación de los modificadores.

La segunda parte consistirá en evaluar usando los mismos textos que usamos en la evaluación del Capítulo 4 y que se pueden consultar en el Apéndice A. Los resultados obtenidos con estos textos serán comparados con los resultados obtenidos al final del Capítulo 4, concretamente, los de la Tabla 4.16.

#### 5.4.2. Resultados

Los resultados de la primera evaluación a realizar son los que encontramos en la Tabla 5.7 donde aparecen las tres frases que queremos evaluar “El muchacho está enfadado”, “El muchacho está poco enfadado” y “El muchacho está muy enfadado”. La intención de esta evaluación es ver si realmente hemos solucionado el problema en estas frases donde la única diferencia que hay entre ellas es el modificador.

Frase	Tristeza	Miedo	Alegría	Enfado	Asco
El muchacho está enfadado	0,92	0,76	0,88	<b>1,6</b>	0,34
El muchacho está poco enfadado	0,92	0,76	0,88	<b>1,6</b>	0,34
El muchacho está muy enfadado	<b>1,56</b>	1,24	0,88	<b>2,69</b>	0,53

Tabla 5.7: Grados emocionales considerando modificadores

También se ha probado para esta primera evaluación con las frases “Él está muy alegre” frente “Él está alegre”, “La mujer iba perfectamente arreglada aquella noche” frente “La mujer iba arreglada aquella noche”, “Ella era realmente feliz” frente “Ella era feliz” y “Sonrió con total libertad” frente “Sonrió con libertad”, cuyos resultados podemos ver en la Tabla 5.8.

Los resultados de la segunda parte de la evaluación se muestran en la Tabla 5.9. Para hacer la evaluación más sencilla hemos añadido en la columna “Sin modificadores” los

Frase	Tristeza	Miedo	Alegría	Enfado	Asco
Él está alegre	0,1	0,23	3,83	0,1	0,03
Él está muy alegre	0,18	0,4	<b>6,7</b>	0,18	0,05
La mujer iba arreglada aquella noche	0,6	0,64	1,9	0,39	0,2
La mujer iba perfectamente arreglada	0,67	0,67	2,32	0,52	0,24
Ella era feliz	0,1	0,2	3,63	0,0	0,0
Ella era realmente feliz	0,18	0,36	<b>6,53</b>	0,0	0,0
Sonrió con libertad	0,21	0,43	3,76	0,23	0,2
Sonrió con total libertad	0,28	0,54	3,88	0,28	0,14

Tabla 5.8: Grados emocionales considerando modificadores para distintas frases con y sin modificadores

resultados que aparecían en el capítulo 4, concretamente, los de la Tabla 4.16 y en la columna “Con modificadores” los resultados de la nueva versión.

### 5.4.3. Discusión de resultados

Vamos a analizar los resultados de la primera evaluación que hemos realizado, en la que usábamos 3 frases similares, que podemos ver en la Tabla 5.7. Con la introducción del tratamiento de los modificadores se ven diferencias en el grado de la emoción enfado que pasa de un 1,6 a un 2,69 para la frase que tiene un modificador aumentativo (“muy”). También vemos que en la frase que usa el modificador “muy” no solo se ha incrementado el valor de enfado, sino que podemos ver como el grado de tristeza aumenta considerablemente pasando de un 0,92 a un 1,56, convirtiéndose la tristeza en una emoción representativa de la oración. Sin embargo, como hemos dicho al principio, no hemos tenido en cuenta los modificadores que disminuyen los valores porque los trataremos como negaciones en el siguiente capítulo y por ello, no hay ninguna diferencia en los resultados entre la frase “El muchacho está enfadado” y “El muchacho está poco enfadado”.

Respecto a los resultados de las otras frases propuestas (Tabla 5.8) vemos que efectivamente el uso de los modificadores aumenta correctamente los grados emocionales de las palabras. También encontramos en las frases “Él está muy alegre” y “Ella era realmente feliz” valores para la alegría superiores a lo que teníamos establecido como límite teórico para los grados emocionales (4), el nuevo método de marcado devuelve para la alegría un 6,7 y un 6,5 respectivamente. Esto se debe a que los modificadores amplían independientemente del grado que tenga asociada la palabra. Evaluando la posibilidad de reducir estos números, haciendo una relación de proporción para convertir el 6,7 en un 4, nos damos cuenta que si lo hacemos podríamos perder otras emociones como en la frase “Él está muy alegre” el asco que llegaría a desaparecer. Por lo tanto, optamos por la opción de que el servicio devuelva los grados emocionales independientemente del valor del resultado obtenido, dado que entendemos que es correcto decir que la emoción mas predominante y por mucho, en la frase anterior, es la alegría. También nos planteamos y evaluamos qué pasaría si el front de la aplicación recibe este valor superior al máximo teórico establecido para una emoción (4) y nos dimos cuenta que es indiferente, dado que la interfaz muestra las emociones en la barra emocional, como vimos en la Sección 3.1, y esta utiliza porcentajes para mostrar los resultados.

Respecto a los resultados de la segunda evaluación, (Tabla 5.9) vemos, que por lo general, no existe una mejora muy significativa. Por ejemplo en el texto “Aladdin” el único

Texto	Sin modificadores					Con modificadores				
	T	M	Al	E	A	T	M	Al	E	A
Aladdin	0,53	0,81	2,54	0,31	0,22	0,53	0,81	2,55	0,31	0,22
Blancanieves	0,85	0,84	1,58	0,85	0,63	0,87	0,88	1,59	0,83	0,57
Blancanieves y sus amigos	0,37	0,46	2,82	0,23	0,18	0,37	0,46	2,82	0,23	0,18
Blog agradecimientos	0,52	0,45	2,14	0,43	0,28	0,56	0,49	2,67	0,5	0,29
Blog enfado	0,41	0,55	1,38	0,34	0,28	0,41	0,5	1,34	0,36	0,3
Blog moda	0,45	0,34	0,89	0,25	0,17	0,45	0,35	0,89	0,26	0,17
Blog muerte	0,87	0,64	1,32	0,57	0,29	0,86	0,64	1,34	0,58	0,3
Blog reflexion	0,72	0,7	1,65	0,58	0,39	0,73	0,73	1,7	0,6	0,38
Cenicienta	0,49	0,5	2,34	0,34	0,31	0,5	0,52	2,46	0,37	0,34
CentroComercial	0,57	0,58	1,26	0,47	0,37	0,56	0,57	1,25	0,47	0,37
Conversación1	0,19	0,22	0,99	0,21	0,16	0,19	0,22	0,99	0,21	0,16
Conversación2	0,38	0,32	1,52	0,27	0,24	0,46	0,39	1,4	0,35	0,28
DobleNegación1	0,97	0,54	1,62	0,7	0,31	0,16	0,27	2,37	0,3	0,09
DobleNegación2	0,61	0,66	2,65	0,4	0,29	0,65	0,74	3,27	0,43	0,31
DobleNegación3	1,62	1,48	0,71	2,36	3,79	1,62	1,48	0,71	2,36	3,79
El castillo de bestia	0,49	0,7	2,51	0,43	0,23	0,48	0,69	2,56	0,43	0,23
El ratoncito perez	0,43	0,61	1,51	0,4	0,33	0,43	0,65	1,51	0,43	0,35
FraseCorta	0,54	0,41	1,69	0,35	0,3	0,54	0,41	1,72	0,32	0,28
FraseCorta2	0,68	0,63	1,44	0,4	0,25	0,66	0,61	1,47	0,39	0,24
La bella durmiente	0,58	0,8	1,96	0,54	0,28	0,59	0,84	2,04	0,56	0,29
La bella y la bestia	0,66	0,95	2,36	0,42	0,21	0,66	0,91	2,3	0,41	0,24
La margarita	0,39	0,38	2,18	0,23	0,17	0,4	0,38	2,19	0,23	0,17
La sirenita	0,62	0,87	2,19	0,35	0,25	0,61	0,89	2,22	0,35	0,25
Mulan	0,78	0,88	1,86	0,84	0,49	0,7	0,77	1,93	0,75	0,46
Negación	0,56	0,59	2,08	0,28	0,19	0,6	0,6	1,97	0,32	0,23
Negación2	0,62	0,5	0,93	0,45	0,21	0,53	0,57	1,08	0,47	0,31
Negaciones	1	0,99	1,39	0,53	0,49	0,69	0,8	1,5	0,37	0,34
Negaciones coordinadas	1,18	0,99	1,29	0,76	0,47	1,18	0,99	1,29	0,76	0,47
Noticia deportes	0,53	0,67	1,4	0,55	0,37	0,55	0,68	1,44	0,57	0,42
Noticia muerte	1,14	1,22	1,13	1,09	0,73	1,08	1,17	1,11	1,04	0,7
Noticia nacimiento	0,46	0,59	1,76	0,35	0,29	0,46	0,58	1,8	0,4	0,29
Noticia OMS	0,82	0,83	1,19	0,64	0,58	0,81	0,82	1,24	0,63	0,57
Noticia política	0,69	0,63	1,42	0,78	0,56	0,7	0,69	1,38	0,75	0,55
Peli aburrida	0,96	0,68	1,38	0,62	0,44	1,11	0,94	1,73	1	0,66
Pinocho	0,6	0,59	1,41	0,41	0,3	0,62	0,59	1,41	0,44	0,33
Pocahontas	0,75	0,84	1,95	0,41	0,27	0,75	0,84	1,95	0,41	0,27
Preguntas triste oración	0,33	0,42	0,97	0,28	0,14	0,33	0,43	1	0,29	0,14
Texto miedo perder amor	0,74	0,94	1,4	0,6	0,38	0,76	0,99	1,48	0,61	0,39
Texto1	0,8	0,88	1,09	0,77	0,51	0,81	0,87	1,23	0,8	0,53
Texto2	0,56	0,61	1,52	0,46	0,31	0,55	0,6	1,52	0,44	0,34
Tipos	0,16	0,31	2,05	0,19	0,09	0,16	0,31	2,07	0,19	0,09
Tristeza	0,83	0,66	1,07	0,46	0,34	0,83	0,69	1,01	0,51	0,36
Tristeza2	0,7	0,76	0,84	0,47	0,36	0,71	0,78	0,88	0,48	0,37

Tabla 5.9: Comparación de los resultados de los textos de prueba después las primeras mejoras y al añadir los modificadores. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).

valor que varía es la alegría y pasamos de un 2,54 a un 2,55; o en el “Blog muerte” donde varían la tristeza de un 0,87 a un 0,86, el miedo de se mantiene, la alegría pasa de un 1,32 a un 1,34, el enfado cambia de un 0,57 a un 0,58 y el asco pasa de un 0,29 a un 0,3. Sin embargo, los resultados varían de manera positiva en algunos casos como en el texto “Peli aburrida”, en el “Blog agradecimiento” o en “DobleNegación2” ya que son textos escogidos para verificar que hay variaciones al añadir modificadores a las evaluaciones, todos aumentan sus emociones porque utilizan modificadores como “muy” y “tan”. En el texto “DobleNegación2”, que podemos analizarlo porque solo posee una frase “Nunca he visto en mi vida un paisaje tan hermoso.”, aparece el modificador “tan” que aumenta la carga emocional de la palabra “hermoso”. Como este texto solo posee una frase y le afecta un modificador, vemos con más facilidad el cambio de los grados emocionales pasando, por ejemplo la alegría de un 2,36 a un 3,27. Por lo tanto, que no se aprecie el aumento de las emociones como en las frases simples de los ejemplos se debe a que, cuanto mayor es el texto, mayor es el número de frases sobre el que dividir la suma de los grados emocionales, y cuanto mayor son las frases mayor es el número de palabras sobre el que dividir la suma de los grados emocionales, haciendo que cada palabra aporte muy poco a un texto. Sin embargo, se nota que es un cambio que no penaliza los resultados y en algunos casos los enriquece y por tanto es un cambio válido para mejorar el marcado emocional de textos.

## Detección de la negación y su papel en la marcación emocional

**RESUMEN:** En este capítulo vamos a ver cómo podemos incluir el tratamiento de la negación en el marcado emocional de textos del EmoTraductor 1.0. Veremos en la Sección 6.1 una clasificación de las negaciones y después, en las Secciones 6.2 y 6.3 los métodos con los que las trataremos en el sistema. Por último, en la Sección 6.4 se contarán las conclusiones del tratamiento de la negación.

Tras no encontrar en ningún trabajo anterior cómo se pueden tratar las negaciones en el marcado de texto usando únicamente categorías emocionales no antónimas entre sí, vamos a plantear distintos métodos para tratar las negaciones en nuestro sistema. Estas estarán basadas en las ideas de otros trabajos que se intentaron aplicar al nuestro. Todos los planteamientos se evaluarán para saber si son válidos y cuál es el más prometedor.

### 6.1. Tipos de negación

Tal y como hacía Carrillo (2011) en su tesis (ver Sección 2.8.3) vamos a dividir las negaciones en dos grupos:

- **Negaciones nominales:** son las negaciones que Carrillo (2011) consideraba negaciones implícitas. Están compuestas habitualmente por adverbios que se encargan de disminuir las emociones del texto. Por ejemplo, en la frase “La comida me gusta poco”, “poco” reduce las emociones de la frase pero no las modifica completamente. En la frase “Mi madre estaba casi enfadada cuando salí de casa”, la palabra “casi” reduce las emociones de “enfadada” pero no llega a ser una negación rotunda de la oración completa.
- **Negaciones sintácticas:** son las negaciones que modifican directamente al verbo de la frase como en “No iremos al cine”, donde el “no” afecta directamente al hecho de ir, o en las frases “Mi madre no está enfadada” o “No me gusta la sandía”, donde el “no” en cada una de ellas está cambiando completamente el significado de la frase ya que la acción del verbo no se está llevando a cabo.

Trataremos de manera diferente estos dos tipos de negaciones ya que afectan de forma distinta al contenido emocional de la frase en la que están contenidas. En las próximas secciones explicaremos cómo tratamos cada una.

## 6.2. Tratamiento de la negación nominal

Las **negaciones nominales** suelen ser adverbios que se encargan de disminuir el contenido emocional de las palabras contenidas en su ámbito. Es por ello que las trataremos de igual manera que hacíamos con los modificadores, ya que podríamos considerar a este tipo de negación un modificador reductor de emociones. Al igual que hicimos con los modificadores, nos hemos basado en el trabajo de Carrillo (2011) para obtener un listado de palabras que se corresponden con negadores nominales. El listado lo podemos ver en la Tabla 6.1 donde aparece la palabra usada en el trabajo de Carrillo (2011), la traducción al español que hemos realizado y el porcentaje de reducción que habría que aplicar a la carga emocional de las palabras contenidas en el ámbito del modificador.

Modificador en inglés	Modificador traducido	% de actuación
Hardly	difícilmente	-85 %
Hardly	duramente	-85 %
Relatively	relativamente	-50 %
Ridiculously	ridículamente	-65 %
Minor	menos	-75 %
Only	solamente	-25 %
Small	pequeño/pequeña	-50 %
Almost	casi	-20 %
Low	bajo/baja	-75 %
Low	escaso/escasa	-75 %
Low	reducido/reducida	-75 %
Few	pocos/pocas	-75 %
Bit	poco/poca	-35 %
Barely	apenas	-75 %
Difficult	difícil	-65 %

Tabla 6.1: Listado de modificadores considerados en la negación nominal de nuestro trabajo.

A pesar de que el algoritmo es el mismo que se usa para los modificadores, explicado en el Capítulo 5, por completitud vamos a hacer un pequeño repaso de cómo actuarían estas negaciones en una frase. Como ejemplo, usamos la frase “Mi madre estaba casi enfadada cuando salí de casa” cuyo árbol de dependencias se puede ver en la Figura 6.1. Esta frase será enviada al servicio de grados emocionales de una frase (ver Sección 3.2.2). Los pasos que llevaría a cabo este servicio serían:

1. Análisis morfosintáctico. Detectamos las categorías gramaticales de cada una de las palabras y decidiremos si son palabras candidatas a ser emocionales.
2. Búsqueda de negaciones. Se recorren las palabras candidatas a ser emocionales para conocer, utilizando el árbol sintáctico, si se ven afectadas por alguna negación. Para saber si son afectadas o no, se busca cada palabra que influye sobre las palabras candidatas y se compara con el listado de palabras que consideramos iniciadoras de

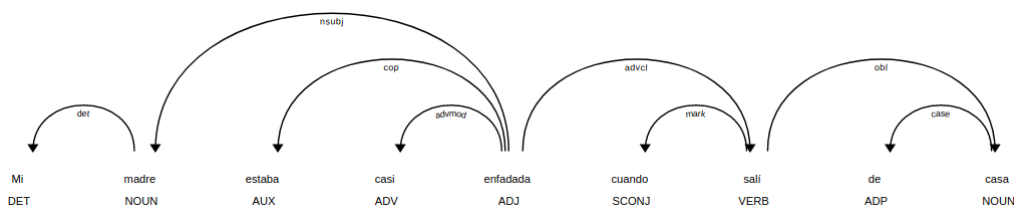


Figura 6.1: Árbol de dependencias creado por Spacy de la frase “Mi madre estaba casi enfadada cuando salí de casa”

una negación. Esta búsqueda genera un listado de tuplas con tres valores donde el primer valor es la palabra emocional, el segundo la posición que ocupa la palabra en la oración y el tercero es la suma de los porcentajes de actuación que llevan las negaciones que afectan a la palabra. En el ejemplo nos quedaría la siguiente lista:

```
[['madre', 1, 0], ['estaba', 2, 0], ['enfadada', 4, -20],
['salí', 6, 0], ['casa', 8, 0]]
```

donde vemos que “enfadada” es la única palabra afectada por la negación nominal.

3. Búsqueda de grados emocionales de cada palabra. Después de todo lo anterior se envía cada palabra candidata a ser emocional al servicio de grados emocionales de una palabra (ver Sección 3.2.1) y los resultados serán almacenados para su posterior procesamiento. Los resultados de los grados emocionales de las palabras de la frase de ejemplo se pueden ver en la Tabla 6.2.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
madre	0,86	0,52	3,79	0,48	0,0
enfadada	1,75	1,4	0,2	2,35	0,08
salí	0,37	0,97	2,63	0,23	0,1
casa	0,33	0,23	2,77	0,23	0,17

Tabla 6.2: Grados emocionales de las palabras emocionales de la frase “Mi madre estaba casi enfadada cuando salí de casa”.

4. Procesamiento de emociones de una frase. Utilizando el listado obtenido en el paso anterior donde tendremos relacionada cada palabra emocional (que aparece en el diccionario) con sus grados emocionales, comprobaremos cada una en el listado de negaciones con el porcentaje que aparezca en la tupla correspondiente. Cuando todos los valores están actualizados se procede a calcular las emociones finales de la frase, sumando para cada categoría emocional los grados de cada palabra y después dividiendo esta suma entre el número de palabras emocionales de la frase. El resultado de los grados emocionales de la frase de ejemplo y el proceso de cómo se calculan se puede ver en la Tabla 6.3.

### 6.3. Tratamiento de la negación sintáctica

Las **negaciones sintácticas** son aquellas que afectan al verbo de la frase. Consideraremos las negaciones sintácticas a las que contienen la palabra “no”. Para poder determinar

Palabra	% Neg. Nominal	Tristeza	Miedo	Alegría	Enfado	Asco
madre	0	0,86	0,52	3,79	0,48	0,0
enfadada	-20 %	1,4	1,12	0,16	1,88	0,064
salí	0	0,37	0,97	2,63	0,23	0,1
casa	0	0,33	0,23	2,77	0,23	0,17
Suma		2,96	2,84	9,35	2,82	0,334
Nº de palabras		4				
Resultado		0,74	0,71	2,34	0,71	0,2

Tabla 6.3: Grados emocionales de la frase “Mi madre estaba casi enfadada cuando salí de casa”.

el ámbito de afectación de este tipo de negaciones nos ayudaremos del árbol de dependencias que realiza Spacy. Recordemos la peculiaridad de que Spacy indica que una palabra es dependiente de otra y no qué palabras dependen de otra, lo que nos hace más costoso descubrir qué palabras se ven afectadas por el “no”, porque tendremos que recorrer todo el árbol entero para ver qué palabras incluyen en sus hijos el “no”. La forma de recorrer el árbol de dependencias y de obtener las palabras en su ámbito dependerá de si estamos tratando una frase con predicado nominal o verbal.

En la frase con predicado verbal “No tenemos dinero”, cuyo árbol mostramos en la Figura 6.2, podemos ver que la negación está afectando directamente al verbo, pero que el verbo es complementado por “dinero” (es un complemento del verbo, sintácticamente hablando). Por ello vamos a considerar que la negación afecta sobre el verbo y sus complementos que serán sus nodos hijos en el árbol de dependencias. Por ejemplo en la frase “No vivimos en esa casa fea”, cuyo árbol podemos ver en la Figura 6.3, vemos que la negación afecta directamente al verbo “vivimos” pero que a su vez “vivimos” es afectado por “casa”. Al igual que en el ejemplo “No tenemos dinero”, consideraremos que la negación afecta al verbo y a sus nodos hijos, que son sus complementos. Por eso diremos que las palabras en el ámbito de la negación en este caso son: “vivimos” y “casa”. A pesar de que “casa” posea más palabras que le afectan, cortaremos el ámbito de la negación en “casa”, es decir, no extenderemos el ámbito a los nodos hijos de las palabras que afecten a “vivimos” porque hemos considerado que la negación sintáctica afecta al verbo y a sus complementos, pero no a los complementos de los adjetivos. Es por eso, que en esta frase no incluimos “fea” en el ámbito de la negación. En la frase con predicado nominal “La niña no es fea”, cuyo árbol podemos ver en la Figura 6.4. Vemos que la negación influye sobre la palabra “fea” pero no sobre el verbo y esto es diferente que en los otros dos casos que hemos visto, ya que el verbo no es la palabra afectada directamente por la negación. Para marcar entonces el ámbito de estos casos seguiremos el mismo método que usábamos antes, eligiendo como ámbito la palabra afectada directamente por la negación y sus nodos hijos, por lo que en este caso estarán afectadas las palabras “es” y “fea”. Un ejemplo más con un verbo copulativo con más complementos es “Ella no es una niña fea”, cuyo análisis vemos en la Figura 6.5, donde al igual que en el ejemplo anterior la negación afecta directamente sobre “niña” y no sobre el verbo porque es copulativo. Siguiendo la misma norma que en el resto de ocasiones, la negación va a afectar a la palabra sobre la que afecta directamente, “niña”, y a sus nodos hijos, en este caso “fea”.

En las siguientes subsecciones explicaremos los distintos métodos que hemos explorado para tratar la negación sintáctica.

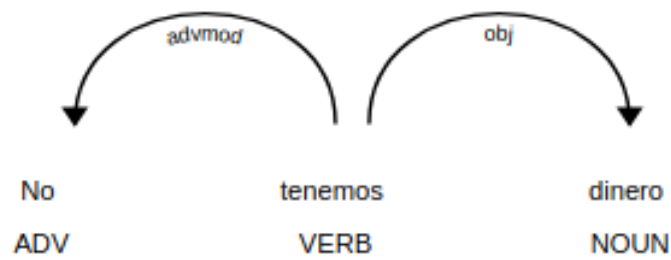


Figura 6.2: Árbol de dependencias creado para la frase "No tenemos dinero"

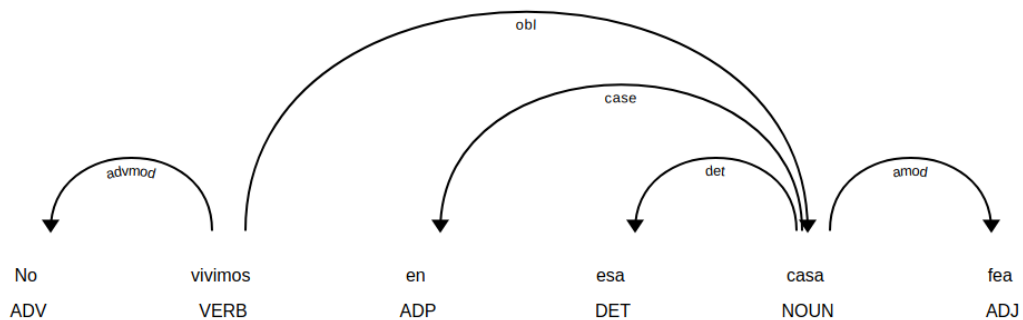


Figura 6.3: Árbol de dependencias creado para la frase "No vivimos en esa casa fea"

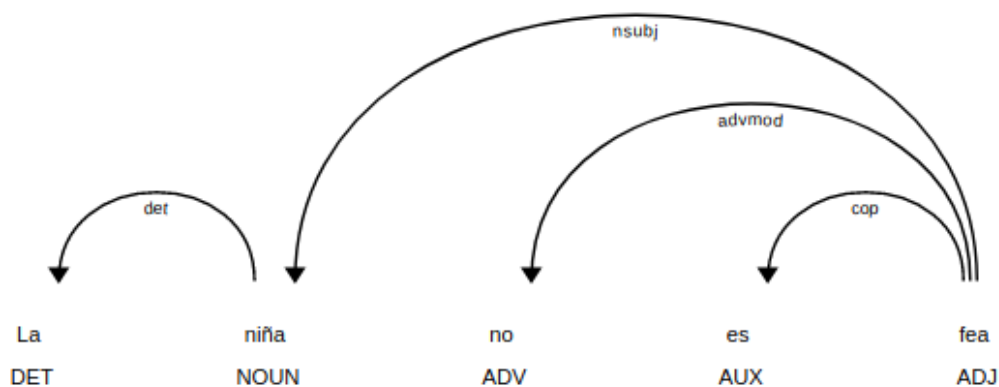


Figura 6.4: Árbol de dependencias creado para la frase "La niña no es fea"

### 6.3.1. Anulación de la aportación emocional de palabras negadas

La primera opción que planteamos consistía en hacer que las palabras que están afectadas por la negación dejen de aportar su carga emocional al conjunto. En la frase "No estoy enfadada", la palabra "enfadada" será la única que aporte emociones a la frase, dado que el verbo "estoy" no es considerada una palabra emocional por nuestro sistema. La palabra "enfadada" además de ser la única emocional también es la única afectada por la negación. Por ello, anularíamos las emociones de esta palabra dando como resultado que la frase es neutra y no transmite emociones de la frase diciendo que la frase es tranquila y no transmite emociones. Este método es una mejora en nuestro sistema, ya que el EmoTraductor 1.0 diría que esta frase transmite enfado cuando realmente es justo lo contrario de lo que

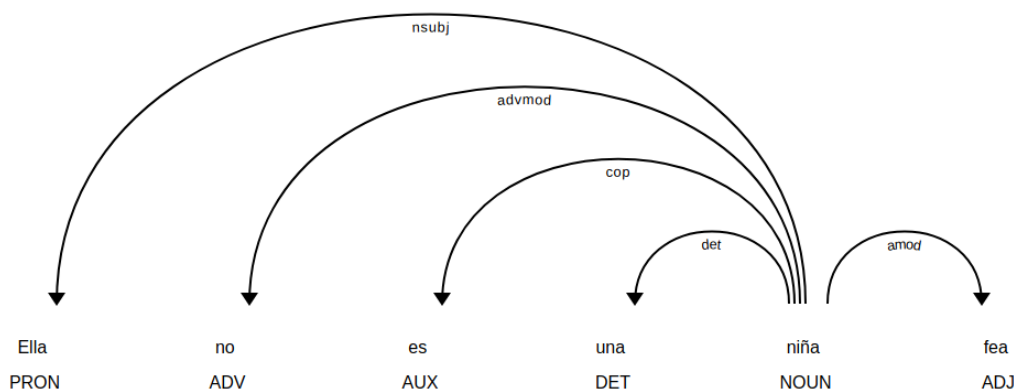


Figura 6.5: Árbol de dependencias creado para la frase “Ella no es una niña fea”

queremos expresar.

Veamos unos ejemplos más complejos con las frases: “Me gustó ir al cine” y “No me gustó ir al cine”. Para ambas frases partimos de las mismas palabras emocionales cuyos valores podemos ver en la primera parte de la Tabla 6.4, donde vemos que “gustó” y “cine” tienen como emoción significativa la alegría e “ir” no posee ninguna emoción con un grado mayor a 1,5 (con el que comenzaríamos a considerarla significativa). Si no tuviésemos implementada la negación el resultado de ambas frases sería el mismo como se ve en la Tabla 6.4, con la alegría como emoción más significativa. Pero al añadir el tratamiento de la negación identificamos que las palabras “gustó” e “ir” están afectadas por la negación y por tanto deben perder sus grados emocionales, quedando el resultado que vemos en la última fila de Tabla 6.5 junto con el proceso completo del cálculo. Ahora la frase con negación no tiene ninguna emoción significativa, lo que supone ya una mejora en nuestro sistema y mucho más cercano a la interpretación que cualquier persona haría de ella.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
gustó	0,17	0,6	3,43	0,2	0,17
ir	0,1	0,2	1,9	0,45	0,05
cine	1,2	1,17	3,53	0,87	0,47

Suma de valores	1,47	1,97	7,41	1,52	0,69
NºPalabras	3				
Resultado final	0,49	1,2	3,53	0,87	0,47

Tabla 6.4: Grados emocionales de la frase “Me gustó ir al cine” y “No me gustó ir al cine” con el EmoTraductor 1.0

A pesar de que vemos un comportamiento interesante en el sistema con este pequeño cambio, hemos analizado el tratamiento de la emoción con los textos que tenemos destinados para ello. Los resultados de esta evaluación se encuentran en la Tabla 6.6 junto con los resultados que teníamos tras las primeras mejoras realizadas en el sistema para poder comparar los resultados. Vemos que precisamente en los textos que hemos escogido para probar la negación (“DobleNegación2”, “Negación”, “Negacion2” o “Negaciones coordinadas”) obtenemos un mejor resultado, más cercano a la realidad, que tras las primeras mejoras. Por ejemplo, en el texto “DobleNegación2”, que dice “No he visto en mi vida un

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
gustó	0	0	0	0	0
ir	0	0	0	0	0
cine	1,2	1,17	3,53	0,87	0,47
Suma de valores	1,2	1,17	3,53	0,87	0,47
NºPalabras	3				
Resultado final	0,4	0,39	1,18	0,29	0,16

Tabla 6.5: Grados emocionales de la frase “No me gustó ir al cine” con la primera aproximación implementada para tratar la negación.

paisaje tan hermoso”, el grado de alegría pasa de un 2,37 a 0,83 dejando de aparecer como emoción significativa en el texto, ya que en este texto corto efectivamente podemos decir que realmente no estamos transmitiendo ninguna emoción o algo de tristeza.

Este método por lo que hemos visto mejora nuestros resultados pero no siempre la negación anula las emociones, existen casos donde una negación de una emoción provoca el aumento de otra emoción. Por ejemplo en la frase “No me gusta el brócoli” al negar el verbo “gustar” disminuimos la alegría, pero en esta frase deberíamos aumentar el asco cuando indicamos que no nos gusta, en la frase “No te quiero ni ver” al negar el verbo “quiero” disminuimos la alegría pero para que la carga emocional sea adecuada deberíamos aumentar los grados de tristeza y de enfado. Para llegar a esta opción exploraremos otros métodos.

### 6.3.2. Negación basada en antónimos de palabras

El primer planteamiento que tenemos para aplicar este tipo de negación es buscar antónimos de las palabras que se ven afectadas por la negación. Para encontrar estas palabras antónimas se utilizó un servicio web (Sheila Plaza y Acosta, 2016) que devuelve para una palabra su antónimo, de tal manera que sustituyéramos los grados de las palabras emocionales afectadas por los grados emocionales de su palabra antónima en caso de existir en el diccionario, y si no, considerarla como no emocional. Por ejemplo, en la oración “No vamos al cine”, “vamos” está afectado por el “no” y por tanto buscaríamos el antónimo de “ir” por ser el verbo en infinitivo de “vamos”. El servicio web de antónimos nos devuelve el verbo “venir” como antónimo. Los grados emocionales de ambas palabras los podemos ver en la Tabla 6.7, donde vemos que a pesar de que “venir” sea el antónimo de “ir” no son emocionalmente contrarios sino más bien semejantes o al menos con la misma emoción mayoritaria. Viendo la situación con otros verbos como por ejemplo “coger”-“soltar”, “tirar”-“coger”, “regalar”-“vender” vemos que pasa lo mismo que con “ir”-“venir”, siendo esto un problema. Otro problema encontrado es que algunos verbos como “bailar”, “ver” o “besar” no tienen ningún antónimo según el servicio web. Un problema aún mayor es qué deberíamos hacer con las palabras que están afectadas también por la negación, en el primer ejemplo “cine” que no es un verbo, no es correcto en todos los casos cambiar la palabra por su contrario, incluso en caso de que tengan antónimos porque ocurriría algo similar a lo que ocurre con los verbos. Por ejemplo, con las palabras “ruido” cuyos grados emocionales son 0,57 de tristeza, 1 de miedo, 0,6 de alegría, 1,73 de enfado y 0,53 de asco y “silencio” cuyos grados emocionales son 1,2 de tristeza, 1,1 de miedo, 1,65 de alegría, 0,35 de enfado y 0,25 de asco; vemos que a pesar de ser antónimas lingüísticamente comparando su carga emocional

Texto	Sin negación					Negación eliminación de emociones				
	T	M	Al	E	A	T	M	Al	E	A
Aladdin	0,53	0,81	2,54	0,83	0,22	0,43	0,64	2,08	0,27	0,21
Blancanieves	0,87	0,88	1,59	0,23	0,57	0,87	0,88	1,59	0,83	0,57
Blancanieves y sus amigos	0,37	0,46	2,82	0,42	0,18	0,37	0,46	2,82	0,23	0,18
Blog agradecimientos	0,5	0,44	2,18	0,35	0,27	0,44	0,35	1,78	0,37	0,23
Blog enfado	0,4	0,49	1,33	0,25	0,3	0,33	0,4	1,15	0,25	0,23
Blog moda	0,44	0,35	0,87	0,56	0,17	0,44	0,35	0,87	0,25	0,17
Blog muerte	0,84	0,62	1,33	0,58	0,28	0,84	0,63	1,31	0,56	0,29
Blog reflexion	0,71	0,71	1,62	0,35	0,37	0,64	0,62	1,39	0,52	0,34
Cenicienta	0,48	0,5	2,34	0,47	0,32	0,49	0,5	2,34	0,34	0,31
CentroComercial	0,56	0,56	1,25	0,21	0,36	0,55	0,56	1,24	0,46	0,35
Conversación1	0,19	0,22	0,99	0,34	0,16	0,19	0,22	0,99	0,21	0,16
Conversación2	0,42	0,37	1,4	0,3	0,28	0,45	0,37	1,3	0,39	0,31
DobleNegación1	0,16	0,27	2,37	0,4	0,09	0,11	0,18	1,58	0,2	0,06
DobleNegación2	0,61	0,66	2,65	2,36	0,29	0,05	0,1	0,83	0,04	0,03
DobleNegación3	1,62	1,48	0,71	0,43	3,79	1,51	1,35	0,2	2,27	3,4
El castillo de bestia	0,47	0,68	2,53	0,43	0,22	0,49	0,7	2,51	0,43	0,23
El ratoncito perez	0,43	0,64	1,5	0,32	0,34	0,41	0,57	1,42	0,38	0,31
FraseCorta	0,54	0,41	1,72	0,39	0,28	0,46	0,34	1,68	0,32	0,28
FraseCorta2	0,66	0,61	1,47	0,54	0,24	0,64	0,61	1,47	0,39	0,25
La bella durmiente	0,58	0,8	1,96	0,4	0,28	0,57	0,78	1,95	0,53	0,27
La bella y la bestia	0,64	0,88	2,22	0,23	0,23	0,65	0,93	2,33	0,41	0,21
La margarita	0,4	0,38	2,19	0,35	0,17	0,41	0,4	2,1	0,25	0,19
La sirenita	0,61	0,89	2,22	0,72	0,25	0,52	0,72	1,76	0,35	0,22
Mulan	0,68	0,76	1,86	0,32	0,46	0,65	0,73	1,63	0,69	0,46
Negación	0,6	0,6	1,97	0,47	0,23	0,18	0,2	1,15	0,1	0,1
Negación2	0,53	0,57	1,08	0,37	0,31	0,35	0,27	0,46	0,24	0,08
negaciones	0,69	0,8	1,5	0,76	0,34	0,34	0,32	0,17	0,18	0,17
Negaciones coordinadas	1,18	0,99	1,29	0,55	0,47	0,62	0,53	0,88	0,46	0,28
Noticia deportes	0,53	0,66	1,4	1,04	0,4	0,51	0,66	1,33	0,54	0,39
Noticia muerte	1,08	1,17	1,11	0,39	0,7	1,04	1,14	1,07	0,99	0,69
Noticia nacimiento	0,46	0,57	1,76	0,63	0,29	0,45	0,55	1,68	0,38	0,28
Noticia OMS	0,81	0,82	1,22	0,75	0,57	0,8	0,81	1,19	0,62	0,56
Noticia política	0,7	0,69	1,37	0,82	0,55	0,61	0,56	1,36	0,7	0,5
Peli aburrida	0,92	0,81	1,4	0,43	0,53	0,95	0,66	1,23	0,64	0,44
Pinocho	0,59	0,57	1,37	0,41	0,32	0,57	0,56	1,33	0,4	0,31
Pocahontas	0,75	0,84	1,95	0,28	0,27	0,75	0,84	1,95	0,41	0,27
Preguntas triste oración	0,32	0,42	0,96	0,59	0,14	0,29	0,36	0,78	0,22	0,11
Texto miedo perder amor	0,74	0,94	1,46	0,78	0,38	0,63	0,82	1,3	0,5	0,33
Texto1	0,8	0,86	1,22	0,44	0,52	0,79	0,84	1,06	0,74	0,49
Texto2	0,54	0,59	1,5	0,19	0,33	0,55	0,59	1,39	0,44	0,33
Tipos	0,16	0,31	2,05	0,49	0,09	0,16	0,3	1,9	0,18	0,08
Tristeza	0,84	0,67	1,01	0,47	0,35	0,68	0,52	0,74	0,36	0,26
Tristeza2	0,7	0,77	0,88		0,36	0,58	0,58	0,71	0,39	0,31

Tabla 6.6: Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir la negación eliminando los grados de las palabras en el ámbito de la negación. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).

no lo son. Por todas estas razones decidimos descartar esta opción del tratamiento de la negación sintáctica.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
Ir	0,1	0,2	1,9	0,45	0,05
Venir	0,1	0,25	1,15	0,25	0,2

Tabla 6.7: Grados emocionales en el diccionario afectivo para las palabras “ir” y “venir”

### 6.3.3. Inversión de los grados emocionales de las palabras en el ámbito de la negación

En el trabajo de Francisco (2008), cuando se encuentra una negación y se quiere evaluar una frase según las dimensiones emocionales, se invierte el valor de cada dimensión. El proceso consiste en utilizar el valor máximo que puede tener cada dimensión emocional (10) y restarle el valor que tiene la dimensión emocional que se desea invertir. Por ejemplo, si tuviéramos las siguientes dimensiones emocionales activación=8, evaluación=5 y control=3, al aplicar una negación quedaría: activación=10-8=2, evaluación=10-5=5 y control=10-3=7.

Siguiendo la idea que utiliza Francisco (2008) con la negación en las dimensiones emocionales, vamos a plantear un método para tratar la negación sintáctica con 5 categorías emocionales y usando los grados de cada una en lugar de dimensiones emocionales.

Nuestra propuesta es invertir el valor de los grados de cada categoría emocional de la misma manera, teniendo en cuenta que el máximo grado que puede tener una categoría emocional es 4. Por ejemplo, en la frase “No iremos al cine” tenemos como palabras afectadas por el “no” a “iremos” y “cine”. Si buscamos estas en el diccionario, donde las palabras están calificadas para cada categoría emocional entre el 0-4, encontramos los valores que podemos ver en la Tabla 6.8 en la columna “Grados emocionales”. Si no tuviéramos en cuenta la negación el resultado de esta frase sería el que podemos ver en la Tabla 6.9 en la columna “Sin considerar negación”, donde vemos que la alegría predomina sobre el resto de emociones. Usando el método explicado para las dimensiones emocionales, transformaríamos los grados de las palabras “iremos” y “cine” haciendo la resta de 4 (máximo grado en el diccionario) menos el valor actual de la emoción, resultando los grados que podemos ver en la Tabla 6.8 en la columna “Grados emocionales invertidos”. Ahí vemos que se han invertido las emociones y la diferencia es bastante grande, y si usamos estos grados para calcular el total de la frase obtendremos los resultados que aparecen en 6.9 en la columna “Considerando negación”, donde se ve claramente que las emociones han sido invertidas y por tanto la negación influye lo suficiente para cambiar todas las emociones de la frase.

Analizando estos resultados nos hacemos la pregunta de si realmente esta inversión es correcta y si realmente la frase “No iremos al cine” produce tanta emoción como podemos ver en el ejemplo, ya que parece que hay cierta exageración de las emociones. Aún así, para evaluarlo correctamente lanzamos las pruebas sobre los textos del Anexo A. Los resultados de las pruebas son los que podemos ver en la Tabla 6.10, donde si nos fijamos en los textos que contienen negaciones (“DobleNegación2”, “DobleNegación3”, “Negación”, “Negación2” y “Negaciones”), podemos ver que cambian significativamente todas las emociones, pero en vez de resultar una emoción o las contrarias a las que teníamos, ahora, todas las emociones son significativas. En concreto, en el texto “DobleNegación2” vemos que los resultados varían pasando la tristeza de un 0,61 a 2,49, el miedo de 0,66 a 2,54, la alegría de 2,65

Palabra	Emociones	Grados emocionales	Grados emocionales invertidos
Iremos	Tristeza	0,1	3,9
	Miedo	0,2	3,8
	Alegría	1,9	2,1
	Enfado	0,45	3,55
	Asco	0,05	3,95
Cine	Tristeza	1,2	2,8
	Miedo	1,17	2,83
	Alegría	3,53	0,47
	Enfado	0,87	3,13
	Asco	0,45	3,53

Tabla 6.8: Grados emocionales del diccionario e invertidos de las palabras “ir” y “cine”

Emoción	Sin considerar negación	Considerando negación
Tristeza	0,65	3,35
Miedo	0,69	3,31
Alegría	2,73	1,29
Enfado	0,66	3,34
Asco	0,26	3,74

Tabla 6.9: Grados emocionales sin considerar y considerando la negación para la frase “No iremos al cine”

a 2,02, el enfado de 0,4 a 2,69 y el asco de 0,29 a 2,77. Ahora este texto tiene todas las emociones como significativas, lo cual no es correcto.

Estos resultados nos muestran como este método que aparentemente tenía bastante sentido se ha convertido en un aumentador de emociones poco significativas y una posible forma de confundir aún más al usuario que esté usando la aplicación, ya que en la aplicación web, en la barra emocional, aparecerían más emociones de las que realmente transmite el texto. Por lo tanto, descartamos esta solución.

#### 6.3.4. Inversión de los grados emocionales significativos de las palabras en el ámbito de la negación

Siguiendo la misma línea que en el método anterior, y evaluando los problemas encontrados, vemos que el posible fallo del método es tener en cuenta para la inversión de valores todos los grados de la categorías. Es decir, para una frase donde las emociones fueran todo ceros excepto la alegría que tuviera un 2,5, al invertirla tendríamos todo 4 excepto en la alegría que tendríamos un 1,5. Sin embargo, nos damos cuenta de que quizá solo debemos invertir las emociones que tengan peso suficiente como para ser mostradas en la aplicación web, las que superen el 1,5. Por ejemplo, si invirtiésemos una palabra con los siguientes valores: tristeza=0,2, miedo=1,6, alegría=2,2, enfado=0,79 y asco=1,4 únicamente invertiríamos las emociones que superen el 1,5, en este caso miedo y alegría, y el resto las dejaríamos igual. El objetivo de este método es no aumentar las emociones no significativas ya que si no aportaban una carga significativa antes de negarlas tampoco deberían aportarla por el simple hecho de ser negadas.

Texto	Sin negación					Negación inversión completa				
	T	M	Al	E	A	T	M	Al	E	A
Aladdin	0,53	0,81	2,54	0,31	0,22	0,91	1,04	2,19	0,81	0,78
Blancanieves	0,87	0,88	1,59	0,83	0,57	0,87	0,88	1,59	0,83	0,57
Blancanieves y sus amigos	0,37	0,46	2,82	0,23	0,18	0,37	0,46	2,82	0,23	0,18
Blog agradecimientos	0,5	0,44	2,18	0,42	0,27	0,99	0,88	2,07	0,95	0,81
Blog enfado	0,4	0,49	1,33	0,35	0,3	0,79	0,84	1,48	0,69	0,7
Blog moda	0,44	0,35	0,87	0,25	0,17	0,44	0,35	0,87	0,25	0,17
Blog muerte	0,84	0,62	1,33	0,56	0,28	0,84	0,63	1,31	0,56	0,29
Blog reflexion	0,71	0,71	1,62	0,58	0,37	1,07	1,04	1,68	0,96	0,81
Cenicienta	0,48	0,5	2,34	0,35	0,32	0,49	0,5	2,34	0,34	0,31
CentroComercial	0,56	0,56	1,25	0,47	0,36	0,64	0,66	1,31	0,56	0,45
Conversación1	0,19	0,22	0,99	0,21	0,16	0,19	0,22	0,99	0,21	0,16
Conversación2	0,42	0,37	1,4	0,34	0,28	0,45	0,37	1,3	0,39	0,31
DobleNegación1	0,16	0,27	2,37	0,3	0,09	0,5	1,14	2,87	1,03	1,13
DobleNegación2	0,61	0,66	2,65	0,4	0,29	2,49	2,54	2,02	2,69	2,77
DobleNegación3	1,62	1,48	0,71	2,36	3,79	2,73	2,55	1,01	3,5	4,34
El castillo de bestia	0,47	0,68	2,53	0,43	0,22	0,49	0,7	2,51	0,43	0,23
El ratoncito perez	0,43	0,64	1,5	0,43	0,34	0,58	0,72	1,54	0,54	0,49
FraseCorta	0,54	0,41	1,72	0,32	0,28	0,46	0,34	1,68	0,32	0,28
FraseCorta2	0,66	0,61	1,47	0,39	0,24	0,64	0,61	1,47	0,39	0,25
La bella durmiente	0,58	0,8	1,96	0,54	0,28	0,57	0,78	1,95	0,53	0,27
La bella y la bestia	0,64	0,88	2,22	0,4	0,23	0,65	0,93	2,33	0,41	0,21
La margarita	0,4	0,38	2,19	0,23	0,17	0,41	0,4	2,1	0,25	0,19
La sirenita	0,61	0,89	2,22	0,35	0,25	0,71	0,89	1,78	0,55	0,4
Mulan	0,68	0,76	1,86	0,72	0,46	0,65	0,73	1,63	0,69	0,46
Negación	0,6	0,6	1,97	0,32	0,23	1,62	1,63	2,03	1,75	1,82
Negación2	0,53	0,57	1,08	0,47	0,31	2,11	1,92	1,87	1,98	1,82
negaciones	0,69	0,8	1,5	0,37	0,34	3,32	3,16	2,18	3,33	3,34
Negaciones coordinadas	1,18	0,99	1,29	0,76	0,47	1,39	1,4	1,8	1,49	1,43
Noticia deportes	0,53	0,66	1,4	0,55	0,4	0,62	0,77	1,39	0,66	0,51
Noticia muerte	1,08	1,17	1,11	1,04	0,7	1,12	1,22	1,17	1,07	0,79
Noticia nacimiento	0,46	0,57	1,76	0,39	0,29	0,66	0,74	1,82	0,58	0,48
Noticia OMS	0,81	0,82	1,22	0,63	0,57	0,92	0,94	1,29	0,75	0,69
Noticia política	0,7	0,69	1,37	0,75	0,55	0,75	0,7	1,49	0,83	0,64
Peli aburrida	0,92	0,81	1,4	0,82	0,53	0,95	0,66	1,23	0,64	0,44
Pinocho	0,59	0,57	1,37	0,43	0,32	0,68	0,66	1,41	0,48	0,41
Pocahontas	0,75	0,84	1,95	0,41	0,27	0,75	0,84	1,95	0,41	0,27
Preguntas triste oración	0,32	0,42	0,96	0,28	0,14	0,84	0,9	1,2	0,76	0,68
Texto miedo perder amor	0,74	0,94	1,46	0,59	0,38	1	1,21	1,64	0,89	0,77
Texto1	0,8	0,86	1,22	0,78	0,52	1,09	1,15	1,22	1,03	0,79
Texto2	0,54	0,59	1,5	0,44	0,33	0,77	0,8	1,53	0,66	0,56
Tipos	0,16	0,31	2,05	0,19	0,09	0,43	0,57	2,03	0,46	0,36
Tristeza	0,84	0,67	1,01	0,49	0,35	1,13	0,98	1,08	0,84	0,77
Tristeza2	0,7	0,77	0,88	0,47	0,36	0,79	0,74	0,91	0,66	0,61

Tabla 6.10: Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir la negación por inversión de todas las emociones. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).

Vamos a ver con la frase “Ella no baila en la fiesta” como evolucionarían los resultados pasando por un primer resultado donde no se tienen en cuenta las negaciones, después un segundo resultado con el método de inversión de todas las emociones, y por último utilizando este nuevo método. En la frase, primero buscaríamos las palabras candidatas a ser emocionales, resultando que son “baila” y “fiesta”. Después buscaríamos a qué afecta la negación. En este caso está afectando a las palabras “baila” y “fiesta”. Y después terminaríamos diciendo que las dos palabras son emocionales por aparecer en el diccionario.

En todos los casos éste sería el punto de partida, y a la hora de interpretar los resultados sería cuando veríamos las diferencias entre métodos. Una vez obtenidos del diccionario los resultados de las palabras “baila” y “fiesta” (que podemos ver en la primera columna de la Tabla 6.11) vamos a buscar aquellos grados que superen el 1,5 para invertirlos (4 menos el valor de la categoría emocional). Para nuestro ejemplo, podemos ver los resultados de esta resta en la segunda columna de la Tabla 6.11, donde solo se han invertido los correspondientes a alegría por ser los únicos que en el diccionario superan el 1,5. Los nuevos grados calculados se utilizarán para calcular la carga emocional de la frase, haciendo como siempre la suma de todos los grados por categoría y dividiendo la suma entre el número de palabras emocionales que posee la frase.

Palabra	Emoción	Sin considerar negación	Considerando negación
Baila	Tristeza	0,05	0,05
	Miedo	0,1	0,1
	Alegría	2,55	1,45
	Enfado	0,14	0,14
	Asco	0,05	0,05
Fiesta	Tristeza	0,13	0,13
	Miedo	0,28	0,28
	Alegría	3,69	0,31
	Enfado	0,07	0,07
	Asco	0,07	0,07

Tabla 6.11: Grados emocionales de la palabras “baila” y “fiesta” según el diccionario y después de aplicar el método de inversión para las emociones que tienen un grado mayor que 1,5

En la Tabla 6.12 vemos los resultados finales para la frase utilizando los distintos métodos (sin negación, invirtiendo todas las emociones e invirtiendo únicamente las emociones significativas). Podemos observar que con el nuevo método regulamos las emociones haciendo que esta frase sea neutra o tranquila porque no transmite emociones. Realmente, si el sistema tuviera todo el conocimiento que puede tener una persona adulta sin ningún tipo de dificultades para reconocer emociones, podríamos decir que verdaderamente esta frase sí que puede connotar un poco de tristeza o alguna otra emoción y no es 100 % neutral. Sin embargo, el paso de un método a otro nos da una mejora considerable.

No obstante, se ha evaluado también este método con los textos que se encuentran en el Anexo A. Los resultados de esta evaluación se encuentran en la Tabla 6.13. Si nos fijamos en los textos con negaciones (“DobleNegación2”, “DobleNegación3”, “Negación”, “Negación2”, “Negaciones” y “Negaciones coordinadas”) vemos que cambian significativamente todas las emociones y en algunos casos encontramos el mismo problema que ocurría con el método de inversión de todas las emociones (ver 6.3.3). Por ejemplo, en el texto “DobleNegación2” con 2,49 en tristeza, 2,54 en miedo, 2,02 de alegría, 2,69 de enfado y 2,77 de asco, una vez

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
Sin negación	0,09	0,19	3,12	0,11	0,06
Inversión todas emociones	3,91	3,81	0,88	3,9	3,94
Inversión emociones >1,5	0,09	0,19	0,88	0,11	0,06

Tabla 6.12: Evaluación de la frase “Ella no baila en la fiesta” sin tener en cuenta la negación, teniendo en cuenta la negación usando el primer método de inversión y con el nuevo método de inversión de emociones superiores a 1,5

más todas las emociones son significativas no asemejándose a la realidad. Lo mismo ocurre con el resto de textos citados.

Este método nos reduce las emociones en algunas frases quedando la mayoría como sin emociones significativas, y en textos completos seguimos teniendo el problema anterior de emociones exageradas.

### 6.3.5. Uso de la antonimia de las categorías emocionales

Basándonos en las ideas de las tesis doctorales de Francisco (2008) y Carrillo (2011), para el tratamiento de la negación en categorías emocionales vamos a crear distintas relaciones de antonimia entre las emociones que contemplamos. Al trabajar únicamente con las cinco emociones básicas, estas relaciones de antonimia son complicadas de elaborar porque no todas nuestras categorías emocionales tienen un antónimo entre ellas, por ejemplo la categoría emocional asco no es antónimo ni de la tristeza, ni del miedo, ni de la alegría ni del enfado. Por esto vamos a establecer relaciones de antonimia completas y parciales. Una relación de antonimia completa significa que las categorías son completamente antónimas, mientras que las parciales indican que cuando negamos una palabra con la categoría emocional antónima parcial, la categoría antónima aumentará parcialmente su emoción. Hemos decidido que la tristeza y la alegría tienen una relación de antonimia completa y la alegría con el enfado en un 50 % y que el miedo y el asco no tienen ninguna relación de antonimia con ninguna otra categoría. Estos porcentajes los hemos decidido usando ejemplos como las frases “Estoy contenta con tu comportamiento” y “No estoy contenta con tu comportamiento”. La primera frase denota alegría mientras que la segunda denotará tristeza y un poco de enfado.

Para aplicar este tipo de relaciones hemos diferenciado los grados emocionales en dos grupos. Si el grado de una emoción está entre el 0 y el 2 lo consideraremos una emoción débil, mientras que si el grado está entre el 2 y el 4 será considerada una emoción fuerte.

El proceso de negación por antonimia funciona de la siguiente manera. Cuando una palabra es detectada como negada procederemos a la inversión de todos sus grados emocionales siguiendo un método u otro si la emoción es considerada débil o fuerte. Si la emoción es menor que 2 (débil), utilizaremos el 2 como grado máximo que tendría esta emoción y haríamos una inversión utilizando este máximo, de manera que tendríamos la resta 2-grados emocionales para calcular el inverso. Si la emoción, por el contrario, es mayor o igual que 2, utilizaremos el 4 como grado máximo que podría tener la emoción y haríamos la inversión utilizando este máximo, de manera que tendríamos la resta 4-grados emocionales para calcular el inverso. Por ejemplo, con la palabra “adicto”, que tiene los grados emocionales: 2,6 de tristeza, 1,97 de miedo, 0,3 de alegría, 1,33 de enfado y 1,23 de asco. Si tuviéramos que invertirla lo haríamos de la siguiente manera:

- Tristeza. Tiene asociado un 2,6 de tristeza. Al ser un número mayor o igual que

Texto	Sin negación					Negación inversión >1,5				
	T	M	Al	E	A	T	M	Al	E	A
Aladdin	0,53	0,81	2,54	0,31	0,22	0,53	0,81	2,19	0,31	0,22
Blancanieves	0,87	0,88	1,59	0,83	0,57	0,87	0,88	1,59	0,83	0,57
Blancanieves y sus amigos	0,37	0,46	2,82	0,23	0,18	0,37	0,46	2,82	0,23	0,18
Blog agradecimientos	0,5	0,44	2,18	0,42	0,27	0,5	0,44	2	0,42	0,28
Blog enfado	0,4	0,49	1,33	0,35	0,3	0,41	0,5	1,27	0,37	0,28
Blog moda	0,44	0,35	0,87	0,25	0,17	0,44	0,35	0,87	0,25	0,17
Blog muerte	0,84	0,62	1,33	0,56	0,28	0,84	0,63	1,31	0,56	0,29
Blog reflexion	0,71	0,71	1,62	0,58	0,37	0,7	0,7	1,6	0,57	0,38
Cenicienta	0,48	0,5	2,34	0,35	0,32	0,49	0,5	2,34	0,34	0,31
CentroComercial	0,56	0,56	1,25	0,47	0,36	0,56	0,57	1,28	0,46	0,36
Conversación1	0,19	0,22	0,99	0,21	0,16	0,19	0,22	0,99	0,21	0,16
Conversación2	0,42	0,37	1,4	0,34	0,28	0,45	0,37	1,3	0,39	0,31
DobleNegación1	0,16	0,27	2,37	0,3	0,09	0,57	0,54	1,62	1,03	0,31
DobleNegación2	0,61	0,66	2,65	0,4	0,29	0,61	0,66	1,59	0,4	0,29
DobleNegación3	1,62	1,48	0,71	2,36	3,79	1,62	1,48	0,47	2,36	3,79
El castillo de bestia	0,47	0,68	2,53	0,43	0,22	0,49	0,7	2,51	0,43	0,23
El ratoncito perez	0,43	0,64	1,5	0,43	0,34	0,43	0,6	1,49	0,42	0,33
FraseCorta	0,54	0,41	1,72	0,32	0,28	0,46	0,34	1,68	0,32	0,28
FraseCorta2	0,66	0,61	1,47	0,39	0,24	0,64	0,61	1,47	0,39	0,25
La bella durmiente	0,58	0,8	1,96	0,54	0,28	0,57	0,78	1,95	0,53	0,27
La bella y la bestia	0,64	0,88	2,22	0,4	0,23	0,65	0,93	2,33	0,41	0,21
La margarita	0,4	0,38	2,19	0,23	0,17	0,41	0,4	2,1	0,25	0,19
La sirenita	0,61	0,89	2,22	0,35	0,25	0,53	0,75	1,78	0,35	0,24
Mulan	0,68	0,76	1,86	0,72	0,46	0,65	0,73	1,63	0,69	0,46
Negación	0,6	0,6	1,97	0,32	0,23	0,49	0,53	1,57	0,28	0,19
Negación2	0,53	0,57	1,08	0,47	0,31	0,59	0,69	1,17	0,51	0,33
negaciones	0,69	0,8	1,5	0,37	0,34	0,86	0,8	1,27	0,37	0,34
Negaciones coordinadas	1,18	0,99	1,29	0,76	0,47	0,98	0,88	1,35	0,76	0,47
Noticia deportes	0,53	0,66	1,4	0,55	0,4	0,52	0,67	1,36	0,55	0,4
Noticia muerte	1,08	1,17	1,11	1,04	0,7	1,06	1,16	1,11	1,02	0,73
Noticia nacimiento	0,46	0,57	1,76	0,39	0,29	0,45	0,57	1,78	0,39	0,29
Noticia OMS	0,81	0,82	1,22	0,63	0,57	0,81	0,82	1,23	0,63	0,57
Noticia política	0,7	0,69	1,37	0,75	0,55	0,67	0,6	1,38	0,77	0,57
Peli aburrida	0,92	0,81	1,4	0,82	0,53	0,95	0,66	1,23	0,64	0,44
Pinocho	0,59	0,57	1,37	0,43	0,32	0,58	0,57	1,37	0,42	0,32
Pocahontas	0,75	0,84	1,95	0,41	0,27	0,75	0,84	1,95	0,41	0,27
Preguntas triste oración	0,32	0,42	0,96	0,28	0,14	0,32	0,42	0,98	0,28	0,14
Texto miedo perder amor	0,74	0,94	1,46	0,59	0,38	0,7	0,92	1,44	0,61	0,37
Texto1	0,8	0,86	1,22	0,78	0,52	0,8	0,86	1,19	0,77	0,51
Texto2	0,54	0,59	1,5	0,44	0,33	0,56	0,61	1,5	0,45	0,34
Tipos	0,16	0,31	2,05	0,19	0,09	0,16	0,31	2,03	0,19	0,09
Tristeza	0,84	0,67	1,01	0,49	0,35	0,76	0,65	0,94	0,45	0,33
Tristeza2	0,7	0,77	0,88	0,47	0,36	0,68	0,65	0,77	0,47	0,36

Tabla 6.13: Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir el segundo método de la negación por inversión en solo las categorías emocionales que superan el 1,5. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).

2, diríamos que es una emoción fuerte y por tanto la invertiríamos tomando como máximo el 4, resultando:  $4-2,6=1,4$

- Miedo. Tiene asociado un 1,97 de miedo. Al ser un número menor que 2, diríamos que es una emoción débil y por tanto la invertiríamos tomando como máximo el 2, resultando:  $2-1,97=0,03$
- Alegría. Tiene asociado un 0,3 de alegría. Al ser un número menor que 2, diríamos que es una emoción débil y por tanto la invertiríamos tomando como mínimo el 2, resultando  $2-0,3=1,7$
- Enfado. Tiene asociado un 1,33 en enfado. Al ser un número menor que 2, diríamos que es una emoción débil y por tanto la invertiríamos tomando como máximo el 2, resultando:  $2-1,33=0,67$
- Asco. Tiene asociado un 1,23 en asco. Al ser un número menor que 2, diríamos que es una emoción débil y por tanto la invertiríamos tomando como máximo el 2, resultando  $= 2-1,23=0,77$

La finalidad de invertir de diferente manera las emociones fuertes y débiles es conseguir no aumentar desmesuradamente una emoción que no estaba presente de manera considerable en la palabra, y tomamos el 2, por ser el valor medio en la escala 0-4, para que en caso de que la emoción estuviera muy poco presente pueda aparecer al negar la palabra. Es decir, si tuviéramos un 1,4, que es casi considerado significativo, al ser invertido no puede ser un número demasiado grande ya que con la negación intentamos reducir las emociones más significativas y si hiciésemos la resta  $4-1,4$  obtendríamos un grado mayor que 1,4; mientras que si tuviéramos un 0,1 en una emoción al invertirla necesitaríamos que ahora fuese un poco más significativo pero no desmesuradamente significativo. Por lo tanto, al invertirlo se convertiría en un 1,9 y sería una emoción significativa en la palabra.

Una vez explicado cómo se van a invertir las palabras, ahora vamos a mezclarlo con la relación de antonimia, que usaremos de la siguiente manera. Cuando negamos una palabra donde una emoción con relación de antonimia es mayor que 1,5 (es significativa) procedemos a invertir la emoción que estamos tratando pero también a aumentar su antónima. Si la relación es total aumentaremos en la categoría antónima el 100 % de los grados que aparecían en la emoción original antes de la inversión. En cambio si la relación es parcial aumentaremos la categoría antónima el 50 % de los grados que aparecían en la emoción original antes de la inversión. Vamos a continuar viendo el ejemplo con la palabra “adicto”. Como vimos en la Tabla 6.14, las emociones de tristeza, miedo, enfado y asco superan el 1,5, pero de estas emociones sólo la tristeza y el enfado tienen relación de antonimia. Empezaremos con la relación de antonimia de la tristeza. Como hemos visto antes, el valor de la tristeza se invertirá para la palabra, pero como esta emoción tiene una relación de antonimia total con la alegría sumaremos al grado de la alegría el valor original que poseía la tristeza, y en este caso sumaríamos 2,6. Por tanto, la tristeza quedaría como 4,3 ( $2-0,3+2,6$ ). Seguimos con la otra relación de antonimia que teníamos, el enfado. Esta relación es parcial, significa que el 50 % del enfado será sumado a la emoción de la alegría después de haber sido invertida. En este caso, sin tener en cuenta la tristeza, tendríamos como grado final de enfado un 0,67 tras su inversión ( $2-1,33$ ) y la alegría quedaría como 2,365 ( $4,3+0,5*1,33$ ).

A continuación con la frase “No es adicto a las drogas” vamos a ver la aplicación de la negación en la frase. Partiendo de que la negación afecta a “adicto” y a “drogas”, y que ambas tienen los grados emocionales en el diccionario que podemos ver en la Tabla 6.14, vamos a calcular cómo afecta la negación.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
Adicto	2,6	1,97	0,3	1,33	1,23
Drogas	2,6	1,85	0,45	2,2	2,25

Tabla 6.14: Grados emocionales del diccionario para “adicto” y “drogas”

En primer lugar se invertirán las emociones de cada palabra. Ya hemos visto el método para la palabra “adicto” y vemos su resultado final en la Tabla 6.15. Para la palabra “drogas” seguiremos el mismo proceso. Primero calculamos la inversión de cada emoción teniendo en cuenta si la emoción es débil o fuerte:

- Tristeza tiene un 2,6. Por lo tanto es una emoción fuerte e invertimos usando:  $4 - 2,6 = 1,4$
- Miedo tiene un 1,85. Por lo tanto es una emoción débil e invertimos usando:  $2 - 1,85 = 0,15$
- Alegría tiene un 0,45. Por lo tanto es una emoción débil e invertimos usando:  $2 - 0,45 = 1,55$
- Enfado tiene un 2,2. Por lo tanto es una emoción fuerte e invertimos usando:  $4 - 2,2 = 1,8$
- Asco tiene un 2,25. Por lo tanto es una emoción fuerte e invertimos usando:  $4 - 2,25 = 1,75$

A continuación miraremos las relaciones de antonimia. Como las emociones tristeza, miedo, enfado y asco superaban el 1,5 debemos ver si alguna tiene relación de antonimia con otra y vemos que la tristeza tiene una relación con la alegría total y el enfado con la alegría parcial. Esto implica que la alegría sea modificada añadiendo a su nuevo valor ya invertido 1,55 el valor original de la tristeza 2,6 y el 50 % del valor original del enfado 50 % de 2,2, es decir 1,1. Quedando la alegría como  $1,55 + 2,6 + 1,1 = 5,25$ . En la Tabla 6.15 vemos los resultados de cada palabra invertida, la suma de las emociones, y el resultado final para la oración. Los resultados los consideramos bastante buenos dado que la emoción mayoritaria es la alegría y en menor medida aparecen el enfado y el asco, ya que hablar de adicciones y de drogas siempre nos producirá algo de enfado y asco.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
Adicto	0,4	1,03	4,965	1,67	1,77
Drogas	1,4	0,15	5,25	1,8	1,75
Suma de valores	1,8	1,18	10,715	3,47	3,52
NºPalabras	2				
Resultado final	0,9	0,59	5,36	1,74	1,76

Tabla 6.15: Grados emocionales de la frase “No es adicto a las drogas”

Creemos que con este método hemos mejorado el resultado devuelto por nuestro sistema para aquellas oraciones que contienen negaciones. No obstante, se evalúa la utilización de este método con los textos que tenemos preparados para este tipo de pruebas. Los resultados

se pueden ver en la Tabla 6.16 enfrentados a los resultados que teníamos antes de aplicar la negación. Vemos que hemos eliminado el caso en el que al aplicar la negación aparecían todas las emociones como significativas. Aún así, hay todavía pequeños problemas que pulir ya que algunos resultados como en el caso de los textos “Negación” y “Negaciones coordinadas”, donde aparecen como emociones significativas la tristeza y la alegría cuando hemos definido que son antónimas. Por otra parte, también vemos una pequeña mejora en los textos de “Aladdín” donde comienza a subir la tristeza por frases como “no estaba enamorada”, y el texto “Negación2” donde hemos conseguido invertir las emociones de mejor manera aunque no nos muestre ninguna emoción significativa.

## 6.4. Conclusiones

Para comenzar, vemos que al igual que los modificadores al añadir el tratamiento de la negación nominal los resultados mejoran en las frases que usan estos tipos de negación. Dado que aportan en ocasiones y no restan al sistema en ninguna situación decidimos que es una buena mejora para el sistema e incluimos a nuestro listado de modificadores nuestras negaciones nominales.

En cuanto al tratamiento de la negación sintáctica tras evaluar los métodos propuestos vemos las limitaciones de cada uno.

El método de anular la aportación emocional de las palabras negadas, vemos que mejora los resultados en algunos textos con respecto al EmoTraductor 1.0 porque reducimos las emociones que aportarían las palabras negadas, sin embargo, notamos que el método tiene una carencia en frases como “No me gusta la comida” o “Tu actitud no es buena” ya que diríamos que son frases sin emociones por haber negado las de las palabras emocionales cuando en realidad no solo hacemos que desaparezcan las emociones principales sino que tenemos que aumentar algunas otras emociones en las frases, en los ejemplos deberíamos aumentar el asco y el enfado y/o la tristeza respectivamente.

El método de la negación basada en antónimos de palabras, tiene como principal problema que muchas palabras antónimas semánticamente hablando no lo son emocionalmente, esto hace que los resultados tras buscar los antónimos y usar sus grados emocionales no varíen los grados que poseía la frase original sin tener en cuenta la negación.

Tanto el método de inversión de los grados emocionales de las palabras en el ámbito de la negación como el método de inversión de los grados emocionales significativos de las palabras en el ámbito de la negación, presentan como problema principal que podríamos confundir al usuario final con los resultados, ya que en algunos textos donde aparecen negaciones en el resultado final aparecen todas las emociones como significativas sin serlo en realidad. También el método de la negación de inversión de los grados emocionales significativos de las palabras en el ámbito de la negación en algún caso presenta como emociones finales del texto emociones antónimas, es decir para el mismo texto son emociones significativas dos emociones incompatibles como la alegría y la tristeza.

Por tanto, viendo las desventajas de cada método hemos decidido adoptar al EmoTraductor 2.0 solo el primer método visto que consiste en la anulación de los grados emocionales de las palabras que se encuentran en el ámbito de la negación. A pesar de que tenga un gran margen de mejora, presenta resultados mejores que sin considerar la negación. No obstante, dejaremos como trabajo futuro investigar y crear nuevos métodos para tratar la negación para categorías emocionales no antónimas entre sí.

Texto	Sin negación					Negación antonimia de cat. emocionales				
	T	M	Al	E	A	T	M	Al	E	A
Aladdin	0,53	0,81	2,54	0,31	0,22	1,08	0,75	2,19	0,76	0,49
Blancanieves	0,87	0,88	1,59	0,83	0,57	0,87	0,88	1,59	0,83	0,57
Blancanieves y sus amigos	0,37	0,46	2,82	0,23	0,18	0,37	0,46	2,82	0,23	0,18
Blog agradecimientos	0,5	0,44	2,18	0,42	0,27	0,97	0,57	1,88	0,78	0,5
Blog enfado	0,4	0,49	1,33	0,35	0,3	0,64	0,57	1,33	0,48	0,44
Blog moda	0,44	0,35	0,87	0,25	0,17	0,44	0,35	0,87	0,25	0,17
Blog muerte	0,84	0,62	1,33	0,56	0,28	0,84	0,63	1,31	0,56	0,29
Blog reflexion	0,71	0,71	1,62	0,58	0,37	1,01	0,8	1,52	0,8	0,56
Cenicienta	0,48	0,5	2,34	0,35	0,32	0,49	0,5	2,34	0,34	0,31
CentroComercial	0,56	0,56	1,25	0,47	0,36	0,61	0,61	1,26	0,51	0,4
Conversación1	0,19	0,22	0,99	0,21	0,16	0,19	0,22	0,99	0,21	0,16
Conversación2	0,42	0,37	1,4	0,34	0,28	0,45	0,37	1,3	0,39	0,31
DobleNegación1	0,16	0,27	2,37	0,3	0,09	0,57	0,48	3,07	0,37	0,47
DobleNegación2	0,61	0,66	2,65	0,4	0,29	2,51	1,04	1,52	1,95	1,28
DobleNegación3	1,62	1,48	0,71	2,36	3,79	2,52	1,89	0,68	3,06	4
El castillo de bestia	0,47	0,68	2,53	0,43	0,22	0,49	0,7	2,51	0,43	0,23
El ratoncito perez	0,43	0,64	1,5	0,43	0,34	0,54	0,65	1,5	0,47	0,39
FraseCorta	0,54	0,41	1,72	0,32	0,28	0,46	0,34	1,68	0,32	0,28
FraseCorta2	0,66	0,61	1,47	0,39	0,24	0,64	0,61	1,47	0,39	0,25
La bella durmiente	0,58	0,8	1,96	0,54	0,28	0,57	0,78	1,95	0,53	0,27
La bella y la bestia	0,64	0,88	2,22	0,4	0,23	0,65	0,93	2,33	0,41	0,21
La margarita	0,4	0,38	2,19	0,23	0,17	0,41	0,4	2,1	0,25	0,19
La sirenita	0,61	0,89	2,22	0,35	0,25	0,79	0,79	1,78	0,54	0,3
Mulan	0,68	0,76	1,86	0,72	0,46	0,65	0,73	1,63	0,69	0,46
Negación	0,6	0,6	1,97	0,32	0,23	1,55	0,84	1,87	1,19	0,91
Negación2	0,53	0,57	1,08	0,47	0,31	1,11	0,92	0,87	0,98	0,82
negaciones	0,69	0,8	1,5	0,37	0,34	2,44	1,5	1,18	2,06	1,67
Negaciones coordinadas	1,18	0,99	1,29	0,76	0,47	1,36	0,94	1,69	0,97	0,76
Noticia deportes	0,53	0,66	1,4	0,55	0,4	0,6	0,71	1,36	0,62	0,45
Noticia muerte	1,08	1,17	1,11	1,04	0,7	1,12	1,19	1,17	1,06	0,73
Noticia nacimiento	0,46	0,57	1,76	0,39	0,29	0,6	0,63	1,71	0,5	0,37
Noticia OMS	0,81	0,82	1,22	0,63	0,57	0,86	0,87	1,22	0,68	0,62
Noticia política	0,7	0,69	1,37	0,75	0,55	0,77	0,6	1,52	0,77	0,55
Peli aburrida	0,92	0,81	1,4	0,82	0,53	0,95	0,66	1,23	0,64	0,44
Pinocho	0,59	0,57	1,37	0,43	0,32	0,64	0,6	1,35	0,44	0,36
Pocahontas	0,75	0,84	1,95	0,41	0,27	0,75	0,84	1,95	0,41	0,27
Preguntas triste oración	0,32	0,42	0,96	0,28	0,14	0,58	0,61	0,91	0,48	0,39
Texto miedo perder amor	0,74	0,94	1,46	0,59	0,38	0,88	1	1,49	0,71	0,53
Texto1	0,8	0,86	1,22	0,78	0,52	1,09	0,99	1,2	0,95	0,63
Texto2	0,54	0,59	1,5	0,44	0,33	0,73	0,69	1,43	0,58	0,44
Tipos	0,16	0,31	2,05	0,19	0,09	0,45	0,43	2,03	0,4	0,22
Tristeza	0,84	0,67	1,01	0,49	0,35	1,07	0,72	0,96	0,67	0,51
Tristeza2	0,7	0,77	0,88	0,47	0,36	0,83	0,64	0,89	0,55	0,43

Tabla 6.16: Comparación de los resultados de los textos de prueba entre las primeras mejoras y al añadir el tercer método de la negación, usando la división de tipos de emociones en fuertes y débiles y las relaciones de antonimia. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).

## EmoTraductor 2.0

**RESUMEN:** En este capítulo vamos a ver el resultado de todas las mejoras planteadas y cómo dan lugar al EmoTraductor 2.0. Veremos en la sección 7.1 las mejoras que finalmente dejamos añadidas al sistema y cómo funciona esta nueva versión del EmoTraductor 2.0 una vez que se han añadido todas las modificaciones. En la sección 7.2 explicaremos algunas pequeñas modificaciones que hemos realizado en la aplicación web y en la sección 7.3 se explicarán los cambios hechos en los distintos servicios web sobre los que trabaja el EmoTraductor. Finalmente, en la sección 7.4 se realizará una evaluación final comparando los resultados del EmoTraductor 1.0 con los del EmoTraductor 2.0.

### 7.1. Mejoras añadidas al sistema

Tras todo lo argumentado, explicado y evaluado en los capítulos 4, 5 y 6, se han dejado implementadas en el sistema las mejoras:

- Normalización de los grados emocionales. Se ha cambiado el rango de grados emocionales de las palabras. El EmoTraductor 1.0 tenía grados del 1-5 y el EmoTraductor 2.0 los tiene del 0-4 (ver Sección 4.2).
- Ampliación del diccionario emocional. Se han añadido palabras nuevas al diccionario afectivo que usa la aplicación (ver 4.3).
- Cambio en el criterio de selección de palabras emocionales. Se sustituye el valor límite establecido (1,5) que debe poseer al menos una emoción para ser considerada emocional (ver Sección 4.4).
- Corrección en el servicio de grados emocionales de un texto. Se modifica el funcionamiento de servicio web para solucionar algunos problemas que había con la ponderación de los tipos de frases (ver Sección 4.5).
- Mejora en el servicio de obtención de grados emocionales de una palabra. Se modifica el funcionamiento del servicio para mejorar las búsquedas de palabras en el diccionario emocional (ver Sección 4.6).
- Incorporación del tratamiento de los modificadores. Se comienzan a tener en cuenta los modificadores aumentativos (ver Capítulo 5).

- Incorporación del tratamiento de la negación. Se divide la negación en negación nominal y sintáctica. Para ambos tipos se añade su tratamiento (ver Capítulo 6). Para la negación nominal se sigue el método explicado en la Sección 6.2, que consiste en utilizar estas partículas de negación como modificadores. Para la negación sintáctica se sigue el método explicado en la Sección 6.3.1 que consiste en eliminar los grados emocionales de las palabras emocionales negadas.

A continuación, vamos a ver el funcionamiento final del EmoTraductor 2.0 paso a paso usando como texto de ejemplo: “No estoy enfadada. ¡Estoy muy alegre!”. Tras cada paso explicaremos al final qué mejoras incorpora, en caso de tenerlas. Al final, compararemos los resultados con lo que obtendríamos con el EmoTraductor 1.0.

1. Separación del texto en frases. Utilizando los signos de puntuación dividiremos el texto en frases. En nuestro caso, el texto se dividiría en dos frases: “No estoy enfadada” y “¡Estoy muy alegre!”. Junto con esta división también se analiza el tipo de frase (enunciativa, exclamativa o interrogativa), ya que dependiendo del tipo tendrán más o menos peso en el resultado final: las enunciativas tendrán el valor resultado obtenido por el servicio de grados emocionales de una frase, las exclamativas tendrán el doble del valor que se obtenga del servicio y las interrogativas la mitad. En nuestro ejemplo vemos que la frase “No estoy enfadada” sería enunciativa, mientras que la frase “¡Estoy muy alegre!” sería exclamativa.
2. División de las frases en palabras. Se enviará cada frase al servicio de grados emocionales de una frase, donde se utilizará el analizador sintáctico Spacy. Spacy evalúa la frase para detectar la categoría gramatical de cada palabra y generar un árbol de dependencias para la frase. Una vez están etiquetadas las palabras se elegirán las palabras candidatas a ser emocionales. Las palabras candidatas a ser emocionales son aquellas que tienen como categoría gramatical verbo, adjetivo o sustantivo. Las palabras candidatas serán almacenadas en un listado de tuplas donde cada tupla tendrá la palabra, la posición que ocupa en la frase y la categoría gramatical que posee. En nuestro ejemplo, en este punto obtendríamos:
  - Los árboles de dependencias que podemos ver en las Figuras 7.1 y 7.2.
  - Las categorías gramaticales de cada palabra y si son candidatas a ser emocionales o no junto con su posición en la frase (se almacenará en un listado de tuplas). Podemos ver los resultados para el texto ejemplo en las Tablas 7.1 y 7.2.

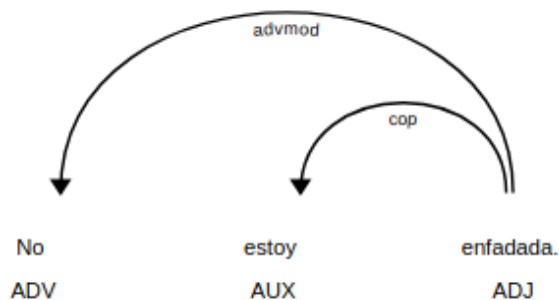


Figura 7.1: Árbol de dependencias creado por Spacy de la frase “No estoy enfadada”

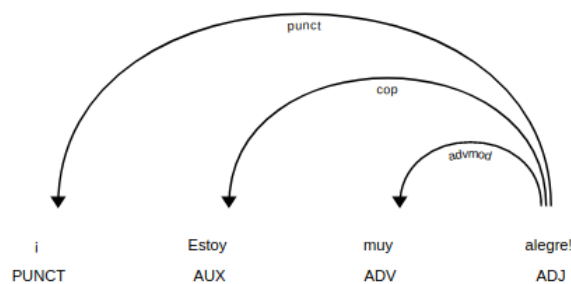


Figura 7.2: Árbol de dependencias creado por Spacy de la frase “¡Estoy muy alegre!”

Palabra	Posición	Categoría gramatical	Candidata
No	0	Adverbio (ADV)	NO
estoy	1	Verbo auxiliar (AUX)	NO
enfadada	2	Adjetivo (ADJ)	SI

Tabla 7.1: Palabras de la frase “No estoy enfadada” con su posición, categoría gramatical y si son candidatas a ser emocionales o no.

En este punto se han añadido la generación de árboles de dependencias para las frases y se ha cambiado el método de almacenamiento de las palabras candidatas a ser emocionales creando el listado de tuplas.

3. Detección de la negación sintáctica. Para cada palabra candidata a ser emocional se recorrerán los nodos hijos de la palabra en el árbol de dependencias de la frase con el fin de conocer si existe la partícula “no” de negación en su ámbito. En caso de existir, en un nuevo listado de tuplas, se almacenarán todas las palabras negadas. En este listado se almacenará cada una de las palabra negadas junto con la posición que ocupan en la frase. En caso de no existir el proceso continuaría. En nuestro ejemplo, únicamente la frase “No estoy enfadada” tiene la partícula “no” de negación, y en este caso el listado de palabras afectadas por la negación sería:

```
[[ 'enfadada', 2 ]]
```

Este punto es totalmente nuevo ya que es la inclusión del tratamiento de la negación sintáctica.

4. Detección de los modificadores y de la negación nominal. Una vez más, para cada palabra candidata a ser emocional se recorrerán los nodos hijos de la palabra en el árbol de dependencias de la frase con el fin de conocer si algunos de los nodos

Palabra	Posición	Categoría gramatical	Candidata
Estoy	0	Verbo (VERB)	SI
muy	1	Adverbio (ADV)	NO
alegre	2	Adjetivo (ADJ)	SI

Tabla 7.2: Palabras de la frase “¡Estoy muy alegre!” con su posición, categoría gramatical y si son candidatas a ser emocionales o no.

corresponden con alguna de las palabras que podemos encontrar almacenadas en el sistema como modificadores o negaciones nominales. En caso de que exista alguna palabra afectada por uno o varios modificadores o negaciones o una combinación de ambas, se sumarán los porcentajes asociados a cada modificador o negación hallada para calcular la modificación que sufrirán las palabras en el ámbito de los modificadores o negadores en su carga emocional. Después de calcular el porcentaje total, se almacenará en una nueva lista de tuplas cada palabra con su posición en la frase y el porcentaje de variación que se aplicará sobre la carga afectiva de las palabras en el ámbito de actuación del negador o modificador. En caso de que las palabras no tengan en sus ámbitos ningún modificador, las palabras también serán añadidas al listado poniendo un 0 en el valor que debe ocupar el porcentaje en la tupla. En nuestro ejemplo solo la frase “¡Estoy muy alegre!” tiene un modificador, quedando el listado de tuplas de la siguiente manera:

```
[[ 'estoy', 0, 0], [ 'alegre', 2, 75]]
```

Este punto es totalmente nuevo ya que es la inclusión del tratamiento de la negación nominal y los modificadores.

5. Análisis morfológico de cada palabra. Todas las palabras candidatas se procesarán usando Spacy y PyStemmer. Estas dos herramientas obtendrán toda la información posible de la palabra: categoría morfológica, género, número, lema con Spacy y lema con PyStemmer. Este procesamiento será realizado analizando solo las palabras que contiene cada frase y no incluirán los datos que hemos obtenido para ellas en los apartados anteriores como la función que cumplen en la frase o el tipo de palabra que eran. Esto será muy importante porque si la categorización que se hizo de la palabra en la frase fue incorrecta, al evaluar la palabra por separado podemos hallar la categoría acertada y así después localizarla mejor. En nuestro ejemplo podemos ver el análisis de cada palabra candidata en la Tabla 7.3. La palabra “contenta”, Spacy la considera un verbo (tercera persona del singular del presente simple del verbo “contentar”).

Palabra	Categoría morfológica	Género	Número	Lema PyStemmer	Lema Spacy
enfadada	Adjetivo (ADJ)	Femenino	Singular	enfad	enfadar
estoy	Verbo auxiliar (AUX)	X	Singular	estoy	estar
alegre	Adjetivo (ADJ)	X	Singular	alegr	alegrar

Tabla 7.3: Palabras del texto introducido y su respuesta ante la primera evaluación que se les realiza con Spacy y PyStemmer

6. Búsqueda de la palabra en el diccionario. Primero se busca en el nuevo diccionario emocional (ver Sección 4.3) la palabra tal y como está escrita en la frase. En caso de encontrar los grados emocionales de la palabra se devuelven y pasaríamos al siguiente punto. Si por el contrario, no se encontró la palabra en el diccionario, si la palabra es plural y acaba en -es o -s quitaremos la partícula del final y volveremos a buscar la palabra en el diccionario. Si la palabra es encontrada en el diccionario finalizaríamos este paso y continuaríamos con el siguiente. Si la palabra era plural acaba en -es o -s y no se encuentra en el diccionario o era una palabra en singular,

usaremos los lemas obtenidos en el paso 5. En caso de que los lemas obtenidos por Spacy y PyStemmer para la misma palabra sean distintos, los subpasos descritos a continuación, se realizarán tanto con el lema obtenido por PyStemmer como con el lema obtenido por Spacy. Primero se realizarán con el lema de PyStemmer y después con el lema de Spacy. En caso de que los lemas obtenidos sean iguales solo usaremos este lema para realizar los siguiente subpasos. Realizaremos la búsqueda en el diccionario usando la información del lema y la categoría gramatical de la palabra. Se buscará en el diccionario usando el par (lema, categoría gramatical). Si encontramos que hay dos palabras en el diccionario con el mismo lema y la misma categoría gramatical añadiremos a los parámetros de búsqueda el género, es decir, usaríamos la tupla (lema, categoría gramatical, género) para buscar en el diccionario. Si tras alguna de estas dos búsquedas se ha encontrado la palabra en el diccionario se continuará el proceso en el siguiente punto. En cambio, si después de estas dos nuevas búsquedas sigue sin encontrarse las emociones de la palabra en el diccionario se continuará buscando el lema como palabra. Es decir, probaremos a usar el lema como palabra para asegurarnos que si la palabra es derivada de otra y se ha obtenido un lema incorrecto, podamos encontrarla. Si en este punto no se encuentra las emociones asociadas a la palabra se devolverá que la palabra no es emocional y se continuará el proceso.

Este proceso ha sido modificado completamente, ya que en el EmoTraductor 1.0 únicamente se buscaba en el diccionario por el lema obtenido por PyStemmer. No se buscaba ni la palabra original ni se utilizando más parámetros de búsqueda (categoría gramatical, género y número).

7. Decisión de palabras emocionales. Si para la palabra que se está evaluando se han obtenido sus grados emocionales diremos que la palabra es emocional. En la Tabla 7.4 vemos nuestras candidatas del texto ejemplo con sus emociones asociadas en caso de tenerlas y por tanto si son consideradas emocionales o no.

Palabra	¿Es emocional?	Tristeza	Miedo	Alegría	Enfado	Asco
enfadada	Sí	1,75	1,4	0,2	2,35	0,68
estoy	No	0	0	0	0	0
alegre	Sí	0,1	0,23	3,83	0,1	0,03

Tabla 7.4: Palabras candidatas a ser emocionales en nuestro texto “No estoy enfadada. ¡Estoy muy alegre!” junto con si son palabras emocionales y sus grados emocionales.

Guardaremos las palabras emocionales en un listado de tuplas de la forma (palabra emocional, posición que ocupa en la frase, categoría gramatical). En caso de que la palabra no sea emocional, se descartará la palabra y se continuará el proceso. Los grados de las palabras emocionales también se almacenarán para su procesado. El listado de palabras emociones para cada una de nuestras frases, junto con los listados de modificadores y negaciones quedarían de la siguiente forma .

- “No estoy enfadada”:

Listado de palabras emocionales:

```
[[ 'enfadada', 2, 'ADJ' ]]
```

Emociones:

```
[[ 1,75 & 1,4 & 0,2 & 2,35 & 0,68 ]]
```

```

Listado de negación:
    [['enfadada', 2]]
Listado de modificadores:
    [['enfadada', 0]]
■ “¡Estoy muy contenta!”:
Listado de palabras emocionales:
    [['alegre', 2, 'ADJ']]
Emociones:
    [[0,1 & 0,23 & 3,83 & 0,1 & 0,03]]
Listado de negación:
    []
Listado de modificadores:
    [['estoy', 0, 0], ['alegre', 2, 75]]

```

En esta fase hemos incluido como novedad el guardado de las palabras emocionales en un listado de tuplas y se ha modificado que palabras se consideran emocionales con la mejora vista en la Sección 4.4

8. Procesamiento de la negación sintáctica. En este paso tenemos el listado de tuplas donde aparecen las palabras negadas (ver paso 3), las emociones asociadas a la palabra y el listado de palabras emocionales (ver paso 7). Vamos a recorrer las palabras emocionales junto con su posición, comprobando si en el listado de negaciones aparece la palabra con la misma posición en la frase. Si es así, modificaremos los grados que tenía asociados la palabra y pondremos todos a cero para no tener en cuenta estas emociones para el cálculo de la carga afectiva de la frase. Si, por el contrario, la palabra no aparece o la palabra aparece pero con otra posición en el listado de negaciones, no se modificarán sus grados emocionales y continuaremos al siguiente punto. En nuestro ejemplo, se cambiarán las emociones asociadas a la palabra “enfadada” poniendo en todos los grados emocionales un 0 como podemos ver en la Tabla 7.5.

Este punto es totalmente nuevo ya que es el cálculo de los grados emocionales teniendo en cuenta la negación sintáctica.

9. Procesamiento de los modificadores y la negación nominal. Se recorrerá de nuevo el listado de palabras emocionales junto con sus posiciones y haremos lo mismo que hemos hecho con la negación. Si la palabra con su posición aparece en el listado de modificadores aplicaremos a los grados emocionales de las palabras el porcentaje que tenían asociado en el listado de tuplas para aquellos cuyo porcentaje sea distinto de 0. Los que tienen en el listado de modificadores un 0 en la posición de la tupla que indicaba el porcentaje, se ignorarán. En nuestro ejemplo, se cambiarán las emociones asociadas a la palabra “alegre” sumando un 75% más a sus emociones tal y como podemos ver en la Tabla 7.5

Este punto es totalmente nuevo ya que es el cálculo de los grados emocionales teniendo en cuenta la negación nominal y los modificadores.

10. Cálculo de los grados emocionales de una frase. Con los datos que ya tenemos para cada palabra de cada frase ahora sumaremos el número de palabras emocionales y sumaremos los grados emocionales de cada una estas por cada categoría. Después, se dividirá cada valor de cada categoría emocional entre el número de palabras emocionales que poseía la frase. En nuestro ejemplo vemos en la Tabla 7.6 el resultado de

Palabra	Qué aplica	Tristeza	Miedo	Alegría	Enfado	Asco
enfadada	Negación sintáctica	0	0	0	0	0
alegre	Modificador +75 %	0,18	0,4	6,7	0,18	0,05

Tabla 7.5: Palabras emocionales en nuestro texto “No estoy enfadada. ¡Estoy muy alegre!” con los efectos de la negación y los modificadores aplicados

la frase “No estoy enfadada” y en la Tabla 7.7 el resultado de la frase “¡Estoy muy alegre!”. Al haber buscado ejemplos sencillos para ver todo el proceso este paso se nos queda reducido a uno muy sencillo.

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
enfadada	0	0	0	0	0
Suma de valores	0	0	0	0	0
NºPalabras	1				
Resultado final	0	0	0	0	0

Tabla 7.6: Grados emocionales de la frase “No estoy enfadada” con el EmoTraductor 2.0

Palabra	Tristeza	Miedo	Alegría	Enfado	Asco
alegre	0,18	0,4	6,7	0,18	0,05
Suma de valores	0,18	0,4	6,7	0,18	0,05
NºPalabras	1				
Resultado final	0,18	0,4	6,7	0,18	0,05

Tabla 7.7: Grados emocionales de la frase “¡Estoy muy alegre!” con el EmoTraductor 2.0

11. Cálculo de los grados emocionales finales del texto. Basándonos en los resultados obtenidos en el paso anterior para cada frase, y utilizando la ponderación explicada en el primer paso por el tipo de frase que estamos tratando, se calculan los grados emocionales del texto completo. Primero se ponderan los grados devueltos para cada frase, después se suman los grados de cada categoría de todas las frases y se divide entre el número de frases, obteniendo así el resultado final. En nuestro ejemplo, debemos tomar el valor de la frase “No estoy enfadada” y dejarlo como está ya que es una frase enunciativa, el valor de la frase “¡Estoy muy alegre! debemos multiplicarlo por 2 por ser una frase exclamativa. Después se sumarán los grados emocionales de las frases y se dividirá entre el número de frases (2). Podemos ver el método completo en la Tabla 7.8.

En este punto se ha cambiado el divisor del cálculo para obtener los resultados finales ya que en la versión anterior el divisor se calculaba teniendo en cuenta el peso de las frases y no el número total de frases (ver Sección 4.5).

Frase	Multiplicador	Tristeza	Miedo	Alegría	Enfado	Asco
No estoy enfadada	x1	0	0	0	0	0
¡Estoy muy alegre!	x2	0,36	0,8	13,4	0,36	0,1
Suma de valores		0,36	0,8	13,4	0,36	0,1
Nº Frases		2				
Resultado final		0,18	0,4	6,7	0,18	0,05

Tabla 7.8: Grados emocionales del texto “No estoy enfadada. ¡Estoy muy alegre!” con el EmoTraductor 2.0

## 7.2. Cambios en la interfaz de la aplicación web

Respecto a los cambios realizados en la interfaz, lo primero que se hizo fue un cambio simplemente estético cambiando el nombre de la aplicación. En la Figura 7.3 vemos cómo aparecía el logo en la versión 1.0 y en la Figura 7.4 vemos como aparece el nombre usando un nuevo logo para la aplicación.



Figura 7.3: Interfaz de la aplicación web EmoTraductor 1.0



Figura 7.4: Interfaz de la aplicación web EmoTraductor 2.0

Por otra parte, se ha tenido que adaptar el límite por el cual una emoción se consideraba significativa en la interfaz para que solo las emociones significativas aparezcan en la barra emocional. En el EmoTraductor 1.0 este valor estaba fijado a 2,5 y en el EmoTraductor 2.0 se ha modificado y bajado a 1,5. En la nueva versión, todas las emociones que tengan un grado mayor al 1,5 aparecerán en la barra emocional.

Como último cambio, se ha tenido que modificar la manera en la que la aplicación obtiene el dato de los JSON que le llega de los servicios cuyas respuestas se han modificado (ver Sección 7.3).

No han sido necesarios realizar más cambios gracias a la arquitectura basada en servicios de la aplicación que hace que la interfaz quede completamente separada de la implementación.

### 7.3. Cambios en los servicios web

A continuación, haremos un repaso de los servicios que hemos modificado en este trabajo para acometer todas las mejoras:

- Servicio de grados emocionales de una palabra. Tal y como se ha contado en la Sección 4.6, se ha modificado la manera de buscar una palabra en el diccionario. Pasando de buscar únicamente el lema de la palabra original a la búsqueda de la palabra en singular si fuera plural, la búsqueda en el diccionario por lema y género, la búsqueda en el diccionario por lema, género y número y la búsqueda en el diccionario usando el lema como palabra. Este cambio ha supuesto un aumento en el número de palabras candidatas a ser emocionales que finalmente pasaban a serlo, lo que se traduce en un aumento de la cobertura de la traducción.
- Servicio de grados emocionales de una frase. Este es el servicio que más cambios ha sufrido. Primero, se ha añadido el tratamiento de los modificadores. Después, se ha añadido el tratamiento de la negación tanto nominal como sintáctica. Por último, este servicio cambia también el resultado final que devuelve. En el EmoTraductor 1.0 devolvía un listado de palabras emocionales y las emociones de la frase. El EmoTraductor 2.0 añade para cada palabra emocional devuelta su posición en la frase y la categoría gramatical de la palabra.

### 7.4. Evaluación del marcado emocional de texto del EmoTraductor 2.0

Después de haber realizado todos los cambios ya comentados, vamos a realizar una evaluación en cuanto al marcado emocional de texto para comparar los resultados que obteníamos usando el EmoTraductor 1.0 con los que obtenemos usando el EmoTraductor 2.0. La comparativa de resultados la podemos encontrar en la Tabla 7.9. Hay que tener en cuenta que los resultados del EmoTraductor 1.0 usan un diccionario con grados emocionales entre el 1 y el 5 y el EmoTraductor 2.0 utiliza uno con grados del 0 al 4. También debemos considerar que en el EmoTraductor 1.0 las emociones significativas son las que superan el 2,5 mientras que en el EmoTraductor 2.0 son las que superan el 1,5.

Aún con estas diferencias nos damos cuenta que existe una mejora en muchos de los textos. Para empezar en aquellos que no se podían evaluar usando el EmoTraductor 1.0: “CentroComercial”, “Noticia deportes” y “Texto1”. Estos no se podían evaluar por la repetición de lemas en el diccionario emocional y por cómo se obtenían los grados emocionales de las palabras a través del servicio web. También se aprecian mejoras en textos como “Blog muerte” donde el EmoTraductor 1.0 devolvía que la emoción mayoritaria y más significativa era la alegría en un texto claramente triste. Ahora el EmoTraductor 2.0 no indica tampoco la tristeza pero es una gran mejora que este texto no se considere alegre sino carente de emociones.

Lo mismo ocurre con los textos seleccionados para evaluar como se comporta el sistema cuando los textos contienen negaciones, como por ejemplo “Negación2”. En dicho texto el EmoTraductor 1.0 no devolvía emociones significativas pero sí dos muy cerca de serlo (el

miedo con un 2,46 y la alegría con un 2,39) mientras que la nueva versión ha reducido las emociones por la negación y ninguna supera el 0,5. Otro ejemplo es el texto “Negacion”, donde la emoción más significativa era alegría, con un 3,33, en el EmoTraductor 1.0 y en la nueva versión la alegría ha desaparecido convirtiendo el texto en algo, que aunque tenga como mayoritaria la alegría con un 1,15, aún le queda para llegar a ser significativo.

También notamos pequeñas mejoras en textos como “Aladdín” donde el cuento, a pesar de seguir siendo alegre con un valor de 2,09, reduce su alegría respecto al EmoTraductor 1.0 donde tenía 3,79.

Por otra parte, fijándonos en el texto “Tipos” vemos que hemos conseguido una emoción significativa, la alegría, y esto se debe a que el EmoTraductor 2.0 ha comenzado a tratar bien las ponderaciones de las frases enunciativas, exclamativas e interrogativas.

También, como cambio genérico, vemos que, por lo general, todas las emociones se suavizan y esto se debe al aumento de entradas en el diccionario emocional. Cada frase ahora tiene más palabras que evaluar y, por tanto, aumenta el divisor a la hora de calcular la media de los grados por frase. Consideramos que esto es un cambio positivo a pesar de que los grados estén suavizados ya que ahora cada grado aportado tendrá más rigor por haber sido resultado de un mayor número de palabras.

Texto	EmoTraductor 1.0					EmoTraductor 2.0				
	T	M	Al	E	A	T	M	Al	E	A
Aladdin	1,44	1,67	3,79	1,26	1,14	0,43	0,64	2,09	0,28	0,21
Blancanieves	1,9	2,02	3,02	1,81	1,51	0,87	0,88	1,59	0,83	0,57
Blancanieves y sus amigos	1,29	1,4	4,18	1,17	1,1	0,37	0,46	2,82	0,23	0,18
Blog agradecimientos	1,29	1,2	2,77	1,04	0,95	0,49	0,39	2,24	0,45	0,26
Blog enfado	0,7	0,95	1,39	0,67	0,6	0,34	0,41	1,15	0,25	0,23
Blog moda	1,1	0,93	2,05	0,96	0,9	0,45	0,35	0,89	0,26	0,17
Blog muerte	2,24	1,91	2,76	1,62	1,28	0,87	0,64	1,32	0,58	0,3
Blog reflexion	2	1,99	3,12	1,8	1,56	0,66	0,65	1,47	0,54	0,35
Cenicienta	1,43	1,53	4,05	1,27	1,16	0,51	0,53	2,46	0,35	0,32
CentroComercial	X	X	X	X	X	0,56	0,57	1,24	0,47	0,35
Conversación1	0,17	0,18	0,63	0,18	0,17	0,19	0,22	0,99	0,21	0,16
Conversación2	0,7	0,68	1,93	0,62	0,56	0,49	0,39	1,31	0,4	0,32
DobleNegación1	1,97	1,54	2,62	1,7	1,31	0,11	0,18	1,58	0,2	0,06
DobleNegación2	1,75	1,77	4,14	1,34	1,26	0,09	0,18	1,46	0,07	0,05
DobleNegación3	1,2	1,2	3,23	1,12	1,4	1,51	1,35	0,2	2,27	3,4
El castillo de bestia	1,41	1,53	3,95	1,26	1,23	0,49	0,7	2,54	0,44	0,24
El ratoncito perez	1,31	1,47	2,25	1,16	1,06	0,41	0,58	1,44	0,38	0,32
FraseCorta	1,74	1,59	3,72	1,28	1,13	0,46	0,34	1,68	0,32	0,28
FraseCorta2	2,31	2,13	2,4	1,6	1,31	0,64	0,61	1,47	0,39	0,25
La bella durmiente	2,06	2,21	2,79	1,96	1,44	0,58	0,82	2,02	0,55	0,28
La bella y la bestia	1,62	2	3,55	1,49	1,24	0,67	0,96	2,41	0,42	0,21
La margarita	1,45	1,48	3,78	1,17	1,13	0,41	0,4	2,1	0,25	0,19
La sirenita	1,89	2,17	3,92	1,34	1,28	0,52	0,72	1,76	0,35	0,22
Mulan	1,34	1,25	2,07	1,18	0,92	0,67	0,74	1,7	0,71	0,46
Negación	2	2,06	3,33	1,37	1,24	0,18	0,2	1,15	0,1	0,1
Negación2	2,21	2,46	2,39	2,04	1,66	0,35	0,27	0,46	0,24	0,08
negaciones	3,02	2,19	2,31	2,09	1,5	0,34	0,32	0,17	0,18	0,17
Negaciones coordinadas	2,29	2,19	1,45	1,56	1,19	0,62	0,53	0,88	0,46	0,28
Noticia deportes	X	X	X	X	X	0,54	0,68	1,37	0,56	0,41
Noticia muerte	2,23	2,36	1,79	2,14	1,67	1,04	1,14	1,07	0,99	0,69
Noticia nacimiento	1,71	2,18	3,32	1,61	1,38	0,45	0,56	1,72	0,38	0,28
Noticia OMS	1,99	2,12	2,57	1,63	1,54	0,8	0,82	1,21	0,62	0,56
Noticia política	1,39	1,32	2,07	1,3	1,08	0,61	0,56	1,36	0,7	0,5
Peli aburrida	1,48	1,16	1,62	1,22	1,12	1,17	0,7	1,45	0,71	0,5
Pinocho	1,37	1,53	2,61	1,14	1,02	0,6	0,57	1,37	0,41	0,31
Pocahontas	1,55	1,54	2,07	1,16	0,99	0,75	0,84	1,95	0,41	0,27
Preguntas triste oración	1,26	1,5	1,83	1,22	0,92	0,29	0,36	0,78	0,22	0,11
Texto miedo perder amor	1,23	1,38	1,25	1,03	0,82	0,65	0,88	1,32	0,52	0,35
Texto1	X	X	X	X	X	0,79	0,85	1,07	0,75	0,5
Texto2	1,42	1,56	2,13	1,18	1,13	0,55	0,59	1,41	0,44	0,33
Tipos	0,47	0,6	1,33	0,43	0,41	0,16	0,3	1,92	0,19	0,08
Tristeza	1,58	1,6	2,77	1,27	1,08	0,68	0,52	0,74	0,36	0,26
Tristeza2	1,29	1,44	1,5	1,02	0,84	0,59	0,59	0,71	0,4	0,32

Tabla 7.9: Comparación de los resultados de los textos de prueba entre lo que devolvía el EmoTraductor 1.0 y el EmoTraductor 2.0. Las columnas corresponden a las siguientes emociones: T (tristeza), M (miedo), Al (alegría), E (enfado), A (asco).



## Conclusiones y Trabajo Futuro

En este capítulo veremos, en la Sección 8.1 las conclusiones de este trabajo valorando si los objetivos planteados en la Sección 1.2 se han cumplido y en caso de ser así de qué manera. En la Sección 8.2 marcaremos unos puntos que pueden servir como base para continuar mejorando la herramienta expuesta en este trabajo.

### 8.1. Conclusiones

Después de finalizar el proyecto y explicar todos los detalles de diseño, implementación y resultados que se han seguido en este documento, es posible extraer las siguientes conclusiones según los objetivos que se marcaron al principio.

Nuestro principal objetivo en este trabajo era mejorar el proceso de marcado emocional de textos que realizaba la herramienta EmoTraductor 1.0. Esta herramienta es una aplicación web que tiene como objetivo ayudar a las personas que tienen dificultades a la hora de interpretar el contenido emocional de un texto mediante el marcado emocional de textos usando categorías emocionales con un grado de intensidad asociado.

Para poder cumplir con este objetivo principal, propusimos pasar por los siguientes objetivos:

- Solucionar los malos funcionamientos encontrados tras la evaluación del EmoTraductor 1.0.
- Aumentar el diccionario emocional para ampliar la cobertura del EmoTraductor.
- Añadir el tratamiento de los modificadores que pueden alterar la carga emocional de las palabras a las que afectan.
- Añadir el tratamiento de la negación en los textos.
- Evaluar todas las mejoras introducidas para decidir cuáles de ellas son finalmente incorporadas a la nueva versión de nuestra herramienta.

Para solucionar los malos funcionamientos encontrados tras la evaluación del EmoTraductor 1.0 y aumentar el diccionario emocional, utilizamos como base de mejora parte del trabajo futuro propuesto en la memoria del TFG donde se implementó el EmoTraductor 1.0. Junto con la memoria del TFG se evalúa el funcionamiento de cada uno de los servicios web que utiliza la herramienta para encontrar malos funcionamientos que no aparecían como trabajo futuro en el TFG.

Con los cambios realizados durante este TFM se ha logrado los siguiente:

1. Mejorar el marcado emocional gracias a la normalización de los grados emocionales del diccionario afectivo. Como vimos en la Sección 4.2, al utilizar el diccionario afectivo inicial donde los valores de las emociones estaban entre el 1 y el 5 (donde el 1 era la ausencia de emoción y el 5 la aparición total de la emoción) cuando ponderábamos las frases según su tipología el tener el valor 1 como ausencia de emoción hacía que los grados emocionales de las frases ponderadas aumentasen emociones que no deberían aparecer. Al normalizar el diccionario cambiando el rango de los grados emocionales para que oscilen entre el 0 y el 4 (donde el 0 es la ausencia de emoción y el 4 la aparición total de esta) evitamos que los textos que contengan distintos tipos de frases con distintas ponderaciones asociadas a cada tipo de frase, incorporen grados inexistentes al texto.
2. Realizar una ampliación del diccionario afectivo. Esta mejora se encuentra explicada en la Sección 4.3 y consiste en añadir más entradas al diccionario afectivo sobre el que se basa la aplicación. Tras este cambio se mejora el marcado de textos ya que con el nuevo diccionario se pueden tener en cuenta muchas más palabras a la hora de determinar el contenido afectivo de los textos haciendo que las emociones resultantes sean más correctas. Junto a esta ampliación de entradas en el diccionario emocional también se realiza una ampliación de información (género, número, categoría gramatical) de cada palabra que se encuentra en él con el fin de realizar búsquedas más correctas.
3. Cambiar el criterio de selección de palabras emocionales. Esta mejora, que encontramos explicada en la Sección 4.4, consiste en la eliminación del límite (2,5) que debía tener alguna categoría emocional de una palabra para que el EmoTraductor 1.0 la considerase emocional y colaborase al cálculo final de las emociones de una frase. Esta mejora acompaña a la anterior dado que, al eliminar el límite hace que se tengan en cuenta todas las palabras que tienen entrada en el diccionario a la hora de marcar emocionalmente un texto y no solo aquellas que tenían para una categoría emocional un grado mayor al establecido.
4. Mejorar el servicio web que obtiene los grados emocionales de una palabra. Gracias a los cambios en este servicio, explicados en la Sección 4.6, que consisten en modificar la manera que tenía el EmoTraductor 1.0 para buscar una palabra en el diccionario, añadiendo la búsqueda por la palabra original o la búsqueda por lema de la palabra, género y número. Con este cambio en la forma de buscar nos aseguramos de que encontramos el mayor número posible de palabras en el diccionario emocional y por tanto aseguramos que los grados emocionales finales de un texto tienen mayor rigor por considerar más palabras.
5. Mejorar el servicio web que obtiene los grados emocionales de un texto. Tal y como se explica en la Sección 4.5, en el EmoTraductor 1.0 existía un problema a la hora de calcular el divisor que se utiliza para la obtención de los grados emocionales finales de un texto, ya que contaba cada frase como su ponderación en vez de como una única frase, es decir si aparecía una frase exclamativa se contaba como dos frases en lugar de una, haciendo que los resultados no sean correctos. Se ha modificado por tanto, la manera de contar las frases que componen un texto.

Además, en este TFM, tal y como se cuenta en el Capítulo 5, hemos introducido el tratamiento de los modificadores en el marcado emocional de textos. Recordemos que lla-

amos modificador a una palabra que modifica la carga afectiva de las palabras a las que acompaña. Para ello hemos incorporado al sistema un listado de posibles modificadores, donde cada modificador que se encuentra en este listado tiene asociado un porcentaje de actuación. Este porcentaje es la modificación que realizará sobre la carga afectiva de las palabras a las que afecta el modificador en cuestión. De manera que si tenemos un modificador como “muy” con un porcentaje de un +75 % la palabra o palabras que sean afectadas por el término “muy” incrementarán sus emociones en un 75 %. Para conseguir determinar el ámbito de un modificador, es decir, para determinar qué palabras se ven afectadas por un modificador, se ha añadido al sistema el tratamiento de árboles de dependencias. Tras introducir esta mejora en el sistema se realizó una nueva evaluación considerando que detectamos y tratamos los modificadores de manera adecuada y que esta mejora mejora los resultados. Por ello damos por cumplido el objetivo de tratar los modificadores en el marcado emocional de textos.

A continuación y como se cuenta en el Capítulo 6, se incorpora al sistema el tratamiento de la negación. Lo primero que realizamos es una clasificación de las negaciones en dos tipos: negaciones nominales y negaciones sintácticas. Las negaciones nominales son las que se identifican por usar palabras como “casi”, “poco” o “menos” y que reducen la carga afectiva de las palabras a las que afecta. Este tipo de negación es tratado por el sistema como modificadores que decrementan la carga emocional de las palabras a las que afecta. Las negaciones sintácticas son las que utilizan la palabra “no” y suelen afectar al verbo de la oración negando así el significado completo de la frase. En este tipo de negaciones solo consideramos aquellas donde aparece la partícula “no”. Para resolver el tratamiento de este tipo de negaciones planteamos cinco aproximaciones diferentes:

- No considerar palabras emocionales a aquellas palabras que se encuentren en el ámbito de la negación. Este método (ver Sección 6.3.1), consiste en eliminar los grados emocionales de las palabras que se ven afectadas por la negación. Simplemente eliminando estos valores ya notamos cierta mejora en el marcado de textos porque dejamos de considerar la emoción equivocada. Por ejemplo, en la frase “Ella no está alegre” dejaríamos de ver la alegría como emoción significativa de la frase ya que “alegre” sería la única palabra emocional de la frase y al estar negada perdería sus grados emocionales. Evaluando este método nos damos cuenta que se mejoran los resultados pero que aún hay posibilidades de mejora, por ejemplo en la oración “No lo estás haciendo bien” no solo deberíamos perder la alegría de la frase sino que necesitaríamos que la tristeza y el enfado aumenten para que el marcado sea más correcto. Por lo tanto, se siguen planteando nuevos métodos.
- Sustitución de los grados emocionales de una palabra negada por los grados emocionales de su antónima. Este método, explicado en la Sección 6.3.2, consiste en encontrar qué palabras son afectadas por la negación y en vez de utilizar sus grados emocionales, buscar un antónimo de la palabra y utilizar los grados emocionales de la palabra antónima para calcular la emoción final de la frase. Este método daba algunos problemas porque había palabras que no tenían antónimos y porque algunas palabras antónimas tenían grados emocionales muy similares a la palabra que teníamos que negar. Por ejemplo, en la frase “No hemos ido al cine” tras buscar la palabra antónima de “ido” y sustituirla para calcular los grados emocionales de la nueva frase, que quedaría “No hemos venido al cine”, el resultado de las emociones es similar, dado que “ido” y “venido” son palabras semánticamente opuestas pero no lo son emocionalmente.
- Inversión de los grados emocionales de las palabras afectadas por la negación. Este

método, explicado en la Sección 6.3.3, consiste en invertir los grados emocionales de las palabras afectadas por la negación. La inversión se realiza restando 4 (valor máximo de una emoción en el diccionario) menos el grado para la emoción de la palabra. Utilizando este método nos damos cuenta de que en palabras con emociones muy bajas (que rondan el 0) la inversión dispara su carga emocional de forma incorrecta. Por ejemplo la palabra “amigo” que tiene como grados emocionales: 0,53 en tristeza, 0,57 de miedo, 3,53 de alegría, 0,2 de enfado y 0,12 de asco si se encontrase en el ámbito de una negación los grados pasarían a ser: 3,47 de tristeza, 3,43 de miedo, 0,47 de alegría, 3,8 de enfado y 3,88 de asco. Vemos en esta palabra que las emociones tristeza, miedo, enfado y asco tras ser invertidas han crecido enormemente sus grados cuando en realidad no deberían haber aumentado tanto. A eso se suma que en algunos de los textos utilizados para la evaluación la inversión conseguía que tuvieran como emociones significativas todas las emociones posibles (tristeza, miedo, alegría, enfado y asco) al mismo tiempo lo cual no es correcto.

- Inversión de los grados emocionales significativos de las palabras afectadas por la negación. En esta nueva aproximación explicada en detalle en la Sección 6.3.4, solo se invierten los grados emocionales de las emociones significativas (valor superior a 1,5). Con este método evitamos que aparezcan grados emocionales elevados en categorías que originalmente no eran significativas. Sin embargo, este tratamiento también provoca un problema similar al anterior y es que a pesar de invertir solo las emociones significativas al ser el minuendo de la inversión el valor máximo del diccionario emocional (4) si una palabra presenta un 1,6 en la emoción que se está invirtiendo al hacer la resta (4-1,6) nos volverá a dar un valor alto cuando en realidad queremos disminuir la emoción. Esto provoca de nuevo que algunos textos tengan como emociones significativas algunas que no corresponderían, llegando incluso a decir que una frase es triste y alegre a la vez, lo cual no tiene mucho sentido.
- Transferencia de grados emocionales de las palabras negadas a las categorías antónimas. Este método, explicado en la Sección 6.3.5, consiste en tener distintas relaciones de antonimia entre las categorías emocionales que se manejan. Para este método, primero se ha decidido invertir los grados emocionales de las palabras que se encuentran afectadas por una negación sintáctica de distinta manera. Los grados emocionales de las categorías que estén entre 0-2, serán consideradas emociones débiles y los grados emocionales de las categorías que están entre el 2-4, serán consideradas como emociones fuertes. La inversión de las emociones débiles consistirá en utilizar el valor medio del rango de valores que puede poseer cada categoría emocional en el diccionario para realizar la inversión, quedando de la siguiente manera: 2-grados emocionales, mientras que la inversión de las emociones fuertes consistirá en utilizar el máximo valor del rango de valores que puede poseer cada categoría emocional en el diccionario para realizar la inversión, quedando de la siguiente manera: 4-grados emocionales. También hemos creado para esta aproximación dos relaciones de antonimia entre las categorías emocionales con las que trabajamos: una antonimia total del 100 % entre la tristeza y la alegría y una antonimia parcial del 50 % entre el enfado y la alegría.

Cuando se localiza una palabra que se encuentra en el ámbito de la negación, lo primero que se hará será resolver las relaciones de antonimia entre las categorías emocionales. Si una palabra tiene un grado original (antes de ser invertido) de tristeza significativo (superior al 1,5) entonces sumaremos a la alegría el valor original que poseía la tristeza o si tiene un grado original de alegría significativo (superior al 1,5) entonces sumaremos a la tristeza el valor original que tenía la alegría. Por otra

parte, respecto a la otra relación de antonimia, la parcial, funcionará de manera similar pero en vez de sumar a la emoción contraria el valor completo que poseía la emoción le sumaremos solo un 50 % de lo que tenía. Tras evaluar esta mejora nos damos cuenta que no invierte siempre de manera adecuada las emociones ocurriendo un problema parecido al que ocurría con otros métodos, que presentamos en algunos textos emociones no adecuadas como la tristeza y la alegría al mismo tiempo.

Tras ver todos los métodos propuestos para tratar la negación concluimos con que son buenos primeros análisis para el tratamiento de la negación a pesar de que en la mayoría de ellos los resultados no son los esperados y por tanto aún quedan cosas por estudiar y mejorar. Sin embargo, el método que obtiene mejores resultados es el primero: no considerar palabras emocionales a aquellas palabras que se encuentren en el ámbito de una negación. Asumiremos que este método es el que varía la obtención de los grados emocionales de la manera más aceptable y es el método que se ha incorporado a la nueva versión del sistema.

Por último, y para conseguir el último objetivo se evalúan todas las modificaciones realizadas. Esto produce una mejora interesante en el marcado emocional de textos, que se comenta en la Sección 7.4 donde vemos que el marcado de textos efectivamente aporta una corrección a los resultados del análisis de los textos que aparecen en el Apéndice A. Aumentando el diccionario afectivo e incorporando el tratamiento de los modificadores y de la negación al sistema ya se aprecia un notable cambio. La suma de estas mejoras con las otras realizadas que ya hemos repasado, conlleva ciertos cambios que se pueden ver en la Sección 7.2 a nivel aplicación web que se han llevado a cabo para que la interfaz siga funcionando correctamente. También se realizan algunos cambios en los servicios web en los que se basa el sistema que se explicaron en la Sección 7.3.

Con todos estos cambios se consigue mejorar el marcado emocional llegando a la versión 2.0 del EmoTraductor.

## 8.2. Trabajo Futuro

A pesar de haber conseguido nuestros objetivos en el trabajo, aún quedan algunos matices y mejoras que se pueden añadir al EmoTraductor 2.0 para crear una herramienta mucho más potente y con mejores resultados.

En primer lugar, aunque hayamos incrementado el volumen del diccionario afectivo que empleamos, este aún se queda corto y debería ser ampliado. Se podría ampliar una vez más con nuevas palabras, lo que haría que de nuevo se aumentase el número de palabras que se pueden marcar y por tanto, mejorase el marcado. Pero también se podría ir más allá introduciendo en el diccionario no únicamente palabras sino frases hechas, expresiones coloquiales (las cuales suelen tener mucha carga emocional) o refranes y dichos del castellano (los cuales no siempre usan el significado original de las palabras). Esta ampliación llevaría consigo el cambio de los servicios web en los que se basa la aplicación porque deberían ser capaces de reconocer y buscar estas expresiones en el diccionario. Ahora mismo el EmoTraductor 2.0 procesa palabra a palabra.

En segundo lugar, otro posible punto de mejora, volverían a ser los modificadores. A pesar de haber implementado un método para tenerlos en cuenta, el listado de modificadores actual se debería aumentar, de esta manera se enriquecería el marcado de texto porque se tratarían modificadores que hasta ahora no se tenían en cuenta, pudiendo añadir muchos adverbios como “puramente” o “simplemente”.

En tercer lugar, se ha hecho un estudio de cómo aplicar la negación utilizando cinco métodos distintos donde sólo uno nos ha parecido adecuado aunque creemos que puede

mejorarse. Es por tanto, un punto a tener en cuenta a la hora de seguir desarrollando la herramienta, se deben seguir estudiando métodos distintos para el tratamiento de la negación. Algunas de las posibilidades para implantar en el sistema y mejorar el tratamiento de la negación podrían ser:

- Añadir al sistema el tratamiento de dimensiones emocionales. Esta vía es bastante factible, dado que los diccionarios sobre los que se basa nuestro diccionario emocional cuentan con las dimensiones emocionales y se excluyeron al realizar el EmoTraductor 1.0.
- Añadir nuevas categorías emocionales. Se puede ampliar el diccionario emocional basándose en ontologías realizadas para considerar más categorías emocionales que se puedan enfrentar entre sí para crear más relaciones de antonimia y por tanto tratar la negación como hemos visto en otros trabajos.

En cuarto lugar, otro posible trabajo futuro sería añadir más tratamientos interesantes al sistema. Aún hay muchos factores que pueden afectar a la carga emocional de un texto que no se están teniendo en cuenta. Por ejemplo, como vimos en la Sección 2.5, se podría añadir el tratamiento de los emoticonos en el sistema para dotar de un mejor marcado de texto a la herramienta. Otra posible mejora sería el reconocimiento de frases subordinadas o las frases coordinadas adversativas que cambiarán las emociones de los textos. También añadir información del contexto de cada texto mejorará el marcado de emociones. Dado que no será lo mismo decir “Su madre se murió” que “La bruja se murió” donde claramente la primera frase es muy triste y la segunda podría ser alegre si estuviéramos hablando de una bruja mala de un cuento.

En quinto lugar, se podría añadir al sistema un método de aprendizaje automático para que a raíz del uso que se de a la aplicación y con los logs de todas las evaluaciones que se han realizado, el sistema sea capaz de incorporar nuevas palabras al diccionario teniendo en cuenta el contexto o el uso que se de a estas palabras en otros textos. De esta manera, no solo enriqueceríamos el diccionario añadiendo nuevas palabras sino que también podríamos modificar los valores de las palabras ya marcadas según su contexto. Para ello se podría crear en el diccionario diferentes entradas de una palabra que varíen sus emociones según el contexto que tenga el texto evaluado.

## Introduction

**RESUMEN:** In this chapter we will make an introduction to the work that we are going to present. First, Section 9.1 will explain the motivation behind this work. Next, in Section 9.2, the proposed objectives for the work will be presented. Finally, Section 9.3 shows the structure that the memory will have.

### 9.1. Motivation

Thanks to the fast growth of the Internet and the fast evolution of communication systems, we increasingly have more information available. This supposes many advantages in our day to day, but it also increases the complication that comes with understanding and interpreting all the information we obtain and knowing how to handle it. The difficulties we discussed, sometimes increase due to the lack of emotional empathy that people with Autism Spectrum Disorders (ASD) have, especially with Asperger Syndrome. This lack of emotional empathy prevents them from recognizing the affective content of a text and may misinterpret its emotional charge. It also prevents them from being able to communicate adequately online with other people since they do not know what emotions can be interpreted in the texts they write.

To solve these problems it is necessary to create tools capable of interpreting the affective load of texts to help the digital integration of people with this type of difficulties. Thanks to affective computing these tools can be developed using Natural Language Processing (NLP) techniques. However, despite the fact that natural language processing is increasingly used in more situations and with success, it is still difficult to find the best method to interpret the affective content of a text due to the degree of subjectivity not only of the emotions, but also of the language itself.

To all the complication of interpreting text through NLP, it is added that most of the studies and work carried out in this field are carried out in English. Thus, we find few works and tools that are used in a standard way or with greater relevance in Spanish and that can serve as a guide for any of the aspects that we need. However, there are some tools that use Spanish to make this type of interpretation. For example, one of these tools is the EmoTraductor 1.0 developed in 2018 at the Faculty of Computing at the Complutense University of Madrid, as part of the TFG by the author of this TFM along with two other colleagues (Eugercios et al., 2018).

EmoTraductor 1.0 is a web application developed with the aim of emotionally marking texts entered by users through a web interface. This application uses emotional categories

to mark the texts, specifically the five basic emotions (sadness, fear, joy, anger and disgust). The presence of each of the categories in the text is shown by what we call emotional degrees, which are numbers that indicate the presence of the emotion in the text (with 0 being the value that indicates its absence). Although the results obtained in EmoTraductor 1.0 were promising, the emotional markup had deficiencies because some complex phenomena of natural language were not considered.

The main objective of this work will be to improve EmoTraductor 1.0 to improve the emotional markup of text by incorporating more advanced natural language processing than that performed in the previous version. In this TFM we will analyze how different phenomena that appear in natural language (negations, modifiers ...) affect the emotional charge of a text and we will incorporate its treatment into the emotional translator in order to improve the results it obtains.

## 9.2. Objectives

The main objective of this work is to improve the emotional marking mechanisms offered by the EmoTraductor 1.0 tool. This tool is a good first approach to an emotion marker in text, although it has some limitations such as a small emotional dictionary or very simple text processing that leaves out important aspects of language that directly influence emotional charge such as negation.

In order to fulfill our main objective, to improve the emotional marking of texts, we will study the operation of EmoTraductor 1.0 detecting the weak points of the marking and trying to improve them, and then we will expand the system to add some improvements.

With this in mind we go on to list our specific objectives:

- Fix the malfunctions found after evaluating the EmoTraductor 1.0.
- Increase the emotional dictionary to expand the coverage of the EmoTraductor.
- Add the treatment of modifiers that can alter the emotional charge of the words they affect.
- Add the treatment of negation in the texts.
- Evaluate all the improvements made to decide which of them are finally incorporated into the new version of our tool.

All the code and tests performed are in the following public GitHub repository:

<https://github.com/NILGroup/NIL-API-Emociones/tree/Paloma>

## 9.3. Memory structure

The structure of this document consists of the following chapters:

- In the chapter 1 the motivation and objectives of this TFM are introduced, as well as the structure of the document.
- In the chapter 2 a review of the state of the art on the subject to be treated is based on a first vision of what is affective computing, the representation of emotions and the emotional marking of texts. Then some of the most relevant emotional dictionaries

for our work are exposed and the importance of emoticons, modifiers and negations in the emotional marking of texts is explained. Finally, some of the existing emotional marking tools are examined.

- Chapter 3 presents the EmoTraductor 1.0 web application, along with its web services, developed prior to this work. It will be our starting point and will be studied to improve its operation.
- The chapter 4 shows the results of a first analysis of the EmoTraductor 1.0 web application and explains a series of improvements that are made and evaluated on the application. These improvements range from changes in the emotional dictionary that you use as a base to the complete change of a web service.
- The chapter 5 explains how modifiers are going to be treated in the emotional markup of texts used by the EmoTraductor, and the results obtained are evaluated.
- The chapter 6 explains how the negations will be treated in the emotional marking of texts used by the EmoTraductor, including the evaluation of the impact of the proposed approaches.
- In the chapter 7 the validity of the new improvements and their application will be discussed, as well as the necessary changes that should have been made in the application.
- Chapter 8 presents the TFM conclusions along with future work.
- The chapters 9 and 10 are the chapters 1 and 8 translated into English.



# Chapter 10

## Conclusions and future work

In this chapter we will see, in Section 10.1 the conclusions of this work evaluating whether the objectives set out in Section 1.2 (or in English 9.2) have been met and, if so, in what way. In Section 10.2 we will mark some points that can serve as a basis to continue improving the tool presented in this work.

### 10.1. Conclusions

After completing the project and explaining all the details of design, implementation and results that have been followed in this document, it is possible to draw the following conclusions according to the objectives that were set at the beginning.

Our main objective in this work was to improve the emotional marking process of texts carried out by the EmoTraductor 1.0 tool. This tool is a web application that aims to help people who have difficulties in interpreting the emotional content of a text by emotionally marking texts using emotional categories with an associated degree of intensity.

In order to fulfill this main objective, we proposed to go through the following objectives:

- Fix the malfunctions found after evaluating the EmoTraductor 1.0.
- Increase the emotional dictionary to expand the coverage of the EmoTraductor.
- Add the treatment of modifiers that can alter the emotional charge of the words they affect.
- Add the treatment of negation in the texts.
- Evaluate all the improvements made to decide which of them are finally incorporated into the new version of our tool.

In order to solve the malfunctions found after the evaluation of the EmoTraductor 1.0 and increase the emotional dictionary, we use as a basis for improvement part of the future work proposed in the memory of the TFG where the EmoTraductor 1.0 was implemented. Together with the TFG memory, the operation of each of the web services used by the tool is evaluated to find malfunctions that did not appear as future work in the TFG.

With the changes made during this TFM, the following have been achieved:

1. Improve emotional marking thanks to the normalization of emotional degrees in the affective dictionary. As we saw in Section 4.2, when using the initial affective dictionary where the values of the emotions were between 1 and 5 (where 1 was the absence of emotion and 5 the total appearance of the emotion) when we would weigh the sentences according to their typology. Having the value 1 as absence of emotion made the emotional degrees of the weighted sentences increase emotions that should not appear. By normalizing the dictionary by changing the range of the emotional degrees so that they oscillate between 0 and 4 (where 0 is the absence of emotion and 4 the total appearance of it) we avoid that texts that contain different types of phrases with different weights associated with each type of phrase, incorporate non-existent degrees into the text.
2. Make an expansion of the affective dictionary. This improvement is explained in Section 4.3 and consists of adding more entries to the affective dictionary on which the application is based. After this change the marking of texts is improved since with the new dictionary many more words can be taken into account when determining the affective content of the texts, making the resulting emotions more correct. Along with this expansion of entries in the emotional dictionary there is also an expansion of information (gender, number, grammatical category) of each word found in it in order to make more correct searches.
3. Change the criteria for selecting emotional words. This improvement, which we found explained in Section 4.4, consists of the elimination of the limit (2,5) that some emotional category of a word must have in order for the EmoTraductor 1.0 to consider it emotional and collaborate in the final calculation of the emotions of a phrase. This improvement accompanies the previous one since, by eliminating the limit, all the words that have an entry in the dictionary are taken into account when emotionally marking a text and not only those that had a degree greater than an emotional category for an emotional category established.
4. Improve the web service that gets the emotional degrees of a word. Thanks to the changes in this service, explained in Section 4.6, which consist of modifying the way EmoTraductor 1.0 used to search for a word in the dictionary, adding the search by the original word or the search by motto of the word, gender and number. With this change in the way we search, we ensure that we find as many words as possible in the emotional dictionary and therefore we ensure that the final emotional degrees of a text are more rigorous because they consider more words.
5. Improve the web service that obtains the emotional degrees of a text. As explained in Section 4.5, in EmoTraductor 1.0 there was a problem when calculating the divisor used to obtain the final emotional degrees of a text, since it counted each phrase as its weighting instead of a single phrase, that is, if an exclamatory phrase appeared, it was counted as two sentences instead of one, making the results not correct. Therefore, the way of counting the phrases that make up a text has been modified.

Furthermore, in this TFM, as explained in Chapter 5, we have introduced the treatment of modifiers in the emotional marking of texts. Remember that we call a word modifier that modifies the affective load of the words it accompanies. To do this, we have incorporated into the system a list of possible modifiers, where each modifier found in this list has an associated performance percentage. This percentage is the modification it will make to the affective load of the words that the modifier in question affects. So if we have a modifier

like “muy” with a percentage of +75% the word or words that are affected by the term “muy” will increase their emotions by 75%. In order to determine the scope of a modifier, that is, to determine which words are affected by a modifier, the dependency tree treatment has been added to the system. After introducing this improvement in the system, a new evaluation was carried out considering that we detect and treat the modifiers appropriately and that this improvement improves the results. For this reason, we consider the objective of dealing with modifiers in the emotional marking of texts as fulfilled.

Next and as it is told in Chapter 6, the treatment of negation is incorporated into the system. The first thing we do is classify negations into two types: nominal negations and syntactic negations. Nominal negations are those that are identified by using words like “casi”, “poco” or “menos” and that reduce the affective load of the words that it affects. This type of negation is treated by the system as modifiers that decrease the emotional charge of the words it affects. Syntactic negations are those that use the word “no” and usually affect the verb of the sentence thus negating the full meaning of the phrase. In this type of denial, we only consider those where the particle “no” appears. To resolve the treatment of this type of denial, we propose five different approaches:

- Do not consider emotional words to those words that are in the scope of the negation. This method (see Section 6.3.1), consists of eliminating the emotional degrees of the words that are affected by the negation. Simply deleting these values we already notice some improvement in the text markup because we stop considering the wrong emotion. For example, in the phrase “Ella no está alegre” we would stop seeing joy as a significant emotion of the phrase since “alegre” would be the only emotional word in the phrase and when it is denied it would lose its emotional degrees. Evaluating this method we realize that the results are improved but that there are still possibilities for improvement, for example in the sentence “No lo estás haciendo bien” not only should we lose the joy of the phrase but we would need sadness and anger increase so that the marking is more correct. Therefore, new methods continue to be proposed.
- Substitution of the emotional degrees of a negated word for the emotional degrees of its antonym. This method, explained in Section 6.3.2, consists of finding which words are affected by the negation and instead of using their emotional degrees, find an antonym of the word and use the emotional degrees of the antonym word to calculate to the final emotion of the sentence. This method gave some problems because there were words that did not have antonyms and because some antonyms had emotional degrees very similar to the word that we had to deny. For example, in the phrase “No hemos ido al cine” after searching for the antonym word of “ido” and replacing it to calculate the emotional degrees of the new phrase, which would read “No hemos venido al cine”, the result of the emotions is similar, since “ido” and “venido” are semantically opposed words but they are not emotionally so.
- Inversion of the emotional degrees of the words affected by the negation. This method, explained in Section 6.3.3, consists of inverting the emotional degrees of the words affected by the negation. The investment is made by subtracting 4 (maximum value of an emotion in the dictionary) minus the degree for the word emotion. Using this method, we realize that in words with very low emotions (around 0), the investment triggers its emotional charge incorrectly. For example, the word “amigo” that has emotional degrees: 0,53 in sadness, 0,57 in fear, 3,53 in happiness, 0,2 in anger and 0,12 in disgust if it were in the field degrees of denial would become: 3,47 of sadness, 3,43 of fear, 0,47 of happiness, 3,8 of anger and 3,88 of disgust. We see in this word that the emotions sadness, fear, anger and disgust after being inverted

have greatly increased their degrees when in fact they should not have increased so much. In addition to this, in some of the texts used for the evaluation, the investment made them have as significant emotions all possible emotions (sadness, fear, happiness, anger and disgust) at the same time, which is not correct.

- Inversion of the significant emotional degrees of the words affected by the negation. In this new approach explained in detail in Section 6.3.4, only the emotional degrees of significant emotions are reversed (value greater than 1,5). With this method we prevent high emotional degrees from appearing in categories that were not originally significant. However, this treatment also causes a problem similar to the previous one, and that is that despite investing only significant emotions, as the investment minuend is the maximum value of the emotional dictionary (4) if a word has 1,6 in emotion that is being invested by doing the subtraction (4-1,6) will give us a high value again when we really want to decrease the emotion. This again causes some texts to have as significant emotions some that would not correspond, even going so far as to say that a phrase is sad and happy at the same time, which does not make much sense.
- Transfer of emotional degrees from the negated words to the antonym categories. This method, explained in Section 6.3.5, consists of having different antonymic relationships between the emotional categories that are handled. For this method, it has first been decided to invert the emotional degrees of the words that are affected by a syntactic negation in different ways. The emotional degrees of the categories that are between 0-2, will be considered weak emotions and the emotional degrees of the categories that are between 2-4, will be considered as strong emotions. The inversion of weak emotions will consist of using the average value of the range of values that each emotional category can possess in the dictionary to make the investment, remaining as follows: 2-emotional degrees, while the inversion of strong emotions will consist of in using the maximum value of the range of values that each emotional category can possess in the dictionary to make the investment, remaining as follows: 4-emotional degrees. We have also created for this approach two antonymic relationships between the emotional categories with which we work: a total antonym of 100% between sadness and joy and a partial antonym of 50% between anger and joy.

When a word is found that is in the field of negation, the first thing that will be done is to resolve the antonymic relationships between the emotional categories. If a word has an original degree (before being inverted) of significant sadness (greater than 1,5) then we will add to the joy the original value that the sadness possessed or if it has an original degree of significant happiness (greater than 1,5 ) then we will add to the sadness the original value that happiness had. On the other hand, with respect to the other antonymic relationship, the partial one, will work in a similar way but instead of adding to the opposite emotion the full value that the emotion possessed, we will add only 50% of what it had. After evaluating this improvement, we realize that emotions are not always appropriately reversed, a problem similar to that occurring with other methods occurs, which we present in some texts inappropriate emotions such as sadness and joy at the same time.

After looking at all the proposed methods for treating negation, we conclude that they are good first analyzes for the treatment of negation, despite the fact that in most of them the results are not as expected and therefore there are still things to study and improve. However, the method that obtains the best results is the first: do not consider emotional

words to be those words that are within the scope of a negation. We will assume that this method is the one that varies the obtaining of emotional degrees in the most acceptable way and is the method that has been incorporated into the new version of the system.

Finally, and to achieve the last objective, all the modifications made are evaluated. This produces an interesting improvement in the emotional marking of texts, which is discussed in Section 7.4 where we see that text marking effectively provides a correction to the results of the analysis of the texts that appear in the Appendix A. Increasing the affective dictionary and incorporating the treatment of modifiers and negation into the system already shows a notable change. The sum of these improvements with the others that we have already reviewed, involves certain changes that can be seen in Section 7.2 at the web application level that have been carried out so that the interface continues to function properly. Some changes are also made to the web services on which the system is based, which were explained in Section 7.3.

With all these changes, it is possible to improve emotional marking, reaching version 2.0 of the EmoTraductor.

## 10.2. Future work

Despite having achieved our goals at work, there are still some nuances and improvements that can be added to EmoTraductor 2.0 to create a much more powerful tool with better results.

First, although we have increased the volume of the affective dictionary that we use, it still falls short and should be expanded. It could be expanded once again with new words, which would again increase the number of words that can be marked and therefore improve the markup. But you could also go further by introducing in the dictionary not only words but made phrases, colloquial expressions (which usually have a lot of emotional charge) or Spanish sayings and sayings (which do not always use the original meaning of the words). This extension would bring about the change of the web services on which the application is based because they should be able to recognize and look for these expressions in the dictionary. Right now EmoTraductor 2.0 processes word by word.

Secondly, another possible point of improvement, would be the modifiers again. Despite having implemented a method to take them into account, the current list of modifiers should be increased, in this way the text marking would be enriched because it would deal with modifiers that until now were not taken into account, being able to add many adverbs such as “puramente” or “simplemente”.

Thirdly, a study has been made of how to apply negation using five different methods where only one has seemed adequate, although we believe that it can be improved. Therefore, it is a point to keep in mind when further developing the tool, different methods must be studied for the treatment of denial. Some of the possibilities to implant in the system and improve the treatment of denial could be:

- Add to the system the treatment of emotional dimensions. This way is quite feasible, since the dictionaries on which our emotional dictionary is based have the emotional dimensions and were excluded when making EmoTraductor 1.0.
- Add new emotional categories. The emotional dictionary can be expanded based on ontologies made to consider more emotional categories that can face each other to create more antonymic relationships and therefore treat denial as we have seen in other works.

Fourth, another possible future job would be to add more interesting treatments to the system. There are still many factors that can affect the emotional charge of a text that are not being taken into account. For example, as we saw in Section 2.5, you could add the treatment of emoticons in the system to provide better text marking to the tool. Another possible improvement would be the recognition of subordinate phrases or the adversarial coordinated phrases that will change the emotions of the texts. Also adding information from the context of each text will improve the marking of emotions. Since it will not be the same to say “Su madre se murió” as ‘La bruja se murió” where clearly the first sentence is very sad and the second one could be happy if we were talking about a bad witch (“bruja”) from a story.

Fifth, a machine learning method could be added to the system so that as a result of the use given to the application and with the logs of all the evaluations that have been carried out, the system is capable of incorporating new words into the dictionary having Consider the context or use of these words in other texts. In this way, not only would we enrich the dictionary by adding new words, but we could also modify the values of the already marked words according to their context. To do this, you could create different word entries in the dictionary that vary your emotions depending on the context of the evaluated text.

# Bibliografía

- AMORES, M., ARCO, L. y ARTILES, M. PosNeg opinion: Una herramienta para gestionar comentarios de la Web. *Revista Cubana de Ciencias Informáticas*, vol. 9(1), páginas 20–12, 2015.
- AMORES, M., ARCO, L. y BARRERA, A. Efectos de la Negación, Modificadores, Jergas, Abreviaturas y Emoticonos en el Análisis de Sentimiento. En *IWSW, 2016*, páginas 43–53. 2016.
- ANDREEVSKAIA, A. y BERGLER, S. CLaC and CLaC-NB: Knowledge-based and corpus-based approaches to sentiment tagging. En *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, páginas 117–120. 2007.
- BACCIANELLA, S., ESULI, A. y SEBASTIANI, F. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. En *LREC*, vol. 10, páginas 2200–2204. 2010.
- BALDASSARRI, S. Computación afectiva: tecnología y emociones para mejorar la experiencia del usuario. *Bit & Byte*, vol. 2, 2016.
- BOSQUEZ, V., SANZ, C., BALDASSARRI, S., RIBADENEIRA, E., VALENCIA, G., BARRAGAN, R., ÁNGEL CAMACHO-CASTILLO, SHAURI-ROMERO, J. y CAMACHO-CASTILLO, L. La computación afectiva: emociones, tecnologías y su relación con la educación virtual. *Revista de Investigación Talentos*, vol. 5(1), páginas 94–103, 2018.
- BRADLEY, M. M. y LANG, P. J. Affective norms for English words (ANEW): Instruction manual and affective ratings. Informe técnico, Technical report C-1, the center for research in psychophysiology, 1999.
- CARRILLO, J. *Un modelo lingüístico-semántico basado en emociones para la clasificación de textos según su polaridad e intensidad*. Tesis Doctoral, Universidad Complutense de Madrid, 2011.
- COWIE, R. y CORNELIUS, R. Describing the emotional states that are expressed in speech. *Speech communication*, vol. 40(1-2), páginas 5–32, 2003.
- CUADRADO, I., MARTÍN-MORA, G., FERNÁNDEZ, I. y MERCHÁN, I. M. Análisis metaperceptivo de los elementos expresivos de internet: clasificación y uso de emoticonos. *Internet Latent Corpus Journal*, vol. 4(1), páginas 29–45, 2014.
- CUEVAS, J. Emociones: calculemos. Las promesas de la computación afectiva. 2015.

- DAS, S. y CHEN, M. Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science*, vol. 53, páginas 1375–1388, 2007.
- DÍAZ, I., SIDOROV, G. y SUÁREZ, S. Creación y evaluación de un diccionario marcado con emociones y ponderado para el español. *Onomazein*, (29), páginas 31–46, 2014.
- DUCHON, A., PEREA, M., SEBASTIÁN-GALLÉS, N., MARTÍ, A. y CARREIRAS, M. EsPal: One-stop shopping for Spanish word properties. *Behavior research methods*, vol. 45(4), páginas 1246–1258, 2013.
- EKMAN, P. Universal and cultural differences in facial expression of emotion. En *Nebraska symposium on motivation*, vol. 19, páginas 207–284. 1972.
- EUGERCIOS, G., GUTIÉRREZ, P. y KALOYANOVA, E. *Análisis emocional para la inclusión digital*. 2018. Trabajo de Fin de Grado en Ingeniería Informática (Universidad Complutense, Facultad de Informática, curso 2017/2018).
- FERRÉ, P., GUASCH, M., MARTÍNEZ-GARCÍA, N., FRAGA, I. y HINOJOSA, J. A. Moved by words: Affective ratings for a set of 2,266 Spanish words in five discrete emotion categories. *Behavior research methods*, vol. 49(3), páginas 1082–1094, 2017.
- FERRÉ, P., GUASCH, M., MOLDOVAN, C. y SÁNCHEZ-CASAS, R. Affective norms for 380 Spanish words belonging to three different semantic categories. *Behavior Research Methods*, vol. 44(2), páginas 395–403, 2012.
- FRANCISCO, V. *Identificación Automática del Contenido Afectivo de un Texto y su Papel en la Presentación de Información*. Tesis Doctoral, Universidad Complutense de Madrid, 2008.
- FRANCISCO, V. y GERVÁS, P. Automated mark up of affective information in english texts. En *International Conference on Text, Speech and Dialogue 2006*, páginas 375–382. 2006.
- GUASCH, M., FERRÉ, P. y FRAGA, I. Spanish norms for affective and lexico-semantic variables for 1,400 words. *Behavior Research Methods*, vol. 48(4), páginas 1358–1369, 2016.
- HINOJOSA, J. A., MARTÍNEZ-GARCÍA, N., VILLALBA-GARCÍA, C., FERNÁNDEZ-FOLGUERIAS, U., SÁNCHEZ-CARMONA, A., POZO, M. A. y MONTORO, P. R. Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behavior research methods*, vol. 48(1), páginas 272–284, 2016.
- JIA, L., YU, C. y MENG, W. The Effect of Negation on Sentiment Analysis and Retrieval Effectiveness. En *Proceedings of the 18th ACM Conference on Information and Knowledge Management 2009*, página 1827–1830. Association for Computing Machinery, 2009.
- JIMÉNEZ, S. M., MARTÍNEZ-CÁMARA, E., MARTÍN, M. y MOLINA, M. Tratamiento de la Negación en el Análisis de Opiniones en Español. vol. 54, páginas 37–44, 2015.
- EL KALIOUBY, R., PICARD, R. W. y BARON-COHEN, S. Affective computing and autism. *Annals of the New York Academy of Sciences*, vol. 1093(1), páginas 228–248, 2006.
- KENNEDY, A. y INKPEN, D. Sentiment classification of movie reviews using contextual valence shifters. *Computational intelligence*, vol. 22(2), páginas 110–125, 2006.

- LANG, P. J., BRADLEY, M. M., CUTHBERT, B. N. ET AL. International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, vol. 1, páginas 39–58, 1997.
- LIN, D. *Dependency-Based Evaluation of Minipar*, páginas 317–329. Springer Netherlands, Dordrecht, 2003.
- MAGNINI, B. y CAVAGLIÀ, G. Integrating Subject Field Codes into WordNet. En *LREC 2000*, páginas 1413–1418. 2000.
- MILLER, G. A., BECKWITH, R., FELLBAUM, C., GROSS, D. y MILLER, K. Introduction to WordNet: An on-line lexical database. *International journal of lexicography*, vol. 3(4), páginas 235–244, 1990.
- MORENO, A., PÉREZ, Á. y TORRES, S. Sentitext: sistema de análisis de sentimiento para el español. *Procesamiento del lenguaje natural*, vol. 45, páginas 297–298, 2010a.
- MORENO, A., PINEDA, F. y HIDALGO, R. Análisis de Valoraciones de Usuario de Hoteles con Sentitext: un sistema de análisis de sentimiento independiente del dominio. *Procesamiento del Lenguaje Natural, Revista nº 45, septiembre 2010*, páginas 31–39, 2010b.
- MORENO-ORTIZ, A. Lingmotif: A User-focused Sentiment Analysis Tool. páginas 153–140, 2017a.
- MORENO-ORTIZ, A. Tecnolengua Lingmotif at EmoInt-2017: A lexicon-based approach. páginas 225–232, 2017b.
- PANG, B., LEE, L. y VAITHYANATHAN, S. Thumbs up? Sentiment Classification using Machine Learning Techniques. En *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, páginas 79–86. Association for Computational Linguistics, 2002.
- PICARD, R. W. *Affective computing*. MIT press, 1997.
- POLANYI, L. y ZAENEN, A. *Contextual Valence Shifters*. Springer Netherlands, 2006. ISBN 978-1-4020-4102-0.
- QUIRK, R., GREENBAUM, S., LEECH, G. y SVARTVIK, J. A comprehensive grammar of the english language. 1985.
- REDONDO, J., FRAGA, I., PADRÓN, I. y COMESAÑA, M. The Spanish adaptation of ANEW (affective norms for English words). *Behavior research methods*, vol. 39(3), páginas 600–605, 2007.
- RUSSELL, J. A. A circumplex model of affect. *Journal of personality and social psychology*, vol. 39(6), página 1161, 1980.
- SEBASTIÁN-GALLÉS, N., MARTÍ, M., CARREIRAS, M. y CUETOS, F. Lexesp: Una base de datos informatizado del español. 2000.
- SHEILA PLAZA, N. R. y ACOSTA, C. *API de servicios web orientados a accesibilidad*. 2016. Trabajo de Fin de Grado en Ingeniería Informática (Universidad Complutense, Facultad de Informática, curso 2015/2016).

- STADTHAGEN-GONZÁLEZ, H., FERRÉ, P., PÉREZ-SÁNCHEZ, M. A., IMBAULT, C. y HINOJOSA, J. A. Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior research methods*, vol. 50(5), páginas 1943–1952, 2018.
- STADTHAGEN-GONZÁLEZ, H., IBAULT, C., PÉREZ, M. A. y BRYLSBAERT, M. Norms of valence and arousal for 14,031 Spanish words. *Behavior research methods*, vol. 49(1), páginas 111–123, 2017.
- STONE, P., DUNPHY, D., SMITH, M. y OGLIVIE, D. *The General Inquirer: A Computer Approach to Content Analysis*. 1996.
- STRAPPARAVA, C. y VALITUTTI, A. WordNet-Affect: an affective extension of WordNet. páginas 1083–1086, 2004.
- TABOADA, M., BROOKE, J., TOFILOSKI, M., VOLL, K. y STEDE, M. Lexicon-based methods for sentiment analysis. *Computational linguistics*, vol. 37(2), páginas 267–307, 2011.
- VASHISHT, G. y THAKUR, S. Facebook as a corpus for emoticons-based sentiment analysis. *Int. J. Emerg. Technol. Advan. Eng.*, vol. 4, páginas 904–908, 2014.
- VILARES, D., ALONSO, M. A. y GÓMEZ-RODRÍGUEZ, C. Clasificación de polaridad en textos con opiniones en español mediante análisis sintáctico de dependencias. *Procesamiento del Lenguaje Natural*, vol. 50, páginas 13–20, 2013.
- WEGRZYN-WOLSKA, K., BOUGUEROUA, L., YU, H. y ZHONG, J. Explore the effects of emoticons on Twitter sentiment analysis. *Comput. Sci. Inf. Technol.*, vol. 2, página 65, 2016.
- WHITELAW, C., GARG, N. y ARGAMON, S. Using appraisal groups for sentiment analysis. En *Proceedings of the 14th ACM international conference on Information and knowledge management*, páginas 625–631. 2005.
- WIEGAND, M., BALAHUR, A., ROTH, B., KLAKOW, D. y MONTOROYO, A. A survey on the role of negation in sentiment analysis. En *Proceedings of the workshop on negation and speculation in natural language processing*, páginas 60–68. 2010.
- WUNDT, W. M. *Grundriss der psychologie*. A. Kröner, 1913.

## Textos para evaluar el mercado emocional

En este Apéndice aparecen todos los textos usados para la evaluación del mercado emocional de textos que realiza la herramienta. Están divididos en dos secciones, en la primera A.1, encontraremos los textos usados en el EmoTraductor 1.0 y en la segunda A.2 encontraremos los añadidos para las evaluaciones de este nuevo trabajo.

### A.1. Textos usados en la evaluación del EmoTraductor 1.0

#### A.1.1. Aladdín

Yasmín era una princesa y vivía en el palacio real. Su padre quería que se casara. Pero ella no estaba enamorada. Yasmín escapó del palacio. Deseaba ser libre. Yasmín conoció a Aladdín en el bazar. Con la ayuda del Genio, Yasmín y Aladdín se enamoraron.

#### A.1.2. Blancanieves

Había una vez una princesa dulce y gentil llamada Blancanieves. La malvada reina estaba celosa de la belleza Blancanieves. Blancanieves tuvo que huir y se refugió en la casa de los siete enanitos. La reina engañó a Blancanieves con una manzana envenenada. Pero llegó el príncipe, despertó a Blancanieves con un beso y vivieron felices para siempre.

#### A.1.3. Blancanieves y sus amigos

La amable y gentil Blancanieves tenía muchos amigos. Los animales del bosque eran sus amigos. Un pequeño pajarito era su amigo. El príncipe azul era su amigo. Pero los siete enanitos eran los mejores amigos de Blancanieves.

#### A.1.4. Blog agradecimiento

Una vez más, gracias. Una vez más me dejáis sin palabras y hacéis que todo merezca la pena. De verdad, gracias. Esto no es solo un regalo o una muñeca más. Quizás para algunos sí, pero para mí es algo muy muy especial. Significa que estáis ahí y que me apoyáis, y no sabéis las ganas de comerme el mundo que siento cuando os noto tan cerca. Os dejo fotos del paso a paso para que veáis todos los detalles y cómo está todo súper cuidado. ¡Menudo trabajazo! Gracias, no solo a la marca que me lo ha hecho, sino a todos por vuestro calor. Gracias a vosotros las PedrocheCampanadas tienen más sentido que nunca. Os adoro.

### A.1.5. Blog enfado

Los que me conocéis y los que me habéis seguido en redes sabréis que las críticas no las he llevado mal y he aguantado mucho. Pero es obvio que no es agradable leer mentiras semana tras semana y aún así hacer oídos sordos y seguir adelante con tu trabajo, tu familia y tus amigos. Claro que al ser 'personajes públicos' tenemos ese riesgo. No es todo de rosa y nos exponemos a que la gente hable. Bien, mal, regular. Como quieran. Pero ¿existe algún tipo de límite? ¿Quién lo pone? Después de la nochecita en Twitter y en prensa he visto que desde luego no existe ningún límite.

### A.1.6. Blog moda

Hoy os dejo este outfit de entretiempo en blanco y negro. llevo una blusa blanca con volantes de Storets que os he enseñado ya alguna vez por instagram, la combiné con LIA Jeans en negro y los zapatos de Dior. El abrigo línea oversize es de paño fino y lo estoy usando mucho estos días de ni frío ni calor. Esto es todo por hoy.

### A.1.7. Blog muerte

Una año mas en este 21 de enero, esta página en blanco se llena de palabras que brotan de mi interior, salpicadas de tantas cosas que contar que se escapan del propio entendimiento. Un año mas ha pasado, y ya son doce, desde aquella mañana tan fría, en que nuestros cuerpos se separaban, pero nuestras almas decidían estar conectadas para siempre. No podré verte, ni escucharte, ni achucharte... pero sin embargo si puedo sentirte a través de tantas personas que van llegando a través de Lágrimas de Vida. A través de tantas historias que se entremezclan con la nuestra, y convierten tantas lágrimas en vida, que puedo besarte, olerte, y achucharte, porque forman parte de mi cada instante.

### A.1.8. Blog reflexión

Y bueno como ya sabéis que me encanta compartir todo con vosotros, mal que le pese a más de uno, hoy quiero hablaros de un tema me lleva preocupando las últimas semanas. ¿Qué entendéis vosotros por autoestima? Quizá pueda contestármelo la gente que critica que algunos no lo tenemos... Yo (y por supuesto mi verdad o mi opinión no son universales no, son simplemente MI forma de ver la vida), creo que puede ser verdad que no tenga mucho autoestima, pero no lo veo como algo negativo del todo. He vivido situaciones en mi vida con determinadas personas que seguro que si me quisiese un poco más ahora mismo no formarían parte de mi vida. Sin embargo lo siguen haciendo, y no me arrepiento de ello. ¿A esto lo podemos llamar no quererse a uno mismo? Yo más bien diría: querer más al otro que a uno mismo. Frase que por cierto me recuerda bastante a la enseñanza de Dios: "ama al prójimo como a ti mismo", aunque se de bastante poco hoy en día. Es diferente no tener autoestima a querer evitar problemas y preferir tener un trato cordial con personas que han sido importantes en tu vida. De todas formas esta es una decisión muy personal de cada uno y me parece injusto que se critique o se comente abiertamente sin ni si quiera conocer a la persona ni saber por qué situaciones ha pasado. No creo que haya nadie que tenga siempre el autoestima alto en todo momento. Las situaciones que vivimos a diario pueden hacer que suba o baje, pero también creo que es de ser inteligente el saber mantener nuestro autoestima dentro de unos límites. Yo no soy una persona muy orgullosa y la verdad es los enfados se me pasan muy rápido, incluso prefiero pedir disculpas aunque no tenga claro del todo que ha sido mi culpa con tal de poner fin a la discusión. Esto es algo que podría definir

como una virtud más que como un "no me quiero a mí misma." que tengo el autoestima baja. Y me encanta que mis padres me hayan enseñado este tipo de valores y educación, de seguir siempre hacina delante te hagan lo que te hagan, perdonando y sin odiar ni guardar rencor a nadie... que eso es lo que acaba consumiendo a muchas personas hoy en día: el rencor y la envidia. En fin, me gustaría saber cómo anda vuestro autoestima y si alguna vez os han dicho que lo tenéis muy alto o muy bajo... Yo me sigo queriendo y valorando a mi misma pero en su justa medida para nunca perder la visión de la realidad.

### **A.1.9. Cenicienta**

Cenicienta trabajaba mucho, día y noche. Y a pesar de todo estaba alegre. Se divertía con sus amigos los ratones. También se ponía contenta cuando cantaba. Pero creer en sus sueños era lo que la hacía sentirse más feliz.

### **A.1.10. El castillo de bestia**

Bestia vivía en un castillo encantado. Plumero, la Señora Potts y Din-Dón dieron la bienvenida a Bella. Todos se esforzaron para conseguir que Bella se sintiera como en casa. Bella disfrutaba de las cenas con Bestia en el gran comedor del castillo. Bella se sintió feliz de que aquel castillo fuera su hogar para siempre.

### **A.1.11. El ratoncito Pérez**

En un barrio pobre, junto al bosque del Quinto Pino, vivía el ratón Pérez con su familia. Sus padres pasaban el día buscando por los rincones un pedazo de queso para alimentar a su familia. No era fácil: un escobazo por aquí, una trampa por allí, cada día llegaban a casa magullados. A Pérez, la vida de los ratones no le parecía nada envidiable. El pequeño roedor se lamentaba: Por qué no habré nacido gato, o pantera. Si al menos fuera rico viviría en otro barrio con más queso y menos trampas. Un día decidió que tenía que dar con un modo de salir adelante. Caminando pensativo, llegó al bosque del Quinto Pino. De entre los árboles le llegó el sonido de un aleteo. - ¿Quién anda ahí? -. susurró, medio escondido bajo una mata. Pérez abrió los ojos como platos cuando una niña se acercó a él sin tocar el suelo. Estaba volando. La niña se presentó: - Soy Pina, el hada de este bosque. Nunca te había visto por aquí. Pérez salió de la mata e, hipnotizado, saludó mientras miraba las alas que se movían sin parar: - Me llamo Pérez, por cierto, ¿qué es un hada? - ¡Es un ser mágico! - respondió Pina, sorprendida -. ¿Acaso no te han contado ningún cuento de hadas? Pérez contestó que sus padres llegaban acasa sin ánimo de contar cuentos. Pina sintió lástima por el pobre ratoncito. En verdad, mucha gente temía a los ratones y no les ponía las cosas fáciles. El hada pensó unos instantes y a continuación le hizo una propuesta más asombrosa que ratón alguno haya escuchado. - Mira, Pérez, podemos hacer un trato. Hay algo que podrías traerme. A cambio, te haré un regalo mágico. ¿Qué me dices? Pérez no podía creer que él pudiera hacer tratos con un hada. Pero en seguida reaccionó: ¿Qué tendría que traerte?, espero que no sea queso. El hada se rió. Cuando le dijo lo que esperaba de Pérez, el roedor se quedó de piedra. - ¿Cómo? - repitió, aturdido -. ¿Quieres que traiga los dientes de las niñas y niños de todo el mundo? Pina se lo explicó: - Querido Pérez, no te asustes. A los niños se les caen los dientes de leche para que pudan salirles los nuevos, más resistentes. Tú sólo tendrás que entrar en sus casas cuando duerman, recoger esos dientecitos y traérmelos. Pérez seguía sin verlo claro. Le preguntó cómo entraría en todas las casas. Al fin y al cabo, sólo era un pequeño ratón. Pina sacó de su bolso una llave dorada y le dijo que con ella podría abrir todas las puertas del mundo. Pérez se relamió:

podría entrar de noche en una gran quesería y... Pero el hada le leyó el pensamiento. - No, Pérez, nada de robar queso. A cambio de los dientes, te daré otra cosa. Y sacó de su bolso un cofre de madera. - De este cofre siempre saldrán monedas y regalos. Los regalos serán para las niñas y niños en agradecimiento por llevarte sus diente-cillos. Y las monedas serán para ti. Con ellas, podrás ayudar a muchos ratones. Pérez apenas podía creerlo. Sería el ayudante de un hada, viajaría por todo el mundo y tendría una gran recompensa. - Pero Pina, ¿para qué quieres los dientes? - Para hacer magia. Con polvo de dientes de leche y lava de volcán hago una pócima que reduce los objetos. Mezclado con trébol de cinco hojas, ese polvo te hace invisible. Con medio diente y agua de luna se puede viajar a cualquier sitio en un segundo... y muchos otros hechizos. Tras decir esto, sacó del bolso un puñado de polvo y lo sopló sobre la llave y el cofre, que se volvieron diminutos. - Entonces, ¿trato hecho? - dijo, ofreciendo al ratón la pequeñísima llave y el diminuto cofre. Pérez los tomó y chocó su mano. - ¡Trato hecho! - dijo el ratón. Pina y Pérez se quedaron charlando hasta el anochecer. Entonces, el ratón frotó la llave, tal como le indicó el hada, y desapareció del bosque. Esa noche entró en las casas donde a algún pequeño se le había caído un diente. Al despertar, el dinete había desaparecido y en su lugar había un regalo. La noticia corrió como la pólvora por todas las casas. Todos los niños y niñas del mundo saben ya quién es el ratoncito Pérez y le dejan sus dientes de leche bajo la almohada para que pase con su cofre mágico. Además, cuando ven un ratón, no se asustan, podría ser él.

#### **A.1.12. La bella durmiente**

Cuando nació la princesa Aurora, las hadas buenas le hicieron maravillosos regalos. Pero el regalo de la malvada Maléfica fue una maldición: al cumplir dieciséis años, la princesa se pincharía en una rueca. Las hadas se llevaron a Aurora a vivir al bosque. Querían protegerla. La maldición de Maléfica era muy fuerte. Hizo que Aurora se pinchara el dedo y se durmiera. El príncipe Felipe rompió el hechizo y despertó a la Bella Durmiente.

#### **A.1.13. La bella y la bestia**

Bella vivía en una pequeña aldea. Le encantaba leer. Bella quería mucho a su padre. Cuando desapareció, Bella salió a buscarlo. Bella encontró a su padre en el castillo de la Bestia. Bella y Bestia aprendieron a ser amigos. Cuando Bella se enamoró de Bestia, él volvió a convertirse en príncipe.

#### **A.1.14. La margarita**

Había una vez una casa de campo con un jardín lleno de flores y una cerca pintada. En medio del bello y verde césped crecía una pequeña margarita. El sol era igual de generoso con la pequeña margarita que con las grandes y suntuosas flores del jardín. Y la margarita crecía, de hora en hora. Allí estaba una mañana la margarita, recibiendo el calor del sol, despreocupada, sin dolerse de ser una pobre flor insignificante. Estaba contenta y, mirando al sol, escuchaba el alegre canto de la alondra, feliz como si fuera un día de fiesta. Los niños estaban en la escuela, y mientras ellos estudiaban, la margarita aprendía a conocer la bondad de Dios en el calor del sol y en la belleza de lo que la rodeaba.

#### **A.1.15. La sirenita**

Ariel era una sirenita que vivía bajo el mar. Ariel se sentía fascinada por los humanos, sobretodo por el príncipe Eric. Ariel pidió a la bruja del mar que la transformara en humana

para poder estar con el príncipe Eric. Si Eric no besaba a Ariel antes de que pasarn tres días, ella volvería a convertirse en sirena. Pero ni la magia de la bruja del mar pudo separar a Eric y Ariel.

#### **A.1.16. Mulán**

Mulán quería llevar el honor a su familia. A veces era difícil. Un día, el padre de Mulán se enteró que tenía que ir a la guerra. Mulán fingió ser un muchacho. Se unió al ejército para que su padre no tuviera que hacerlo. Mulán se entrenó duramente. Ayudó al ejército chino a ganar la guerra. La familia de Mulán se sintió muy orgullosa de ella.

#### **A.1.17. Noticia OMS**

Expertos de OMS recomiendan declarar a Paraguay "país libre de malaria". Un grupo de expertos de la Organización Mundial de la Salud (OMS) recomendó declarar a Paraguay como "país libre de malaria" después de verificar la situación de esa enfermedad sobre el terreno durante dos semanas, informó el domingo el Ministerio de Salud Pública. La cartera sanitaria explicó a través de un comunicado que los expertos de la OMS comprobaron que en el país suramericano que "no se registran casos de transmisión local de malaria en los últimos años y que el sistema sanitario puede prevenir eventuales casos de esta enfermedad". Esta recomendación cierra el esfuerzo del sistema sanitario paraguayo en los últimos seis años, que desde 2012 no registra casos autóctonos de la enfermedad. Este gran logro nacional posibilita que Paraguay hoy esté en el privilegiado sitial de primer país sudamericano en erradicar la enfermedad", indicaron desde el Ministerio de Salud Pública. La malaria pasó a la historia de los registros médicos del país en 2011, con el último caso registrado en Paraguay, según datos del Servicio Nacional de Erradicación del Paludismo (Senepa) presentados en septiembre de 2017. El Senepa se fundó a raíz del brote de malaria que sufrió Paraguay en 1957, cuando unas 90.000 personas en el 90 % del territorio del país contrajeran la enfermedad. El paludismo o malaria es una enfermedad letal causada por parásitos del género plasmodium, que se transmiten al ser humano por la picadura de mosquitos infectados. El paludismo se mantiene todavía presente en 91 países y se calcula que en 2016 hubo 445.000 fallecidos en todo el mundo y 216 millones de casos, según datos de la OMS.

#### **A.1.18. Noticia muerte**

Un brutal caso de violencia machista paraliza uno de los centros comerciales más concurridos de la Ciudad de México. Burló todos los controles de seguridad. Subió a la segunda planta del centro comercial. Dos disparos. Uno al pecho, el otro al abdomen. Después dejó una nota en el mostrador de la tienda en la que trabajaba su esposa y se disparó en la cara. El feminicidio e intento de suicidio paralizó este lunes Reforma 222, uno de los centros comerciales más concurridos de la Ciudad de México. Selene, de 28 años, falleció horas más tarde. Alejandro, el agresor de 35, sobrevivió y se encuentra grave, pero estable en un hospital de la capital. Después de las detonaciones, los trabajadores de las tiendas y los clientes corrieron para protegerse y resguardarse. Los hechos se produjeron a las tres de la tarde, en lo que había sido una anodina tarde de puente. Las primeras versiones barajaban que se había tratado de un robo a la zapatería Nine West de la plaza comercial, que se encuentra en el Paseo de la Reforma, una de las avenidas más transitadas e icónicas de la capital. Después las autoridades dieron a conocer que había sido un homicidio. La Procuraduría (Fiscalía) local ha iniciado las investigaciones por el delito de homicidio calificado. Selene

había pedido a Alejandro el divorcio y la custodia de su hijo. El agresor había estado preso tres veces por robo en 2002, 2004 y 2012 y ya había amenazado unos días antes a Selene. "No vas tú a destruir mi vida y después como si nada. Por puta te pasó esto", se lee en la supuesta nota que dejó el feminicida y que ha sido divulgada por la prensa mexicana. "La mujer lesionada manifestó que el hombre la había amenazado a causa de un conflicto pasional", señaló la policía capitalina. En un país en el que son asesinadas siete mujeres a diario por la violencia machista, los crímenes pasionales aún son vistos como excusas y detonantes para atacar a las mujeres por el hecho de ser mujeres.

#### **A.1.19. Noticia nacimiento**

Chiara Ferragni ya es madre de su primer hijo y sabemos el nombre. La influencer todavía no se ha pronunciado en sus redes sociales pero los medios italianos confirman el nacimiento del pequeño Leone. Chiara Ferragni por fin es mamá y, aunque la influencer italiana todavía no se ha pronunciado a través de sus cotizadas redes sociales, el portal italiano Tabloit ya ha confirmado la noticia. El medio italiano ha anunciado que el hijo de Chiara con el raper Fedez llegó al mundo en la noche del 19 al 20 marzo. La italiana dio a luz en el hospital Cedars-Sinai de LA, el mismo en el que dio a luz Beyoncé, a través de un parto programado. Los 12 millones de seguidores de la influencer apuntaban en las últimas horas que Ferragni podía estar dando a luz por su desaparición en redes sociales. Al final la noticia se confirmaba horas después, apodando ya en las redes sociales al pequeño como baby Raviolo, por el origen de sus padres. Pero ambos ya anunciaron en su Instagram en nombre del pequeño Leone. Según el mismo medio, el parto se habría adelantado dos o tres semanas, pero tanto el bebé como la madre se encuentran en perfecto estado. Era la misma Chiara la encargada en desvelar el sexo y el nombre de su primer hijo ante todos sus seguidores, considerada por Forbes como la influencer más poderosa del mundo de la moda dándose a conocer gracias a su web The Blonde Salad.

#### **A.1.20. Noticia política**

El estancamiento político que agobia a España desde hace 10 meses podría llegar a su fin pues se espera que el parlamento apruebe un gobierno liderado por el presidente interino Mariano Rajoy. Rajoy necesitaba una mayoría en el parlamento de 350 escaños durante la primera votación realizada el jueves, pero no la logró. El sábado se realiza una segunda votación en que sólo necesita más votos a favor que en contra. Cuenta con el apoyo de 170 legisladores, de los cuales 137 son de su propio partido, el centroderechista Partido Popular. Todo dependerá del Partido Socialista. Ha accedido a abstenerse en la votación del sábado aunque 11 diputados han dicho que votarán en contra. De ser elegido, Rajoy será juramentado por el rey Felipe VI el domingo. Si para el lunes no hay gobierno, se convocará a una tercera elección.

#### **A.1.21. Pinocho**

Había una vez un artesano que labraba en madera juguetes tan bellos que todos lo tenían por un escultor y un maestro. Maese Geppetto, que así se llamaba, era muy anciano, y entre sus muchas virtudes no figuraba el buen carácter. Por eso, había envejecido solo. Un día, sinténdose más viejo y más solo que de costumbre, para ahuyentar la soledad decidió tallar en madera un muñeco que le hiciera compañía. Eran tantos los que había construido para transformar el aburrimiento de los niños en momentos maravillosos. Ahora se fabricaría uno para él. En el taller de un colega carpintero consiguió un tronco de pino y se dispuso

a fabricar el muñeco. Corta un poco aquí, talla un poco allá y he aquí la cabeza, la nariz y la boca... En cuanto hubo terminado de modelar la boca de madera, el muñeco rompió a hablar, exclamando: Quién eres, eres mi abuelito. Superado el asombro, maese Gepetto se apresuró a completar su obra y, como había tallado al muñeco en un tronco de pino, pensó llamarlo Pino. Pero después de reparar en su traviesa mirada, decidió que le iría mejor el nombre de Pinocho. Y en cuanto Pinocho tuvo un par de piernas de madera, se incorporó de un salto, cogió impulso y se marchó a todo correr.

### A.1.22. Pocahontas

Pocahontas era una indígena de Norteamérica. Amaba la naturaleza. Un día, un barco de Inglaterra llegó a Norteamérica. Algunos ingleses eran codiciosos y malvados, pero John Smith era bondadoso y amable. La tribu de Pocahontas quería que los ingleses se fueran. Los ingleses se fueron en su barco, pero Pocahontas y John Smith nunca se olvidarían uno de otro.

### A.1.23. Tipos

¿Has ido al parque esta mañana? Sí, he ido. ¿Sabes a quién he visto? A Luis. ¡Vaya sorpresa! No sabía que estaría ahí. ¿Verdad? ¡Me he quedado flipando! Llevaba mucho sin verle. - Buenos días, ¿qué tal? - me dijo-. ¡Estas genial! - ¡Bien!

## A.2. Textos añadidos en la evaluación del nuevo EmoTraductor

### A.2.1. Centro comercial

El Centro Comercial había quedado en penumbra, en el límite mismo entre universos. Tanto al este como al oeste seguían alternándose el día y la noche, pero por sus ventanas sólo se filtraba al interior la luz sucia que nacía de su mezcla: un eterno crepúsculo, una promesa de amanecer que nunca terminaba de fructificar. Por puro azar, su planta se alineaba de forma tal que la Grieta lo partía por la mitad. No había ninguna hendidura visible, claro está, pero resultaba patente por la disparidad de los establecimientos que podían encontrarse a cada lado.

El Centro Comercial estaba siempre abarrotado. Las tiendas trabajaban de forma ininterrumpida, mezclándose los trabajadores del turno día con los de noche, pero cada uno en su parte; las tímidas pruebas efectuadas para contratar empleados del universo opuesto habían acabado en locura.

Los respectivos consejos de dirección se habían puesto de acuerdo para cobrar y limitar la entrada. En un momento dado sólo se admitían veinte mil clientes, que se agolpaban asombrados ante los escaparates, adquiriendo algún producto cuya utilidad desconocían o, si eran osados de verdad, escabulléndose en las salas de cine para echar un fascinado y asqueado vistazo a lo que aguardaba agazapado más allá de la penumbra. Sin embargo, en ocasiones, el visitante sentía despertar algo distinto en su interior, como un recuerdo recién formado, un anhelo antinatural. Entonces salía del Centro Comercial por la puerta opuesta a aquella por la que había accedido y desaparecía para siempre del mundo que lo había visto nacer.

Quizá la Grieta sirva para eso, para corregir algún trágico error cósmico y devolver a casa a los hijos pródigos. ¿Quién podría asegurarlo? Lo único cierto es que está allí, en la

Penumbra, y cumple una función, que tal vez sea revelada cuando, por fin, despunte el día.

### **A.2.2. Conversación1**

– Disculpe, ese es mi asiento. – ¿Está seguro? – Sí, mi entrada dice fila seis, asiento doce. Es ese mismo. – Disculpe, había visto mal mi entrada. Mi asiento es el dos. Ya le dejo su asiento. – Muchas gracias. – De nada.

### **A.2.3. Conversación2**

– Buen día. – Buen día. ¿En qué puedo ayudarla? – Necesito dos kilos de pan, por favor. – Dos kilos de pan. Aquí están. ¿Algo más? – Nada más. ¿Cuánto le debo? – Treinta pesos. – Aquí tiene. – Muchas gracias. Buenas tardes. – Buenas tardes.

### **A.2.4. DobleNegación1**

A mí no me parece mal ese plan y a mi hermana tampoco le disgusta.

### **A.2.5. DobleNegación2**

No he visto en mi vida un paisaje tan hermoso

### **A.2.6. DobleNegación3**

No comeré jamás hígado de cerdo; ¡qué asco!

### **A.2.7. FraseCorta**

Con los soles de finales de marzo mamá se animó a bajar de los altillos las maletas con ropa de verano. Sacó camisetas, shorts, sandalias. . . , y aferrado a su cubo y su pala, también sacó a mi hermano pequeño, Jaime, que se nos había olvidado.

### **A.2.8. FraseCorta2**

Todas las mañanas llego a la oficina, me siento, enciendo la lámpara, abro el portafolios y, antes de comenzar la tarea diaria, escribo una línea en la larga carta donde, desde hace catorce años, explico minuciosamente las razones de mi suicidio.

### **A.2.9. Negación**

A Cristina le encanta pasear, pero no le gusta la lluvia. Cuando no hace sol, no desiste y sale de todos modos con su perro. Como no quiere mojarse, no olvida nunca coger un paraguas. A su perro tampoco le gusta el agua, por eso le pone un sombrero.

### **A.2.10. Negación2**

No nos ha informado las especificaciones técnicas de lo que necesita; por lo tanto, no podemos enviarle los productos que ha solicitado.

### A.2.11. Negaciones

No le digas a nadie nada de lo que no le dije a Marcos el martes cuando al final no vimos la película.

### A.2.12. Negaciones coordinadas

No, no puedo olvidar esa tarde tu cara como si me dejaras. Pero, adivino que es así como va a ser la historia. Siempre sonríes pero en los ojos se te nota el dolor, sí se te nota. No, no puedo olvidar el mañana, cuando piense en todas mis penas, en cuando te tuve allí pero te dejé ir.

### A.2.13. Noticia deportes

El Real Madrid se impuso en un Clásico vibrante y recupera el liderato de la Liga. Los blancos derrotaron al Barcelona en una gran segunda mitad, en la que los goles de Vinicius Jr. en el 71' y Mariano en el 92' sentenciaron el encuentro. En la primera parte hubo ocasiones para ambos equipos. El Real Madrid fue el primero en intentarlo en el minuto 6, pero la volea de Benzema dentro del área se marchó alta. La réplica la puso Griezmann, que también remató elevado en el 21'. La presión alta de los blancos hacía daño a los visitantes, que trataban de contrarrestarlo con posesiones muy largas en busca del espacio. Lo encontró Arthur en el 34', pero Courtois estuvo más acertado en el mano a mano. La última jugada de peligro de los primeros 45 minutos llegó después. Pase de Busquets a la espalda de la defensa y Messi se anticipó a Varane por centímetros. Courtois desvió el potente disparo. La segunda mitad fue del Real Madrid, que en el 56' gozó de su ocasión más clara hasta ese momento. Isco recibió escorado en la frontal, levantó la cabeza y se inventó un bonito disparo de rosca que obligó a lucirse a Ter Stegen con una gran estirada. Eran los mejores minutos de los locales, que encerraban a su rival y empezaban a sumar jugadas de peligro. Centró Carvajal algo escorado en el 61' y Piqué salvó a su equipo sobre la línea de gol tras el testarazo de Isco, que ya había superado a Ter Stegen. Tres minutos después recuperó el esférico el lateral derecho madridista y volvió a centrar. La volea de Benzema dentro del área no encontró portería por muy poco. Transcurrían los minutos y el Real Madrid estaba más cerca del gol. Antes llegó el susto. Braithwaite, que acababa de entrar al campo, se fue por banda derecha en el 70', pero no pudo superar a Courtois en el mano a mano. Un minuto después llegó el primer tanto madridista. Kroos marcó el desmarque a Vinicius Jr. y le asistió entre líneas. El brasileño se internó en el área y su disparo tocó lo justo en Piqué para superar a Ter Stegen. El dominio local tenía premio y de ahí hasta el final apenas se vieron ocasiones hasta que salió Mariano al campo. El delantero arrancó en el minuto 92 y se plantó ante Ter Stegen, al que batió en la definición. El Real Madrid recupera el liderato.

### A.2.14. Peli aburrida

La película era muy aburrida, tan aburrida que me quedé dormido rápidamente y profundamente. Al despertarme me dio rabia la gran cantidad de dinero que me había gastado en la entrada de cine. Una vez más, torpemente, me volví a equivocar rotundamente.

### A.2.15. Preguntas tristes (oración)

¿Dónde lloran los océanos, si ya ellos están llenos de todas las gotas del infinito? Ya no le caben más gotas, está lleno, y se queja, y llama al viento y a las nubes para que no

descarguen más. ¿Dónde van esos vientos tempestuosos llenos de ira, si ya están a tope en su cabida terrenal? Ellos les dicen al mundo que no soplen más, que ya no tiene cabida. ¿Dónde van esos fuegos tronidos de arriba, cuando rugen como barabunta de fuegos sobre la tierra? Ellos les dicen al cosmos, que no los ennegrezcan más porque ya no pueden tronar más fuerte. ¿Dónde van esos seres que ya descansan en sus moradas, en la tierra que los vio nacer, esa tierra, que ya no puede más y le dice al Universo NO me mates más, que ya no me queda terreno en mi seno para acoger a tanta alma? ¿Dónde van tantas almas?, al cielo y el cielo dice, Si aquí cabéis todos. A ti Creador, te hablo, por favor, súplicas tienes a montones a miles de millones, ten piedad de nosotros y danos una oportunidad para vivir. Ayúdanos a todos.

### A.2.16. Texto1

Con una persistencia propia de la tozudez el gobierno boliviano tiende a negar la realidad y los hechos convirtiéndolos en el lado opuesto a la verdad. La economía nacional muestra claros síntomas de apertura hacia una crisis de dimensiones muy costosas en términos sociales. Un gobierno que se reclama “de los pobres” tendría que asumir estos síntomas con la seriedad debida, pero lejos de ello nos cuenta, ahora, historias que solo la imaginación afebrada puede generar. Ante las interrogantes que periodistas con mucho esfuerzo hacen a las autoridades del sector económico, las respuestas no son solo evasivas sino abusivas, pues tienden a tratar al ciudadano como a un tonto que poco sabe de lo que se habla. La pregunta de cómo afectará la caída de los ingresos por las exportaciones de gas, y de su impacto en los presupuestos de las gobernaciones y alcaldías, es absolutamente válida, cuando al fragor de la campaña política se habla de estar camino a ser una potencia económica en la región. El gobierno comienza afirmando que la venta de menores volúmenes de gas a Brasil y Argentina, se debe a que ambas naciones están ejecutando políticas neoliberales y que el impacto de estas ocasiona una menor demanda de gas. La causa es la demanda no la oferta. Esta aseveración, induce a pensar que la oferta existe y que menores volúmenes al calor de los contratos firmados en la década de los noventa, tendrán como efecto sanciones que Bolivia aplicará a sus clientes. La realidad es distinta, tanto Brasil como Argentina saben que el gas boliviano está en declive y que sus reservas ya no garantizan una oferta permanente en los próximos 15 a 20 años. Por ello tomaron sus previsiones y ambos han demandado menores volúmenes a precios menores. La caída de los ingresos provenientes del gas entre 2014 y 2018 es de 6.500 millones a 1.500 millones por efecto precios en los mercados internacionales, este impacto tiene que afectar los ingresos fiscales del IDH destinado al gobierno central y sub gobiernos nacionales. Para mediatizar esto el gobierno se esfuerza en convencernos de que es hora que tanto gobernaciones como alcaldías busquen otras fuentes de ingreso. Notable manera de torcer las cosas. Durante trece años el gobierno central tuvo en sus manos más de 60 mil millones de dólares para reconvertir la economía nacional y sacarla de su condición mono exportadora de materias primas. Fracasó. Y ahora pretender exigir a los demás lo que ellos no pudieron hacer. Rehúye responsabilidades, y así como culpa a Brasil y Argentina de sus políticas que ocasionan menor demanda de gas, culpa a los gobiernos sub nacionales de no haber diversificado sus ingresos. Esta forma de torcer las cosas, de construir un relato de negación permanente sobre sus responsabilidades en el gobierno es la mayor muestra de la incapacidad para enfrentar los problemas y darles soluciones adecuadas. Niegan la caída en los volúmenes de gas como niegan la ausencia de nuevos descubrimientos gasíferos, niegan haber ideologizado el uso de los hidrocarburos como propaganda política y varita mágica del derroche, con la pérdida de mercados y proyectos de plantas hidrocarburíferas sin estudios de factibilidad y que ahora son monumentos a la ineficiencia. La negación es

la manera de deslindar responsabilidades, como aquella que hizo el Presidente de afirmar que tuvo un hijo, que se murió y que ahora dice ser una mentira fabricada por otros. La verdad es para el gobierno una impostura.

### A.2.17. Texto2

Nunca fui el mismo una vez llegué de la selva amazónica. Mi vida solía ser algo monótona, pero ese viaje cambió mi vida. Pensé que lo más emocionante sería ver los paisajes o los efectos de algún alucinógeno, pero ese encuentro me reveló que aún hay mucho por explorar, por recorrer, por descubrir. Estuve al borde de la locura, casi como si una fuerte corriente marina cortara con el sentido de un cardumen. Fue toda una travesía y por poco pierdo la vida. Sin embargo, gracias a Dios estoy acá hoy para compartir la historia. Eran apenas las 5:00 a.m. cuando nos despertamos en aquel extraño y largo día. El cielo se preciaba sobre nosotros. El olor matutino de las tostadas era engañoso. Él no tenía una tostadora. Corrió las cortinas y la luz del día nunca entró. Hacía frío, de ese que te cala en los huesos. La luz se negaba a entrar donde estábamos, eso nos pareció extraño, pero excitante. Terminamos de desayunar unas arepas viejas, había que estar bien alimentado, nos pusimos unos impermeables y salimos. La jornada apenas empezaba. El ambiente que nos rodeaba era tenso, pero lindo. Un chamán me dio una planta, dijo que tenía propiedades curativas, que me revelaría cosas sobre mi subconsciente. Seguramente la búsqueda del último T-Rex de la montaña sería tarea difícil. De pronto recordé la excitante y brutal noche de ayer con aquellas mujeres tan liberales, sonreí y observe la lluvia. Encontré la misma calle que me vio crecer y con mi lente de fantasía puede captar cosas que jamás había visto. ¿Era acaso Keiko, la primogénita, una alucinación amazónica? ¡Era ella! Y a pesar que el cielo se desvanecía sobre nosotros, su presencia nos iluminaba el camino y me daba confort. Era una mujer muy hermosa y la oscuridad le quedaba tan bien, que formaban un todo y un nada a la vez. Tenía largas piernas y en sus caderas se reflejaba la poca luz que había en el cuarto. Estábamos en un cuarto grande y oscuro. Hacía frío, a pesar de la hora. Y mientras eran peras o manzanas, no dejaba pensar en la expectativa que me había creado alrededor de esa inmensa nada. De golpe recordé las páginas de Delirio, de Laura Restrepo. Todo empezaba a cobrar sentido. Ya como que mis pensamientos no tenían sentido ni orden. Pero volví y cuando lo hice... era otro ser. Empecé a creer que mi vida estaba de vuelta, mi indiferencia al entorno no existía más, me prometí repetirlo otra vez. Cuando todo terminó, todo se puso en silencio y ya no me acordaba por qué había ido. Y cuando despertó el fujimorismo seguía allí.

### A.2.18. Texto miedo a perder el amor

Tengo miedo de perderte del todo. Tengo miedo de que acaricies a otra como lo hacías conmigo, Que compartáis películas, que la mires como me mirabas a mi. Que la escribas y le dediques canciones. Te tengo tan cerca y a la vez tan lejos. Te quiero tanto y quisiera tanto darte todo este amor que me está matando.

Estoy perdiendo a mi único pedazo, el único que faltaba. Echo de menos tu olor, tu piel, tu casa, tu sonrisa, tus ojos puros. Te echo de menos. Aquí encerrada. Sin ti.

Me estoy volviendo loca. Hace tiempo que no duermo como debería, y tampoco me cuido como debería. Me estoy dando igual, te has llevado el último pedazo. No quiero ser una persona puente más. Tengo miedo. Mucho miedo. Te quiero. Te quiero. Te quiero.

Lo único que me queda de ti, es un trozo de papel impregnado aún por tu puño y letra, aún con tu fragancia. No me cansaré nunca de ella. No me olvidaré de la caja roja, ni de tus palabras, ni de tu ayuda. Jamás olvidaré el amor tan crudo, tan puro, tan rudo que

me diste. Me sangra el cuerpo y creo morir. Me siento exhausta, ya perdí la batalla. Sigo. Sigo. Sigo. Estoy bien. No pasa nada. Sigo. Sigo. Sigo. Me doy la vuelta, ya te has ido. Me doy la vuelta, está todo destruido.

Dónde ir. Dónde mirar. Si a cada paso que doy, estás tú para recordar. Quiero tenerte entre mis brazos, no soltarte jamás. No perdí la fe contigo, La perdí conmigo. Me rendí yo. Toda yo. Me di por vencida. Qué difícil es avanzar con pedazos que sientes que no son tuyos. Qué difícil es luchar por algo que ya no tiene sentido.

Abrázame y no me sueltes jamás. Me caigo, me voy a estampar y acelero, acelero, acelero, cierro los ojos y dejo de respirar. Da igual, de verdad que todo da igual.

Tengo miedo. Mucho miedo.

#### **A.2.19. Tristeza**

No hay sol, solo llueve, golpes en el suelo, marcas convertidas en simbolos de lo que alguna vez sucedió... Distancia, marcada a pasos. Los sueños no son sueños, porque no he despertado y a esto, aún, se le llama realidad. No hay nadie, solo viento. -¿Y la calma? Notaba ahí fuera el temporal que habitaba en mi cabeza hace tiempo, eso me entristeció; no había salido para hacerme más sufrir.

#### **A.2.20. Tristeza2**

La verdad, había sido un mal sueño, pero cuando desperté y vi que era verdad; volví en mí. El silencio hacia catacumbas las paredes y yo estaba dentro. Soñé con volver a dormirme, pero con la cabeza rota de pensamientos no podía, tenía miedo de tener la más miserable pesadilla. Acaso:¿no era ya, una pesadilla mi vida? Me sentí rota y miré en silencio al vacío que se reflejaba en el paisaje invernal y allí esperaba que pasase el tiempo y todo volviera a la normalidad. Incluso yo.