Annotating Expressions of Engagement

in online book reviews:

A contrastive (English-Spanish) corpus study for

computational processing

MASTER in English Linguistics, MASTER DISSERTATION

Faculty of English Philology I, UCM, Spain

Tutor: Dr. Julia Lavid

Student: Natalia Mora

September 2011

Resumen en castellano

El estudio del lenguaje evaluativo en lengua inglesa cuenta ya con importantes contribuciones desde hace más de una década (Chafe y Nichols, 1986; Biber and Finnegan, 1989; Ochs and Schiefflen, 1989; Bybee and Fleischman, 1995; Niemeier & Dirven 1997; Wierzbicka, 1990; Conrad and Biber, 2000; Hunston, 2000; Hunston and Thompson, 2000 *inter alia*). En la escuela de Lingüística Sistémico-Funcional, los investigadores Martin y White propusieron en 2005 un modelo para el estudio de lo que denominaron 'Appraisal' (Valoración, en traducción propia), que engloba la gran parte de los aspectos que se habían estado discutiendo bajo distintos nombres y acercamientos. Este modelo se centra en "los recursos lingüísticos utilizados para negociar emociones, juicios y valoraciones, junto con los recursos que sirven para amplificar y comprometerse con estas evaluaciones" (ibid: 145, traducción propia) y se pueden distinguir tres ejes: 'Attitude', 'Graduation' y 'Engagement' (Actitud, Graduación e Intervención, en traducción propia). La primera se considera el eje central, en tanto trata las expresiones de sentimientos, juicios y valoraciones. La segunda se utiliza para modificar la intensidad o especificar el nivel de prototipicalidad de un elemento, y puede afectar a cualquiera de los otros dos aspectos. Por último, la Intervención, que es el subsistema analizado en el presente estudio, se centra en la presencia en los textos de otras voces y puntos de vista distintos a la del autor (Heteroglosia) en oposición a los textos que carecen de ellos (Monoglosia). Con el propósito de hacer una adaptación más específica al género analizado en este estudio, este modelo se ha modificado recientemente por investigadoras como Carretero and Taboada (submitted). Como afirma Thompson (2003: 6) con razón, dada la alta frecuencia de uso de estas expresiones en el lenguaje cotidiano, su análisis puede mostrar importantes conclusiones sobre la sociedad. Es por esto que una gran cantidad

de literatura se ha escrito sobre el tema, ya que puede aplicarse a muchos y variados campos, como son la adquisición del lenguaje (Painter, 2003), el lenguaje periodístico (White, 1998), o los textos literarios (Love, 2006). Sin embargo, la mayor parte los trabajos mencionados se han centrado en el análisis de la Actitud, mientras que la Graduación y la Intervención han recibido menos atención, especialmente desde el punto de vista contrastivo. Aunque algunos trabajos han tratado la Intervención dentro del marco del modelo sistémico de la Valoración (White, 1998, 2003; Miller, 2004; Brooke et al, 2009; Carretero and Taboada 2010a, 2010b, *inter alia*), todavía se necesita una investigación más extensa y profunda que pueda proporcionar nuevas conclusiones sobre su uso, realizaciones, límites y aspectos contrastivos entre lenguas.

Si pensamos en la importancia y la ubicuidad de las nuevas tecnologías en la sociedad de la información en la que vivimos, no es de extrañar que el análisis de aspectos relacionados con la expresión de la opinión en Internet —incluyendo la Valoración y los subtipos antes mencionados— cobre cada vez más importancia desde campos como la política, la publicidad y para el público en general. En el ámbito del Procesamiento del Lenguaje Natural el estudio del lenguaje evaluativo y la expresión de la opinión cuenta ya con cierta tradición en lo que se conoce como el análisis de la subjetividad (Wiebe et al, 2004; Wiebe et al, 2005; Wilson and Wiebe, 2005; Somasundaran et al, 2008) o el llamado análisis de sentimiento (Pang et al, 2002). En este sentido, investigadores como Taboada y Grieve (2004) se han centrado en el estudio y procesamiento computacional del sistema de Actitud dentro del modelo de la Valoración.

Este tipo de estudios computacionales requiere el desarrollo de bases de datos textuales (corpora) que hayan sido objeto de anotaciones manuales fiables y consistentes, de forma que puedan ser luego procesados computacionalmente por

algoritmos de aprendizaje automático. Hasta la fecha no existen textos anotados con categorías de Valoración en lengua inglesa y española, con la excepción de los textos anotados por investigadores dentro del proyecto CONTRANOT.[1] Mientras estos recursos no estén disponibles, los estudios sólo pueden basarse en muestras limitadas de géneros diversos de los que no se pueden obtener resultados consistentes.

En consecuencia, es necesario desarrollar metodologías más avanzadas como las desarrolladas en la comunidad de Procesamiento del Lenguaje Natural, con el propósito de conseguir fiabilidad y rigor en el proceso de anotación. Para ello se necesitan medidas de acuerdo interno y entre anotadores que garanticen la calidad de los datos para que sean procesados posteriormente por algoritmos de aprendizaje automático.

Con este propósito, el presente estudio se plantea con tres objetivos principales:
1) comprobar empíricamente dos aspectos centrales de la Intervención mediante dos estudios de fiabilidad basados en las anotaciones de dos anotadores humanos;
2) extraer aspectos contrastivos relevantes sobre el uso de la Intervención en inglés y en español en reseñas de libros en Internet;
3) crear una base de datos de muestras anotadas de forma contrastiva en inglés y en español con diferentes tipos de Intervención. La idea es que estas muestras puedan ser luego procesadas computacionalmente y estén disponibles para la comunidad científica.

Para la consecución de estos objetivos se ha utilizado una metodología inspirada en los trabajos sobre anotación de corpus realizados en el campo del Procesamiento del Lenguaje Natural. Esta metodología consta de dos fases: en una primer fase, se han realizado dos estudios de fiabilidad ('Agreement studies') con el fin de evaluar el grado

---

[1] El proyecto CONTRANOT se está desarrollando en la Universidad Complutense de Madrid bajo la dirección de la Dra. Julia Lavid, y está financiado por el Ministerio de Ciencia e Innovación. Investigadores del proyecto, como M. Taboada y M. Carretero están trabajando en la creación de esquemas de anotación de categorías de Valoración en textos comparables en inglés y español. Véanse los artículos de Taboada (2009) y Carretero & Taboada (2009, 2010a, 2011).

de acuerdo entre dos anotadores independientes en dos aspectos altamente relevantes para el análisis lingüístco de la categoría de la Intervención en inglés y español. Uno de estos aspectos es la identificación del segmento lingüístico que expresa el concepto de Intervención, tanto en inglés como en español. El segundo aspecto es la clasificación del segmento identificado por los dos anotadores como perteneciente a un tipo o a otro dentro de los posibles subtipos de Intervención estudiados.

Una vez comprobada empíricamente la fiabilidad de las anotaciones realizadas sobre una pequeña muestra inicial, tanto en cuanto a la identificación del segmento como del subtipo seleccionado, la segunda fase de la metodología se centra en la anotación de una muestra de mayor tamaño (28 reseñas de libros, 14 en inglés y 14 en español) por un único anotador, en este caso, la autora de este trabajo.

Los resultados de los dos estudios de fiabilidad realizados mostraron que se producen desacuerdos recurrentes en la longitud de la expresión anotada, así como en la anotación de las categorías de Declaración (Pronounce), perteneciente a Proclamación (Proclaim), y Oposición (Counter), dentro del sistema de Negación (Disclaim), que estaban más abiertas a la interpretación, así como al solapamiento con otras categorías, debido a su alto nivel de dependencia del contexto.

La anotación de la muestra mayor (28 reseñas) demostró la existencia de similitudes formales entre inglés y español en cuanto a las expresiones utilizadas, dándose las diferencias más significativas en la frecuencia de uso de diferentes tipos de Intervención.

La interpretación de los datos también mostró que se encontraron ciertas limitaciones, principalmente producidas por la extensión del corpus. Por tanto, futuras líneas de investigación podrían proporcionar una muestra mayor, ya que es muy probable que ayudase a mejorar los resultados significativamente. Sin embargo, no sólo

sería conveniente una mejora cuantitativa, sino que el análisis específico de cada subtipo de Intervención y de los de éstos a su vez, así como de las expresiones concretas más conflictivas, podrían ayudar a solventar problemas en la interpretación y anotación de tales unidades y categorías, como pueden ser esclarecer patrones fijos y los límites que separan un tipo de otro. Por último, se ha sugerido que podría ser significativa la influencia de colocaciones y del registro y el género de los textos sobre las expresiones, por lo que el estudio de estos factores podría también ayudar en el análisis de la Intervención.

Abstract

This dissertation studies the expression of Engagement and alternative points of view in English and Spanish online book reviews, following the Appraisal model designed by Martin and White (2005). The study has three main aims: 1) to test two main aspects of the linguistic category of Engagement empirically, namely the identification of span realising Engagement and the classification of Engagement into different subtypes; 2) to extract relevant contrastive features of the use of Engagement in English and Spanish in online book reviews; 3) to create a bilingual (comparable) machine-readable annotated corpus with Engagement features in English and Spanish which can serve as the training corpus for machine learning algorithms and be offered to the scientific community for further research. Following standard methodologies in the field of Natural Language Processing, two agreement studies are carried out, designed to measure inter-annotator agreement based on an initial set of 10 reviews. A larger set of 28 reviews (14 English, 14 Spanish) is further annotated by one single human coder in order to extract relevant results on contrastive aspects and provide publicly-available machine-readable annotated texts with Engagement categories. The findings reveal disagreement mainly on span length and the annotation of some specific categories, namely Pronounce and Counter. In addition, differences regarding frequency in the use of Engagement types were found in both languages, although the expressions employed were formally similar. Finally, the results of the annotation of the larger data set showed that more expressions than what was initially expected can be annotated context-independently, although regarding some other expressions, register and collocations were seen to have a decisive influence on their interpretation of some expressions, in the same way that genre has on their frequency of use, for resources aimed at emphasising reviewer's personal opinion were more frequent than those who acknowledged and evaluated external sources.

Table of Contents

1. Introduction

Under the scope of Systemic Functional Linguistics (hereafter SFL), the theory of Appraisal, mainly developed by Martin (2000) and White (2002, 2003a), has brought a comprehensive way of approaching evaluative language. It follows Halliday's (Halliday, 1985/2004) distinction of ideational, textual and interpersonal metafunctions, and focuses on interpersonal meaning. Studies on this aspect have traditionally been oriented to interaction (Halliday 1984, Martin 1992, Eggins and Slade 1997); however, Appraisal pays special attention to feeling, as it emerged as the analysis of affect and evaluation in texts such as narrative (Martin 1996, 1997), administrative and history discourse (Iedema 1995; Martin 2001), as well as the print media (Iedema, Feez and White 1994).

According to Martin (2000: 145), Appraisal refers to "the linguistic resources used to negotiate emotions, judgements and valuations, alongside resources for amplifying and engaging with these evaluations". These resources are used by everybody on an everyday basis, so they are constantly present in our discourse. Thompson (2003: 6) rightly claims that the analysis of evaluation is of paramount importance because it does not only express an opinion, but it also maintains relations between writer and reader and even organizes discourse. In addition, he argues that opinionated texts express communal value-systems, which are a part of the ideology of the society in which the text has been produced. Thus, the analysis and identification of what the writer thinks may reveal the ideology of the society.

The Appraisal framework, as outlined in Martin and White (2005), proposes that these linguistic expressions of evaluative meanings such as emotion, attitude and opinion can be divided into three different axes: Attitude, Engagement and Graduation. Attitude is the term used for emotional, ethic and aesthetic opinions, which conform to the categories of Affect, Judgement and Appreciation respectively. Engagement is

concerned with the existence and alignment (heterogloss) or absence (monogloss) of alternative viewpoints with respect to an expressed opinion. Finally, Graduation deals with the ways in which the strength of opinions is intensified or weakened. This can be done according to intensity or amount, or according to prototypicality, that is, the degree to which something matches a given category.

As a result of the great impact of this theory, much research is done in order to prove the validity and expand the categories proposed so far. However, although considerable research has been devoted to Attitude (e.g. Macken-Horarik, 2003; Painter, 2003; Page, 2003; White, 2004), less attention has been paid to Engagement and Graduation, which require further study.

The interest of politicians, marketers and the public in general in opinions in the Web has led to a growing literature and research on the automatic identification and extraction of this type of units using computational tools and techniques. The Natural Language Processing (NLP) scientific community is currently investigating the accurate automatic analysis of these aspects, due to the significant improvement it would bring to fields such as subjectivity (Wiebe et al., 2004) and sentiment analysis (Pang et al., 2002) for computational applications like, for example, opinion mining. Such automatic analysis requires, however, the existence of publicly-available machine-readable annotated texts with Appraisal features to be used as test data for the development of machine learning algorithms. At present, there are no machine-readable Appraisal-annotated texts publicly available, though some research efforts within the CONTRANOT project[2] are making progress in this direction (Taboada, 2009; Carretero and Taboada, 2009, 2010a, 2011).

---

[2] The CONTRANOT project is financed by the Spanish Ministry of Science and Innovation under the I+D Research Projects Programme (reference number FFI2008-03384), with Prof. Julia Lavid as principal investigator. It aims at the creation and validation of contrastive functional description through corpus analysis and annotation.

In the Linguistics community, studies on Appraisal are limited to example extracts which come from a wide variety of genres such as those aforementioned (i.e. literary texts, new reports, etc.) and, even though they are useful for the purpose of demonstrating the theory, they can only be employed in a qualitative analysis and would bring about inconsistencies if analysed collectively. Therefore, it is necessary to develop more rigorous methodologies, such as those developed in the NLP community, to ensure the consistency and reliability of the analysis. Current work in NLP on corpus annotation shows the need to use inter- and intra-annotator agreement measures to ensure the quality of the data for further processing, and recent interdisciplinary work —such as the work within the CONTRANOT project— shows the potential of human-coded annotation to test aspects of linguistic theories/categories empirically (see Hovy and Lavid 2010).

In addition, with the exception of the work comparing Appraisal features in English and Spanish carried out within the CONTRANOT project, and other recent work on Appraisal in Japanese (Sono 2011), there are few studies comparing the use of Engagement strategies in two or more languages under the Appraisal framework[3]. Therefore, a systematic and corpus-based contrastive (English-Spanish) study on the use of Engagement resources in online book reviews will hopefull shed light both on the features of Engagement but also on the generic characterisation of these type of reviews.

2. Aims of the study

The current study has been designed with three main aims in mind:

1. To test two main aspects of the linguistic category of Engagement empirically: 1) the identification of span realising Engagement; 2) the classification Engagement into different subtypes. These two aspects were tested through the design and performance

---

[3] An remarkable exception is the articile by Marín Arrese and Núñez Perucha (2006), but it is focused on the use of Engagement strategies in English and Spanish journalistic texts.

of two agreement studies carried out through the independent annotation of a small initial data set by two independent coders.

2. To extract relevant contrastive features of the use of Engagement in English and Spanish in online book reviews. Once Engagement features were empirically and contrastively validated through the above-mentioned agreement studies, the research focused on the quantitative and qualitative analysis (by one single annotator) of a larger set of bilingual texts.

3. To create a bilingual (comparable) machine-readable annotated corpus with Engagement features in English and Spanish which can serve as the training corpus for machine learning algorithms and to be offered to the scientific community for further research.

The study is structured as follows. Section 3 offers an overview of the theoretical framework underlying the current study, namely, the Appraisal Framework within Systemic Functional Linguistics (SFL). Section 4 presents the materials (data sets) used and the methodological steps adopted to achieve the aims presented above. Section 5 explains the agreement studies carried out to test the two main aspects of the linguistic category of Engagement empirically, including the annotation procedure, the annotation scheme and guidelines and the statistical results obtained. Section 6 describes the annotation of the larger set of texts, and Section 7 reports on the results obtained in this annotation both in quantitative and in qualitative terms. Finally, Section 8 summarises the study, draws some preliminary conclusions and outlines some extensions and future work.

3. Theoretical Framework

This section provides an overview of the theoretical field and concepts in which the study is framed. The main notions explained are Appraisal and Engagement. Appraisal

is the linguistic theory in which Engagement is included along with Attitude and Graduation. These last two axes will be explained, so that the former can be fairly differentiated from those other aspects of language. Engagement is the specific system analysed in the present study. First, a comprehensive description of the subsystems of Engagement which will be used in the analysis of the corpus will be provided; secondly, a couple of issues with respect to this system will be clarified; and then, related work on the topic will be discussed.

3.1. Appraisal theory

Appraisal Theory has been mainly developed in Australia, as their most important contributors are J.R. Martin and P. White (Martin, 2000; White, 2003; Martin and White, 2005). As said above, Martin and White's model of evaluation emerged within the theoretical framework of SFL. More specifically, they describe it "as an interpersonal system at the level of discourse semantics" (Martin and White, 2005: 33). It is an interpersonal system as far as it evolves within the interpersonal metafunction (i.e. the negotiation of social relationships), and discourse semantics refers to the meaning beyond the clause. Therefore, Appraisal deals with interpersonal meaning above the clause level.

Appraisal is realised in the form of evaluative language, which is created by expressions that indicate "the subjective presence of writers/speakers in texts as they adopt stances towards both the material they present and those with whom they communicate" (Martin and White 2005: 1). As Hunston (2010: 20) points out, even though most of examples used by these authors are lexical items, not only isolated words convey these meanings[4]. Following the systemic functional tenets, appraisal

---

[4] Actually, some attempts have been made to classify the occurrence of Attitude (Affect, Judgement and Appreciation) into fixed patterns (Martin 2003, Taboada and Grieve 2004).

resources are considered as a system of their own within language, and can be divided into three subsystems: Attitude, Graduation and Engagement, presented in figure 1.
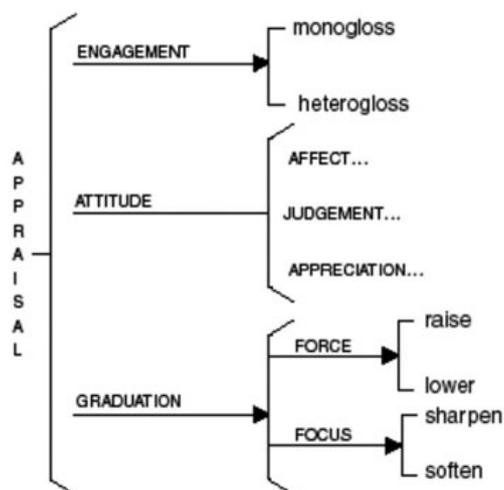

Figure 1. Appraisal System (Martin and White, 2005: 38)

Attitude is used to communicate three types of meaning: emotional (Affect), ethical (Judgement) and aesthetical (Appreciation) (Martin and White: 2005: 42). This system may be considered as the core one, as far as Graduation strengthens or weakens these meanings, and Engagement grounds attitudes and opinions. In a similar way, following Painter (2003), they (ibid) claim that, considering Affect, Judgement and Appreciation, the three subsystems of Attitude, it is the first one the most central one, inasmuch as we are born with and embody physiologically this expressive resource. Martin and Rose (2003: 24) summarize them in the following clear and brief way: Affect refers to "resources for expressing feelings", Judgment to "resources for judging character" and Appreciation to "resources for valuing the worth of things".

As has been said, Affect, in general terms, expresses feelings[5]. It conveys emotional responses such as anger, happiness, disgust, etc. about somebody's (either the

---

[5] Some further modifications have been proposed for this system. One of the most recurrent proposals is a division of Affect into two different types. Bednarek (2009) proposes the classification of Overt Affect and Covert Affect. In her own words (ibid: 166) "OVERT AFFECT directly name an emotional response

speaker or other person) reactions. Examples [1] and [2] show a couple of Affect resources.

[1]   I **like** a story that is based on facts. [no19][6]

[2]   La historia contada desde el punto de vista de un niño me **disgusta** menos … [no2_19]

Judgement provides moral evaluations about people and their behaviour. Instead of being oriented towards the individual, as Affect is, Judgement is more socially-oriented, that is, the opinions participate of social code or criteria. This is illustrated by examples [3] and [4].

[3]   A **good-natured** bunch of gringos … [no4]

[4]   … la protagonista es **patética** … [no1_14]

Finally, Appreciation consists of meanings related to the intrinsic qualities of something, when the speaker refers to their composition, their value, or a reaction they produce.

[5]   … you will discover that there are no **original** plots … [yes24]

[6]   Un libro **sorprendente** escrito por la pluma de este especial autor … [no2_23]

Graduation consists of the expressions which intensify or soften an opinion[7]. It affects both Attitude and Engagement, for they both can be up-scaled and down-scaled. The former may develop higher or lower degrees of positivity or negativity. The latter may vary in the degree of the speakers' intensity or their commitment to the message (Martin

---

of Emoters (*fear*, *love*, *hate*) whereas resources of COVERT AFFECT only indirectly denote such an emotional response", e.g. *after the **distressing** events of 1887* (example taken from Hunston (2003: 353) who creates a roughly similar division: 'reflective' and 'constitutive' Affect.)

[6] Even though Attitude and Graduation are not analysed in the present study, all the examples are taken from the same corpus, unless said otherwise. The bold emphasis used in the examples is always my own.
[7] For an exemplary illustration of how Graduation is applied in research, see Hood and Martin (2007).

and White, 2005: 135-6). Graduation is further divided into the subsystems of Force and Focus. Force refers to the specific assessment of intensity that operates over qualities, that is, the degree of evaluation is strengthened or weakened, as exemplified in [7] and [8]. Focus, on the other hand, modifies those units which are non-scalable, but with respect to their prototypicality; more plainly, expressions functioning inside Focus sharpen or soften the boundaries between categories of non-gradable items, as examples [9] and [10] illustrate.

[7]   … in a way that seems **a bit** unlike a 7 year old. [no4]

[8]   … me alegro de tener en mi pueblo una Biblioteca Municipal **medianamente** decente … [no2_19]

[9]   … it is **sort of** an indulgence for me to pay the price for a hard bound book … [no3]

[10]   … pone en **verdadero** peligro a su abuela al mandarle a su casa un desconocido. [no2_10]

Finally, together with Attitude and Graduation, the third axis is that of Engagement, which conceives language as a compound of different perspectives and divides language into *monogloss* and *heterogloss*. As this is the focus of the study, it is explained in detail in the following section.

3.2. Engagement

Engagement in Appraisal implies other voices than the speaker's one, either their presence or absence in texts. It deals with sourcing opinions and the speaker's alignment with respect to them, that is, the way in which the speaker positions him/herself with respect to these opinions as well as hypothetical responses. As Read et al (2007: 94) point out, Engagement provides the resources through which speakers construe their

point of view and adopt stances towards others' opinions. These resources include all items "by which the textual or authorial voice is positioned intersubjectively" (White, 2001: 14). Speakers have at their disposal a number of options to engage or disengage with opinions, either their own or others', which represent the types of Engagement (e.g. acknowledging, denying or countering other possibilities).

Martin and White (2005: 92-94) acknowledge some specific influences in the composition of the axis of Engagement. These are mainly three: Stubbs (1996), Bakhtin (1981) and Voloshinov (1995). Stubbs (1996: 197) assumes that "whenever speakers (or writers) say anything, they encode their point of view towards it", that is, every utterance conveys an opinion. Both Bakhtin and Voloshinov argue that verbal communication takes into account what has been said before, and anticipates what may be said afterwards. In this sense, every utterance is not only opinionated, but also 'dialogic'. However, Martin and White show that there are cases in which other voices are not referred to, such as bare assertions (e.g. *It **is written** through the eyes of a seven year old boy* [yes17]), imperative mood (e.g. ***Take** this novel for example* [no22]) and future tense with *will* and *be going to* (e.g. *this is one book that **will challenge** them all* [yes14]; *I could scarcely put the book down for fear that something **was going to happen** between the pages in my absence* [yes2]. These assertions contrast with dialogistic or heteroglossic utterances in that they do not overtly recognise dialogistic alternatives, but are presented as facts, so they are monoglossic. The distinction between monogloss and heterogloss is the first one for analysing Engagement, for all monoglossic expressions are excluded, the coding scheme being restricted to heterogloss. Figure 2 summarizes all further heteroglossic subsystems.
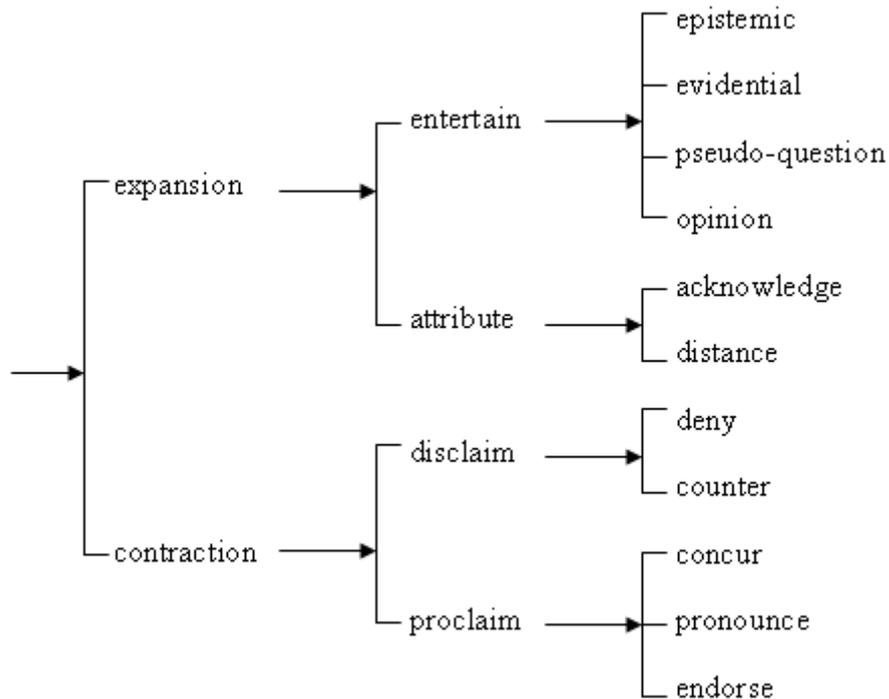
expansion
entertain
epistemic
evidential
pseudo-question
opinion
attribute
acknowledge
distance

contraction
disclaim
deny
counter
proclaim
concur
pronounce
endorse

Figure 2. The Engagement System (adapted from Martin and White, 2005: 134)

In the first place, heteroglossic utterances are subdivided according to

> the degree to which an utterance, by dint of one or more of these locutions actively makes allowances for dialogically alternative positions and voices (dialogic expansion) or alternatively acts to challenge, fend off or restrict the scope of such (dialogic contraction). (Martin & White, 2005: 102)

Thus, Expansion and Contraction stand for expressions that open or acknowledge different opinions or dialogic voices, and those that limit or challenge those alternative viewpoints.

3.2.1. Expansion

Entertain offers two ways in which the dialogical voice can be opened up to alternative viewpoints: Entertain resources, which present the author's position as one inside a range of possible options; and Attribute expressions, which provide an external source for a given opinion.

a) Entertain

Entertain is divided into four subtypes which refer to possibility and probability, evidentials and questions. Martin and White (2005) add Deontic units (those dealing with obligation and permission); however, in this category, Attitude and Engagement are interwoven due to the former's moral content. As Carretero and Taboada (to be published) point out, "these expressions could also be considered as part of the Judgement subtype of Attitude, since the expression of obligation implies that its fulfilment is morally desirable, and the expression of permission that it is morally acceptable." For this reason, Deonticity has been excluded from the present study.

Epistemic resources are associated to those traditionally thought to refer to the degree of the authorial voice's certainty about the probabilities and possibilities of something to be true. Martin and White's (2005: 104-5) reinterpretation under the heteroglossic perspective understands these elements as those that recognise the proposition as "one among a number of propositions available in the current communicative context". Examples [11] and [12] show a couple of these resources

[11] **If** I give you more than that I just **may ruin** the book for you [yes11]

[12] contado de esta forma **tal vez** no parezca demasiado atractivo, [yes5_2]

Epistemic markers have been distinguished in this paper from those labelled as Opinion, following Carretero and Taboada's (personal communication) model for analysing Engagement. Opinion differs from Epistemic in that the former deals with non-verifiable utterances. The necessity for creating this category emerges from the importance these resources have in consumer-oriented reviews, as well as they do not adapt to any of those proposed by Martin and White (2005). For example, it includes

items where the fact that a proposition belongs to the authorial voice is explicitly expressed, as in [13] and [14].

[13] I **don't believe** it would be too much for anyone to handle. [yes11]

[14] Por lo demás **para mí** la grandeza del libro reside en como cuenta, como define esos sentimientos, esos instintos, esa pasión [yes4_17]

Evidential markers deal with the type of source of evidence the speaker can make use of, as well as asserting that the speaker believes the proposition has probabilities to be true, as illustrated by [15] and [16].

[15] As I read the book, I **got the impression** of the main female character to be smart, clever, fun and sporadic. [no21]

[16] La decrepitud de ender **parece coincidir** con la decrepitud de su saga, [no1_1]

Pseudo-questions include those questions used to play as one possible view of a proposition in non-interactive texts, so, as examples [17] and [18] show, they are not restricted to those in which the addressee is positioned to provide a particular answer (rhetorical questions).

[17] However, Thomas is young and needs someone to help rule the kingdom. **Who would** that be? [yes9]

[18] me pregunto, **¿es que no se puede una** sentir mujer e igual de sexy con unas braguitas de algodón, una camiseta, unos vaqueros y unas deportivas… [yes4_17]

b) Attribution

Attribution deals with "those formulations which disassociate the proposition from the text's internal authorial voice by attributing it to some external source" (Martin and White, 2005: 111). Thus, the authorial voice is positioned and expresses its attitude with respect to an external source which is introduced in the discourse. This can be done in

two ways: merely acknowledging the external source, or distancing from it, typically by means of a reporting item.

Speakers *acknowledge* another source when their point of view is not included in the discourse, that is, speakers only take into account a different voice, without aligning with it, as exemplified in [19] and [20].

[19] **According to the New York Times**, Doubleday distributed 10,000 advance copies of the book to booksellers and the media, [no1]

[20] es **según la contraportada de mi edición** "Una fantasía erótica ajena a la represiva educación sexual contranatura todavía imperante. [yes4_17]

Speakers *distance* from others' propositions when they explicitly disagree with them, as examples [21] and [22] show. Though the typical realisation involves the verb *claim*, some other resources may be found in which it is the content of the whole proposition which provides a distancing meaning.

[21] The are too many problems that drag this book down, preventing it from being the masterpiece that **many are claiming** it to be. [no25]

[22] … acaba, para mi gusto dejándose llevar por las tonterías que le cuenta una tía loca que **dice ser** la Estatua de la libertad en persona y pone en verdadero peligro a su abuela al mandarle a su casa un desconocido. [no2_20]

3.2.2. Contraction

As said above, dialogic contraction restricts or challenges alternative positions, that is, the authorial voice acknowledges other viewpoints but does not support them. More specifically, it refers to expressions in which speakers disclaim or proclaim a proposition.

a) Disclaim

Disclaiming involves that the speaker disagrees or rejects contrary positions from other voices. In turn, it can take the form of Deny and Counter.

Deny is an overt negation of a proposition. However, not all cases in which a negative item is used are instances of denying. For example, a test to check if an expression acts as deny is question tags. When a negation triggers a question tag it can be considered as deny, as is the case of examples [23] and [24]. Deny also includes items whose meaning, though used in a positive proposition, refers to the non-performance of some action. These are typically verbs like *prevent* or *avoid* in English, and *impedir* or *evitar* in Spanish.

[23]   Mr. Brown - who had **never** before sold more than a few thousand copies of his earlier efforts … [no1]

[24]   … **no capta** mi atención. [no2_19]

Similarly to Deny, Counter involves a contrary position to one that has already been introduced. However, it plays with previous expectations about a coming proposition, which is not simply negated, but substituted by a different one. Thus, the main meanings that they convey are concession and counter-expectation. For instance, examples [25] and [26] show some of the typical expressions which convey this kind of meaning.

[25]   **However**, in other areas the book **just** falls flat on its face. [no3]

[26]   **Pero** para quienes les gusten los libros de intriga o **incluso** de novela rosa… [yes4_24]

b) Proclaim

Proclamation refers to resources used to agree with a proposition from an external source. Thus, the proposition is presented as valid and reliable. Proclamation is achieved through concurrence, pronouncement or endorsement of alternative viewpoints.

Concurring means that there is partial or complete agreement between the authorial voice and the message, or that some knowledge is shared with a majority of voices. For example, [27] and [28] show one of the most common patterns.

[27] Luke witnesses the violence, **of course**, and must not talk about it …
[no4]

[28] … y, **cómo no**, allí se encontrará con el lobo … [no2_10]

Another frequent way of proclaiming others' propositions is Pronounce. When an item can be considered as Pronounce, it means that the author emphasizes or asserts a message. The author may intervene explicitly to express that his or her opinion is firm, without referring to other voices. Examples [29] and [30] illustrate Pronounce.

[29] These are factoids, **NOT** descriptive passages! [no1]

[30] **lo cierto es** que si yo me topara con una opinión como esta, iría corriendo a buscar el libro… [no2_19]

Finally, Endorse refers to those propositions from external sources which are presented as correct, unquestionable and undeniable, thus excluding other possible options, as is the case of [31] and [32].

[31] **Eyes of the Dragon shows**, yet again, that he can jump from genre to genre as he's always done. [yes7]

[32] Eso el autor lo **define a la perfección**, [yes4_17]

3.2.3. Other categories and approaches

This theoretical background has followed the model presented in Martin and White (2005). However, it must be remarked that some variations have been introduced in heteroglossic Engagement outside this work. One of the most important ones is the subsystem of Justification (White 2003, 2011), which, along with Pronounce, would be included in a category called Reinforce, which runs parallel to Concur and Endorse. Justification refers to those formulations "which construe a particular type of consequentiality, namely those by which non-'factual' propositions (for example,

attitudinal evaluations, directives/recommendation, predictions and so on) are justified, substantiated or otherwise argued for" (White, 2003: 274). However, this paper will only take into account the specific theory presented in Martin and White (2005), which is widely accepted and established.

It should also be noted that, although some more delicate categories can be found (e.g. Pronounce could be further divided into subjective and objective), the present paper will only consider those explained in this section, as they all belong to the same level of analysis, and other distinctions could not help to provide better results, but make them more complex.

It cannot remain unsaid either that there are other traditions that have approached evaluative language. The brief and descriptive summary of this literature from a chronological perspective is provided by Martin and White (2005: 39) themselves, in the form of the following table:

Approaches to evaluation	'entity focussed'	'proposition focussed'
Chafe&Nichols 1986		evidentiality
Ochs & Schiefflen 1989	affect specifiers	affect intensifiers
Biber & Finnegan 1989	affect	evidentiality
Wierzbicka 1990	emotion	
Bybee & Fleischman 1995	evaluation	modality
Niemeier & Dirven 1997	emotion	
Conrad & Biber 2000	attitudinal stance	epistemic stance
Hunston & Thompson 2000	opinions about entities	opinions about propositions
Hunston 2000	'status' and 'value' on the 'autonomous plane'	'status' and 'value' on the 'interactive plane'

Table 1. Approaches to evaluation

Chafe and Nichols (1986) worked on evidentiality focusing on the language of the Americas. Their study discussed language from the perspective of the types of evidence the authorial voice has for making assertions. Biber and Finnegan (1989) distinguish

between epistemic stance (i.e. modalization) and attitudinal stance (i.e. appraisal).

Bybee and Fleischman (1995) study the grammar of modality and separate epistemic

(which refers to probabilities) and deontic (more related to obligation) modality from

evaluation (which deals with the perception of something as desirable or undesirable).

Together with these two works, Ochs and Schiefflen (1989) and Conrad and Biber

(2000) share their perspective on the difference between affect and modality. These two

notions can also be associated with Hunston's (2000) further distinctions between

'value' regarding world phenomena usually in a 'good' or 'bad' dichotomy (i.e. affect)

and 'value' regarding propositions evaluated by other propositions in the same text (i.e.

modality). Wierzbicka (1990) and Niemeier & Dirven (1997) work on emotion and

orient their studies towards the analysis of attitude. Hunston and Thompson (2000)

make a distinction between the target of opinions, namely, entities (which are associated

with attitude) or propositions (more oriented to epistemicity).

In conclusion, several authors have been dealing with similar notions and a

number of terms have been proposed for distinctions which reflect almost the same

phenomena. Martin and White's Appraisal model comprises, classifies and analyses

most of them, providing a useful tool for analysing evaluative language.

3.2.4. Previous and related research on Appraisal and Engagement

The popularity of the analysis of evaluative units in opinion texts has grown in the last

years. This theory has been applied to a wide range of types of texts, such as school

children's narratives (Coffin, 1997; Martin 1996), language acquisition (Painter, 2003),

news reports (White, 1998), as well as literary texts (Love, 2006). However, most of the

studies carried out focus on attitudinal expressions, leaving aside the categories of

Engagement and Graduation (e.g. Macken-Horarik, 2003; Page, 2003; White, 2004, etc.).

After the theoretical study of evaluative language, the following steps have been the implementation of appraisal analysis by means of corpus annotation and automatic extraction. On the one hand, Wiebe et al (2005) and Wilson and Wiebe (2005) provide a detailed annotation scheme for 'private states' (Quirk et al, 1985), such as opinions, sentiment, emotions, etc. Somasundaran et al (2008) introduces the concept of 'opinion frame', which are individual opinions that are linked by a shared target, in order to overcome sentence-level analysis. On the other hand, the automatic extraction of these opinion expressions (what has been called Sentiment Analysis or Opinion Mining[8]) has obtained encouraging results. Taboada and Grieve (2004) initiated computational experimentation with the Appraisal framework, assigning adjectives into one of the three broad attitude classes, and providing a prominence schema which revealed that, when dealing with reviews, most adjectives are found in the middle and the end. Whitelaw et al. (2005) show how Appraisal taxonomies can significantly improve sentiment classification and extraction. Dong (2006) and Wang and Dong (2007) focused on the analysis of the semantic orientation of adjectives, and created a computational model that is able to identify grammatical and semantic structures of appraisal in design texts; in addition, they developed a support vector machine based system able to classify semantic orientation as well as human coders. Bloom, Garg and Argamon (2007) and Bloom, Stein and Argamon (2007) are another example of how research makes progresses on the automatic extraction of Appraisal, but focused only on Attitude realised in the form of adjectives.

There are also specific studies on the linguistic analysis of Engagement. They pay special attention to academic discourse and the acknowledgement of others' opinion by means of reporting verbs (Chatterjee, 2007), though sometimes, despite the

---

[8] See Pang and Lee (2008) for an overview of the field.

similarities, they do not adhere explicitly to the system of Appraisal (Hyland, 1999). Engagement has also shown to be an enlightening tool in the analysis of journalistic and political discourse (White, 1998, 2003; Marín-Arrese and Núñez, 2006; Miller, 2004) and Multimodality (Tan, 2010). The work of Brooke et al (2009), who adapted an existing English semantic orientation calculator to Spanish, deserves special mention.

As a conclusion, although some initial efforts have been made in the analysis of Engagement, further study is needed to get new findings on their use, realisations, and boundaries (e.g. its differences with Attitude, a topic deeply studied by Carretero and Taboada (2010a; 2010b) inside the CONTRANOT Project) and Graduation (for example, Paradis' (2003) study on *really*). Moreover, few studies have been carried out contrastive English and Spanish within the Appraisal framework, so a specific emphasis should be made on this aspect, either in monolingual or contrastive analyses.

4. Materials and Methods

This section describes the materials used in the study as well as the methods used to achieve the aims described in section 2 above.

4.1. Materials

The materials used for this study consist of a sample of 38 unedited book reviews posted online, 19 in English and 19 in Spanish extracted from the Simon Fraser University Review Corpus (Taboada 2008), which, when finally annotated, will form part of the larger CONTRASTES corpus (see Lavid 2008), currently compiled by members of the UCM Research Group on English-Spanish Functional Linguistics and its Applications.

The Simon Fraser U. Review Corpus —consisting of 1,600 movie, book, and consumer product reviews— was collected to do research on the automatic extraction of opinion (Taboada and Grieve, 2004; Taboada et al., 2006). 800 reviews are written in English and were extracted from Epinions.com between the end of 2003 and the beginning of 2004. Another pack of 800 reviews collected in 2008 are written in Spanish and come from Ciao.es and Dooyoo.es. In these three webpages, the goal of writing a review is to help other readers and consumers. In addition, in the Spanish sites, reviewers also have a potential monetary compensation, provided that their comments are marked as "useful" by other people. In these reviews, the name of the author (or user name) is provided, the posting date and a star rating, which is accompanied by a final mark as "recommended" or "not recommended", plus a star-value in a range from 1 to 5.

The online book reviews used in this study are usually written by non-professional users in an informal way, and addressed to peers. Some of them are regular visitors and users of those websites, and therefore they know each other. For this reason, personal experience and emotional comments may play an important role in the reviews. These online reviews differ from other types of reviews, such as professionally-written reviews in newspapers and specialized magazines, or school-based critique of stories and books.

The sample of 38 unedited book reviews (19 English and 19 Spanish) was divided as follows:

1.- An initial data set was first used for the two agreement studies described in section 5 below. This data set consisted of 10 online book reviews (5 in English and 5 in Spanish). Table 2 displays the name of the reviews of this initial data set and their length.

English		Spanish	
Book review	Length (words)	Book review	Length (words)
no7	242	no_1_1	138
no21	288	no_1_11	129
no25	308	no_1_12	322
yes9	470	yes_4_17	1313
yes11	658	yes_5_2	256
Total	1,966	—	2,158

Table 2. Initial dataset reviews

As shown in the table, the English set amounts to 1,966 words, while the Spanish set consists of 2,158 words, with an average of approximately 400 words for each online review. They were selected randomly with respect to their content, only taking into account their length, so there could be some balance between languages. Three reviews with a negative opinion of the book were selected in each language (no7, no21, no25; no_1_1, no_1_11, no_1_12), plus two positive ones (yes9, yes11; yes_4_17, yes_5_2). The English names stand for the polarity (yes/no) and the number inside the Simon Fraser University Review Corpus. The Spanish names show the polarity, the number of stars given (1 or 2 in the negative ones, and four or five in the positive ones, as those valued with three stars were not included in that corpus for they could be unclear or neutral regarding their point of view) and their ordinal number corresponding to their position in the corpus.

2.- A larger data set consisting of 28 reviews (14 English and 14 Spanish, 7 recommended texts, and 7 not recommended in each language) was annotated by the author of this paper once the annotation procedure proved consistent and reliable. Table 3 displays their identifier and their word length.

	English				Spanish			
	no		yes		no		yes	
	B.Rev.	Words	B.Rev.	Words	B.Rev.	Words	B.Rev.	Words
	no1	623	yes2	487	no1_14	328	yes4_9	409
	no3	738	yes5	436	no2_7	240	yes4_22	534
	no4	746	yes7	610	no2_10	819	yes4_23	795
	no5	397	yes12	418	no2_19	827	yes4_24	403
	no18	529	yes14	516	no2_20	438	yes5_5	363
	no19	369	yes17	698	no2_23	406	yes5_6	594
	no22	483	yes24	581	no2_24	510	yes5_16	737
$Total_1$	3885		3746		3568		3835	
$Total_2$	7631				7403			
$Total_3$	15034							

Table 3. Larger dataset reviews

As the table shows, the two languages were balanced, the total length of the corpus was about 15,000 words and the average per text was 500 words approximately. The results of the annotation of this larger set are presented in section 7.

4.2. Methods

Given the aims of the study described in Section 2 above, the following methodological steps were carried out sequentially:

First, two agreement studies were designed and performed in order to test the two main aspects of the linguistic category of Engagement empirically: 1) the identification of span realising Engagement; 2) the classification Engagement into different subtypes. The two agreement studies are described in detail in section 5 below.

Second, an intensive annotation of the larger data set was performed individually by the author of this paper to extract relevant quantitative and qualitative contrastive features of the use of Engagement in the bilingual online book reviews. The results of this second step are described in section 6 below.

5. Agreement studies

Agreement studies (also called reliability studies) are common in the NLP community where the quality of the annotations is essential for the success of the annotation project. As explained in Hovy and Lavid (2010: 23):

> It is taken as axiomatic that any annotation must be performed by at least two, and usually more, people acting independently, so that their tagging decisions can be compared; if they do not agree with enough reliability then the whole project is taken to be ill-defined or too difficult … The underlying premise of annotation is that if people cannot agree enough, then either the theory is wrong (or badly stated or instantiated), or the annotation process itself is flawed. In any case, training of computer algorithms is impossible on inconsistent input.

In the Linguistics community, and within the CONTRANOT project, agreement studies are used to test aspects of linguistic categories empirically (Hovy and Lavid 2010). The reliability of the annotation is based both on intra-annotator agreement (measures stability: consistency of each annotator alone over time), and inter-annotator agreement (measures reproducibility: different annotators on the same problem).

Following this methodology, two agreement studies were designed and carried out in the first phase of this study on the initial data set of 10 online reviews (5 English and 5 Spanish). The first study measured inter-annotator agreement on span selection, while the second measured inter-annotator agreement on the type of label chosen by the annotators on the previously selected spans. The annotation procedure for both agreement studies is described below.

The statistical test used to measure inter-annotator agreement in both studies was the *kappa coefficient* (K), which measures agreement among coders in any situation where at least two independent coders are analysing the same element. The operation is based on the difference between the actual agreement and the expected agreement by chance. The K value moves between 0 (the agreement is no other than the expected by

chance) and 1 (there is total agreement). The following table shows what the values gradually stand for.

	Poor	Slight	Fair	Moderate	Substantial	Almost perfect
Kappa	0.0	.20	.40	.60	.80	1.0

Kappa	Agreement
< 0	Less than chance agreement
0.01–0.20	Slight agreement
0.21– 0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

Table 4. Interpretation of Kappa (from Viera and Garret, 2005: 362)

The results will show the level of agreement for each analysis, and the reasons for disagreement will be discussed. However, a detailed explanation of why some markables were labelled in a specific way is not provided, as that is beyond the purpose of the section, as agreement studies are only the support for the individual annotation of the larger set.

5.1. Annotation procedure

Two annotators, #1 and #2, were asked to analyse Engagement expressions in the texts individually, both having studied Martin and White's Appraisal framework in depth.[9] They were provided with a specific annotation scheme based on the existing system network specification for Engagement in SFL, shown in Figure 3 above (reproduced below for convenience).

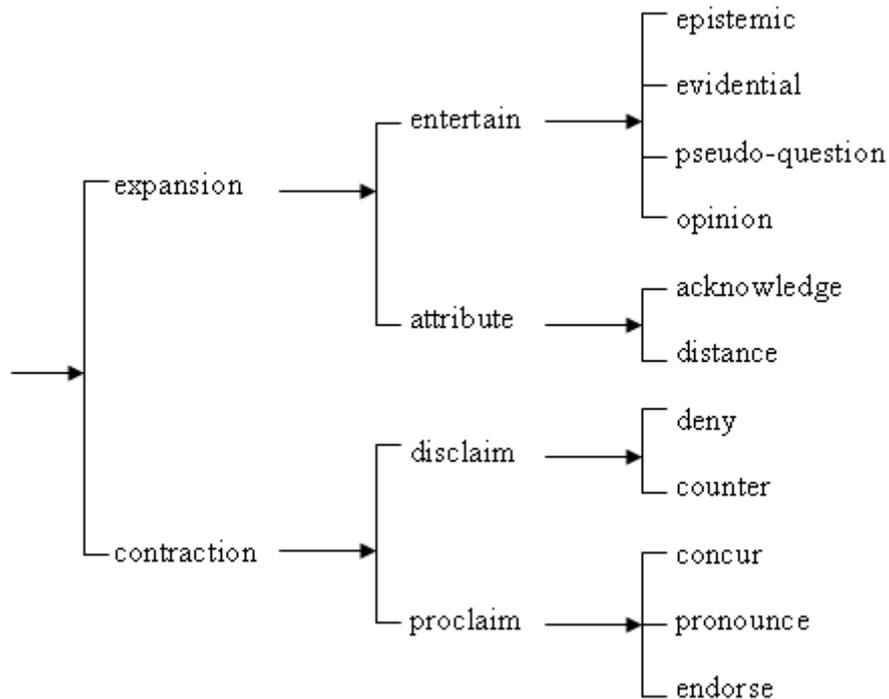[9] The two annotators were Dr. M. Carretero and the author of this paper.

Figure 3. The Engagement System (adapted from Martin and White, 2005: 134)

The annotators were given two types of annotation sheets with instructions, each corresponding to a different type of tasks:

- For the first task they had to identify the span realising the Engagement expression (see Appendix 1). After this initial analysis, their results were shared and measured, and the cases of disagreement were discussed in order to come to a common understanding and decide which expressions were actual examples of Engagement and how long they were, and, consequently, which specific spans were going to be annotated and classified in a second round.

- For the second task they had to classify the selected span as one of the possible subtypes of Engagement. In this phase, the initial template had been previously modified and the agreed spans had been highlighted, so the annotators would only have to think of the type they belonged to (see Appendix 2).

To sum up, the process consisted of two rounds in which the annotators worked individually and agreement results were measured; the first one based on span selection, and the second one on engagement type classification, with an intermediate phase in which results were discussed and the spans were selected by the two annotators together. The results for each of the agreement studies and further explanations on how they were obtained are presented in the following sections.

5.2. Results of Agreement study 1

The first agreement study focussed on the selection of the span expressing Engagement both in English and in Spanish. Span selection was performed individually by each annotator. Following the annotation procedure described above, annotator #1 marked 232 spans, and annotator #2 230. However, though the total number is quite similar, differences arise between languages, as table 5 shows.

English			Spanish		
Book Rev.	Annot. #1	Annot. #2	Book Rev.	Annot. #1	Annot. #2
no7	19	17	no1_1	9	8
no21	21	24	no1_11	10	11
no25	19	21	no1_12	18	19
yes9	16	21	yes4_15	80	69
yes11	24	24	yes5_2	15	16
Total	99	107	Total	132	123

Table 5. Annotated spans per language and annotator.

Clearly, while #1 found less markables in English than #2, in Spanish the situation was the opposite. The table also shows that while in English the difference in annotation was distributed evenly, in Spanish most of the disagreement is focused in the review yes4_15, which was the longest one. These data are complemented with tables 6 and 7.

			Value	Typ. asin. error	Approximated T
English	Agreement measure	Kappa	0.559	0.095	7,882
	No. of valid cases		112		
Spanish	Agreement measure	Kappa	0.497	0.077	7,826
	No. of valid cases		143		

Table 6. Span Selection. Symmetric measures (English and Spanish)

English

		Annotator #2			
		.00	1.00	2.00	Total
Annot. #1	.00	0	12	0	12
	1.00	4	83	0	87
	2.00	0	0	13	13
Total		4	95	13	112

Spanish

		Annotator #2			
		.00	1.00	2.00	Total
Annot. #1	.00	0	12	0	12
	1.00	18	91	0	109
	2.00	0	0	22	22
Total		18	103	22	143

Table 7. Span Selection. Contingency Annotator #1 * Annotator #2 (English and Spanish)

Table 6 shows the *kappa* coefficient obtained in English and in Spanish. Although both of them are classified as moderate agreement, the English texts (0.559) present a higher coefficient than the Spanish ones (0.497). It may seem lower than expected, but the examination of the cases of disagreement has shown that it was due to subtle differences in the analysis, which were solved in the discussion phase. If another individual analysis of span selection had been carried out, it is highly probable that the coefficient had been significantly closer to 1.

Table 7 displays the results obtained with respect to matches: 0 means that the annotator did not marked anything; 1 means that the annotator found a token representative of Engagement; if an annotator has 1 and the other 0, then there was disagreement; if both of them are marked with 1, the match was perfect; 2 means that the match was only partial, but both of them found Engagement. Perfect agreement is the most common result in both languages; however, the cases of partial agreement and disagreement require further explanation. Even though the results have been presented

per language for the sake of clarity, the reasons can be applied to both English and Spanish results, with some subtle differences which will be pointed out.

Disagreement (0-1 or 1-0) is mainly based on the following problems. In the first place, annotator #1 always considered the verb *seem* a markable, irrespective of what type of units was next, while annotator #2 considered only those cases in which it was followed by an infinitive. Finally, it was decided to mark it in all the cases, as the evidential meaning was thought to remain in the verb.

Some items in the boundary between categories (in this case Attitude and Engagement) were a source for disagreement. More specifically, a verb like *sorprender* was included in Engagement by #1, while #2 excluded it. It was agreed to include it, as it meant a kind of unexpected outcome.

As explained in section 5.1., the sentences which were considered part of the plot were not analysed. This was another source of disagreement, as the line between opinion and objective description may be very thin. In these cases, #1 had included more spans than #2, and it was decided to exclude them, for author's intention was thought to be describing.

Discursive markers, mainly in Spanish (e.g. *eso sí, es que, por lo demás, con todo eso…*) were another unclear resource. Some of them had been marked by #1 and others by #2, and the annotators decided to include them all, for they had actually gone unnoticed due to their untypical occurrence.

Partial disagreement (2-2) refers to cases in which both annotators found an expression representative of Engagement, but there was disagreement regarding the length of the span, even though they were marking essentially the same item. Wiebe et al (2005) and Read et al (2007) considered intersecting annotations as matches. The same criterion has been followed in this study when calculating *kappa*: although 1-1

represents exact matches and 2-2 only partial matches, in the statistical result they both are considered matches in the same way.

The most typical cases of partial match were found in the word in which the span ended (e.g. *la verdad es* vs. *la verdad es que*). Annotator #1 did not include the last word while Annotator #2 did. In most of the spans, the last word was decided to be included, as more often than not it cannot be substituted by any other word, that is, it is part of the expression.

Discontinuous spans were another problem to discuss. #1 divided a span when an item which did not represent Engagement or belonged to a different type was inside, while #2 marked them as a whole. For example, an expression like *am not even going to talk* was annotated as two spans by #1 ("am not…going to talk", "even") while #2 did not divide it.

The scope of negation also raised some doubts. In examples such as *sin leérselo a nadie*, #1 only marked "sin leérselo", while #2 included the object. In order to achieve symmetry with the English cases, where the object is positive, Spanish objects were excluded from the span. When negation affected verbs, #1 marked not only the negative element, but also the verb, while #2 did not (e.g. *no es* vs. *no*). Similarly to the previous case, it was decided to include the verb like it had been done in the English texts.

In conclusion, most of disagreement was caused by ambiguity and small gaps which where analysed differently, such as the case of Pronounce, which is especially open to interpretation, as shown in next section.

5.3. Results of Agreement Study 2

The second agreement study focused on the labelling of the spans according to the types of Engagement included in the annotation scheme. Tables 8, 9 and 10 present the results, with each subtype of Engagement category numbered as follows: 1. Epistemic,

2. Evidential, 3. Pseudo-question, 4. Opinion, 5. Acknowledge, 6. Distance, 7. Deny, 8. Counter, 9. Concur, 10. Pronounce, 11. Endorse.

		Annotator #2									Total
		1.00	2.00	3.00	4.00	5.00	7.00	8.00	9.00	10.00	
Annot. #1	1.00	14	0	0	0	0	0	2	0	2	18
	2.00	1	3	0	0	0	0	0	0	0	4
	3.00	0	0	5	0	0	0	0	0	0	5
	4.00	0	1	0	3	0	0	0	0	0	4
	5.00	1	0	0	0	4	0	0	0	0	5
	7.00	0	0	0	0	0	28	0	0	0	28
	8.00	1	0	0	0	0	0	33	0	0	34
	9.00	0	0	1	0	0	0	0	2	0	3
	10.00	0	0	0	0	0	0	5	0	3	8
Total		17	4	6	3	4	28	40	2	5	109

Table 8. Engagement subtypes. Contingency Annotator #1 * Annotator #2 (English)

		Annotator #2									Total
		1.00	3.00	4.00	5.00	7.00	8.00	9.00	10.00	11.00	
Annot. #1	1.00	14	0	0	0	0	1	0	1	0	16
	2.00	0	0	0	0	1	0	0	0	0	1
	3.00	0	7	0	0	0	0	0	0	0	7
	4.00	0	0	11	0	0	0	0	0	0	11
	5.00	0	0	0	10	0	0	0	0	1	11
	7.00	0	0	0	0	33	1	0	0	0	34
	8.00	0	0	1	0	0	33	0	0	0	34
	9.00	0	0	0	0	0	0	3	0	0	3
	10.00	1	0	0	0	0	1	0	15	0	17
	11.00	0	0	0	0	0	0	0	0	1	1
Total		15	7	12	10	34	36	3	16	2	135

Table 9. Engagement subtypes. Contingency Annotator #1 * Annotator #2 (Spanish)

			Value	Typ. asin. error	Approximated T
English	Agreement measure	Kappa	0.836	0.040	18.099
	No. of valid cases		109		
Spanish	Agreement measure	Kappa	0.928	0.025	24.664
	No. of valid cases		135		

Table 10. Engagement subtypes. Symmetric measures (English and Spanish)

In tables 7 and 8, there are some numbers missing in the columns and lines, due to the fact that no matches were found in those cases. The results, both in English and in Spanish, clearly show an almost perfect agreement, in which only remote cases were tagged in a different way. Most of disagreement was isolated; however, sometimes it was recurrent, like in 10-8, where there are five cases in English and one in Spanish in which annotator #1 regarded a span as Pronounce, but #2 marked it as Counter. Similarly, there are two occurrences of 1-8 (#1 marked Epistemic and #2 Counter) in English and one in Spanish, the same proportion as in 1-10 (#1 labelled the span as Epistemic and #2 Pronounce). It must be noted that a significant part of this disagreement was caused by the analysis of the adverb *really*, which also explains why the *kappa coefficient* was lower in English than in Spanish. Section 6, which will discuss tag assignation of the larger set of texts in detail, will provide further information on why some spans were classified in that way.

As a conclusion, it can be deduced from the distribution of Engagement shown in these tables that some classes are easier to identify than others, and that Pronounce and Counter are the most conflictive types, sharing to some extent the results obtained by Read et al (2007).

6. The annotation of the larger set

After the agreement studies were completed, the analysis of a larger set was carried out by one single annotator, the author of this paper. The purpose of the annotation of this larger data set was to discover contrastive patterns of use of Engagement categories in English and Spanish online book reviews, and to create a bilingual (comparable) machine-readable annotated corpus with Engagement features in English and Spanish which can serve as the training corpus for machine learning algorithms.

The annotation was performed with the UAM Corpus Tool (2011), a free and openly available state-of-the-art tool designed for annotating text corpora. The reasons for choosing this tool were the following: possibility to annotate multiple texts using the same annotation schemes, annotation of each text at multiple levels, possibility to search for instances across levels, comparative statistics across subsets, and the fact that all annotations are stored in XML files, so that annotations can more easily be shared with other applications. The UAM Corpus Tool also uses "stand-off" XML, which means that the annotation files do not contain the text, just pointers to the text, thus allowing for multiple overlapping analysis of the same text, not so easy in standard XML.

7. Annotation Results

In this section the results of the annotation of the larger set are described in quantitative and qualitative terms. The quantitative results are displayed in frequency tables, as obtained from the annotation with the UAM Corpus Tool. The qualitative results are summarised and illustrated with examples taken from the annotated data set.

7.1. Quantitative results

Regarding the quantitative results, the texts have been processed in two sets, according to language, displayed in tables 11 and 12. The annotated spans are presented in raw numbers and in ratio per total of marked spans, and the level of delicacy increases from left to right, so that every level, irrespective of delicacy, can be analysed.

Type	%	N.	Subtype₁	%	N.	Subtype₂	%	N.
expansion	35.4	121	entertain	33.3	114	epistemic	15.5	53
						evidential	6.1	21
						pseudo-questions	7.9	27
						opinion	3.8	13
			attribute	2.0	7	acknowledge	2.0	7
						distance	0.0	0
contraction	64.6	221	disclaim	51.5	176	deny	21.6	74
						counter	29.8	102
			proclaim	13.2	45	concur	2.0	7
						pronounce	10.5	36
						endorse	0.6	2
Total	100	342		100	342		100	342

Table 11. Global results of English reviews

Type	%	N.	Subtype₁	%	N.	Subtype₂	%	N.
expansion	29.9	102	entertain	29.0	99	epistemic	17.6	60
						evidential	1.8	6
						pseudo-questions	2.3	8
						opinion	7.3	25
			attribute	0.9	3	acknowledge	0.6	2
						distance	0.3	1
contraction	70.1	239	disclaim	62.8	214	deny	36.1	123
						counter	26.7	91
			proclaim	7.3	25	concur	1.5	5
						pronounce	5.9	20
						endorse	0.0	0
Total	100	341		100	341		100	341

Table 12. Global results of Spanish reviews

In addition, a further distinction has been done inside each language: the reviews which gave a positive mark have been separated from those with a negative one. Similarly to the previous results, these are presented in tables 13 and 14 for the English texts, and 15 and 16 for the Spanish ones.

Type	%	N.	Subtype₁	%	N.	Subtype₂	%	N.
expansion	39.2	83	entertain	36.8	78	epistemic	17.0	36
						evidential	7.5	16
						pseudo-questions	8.0	17
						opinion	4.2	9
			attribute	2.4	5	acknowledge	2.4	5
						distance	0.0	0
contraction	60.8	129	disclaim	49.5	105	deny	19.3	41
						counter	30.2	64
			proclaim	11.3	24	concur	0.9	2
						pronounce	10.3	22
						endorse	0.0	0
Total	100	212		100	212		100	212

Table 13. Results of Negative English Reviews

Type	%	N.	Subtype₁	%	N.	Subtype₂	%	N.
expansion	29.2	38	entertain	27.7	36	epistemic	13.1	17
						evidential	3.8	5
						pseudo-questions	7.7	10
						opinion	3.1	4
			attribute	1.5	2	acknowledge	1.5	2
						distance	0.0	0
contraction	70.8	92	disclaim	54.6	71	deny	25.4	33
						counter	29.2	38
			proclaim	16.2	21	concur	3.8	5
						pronounce	10.8	14
						endorse	1.5	2
Total	100	130		100	130		100	130

Table 14. Results of Positive English Reviews

Type	%	N.	Subtype₁	%	N.	Subtype₂	%	N.
expansion	36.4	76	entertain	34.9	73	epistemic	21.5	45
						evidential	2.9	6
						pseudo-questions	1.4	3
						opinion	9.1	19
			attribute	1.4	3	acknowledge	1.0	2
						distance	0.5	1
contraction	63.6	133	disclaim	56.0	117	deny	32.1	67
						counter	23.9	50
			proclaim	7.7	16	concur	1.9	4
						pronounce	5.7	12
						endorse	0.0	0
Total	100	209		100	209		100	209

Table 15. Results of Negative Spanish Reviews

Type	%	N.	Subtype₁	%	N.	Subtype₂	%	N.
expansion	19.7	26	entertain	19.7	26	epistemic	11.4	15
						evidential	0.0	0
						pseudo-questions	3.8	5
						opinion	4.5	6
			attribute	0.0	0	acknowledge	0.0	0
						distance	0.0	0
contraction	80.3	106	disclaim	73.5	97	deny	42.4	56
						counter	31.1	41
			proclaim	6.8	9	concur	0.8	1
						pronounce	6.1	8
						endorse	0.0	0
Total	100	132		100	132		100	132

Table 16..Results of Positive Spanish Reviews

In order to describe these results in a comprehensive and descriptive way, they have been summarized and arranged in three sets: most used types, least used types, and intermediate types (types left in between).

7.1.1. Most used types

As shown in tables 11-15, and summarized in tables 16 and 17, the most frequently used Engagement subtypes in English are Counter, Deny and Epistemic (in that order) while in Spanish the most frequently used subtypes are Deny, Counter and Epistemic. These frequencies are listed in tables 17 and 18 below.

	1st		2nd		3rd	
	Type	%	Type	%	Type	%
Global	Counter	29.8	Deny	21.6	Epistemic	15.5
Negative	Counter	30.2	Deny	19.3	Epistemic	17.0
Positive	Counter	29.2	Deny	25.4	Epistemic	13.1

Table 17. Most frequent subtypes in English

	1st		2nd		3rd	
	Type	%	Type	%	Type	%
Global	Deny	36.1	Counter	26.7	Epistemic	17.6
Negative	Deny	32.1	Counter	23.9	Epistemic	21.5
Positive	Deny	42.4	Counter	31.1	Epistemic	11.4

Table 18. Most frequent subtypes in Spanish

These similarities and differences between English and Spanish seem to indicate that while English reviews tend to turn over or modify previous or posterior claims, Spanish reviewers simply deny more systematically how things are. In other words, English authors prefer to establish a contrast between how things may seem and what they think, emphasising the latter, whereas Spanish reviewers overtly deny without making any concession. This characteristic is illustrated by examples [33] and [34].

[33] **though** it has been done again and again … the perspective offered in the tale is fresh enough to be thoroughly enjoyable [yes24]

[34] **No he encontrado** por ninguna parte "La agudeza, ironía y picaresca" [no2_19]

A high presence of the Disclaim types points to the author's interest in communicating what books (or the author, the reviewer or any entity) are not or do not do and introducing unexpected information (in the sense of contrary to what was previously claimed). Examples [35]-[38] illustrate these meanings in a very clear way.

[35] **Although** prolix, author and former schoolmaster Dan Brown **fails to create** adequate descriptions of people …, or places …, or objects …. [no1]

[36] **No se** se[10] recomendaría el libro, a mi por lo menos **no me ha gustado**, **ahora**, se lee rápido y sobre todo si el tema te gusta. [no2_20]

[37] This book **wasn't only** another good John Grisham tale, [yes2]

[38] **No esperéis** en ella Gheishas y Samurais **sino** japoneses actuales. [yes5_6]

In [35], "although" has been annotated as Counter and "fails to create" as Deny. In [36], "No se" is Deny, the same as "no me ha gustado", and "ahora" is Counter (though it is not a very used item). In [37], "wasn't" functions as Deny and "only" as Counter. In [38], "No esperéis" has been classified as Deny and "sino" as Counter. These examples show that similar strategies can be found in either recommended or not recommended books.

Regarding Epistemic items, they are the third most used type both in English and in Spanish. These expressions are often employed to present an opinion as not overbearing so, to a certain extend, it is tied to politeness strategies, as in [39] and [40].

[39] – while a touch of romance **might have been** nice for the story, Grisham made Luke 7 instead of, say, 15. [no4]

[40] Como curiosidad, el título original era "tierra por medio", menos espectacular y comercial pero **tal vez** más adecuado por el sentido que tiene en la trama. [yes4_23]

---

[10] As all the examples are transcribed true to the original, some mistakes may appear. In this case, it is understood that the appropriate form would be "No sé si recomendaría".

Here, instead of bare assertions of what the author thinks, these elements have been introduced to add probably untrue insecurity about the statements. In addition, it must be noted that there is a high frequency of conditional clauses. For example, a quite common strategy is illustrated by [41] and [42].

[41] **If** you want a fast-paced court-room drama, read another of his books first. But **if** you are tired of that sort of thing but what something different and deeper, read The Street Lawyer. [yes2]

[42] **Si** lo que quereis es percibir la dureza de ese tiempo y sociedad os lo recomiendo, pero **si** quereis pasar un rato entretenido no os lo recomiendo. [no2_7]

It is usual to find as closure of the texts this type of recommendations in which conditional sentences are used to diminish the strength of monoglossic assertions by means of an epistemic item. Actually, *if* and *si* represent the 41.5% and 38.3% respectively in English and Spanish with respect to the total number of epistemic occurrences.

Contrary to the case of Counter and Deny, Epistemic items appear mostly on negative texts. This is linked to the aforementioned fact of softening negative comments: if a reviewer wants to criticize a book, s/he will try to introduce more epistemic expressions to mark that his/her point of view is not be an unquestionable truth. In section 7.2 below some of these elements will be compared to those inside the category of Opinion, which is very related to the same phenomenon.

Some differences can be observed in percentages, which reflect how many spans were marked as each of these types with respect to all the spans annotated. In Spanish, the occurrence of Deny, Counter and Epistemic spans is significantly higher than in English[11], which means that the use of Engagement in Spanish is much more focused on some specific types than in English, where it is more distributed. Moreover, they are not

---

[11] There is only one exception, which is epistemicity in positive Spanish texts.

only higher, but, generally, there is also more distance between one rank and the next one, what delimits the three clearly differentiated dominant classes in the Spanish texts.

7.1.2. Least used types

The least used types have obtained between 0% and 2% of representation in the annotated spans. Such a low amount of occurrences implies problems when processing and training a corpus automatically, for there are not enough examples to base in and extract relevant features. Tables 19 and 20 list the three least used types in English and in Spanish.

	1st		2nd		3rd	
	Type	%	Type	%	Type	%
Global	Distance	0.0	Endorse	0.6	Acknowledge Concur	2.0
Negative	Distance Endorse	0.0 0.0	—	—	Concur	0.9
Positive	Distance	0.0	Acknowledge Endorse	1.5 1.5	—	—

Table 19. Least used types in English

	1st		2nd		3rd	
	Type	%	Type	%	Type	%
Global	Endorse	0.0	Distance	0.3	Acknowledge	0.6
Negative	Endorse	0.0	Distance	0.5	Acknowledge	1.0
Positive	Evidential Acknowledge Distance Endorse	0.0 0.0 0.0 0.0	—	—	—	—

Table 20. Least used types in Spanish

These percentages reflect what was seen in the most used types. It can be said that the least used Engagement type in English is the second one in Spanish, and the other way round. However, there is more variation than in the previous tables.

In the first place, Distance is the least common tag in English, both in negative and positive reviews. In Spanish, on the other hand, Distance has some representation, but Endorse does not appear at all, though in the positive reviews other types (Evidential, Acknowledge and also Distance) do not either. Acknowledge and Concur are also included, the former in the two languages, but the latter only in English.

These categories are classified in two more general groups: Attribute (Acknowledge and Distance) and Proclaim (Concur and Endorse). There are two exceptions: Evidential that appears in the Spanish table, belongs to the category called Entertain; and Pronounce, inside Proclaim, has obtained a significantly higher percentage than the other two related categories.

It is interesting to point out that all the types except one (Evidential) share a similar function in discourse. They are used to refer to other's ideas at the same time that the reviewer opines about them. The difference is that in the Attribute type it may be either neutrally (Acknowledge) or negatively (Distance), and the Proclaim categories present agreement with others' opinion. This means that reviewers do not use strategies to include others' comments, as they are mainly interested in presenting only their own point of view. Actually, Pronounce, which is used to reinforce one's words, is more used than these categories. The virtually nil occurrences of Endorse, specifically, can be explained by lack of appropriateness regarding the genre of the corpus. Book reviews deal with opinions and not with proved facts, and if reviewers want to transmit that something is really true, they tend to use Pronounce resources.

The few realisations that can be found of Acknowledge and Concur (the only two categories inside this group which appear in English and Spanish texts) are more or less similar across languages, as examples [43]-[46] illustrate. The other types cannot be compared, as there is no evidence to do so.

[43] I guess – **according to Grisham**, they pick cotton without complaint, enjoy a lively game of baseball, and help paint a house. [no4]

[44] **Como dice mi amiga Tifoidea**, esta lectura puede resultar algo espesa … [no2_24]

[45] Luke, the narrator, gives me the feeling he is 7 going on 22. **Of course**, that's probably because the diction of these hard working, poorly educated sodbusters is not written so as to be really "twangy". [no5]

[46] ¿Recomendable? **Por supuesto**, aunque si primero os habeis leído "Cien años de soledad" esta os sabrá a poco. [yes4_9]

The two first examples represent cases of Acknowledge, while the other two have been annotated as Concur. As can be seen, equivalent expressions are used to include the same Engagement category in the discourse.

The case of Evidential markers seems to be marginal, as such a low frequency has only been obtained in positive Spanish reviews, while the rest of the texts show different results. Table 21 summarizes Evidential occurrences.

English			Spanish		
Global	Negative	Positive	Global	Negative	Positive
6.1	7.5	3.8	1.8	2.9	0.0

Table 21. Results of Evidential annotations

Clearly, the English texts tend to include more often evidential items than the Spanish ones, though in both cases they mainly appear in negative texts. This higher use of evidentiality especially in negative reviews may point to the same conclusion drawn with respect to epistemicity. These expressions are often used in the first person, so they stress that the source of the information is one's mind, that is, it emphasizes that the reader is facing reviewer's opinion. If not, they are used in the third person, talking about what something *seems*, which is in the boundary between the something's inherent quality and individual's interpretation of it. Therefore, the attention shifts from

the object to the subject to soften negative comments and incredulity about facts, as exemplified in [47] and [48].

[47] So, **apparently** Luke was pretty short or his dad was pretty tall. [no4]

[48] Hay pasajes en los que **parece** que la cosa se va a animar, pero decae al instante. [no2_19]

The difference regarding frequency is due to the fact that both languages share some evidential expressions, but English adds some more that do not appear in Spanish, even though they could. The main resource for introducing evidentiality in these texts is the verb *seem* / *parecer*, and also *suppose* / *suponer*. However, English also employs the verb *feel* and *wonder*, as well as some adverbs such as *apparently*, but no equivalent is found in Spanish.

7.1.3. Intermediate types

The types which have not been discussed yet have obtained intermediate results regarding frequency. As the most and least used categories are very similar, these intermediate cases may be thought to be shared to some extend. Tables 22 and 23 summarize these results.

	1st		2nd		3rd		4th		5th	
	Type	%	Type	%	Type	%	Type	%	Type	%
G	pronounce	10.5	pseudo-questions	7.9	evidential	6.1	opinion	3.8	—	—
N	pronounce	10.3	pseudo-questions	8.0	evidential	7.5	opinion	4.2	acknowledge	2.4
P	pronounce	10.8	pseudo-questions	7.7	evidential concur	3.8	—	—	opinion	3.1

Table 22. Intermediate types in English

	1st		2nd		3rd		4th		5th	
	Type	%	Type	%	Type	%	Type	%	Type	%
G	opinion	7.3	pronounce	5.9	pseudo-questions	2.3	evidential	1.8	concur	1.5
N	opinion	9.1	pronounce	5.7	evidential	2.9	concur	1.9	pseudo-questions	1.4
P	pronounce	6.1	opinion	4.5	pseudo-questions	3.8	concur	0.8	—	—

Table 23. Intermediate types in Spanish

These tables do not show the same balance than the previous ones. Some categories are mixed across languages and others across review type. The most used strategy of these is unanimously Pronounce in English, while in Spanish the slot is shared with Opinion, the same as in the second one.

Pronounce is more used in the positive texts than in the negative ones, though it is not a substantial difference, so it can be said that it is associated indistinctly with either praise or critic. Regarding language, it is slightly more used in English than in Spanish. This is due to the fact that the insertion of the non-contracted form of auxiliary verbs is frequent in English when this meaning is intended, while in Spanish other strategies must be used for expressing the same emphasis, as shown in examples [49] and [50].

[49] Unfortunately, the book **does not do** much more. [no5]

[50] pero **de verdad** ya empieza a exasperarme, [no1_14]

Additionally, the English reviewers make use of capital letters for realising Pronounce (e.g. [51-53]), while this has not been observed in any of the Spanish texts.

[51] Or next summer at the lake when there is **ABSOLUTELY NOTHING** to do. [no5]

[52] I **DO like** Chrichton. [no19]

[53] Usually in King's books, he writes slowly and boringly so that all it does is get me **NOT** interested in one of his books [yes7]

The use of capital letters gives the impression of a more exaggerated and nearer to oral communication language. Spanish reviewers, on the other hand, focus more frequently on sincerity, the truth of the words and awareness of something. Examples [54]-[56] make these concepts easier to understand.

[54] En esta ocasión, **todo hay que decirlo**, no he quedado desilusionado con la lectura [no5]

[55] **La verdad es** que es una novela de dos ciudades… [yes4_23]

[56] **Sé** que es una visión particular y **sé** que me salgo de la fantasía propia del cuento [no2_10]

Opinion occurrences are mainly concentrated in the negative reviews. Actually, together with Epistemic and Evidential, Opinion is the third category that helps to soften negative comments. These three types belong to the more general one called Entertain, so some relation can be established between this category and that function. These expressions make explicit the fact that the utterance is one's opinion, thus tied to subjectivity and the conflict between individual's ideas and others' is minimised, as illustrated by [57] and [58].

[57] Here are some better titles for your summer reading list, **in my humble opinion**. [no4]

[58] **Desde mi punto de vista** el final no te aporta nada. [no2_7]

These markers are more used in Spanish than in English. This could be linked to the impersonality typically associated with English; however, the corpus reveals that these strategies are substituted by others equally personal, but that belong to different categories, predominantly evidential, such as *I get the feeling* (no4), *I feel* (no3), etc., which also explains why evidential markers are more common in English.

Pseudo-questions are far more frequent in the English language. Although the variation in the position of the ranking is only one slot, the percentages show a significant distance. This type of questions is addressed to readers, even though, obviously, they cannot give an answer. They differ considerably from the rest of Entertain categories with respect to the function previously explained. Typically, the reviewer makes questions about the development of the plot.

The distribution in review types is worth mentioning. While in English they are more common in negative texts, in Spanish the percentage is higher in the positive ones. This is due to a difference in the function they perform. When used in positive texts, questions are used to create mystery about the plot in order to encourage the reader to buy or read the book ([59]); in negative texts, they usually refer to something that the reviewer thinks it has not been appropriately accomplished in the book ([60]).

[59] **Will Peter** become King like he rightfully should? **Will Flagg** get his deserved punishment? **Will Thomas** apologize to his father's grave and to his imprisoned brother? FIND OUT AND BUY THE BOOK [yes7]

[60] … a mí personalmente me gusta poder identificarme con los personajes, y en este caso no encontré ninguno con el que hacerlo, **una abogada super rica, fashion, y genial?**, **un caballero templario?**, **o tres homosexuales?** [no1_14]

The remaining categories, Evidential, Concur and Acknowledge, were already discussed in the previous section. All these markers are more used in English and in negative texts, except for Concur, which in Spanish is more common in positive reviews. However, the occurrences are so few that it is difficult to claim that the differences are significant. A larger corpus would be required in this case to draw a more detailed and explanatory conclusion.

7.2. Qualitative Results

With a view to a plausible future automatic processing of Engagement expressions, the patterns found in the corpus are described according to each type from a contrastive English-Spanish perspective. The sketch of patterns is one of the most important issues in automatic analysis, inasmuch as they are the basis of algorithms searches; therefore, a detailed description of patterns' features may help to make progress in the field. In this paper, only some basic coincidences among the occurrences are commented, as a larger corpus would be needed for more precise results, and each conflictive case would require an individual study to clarify its uses, which is very far from the aims of the present study.

In general, expressions that represent Engagement can be divided into two groups: those which are systematically annotated as Engagement items and those which depend on the context (either to be analysed as Engagement or not, or to be inside a category or another). Some belonging to the former group are cited in Martin and White (2005)[12], so they do not cause major problems. However, those expressions which depend on the context are the ones that deserve a more thorough study. Some comments are also included about the annotation of specific expressions which are doubtful or conflictive examples and some clues on recursive items or patterns are suggested for future computational processing. In order to give a complete overview of the results of the annotation, the two types of items (those easily analysed and the controversial cases) are presented.

7.2.1. Epistemic

As the annotations have shown, online book reviews are very rich in Epistemic resources, and at the same time these are some of the most fixed expressions, which is also reflected in the similarities between languages.

---

[12] The expressions that they give as examples of Engagement are only in English.

Well-known cases of epistemicity are the conjunction *if* (*I would have been disappointed **if** I had purchased it* [no18]), phrases which include an epistemic word (*Fans of John Grisham will **no doubt** want to read this* [no4]), adjectives such as *possible* (***possible** spoilers ahead* [no1]) and *certain* (*Have you ever picked up a paperback ... not even **certain** that you will finish the book*? [yes14]), and adverbs such as *perhaps* (***Perhaps** it is the boring life on the cotton farm* [no3]), *probably* (*Which **probably** explains why it wasn't all about write-an-awesome-story-to-sell excitement* [yes2]), and *undoubtedly* (*he will **undoubtedly** be a therapist's field day in the future* [no4]). All these resources are not context dependent, so their analysis is always the same.

In addition, there are some adverbs that cause problems in the annotation: *certainly* (*This one was **certainly** different from the rest of them* [yes2]), *definitely* (*Mr. John should **definitely** return to writing about dealings with shyster lawyers, big money, crooked cops, righteous do-gooders, bad folks and such* [no5]) and its Spanish counterpart *definitivamente* (*No, **definitivamente**, creo que ni siguiera en película* [no2_19]). *Certainly* may be ambiguous, as it could be both Epistemic and Concur. Basing on Martin and White (2005: 124-5), who briefly discuss the case of *certainly*, it can be said that this adverb typically appears as Concur when used at the beginning of the sentence and separated by a comma, as well as in Concurring concession + Counter pairings (e.g. ***Certainly**, he was badly behaved, **but** look at what he has achieved* (ibid: 125)). These restrictions on their use regarding the category it belongs to may be essential for automatic analysis.

Definitely and *definitvamente* refer to a high level of certainty. However, they can also be used to indicate that something has been done categorically. The two meanings are distinguished mainly by the position these adverbs occupy in the sentence. If used as epistemic, they appear before or inside the verbal phrase, while if used as

manner adverbs they appear after the verbal phrase. It could happen otherwise if they are used as modifiers of another phrase (nouns, adjectives, etc.), where they appear after the verb but preceding the phrase they modify, and therefore, a dependency or a syntactic analysis (i.e. a parser) could shed light on the tagging.

The Spanish corpus has yielded very similar results to the English one. Mostly, the epistemic resources used can be directly translated from one language into the other, such as *si* (***si** queremos leer Caperucita, vayámonos a los dos cuentos de toda la vida* [no2_10]), *quizá* (***quizá** incluso algo menos* [no2_24]), *seguramente* (***Seguramente** se trate de esto último* [no2_10]), and *tal vez* (*menos espectacular y comercial pero **tal vez** más adecuado* [yes4_23]). As in English, these expressions can be analysed systematically as Epistemic. It should only be noted that *si* can work either in conditional clauses or in subordinate substantive clauses, typically functioning as direct object[13], therefore, a syntactic analysis should be carried out previously.

Other frequent epistemic item is the verb. Verbs which are not epistemic alone can be combined with phrases to create epistemic units. For instance, the fixed expression *to have no idea* (***I have no idea** why the KGB would be that interested in her* [no18]) is a clear example. These expressions can be analysed automatically in the same way as single units.

When epistemic by their own, some verbs are less ambiguous than others with respect to the Engagement type they belong to. For example, *suppose* (*Develops somewhat of a plot, **I suppose*** [no22]), *presume* (***I presume** this was to create tension* [yes24]) and *guess* (***I guess** that is my major issue with this book* [no3]), when they appear in the first person, are always classified as epistemic. Therefore, automatic

---

[13] This distinction has not been made when discussing *if*, as in English two different conjunctions are used — *if* and *whether*. However, it is true that *if* may be used in both cases. Then, a similar analysis to that of *si* should be followed.

processing would need to do a syntactic previously analysis in order to tag them in the right way.

On the other hand, a verb apparently simple as is the case of *think* may cause some problems. This verb can be analysed either as Epistemic or Opinion when used in the first person, depending on the meaning of the proposition. Examples [61] and [62] represent these two categories:

[61] **I think**, as others have mentioned, that I would have liked to know more at the end of the book [yes17]

[62] **I think** his work should be removed from the shelf [no22]

In [61], the sentence implies a probability and consequently *think* is epistemic. In [62], however, the verb expresses an opinion about carrying out an action and could be substituted by a phrase such as *in my opinion*, while the former could not or, at least, it would not be so natural. At present, computers have some difficulties when analysing the semantics of sentences, that is, a program would face a hard task if it were asked to understand the sentence as people do and identify whether it deals with probabilities or opinion. However, a plausible solution would be the statistical analysis of collocations. While in [61] the verb after *think* (taking into account only the main sentence) presents a conditional form with the modal *would*, [62] uses *should*. If these elements were interchanged, the former would be classified as Opinion, and the latter as Epistemic. Thus, collocates may be a decisive factor to be taken into account in this case, and they can be measured and processed by computers when tagging these categories.

Verbs are the other main source of epistemicity in Spanish as well. As in English, some verbs have an epistemic meaning by default, such as *suponer* (***Supongo que Beatriz, no llega a ser*** *una chica lesbiana* [no2_20]). The verb *creer* presents the

same problems as *think*. When used in the first person it may convey an epistemic or opinion meaning. As before, two propositions ([63] and [64]) can be compared.

[63] **Creo** que sus libros suelen estar relacionados con temas médicos [yes5_5]

[64] **creo** que lo ha dicho muy acertadamente [no2_24]

Example [63] has been classified as Epistemic (it deals with probabilities), while [64] could be considered Opinion and the verb could be substituted by *en mi opinión*. They support the need for the analysis of collocations when trying to distinguish between these two types. In the first sentence, the verb *soler* appears, which, if found with *creer*, typically implies epistemic propositions. The second example includes an element that could be classified as Attribute (*muy acertadamente*). Looking at the corpus analysed in the present study, valuations tend to be included in Opinion sentences[14]. If the sentences had been like those analysed in English, the same criterion could have been applied, and the other way round.

Modal verbs are another common resource for epistemic meaning. In the corpus, some of them have been found, such as *have to* (*This **has to be** the worst I've ever read from Michael Crichton* [no19]), *must* (*Dean Koontz **must be** the master...the master of overwriting* [no22]) or *might* (*readers **might enjoy** the trip back in time* [no3]). Nevertheless, modal verbs are not only used as Epistemic, but they can also convey a deontic meaning. Bearing in mind that distinguishing between these two meanings (which can even be overlapped in an utterance) causes linguists many headaches, it seems to be far from the capabilities of automatic analysis, inasmuch as pragmatic and

---

[14] This does not mean that it always happens in this way. For example, such a simple sentence as ***Creo que es muy alto*** could be classified as both Epistemic (if talking about someone the speaker has not met yet) and Opinion (if the speaker already knows the other person). Obviously, the solutions discussed here are only some suggestions which can be studied more in detail, and there are more conflictive cases than those that have appeared in the corpus.

contextual elements (e.g. whether someone has the authority to permit or forbid something), which sometimes have much to do with the meaning implied in this type of sentences, are still beyond what natural language processing can do. Probably, some progress should be made in this aspect before undertaking this task.

Finally, modal verbs are also common in Spanish, but they are much less numerous than in English, as they are mainly but not only restricted to *poder* (***Puede que ... me atreva*** *con otro libro de esta misma autora* [yes4_23], *tener que* (*mi optimismo natural me hace pensar que la cosa **tiene que mejorar*** [no2_19] and *deber* (*este fiasco no **debería pillarme** de sorpresa* [no2_19]. As their English counterparts, they can be used either as Epistemic or Deontic, and the same problems aforementioned arise.

7.2.2. Evidential

As explained before, some of the evidential expressions are shared but English is qualitatively and quantitatively more profuse. Shared expressions are verbs, mainly *seem* (*Grisham really **seems to take** a lot of time setting the scene and developing the characters* [no3]) and *parecer* (*Caperucita (Sara) al principio **parece buscar** la manera de ser feliz* [no2_10]), used in the third person. Evidential verbs in the first person also appear in the two languages, though the English ones focus more directly on feelings than the Spanish verbs. For instance, in English one can find *feel* (***I feel** that this book was very unrealistic* [no18]), while in Spanish one sees *intuir* (*Simplemente con leer el argumento, **intuí** que no me gustaría* [no2_19]). To verbs, English adds adverbs which do not appear in the Spanish texts, such as *apparently* (*So, **apparently** Luke was pretty short or his dad was pretty tall* [no4]).

There seems to be no problem to classify these items as Evidential systematically. Only the verb *seem* could depend on what comes after it. In the

reliability studies it was decided to mark it always, but it could also be thought to be Evidential only when followed by an infinitive with *to*. In that case, it would just be needed a part-of-speech parser. *Parecer* could be analysed in the same way.

7.2.3. Pseudo-questions

Pseudo-questions are realised in four different ways in the corpus. In the first place, they can be complete and traditional questions, where the subject and the verb are inverted and there is a question mark at the end. In these, the span selected as Engagement is only the mood, that is, the subject plus the operator (e.g. ***why doesn't the baddy*** *just shoot the goody instead of talking at him and then letting him escape?* [yes14]). In Spanish, some traditional questions (which follow the grammatical Spanish rules of interrogative mood) have also been found (e.g. ***¿Y qué nos cuenta Haruki*** *en esta novela?[15]* [yes5_6]). A syntactic parser would be enough for tagging this type of questions. In addition, the final question mark would be a significant indicator of a question taking place, and it should be noted that in Spanish, even though the rule says that two question marks are needed in a sentence to be spelt properly, it is common to find only the last one in informal texts, like those studied in the present paper.

Secondly, pseudo-questions can be realised by affirmative sentences (i.e. there is no subject-verb inversion) to which a question mark has been added, which is also usual in English informal speech (e.g. ***A child of 7 would*** *think in those terms?* [no5]). As in the previous case, the question mark would trigger a reinterpretation of the sentence.

Thirdly, pseudo-questions can consist of simple phrases, for instance, when some questions are made successively and the rest of the sentence is omitted, one could only find nominal/adjectival/etc. phrases. These also include question marks, but, as there is

---

[15] In Spanish, it may happen that some elements are included between the operator and the subject. In order to do the most appropriate analysis in the UAM Corpus Tool, the span was selected without breaks.

no mood, the all the phrase is tagged either in English (e.g. ***Stained glass*** *the smell of* ***incense or centuries of candle wax*** ***Pigeons in the rafters*** ***Bats in the belfry*** [no1]) or in Spanish (e.g. ***una abogada super rica, fashion, y genial*** ***un caballero templario*** ***o tres homosexuales*** [no1_14]). In these cases the question mark would again be a decisive indicator for the computer to backtrack and reinterpret the sentence.

Finally, pseudo-questions can be realised by means of indirect questions, that is, they are not made in a direct way with a question mark. These are the most difficult to recognise. However, these sentences do require an answer, which may be provided by reviewers themselves. These should not get confused with questions transformed into declarative sentences, for example in reported speech. Only one example was found in the corpus, more specifically in the English texts: *Now **how much can** develop between a 7 year old boy and 17 year old girl. Fortunately, not much.* [no5]. The analysis of this sentence as a pseudo-question may be controversial, due to the very indirect nuances aforementioned. However, the fact that an answer is provided points to the reviewer's intention of creating a dialogue with readers and, actually, readers expect an answer after the first sentence, for the reviewer is talking about the plot of the book, and the reader does probably want to know what happens. The difference between this sentence and prototypical indirect questions is that, in this case, it is not part of a subordinate clause, but a main one, so a parser could help to identify these units.

7.2.4. Opinion

Opinion expressions have not shown a wide variety of resources. They are mostly realised by verbs in the first person and phrases, irrespective of the language. These verbs can be divided into two groups: those whose meaning is more transparent, and those which are less. The former group includes the verbs *think* (***I think*** *I'd better check it out before*

buying it [no5]) and *creer* (**creo** *que el autor no sabe conducir la trama, pues realmente no sucede gran cosa* [no2_7]). The difference that exists between these verbs functioning as Epistemic or Opinion was discussed above. Collocations seem to play an important role in deciding which option is appropriate. The kind of verbs that comes after these expressions is probably the key to annotate them in the proper way —for instance, it would be difficult to imagine an example where *I think I'd better...* could be Epistemic. In Spanish, after examining the corpus, it can be said that the judgement on other's capabilities (e.g. *no sabe*) tends to be more associated with opinion than epistemicity. In addition, other type of parts of speech can also be important, such as *realmente* in the Spanish sentence above. The most obvious problem there is the range of collocations, for *realmente* is probably too far from *creo* to be in an acceptable scope.

Verbs which are less transparent are those such as *find* (**I found** *the book as entertaining* [no4]) and *feel* (**I feel** *that this book was very unrealistic due to the fact that Susan was a physist, yet she was not a spy* [no18]) in English, and *parecer* in Spanish, which is used in the third person and the speaker is included as indirect object (*La edición de bolsillo, de la editorial Tusquets, no **me pareció** cara* [no2_24]). What all these verbs have in common is that soon after them one can find adjectives or Attitude elements and, as a result, they change their original meaning into opinions. So far, *parecer* had been analysed only as Evidential. The difference now is that the first person is included before the verb. Therefore, in order to be automatically annotated, this pattern should only be introduced without further problems. The analysis of the verb *parecer* proposed in this paper can be summarized in the following way:

Use of *parecer*	Engagement	Example	Meaning
Me parece + noun	Opinion	Me parece (que es) un pájaro	Creo que es un pájaro
Me parece + adjective	Opinion	Me parece interesante	Creo que es interesante
Me parece + infinitive	Opinion	Me pareció ver algo detrás de la cortina	Creí ver algo detrás de la cortina
Me parece que + subordinate clause	Opinion	Me parece que estás equivocado	Creo que estás equivocado
Parece + noun	Evidential	Parece (que es) un pájaro	Da la impresión de que es un pájaro
Parece + adjective	Evidential	Parece tonto	Da la impresión de que es tonto
Parece + infinitive	Evidential	Parece buscar la manera de…	Da la impresión de que busca la manera de…
Parece + subordinate clause	Evidential	Parece que va a llover	Da la impresión de que va a llover

Table 24. Use of *parecer*

As well as verbs, some phrases are also typical realisations of Opinion, such as *in my (humble) opinion* (*This one was certainly different from the rest of them **in my opinion**, in both good and bad ways* [yes2]). In Spanish the variety is wider, as one can find, apart from those which were already included in previous examples, *a mi entender* (*el fallo que tiene, **a mi entender**, es el no desarrollo del argumento* [no2_7]) and *para mi gusto* (*acaba, **para mi gusto** dejándose llevar por las tonterías que le cuenta una tía loca* [no2_10]). These phrases can be analysed systematically, but the list can be too long to be considered individually. Nevertheless, all the phrases found share a similar syntactic form: preposition + (modifier/adjectival phrase) + noun phrase, in which the possessive *mi* is used before the noun. Therefore, it is more than likely that they can create a fixed group.

7.2.5. Acknowledge

The Attribute category is the least used type of Engagement in this corpus and very few occurrences were obtained, although Acknowledge instances were found both in English and in Spanish.

These expressions respect the fixed expressions *according to* (***According to the New York Times***, *Doubleday distributed 10,000 advance copies of the book to booksellers and the media* [no1]) and *as X has said/mentioned/etc.* (*I think,* ***as others have mentioned****, that I would have liked to know more at the end of the book* [yes17]). In Spanish, a similar example was found (***Como dice mi amiga Tifoidea****, esta lectura puede resultar algo espesa para una persona que se enfrenta a su primer libro de autor japonés* [no2_24]). In addition, verbs of saying in the third person were also tagged as Acknowledge (e.g. ***the book jacket says*** *that is was written by Dean Koontz* [no18]), as well as the verb *refer*, in the structure *refer + to +* noun phrase + *as* (***Fiction writers often refer*** *to this as "maid and butler talk"* [yes24]). Analysing these expressions automatically seems to be quite easy, as only the patterns aforementioned should be looked for. However, a larger corpus could reveal different structures.

7.2.6. Distance

Only one example of Distance was tagged in the corpus, more specifically, in the Spanish texts. This example also includes a verb of saying, *decir* (*pero acaba, para mi gusto dejándose llevar por las tonterías que le cuenta una tía loca que* ***dice ser*** *la Estatua de la libertad en persona* [no2_10]). The structure *decir +* infinite used in the third person seems to convey a meaning of unbelief which is not implied in other structures and, consequently, it is classified as Distance instead of Acknowledge. Just like for the latter type, a higher variety and quantity of realisations would be needed to reach more significant conclusions.

7.2.7. Deny

In opposition to the previous category, those in Disclaim were the most used. The simplest strategies are the negation of the verb (*The chapters* ***aren't*** *really long* [yes7];

*Supongo que Beatriz, **no llega a ser** una chica lesbiana* [no2_20]), and negative particles such as *never* (***If** you have **never** had the fun of reading a book in this series **do not start** with this one* [yes5]) and *nunca* (***Nunca** es tarde* [yes4_9]), *nothing* (*it really was **nothing*** [no3] and *nada* (*cuando **nada** se espera* [yes4_9]), etc. As explained in section 3.2.2.a), not all cases of negation can be considered as Deny. In order to annotate Deny appropriately, the scope of negation must be analysed. If it affects a group, but not the verb (which can be checked with question tags), then it is not Deny. Seemingly, this type of grammatical correctness can be analysed automatically[16]; therefore, analysing at least the instances found in the corpus would be plausible.

Some verbs whose meaning is negative, though used in affirmative, were also found, such as *fail* (*the characters … and **failed to react** in a believable manner to the situation* [yes24]), *neglect* (*There are authors that **neglect** to do this* [no3]), *evitar* (*me alegro de tener en mi pueblo una Biblioteca Municipal medianamente decente y **evitar** así gastarme una media de 20 € en cada libro* [no2_19]) and *impedir* (*mi orgullo me **impide** dejar un libro a medias* [no2_19]). Unless the automatic analysis could make use of a successful semantic analysis of verbs in which the fact that a resulting action has not been accomplished is specified, these verbs could be introduced as a fixed list in order to be analysed, as far as they are always analysed in the same way.

Finally, some correlative particles were also found, mainly in Spanish. These included items such as *ni* (*no es una lectura para todos los públicos **ni** para todos los estados de ánimo* [no2_24]) and *tampoco* (*Me ha recordado mucho a "Si te dicen que caí", de Joan Marsé, que **tampoco** me gustó mucho* [no2_19]). These particles consist of such a limited number that do not present any apparent problem for systematic analysis.

---

[16] For example, texts editors can to recognise errors in negation, such as double negation, and consequently, they are able to tell whether the verb is affected by the negative particle or not.

7.2.8. Counter

Counter is the other Disclaim type and, though it was less frequent than Deny, the variety of occurrences is richer, mainly due to the different adverbs and phrases which can be employed. In the English and the Spanish texts, the Counter resources which did not present any problem for the process of analysis were *instead* (***Instead*** *this is the tale of Luke Chandler and his family* [no3]), *al/todo lo contrario* (*Que nadie piense que es una especie de relato erótico novelado,* ***todo lo contrario****, apenas hay sexo* [yes4_23]), *en cambio* (***En cambio****, a Aurea le gustó mucho* [no2_24], *however* (***However*** *on the positive side, this book contained a much deeper message than those other books did* [yes2]), *sin embargo* (*Yo* ***sin embargo****, discrepo* [no2_19]), *but* (*The chapters aren't really long...****but*** *the artworks amazing* [yes7]), *pero/mas* (*Es de bolsillo,* ***pero*** *el tamaño de la letra es legible* [yes5_6]), *sino* (*No espereís en ella Gheishas y Samurais* ***sino*** *japoneses actuales* [yes5_6]), *only* (*When the story does finally pick up a little bit it is* ***only*** *temporary* [no3]), *sólo* (*Tal vez (****sólo*** *tal vez) si me explican la historia con dibujitos...*[no2_19]), *lo único* (***lo único*** *que me aportó fueron pesadillas varias, y muchísimo sueño* [no1_14]), *except (for)/with the exception of* (*A good-natured bunch of gringos,* ***except for*** *the switchblade brandishing one called Cowboy* [no4]), *(al)though* (***Although*** *prolix, author and former schoolmaster Dan Brown fails to create adequate descriptions of people* [no1]), *aunque* (*no esperéis encontrar epidemias ni virus,* ***aunque*** *no creo que decepcione a sus seguidores* [yes5_5]), *actually* (*Well* ***actually****, you have until Saturday Night* [no19]), *in fact* (***In fact****, he is so bad that I think his work should be removed from the shelf* [no22]), *even* (*It might* ***even*** *make a good movie since the movies tend to cut out all the technical stuff anyway* [no19]) and *incluso* (*Con esto quiero decir y reconocer que he dejado el libro* ***incluso*** *antes de llegar a la mitad* [no2_24]). These resources are repeated many times throughout the texts and do not present interpretation problems.

Some other expressions are more open to different readings, more specifically, the words *still*, *yet* and *just*. *Still* should only be analysed when functions as an adverb (*But **still**, this story would have us believe that ...* [no19]). *Yet* when it is either an adverb (*Thus begins a summer full of adventure, the likes of which I have **yet** to see in all of my 38 years* [no4]) or a conjunction (*A grotesque, **yet** funny, look at Southern poverty* [no4]). And *just* when is used as an adverb (*it's as if he didn't know how to end it, and so he **just** stuck in whatever* [no22]), but there are some exceptions in this case. A dictionary may help to distinguish the meanings which should be annotated. According to the *Macmillan English Dictionary for Advanced Learners* (2002: 778), *just* as an adverb can be used in the following contexts: 1) when something happens; 2) as 'only'; 3) as 'exactly'; 4) for emphasis in spoken language; 5) when something almost does not happen; and 6) for downtoning the speech. Only the contexts 2, 4, 5 and 6 should be analysed as Counter. The fact that these words (*still*, *yet* and *just*) are only annotated as Engagement when they belong to a particular word class implies that a part-of-speech analysis is previously required. The special case of *just* may cause additional complications for automatic analysis, which are even increased when the numerous phrases in which *just* may appear are considered. It could be easier to restrict the cases in which it should not be annotated than the other way round. Although a specific and detailed corpus-based study of this word, collocations and uses should be needed[17], the dictionary gives some clues about some patterns that can be left out, such as *just* + deictic temporal expressions (*now*, *then*, *last week*, *at the moment*, etc.), *only just, to be just going/about to, just the* + place/thing/person, *just on*, etc. Therefore, it seems plausible to limit, at least to a certain extent, the analysis.

---

[17] Aijmer (1985) could be a helpful study for this aspect.

In Spanish, similar units to these three discussed above are *ya, aún* and *simplemente*. *Ya* could be considered as Engament only when it functions as an adverb (*creo que **ya** ha terminado la versión cinematográfica de ésta novela* [yes4_22]), so the use of a parser is recommended; *aún* does always belong to this category (***Aún** siendo reacio a volver a caer en este tipo de temáticas tan recurridas últimamente, decidí darles otra oportunidad* [yes5_16]); as either does *simplemente* (***Simplemente** con leer el argumento, intuí que no me gustaría* [no2_19]). These Spanish items can be as easily analysed as the English ones.

The main clause of conditional sentences in which a fact that could have been carried out but it actually did not can also be analysed under the label Counter. These typically consist of a modal verb (would/could) followed by a present perfect (e.g. ***I would have been disappointed** if I had purchased it* [no18]). No instance of a conditional clause analysed as Counter that did not correspond to this scheme was found in the corpus, as well as there were no cases in which this scheme were not analysed. Therefore, a parser can help to improve the automatic analysis of this group of expressions. No equivalent example was found in Spanish, though they are expected to appear in a larger corpus.

Two cases were found in Spanish with no direct equivalent in the English texts: *ahora* and *es que*. *Ahora* was found once in the corpus where it did not conveyed a temporal meaning: *No se se recomendaría el libro, a mi por lo menos no me ha gustado, **ahora**, se lee rápido y sobre todo si el tema te gusta* [no2_20]. The use of the commas before and after this adverb changes its meaning clearly, so an automatic interpretation could be restricted by this structure. On the other hand, *es que* (*Sé que es una visión particular … pero **es que** … los personajes son superficiales y muy poco perfilados* [no2_10]) is a very difficult unit to analyse. It can be used in such a wide variety of contexts that it is a hard task to delimit the specific contexts in which it could

be analysed as Engagement, and as far as these criteria cannot be specified and featured, automatic analysis does not seem possible. Together with other expressions discussed in this section, *es que* is a very good candidate for a detailed analysis under the approach of Appraisal. At first sight, it can only be suggested that, in the example, the use of a preceding Counter unit such as *pero* may trigger this interpretation and that the expressions must be taken as a whole, and consequently those cases in which *que* is used a conjunction to introduce subordinate substantive clauses are excluded. In addition, *es que* can be considered to be in the boundary between Counter and Pronouncement, but in this paper it has been annotated as the former because the surrounding clauses emphasized this meaning more than the other.

There is one resource in English that could almost be included in Counter: *and then*. It has not been annotated because it is actually in the boundary, but it is worth mentioning. Examples [65]-[67] show some of the cases found.

[65] I told my husband there is about 30 pages of boring **and then** about 3 pages of excitement [no3]

[66] It spent about 20 pages setting up the ending **and then** it took about three pages to pull it off [yes7]

[67] It sucked because you dedicated 380 pages to all of the characters **and then** the ending just sucks [yes7]

Albeit *then* is a temporal adverb, they not simply reflect a relationship between two consecutive events. These instances show a meaning of counter expectancy and contrast between what is said in the first clause and what is introduced by *and then*. Although this relationship is purely semantic, some parallelisms can be found in the structure of the sentences: in [65] there are two similar noun phrases (*30 pages of boring* and *3 pages of excitement*); in [66], one finds almost the same structure (*30 pages setting up the ending* and *three pages to pull it off*); in [67], there is greater variation but inside two complete clauses (*you dedicated 380 pages to all the characters* and *the ending just*

sucks). What may also play a significant role in creating a different specific meaning in these structures is polarity. While one clause can be seen as positive (or maybe neutral), the other expresses a negative judgement. However, the polarity of the two sentences put together is negative due to the result of the comparison they make. In conclusion, a specific analysis of this structure and its implications may be significant for adding it as a Counter resource.

7.2.9. Concur

Concur is the first of the last type of Engagement, Proclaim. There were very few occurrences in this category, but some parallelisms can be established across languages. The most ordinary expressions were *of course* (*Except, **of course,** for the female interest that show's up in the person of Tally* [no5]) and *por supuesto* (*¿Recomendable? **Por supuesto**, aunque si primero os habeis leído "Cien años de soledad" esta os sabrá a poco* [yes4_9]). Adverbs were also used, such as *claro* (*Tal vez (sólo tal vez) si me explican la historia con dibujitos... **Claro**, y si no la protagoniza Banderas...* [no2_19]). While *of course* and *por supuesto* are unambiguous regarding their analysis, *claro* would need to be distinguished from the adjectival homographic form by using a part-of-speech tagger. In addition, when used as an adverb with a Concur meaning (in opposition to its use as a manner adverb complementing the verb), it is usually preceded and/or followed by a comma, though this may be omitted in informal speech, and it can also be used together with *que*. These two criteria could help to mark the boundary between the two meanings.

A verb was found in the corpus that if used in the first person in an affirmative sentence can be seen as Concur: *agree with*[18] (***I would agree with** the quotes of the cover from Norman Mailer ("*[quote]*")* [yes14]). Actually, this structure makes explicit

---

[18] If the subject were not *I/we*, then it could be classified as Acknowledge, and If the verb is negated, it would probably be labelled as Distance.

the underlying meaning of Concur expressions. In a similar way, the verb *saber* (*Compré este libro porque estaba publicitándose en la Fnac, ya **sabeis** que de vez en cuando ponen varios libros en paneles ...* [no2_24]) was found in the Spanish texts, and if used in the second person as well as in affirmative clauses, the Concur meaning appears. In addition, they must be followed by a substantive subordinate clause (in Spanish, the conjunction *que* shows it clearly, though in English it may be omitted) instead of a verb, where the meaning changes to 'be able to'. These verbs should not cause interpretation problems when these conditions are given.

7.2.10. Pronounce

Several strategies have been employed to introduce Pronounce in the texts. The most obvious one is the use of capital letters (*The only real thing that I didn't **AT ALL** like about this book is that the ending was horribly rushed* [yes7]). As said in section 7.1.3., capital letters used for emphasis do only appear in the English texts, though the usage in Spanish could be expected in a larger corpus. When analysing automatically, capital letters should not get confused with acronyms. While ordinary words can be used both in lowercase and uppercase letters, acronyms can only be spelt in capitals.

Adverbs, adjectives and phrases can also convey Pronounce meaning, mainly in the Spanish texts. For example, they can use the word *fact/hecho* for emphasising that something is just in the way it is said, in phrases such as *the fact that* (*I feel that this book was very unrealistic due to **the fact that** Susan was a physist, yet she was not a spy* [no18]) and *el hecho de que* (*no sé si es que nos escritores desconocen **el hecho de que** existieron más órdenes de caballería, tanto o mejores que los templarios* [no1_14]). A similar function is performed by *lo que ocurre es que* (***Lo que ocurre es que** mi orgullo me impide dejar un libro a medias* [no2_19]). These expressions are so fixed in the language that they are easily systematically analysed.

Other phrases conveyed Pronounce by making explicit reference to the truth of the message. These occurrences were only found in Spanish, but there is no prior cause to think they cannot appear in English as well. These include *sinceramente* (***sinceramente**, me parece patético* [no1_14]), *de verdad* (***de verdad** ya empieza a exasperarme* [no1_14]), *la verdad es que* (***La verdad es que** es una novela de dos ciudades* [yes4_23]), *todo hay que decirlo* (***todo hay que decirlo**, no he quedado desilusionado con la lectura* [yes5_16]), *lo cierto es que* (***lo cierto es** que si yo me topara con una opinión como esta, iría corriendo a buscar el libro* [no2_19]). Again, these all are used ordinarily in everyday speech and are prone to be found exactly in the same way in most of the texts analysed.

As has happened before, some expressions may be in the boundary between two categories. This is the case of *eso sí* (*lo único que me aportó fueron pesadillas varias, y muchísimo sueño, **eso sí**, contra el imsomnio es infalible....* [no1_14]). Although it has been annotated as Pronounce, it is worth noting that it may also imply some nuances of Counter meaning, as the example shows. However, it seems that the cotext (i.e. the surrounding text) probably has an influence, while the Pronounce meaning may be more inherently rooted in the expression.

Another Pronounce resource is based on verbs of knowledge, but with a first person subject, such as *know* (***I know** this is nit-picky* [no4]) and *saber* (***Sé** que es una visión particular y **sé** que me salgo de la fantasía propia del cuento* [no2_10]). This use contrasts with the one seen in Concur, where the subject appeared in the second person. While the former focuses on the fact that something is true because the authorial voice is aware of it, the latter refers to a knowledge that is shared by everybody or, at least, the interlocutors. Although few occurrences of this type were found in the corpus, there was no case in which this pattern was not followed; therefore, it seems that this distinction can be made systematically.

As explained in section 7.1.3., the use of the non contracted form of auxiliary verbs in affirmative (*When the story **does finally pick up** a little bit it is only temporary* [no3]) and negative clauses (*he took a perfectly good story, slapped on a fancy ending- that **did not go** at all with the storyline and called it a book* [no18]) is another source of Pronounce meaning, but this is only found in the English texts. As the examples show, satellite items can be included to reinforce the emphasis already introduced by the absence of contraction, such as *finally* and *at all*, although these are not considered inside Engagement[19]. The use of the non-contracted forms is not exclusive of emphasis. It is also typical in formal writing, for example. However, the book reviews analysed in the present paper were said to belong to an informal register, what explains the number of misspellings. Therefore, the contrast between this register and the use of complete auxiliary items justifies the meaning of Pronouncement that they convey. Thus, register, as well as genre, is one decisive factor to be taken into account in the analysis of some Engagement resources.

The most conflictive expression annotated as Pronounce in the corpus are the adverbs *really* and *realmente*. They are problematic because of the several denotations they have and the fact that they are differently annotated: they could be classified as Graduation (*The chapters aren't **really** long like most of Kings' books* [yes7]) or Engagement (Pronounce) (*I felt Grisham **really** took too much time here* [no3]; *el autor no sabe conducir la trama, pues **realmente no sucede** gran cosa* [no2_7]) and The important role of context and cotext in the analysis of this word[20] seems to move it away from an easy automatic analysis.

---

[19] *Finally* has been spelt in bolds because the spans have not been broken even though other units were introduced.
[20] The reader is remitted to, for example, Paradis (2003).

7.2.11. Endorse

The last type of Engagement is Endorse. As the meaning of proving something to be true is not much related to the genre analysed here, as explained in section 7.1.2, only two occurrences were found in the corpus which could be annotated under this label, and they both were in the English texts: *In Eyes of the Dragon,* **SK proves** *... that he is at the top of his game still at his old age* [yes7]; **Eyes of the Dragon shows** *... that he can jump from genre to genre as he's always done* [yes7]. The use of the verbs *prove* and *show* followed by a *that*-clause or a substantive subordinate clause is typically associated with the category Endorse. However, in order to get more elaborated results on this type, probably a different genre should be analysed, as the length of the corpus (unless it is a really large sample), will not probably provide enough and explanatory occurrences.

8. Summary and concluding remarks

This study has been carried out with three main aims in mind. The first aim, testing two aspects of the category of Engagement empirically through manual corpus annotation, has been attained through the performance of two agreement studies based on: 1) the identification of the Engagement spans; and 2) the classification of the subtypes of Engagement in the bilingual texts. Some mismatches between the annotators were recurrent. Regarding the Agreement study 1 (span selection), the main causes of disagreement were the length of the span, that is, how long the Engagement expression is; how to deal with discontinuous spans (those which are broken by a word which does not belong to the same tag); the selection of those parts of the texts which were a description of the plot (and consequently, they would not be analysed); the scope of negation when tagging Deny, more specifically, the introduction of secondary elements

(e.g. objects) in annotations of Deny in Spanish. With respect to the Agreement study 2 (labelling subtypes), the *kappa coefficient* indicated an almost perfect agreement. From the few mismatches found it can be concluded that the subtypes Pronounce and Counter are open to interpretation and may overlap with other categories in a higher degree than the rest of subtypes because these two are more context-dependent.

The second aim was to extract relevant contrastive features of the use of Engagement in English and Spanish in the online book reviews. The third aim was to create a bilingual (comparable) machine-readable annotated corpus with Engagement features in English and Spanish which can serve as the training corpus for machine learning algorithms and to be offered to the scientific community for further research.

These last two aims were achieved through the contrastive annotation of a larger set of bilingual texts by a single annotator. The results of this annotation showed that the use of Engagement categories is formally very similar in both languages, though more significant differences were found in the functional aspect, as shown in the qualitative analysis of the specific occurrences found and the frequency tables.

The main conclusions which can be drawn from the contrastive annotation are the following:

It can be stated that, in general, Counter, Deny and Epistemic are the most typical Engagement resources found in online book reviews, though the English reviews tend to use Counter more than Deny, while the opposite situation is found in the Spanish ones. This may be due to a higher concern of the English reviewers for making some concessions about facts that do not adjust to their opinion, in order to avoid such a direct claim as overtly denying something, which is more accepted in the Spanish texts.

The least used types are those whose function is inserting external sources in the texts, mainly due to the fact that the genre of book reviews focuses on individual's

personal opinion. These categories obtained such a low number of occurrences that both languages can be said to share their results.

Intermediate cases have shown differences across languages, such as the preference of Spanish reviews to make explicit their subjective point of view (Opinion markers) versus the tendency to emphasise the author's voice in the English ones.

As well as showing some tasks which can be easily systematized and automated, the qualitative discussion has outlined the most significant problems for annotating Engagement automatically which can be deduced from the corpus analysed:

1) A high amount of Engagement resources (actually more than expected) are fixed expressions which can be easily automated because they are not context-dependent, and every time the word appears it can be tagged in the same way.

2) The study of the collocations or surrounding words that accompany some specific units can entail great advances in the automatic analysis of complex expressions.

3) A prior syntactic and semantic analysis of the corpus would help to improve the level of precision of an automatic analysis.

4) Although a larger corpus would reveal a wider variety of Engagement resources, those found in the present texts can be said to be typical of the genre.

As pointed out repeatedly throughout the discussion, this research has involved important limitations regarding the length of the corpus. Therefore, future research lines may focus on the creation of a larger corpus, which is thought to improve the results considerably. Nevertheless, not only a quantitative difference is recommended, but also the specific analysis of each category or even expressions could solve some significant problems of interpretation and annotation, such as categories' boundaries and fixed patterns.

9. Acknowledgements

I would like to thank my supervisor, Dr Julia Lavid, who has given me valuable guidance and advice at different phases of the work: from the initial planning and design to the development of the methodology to be used, including the guidance with useful computational resources such as SPSS for the statistical data analysis. She also revised the final manuscript thoroughly and provided substantial input to improve its final presentation.

I'm also grateful to Dr Marta Carretero for her assistance and cooperation in the initial annotation tasks and her always kind disposition to help.

Obviously, possible mistakes, errors of judgement, misrepresentations of the facts or other failings of this work are only mine. It is my hope that the current study, though modest and incipient, will pave the way for further stimulating insights and future work into the cross-linguistic analysis and annotation of Engagement.

References

Aijmer. K. (1985). Just. In: S. Bäckman and G. Kjellmer (eds.), *Papers on language and literature presented to Alvar Ellegård and Erik Frykman,*. Gothenburg: Gothenburg Studies in English 60. 1-10.

Bakhtin, M.M. (1981). *The Dialogic Imagination.* C. Emerson and M. Holquist (trans.). Austin: University of Texas Press.

Bednarek, M. (2009). Language patterns and ATTITUDE. *Functions of Language*, 16(2), 165-92.

Biber, D. and E. Finegan. (1989). Styles of stance in English: lexical and grammatical marking of evidentiality and affect. *Text*, 9(1) (Special Issue on the pragmatics of affect). 93-124.

Bloom, K., N. Garg and S. Argamon. (2007). Extracting Appraisal Expressions. *Proceedings of NAACL HLT*. Rochester. 308–315

Bloom, K., S. Stein and S. Argamon. (2007). Appraisal Extraction for News Opinion Analysis at NTCIR-6. *Proceedings of NTCIR-6 Workshop Meeting*. Tokyo, Japan.

Brooke, J., M. Tofiloski and M. Taboada. (2009). Cross-Linguistics Sentiment Analysis: From English to Spanish. In *Proceedings of RANLP 2009, Recent Advances in Natural Language Processing*. Borovets. 50-54.

Bybee, J. and S. Fleischman. (1995). *Modality in Grammar and Discourse*. Amsterdam: Benjamins.

Carretero, M. and M. Taboada (2009) Contrastive analyses of Evaluation in text: Key issues raised by the application of Appraisal Theory to corpora of consumer-generated product reviews in English and Spanish. *I Congreso Internacional de Lingüística de Corpus*. Murcia, Spain.

——. (2010a) Products, consumers and evaluation: a proposal of solutions to problematic issues of Attitude in English and Spanish consumer reviews. *22nd European Systemic Functional Linguistics Conference and Workshop*. University of Primorska, Slovenia. July 2010.

——. (2010b) The annotation of Appraisal: How Attitude and epistemic modality overlap. *4th International Conference on Modality in English (ModE4)*. Madrid, Spain.

——. (2011). Annotating Appraisal: contrastive issues raised in the analysis of consumer reviews in English and Spanish.*38 th International Systemic Functional Linguistic Congress*. Lisbon, Portugal. July 2011.

——. (to be published). The annotation of Appraisal: How Attitude and Epistemic Modality overlap in English and Spanish consumer reviews. In J.R. Zamorano-Mansilla, E. Domínguez, C. Maíz and M.V. Martínez de la Rosa, *Thinking Modally: English and Contrastive Studies on Modality*. Cambridge Scholars.

Chafe, W. and J. Nichols (eds.). (1986). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, N.J,: Ablex. (Advances in Discourse Processes XX).

Chatterjee, M. (2007). Textual engagements of a different kind?. *ASFLA 2007 proceedings*. Available to download at: http://www.asfla.org.au/category/asfla2007/2007appraisal/

Coffin, C. (1997). Constructing and giving value to the past: An investigation into secondary school history. In F. Christie and J. R. Martin (Eds.), *Genre and Institutions: Social Processes in the Workplace and School*. London: Cassell. 196-230.

Conrad, S. and D. Biber. (2000). Adverbial marking of stance in speech and writing. In Hunston and Thompson (eds,). 56-73.

Crystal, D. (2003). *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell.

Dong, A. (2006). How am I doing? The language of Appraisal in design. In J.S. Gero (ed.), *Design Computing and Cognition*. Dordrecht: Springer. 385-404.

Eggins, S. and D. Slade. (1997). *Analysing Casual Conversation*. London: Cassell.

Halliday, M. A. K. (1984). Language as code and language as behaviour: a systemic-functional interpretation of the nature and ontogenesis of dialogue. In R. Fawcett, M. A. K. Halliday, S. M. Lamb and A. Makkai (eds.), *The Semiotics of Language and Culture Vol 1: Language as Social Semiotic*. London: Pinter. 3-35.

Halliday, M. A. K. and C.M.I.M. Matthiessen. (1985/2004). *An Introduction to Functional Grammar* (3rd ed.). London: Arnold.

Hood, S.E. and J.R. Martin. (2007). 'Invoking attitude: the play of graduation in appraising discourse'. In R. Hasan, C. Matthiessen and J. Webster (eds), *Continuing discourse on language: a functional perspective. Vol 2*. London: Equinox. 739-764.

Hovy, J, and Lavid, J. (2010). 'Towards a science of corpus annotation: a new methodological challenge for Corpus Linguistics', *International Journal of Translation* 22(1): 13-36.

Hunston, S. (2000). Evaluation and the planes of discourse: status and value in persuasive texts. In Hunston and Thompson (eds.). 176-207.

——. (2003). Frame, phrase or function: a comparison of frame semantics and local grammars. In Archer et al (eds.), *Proceedings of the Corpus Linguistics 2003 Conference*. UCREL Technical Papers, vol. 16. Lancaster: UCREL.

——. (2010). *Corpus Approaches to Evaluation. Phraseology and Evaluative Language*. London: Routledge.

Hunston, S. and G. Thompson (eds.). (2000). *Evaluation in text: authorial stance and the construction of discourse*. Oxford [etc.]: Oxford University Press.

Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics,* 20. 341-367.

Iedema, R. (1995). Literacy of Administration (Write it Right Literacy in Industry Research Project – Stage 3). Sydney: Metropolitan East Disadvantaged Schools Program.

Iedema, R., S. Feez and P.R.R. White. (1994). Media Literacy (Write it Right Literacy in Industry Research Project – Stage 2). Sydney: Metropolitan East Disadvantaged Schools Program.

Lavid, J. (2008). CONTRASTES: an Online English-Spanish Textual Database for Contrastive and Translation Learning. In B. Lewandowska-Tomaszick (ed.) Corpus Linguistics, Computer Tools and Applications: A State of the Art. Frankfurt: Peter Lang 431-443.

Love, K. (2006). APPRAISAL in online discussions of literary texts. T*ext and Talk,* 26 (2), 217-244.

Macken-Horarik, M. (2003). APPRAISAL and the special instructiveness of narrative. *Text*, 23 (2), 285-312

Marín-Arrese, J.I. and B. Núñez Perucha. (2006). Evaluation and Engagement in Journalistic Commentary and News Reportage. *Revista Alicantina de Estudios Ingleses*, 19. 225-248.

Martin, J.R. (1992). *English Text: System and Structure*. Philadelphia, Amsterdam: John Benjamins.

——. (1996). Evaluating disruption: symbolising theme in junior secondary narrative, in R. Hasan and G. Williams (eds.) (1996), 124-71.

——. (1997). Register and genre: modelling social context in functional linguistics – narrative genres. In E. R. Pedro (ed.), *Proceedings of First International Conference on Discourse Analysis*. Lisbon: Colibri/Portuguese Linguistics Association, 305-44.

——. (2000). Beyond exchange: Appraisal systems in English. In S. Hunston and G. Thompson (Eds.), *Evaluation in Text: Authorial Distance and the Construction of Discourse*. Oxford: Oxford University Press. 142-175.

——. (2001). Writing history: construing time and value in discourses of the past. In C. Colombi and M. Schleppergrell (eds.), Developing Advanced Literacy in First and Second Languages. Mahwah, N.J.: Erlbaum.

——. (2002). Meaning Beyond the Clause: SFL Perspectives. *Annual Review of Applied Linguistics*, 22, 52-74.

——. (2003). Introduction: Special issue on Appraisal. *Text*, 23 (2), 171-181.

Martin, J.R. and D. Rose. (2003). *Working with Discourse: Meaning beyond the Clause.* London, New York: Continuum.

Martin, J.R. & P.R.R. White (2005) *The Language of Evaluation. Appraisal in English.* New York: Palgrave.

Miller, D. (2004). "...to meet our common challenge": ENGAGEMENT strategies of alignment and alienation in current US international discourse. *Textus, XVIII*(1), 39-62.

Newmeyer, F. (1998). *Language Form and Language Function*. Cambridge, MA: MIT Press.

Niemeier, S. and R. Dirven (eds.). (1997). *The Language of Emotions: Conceptualisation, Expression, and Theoretical Foundation*. Amsterdam: Benjamins.

O'Donnell, M. 2011. UAM Corpus Tool (version 2.8.1), (http://www.wagsoft.com/CorpusTool/ July 2011,).

Ochs, E. and B. Schiefflen. (1989). Language has a heart. *Text*, 9(1) (Special Issue on the pragmatics of affect). 7-25.

Page, R.E. (2003). An analysis of APPRAISAL in childbirth narratives with special consideration of gender and storytelling style. *Text*, 23 (2), 211-237

Painter, C. (2003). Developing attitude: an ontogentic perspective on APPRAISAL. *Text* 23(2). 183-210.

Pang, B. and L. Lee. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2). 1–135

Pang, B., L. Lee and S. Vaithyanathan. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*, Philadelphia, PA, USA.

Paradis, C. (2003). Between epistemic modality and degree: the case of *really*. In Facchinetti, Roberta; Krug, Manfred, and Palmer, Frank (eds.). *Modality in Contemporary English. Topics in English Linguistics 44*. Berlin: Mouton de Gruyter.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. (1985). A *Comprehensive Grammar of the English Language*. New York: Longman.

Read, J., D. Hope and J. Carroll. (2007). Annotating Expressions of Appraisal in English. *Proceedings of the Linguistic Annotation Workshop*, Prague. 93–100.

Somasundaran, S., J. Ruppenhofer and J. Wiebe. (2008). Discourse Level Opinion Relations: An Annotation Study. *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*. Columbus. 129–137.

Sano, M. (2011). 'Reconstructing the English system of Attitude for the application to Japanese: an exploration for the construction of a Japanese Dictionary of Appraisal.' Paper presented at the 38th International Systemic Functional Congress.

Stubbs, M. (1996). Towards a modal grammar of English: a matter of prolonged fieldwork. In M. Stubbs, *Text and Corpus Analysis*. Oxford: Blackwell.

Taboada, M. (2008) SFU Review Corpus [Corpus]. Vancouver: Simon Fraser University. http://www.sfu.ca/~mtaboada/research/SFU_Review_Corpus.html

——. (2009) Comparable corpora for cross-linguistic sentiment analysis. *Conference of the American Association for Corpus Linguistics*. Edmonton, Alberta. October 2009.

Taboada, M. and J. Grieve. (2004). Analyzing appraisal automatically. *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text (AAAI Technical Report SS-04-07)* (pp. 158-161). Stanford University, CA.

Taboada, M., C. Anthony and K. Voll. (2006). Creating semantic orientation dictionaries. *Proceedings of 5th International Conference on Language Resources and Evaluation (LREC)* (pp. 427-432). Genoa, Italy.

Tan, S. (2010). Modelling Engagement in a Web-based Advertising Campaign. *Visual Communication* 9(1).

Thompson, G. (2003). Evaluation: An Introduction. In Hunston and Thompson (eds.) 2003. 1-27.

Thompson, G. and S. Hunston. (2006). System and Corpus: two traditions with a common ground. In Thompson and Hunston (eds.) 2006. 1-14.

Thompson, G. and S. Hunston (eds.). (2006). *System and Corpus: Exploring Connections*. London; Oakville, Conn.: Equinox.

Viera, A.J. and J.M. Garret. (2005). Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*, 37(5), 360-3.

Voloshinov, V.N. (1995). *Marxism and the Philosophy of Language, Bakhtinian Thought – an Introductory Reader*. S. Dentith, L. Matejka and I.R. Titunik (trans.). London: Routledge.

Wang, J. and A. Dong. (2007). How am I doing 2: Computing the language of Appraisal in design. I*nternational Conference on Engineering Design*, Paris.

Wiebe, J., T. Wilson, R. Bruce, M. Bell, and M. Martin. (2004). Learning subjective language. *Computational linguistics*, 30(3). 277–308.

Wiebe, J., T. Wilson and C. Cardie. (2005). Annotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation,* 39(2-3), 165-210.

Wierzbicka, A. (1990). The semantics of emotions: fear and its relatives in English. *Australian Journal of Lingusitics*,10(2) (Special Issue on the Semantics of Emotions). 359-75.

Wilson, T. and J. Wiebe. (2005). Annotating attributions and private states. *Proceedings of the Workshop on Frontiers in Corpus Annotation II: Pie in the Sky*. Ann Arbor. 53-60.

White, P.R.R. (1998). *Telling Media Tales: The News Story as Rhetoric.* Unpublished Ph.D. dissertation. Sydney: University of Sydney.

——. (2001). *An introductory tour through Appraisal Theory* [on line]. (www.grammatics.com/Appraisal/AppraisalOutline/AppraisalOutline.doc, August 2011)

——. (2002) Appraisal. In J. Verschueren et al. (eds.). *Handbook of pragmatics*, 1–27. Amsterdam: Benjamins.

——. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text*, 23 (2), 259–284.

——. (2004). Subjectivity, Evaluation and Point of View in Media Discourse. In C. Coffin, A.Hewings and K. O'Halloran (eds), *Applying English Grammar*. London: Hodder Arnold.

——. (2011). Language as dialog – Engagement and the analysis of interpersonal positioning in student writing and media commentary (Workshop). *38th International Systemic Functional Congress (ISFC38): Negotiating difference: languages, metalanguages, modalities, cultures*. Lisbon, Portugal.

Whitelaw, C. et al. (2005). Using appraisal taxonomies for sentiment analysis. In *Proceedings of MCLC-05, the 2nd Midwest Computational Linguistic Colloquium*, Columbus.

List of tables

List of figures

Appendix 1
Annotation sheets for Agreement study 1 (Span selection)

Instructions:

1- Read the complete text.
2- Identify examples of Engagement, without assigning them a specific category.
3- Place the examples in the box on the right.

Book review "no1_1"

		Annotator #
1	La saga de Ender comenzo con un libro impresionante de principio a fin como es el juego de Ender	
2	y despues fue bajando la calidad hasta llegar a su final en un libro olvidable que parece más bien el propio card hablando con sigo mismo sobre filosofia y ética.	
3	La decrepitud de ender parece coincidir con la decrepitud de su saga,	
4	de hecho el pesonaje de ender sobra, sobre por todas partes	
5	y parece metido en la historia con calzador junto con algunos personajes del primer libro pretendiendo ser como el primer libro pero que parecen calcomanias planas sin niguna gracia	
6	tarde cuatro intentos en acabarlo	
7	y lo hice solo por las 5 veces que me he leido el juego de ender.	
8	Vacio, completamente vacio sin mas objetivo que acabar con ender de la forma más insulsa posible	

Book review "no1_11"

		Annotator #
1	Es lo peor que he leído en mucho tiempo	
2	(bueno, la verdad, es que ni siquiera pude terminar de digerirlo).	
3	Es sospechoso que saliera justo antes del día del libro por lo que no puedo dejar de imaginarme que le	

	encargaron que escribiera algo para sacarlo a tiempo y ella no tuvo vergüenza torera y se plantó con unos escritos infumables, pretenciosos, con frases tópicas y pensamientos filosóficos de salón de belleza de capital.	
4	Parece mentira que la editorial se lo aceptara, aunque imagino que querían aprovechar el tirón mediático que tiene esta muchacha.	
5	Realmente prometía con Amor, curiosidad, prozac y dudas	
6	pero luego intentó vivir de rentas	
7	y en este mundillo: renovarse o morir.	
8	En fin, una chapuza indigesta que te cabrea (porque tampoco es que sea barato).	

Book review "no1_12"

		Annotator #
1	Tardé muchísimo en leerme este libro.	
2	De hecho, es el libro que más he tardado en leer.	
3	Que conste que he leído muchos libros, muchos de escritura más densa que la de Gala,	
4	y me los terminé en un santiamén.	
5	¿Por qué?	
6	A lo mejor, porque sus autores van a lo que van.	
7	Porque saben que necesitan un ritmo narrativo, vital en toda obra literaria, se diga lo que se diga.	
8	"El manuscrito carmesí" trata sobre el último sultán, Boabdil, el desventuradillo, y su formación para llegar a ser príncipe, sus desamores, etc., bajo un trasfondo histórico muy cuidado.	
9	A mi, personalmente, me decepcionó esta historia sobremanera.	
10	Esperaba una trama narrativa y un ritmo que no existen,	
11	y me encontré con un poeta que intenta hacer narrativa de un conjunto de frases bonitas.	
12	Para mí, el escritor es aquel que, a partir de una idea absurda o genial, es capaz de atrapar al lector mediante un empleo de buena literatura.	
13	Creo que Antonio Gala se dejó llevar demasiado con este libro (600 páginas),	
14	y lo convirtió en su particular meca de una visión histórica de Granada, por aquellos tiempos, mezclada, cómo no, con sus obsesiones sexuales, que de vez en cuando asoman, y de qué manera.	

15	Hay escritores, como Faulkner, que pueden hacerte no entender nada de lo que lees, pero atraparte con intensidad en la lectura.	
16	Y hay escritores, como Gala, que emplean sus obsesiones como artefacto de sus pretensiones para escribir una novela que no es una novela, sino un extensísimo tratado de historia que podría pasar por obra de referencia (si tuviera un buen sumario).	
17	No es suficiente con una prosa poética, soñadora, melancólica, iluminadora, cuando se deja de lado el discurso narrativo, abandonado al tedio.	
18	Leí "el manuscrito carmesí" impulsado por la comercial idea de que era premio planeta.	
19	Desde entonces, recelo de este premio,	
20	y dudo de la veracidad con la que el jurado aplica sus valoraciones.	

Book review "yes4_17"

		Annotator #
1	Sampedro ya me conquistó hace años con "La sonrisa etrusca", cuánta ternura en esa novela,	
2	pero no es de ella de la que quiero hablaros en esta opinión.	
3	Sino de "El amante lesbiano" mi último descubrimiento.	
4	Lo cierto es que han pasado muchos años desde que descubrí a Sampedro	
5	y no había leído nada más de él,	
6	me gustó la primera que nombro,	
7	pero nada más,	
8	pero el otro día me acerqué al la biblioteca del barrio a apuntarme (que no está el horno para comprar tantos libros, ahora compraré sólo aquellos que me gusten)	
9	y al verlo entre las estanterías decidí traérmelo a casa,	
10	y lo cierto es que me ha sorprendido gratamente.	
11	El libro empieza con dos citas a modo de presentación:	
12	"Entremos más adentro en la espesura" de San Juan de la Cruz	
13	y "Ama y haz lo que quieras" de San Agustín.	
14	Lo cierto es que después de leerlo podría decir que estas dos simples frases resumen a la perfección la	

	esencia del libro.	
15	La historia empieza cuando Mario al salir de su rutinaria visita al cardiólogo coge un taxi	
16	y sin darse cuenta aparece en otro lugar, en "Las Afueras", un lugar desconocido donde todo es posible, y donde nuestro protagonista se encuentra con su pasado, con su familia, con verdades, con secretos guardados durante años y lo más importante consigo mismo.	
17	El libro no es más que eso,	
18	a ver cómo lo definiría,	
19	pues lo haría como una historia en la que se nos presenta a un personaje que ha vivido una vida soñada por otros, que no es la que él hubiese deseado, una vida encorsetada, marcada por los clichés que su familia, la sociedad, las costumbres y él mismo le acaban imponiendo.	
20	Sabe que no es feliz	
21	pero se deja llevar	
22	y no hace nada para cambiarlo.	
23	Pero entonces aparece alguien que "trastocará" todos sus esquemas y que le hará descubrir al verdadero Mario, al que nunca se atrevió a ser.	
24	La historia me enganchó desde el principio,	
25	en las primeras líneas apunta como nuestros padres, y nuestra familia más cercana en general, nos influyen desde muy pequeños en nuestro modo de ver la vida, como nos van condicionando y nos van moldeando a su antojo, ya sea intencionadamente o no, y como desde muy temprana edad eso nos va marcando, como sin apenas darnos cuenta todo ello va dando forma al adulto en el que nos convertimos.	
26	Luego es cierto que unos se apartan de la norma, de lo establecido,	
27	son aquellos que deciden vivir sus propias vidas y deciden hacerlo en plenitud	
28	y luego están los que se conforman, o los que son felices de esa forma.	
29	Eso más que decirlo el libro es una conclusión mía al leerlo, es uno de mis múltiples puntos de vista sobre este libro.	
30	Si bien es cierto que no todas las partes del mismo me gustan, porque cuando intentan que Mario se acepte tal cual es, y cuando le están "enseñando" a asumir su género femenino, se le dice que debe usar braguitas de raso, medias finas, zapatos de tacón, minifalda, llevar el pelo largo…en fin convertirse en un estereotipo,	
31	y digo yo,	
32	o más bien me pregunto,	
33	¿es que no se puede una sentir mujer e igual de sexy con unas braguitas de algodón, una camiseta, unos vaqueros y unas deportivas,	

34	o es que necesariamente el género viene marcado por llevar una determinada indumentaria.	
35	Me sorprende porque un libro que pretende salirse de los estereotipos y ofrecer una visión distinta de la sexualidad, una apertura de mente,	
36	por otro lado define a la mujer siguiendo unos clichés algo obsoletos.	
37	Qué el hábito no hace al monje,	
38	ni a la inversa.	
39	Y eso no es todo, porque para aceptarse además de vestirse con estas prendas toma clases de costura, aprende a realizar tareas domésticas, en fin, "cosas de chica" según Sampedro,	
40	o tal vez yo lo entendí mal, pero esa parte del libro no me gusta.	
41	Por ejemplo en una parte del libro se dice: "Y ese deseo te lo voy a arrancar de cuajo.	
42	Haré que desees de otro modo. Que ames poseída, según tu género",	
43	vamos que para ser de género femenino hay que desear de forma sumisa.	
44	Por lo demás para mí la grandeza del libro reside en como cuenta, como define esos sentimientos, esos instintos, esa pasión,	
45	y es que ya lo comprobé en la anterior obra (La Sonrisa Etrusca)	
46	Sampedro es un genio describiendo sensaciones.	
47	Tiene algunas escenas de contenido erótico y de contenido sexual, algunas más explícitas, otras menos, pero todas ellas llenas de pasión "a lo bestia", de puro deseo, de puro instinto	
48	y me gusta como lo describe,	
49	como lo narra,	
50	casi consigue hacernos sentir ese deseo.	
51	Cuando explica como es el sexo cuando hay un amor pasional, turbulento, cuando sientes adoración por la otra persona, cuando todo es DESEO, puro instinto animal	
52	(que lo de animal lo digo yo, que no Sampedro, porque así lo entiendo)	
53	y es que el deseo, la pasión, no puede ser de otra forma.	
54	Había partes con las que sentí muy identificada,	
55	cuando amas tanto a alguien, cuando hay tanto deseo que los cuerpos se funden, se entremezclan, se confunden el uno con el otro es una mezcla de olores, sabores, sensaciones.	
56	Cuando todo es posible, y permitido,	
57	cuando uno "encaja" literalmente con el otro.	
58	Eso el autor lo define a la perfección,	

59	por fuerza que debe haberlo sentido en algún momento, porque no hay palabras.	
60	Define el acto sexual no como sólo el acto de la cópula, como el simple contacto de órganos genitales,	
61	lo define como algo mucho más SUBLIME	
62	(sé que estos días repito la palabra, pero es la que lo define),	
63	como el contacto más allá de lo carnal	
64	(y no estoy hablando de un amor místico ni de puro sentimiento) voy más allá, o va más allá,	
65	hablo de la unión perfecta, esas ansias de que el otro esté dentro de ti, ese deseo que no se puede contener, ese querer vaciarte y llenarte del otro, ese querer estar dentro o que esté dentro,	
66	en fin, que podría ser más explícita, pero seguro que todos me habéis entendido.	
67	Os dejo alguna frase que describe lo que quiero decir:	
68	"Parece de ahora mismo aquella revelación de su carne que me alzó en lo más alto, porque mi obsesión no descansa, no deja de estar siendo".	
69	"Todo mi cuerpo se agolpa en mi boca, donde sus labios y su lengua y sus dientes me invaden, me mordisquean, me gozan, me electrizan, me poseen…	
70	Cierro los ojos: no hay más mundo que ese beso y mi ser volcado en él".	
71	"Poco a poco …	
72	(omito nombres aquí por si alguien lo lee, no desvelar nada)	
73	que ha descansado su cuerpo sobre mis espaldas arañadas, retira su presencia de mi hondura, con mimo, con reticencia.	
74	Lo que me duele entonces es la ausencia progresiva, el resbalar hacia ese vacío que corta nuestra unión umbilical…"	
75	Podría dejaros muchas frases más, pero mejor os invito a que leáis el libro,	
76	desde mi punto de vista no tiene desperdicio,	
77	eso sí, si tenéis algún tipo de prejuicio en cuestiones sexuales no lo hagáis porque os puede chocar.	
78	En el libro propone mil formas de amor, de amar y ser amado, formas distintas, diferentes, pero todas ellas igual de válidas,	
79	es según la contraportada de mi edición "Una fantasía erótica ajena a la represiva educación sexual contranatura todavía imperante.	
80	Una indagación en las múltiples variantes cerebro-genitales del amor".	
81	Una vez dije a alguien (y esto creo que lo dejé por algún test) "Ojalá pudieras ser yo por unos momentos para que sintieras lo maravilloso que es tenerte dentro de mí".	

		Annotator #
82	Y lo dije porque lo sentí así,	
83	pues esto y mil cosas más me hizo recordar esta historia.	
84	Un saludo y feliz y erótica noche, jeje.	

Book review "yes5_2"

		Annotator #
1	Por fin un libro infantil pensado para niños (es decir personas pequeñas) y no para idiotas bajitos.	
2	Me explico,	
3	por alguna extraña razon se tiende a tratar a los niños como si fueran imbéciles	
4	y esto es, salvo honrosas excepciones extensible a la literatura infantil.	
5	Harry Potter y la piedra filosofal (y todos los demás de la serie) es un libro delicioso que resulta entretenido a cualquier edad.	
6	Consigue que un padre que se lo lea a sus hijos se lo pase tan bien como ellos	
7	(incluso sin leerselo a nadie a mi me encantó)	
8	Por si alguien no ha oido aun hablar de el (unas vacaciones en el polo norte quizas) Harry es un niño huerfano que vive con sus tíos, que por supuesto son odiosos hasta que un dia descubre que es un mago y que va a ir a estudiar a la mejor escuela de magia del mundo.	
9	contado de esta forma tal vez no parezca demasiado atractivo,	
10	un libro de aventurillas con malos malísimos y buenos buenísimos.	
11	Sin embargo con todo eso los personajes son creibles	
12	y el argumento está muy bien estructurado.	
13	Además partiendo de una muy bien documentada base mitológica y de la sabiduría popular sobre brujas, la escritora describe el colegio de magia y la vida de los propios magos de forma que resulta fácil imaginar que en realidad existen.	
14	Incluso inventa un deporte, el quidich, que es algo así como el futbol para nosotros, al que si no fuera porque aún no he encontrado escobas voladoras jugaría encantada.	

Book review "no7"

		Annotator #
1	I bought and read "A Painted House" as soon as it came out as I do with all John Grisham books.	
2	I was interested to read how he treated a story that did not revolve around a legal case.	
3	He did not treat it well,	
4	and I will not buy another one of his books unless it is a legal thriller like his others.	
5	The setting of the story seemed interesting enough	
6	--a rural Arkansas farming town during the Korean War.	
7	The story takes place over the length of a cotton-picking season.	
8	It is an observation of the picking season as told by a seven year old son of a sharecropper, Luke.	
9	Luke's character is very well developed,	
10	and his is the only character that is.	
11	There are many other characters in the story,	
12	but the reader is never given enough information to really care about them.	
13	The weak character definition made for a long and boring read, as I really didn't care what became of most of the characters in the story.	
14	The book was a very long read.	
15	The one redeeming quality was that it was very descriptive.	
16	I had clear visuals of what the characters and the town looked like.	
17	The descriptive nature of the story, however, was not enough to make this an interesting story.	
18	I am not even going to talk about the plot because I really don't see a point.	
19	To paraphrase Nike, "Just Skip It!"	

Book review "no21"

		Annotator #
1	I enjoyed reading Dean Koontz's book "Tick Tock" very much.	
2	What I didn't enjoy doing, however, was listening to this book on audiocassette.	
3	I had never before listened to a book on tape, so this was a first.	

		Annotator #
4	Just because I give my opinion of this book on audiocassette a bad review, it does not mean I'm bashing audiocassette books in general.	
5	I'm sure this book on tape has got to be one of the worse.	
6	I listened to the taped version a week after finishing the book, so the book was still fresh in my mind.	
7	The person who read the book complete stripped the personality out of the characters by his interpretations of the voices.	
8	As I read the book, I got the impression of the main female character to be smart, clever, fun and sporadic.	
9	The reader, however, made the female's voice whiny and obviously fake.	
10	His reading of the book put it into a whole new perspective,	
11	and not a likeable one either.	
12	He instead made the whole book sound ridiculous.	
13	I am a very big fan of Dean Koontz books, therefore I would hate for someone who has never read or listened to a Dean Koontz book before to pick "Tick Tock" on audiocassette as their first choice.	
14	Should that happen, there is a very likely chance that the person would never again try any piece written by Dean Koontz.	
15	I suggest that if you plan on listening to Dean Koontz's "Tick Tock" on audiocassette that you read the book first, for it is much different even though it is exactly the same.	
16	PS--Please keep in mind that my ratings are not for the book itself,	
17	but for the book on audiocassette	

Book review "no25"

		Annotator #
1	A teacher of mine recommended this book to me, saying that all her friends loved it and couldn't put it down, so when it came time to do a critique in her class, I chose this book.	
2	Unfortunately, Mr. Pears book was not all that it was cracked up to be.	
3	The are too many problems that drag this book down, preventing it from being the masterpiece that many are claiming it to be.	
4	First, there is the narration style:	
5	4 different narrators telling the events the way they saw them, each fingering a different suspect.	

6	If the plot had been more gripping, more intense, this would have worked perfectly.	
7	Unfortunately, it wasn't,	
8	and by the beginning of the third narrator, I did not feel like reader the remaining two parts.	
9	Then there is the narrators themselves.	
10	Again, this could have been excellent, save the fact that 2 of the narrartors are so arrogant and full of themselves that it made it hard for me to want to keep reading it.	
11	Finally there was all the historical detail.	
12	Mr. Pears obviously spent a lot of time researching to make his book as accurate as possible, including his writing style.	
13	I commend him for that, however the details he gives are so numerous that they begin to over shadow the plot.	
14	If you like the history behind what each narrator does (i.e. the development of science for Da Cola's tale) then you will enjoy it.	
15	If you find any one of the subjects offered (say cryptography) then it will hinder your enjoyment immensly.	
16	I did enjoy reading some of this book, namely Da Cola's tale, and the final tale,	
17	but the two tales in the middle dragged too much for me to be able to really recommend this book.	
18	It is another case of hype building up to nothing	

Book review "yes9"

		Annotator #
1	Once upon a time, Stephen King wrote a fairy tale.	
2	The eyes of the Dragon is a truly mesmorizing and fascinating story.	
3	A fairy tale for adults,	
4	but I think this could be for older children also.	
5	It doesn't have the customary violence and language that most Stephen King novels have,	
6	and it isn't so intense that it will frighten you.	
7	"Why then would I want to read it?" You may be asking yourself.	
8	Well you want to read it because it is a great novel.	
9	The story is set in an ancient kingdom called Delain.	

10	The story centers around two young boys, princes, the son of the King.	
11	The two brothers are Peter and Thomas.	
12	Peter, the oldest, is in line to be the next king.	
13	Peter is the perfect child.	
14	Everyone loves Peter,	
15	he is handsome, intelligent, and caring.	
16	Thomas, on the other hand is none of these.	
17	He is a lonely little boy disliked and ignored by all, even his father.	
18	Thomas is ignored by all, that is, except for one person:	
19	the kings magician Flagg.	
20	Flagg has been in the kingdom for years; as long as anyone can remember.	
21	Before Peter can become King, his father is murdered and he is framed for it.	
22	Since Peter is royalty, he can't be killed, so he is sent up into a tower to live out the rest of his life.	
23	The only people he sees are his guards.	
24	The only thing Peter has to pass his time is his dollhouse that he played with when he was little,	
25	the only other convience he was allowed was a napkin to eat with,	
26	these he specially asked for.	
27	Thomas is now the King.	
28	However, Thomas is young and needs someone to help rule the kingdom.	
29	Who would that be?	
30	Why Flagg of course.	
31	Thomas knows a secret that could destroy him and is slowly driving him mad.	
32	This secret could not only destroy him,	
33	it would definitely destroy Flagg.	
34	What is this secret?	
35	Read the story:	
36	find out.	
37	Will Peter ever get out,	
38	or will he die in the tower?	
39	Once again read the story find out.	
40	I love Stephen King.	

41	The scarier the story the better.	
42	However, this is just about my favorite Stephen King story.	
43	It isn't scary,	
44	but it is enthralling.	
45	For the more avid King readers out there;	
46	check out the reference to this story in the Dark Tower series,	
47	Roland mentions some of the characters in this story that passed through his town when he was young.	
48	For anyone out there who just wants to read a good story, check out this book.	
49	It isn't violent	
50	(well there are a few parts,	
51	but not usual King violence).	
52	It isn't vulgar or sexual.	
53	It is just a great read.	
54	I recommend this to anyone and everyone.	

Book review "yes11"

		Annotator #
1	Murder, Mutants, Abuse, Torture, Doctors, Scientists, Secrets, Gun Shots, Betrayal, Sex, Love....	
2	these are just some of what you'll find in James Patterson's book, When the Wind Blows.	
3	General Information	
4	When the Wind Blows is written by author James Patterson.	
5	It was published in 1998 by Little, Brown And Company.	
6	The hard cover version that I read had a total of 416 pages	
7	and unlike the other two books I've read by Patterson, I could actually find places where I could force myself to stop reading for the night.	
8	I've only recently started reading books he's written	
9	and confess that I'm already a huge fan.	
10	ISBN: 0-316-69332-4	
11	The Story	

12	Dr. Frances O'Neill, a well loved vet in Colorado, was working as usual in her place of business - the Inn Patient.	
13	Frannie as she's called by her friends, has had her share of heart break and loss.	
14	Her husband David, a doctor at the local hospital was murdered less than 2 years earlier.	
15	Frannie is a real softy which is what many of her clients and their owners loved about her.	
16	So there she was, minding her own business when one day this gorgeous hunk of man walks in and rents her cabin [which is being the Inn Patient].	
17	His name is Kit Harrison	
18	and while she can't stand him and at the same time she can't help but like him - just a little.	
19	It takes quite awhile for Frannie to find out but Kit is not who she thinks he is.	
20	Then, Frannie's good friends husband is killed.	
21	They say it was an accident, that he hit his head in the pool;	
22	but it's very hard for those who knew him to believe.	
23	After much crying and several drinks, Frannie drives back home to the Inn Patient where she not only works, but has lived ever since David's death.	
24	On her drive home she thinks she sees something.	
25	She stops the car.	
26	Wait.	
27	What is that?	
28	Oh my God!	
29	It looks like a human girl.	
30	A beautiful little human girl - no more than 12 years old.	
31	But she has wings...	
32	and oh my God!	
33	She can fly!	
34	Frannie tries to gain her composure	
35	and then resumes the drive home.	
36	By the time she reaches the Inn Patient she's convinced herself that she couldn't have really seen a girl with wings.	
37	It just wasn't possible.	
38	It must have been the alcohol	

39	...... right?	
40	That's all I'll give you for the story.	
41	Sorry, but if I give you more than that I just may ruin the book for you.	
42	That's how Patterson writes.	
43	And in case you're wondering all of this happens in the first few chapters of the book!	
44	About the writing	
45	Like every other book I've read by Patterson, When the Wind Blows was thrilling and very hard to put down - making it very easy to read.	
46	Patterson captivates the reader from page one	
47	and if you think you'll find a stopping place at the end of a chapter - think again!	
48	Unlike the last book I read by Patterson Cradle and All, I was actually able to put down When the Wind Blows so I could get some sleep.	
49	It wasn't easy	
50	but I did manage.	
51	In the beginning of the book is an Author's Note.	
52	In this note, Patterson states that what he writes in When the Wind Blows is much closer to the truth than even he thought it would be.	
53	While in part, I have a hard time believing that - there's a part of me that knows there must be some truth to the matter.	
54	Overall	
55	If you're looking for an exciting book of fiction and science and doctors interest you, I recommend you give When the Wind Blows a try.	
56	There are some graphic parts which made me mad to read,	
57	but I don't believe it would be too much for anyone to handle.	
58	Thanks for reading ~	
59	Melissa	

Appendix 2
Annotation sheet for Agreement study 2 (Labelling Subtypes)

Instructions:

1- Read the complete text.
2- Focus on the examples of Engagement and assign them a specific category (one from the most delicate level: EP, EV, PQ, O, ACK, DIS, DE, COUN, CON, PR, END)
3- Place each example with the corresponding tag in the box on the right.

Book review "no1_1"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	La saga de Ender comenzo con un libro impresionante de principio a fin como es el juego de Ender				
2	y despues fue bajando la calidad hasta llegar a su final en un libro olvidable que **parece más bien el propio card** hablando con sigo mismo sobre filosofia y ética.				
3	La decrepitud de ender **parece coincidir** con la decrepitud de su saga,				
4	**de hecho** el pesonaje de ender sobra, sobre por todas partes				
5	y **parece metido** en la historia con calzador junto con algunos personajes del primer libro pretendiendo ser como el primer libro **pero** que **parecen calcomanias** planas **sin nniguna gracia**				
6	tarde cuatro intentos en acabarlo				
7	y lo hice **solo** por las 5 veces que me he leido el juego de ender.				
8	Vacio, completamente vacio **sin mas objetivo que** acabar con ender de la forma más insulsa posible				

Book review "no1_11"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	Es lo peor que he leído en mucho tiempo				
2	(bueno, **la verdad,** es que **ni siquiera** pude terminar de digerirlo).				
3	Es **sospechoso** que saliera justo antes del día del libro por lo que **no puedo dejar de imaginarme** que le encargaron que escribiera algo para sacarlo a tiempo y ella **no tuvo** vergüenza torera y se plantó con unos escritos infumables, pretenciosos, con frases tópicas y pensamientos filosóficos de salón de belleza de capital.				
4	**Parece mentira** que la editorial se lo aceptara, **aunque** **imagino** que querían aprovechar el tirón mediático que tiene esta muchacha.				
5	**Realmente** prometía con Amor, curiosidad, prozac y dudas				
6	**pero** luego intentó vivir de rentas				
7	y en este mundillo: renovarse o morir.				
8	En fin, una chapuza indigesta que te cabrea (porque **tampoco** es que sea barato).				

Book review "no1_12"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	Tardé muchísimo en leerme este libro.				
2	**De hecho**, es el libro que más he tardado en leer.				
3	**Que conste** que he leído muchos libros, muchos de escritura más densa que la de Gala,				
4	y me los terminé en un santiamén.				

5	**¿Por qué?**				
6	**A lo mejor**, porque sus autores van a lo que van.				
7	Porque **saben** que necesitan un ritmo narrativo, vital en toda obra literaria, **se diga lo que se diga**.				
8	"El manuscrito carmesí" trata sobre el último sultán, Boabdil, el desventuradillo, y su formación para llegar a ser príncipe, sus desamores, etc., bajo un trasfondo histórico muy cuidado.				
9	A mi, **personalmente**, me decepcionó esta historia sobremanera.				
10	Esperaba una trama narrativa y un ritmo que **no existen**,				
11	y me encontré con un poeta que intenta hacer narrativa de un conjunto de frases bonitas.				
12	**Para mí,** el escritor es aquel que, a partir de una idea absurda o genial, es capaz de atrapar al lector mediante un empleo de buena literatura.				
13	**Creo** que Antonio Gala se dejó llevar demasiado con este libro (600 páginas),				
14	y lo convirtió en su particular meca de una visión histórica de Granada, por aquellos tiempos, mezclada, **cómo no**, con sus obsesiones sexuales, que de vez en cuando asoman, y de qué manera.				
15	Hay escritores, como Faulkner, que pueden hacerte **no entender nada** de lo que lees, **pero** atraparte con intensidad en la lectura.				
16	Y hay escritores, como Gala, que emplean sus obsesiones como artefacto de sus pretensiones para escribir una novela que **no es** una novela, **sino** un extensísimo tratado de historia que podría pasar por obra de referencia (si tuviera un buen sumario).				
17	**No es** suficiente con una prosa poética, soñadora, melancólica, iluminadora, cuando se deja de lado el discurso narrativo, abandonado al tedio.				
18	Leí "el manuscrito carmesí" impulsado por la comercial idea de que era premio planeta.				
19	Desde entonces, **recelo** de este premio,				

20	y **dudo** de la veracidad con la que el jurado aplica sus valoraciones.				

Book review "yes4_17"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	Sampedro ya me conquistó hace años con "La sonrisa etrusca", cuánta ternura en esa novela,				
2	**pero no es** de ella de la que quiero hablaros en esta opinión.				
3	**Sino** de "El amante lesbiano" mi último descubrimiento.				
4	**Lo cierto es** que han pasado muchos años desde que descubrí a Sampedro				
5	y **no había leído** nada más de él,				
6	me gustó la primera que nombro,				
7	**pero nada más,**				
8	**pero** el otro día me acerqué al la biblioteca del barrio a apuntarme (que **no está** el horno para comprar tantos libros, ahora compraré **sólo** aquellos que me gusten)				
9	y al verlo entre las estanterías decidí traérmelo a casa,				
10	y **lo cierto es** que me **ha sorprendido** gratamente.				
11	El libro empieza con dos citas a modo de presentación:				
12	"Entremos más adentro en la espesura" de San Juan de la Cruz				
13	y "Ama y haz lo que quieras" de San Agustín.				
14	**Lo cierto es** que después de leerlo **podría decir** que estas dos simples frases resumen a la perfección la esencia del libro.				
15	La historia empieza cuando Mario al salir de su rutinaria visita al cardiólogo coge un taxi				

16	y sin darse cuenta aparece en otro lugar, en "Las Afueras", un lugar desconocido donde todo es posible, y donde nuestro protagonista se encuentra con su pasado, con su familia, con verdades, con secretos guardados durante años y lo más importante consigo mismo.				
17	El libro **no es más que** eso,				
18	**a ver cómo lo definiría**,				
19	pues lo **haría** como una historia en la que se nos presenta a un personaje que ha vivido una vida soñada por otros, que **no es** la que él hubiese deseado, una vida encorsetada, marcada por los clichés que su familia, la sociedad, las costumbres y él mismo le acaban imponiendo.				
20	Sabe que no es feliz				
21	pero se deja llevar				
22	y no hace nada para cambiarlo.				
23	Pero entonces aparece alguien que "trastocará" todos sus esquemas y que le hará descubrir al verdadero Mario, al que nunca se atrevió a ser.				
24	La historia me enganchó desde el principio,				
25	en las primeras líneas **apunta** como nuestros padres, y nuestra familia más cercana en general, nos influyen desde muy pequeños en nuestro modo de ver la vida, como nos van condicionando y nos van moldeando a su antojo, **ya sea intencionadamente o no**, y como desde muy temprana edad eso nos va marcando, como **sin apenas darnos cuenta** todo ello va dando forma al adulto en el que nos convertimos.				
26	Luego **es cierto** que unos se apartan de la norma, de lo establecido,				
27	son aquellos que deciden vivir sus propias vidas y deciden hacerlo en plenitud				
28	y luego están los que se conforman, o los que son felices de esa forma.				
29	Eso **más que** decirlo el libro es **una conclusión mía** al leerlo, es **uno de mis múltiples puntos de vista** sobre este libro.				

30	**Si bien es cierto** que **no todas** las partes del mismo me gustan, porque cuando intentan que Mario se acepte tal cual es, y cuando le están "enseñando" a asumir su género femenino, se le dice que debe usar braguitas de raso, medias finas, zapatos de tacón, minifalda, llevar el pelo largo…en fin convertirse en un estereotipo,				
31	y **digo** yo,				
32	o más bien **me pregunto**,				
33	**¿es que no** se puede una sentir mujer e igual de sexy con unas braguitas de algodón, una camiseta, unos vaqueros y unas deportivas,				
34	**o es que necesariamente** el género viene marcado por llevar una determinada indumentaria.				
35	Me **sorprende** porque un libro que pretende salirse de los estereotipos y ofrecer una visión distinta de la sexualidad, una apertura de mente,				
36	por otro lado **define** a la mujer siguiendo unos clichés algo obsoletos.				
37	Qué el hábito **no hace** al monje,				
38	**ni** a la inversa.				
39	Y eso **no es** todo, porque para aceptarse además de vestirse con estas prendas toma clases de costura, aprende a realizar tareas domésticas, en fin, "cosas de chica" **según Sampedro**,				
40	o **tal vez** yo lo entendí mal, **pero** esa parte del libro **no me gusta**.				
41	Por ejemplo en una parte del libro se dice: "Y ese deseo te lo voy a arrancar de cuajo.				
42	Haré que desees de otro modo. Que ames poseída, según tu género",				
43	**vamos que** para ser de género femenino hay que desear de forma sumisa.				
44	**Por lo demás para mí** la grandeza del libro reside en como cuenta,				

	como define esos sentimientos, esos instintos, esa pasión,				
45	y **es que** ya lo **comprobé** en la anterior obra (La Sonrisa Estrusca)				
46	Sampedro es un genio describiendo sensaciones.				
47	Tiene algunas escenas de contenido erótico y de contenido sexual, algunas más explícitas, otras menos, **pero** todas ellas llenas de pasión "a lo bestia", de puro deseo, de puro instinto				
48	y me gusta como lo describe,				
49	como lo narra,				
50	**casi** consigue hacernos sentir ese deseo.				
51	Cuando explica como es el sexo cuando hay un amor pasional, turbulento, cuando sientes adoración por la otra persona, cuando todo es DESEO, puro instinto animal				
52	(que lo de animal **lo digo yo**, **que no Sampedro**, porque así lo **entiendo**)				
53	y **es que** el deseo, la pasión, **no puede** ser de otra forma.				
54	Había partes con las que sentí muy identificada,				
55	cuando amas tanto a alguien, cuando hay tanto deseo que los cuerpos se funden, se entremezclan, se confunden el uno con el otro es una mezcla de olores, sabores, sensaciones.				
56	Cuando todo es posible, y permitido,				
57	cuando uno "encaja" literalmente con el otro.				
58	Eso el autor lo **define a la perfección**,				
59	**por fuerza** que **debe haberlo sentido** en algún momento, porque **no hay** palabras.				
60	**Define** el acto sexual **no como sólo** el acto de la cópula, como el simple contacto de órganos genitales,				

61	lo **define** como algo mucho más SUBLIME				
62	(**sé** que estos días repito la palabra, **pero** es la que lo **define**),				
63	como el contacto más allá de lo carnal				
64	(y **no estoy hablando** de un amor místico ni de puro sentimiento) voy más allá, o va más allá,				
65	hablo de la unión perfecta, esas ansias de que el otro esté dentro de ti, ese deseo que **no se puede contener**, ese querer vaciarte y llenarte del otro, ese querer estar dentro o que esté dentro,				
66	en fin, que podría ser más explícita, **pero seguro que** todos me habéis entendido.				
67	Os dejo alguna frase que describe lo que quiero decir:				
68	"Parece de ahora mismo aquella revelación de su carne que me alzó en lo más alto, porque mi obsesión no descansa, no deja de estar siendo".				
69	"Todo mi cuerpo se agolpa en mi boca, donde sus labios y su lengua y sus dientes me invaden, me mordisquean, me gozan, me electrizan, me poseen…				
70	Cierro los ojos: no hay más mundo que ese beso y mi ser volcado en él".				
71	"Poco a poco …				
72	(**omito** nombres aquí por si alguien lo lee, **no desvelar** nada)				
73	que ha descansado su cuerpo sobre mis espaldas arañadas, retira su presencia de mi hondura, con mimo, con reticencia.				
74	Lo que me duele entonces es la ausencia progresiva, el resbalar hacia ese vacío que corta nuestra unión umbilical…"				
75	Podría dejaros muchas frases más, **pero** mejor os invito a que leáis el libro,				
76	**desde mi punto de vista no tiene** desperdicio,				

#		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
77	**eso sí**, **si** tenéis algún tipo de prejuicio en cuestiones sexuales **no lo hagáis** porque os puede chocar.				
78	En el libro propone mil formas de amor, de amar y ser amado, formas distintas, diferentes, **pero** todas ellas igual de válidas,				
79	es **según la contraportada de mi edición** "Una fantasía erótica ajena a la represiva educación sexual contranatura todavía imperante.				
80	Una indagación en las múltiples variantes cerebro-genitales del amor".				
81	Una vez **dije** a alguien (y esto **creo** que lo dejé por algún test) "Ojalá pudieras ser yo por unos momentos para que sintieras lo maravilloso que es tenerte dentro de mí".				
82	Y lo **dije** porque lo sentí así,				
83	pues esto y mil cosas más me hizo recordar esta historia.				
84	Un saludo y feliz y erótica noche, jeje.				

Book review "yes5_2"

#		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	Por fin un libro infantil pensado para niños (es decir personas pequeñas) y **no** para idiotas bajitos.				
2	Me explico,				
3	por alguna extraña razon se tiende a tratar a los niños **como si** fueran imbéciles				
4	y esto es, salvo honrosas excepciones extensible a la literatura infantil.				
5	Harry Potter y la piedra filosofal (y todos los demás de la serie) es un libro delicioso que resulta entretenido a cualquier edad.				

6	Consigue que un padre que se lo lea a sus hijos se lo pase tan bien como ellos				
7	(**incluso sin leerselo a nadie** a mi me encantó)				
8	**Por si** alguien **no ha oido** aun hablar de el (unas vacaciones en el polo norte **quizas**) Harry es un niño huerfano que vive con sus tíos, que **por supuesto** son odiosos hasta que un dia descubre que es un mago y que va a ir a estudiar a la mejor escuela de magia del mundo.				
9	contado de esta forma **tal vez no parezca** demasiado atractivo,				
10	un libro de aventurillas con malos malísimos y buenos buenísimos.				
11	**Sin embargo con todo eso** los personajes son creibles				
12	y el argumento está muy bien estructurado.				
13	Además partiendo de una muy bien documentada base mitológica y de la sabiduría popular sobre brujas, la escritora describe el colegio de magia y la vida de los propios magos de forma que resulta fácil imaginar que en realidad existen.				
14	**Incluso** inventa un deporte, el quidich, que es algo así como el futbol para nosotros, al que **si no fuera** porque **aún no he encontrado** escobas voladoras jugaría encantada.				

Book review "no7"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	I bought and read "A Painted House" as soon as it came out as I do with all John Grisham books.				

2	I was interested to read how he treated a story that **did not revolve** around a legal case.				
3	He **did not treat** it well,				
4	and I **will not buy** another one of his books **unless** it is a legal thriller like his others.				
5	The setting of the story **seemed** interesting enough				
6	--a rural Arkansas farming town during the Korean War.				
7	The story takes place over the length of a cotton-picking season.				
8	It is an observation of the picking season as told by a seven year old son of a sharecropper, Luke.				
9	Luke's character is very well developed,				
10	and his is the **only** character that is.				
11	There are many other characters in the story,				
12	**but** the reader is **never** given enough information to really care about them.				
13	The weak character definition made for a long and boring read, as I **really didn't care** what became of most of the characters in the story.				
14	The book was a very long read.				
15	**The one** redeeming quality was that it was very descriptive.				
16	I **had clear visuals** of what the characters and the town looked like.				
17	The descriptive nature of the story, **however, was not** enough to make this an interesting story.				
18	I **am not even going to talk** about the plot because I **really don't see** a point.				
19	To paraphrase Nike, "**Just** Skip It!"				

Book review "no21"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	I enjoyed reading Dean Koontz's book "Tick Tock" very much.				
2	What I **didn't enjoy** doing, **however,** was listening to this book on audiocassette.				
3	I had **never** before listened to a book on tape, so this was a first.				
4	**Just** because I **give my opinion** of this book on audiocassette a bad review, it **does not mean** I'm bashing audiocassette books in general.				
5	**I'm sure** this book on tape **has got to be** one of the worse.				
6	I listened to the taped version a week after finishing the book, so the book was **still** fresh in my mind.				
7	The person who read the book complete stripped the personality out of the characters by his interpretations of the voices.				
8	As I read the book, I **got the impression** of the main female character to be smart, clever, fun and sporadic.				
9	The reader, **however,** made the female's voice whiny and **obviously** fake.				
10	His reading of the book put it into a whole new perspective,				
11	and **not** a likeable one either.				
12	He **instead** made the whole book sound ridiculous.				
13	I am a very big fan of Dean Koontz books, therefore I would hate for someone who has **never** read or listened to a Dean Koontz book before to pick "Tick Tock" on audiocassette as their first choice.				
14	**Should** that happen**, there is a very likely chance that** the person **would never again try** any piece written by Dean Koontz.				

15	I suggest that **if** you plan on listening to Dean Koontz's "Tick Tock" on audiocassette that you read the book first, for it is much different **even though** it is exactly the same.				
16	PS--Please keep in mind that my ratings **are not** for the book itself,				
17	**but** for the book on audiocassette				

Book review "no25"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	A teacher of mine recommended this book to me, **saying that** all her friends loved it and **couldn't put** it down, so when it came time to do a critique in her class, I chose this book.				
2	Unfortunately, Mr. Pears book **was not** all **that it was cracked up to be**.				
3	The are too many problems that drag this book down, **preventing** it from being the masterpiece that **many are claiming** it to be.				
4	First, there is the narration style:				
5	4 different narrators telling the events the way they saw them, each fingering a different suspect.				
6	**If** the plot had been more gripping, more intense, this **would have worked** perfectly.				
7	Unfortunately, it **wasn't**,				
8	and by the beginning of the third narrator, I **did not feel** like reader the remaining two parts.				
9	Then there is the narrators themselves.				
10	Again, this **could have been** excellent, **save the fact that** 2 of the narrartors are so arrogant and full of themselves that it made it hard for me to want to keep reading it.				
11	Finally there was all the historical detail.				

12	Mr. Pears **obviously** spent a lot of time researching to make his book as accurate as possible, including his writing style.				
13	I commend him for that, **however** the details he gives are so numerous that they begin to over shadow the plot.				
14	**If** you like the history behind what each narrator does (i.e. the development of science for Da Cola's tale) then you will enjoy it.				
15	**If** you find any one of the subjects offered (say cryptography) then it will hinder your enjoyment immensly.				
16	I **did enjoy** reading some of this book, namely Da Cola's tale, and the final tale,				
17	**but** the two tales in the middle dragged too much for me to be able to **really** recommend this book.				
18	It is another case of hype building up to **nothing**				

Book review "yes9"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	Once upon a time, Stephen King wrote a fairy tale.				
2	The eyes of the Dragon is a truly mesmorizing and fascinating story.				
3	A fairy tale for adults,				
4	**but** **I think** this could be for older children also.				
5	It **doesn't have** the customary violence and language that most Stephen King novels have,				
6	and it **isn't** so intense that it will frighten you.				
7	"**Why** then would I want to read it?" You **may be asking** yourself.				
8	Well you want to read it because it is a great novel.				
9	The story is set in an ancient kingdom called Delain.				

10	The story centers around two young boys, princes, the son of the King.				
11	The two brothers are Peter and Thomas.				
12	Peter, the oldest, is in line to be the next king.				
13	Peter is the perfect child.				
14	Everyone loves Peter,				
15	he is handsome, intelligent, and caring.				
16	Thomas, on the other hand is none of these.				
17	He is a lonely little boy disliked and ignored by all, even his father.				
18	Thomas is ignored by all, that is, except for one person:				
19	the kings magician Flagg.				
20	Flagg has been in the kingdom for years; as long as anyone can remember.				
21	Before Peter can become King, his father is murdered and he is framed for it.				
22	Since Peter is royalty, he can't be killed, so he is sent up into a tower to live out the rest of his life.				
23	The only people he sees are his guards.				
24	The only thing Peter has to pass his time is his dollhouse that he played with when he was little,				
25	the only other convience he was allowed was a napkin to eat with,				
26	these he specially asked for.				
27	Thomas is now the King.				
28	However, Thomas is young and needs someone to help rule the kingdom.				
29	**Who** would that be?				
30	**Why** Flagg **of course**.				
31	Thomas knows a secret that could destroy him and is slowly driving him mad.				
32	This secret could not only destroy him,				
33	it would definitely destroy Flagg.				
34	**What** is this secret?				

35	Read the story:				
36	find out.				
37	**Will** Peter ever get out,				
38	**or will** he die in the tower?				
39	Once again read the story find out.				
40	I love Stephen King.				
41	The scarier the story the better.				
42	**However,** this is **just** about my favorite Stephen King story.				
43	It **isn't** scary,				
44	**but** it is enthralling.				
45	For the more avid King readers out there;				
46	check out the reference to this story in the Dark Tower series,				
47	Roland mentions some of the characters in this story that passed through his town when he was young.				
48	For anyone out there who **just** wants to read a good story, check out this book.				
49	It **isn't** violent				
50	(well there are a few parts,				
51	**but not** usual King violence).				
52	It **isn't** vulgar or sexual.				
53	It is **just** a great read.				
54	I recommend this to anyone and everyone.				

Book review "yes11"

		ENTERTAIN (EP, EV, PQ, O)	ATTRIBUTE (ACK, DIS)	DISCLAIM (DE, COUN)	PROCLAIM (CON, PR, END)
1	Murder, Mutants, Abuse, Torture, Doctors, Scientists, Secrets, Gun				

	Shots, Betrayal, Sex, Love....				
2	these are **just** some of what you'll find in James Patterson's book, When the Wind Blows.				
3	General Information				
4	When the Wind Blows is written by author James Patterson.				
5	It was published in 1998 by Little, Brown And Company.				
6	The hard cover version that I read had a total of 416 pages				
7	and unlike the other two books I've read by Patterson, I could **actually** find places where I could force myself to stop reading for the night.				
8	I've **only** recently started reading books he's written				
9	and **confess** that I'm already a huge fan.				
10	ISBN: 0-316-69332-4				
11	The Story				
12	Dr. Frances O'Neill, a well loved vet in Colorado, was working as usual in her place of business - the Inn Patient.				
13	Frannie as she's called by her friends, has had her share of heart break and loss.				
14	Her husband David, a doctor at the local hospital was murdered less than 2 years earlier.				
15	Frannie is a real softy which is what many of her clients and their owners loved about her.				
16	So there she was, minding her own business when one day this gorgeous hunk of man walks in and rents her cabin [which is being the Inn Patient].				
17	His name is Kit Harrison				
18	and while she can't stand him and at the same time she can't help but like him - just a little.				
19	It takes quite awhile for Frannie to find out but Kit is not who she thinks he is.				
20	Then, Frannie's good friends husband is killed.				
21	They say it was an accident, that he hit his head in the pool;				

22	but it's very hard for those who knew him to believe.				
23	After much crying and several drinks, Frannie drives back home to the Inn Patient where she not only works, but has lived ever since David's death.				
24	On her drive home she thinks she sees something.				
25	She stops the car.				
26	Wait.				
27	What is that?				
28	Oh my God!				
29	It looks like a human girl.				
30	A beautiful little human girl - no more than 12 years old.				
31	But she has wings...				
32	and oh my God!				
33	She can fly!				
34	Frannie tries to gain her composure				
35	and then resumes the drive home.				
36	By the time she reaches the Inn Patient she's convinced herself that she couldn't have really seen a girl with wings.				
37	It just wasn't possible.				
38	It must have been the alcohol				
39	...... right?				
40	That's all I'll give you for the story.				
41	Sorry, **but if** I give you more than that I **just may ruin** the book for you.				
42	That's how Patterson writes.				
43	And **in case you're wondering** all of this happens in the first few chapters of the book!				
44	About the writing				
45	Like every other book I've read by Patterson, When the Wind Blows				

	was thrilling and very hard to put down - making it very easy to read.				
46	Patterson captivates the reader from page one				
47	and **if you think** you'll find a stopping place at the end of a chapter - think again!				
48	Unlike the last book I read by Patterson Cradle and All, I was **actually** able to put down When the Wind Blows so I could get some sleep.				
49	It **wasn't** easy				
50	**but** I **did manage**.				
51	In the beginning of the book is an Author's Note.				
52	In this note, **Patterson states** that what he writes in When the Wind Blows is much closer to the truth than **even he thought** it would be.				
53	While in part, I have a hard time **believing** that - there's a part of me that knows there **must be** some truth to the matter.				
54	Overall				
55	**If** you're looking for an exciting book of fiction and science and doctors interest you, I recommend you give When the Wind Blows a try.				
56	There are some graphic parts which made me mad to read,				
57	**but I don't believe** it **would be** too much for anyone to handle.				
58	Thanks for reading ~				
59	Melissa				