Table of Contents

Introduction

A corpus-based approach

One of the major issues in science is that of evidence and its nature. While empiricists believe that knowledge can be attained through sense experience, rationalists believe that the source of our concepts and knowledge is human intellect. This debate has also characterised the field of linguistics. Empirical approaches to the study of language are based on the observation of natural language use, i.e. the language that speakers and writers actually employ in a natural context. On the other hand, rationalism in language studies is characterised by the formulation of theoretical statements about language on the basis of introspective judgements that native speakers can make about their mother tongue.

Between the 1950s and the 1970s, rationalism dominated the field of language studies under the influence of Noam Chomsky (see Chapter 1, section 1.1.1); however, the increasing availability of personal computers, which enable linguists to collect and store numerous instances of naturally occurring language in the form of a "corpus", was one major source which contributed to the gradual resurgence of interest in empirical data. Nowadays, "corpus linguistics" is an established approach to the study of language (see Chapter 1, section 1.2.2 for a discussion of whether corpus linguistics is a methodology or a discipline) and the importance of empirical data in language description is no longer questioned.

The focus of analysis of “corpus-based” language studies is actual patterns of language use, that is, «the systematic ways in which linguistic features are used in association with other linguistic and non-linguistic features» (Biber et al., 1998: 5). The investigation of patterns of association not only allows researchers to verify their intuitions about language, but also to investigate important phenomena which normally escape human introspection such as the frequency of distribution of those patterns.

In the present work, I adopt a corpus-based approach to investigate recurrent patterns in the expression of opinion and point of view. I focus on the use of adverbials marking “epistemic stance” (e.g. “obviously”, “possibly”, “actually”), focussing in particular on items which convey the writer’s comments on the degree of certainty (or doubt) of a proposition and its status as a real-life fact (Biber et al., 1999) (see Chapter 3 for a theoretical treatment of stance adverbials). I concentrate on these devices because they are one of the primary lexical markers of stance in English (Biber and Finegan, 1988: 1).

The group of speakers which I focus upon is composed of Italian university students who are learning English as a foreign language. I try to describe how stance adverbials are employed in terms of their rate of incidence and the contexts in which they appear. The aim is to verify whether Italian learners employ stance adverbials in the same way as native speakers.

In the following sections, I briefly introduce the phenomenon of language expressing opinion and point of view and relate it to argumentation and learner writing. I subsequently illustrate the corpora which I use in the present study and finally, I provide an outline of the book.

The expression of opinion and point of view

When we write or speak we do not communicate factual information only; on the contrary, along with facts, we concurrently provide information concerning our opinions, feelings and assessments of those facts. Example 1 illustrates the pervasiveness of this phenomenon.

> (1) *How big is the world economy?* That *sounds like* a *straightforward* question. *Simply* to add up the size of all the world's national economies *would seem* to be the *obvious* way to answer it. *But* how that is done yields *radically* different results, and *therein lies a tale*. [EMA][1]

Example 1 is about the size of the world economy and the writer aims to show that determining it is not as easy as it may appear. The example opens with a direct question, which is a way to engage readers, but also to cast doubt on the very object of the question, that is, the size of the world economy. The two following statements offer a particular view on this issue, which is being represented by the writer as a sharable, common sense opinion. This can be noted in the use of lexical items such as "straightforward", "simply" and "obvious", and the choice of grammatical constructions such as the simple present tense for the verb "to sound", which conveys the impression of factuality. However, these two sentences contain other lexical features which pre-empt the evaluation of this apparently sharable point of view as problematic. These features are the expression "sounds like" and the verb "seem", which are characterised by a lower degree of certainty than the verb "to be" (e.g. "This *is* a straightforward question"; "Simply to add up the size of all the world national economies *is* the obvious way to answer it"). This impression is strengthened by the use of the modal auxiliary "would" which frames the sentence as hypothetical and debatable.

The third statement starts with the conjunction "but" which indicates contrast with respect to the previous statement, a contrast which was announced by the expressions of uncertainty noted above. "But" introduces the point that the writer wishes to

---

[1] This example is taken from EMA ("editorial material"), one of the corpora which I use for analysis. It collects columns and editorials. See Chapter 3 for more detailed information on the corpora used in this study.

make: the way national economies are added up can make a difference in terms of the overall size of the world economy. Unlike the previous sentences, which were characterised by a degree of uncertainty, the third statement is characterised by an assertive tone indicating that the writer is committed to this point of view. The use of the adverb “radically” not only allows the writer to convey his/her opinion on the results (i.e. they are not simply different, but extremely different); “radically” also serves to reinforce the assertiveness of the statement. In the light of the opinion introduced by “but”, words such as “simply” and “obvious” acquire slightly negative overtones as they reflect a rather naïve or even superficial attitude on the part of anybody thinking that to determine the size of the world economy it is sufficient to simply calculate the sum of the different national economies. Example 1, therefore, shows that the expression of opinion and point of view is highly context-dependent. Although the words “simply” and “obvious” always encode an opinion, they acquire positive or negative connotations depending on the context in which they occur.

Finally, the third statement ends with “and therein lies a tale” which is a humorous expression normally used to say that something needs a long explanation. The choice of this expression indicates the playful attitude of the writer towards the explanation which follows and towards the readers who not only are engaged through a humorous remark, but they are also asked to access their background knowledge in order to recognise the Shakespearean reference[2]. Familiarity with this expression helps establish common ground between the writer and the readers thus encouraging solidarity towards the writer's opinion. This expression also works as a metadiscoursal comment (“therein”) that serves to pull the reader along with the argument.

[2] The expression used by Shakespeare is “thereby hangs a tale” which means “there's a story about that” (Honigmann, 1997: 205). This expression appears in *As you like it* (Act 2, Scene 7, line 922) (text downloaded at http://www.opensourceshakespeare.org/). It also appears in *Othello* (Act 3, Scene 1, lines 7-8).

The analysis of example 1 shows that the expression of opinion and point of view is a very complex phenomenon, which plays a role on different language levels, from lexis to grammar, from single sentences to the whole discourse. In addition, it is not confined to a specific part of the text, but it is a pervasive type of information.

The complexity of the phenomenon of language expressing opinion and point of view is reflected in the heterogeneity of approaches that have been employed to investigate it. Consequently, a large number of terms have been used: “affect” (Besnier, 1993; Leech, 1974), “appraisal” (Martin, 2000; White, 2003), “attitude” (Halliday, 1994), “evaluation” (Hunston and Thompson, 2000), “evidentiality” (Chafe, 1986), “metadiscourse” (Crismore, 1989; Hyland, 2005a), “modality” (Palmer, 1979; Halliday, 1994; Stubbs, 1996) and “stance” (Biber and Finegan, 1988; 1989).

Most of these terms cannot be considered synonyms because they do not share the same perspective and/or because they describe different levels of language. For instance, the marking of the speakers’ judgements of desirability has been approached either by focusing on the language user (it is people who assess and judge the world through the means of language) or by concentrating on the language item (it is words which convey “connotative meanings”) (Thompson and Hunston, 2000: 2). Accordingly, this function has been defined as “attitude” or “affect” in the former case, and “connotation” in the latter (*ibid.*: 2). Leech’s distinction between “connotative meaning”, i.e. «the communicative value an expression has by virtue of what it refers to, over and above its purely conceptual content» (1974: 14) and “affective meaning”, i.e. language expressing «the personal feelings of the speaker» (*ibid.*: 18) can be seen as reflecting in part this difference in perspective.

Chafe (1986) employs the label “evidentiality” to refer to the speaker’s attitudes towards knowledge and to the source of information. Stubbs (1996), instead, refers to the speaker’s degrees of commitment to and detachment from propositional information using the term “modality”. “Modality” is also the label through which Halliday studies the ways in which «the intermediate degrees between positive and negative poles» (1994: 88) can be expressed. Stubbs and Halliday, however, use the term "modality" to cover different areas of language. Halliday concentrates on modal auxiliary verbs (e.g. "they *must* have known"), modal adjuncts (e.g. "they *certainly* knew"), passive verb predicators (e.g. "You *are required* to be patient") and adjective predicators (e.g. "I am *determined* to win") (*ibid*.: 91). Stubbs uses the term modality to cover a wider area which also includes expressions of lexical commitment (e.g. "so to speak", "what is often called", "so-called"), vague lexis (e.g. "kind of", "sort of", "and things like that") and syntactic phenomena such as "simple forms" and "ing- forms" of verbs, and the use of connectors[3].

Modality has traditionally been studied as a separate phenomenon from the expression of attitudinal (Halliday, 1994) or affective (Leech, 1974) meanings. However, recent approaches to the study of opinion and point of view have taken a "unifying perspective" (Thompson and Hunston, 2000: 4). According to Hunston (1994), modality is only one of the different types of assessment that speakers can make. In addition to conveying their degree of doubt and certainty about the propositions they utter, speakers also comment on whether something is good or bad and whether it is important or unimportant. Moreover, the expression of modality may

---

[3] According to Stubbs (1996: 216), the choice between "simple form" and "ing-form" of verbs reveals the degree of truth and certainty attached to a proposition. "Simple forms", such as "I promise", express certainty and permanence, whereas "ing-forms", such as "I'm suspecting that he might not win", express uncertainty and change. Stubbs also points out that logical and pragmatic connectors ("because", "but", "if") are used not only to perform logical functions, but also to justify the confidence in the truth of what is being said (ibid.: 224). He provides the following examples to illustrate this interaction of syntactic and pragmatic functions:

> He was drowned because he fell off the pier
> He was drowned, because he fell off the pier

The first sentence relies on the structure "effect + cause", while the second has the structure "assertion + justification" (ibid.: 224).

anticipate other types of assessment such as that of desirability (*ibid*.: 197). This unifying perspective is referred to by Thompson and Hunston as “evaluation”, i.e. «the expression of the speaker or writer’s attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about» (2000: 5).

According to Thompson and Hunston (2000: 6-13), evaluation performs three main functions: to express opinions, to build up and maintain relations and to organise the discourse. This definition overlaps significantly with that of “metadiscourse” which is the way writers «express a viewpoint and engage with readers as members of a particular community» (Hyland, 2005a: 37). Metadiscourse also accounts for the writers’ awareness of the audience’s need for guidance through the text. Therefore, metadiscourse includes language resources which are used to overtly organise discourse in recognisable sequences of meaning (*ibid*.: 49)[4].

A further unifying perspective that overlaps considerably with the frameworks of evaluation and metadiscourse is that of “appraisal” (Martin, 2000; White, 2003; Martin and White, 2005). The focus of this model is on 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). The study of appraisal, however, unlike that of evaluation or metadiscourse, is rooted in the theoretical tradition of Systemic Functional Linguistics whose major exponent is M.A.K. Halliday (1994). The appraisal framework takes into account both explicit and implicit evaluation and is not confined to the study of specific grammatical constructions.

In this work, I take the unifying approach and investigate the expression of opinion and point of view using Hunston and

[4] The overt or explicit marking of opinion and point of view, which is a prerequisite for inclusion of evaluative devices in frameworks such as metadiscourse, raises questions as to how we can best account for implicit evaluation which is as important and pervasive as explicit evaluation. This issue has been discussed in Douthwaite (2007).

Thompson's (2000) evaluation framework. One major reason behind the adoption of this approach is that my analysis includes a study of epistemic stance adverbials in terms of the three functions which Hunston and Thompson identified as characterising evaluation, i.e. to give opinion, to build and maintain relationships, and to organise discourse. However, I also employ the terminology used by Biber and Finegan (1988; 1989) who have studied stance adverbials extensively thus providing functional categories and metalanguage to classify and describe them.

Opinion and point of view in argumentative writing

The study of opinion and point of view is particularly significant when dealing with argumentative texts. One of the main concerns of anyone who is producing a piece of argumentative writing is to persuade readers. Readers will be persuaded if the logic of the argument appears convincing to them. One way of achieving this result is to strategically weave writer visibility and detachment by overtly taking a position (e.g. “I believe that”) or by disguising personal involvement (e.g. “it may argued that”) (Hyland, 2005b; Molino, 2010). When writers take an overt position, they emphasise their contribution to the debate. In academic discourse, for instance, self-mention helps «construct an intelligent, credible and engaging colleague, by […] reflecting an appropriate degree of confidence and authority» (Hyland, 2001: 216). On the other hand, by choosing to background their role in research and adopt an impersonal stance, writers construct for themselves a modest authorial self, which in some disciplinary fields (e.g. geology) is considered as more appropriate and convincing (Dressen, 2003: 278).

In addition, writers also need to establish the most appropriate relationship with readers, for instance, by choosing the right words and structures not to offend them in case the argument diverges from shared assumptions.

Epistemic stance adverbials contribute to achieving the aim of persuading readers as they may be employed by writers to emphasise their individual positions (e.g. "really"); to distance themselves from the truth-value of a proposition (e.g. "probably"); to assume the existence of common ground with the audience (e.g. "clearly"); and to concede points to readers anticipating their possible objections (e.g. "certainly") (Biber and Finegan, 1988; 1989) (see Chapter 3 for a discussion of the use of stance adverbials).

Text corpora

The data for the present investigation are taken from three different sources: the "International Corpus of Learner English" (ICLE), a corpus of persuasive opinion articles from newspapers and magazines, which is given the notation EMA ("editorial material"), and the "Louvain Corpus of Native English Essays" (LOCNESS).

ICLE is a 3.7-million-word learner corpus (Granger et al., 2009). It mainly collects argumentative essays written by university students from different mother tongue backgrounds. In this work, I am only concerned with Italian students using English as a foreign language. Therefore, I only analyse the Italian subcorpus of ICLE. The data provided by this subcorpus is compared to data taken from EMA, a corpus which I personally collected. EMA consists of editorials and columns in English taken from authoritative newspapers and magazines. The aim of this comparison is to identify areas in which non-native writing diverges from native writing standards in the marking of epistemic stance through the use of adverbials. However, as a number of scholars has pointed out (Flowerdew, 1998; Granger et al., 2002b; Neff et al., 2003), the issue of "nativeness" may interact with that of "expertise". In other words, the differences noted in the use of stance adverbials between native writers and learner writers may also be due to the lack of experience in argumentative writing on the part of young learners. For this reason, I also investigate the use of epistemic stance adverbials in LOCNESS, a comparable corpus of argumentative essays by British and American students. The combination of the results obtained from the comparison between the two native corpora on the one hand, and the non-native corpus on the other, may help better understand the complex interplay of nativeness and expertise in the use epistemic stance adverbials in argumentative writing by Italian learners of English (for a detailed description of the three corpora and a discussion of methodological issues see Chapter 3, section 3.3).

Outline of the book

The first chapter illustrates the theoretical and methodological context of the present work, focusing on corpus linguistics and computer learner corpora. After an account of the birth of corpus linguistics, I discuss its status in modern linguistics. Some key-issues related to the notion of "corpus" are dealt with. Subsequently, I describe the role which learner corpus data play in Foreign Language Pedagogy and Second Language Acquisition research. Criticisms levelled at learner corpus research is also discussed. There follows a general description of computer learner corpora according to their design criteria. The methodological approaches to the study of learner language are also illustrated and a brief overview of the software programmes most commonly used for learner corpus analysis is furnished.

In the second chapter, I present the phenomenon of language expressing opinion and point of view according to the framework of evaluation. The functions of evaluation (i.e. expressing opinion, maintaining relations and organising discourse) are discussed and illustrated. I subsequently describe the most common parameters that speakers/writers adopt for their judgements, i.e. desirability, certainty, expectedness and importance. A discussion of the linguistic identification of the phenomenon is provided.

The third chapter is concerned with the actual linguistic analysis. The chapter opens with an illustration of the structure and functions of epistemic stance adverbials. I subsequently give an account of the corpora involved in the research (ICLE, LOCNESS, EMA), I expound the methodology adopted and I illustrate and discuss the findings.

The conclusion provides the answer to the research questions and offers suggestions for pedagogical applications.

Chapter I

Corpus Linguistics and Computer Learner Corpora

1.1. From early corpus-based approaches to modern corpus linguistics

In a broad sense, a corpus can be defined as «any collection of more than one text» (McEnery and Wilson, 1996: 21), either of written or of transcribed spoken language, which is meant to provide empirical data for linguistic analyses (see 1.2.1 for a more detailed definition of corpus according to its constitutive features). Linguists have been using empirical data since as early as the end of the nineteenth century. For example, longitudinal studies concerned with language acquisition have always relied upon collections of utterances representing children's language as sources of data (Brown, 1973). Historical linguists have also needed empirical evidence as the basis for their statements on a particular historical period or a "dead" language. Examples of historical corpora include the "Helsinki Corpus" (Kytö, 1996) containing data from the earliest Old English period to the end of the Early Modern English period, and the "Corpus Taurinense" (Barbera, 2009) collecting samples of XIII century Florentine.

During the 1950s the use of empirical data for linguistic analysis was gradually abandoned as mainstream linguistics relied almost exclusively on intuition for its theoretical statements. This shift away from empiricism towards rationalism took place under the influence of Noam Chomsky (1957; 1965).

1.1.1. Chomsky's critique of corpus-based research

Chomsky elaborated a model known as transformational-generative grammar. This model relied on the idea that only native speakers' intuitions can account for grammatical phenomena which do not appear on the "surface" of language. Chomsky and the generativists were convinced that human beings possess an innate set of linguistic concepts which they can access through introspection and which empirical evidence is not able to account for. In addition, they believed that the sentences of a language are potentially infinite, even though they combine according to a limited number of rules. The aim of Chomsky and the generativists was to define the core of language knowledge and to state abstract principles which could be applied to a "universal grammar". They employed this label to refer to a grammar which is common to any language, and which can explain all languages using the same basic parameters. In order to achieve that aim, they concentrated on language competence as revealed by the judgements that native speakers of a language can make about their own language, rather than on language performance, i.e. what native speakers actually say in their language.

Chomsky levelled three main criticisms at early corpus-based research (McEnery and Wilson, 1996: 4-10). The first one had to do with the fact that his theory of language was based on the dualism of performance and competence. Because corpora are by their very nature collections of natural language use, that is, of linguistic performance, a rationalist theory of language cannot rely on the information provided by corpora. According to Chomsky, performance is determined by factors which do not relate directly to competence, such as the physical and mental situation of language production. As a consequence, performance does not mirror competence and is consequently an unreliable source of evidence.

The second objection to corpus-based research stems from Chomsky's idea that the sentences of a natural language cannot be gathered in a corpus because they are not finite. Consequently, a corpus, however large, will always be incomplete thus providing only partial information about a natural language.

The third criticism concerned statistical evidence. Chomsky argued that since frequency is fundamental to the inclusion of an utterance in a corpus, it is almost impossible to distinguish between ungrammatical features (e.g. *He passed me the salt* vs. **He broke me the salt*) and features which do not occur in the corpus because they are rare. This led him to the conclusion that only introspection could account for sentences which were grammatically incorrect.

The effects of such criticism were extremely damaging for the corpus-based approach, to the extent that it is commonly believed that corpus-based studies were suspended during the 1960s and the 1970s, and that they resurged in the 1980s (McEnery and Wilson, 1996: 17) (see section 1.1.2). The reason why corpus linguistics did not collapse completely was that the drawbacks of rationalism had become evident. Researchers working on computer corpora could counter-argue Chomsky's position relying on a set of ideas which had been circulating since the 1930s, and which, in turn, had gained strength thanks to corpus work since the 1960s. In particular, the advantages of naturally-occurring data over introspection became increasingly evident.

As has been argued by a number of corpus linguists (McEnery and Wilson, 1996; Stubbs, 1996; Pullum and Scholz, 2002; Sampson and McCarthy, 2004), an introspective judgement on a thought process is not verifiable by other analysts as it is a personal, private viewpoint. On the contrary, when sentences are actually uttered, it is possible to record them and they become observable by any linguist willing to examine them. Corpora, therefore, have the advantage of rendering public the evidence which supports any theoretical statement about language.

Another advantage of corpus data is that they are authentic whereas introspective data are artificial and decontextualised. When a linguist asks the informant whether a construct is possible in his/her mother tongue, the risk is that of manipulating the answer because as McEnery and Gabrielatos observe, «what might sound awkward and ungrammatical out of context can become quite grammatical in context» (2006: 98).

The partiality of the data contained in a corpus, which Chomsky considered a major problem, has been seen by corpus linguists as an aspect which is interesting in itself because the absence of a particular language feature or the recurrent occurrence of another feature offers linguists evidence of the frequency of a construct or a word. Frequency is a crucial datum of which native speakers of a language are only vaguely aware, but which is of paramount importance when accounting for how naturalness is achieved (Hoey, 2005). Finally, corpus-based analyses allow linguists to carry out systematic investigations which are not possible using the process of introspection (Leech, 1992).

1.1.2. Three stages of penetration of the computer corpus in linguistics

Tognini-Bonelli identifies «three stages of penetration of the computer corpus in linguistics» (2001: 45-46). The 1960s represent the starting point of corpus linguistics. It was in this decade that important pioneering enterprises were planned. In 1961, at University College London, Randolph Quirk (1960) began to compile the "Survey of English Usage" (SEU), a collection of written and spoken British English. In the same year, at Brown University, Providence, Rhode Island, W. Nelson Francis and Henry Kucera started working on the "Brown Corpus" collecting American English prose (Francis and Kucera, 1979).

In the early 1980s, corpus linguistics became increasingly influential. In the main, this was due to the availability of both institutional and private computing facilities which contributed to a second stage of penetration of the computer corpus. The computer corpora of the 1980s were bigger than those of the previous generation, and software improvements allowed researchers to carry out quicker and more accurate analyses. This had the effect of refining and systematising the methodology and also of making linguists aware of the value of sophisticated software programmes available to process the data.

The 1990s are seen as a third phase. The amount of data available was so large that it enabled researchers to obtain new insights into language use. For instance, they were able to provide evidence for phenomena such as the "collocation" (Firth, 1957) of words (see section 1.2.2.2). Some linguists started to look at the computer corpus in a different way. In their view, the computer corpus could no longer be considered as a simple tool, but as a resource capable of changing existing theoretical assumptions. This position is referred to as the "corpus-driven approach" and it is often contrasted with the "corpus-based approach" (see section 1.2.2). In the following section, I provide a definition of corpus according to current views of this resource and discuss the differences between the corpus-based and the corpus-driven approach.

1.2. Modern corpus linguistics

1.2.1. Definition of "corpus"

It is common practice among scholars who provide book-length accounts of corpus linguistics to define the concept of "corpus" (Sinclair, 1991; 1996; Atkins et al., 1992; McEnery and Wilson, 1996; Stubbs, 1996; Biber et al., 1998; Tognini-Bonelli, 2001; Granger et al., 2002b; Hunston, 2002; Meyer, 2002; Sampson and McCarthy, 2004; Sinclair, 2004). A corpus is considered a finite size collection of authentic language utterances which is representative of a particular language variety. A corpus has to be compiled according to specific design criteria and it functions as a standard reference for research.

Below, making reference to the existing literature on the topic, I focus on the major issues connected with the modern conception of corpus and how to build one.

1.2.1.1. Authenticity

The issue of authenticity has been under discussion since the beginning of corpus linguistics, notably as a key principle against rationalism (Sinclair, 1991; 1996; Stubbs, 1996; Biber et al., 1998; Tognini-Bonelli, 2001; Granger et al., 2002b; Hunston, 2002; Sampson and McCarthy, 2004). Corpus linguistics is the study of natural language use. The label "natural language" is adopted by scholars to mean utterances actually produced by speakers or writers in communication, in contrast to synthetic sentences derived for introspection and teaching. In order to study a natural language, it is necessary to make use of entire texts rather than isolated invented sentences, as was typical among rationalists. If the unit of study is complete texts, the examples provided by the corpus can be considered authentic in the name of an existing context and by dint of constituting a *bona fide* speech event. For this reason, the data provided by corpus analysis can be seen as a valid kind of evidence for linguistic statements (Sinclair, 1991: 5; Stubbs, 1996: 32). The term "use" added to "natural language" reinforces this stance, because it means that the object of study is authentic utterances rather than what is theoretically possible in a language (Biber et al., 1998: 1).

The concept of authenticity may be debatable when applied to texts which are not produced spontaneously. This is the case of "learner corpora", the type of corpus which I employ in the present study. Texts produced by language learners are often elicited by the teacher or the researcher in order to obtain and analyse particular lexical items or syntactic constructs. A restrictive interpretation of the concept of authenticity would exclude this type of corpus form the typology of corpora which are suitable for providing evidence for theoretical statements (see Oostdijk, 1991). With regard to this issue, however, Sinclair (1996) suggests making the criteria of corpus design explicit and using the label of "special corpora" for corpora containing pieces of language which have arisen under experimental conditions. This precaution serves to prevent researchers from making statements about ordinary language by using corpora not designed to record genuine linguistic behaviour.

1.2.1.2. Representativeness

Representativeness is a central feature if the corpus is considered a collection of evidence from which generalisations about language may be made (Sinclair, 1991; Leech, 1992; Biber et al., 1998; Tognini-Bonelli, 2001; Meyer, 2002). Linguists should sample texts in such a way as to render the corpus as representative as possible of the language variety under analysis (see section 1.2.1.3 for a discussion of corpus size). In particular, if a corpus is used to provide theoretical statements about a language as a whole, linguists should pay considerable attention to the issue of diversity, i.e. the variety of text types included in the corpus. According to Biber et al. (1998: 249) a large corpus is well-designed if it contains as many registers as possible of a given language.

1.2.1.3. Finite size

The condition of finite size is closely interconnected to the issue of representativeness and the use of statistical methods to analyse the corpus. Finite size is linked to representativeness because this latter concept implies selection, which in turn implies choosing a finite number of elements. Finite size is also linked to statistical analysis which can be carried out only if the elements on which statistics operates are finite (Barbera et al., 2007: 51).

A notable exception to the issue of finite size is the monitor corpus in which texts are constantly added so as to provide up-to-date information and, consequently, a historical perspective on language use (Sinclair, 1991: 25). An example is the "Bank of English" (currently standing at 450 million of words), an international project sponsored by HarperCollins Publishers and conducted by the COBUILD team which was originally lead by

Prof. John Sinclair at the University of Birmingham, UK[1].

The issue of size rises questions as to whether the corpus should be large or small. Sinclair argues that «there's no maximum to size» (2004: 9); rather we should talk of the minimum size which is determined according to «1) the kind of query that is anticipated from users, 2. the methodology they use to study data» (*ibid*.: 9). Corpora, such as the "Bank of English" are suitable for studying language as a whole and for providing data for the most varying linguistic constructions, such as conditional clauses (Meyer, 2002: 33). Small corpora (e.g. the Italian subcorpus of ICLE used in this study which totals approximately 200,000 words) are preferable for the study of restricted varieties (Bondi et al., 2004).

1.2.1.4. Principled choice

The concept of principled choice (Leech, 1992; Biber et al, 1998; Tognini-Bonelli, 2001) helps distinguish between a corpus proper and a simple text archive or database in which texts are gathered which do not necessarily have a connection with each other (Kennedy, 1998: 4). Corpora have to be compiled according to specific design criteria. Design criteria are established on the basis of the type of linguistic inquiry.

1.2.1.5. Standard reference

The concept of standard is fundamental to an approach to the study of language which aims to be systematic and scientific in its analyses. According to Stubbs «the most important implication(s) of corpus study [is] that the data of linguistics become publicly accessible» (1996: 234). Public accessibility implies that linguists working on the same area of study may use a shared database of natural language or that they may have access to information regarding the way the corpus used by a col-

---

[1] For more information on the "Bank of English", see the official website http://www.titania.bham.ac.uk/

league has been compiled. As a consequence, findings can be compared and in case of significant differences, those discrepancies can plausibly be ascribed to the use of a particular methodology rather than to differences in the data analysed (McEnery and Wilson, 1996: 24).

1.2.1.6. Machine readable format

The widespread availability of corpora, which is at the basis of the public accessibility of data, has become possible thanks to the use of machine-readable texts. Corpora in electronic format can be processed by sophisticated software programmes that enable researchers to carry out quick and precise analyses. In addition, they can be annotated (see section 1.2.2.1) by using, for instance, grammatical tags which classify the words in a corpus according to the part of speech to which they belong (e.g. the word "cat" would be tagged as a *noun*). Finally, machine-readable corpora make storage less difficult.

1.2.2. Corpus-based vs. corpus-driven approach

In recent years two theoretical positions concerning the way corpora can be used in research have emerged, the corpus-based and the corpus-driven approach. These can be distinguished in terms of the methodology adopted and the view of the type of insights that a corpus can provide. In this section, these two positions are briefly described and discussed.

1.2.2.1. The corpus-based approach

Linguists who adopt a corpus-based approach analyse corpora with pre-existing categories in mind, which they derive from established linguistic theories. The purpose of using corpora is to test and exemplify language descriptions that were formulated before the use of large corpora (Tognini-Bonelli, 2001: 65). Corpora are therefore employed to quantify language phenomena and to refine existing categories.

Corpus-based linguists often believe that a corpus of “raw” data should be enriched by adding annotation. Annotation, or tagging, is carried out by attaching a code (label or tag) to each single linguistic item in the corpus. Various types of annotation are available. Part-of-Speech (POS) Annotation (or Grammatical Tagging, or Morphosyntactic Annotation) is the most common mode for tagging. It makes use of labels indicating the part of speech to which a particular lexical unit belongs. For example, the item *house* may be tagged as singular common noun. Parsing or Syntactic annotation, instead, marks the syntactic relationship between the constituents of a sentence. Semantic annotation marks either the semantic relationships among items in a text or the semantic area to which one of the senses of a word belongs.

According to corpus-based linguists, corpus linguistics is seen «more [as] a way of doing linguistics […] than a separate paradigm within linguistics» (Meyer, 2002: xi). Evidence to support this opinion is provided by the fact that the majority of corpora now available have been compiled in such a way as to address the research needs of various areas of study. Applied linguistics (Hunston, 2002), discourse analysis (Baker, 2006), contrastive linguistics and translation studies (Granger et al., 2003), language education (Tan, 2002), cognitive linguistics (Gries and Stefanowitsch, 2006), stylistics (Semino and Short 2004) and lexical semantics (Stubbs, 2001) have all taken advantage of the corpus-based approach.

1.2.2.2. The corpus-driven approach

The method adopted by corpus-driven linguists is inductive. The corpus is investigated to observe recurrent linguistic behaviour without having already formed a hypothesis before approaching the data. Hypotheses are formulated after observation. The subsequent testing of these hypotheses leads to a generalisation about a given language phenomenon and finally to a theoretical statement (Tognini-Bonelli, 2001: 85).

According to Tognini-Bonelli, who offers one of the most extensive and detailed accounts of the two approaches, a crucial difference between the corpus-based and the corpus-driven approach is the fact that in the former, corpus data are used to make minor adjustments to existing linguistic models, while in the latter, data are used to challenge existing models if the data do not fit those descriptions.

The aspects which attract the attention of researchers in a corpus-driven study are: 1) the presence of recurrent patterns of language; 2) their frequency of distribution and 3) the absence of a pattern. The two most easily observable patterns are "collocation" and "colligation". Collocation is the tendency of specific lexical items to co-occur. Examples are the expressions "to *make* progress" (and not "to *do* progress") or "*heavy* smoker" (and not "*big* smoker"). Colligation is the grammatical pattern in which a given word or expression is likely to appear. For example, Hunston (2001: 27-28) shows that the expression "up to a point" tends to occur at the end of a sentence before the full stop, and functions as a contrasting response to an opinion which is only partially accepted (e.g. "And so it does, *up to a point*").

Corpus-driven linguists prefer to use "raw", non-annotated corpora because they believe that annotation may induce linguists to focus on language categories (e.g. the part of speech of lexical items) at the expenses of the distinctiveness of words and their contextual features (Tognini-Bonelli, 2001: 92-92) or to construct hypotheses prior to examining the raw data.

1.3. Learner corpus research

1.3.1. The background to learner corpus research

The great variety of corpora now available to researchers and the increasing concern of corpus compilers with the fact that corpora should be carefully built up have made it possible to answer very different research questions and to concentrate on both the language as a whole and restricted varieties of language[2]. The data of corpus-based analyses have thus been exploited in several areas of linguistics, ranging from lexicography to stylistics, from critical discourse analysis to sociolinguistics. However, corpus linguistics entered the educational field fairly late when compared to other areas (Leech, 1998: xvi-xvii).

According to Leech there are two reasons for this delay, one practical and the other theoretical. The practical reason is that researchers operating within the fields of Foreign Language Pedagogy and Second Language Acquisition (SLA) research could not always afford to buy computers or carry out expensive research projects. This problem was greatly reduced with the advent of personal computers in the mid 1980s. The theoretical reason is that over the past 20 years learners' productions have been disregarded in favour of introspective data. This was a consequence, on the one hand, of the importance attributed by the communicative approach to learning as a process and, on the other hand, of the tendency of SLA researchers to consider the task of learning a foreign/second language as the acquisition of competence (see section 1.3.2).

In the late 1980s learner corpora began to be collected, even

---

[2] The "British National Corpus" (BNC) (100 million word) is an example of a general purpose corpus collecting samples of written and spoken British English from a wide range of sources (see official website: http://www.natcorp.ox.ac.uk/). The "Bergen Corpus of London Teenage Language" (COLT), on the other hand, is an example of a corpus focusing on a particular language variety, namely the speech of teenagers from different boroughs of London (see official website: http://www.hd.uib.no/colt/)

though their theoretical and practical qualities were acknowledged only in the early 1990s (Granger, 1998: 4; Granger, 2002: 4). The field of research connected to learner corpora is known as "learner corpus research". Its aim is to provide answers to either research questions concerning restricted groups of learners or more general questions which may suggest how to learn a foreign/second language successfully (Leech, 1998: xv).

A data-oriented approach to learner language analysis is not new. In the 1960s and 1970s studies in Error Analysis (EA) focused on learners' performance, though exclusively in terms of errors. The purpose of EA was twofold: on the one hand, actual learner mistakes were analysed to provide feedback about teaching methods and materials; on the other hand, learner erroneous performance was studied to understand how languages are learned and produced (Ringbom, 1999: 492).

The procedure of describing and explaining learner language exclusively from the point of view of shortcomings was subsequently criticised (*ibid.*: 492). One obvious limitation was that the concentration on errors provided a very partial picture of learner language. However, what is of equal, if not greater interest is how learners produce correct forms.

A further risk was to underestimate the positive influence that the first language may have on learner output, a risk which is connected to that of ignoring the phenomenon of avoidance. Avoidance occurs when learners do not employ a structure which they find difficult. This phenomenon remained unaccounted for by EA as the focus was exclusively on what learners do (Ellis, 1994: 68). Finally, the adoption of the concept of "error", which may be convenient for analysing the language produced by beginners and intermediate students, cannot fully account for the language produced by advanced learners, as their performance is often characterised by a certain degree of non-nativeness and it may be difficult even for native speakers to identify what in a text is "erroneous" (Ringbom, 1999: 492). According to Leech (1998: xvii), the increasing awareness of these limitations contributed to reinforcing prejudices against data-oriented studies of learner language.

Learner corpus research, however, is different from EA. First of all, it makes use of powerful and sophisticated computer programmes that can investigate language quickly and systematically. Secondly, learners' linguistic behaviour can easily be studied not only in terms of errors ("misuse"), but also from the point of view of what learners do correctly. Moreover, contrary to EA, computer learner corpora allow researchers to study what learners avoid, for instance by comparing native and non-native speakers' productions. Finally, learner corpus research enables linguists to describe learner language according to the frequency with which certain features are used. In other words, a linguistic feature may be labelled as "overused" or "underused", depending on whether learners use it more or less frequently than native speakers.

1.3.2. Empirical data in Foreign Language Pedagogy and Second Language Acquisition (SLA) research

Foreign language pedagogy has to do with the task of teaching and learning a foreign language. With the term "foreign language" scholars mean that the process of acquisition takes place in an educational environment and that the target language does not play an institutional role in the country where it is learnt. Second Language Acquisition research, on the other hand, is concerned with the study of «any language other than the first language […] [which] plays an institutional and social role in the community» (Ellis, 1994: 11-12). A second language, in general, is not learnt in a classroom setting, but in a natural context without explicit instruction. However, the term "second language acquisition" is sometimes used as a general term that embraces both untutored and tutored acquisition (*ibid.*).

Although the two fields of research have developed independently, they share a number of similarities. The formulation of the notion of interlanguage, which was first coined by Selinker (1972), was fundamental to both lines of research. Interlanguage is the intermediate language system, which is different from both the mother tongue and the target language, which learners develop by degrees. According to Tarone, the concept of interlanguage «provided the initial spark which ignited a field of research on second language acquisition/learning» (1999: 512). In the case of foreign language pedagogy the notion of interlanguage has allowed teachers and researchers to approach learners' productions in terms of the whole language system which characterises their knowledge of the target language at a given point in time.

Foreign language pedagogy and SLA research have another point in common: they have traditionally disregarded empirical data (Granger, 2002: 6). In the field of foreign language pedagogy, despite the interest in learner variables such as motivation (Oxford, 1996), learner needs (Brown, 2009) and learning styles (Oxford and Ehrman, 1993), empirical data concerning learner language has long remained in a peripheral position (Mark, 1998). Recently, however, there have been many claims for the importance of empirical data, both native and learner data (Johns and King, 1991; Aston, 1995; Wichmann et al., 1997).

According to Murison-Bowie (1996) it is possible to distinguish "weak" and "strong" claims in favour of empirical data in foreign language pedagogy. Researchers supporting the "weak" claim view corpora as sources of authentic language data to be exploited to offer learners genuine descriptions of the target language. According to Biber et al. (1994: 174), computational methods should be used to expose learners to different varieties of language; in this way, they will become aware that different registers favour different linguistic features. According to Bernardini (2000: 134), the concept of idiomaticity seems to be one of the main concerns of the supporters of the "weak" claim. In order to achieve naturalness, learners are presented with culturally specific ways of saying thanks to the use of authentic material.

Native language data have already been exploited to improve dictionaries, grammars and textbooks aimed at students of English as a foreign language. Dictionaries, in particular, are the tools which have benefited most from more accurate descriptions of authentic native English[3]. In addition to native English data, learner corpus data, too, can be exploited in the areas of materials design, classroom methodology, curriculum design and language testing.

The supporters of the "strong" claim for the importance of empirical data in foreign language pedagogy take a "process-oriented perspective" (Bernardini, 2000: 132), that is, they consider corpora not only as sources of authentic language but also as tools which can help improve teaching and learning methods. Their aim is to stimulate learners to develop inductive abilities and to become increasingly autonomous in their learning process. One way could be to integrate «form-based activities into the communicative teaching framework» (Granger and Tribble, 1998: 199) and provide corrective feedback, i.e. explicit instruction in areas where learners are more prone to errors. Form-focussed instruction and corrective feedback can be put into effect through a concordancer, i.e. «a collection of the occurrences of a word-form, each in its own textual environment» (Sinclair, 1991: 32). This teaching practice, which is referred to as "data-driven learning" (DDL), is intended to make learners aware of patterns of use and to teach them how to form generalisations. In this way the student is seen as a researcher and «the methodology of research becomes the methodology of learning» (Murison-Bowie, 1996: 190).

The link between corpus linguistics and SLA research is somewhat more controversial. Granger (1998: 4) points out that current SLA research has relied primarily on introspection and elicited data. There are three main reasons why natural language use has been disfavoured. First, some language features are not frequent and may occur only if elicited. Second, it is almost impossible to attain a systematic account of the variables that influence learners' output. And third, learners' output does not mirror their entire linguistic competence because learners tend to avoid the features that they do not master (*ibid*.: 4).

According to Granger, however, introspective and elicited

[3] The first pioneering project was launched by Collins Cobuild.

data also have a number of drawbacks. On the one hand, learners may produce a language that diverges from the language used in a natural context; on the other hand, generalisations of findings are not possible because of the limited quantity of empirical data used in SLA research. As a solution to the need for authentic data, Granger proposes the learner corpus. While native corpora mainly provide information on the frequency and use of items, learner corpora can help identify what is difficult for learners in general or for a given category of learners in particular (Granger, 2002: 21). Learner data may also provide information on the influence of the mother tongue on learning a target language.

Learner corpus research is seen by Granger as the approach to language learning which could make SLA specialists adopt empirical data in a more extensive way and which could act as a spur to more interdisciplinary collaboration between foreign language pedagogy and SLA research. Interdisciplinarity is one of the most urgent challenges facing learner corpus research. Granger states that «for learner corpus research to realise its enormous potential, cooperative involvement on the part of SLA [and] ELT [English Language Teaching] […] researchers would seem to be essential» (*ibid.*: 28).

1.4. Computer learner corpora

Different categories of people may be interested in collecting learner data. A teacher aiming at adapting teaching to the needs of her/his students may decide to gather their written productions in order to discover the areas in which they require further instruction. SLA researchers and linguists dealing with the use of language for specific purposes may also find learner corpora extremely useful. By comparing data across learner languages, SLA researchers may determine to what extent the mother tongue influences the learners' output. Linguists may collect specific learner corpora to answer questions concerning restricted varieties of learner language. Finally, learner corpora may be collected by publishing houses which intend to produce more learner-aware materials.

Learner corpus research adopts the same principles, tools and methods as other corpus linguistics inquiries (Granger, 2002: 4). Section 1.4.1. illustrates significant aspects of learner corpus research such as learner corpus design criteria, methodological approaches and software tools used for learner language analysis.

1.4.1. Design criteria

Learner language is not a homogeneous variety of language (Granger, 1998: 7). Therefore, when collecting a learner corpus it is necessary to identify and encode as many variables as possible. Sylviane Granger (1998; 2002), the founder and coordinator of the "International Corpus of Learner English" (ICLE), has identified the main features that need to be considered in assembling a learner corpus (see Table 1). She distinguishes between those variables which pertain to the learner and those variables which pertain to the language situation.

Table 1. Design criteria: variables which pertain to the learner and variables which pertain to the language situation.

Learner variables	*Language situation*
Age	Medium
Sex	Genre
Mother tongue	Topic
Region	Technicality
Other foreign languages	Task setting (exam, use of reference tools, etc.)
Proficiency level	
Learning context (second or foreign language)	
Practical experience (period of time spent in a target language-speaking country, etc.)	

Compiling a learner corpus according to explicit design criteria has two important practical implications. The first one is that it is possible to assemble subcorpora to accommodate different research purposes. For instance, a researcher may be interested in comparing the output of learners who used reference tools with that of learners who did not, or the productions of learners belonging to different age groups in order to check how learner language evolves. The second practical implication is that learner data can be shared among researchers. Without standardisation and documentation of data, findings can be neither verified nor compared; therefore, it is very important to provide clear information about the criteria adopted in compiling a learner corpus. Moreover, data sharing is also at the basis of international collaboration, which is essential in answering questions involving comparisons among non-native varieties of a language.

1.4.2. Methodological approaches

A computer learner corpus can be exploited in two main ways: Contrastive Interlanguage Analysis (CIA) and computer-aided error analysis (Granger, 2002: 11-14).

CIA involves the comparison between native (NS) and non-native (NNS) productions or among non-native data produced by learners with different mother tongue backgrounds. CIA has been proposed by Granger (1998) as an improved alternative to early contrastive studies known as Contrastive Analysis (CA) (Lado, 1957). CA aimed at predicting learner difficulties through detailed contrastive studies of the first and the target language. The results of CA would subsequently be applied to language teaching, syllabuses and tool design (Ringbom, 1999: 489). However, CA failed to offer real help to foreign language pedagogy because it was based on flawed theoretical and methodological assumptions (*ibid*.: 490).

In CA research the errors made by learners were hypothesised as being the result of the interference of their mother tongue. However, learners make errors which are not necessarily related to differences between the target and the native language (*ibid*.: 490). In addition, it became apparent that differences across languages do not equal learning difficulties. A second problem was the idea in CA to compare whole language systems turned out to be impossible to put into practice. Comparison was only possible on a micro-level, such as the level of phonology or morphology. Finally, another problem in CA was that the comparison of language items or constructions was carried out by ignoring the context of use of those features. In other words, early CA focused on the language system disregarding the actual output of learners (*ibid*.: 490).

The approach proposed in CIA can address most criticisms levelled at CA. First of all, CIA is not concerned with abstract language systems but aims at comparing how native and non-native speakers use language in similar situations. Therefore, language items and constructions are analysed in terms of the functions that they perform in context.

NNS/NNS comparisons have as their objective an improved knowledge of learner language. By comparing interlanguages, researchers are able to determine the areas which are affected not only by negative but also by positive transfer from the mother tongue.

NS/NNS comparisons, on the other hand, can help detect areas of non-nativeness, not only in terms of errors but also in terms of over- and underuse. NS/NNS contrastive analysis needs a control native corpus to measure the divergence between the output of learners and native speakers. The choice of the most appropriate control corpus requires particular care. Although many native corpora are now available it is important to know that not all text types are comparable. The reason is that certain language features are style-sensitive, that is, they may vary according to text types and context. Therefore, choosing a control corpus means paying attention to the text type it contains; if not, distorted findings may be reported. The monitor corpus selected for comparison, however, should not be considered the only possible model against which learner language is measured. As a consequence, labels such as "overuse" and "underuse" should not be interpreted prescriptively but rather descriptively (Leech, 1998: xix). However, as the aim of learner corpus analysis might be to provide data to be exploited by teachers and EFL tool designers, it is essential to define a norm to which learners should conform (Granger, 2002: 13). Thus, if properly chosen, the monitor corpus may be one of the acceptable norms.

The second type of methodological approach is computer-aided error analysis. As mentioned in section 1.3.1, the error-oriented approach of learner corpus research is not the same as previous EA studies. At present, corpus researchers have two methods available. On the one hand, they can use software programs to scan the corpus and identify all instances of misuse of a selected item. This method is consistent and quick, but it obliges linguists to focus only on error-prone items which are already known to them. On the other hand, researchers may opt for error tagging the learner corpus, by either using an existent error editor, i.e. a software tool which supports the activity of tag assignation, or devising their own standardised system. This method is extremely demanding and time-consuming, but, once the corpus has been error tagged, it has the advantage of allowing researchers to find out where learners have difficulties (Granger, 2002: 14).

1.4.3. Software tools for learner language analysis

1.4.3.1. Learner corpus annotation

The annotation of interlanguage raises the methodological question of whether specific taggers have to be created. Meunier (1998: 21) argues that advanced data does not need specific taggers and can be easily annotated using software created for native language corpora. Interpretative labels can be added automatically or semi-automatically, depending on the linguistic aspects for which a corpus is annotated.

Part-of-Speech taggers are fully automatic; this means that researchers cannot modify software parameters. As a consequence, if researchers aim at carrying out a very refined analysis they will have to choose a POS tagger with a fine-grained tagset in terms of the number and types of tags provided by the programme (*ibid.*: 20). This kind of annotation allows researchers to search for a given part of speech in learners' output and in particular for error-prone categories. A linguistic analysis carried out on a POS tagged learner corpus may also highlight the main shortcomings of interlanguage in terms of style and syntax (Granger, 2002: 18).

The syntactic constituents of a text can be further analysed with the help of a syntactic tagger or "parser". Syntactic tagging is generally a semi-automatic kind of annotation; therefore the process of adding labels is more time-consuming. Nevertheless, parsers can provide more sophisticated syntactic analyses (Meunier, 1998: 21). Semantic tagging and discoursal tagging are both in the initial stages of development. Semantic tagging is seen as potentially very useful for identifying areas of lexical poverty. Discoursal taggers still do not have enough evidence to prove their potential, because they have only begun to be developed. An example, however, is the software "Xanadu" (University of Lancaster) which has been used to mark cohesive relationships (*ibid.*: 25).

Error tagging is the only type of annotation which has been devised exclusively for learner language. It is manual, but it is considered fundamental if one wishes to provide adequate analyses of the peculiar nature of interlanguage. The grammar and style checkers which have been designed to exploit native language data fail to account for the errors of learners; hence, special checkers have to be developed. In order to design checkers suitable for learner language «one needs to have access to comprehensive catalogues of authentic learner errors and their respective frequencies in terms of types and tokens» (Dagneaux et al., 1998: 165). In other words, researchers need to work on error tagged corpora if they wish to provide adequate data for software designers. Error categories differ depending on whether errors are described from the point of view of their nature (e.g. spelling, grammar, vocabulary) or their source (e.g. L1-

induced errors, overgeneralization). For this reason, there are various possible error editors, even though it is preferable to analyse errors according to their nature because of the subjectivity involved in describing the source of the error (*ibid.*: 166). The information that an error tagged learner corpus can provide may be exploited to design not only style and grammar checkers, but also most types of pedagogical tools like learner dictionaries, grammars and textbooks.

1.4.3.2. Automatic approaches to analysing learner language

There are different types of analysis that automatic approaches can carry out. Meunier (1998) illustrates them in detail distinguishing between statistical, lexical, grammatical and syntactical analyses. There are also different software tools that researchers can use, such as style and grammar checkers and text retrieval software. I will concentrate on text retrieval software because they are «the type of software which has achieved the most startling results» (Granger, 2002: 15) and because this is the type of computer tool I used for my analysis.

The majority of text retrieval software products can provide general statistics such as the number of words in a text, the number of sentences and the average number of words per sentence. However, it is lexical analysis that may benefit the most from this kind of tool. With the help of various facilities, such as wordlists, concordances, type/token ratio calculation and distribution or dispersion graphs, text retrieval software allows linguists to carry out frequency analyses, context analyses and variation analyses.

Wordlists and dispersion graphs provide information concerning the frequency of occurrence of a given language feature. Wordlists display all the words of a corpus in various ways (alphabetical order, from the most frequent to the less frequent word, etc.) and are used by learner corpus researchers to compare native and non-native data. In this way, it is possible to determine which items are over- or underused.

Dispersion graphs, on the other hand, measure the distribution or dispersion of a selected item in the whole corpus. For instance, if a learner corpus is made up of samples from learners with different first language backgrounds, dispersion graphs show which group of learners uses that particular item most frequently. This quantitative information may reveal the role of the influence of the mother tongue on the learners' output. It is probable that learners who overuse the selected item have been influenced by their first language, if their group is the only one to show a significant overuse.

Concordances are used to study items in context, with the possibility of varying the size of the context and sorting the lists differently. This facility is very useful because «it throws light on the collocates or patterns that learners use, correctly or incorrectly» (Granger, 2002: 15).

Type/token ratio calculations may suit lexical frequency studies, especially to determine the degree of lexical variation. The number of types, i.e. the number of different words in a text, is divided by the number of tokens, i.e. the number of running words. If the type/token ratio is high, it means that there is a great deal of lexical variation. It should be pointed out, however, that the type/token ratio varies greatly according to the length of texts. Moreover, it is important to remember that "[a]dvanced learners' lexical problems are not due to a lack of vocabulary but rather to the inappropriate use they make of the words they know» (Meunier, 1998: 32). Therefore, the type/token ratio needs to be interpreted in the light of results obtained through the use of wordlists and concordances which are probably more reliable in providing significant quantitative data to complement qualitative interlanguage analyses.

Chapter II

Evaluation: The Study of Language Expressing Opinion

2.1. Concepts and terminology

As pointed out by Stubbs «utterances always encode a point of view» (1996: 197). In addition to propositional information, writers/speakers also convey their opinion concerning that information. Every act of evaluating something can be done along a positive-negative scale. However, the positive-negative parameter is a basic parameter and there are several different sets of values that can be related to it. Speakers can signal how desirable or undesirable something is; to what extent something is certain or uncertain; expected or surprising; important or irrelevant (Thompson and Hunston, 2000: 25) (see section 2.2.4). This is one of the reasons why the ways of expressing opinion have been studied from various perspectives as illustrated in the Introduction.

In this study, following Hunston and Thompson (2000), I take a "combining" approach and I use "evaluation" as the superordinate term. As Thompson and Hunston admit, this label is «as slippery as any of the others in this field» (*ibid.*: 5). The reason is that it has already been employed to refer to particular aspects of language expressing opinion, and thus it could be confusing. On the one hand, it has been used in a similar way as "attitude" in relation to lexis (Carter, 1987). On the other hand, it has been adopted in the area of discourse analysis. "Evaluation" is a term which has been used to refer to a constituent of the macro-structural configuration of a text (Labov, 1972; Hoey 1979; 1983) (see section 2.3.3)[1].

Thompson and Hunston, however, point out that the term “evaluation” has an advantage over other cover terms, that is, its flexibility of use. They argue that this label not only implies that it is the user who evaluates, «but it also allows us to talk about the values ascribed to the entities and propositions which are evaluated» (2000: 5). This flexibility is particularly welcome as it allows researchers to account for a wide range of aspects such as «why, when, how, and what speakers and writers evaluate» (*ibid.*: 6).

Since the focus of my study is on adverbs functioning as adverbials (see Chapter 3 for a detailed account of these features), I also keep the terminology used by Biber and Finegan (1988: 1989). Biber and Finegan's (1988) study, in particular, is the source from which the list of stance adverbs was taken for the present analysis (see Appendix 1). The label “stance” has also been adopted in the *Longman Grammar of Spoken and Written English* (*LGSWE*) (Biber et al., 1999), which I follow for its classification of these adverbials in three categories: epistemic, attitudinal and style stance adverbials.

The present chapter is organised as follows. Section 2.2 illustrates the phenomenon of evaluation in terms of the functions that it may perform and the scales of values along which speakers assess or judge a proposition or an entity. Section 2.3 provides information about the linguistic devices used to convey evaluative meanings. The most important aspects of evaluative lexis are dealt with, and information on the grammatical marking of evaluation is provided. Section 2.3 also shows that evaluation is a complex and pervasive phenomenon which occurs at the clause, sentence and discourse levels.

---

[1] Discourses are organised through relationships of macrostructures that constitute conventionalised patterns (e.g. the Problem-Solution-Evaluation pattern).

2.2. Functions and parameters of evaluation

Thompson and Hunston (2000: 6-13) claim that the three main functions that “evaluation” performs are to express opinions, to build up and maintain relations and to organise the discourse.

Every act of evaluation may perform more than one function at the same time. For instance, the conjunct *though* in example 2 is used to express contrast with what has been said, thus indicating the partial attachment to the truth-value of the previous proposition on the part of the writer. In other words, it indicates that the writer judges the previous proposition, and the opinion expressed in it, as only partially sharable. At the same time, the conjunct also to structures the text by making the relationship between sentences explicit and guide readers through the argument (see example 2)

> (2) The precise size of the world economy may not matter much from a policy point of view, though a $14 trillion difference is hardly small change. [EMA]

It is also possible that a single evaluative function may be performed by various linguistic devices operating at different levels of language. For example, the writer's degree of attachment to the truth-value of a proposition can be expressed using modal auxiliary verbs (e.g. “it *may* be true that....”), adverbs (e.g. “it is *probably* true that...”) or modal lexical verbs (e.g. “the results *suggest* that....”).

2.2.1. Expressing opinion

The first function of evaluative language is to express opinions. The opinions expressed by individuals certainly reflect their personal points of view; however, individuals' opinions are also rooted in the value-system of the socio-cultural context in which individuals live and communicate. Therefore, the study of the expression of opinion is not simply the study of how an individual takes a position; it also is the study of how every act of evaluation builds on a communal value-system and how it reinforces or undermines that value-system (Van Leeuwen and Wodak, 1999). By identifying instances of evaluation it is possible to define the value-system, or ideology, of the society in which a text has been produced.

The examples which I use here to illustrate this phenomenon are taken from Stubbs (1996: 98-100) who investigated how sexist ideological positions were conveyed by the British general Baden-Powell in his last message to Boy Scouts and Girl Guides. These messages were published in 1942 and 1941 respectively, and had a very wide circulation (*ibid.*: 83).

In a rather patronizing tone, the two messages express the idea that men and women have very different roles in society. Men are described as belonging to the public sphere, where happiness is seen as linked to career and money (example 3, words in italics). On the other hand, women are defined in their social role as acting within the private sphere as wives or mothers (example 4, words in italics).

> (3) I believe that God put us in this jolly world to be *happy* and *enjoy* life. *Happiness* doesn't come from being *rich,* not merely from being *successful in your career*, not by *self-indulgence*. One step towards *happiness* is to make yourself healthy and strong while you are a boy, so that you can be useful and so can *enjoy* life when you are a man.

> (4) Later on, when you have a *home* on your own, by making it a bright and cheery one you will make *your husband* a *happy* man. If all homes were bright and cheery, there would be fewer public houses and the men would not want to go out to them but would stay at home. It may mean *hard work* for you, but it will bring its own rewards then, if you keep *your children* healthy and clean and busy they will be *happy*. *Happy* children love their parents. There is nothing can give you greater *joy* than a loving child.

Ideologies can be recognised in the use of lexical items which are value-laden at a specific moment for a given society. For instance, as Stubbs observes (1996: 179), in English the word “career” (example 3) tends to have a strong positive semantic prosody[2]. It generally co-occurs with words such as "political", "international", "managerial", "brilliant", "distinguished", "promising" which convey the idea that the word "career" indicates a high-prestige profession. "Career" is used by Baden-Powell in the message to Boy Scouts and not to Girl Guides. This choice indicates that the value-system which influenced the British general and which, in turn, was reinforced by his words, is one in which it is desirable for (or even expected from) men – but not so for women – to fulfil themselves professionally.

Stubbs also provides evidence for the sexist ideology of the two texts by studying how words that are not intrinsically sexist, such as "happy" and "happiness", are used in Baden-Powell's texts. Through a concordancer, Stubbs finds the grammatical frames in which "happy" and "happiness" occur. Their syntactic contexts vary according to whether Baden-Powell addresses girls or boys. In the former case, the concept of "happiness" involves other people, notably husband and children (e.g. "you will make your husband a *happy* man"; "if you keep your children healthy and clean and busy they will be *happy*"). In the case of boys, the concept of happiness appears as something that has to be sought for one's own sake (e.g. "one step towards *happiness* is to make yourself healthy and strong").

Stubbs' analysis suggests that ideologies can be conveyed directly, by socially significant words, but also indirectly, by patterns of co-occurrence of lexis and grammar.

---

[2] "Semantic prosody" is an expression principally theorised by Louw (1993) to refer to the meanings that typically occur with a word or phrase. These meanings may be positive or negative in their "evaluative polarity" (Channell, 2000: 41), that is, a given lexical item may assume a positive or negative association. A frequently cited example is Sinclair's (1991: 74) analysis of the phrasal verb "set in" which assumes a negative association because it generally occurs with subjects indicating something which is undesirable or unattractive, such as "rot", "decay", "infection" and "prejudice".

2.2.2. Maintaining social relations

The second function of evaluation is to establish and maintain relationships. According to Thompson and Hunston (2000: 8), this function has mainly been investigated in relation to the study of persuasion and manipulation, and the study of "hedging" and "politeness" (Brown and Levinson, 1978). These areas have often been explored in connection with the study of academic writing (Myers, 1989; Thompson and Ye, 1991; Thompson, 2001; Charles, 2003; Hyland, 2005a).

In academic writing, scholars need to engage in a dialogue with their audience in such a way as to be perceived as credible and trustworthy. This is one of the conditions for writers to succeed in their attempt to persuade readers (Hyland 2005b)[3]. An example of the way scholars try to guide readers towards preferred interpretations is the use of anaphoric nouns (Francis, 1986; Charles, 2003). The anaphoric noun "this claim", for instance, summarises the previous stretch of text and evaluates what has been said as being "a claim"; at the same time, it guides readers through the content of the text by enhancing cohesion. Anaphoric nouns, therefore, can be considered resources to express opinion and, at the same time, to help maintain the relationship between writers and readers by guiding readers through the text and by elaborating propositional meaning according to the writer's point of view (Hyland, 2005a: 49).

The persuasion of readers can also be obtained by placing the expression of evaluation in the less prominent position in the clause, so that readers will simply accept its validity as a

---

[3] Hyland (2005b) devised a model, stance and engagement, which accounts for how writers convey their opinions, judgements and commitments (i.e. stance), and for how writers relate to their readers (i.e. engagement). Although Hyland treats stance and engagement separately, he observes that the interplay of writer stance and reader engagement enables writers to be persuasive and to connect with the value-system in which they operate (*ibid.*: 175). Engagement resources are evaluative in that they reflect the writer's assessment of the audience in terms of their need for guidance, their background knowledge and their possible reactions to the message being conveyed.

matter of fact. Hoey (2000: 33) carries out an analysis of Chomsky's writing to show how the American linguist tried to persuade readers and pre-empt criticisms towards his position. Hoey points out that Chomsky's evaluations are placed in the structure of the clause in places which leave little room for negotiation. This strategy is extremely successful in manipulating the reader. For instance, Chomsky places his opinion in the "given" position of the sentence or in premodifier position within nominal groups thus treating it as something which does not come into question (example 5).

> (5) The answer to this *essentially terminological* question seems to have no bearing at all on any *serious* issue. [Example taken from Hoey (2000: 33)].

In the study of academic writing, a great deal of attention has been paid to "hedging" (Myers, 1989; Hyland, 1998). Hedging is an epistemic device which, on the surface, allows writers to express likelihood and to show that, to some extent, a proposition is imprecise. However, Myers (1989) considers hedging as a politeness strategy which can also be used to minimise the possibility of criticism and establish an appropriate relationship with readers rather than to express probability or uncertainty.

He provides the following examples: "Thus, a common short sequence of RNA *might* be attached to several mRNAs [...]" and "The three short segments are *probably* spliced to the body of this mRNA" (*ibid.*: 13). Myers points out that in the above examples the hedging does not reflect the personal doubt of the author, but rather the most appropriate way of offering a claim so that it can be accepted by the other members of the scientific community (*ibid.*: 12).

2.2.3. Organising discourse

The third function that evaluation performs is to organise discourse. As mentioned above, writers construct their texts according to the supposed needs and reactions of readers; this awareness of the audience allows writers to establish the most appropriate relationship with readers in order to persuade them. Some of the resources used for this purpose, however, may also signal the way in which a text is organised. The stance adverbial “actually” is a case in point. Tognini-Bonelli (1993) shows that “actually” and the corresponding adjective “actual” have a structural role in discourse. In other words, “actually” and “actual” may be used to link two elements and, at the same time, to differentiate them in terms of their evaluative status.

For instance, in “innovations that *actually* address health priorities” the introduction of “actually” in the relative clause implies a differentiation between “innovations in the field of public health” in general and those “that *actually* address health priorities” on the other. Readers are in this way warned implicitly of the fact that there may be innovations in the field of public health that not necessarily address health priorities. Therefore, “actually” has an evaluative role in that it expresses the writer's judgements on innovations in the field of public health. At the same time “actually” also establishes a correspondence between two distinct elements, i.e. the two types of “innovations”. Therefore, “actually” can also be considered as a device which maintains discourse coherence.

Evaluative comments contributing to the organisation of discourse tend to occur at boundary points. They enable writers to signal the beginning and the end of a paragraph or a discourse section, and to assume that readers are following the progression of the text. Thus, evaluation organises the text in such a way that a monologic piece of writing can be seen as a dialogue between the author and her/his readers.

2.2.4. Parameters of evaluation

When speakers express their evaluations they may be assessing an entity (the city of Barcelona in example 6) or a proposition (all that comes before the comma in example 7). This evaluation reflects the parameters of their judgement.

Thompson and Hunston (2000: 22-26) distinguish among four main parameters: “goodness”, “certainty”, “expectedness” and “importance or relevance”.

(6) Barcellona is *perfect* for a summer break […]. [Example taken from Thompson and Hunston (2000: 1)].

(7) The new 1.8 engine is worth a second glance, *it seems*. [Example taken from Biber et al. (1999: 865)].

Conrad and Biber (2000) have provided evidence in favour of the hypothesis that different parameters may be associated with different registers. For instance, epistemic stance adverbials such as “perhaps” and “probably” are more frequent in academic writing than in news reportage (*ibid.*: 63). An explanation for this may be that academics usually pay attention to the degree of certainty attached to their knowledge claims (*ibid.*: 65). Evaluation of “goodness” (“amusingly”, “delightfully”) is typical of genres such as film or book reviews whose main purpose is to assess whether a film is worth seeing or a book is worth reading. Both evaluations of “goodness” and “certainty” can be described as having a “real-world-oriented” function (Thompson and Hunston, 2000: 24), that is, they depend on the sets of values that underlie the discourse and on the experiential knowledge of speakers. Evaluations of “importance” (“more importantly”, “significantly”) and “expectedness” (“surprisingly”, “as may be expected”), too, are based on experience, but they have an additional function that is defined as “text-oriented” (*ibid.*: 24). By signalling that something is significant or that the parts of a discourse are related in terms of what is expected, speakers indicate to listeners the intended coherence of what they are saying (*ibid.*: 24). Evaluations of “importance” and “expectedness”, therefore, help speakers to organise their discourse. Indications of relevance, in particular, as they tend to occur at boundary points, enable writers to guide readers through the content of the text. For example, in Swales' (1990) well-known *Create a*

Research Space model which accounts for the semantic macro-structure of research article Introductions, the first rhetorical Step that academic writers may take is claiming centrality for their research topic.

Although it is possible to distinguish among these four parameters, "goodness" is the parameter to which the others can be related. Evaluations of "certainty", "expectedness" and "importance", in turn, may be assessed as good or bad. However, the assessment of each of the four parameter as good or bad is a cultural fact. For instance, it is culture that determines whether the expression of certainty when constructing an argumentative line is something positive or negative (Salager-Meyer et al., 2003). No parameter of "evaluation", therefore, is «neutral with respect to cultural value» (Thompson and Hunston, 2000: 25).

2.3. The identification of the ways in which evaluation is expressed in English

From a conceptual point of view, identifying instances of evaluation means identifying signals of comparison, subjectivity and social value. Linguists, however, have tried to recognise evaluative meanings not only conceptually but also in terms of their linguistic realisation at the lexis, grammar and discourse levels. The following sections deal with each of these language levels individually.

2.3.1. Lexis

At the lexical level it is possible to distinguish between adjectives, adverbs, nouns and verbs which are clearly evaluative and lexical items whose main function is not evaluation. Adjectives such as "excellent", "certain", "possible", and adverbs such as "astonishingly" belong to the group of items whose chief purpose is to evaluate. In many cases it is possible to distinguish between lexical items which are clearly value-laden (e.g. "spinster") and other which are not (e.g. "single woman"), although in some contexts they may acquire evaluative status.

It is also possible to identify lexical items which, along with evaluative meanings, provide other types of information. In that case, it may be more difficult to determine whether the status of evaluation is positive or negative. Consequently, the context in which those items are used plays a key role. Thompson and Hunston (2000: 17) illustrate this aspect of evaluative lexis through the example of verbs such as "collaborate", "participate", "engage", "join", "collude", "interfere", "meddle", "assist" and "help". When these verbs are followed by the preposition "in", they have the same chief function, that is, to indicate involvement and participation in some activity. The verbs "help" and "assist" generally evaluate the participation positively. However, when the action in which the people are involved has negative consequences, the evaluation of the act of helping may be positive for those who have received help and negative for those who have been damaged by the action. For instance, "to assists in a bank robbery" is a positive action from the perspective of the robbers, but negative from the point of view of the bank staff. As Leech points out, «the "putative properties" of the referent [are] due to the viewpoint adopted by any given individual» (1974: 14). As a consequence, it may appear that evaluation is «somehow incidental to language rather than an essential part of it» (*ibid.*: 15) and that it is «somewhat randomly dispersed across a range of structural options shared with non-evaluative functions» (Thompson and Hunston, 2000: 74). However, as I illustrate below, evaluation may be explored not only in lexical terms but also from the point of view of grammar, and this will provide a more systematic account of the phenomenon.

2.3.2. Grammar

Linguists have tried to identify the grammatical categories whose main function is to express evaluation. Biber and

Finegan (1989) analyse to what extent different types of texts favour different grammatical categories for the marking of stance. They identify «markers of affect (both positive and negative) and of evidentiality (both certainty and doubt) in the four grammatical categories of verbs, adjectives, adverbs and modals» (*ibid.*: 95). Following semantic and grammatical criteria they divide stance markers into twelve categories:

1. Affect markers (adverbs, verbs, and adjectives): e.g. "astonishing", "can't stand", "fascinated", "afraid";
2. Certainty adverbs: e.g. "decidedly", "certainly", "obviously";
3. Certainty verbs: e.g. "deduce", "know", "prove", "realise";
4. Certainty adjectives: e.g. "evident", "well-known", "incontestable";
5. Doubt adverbs: e.g. "apparently", "presumably", "supposedly";
6. Doubt verbs: e.g. "feel", "imagine", "gather";
7. Doubt adjectives: e.g. "arguable", "possible", "unlikely";
8. Hedges: e.g. "almost", "sort of", "more or less";
9. Emphatics: e.g. "a lot", "most", "such a";
10. Possibility modals: "can", "might", "could", "may";
11. Necessity modals: "ought", "should", "must";
12. Predictive modals: "will", "would", "shall".

Traditionally, the grammar of modality has been far more investigated than the grammar of affect. Modality is realised by features which are more closely linked to the structure of the clause, such as modal verbs (e.g. "it *may* be argued that..."). Opinions of desirability, on the other hand, depend on lexical items, such as certain adjectives (e.g. "beautiful", "great", "interesting", "terrible") and nouns (e.g. "beauty", "interest", "happiness"). It is often believed that a structural description of affect is not feasible because lexical items do not constitute a closed class (Hunston and Sinclair, 2000). Recently, however, some researchers have tried to account for the sets of structures that are involved in the expression of affect too.

Hunston and Sinclair (2000) attempt at providing a structural account of evaluation which also includes affective meanings. They work with data taken from the "Bank of English" and their aim is to provide a "local grammar" of evaluation. A local grammar may be defined as a set of rules and categories that can describe certain constructions and their functions in a more accurate way than a general grammar. One reason is that general grammars need to cover most language patterns; therefore they tend to focus on highly general features and to disregard those language phenomena which apparently do not show significant regularities. Hunston and Thompson argue that except for modality, evaluation does not have its own grammar because it is a function of language that "appears parasitic on other resources" (2000: 74). However, according to Hunston and Sinclair (2000) the perspective of a local grammar can gather evidence that evaluative meanings, too, can be accounted for in a systematic way. For instance, in the sentence "it seemed important to trust her judgement", the adjective "important" is categorised by the two scholars as "evaluative category" and the non-finite clause "to trust her judgement" as the "thing evaluated". These labels are applicable to a number of constructions or patterns, as shown in Table 2.

Table 2. Example of descriptive categories of a local grammar of evaluation: "Evaluative category" and "Thing evaluated"

		Evaluative category	*Thing evaluated*
It	was	certain	that he was much to blame
They	would be	sensible	to say "yes"
You	are	right	that he didn't go to the apartment when he said he did

Hunston and Sinclair's first objective is to identify all the patterns in which evaluative adjectives occur[4]. They observe that the adjectives that share a pattern also share meaning, which indicates that all the adjectives which may be used in a given pattern can be considered evaluative adjectives, although the type of evaluative judgement may vary.

For instance, the adjectives that are likely to occur in the pattern "it+ link verb+adjective group+ clause" (e.g. "it was wonderful talking to you the other day") are: "fortunate", "heartening", "splendid", "wonderful"; "awful", "stupid", "terrible"; "important", "necessary"; "common", "odd", "interesting", "relevant", "significant", "surprising"; "certain" (*ibid.,* 84-85). All these adjectives are evaluative, but the judgements they express are varied, namely judgements of desirability, importance, expectedness and certainty. This flexibility in accommodating different types of judgements implies that items expressing modal meanings share this pattern with items expressing affective meanings. Therefore, not only are there areas of overlap between the two semantic spheres, but affective meanings, too, seem to be organised around a definable number of constructions which can be studied systematically.

The identification of patterns in which evaluative adjectives occur is the first step towards developing the descriptive categories of a local grammar of evaluation. Hunston and

---

[4] The patterns are:

1. "it + link verb + adjective group + finite or non-finite clause" (e.g. It is *disturbing* to see the line of vehicles filled with tourists driving along the sand on busy weekends [...] [EMA]);
2. "there + link verb + something/anything/nothing + adjective group + about/in + noun group/ing- clause" (e.g. There is nothing *new* about young people not voting [...] [EMA]);
3. "link verb + adjective group + to-infinitive clause" (e.g. [...] we are *right* to be afraid of the stranger [...] [EMA]);
4. "link verb + adjective group + that-clause" (e.g. [...] the odds are *good* that such records will become the national standard [...] [EMA]);
5. pseudo-cleft clauses beginning with "what" (e.g. What is *remarkable* is that these two rare species live together in the same cave. [EMA]).
6. patterns with general nouns like "thing" or "point" (e.g. The key *thing* is that it does not respect familiar boundaries [...] [EMA]).

Sinclair's ultimate goal is to use the categories thus identified to parse computer corpora. They argue that in so doing, researchers will be able to retrieve evaluative adjectives which otherwise could not be distinguished from non-evaluative adjectives because traditional grammar tags ignore the function of evaluation.

2.3.3. Discourse

Finally, the phenomenon of evaluation has also been approached from the perspective of discourse analysis. In particular, the study of the ways in which subjectivity is realised in discourse is one of the main concerns of linguists dealing with different discourse types. In narrative texts, evaluation plays a central role, as it conveys the «narrator's attitudes and feelings towards the sequence of events narrated» (Georgakopoulou and Gustos, 1997: 129). In Labov's (1972) well-known structure for oral storytelling, Evaluation is the passage which clarifies why a particular narrative is worth telling (the pattern comprises six elements: Abstract, Orientation, Complicating Action, Resolution, Evaluation and Coda).

Evaluation is also part of Hoey's (1979; 1983) Problem-Response (or Problem-Solution) rhetorical pattern of non-narrative texts. The Problem-Response pattern is one of the most common patterns in non-narrative texts. This pattern, which may be considered a minimum discourse (Hoey 1983: 31), tends to be preceded by a part which illustrates the Situation and it is followed by a part, known as Evaluation, which assesses the Response as successful or unsuccessful (Situation-Problem-Response-Evaluation)[5].

A number of scholars, however, agree that evaluation is not confined to a separate section of rhetorical patterns (Hoey: 1979; Jordan 1984; Hunston, 1994; 2000; Thompson and

[5] Given that most Responses are evaluated positively, this pattern is generally referred to as the Problem-Solution pattern (Hoey, 1983: 52).

Hunston, 2000). Hoey (1979: 54) points out that the evaluation that occurs at the discourse level as a member of the basic textual pattern Problem-Solution-Evaluation depends in part on the evaluation at the paragraph level, which is conveyed by a particular kind of clause relation which he calls “evaluation-basis”. The “basis” element is a proposition that supports the evaluative claim, providing intermediate justification for the final statement of stance.

In line with Hoey's view, Jordan (1984) argues that the Problem or Solution section may contain lexical items that anticipate what readers will find in the Evaluation section. He calls these signals of the writer’s opinion “pre-evaluation” (*ibid.*: 110). For instance, if a writer describes something as a “solution”, she/he is anticipating that a problem has successfully been overcome. “Pre-evaluation” may also function as a device to encourage the reader to find out more about the issue.

These studies suggest that although evaluation may occur in particular strategic points in the semantic macro-structure, it is a pervasive phenomenon which is built up progressively and cumulatively across the text, and which occurs in all clauses.

Chapter III

Stance adverbials in English learner writing

The present study aims to compare argumentative texts produced by Italian learners of English with texts produced by native speakers in terms of the expression of the opinion of writers, henceforth referred to as "writer stance". The linguistic features focused upon are stance adverbials made up of single adverbs commenting on the degree of certainty, doubt and actuality of propositions. An account of these features is provided in section 3.1. The rest of the chapter offers a description of the stance adverbials that most often occur in the argumentative text type (section 3.2). It also provides a detailed account of the corpora used in this study and the methodology adopted (section 3.3). Finally, section 3.4 presents the results of the analysis and a discussion of findings.

3.1. Stance adverbials: a structural and functional account

A clause containing a stance adverbial can be seen as comprising two distinct structural elements: one is the proposition that is being evaluated by the stance adverb, and the other is the adverbial expressing stance. Adverbials are always optional, in the sense that the sentence is grammatically correct even if they are omitted (Greenbaum, 1996: 142; Biber et al., 1999: 766). Their use, therefore, reflects the choice of writers to express their comments overtly.

In the *LGSWE* (Biber et al., 1999: 762-892), adverbials are classified in terms of their semantic function. Adverbials may comment on the circumstances in which an action is being carried out (i.e. circumstance adverbials) (example 8); they may connect the clause to other parts of the text/discourse (i.e. linking adverbials) (example 9); or they may express speaker/writer stance towards the clause (i.e. stance adverbials) (example 10). In this study, I only concentrate on adverbials expressing stance, as I am interested in the expression of the opinion by writers[1].

(8) I'll see you all *tomorrow night.*

(9) *As a result,* more and more families are becoming bitter toward welfare.

(10) That sort of gossip *certainly* should be condemned[2].

Stance adverbials, too, may be classified on the basis of their semantic category (Biber et al., 1999: 854-857). Epistemic stance adverbials convey the writer's comments on:

1. The degree of certainty of a proposition:

(11) *No doubt* his bifocals added to this impression, as did his nonchalant gait and slouchy posture.

(12) *Maybe* it's true, *maybe* it isn't.

---

[1] Arguably, evaluative meanings may be conveyed not only by stance adverbials, but also by linking adverbials, which may be interpreted as fulfilling the function of organising discourse (see Chapter 2, section 2.2.3). Moreover, evaluative meanings can also be identified in circumstance adverbials (e.g. "After *intensive* test, they believe that AIDS virus had been eradicated from the patient's body" – example taken from Biber et al. 1999: 763). Since in this study I am concerned with items which reflect the opinion of the writer on the *proposition* expressed in the clause, I judged the classification scheme provided by Biber et al. (1999) as fit for the purpose.

[2] The examples in section 3.1 are taken from Biber et. al (1999: 762-892). Adverbials are in italics.

2. The status of a proposition as a real-life fact:

(13) *In fact*, I'm taller than the doors.

3. The source of a proposition:

(14) Egypt's nuclear power industry is still in the design phase, but *according to Mr. Kandil,* nuclear power was the only clean energy alternative for Egypt.

4. The limitation of a proposition:

(15) *In most cases* he would have been quite right.

5. The point of view from which a proposition is true:

(16) *In our view* it would be a backward step.

6. The degree of precision of a proposition:

(17) Indeed, the only real drawback, *if you can call it that,* is that people are continually coming up and congratulating us on our victory over England.

Attitude stance adverbials convey the writer's attitudes towards the content of a proposition. The meanings that are most commonly expressed are value judgements (example 18) and assessments of expectations (example 19).

(18) *Unfortunately* I have too many of them but someday you'll be old too.

(19) And *most surprising of all*, much farther away still in western Australia, we have the gingin chalk of late cretaceous age.

Style stance adverbials comment on the manner of conveying the message (example 20).

(20) *More simply put*, a feedback system has its inputs affected by its

outputs.

Stance adverbials can also be classified in terms of their syntactic form. The possible syntactic forms are:

1. Single adverb:

(21) They had *evidently* been too scared of their autocratic director to record such an unlikely phenomenon.

2. Adverb phrase:

(22) But *quite frankly* I can't see myself ever getting – given the same sort of circumstances.

3. Prepositional phrase:

(23) His bedside manner was, *in a word,* menacing.

4. Noun phrase:

(24) Some will *no doubt* accuse Jarman of shock tactics along Warhol lines.

5. Finite clause:

(25) Well, then, I have come here to heal myself, *if you like to put it that way.*

6. Non-finite clause:

(26) *Based on studies of crop plants and native species grown under controlled conditions,* root growth responds at least as much, and perhaps more, to elevated carbon dioxide than does shoot growth.

In the present study, the semantic category selected for analysis is that of epistemic stance adverbials and, within this category, I focus on adverbials expressing doubt, certainty and actuality. As regards syntactic form, I concentrate on single adverbs.

These choices were made on the basis of the incidence of use of items in the four registers considered in the *LGSWE*, namely conversation, fiction, newspaper language, and academic prose. Among the three categories of stance adverbials, epistemic stance adverbials are by far the most common group in all four registers (*ibid.*: 859). In addition, if we exclude conversation, epistemic stance adverbials are employed most frequently in academic prose, a register in which the argumentative text type plays a major role (*ibid.*: 859)[3]. The most common epistemic stance adverbials are those marking doubt and certainty, followed by adverbials indicating actuality (*ibid.*: 869). Finally, single adverbs are the most common syntactic form across all four registers (*ibid.*: 862). The reason for focussing on the most common types of stance adverbials is that these are arguably the items that are more likely to be used by learners in their written productions and those for which they may be required to develop a good command of use.

3.2. The argumentative text type as marked by stance adverbials

Some genres, such as academic book reviews (Hyland, 2000), may favour the overt expression of writer stance while others, such as the methodology section of an experimental research article, may not, or at least not to the same degree. In addition, different genres may favour different sets of evaluative markers. For instance, personal memoirs are mainly character-

[3] The academic prose subcorpus is composed of book extracts and research articles. Almost all the texts are written for an expert readership. The sole exception is 16 book extracts which are written for a lay audience (Biber et al. 1999: 33). It is widely acknowledged that academic discourse in general and research articles in particular are argumentative and persuasive and display a significant use of evaluative language (Bondi and Mauranen, 2003; Mur Dueñas, Lorés Sanz and Lafuente Millán, 2010).

ised by attitude stance adverbials, whereas academic prose is characterised by a large use of epistemic stance adverbials (Biber and Finegan, 1988). Therefore, it is possible to describe a given genre in terms of a particular stance style.

Biber and Finegan (1988) identify eight "speech styles" of English as marked by stance adverbials analysing the "LOB" (Lancaster-Olso/Bergen) and "London-Lund" corpora (410 written and spoken texts). With the term "speech style" the two scholars refer to the frequent occurrence and co-occurrence of several linguistic features which characterise a particular set of texts in opposition to other texts characterized by the absence of those features (*ibid.*: 3). In their (1988) study, Biber and Finegan focus exclusively on stance adverbials (in a variety of syntactic forms) and use a multivariate statistical technique called "cluster analysis" which enables them to group the texts that are maximally similar in their frequency of use of stance adverbials. Each cluster, however, is interpreted by the two scholars as a speech style by taking into account not simply the frequent use of a particular type of stance adverbial, but also by considering the discourse functions of items in their sentential context and the overall communicative situation of the texts. This analysis enables Biber and Finegan to identify the following speech styles of English as marked by stance adverbials: "secluded from dispute", "emphatic shared familiarity", "faceless", "emphasis of individual position", "generalized content", "cautious" and "concession to reader/listener".

The identification of speech styles is a way of classifying texts in terms of shared linguistic forms. Biber and Finegan argue that the focus on linguistic form reveals similarities across texts that other types of classifications, such as that based on genre distinctions, may not account for. More specifically, they claim that «for some genre categories, greater linguistic differences exist among texts within the categories than across them» (*ibid.*: 3). For instance, a speech style as marked by stance adverbials that is characterised by "emphatic shared familiarity" seems to be typical of face-to-face and telephone conversation, but it may also characterise the stance style of interviews, spontaneous speech and the press (*ibid*.: 15).

Biber and Finegan's (1988) study includes the analysis of newspaper editorials. Editorials are characterised by a high degree of epistemic comments, in particular of certainty, doubt and actuality. However, some articles appear to be "faceless", i.e. they do not show a significant presence of stance adverbials[4].

In editorials, adverbials of certainty (called *surely* adverbials in Biber and Finegan, 1988) generally assume the existence of common ground between writers and readers, inducing readers to accept assertions and limit objections ("certainly" in example 27). In particular, they tend to solicit the readers' empathy and «seclude certain assertions from polite dispute» (*ibid*.: 19) (e.g. "there can be no doubt" in example 27).

> (27) It was a heavy, distasteful task that fell to Mr. Frank Foulkes yesterday. For Mr. Foulkes is, *of course*, the president of the Electrical Trades Union, and it was in that Union, and it was in that capacity that he announced the results of the elections for the membership of the union's general executive, in which the Communists have suffered an overwhelming defeat […] [he remarked] "I have always said that our members are always right until they have been proved wrong, even when they have taken unofficial actions against an employer."
>
> And this, *certainly*, is in accordance with the Communist creed, but now the members, that is, the rank-and-file members, have cut right across it. They have taken the democratic path, as a result of which it is very possible – but no more than possible at the moment – that the ETU may be readmitted to membership of the Trade Union Congress. Here, however, much, if not indeed all, may depend on the attitude of Mr Foulkes […] [he said] "if affiliation to Congress depends on my resignation, we will not be affiliated, I can assure you." This was, *of course*, a reference to the directive by the General Council […]
>
> *There can be no doubt*, however, that readmission is what the members, or at least the vast majority of members, of the ETU want

[4] The fact that certain newspaper editorials do not present a significant number of stance adverbials does not mean that these articles do not show other overt or covert signs of evaluative language. Biber and Finegan's "speech styles" refer to the use of stance adverbials and consequently should not be considered as representing the overall argumentative style of newspaper editorials.

[…][5].

Sometimes, certainty adverbials acknowledge that most readers might know the information being provided, but, at the same time, they supply that information for those readers who, it is believed, are still not cognizant with it (e.g. "of course", especially the first occurrence, in example 27).

Certainty adverbials also have a more specialised function. These items can be used to set up concessive clause-relations with the aim of conceding points to readers. The points conceded, however, are obvious objections that hypothetical readers might raise, and which oppose the writer's point of view. By conceding those points, the writer shows that opposing views cannot diminish the force of her/his argument (example 28).

> (28) It is, *of course*, true that the number of places – especially science places – in a university or college is, in the short run, fixed by physical conditions. But it is also true that, in the long run, […].

Doubt adverbials are typical of a cautious style that marks the stance that the evidence of an assertion is disputable (example 29). According to Biber and Finegan, this stance style is typical of academic prose and fiction, but may also characterise newspaper editorials (*ibid*.: 15-16).

> (29) *Presumably* domestic ritual objects began to be made at much the same time. The name of Rabbi Meir of Rothenburg […] [became well known] and *it may be assumed* that by his day Jewish ritual art as we know it now had begun to assume its form.

Actuality adverbials are used for two different purposes. On the one hand, they convey a sense of emphatic conviction (e.g. "it's *really* annoying"); on the other hand, they emphasise the individual position of the speaker or writer in opposition to other possible positions (example 30). Often actuality adverbials

---

[5] Examples 27-30 are taken from Biber and Finegan (1988).

occur in sentence-initial position and have a cohesive function (example 30, first occurrence of "in fact").

> (30) The point I wish to make is that all these casualties did not occur among the non-swimming members of our population. *In fact*, I venture to suggest it is probable that the grater proportion of these unfortunate people could swim and *in fact* might not have drowned had they been non-swimmers.

3.2.1. Research questions

Given that the texts contained in the Italian subcorpus of ICLE are argumentative texts, I would expect them to be characterised by the same stance styles as newspaper editorials, which belong to the argumentative text type too. However, there have been studies of learner language demonstrating that non-native speakers seem to have difficulties in modulating their writing style in accordance to the text type (Altenberg and Trapper, 1998: 91). Therefore, the hypothesis is that Italian learners will show a different argumentative style in terms of writer stance from editorialists.

The comparison between learner English and native English as used in editorials may produce results that may also be due to the writing inexperience of Italian learners. This is because the texts that constitute the Italian subcorpus of ICLE have been produced by university students who may be lacking in expertise in writing argumentative essays. For this reason, I also compare the language output of Italian students to that of native British and American university students. A comparison between native expert writing and the type of non-native writing collected in ICLE, which would be justified in terms of similarity of text type, would provide data for which it would not be possible to distinguish between the variables of non-nativeness and lack of expertise. The same would apply to a comparison between native and non-native novice writers. The triple comparison that I carry out in this study is therefore intended to arrive at a better identification of the influence of the variable of non-nativeness on areas of underuse, overuse and misuse of stance adverbials.

The present analysis does not intend to prescribe the use of stance adverbials in written texts. The main aim is to verify whether Italian students construct writer stance by exploiting this particular grammatical device, and whether they have a good command of the interpersonal and cohesive functions that stance adverbials can perform. The questions that this study tries to answer are the following:

1) Are there any differences in the use of epistemic stance adverbials made up of single adverbs marking certainty, doubt and actuality in the productions of native and non-native writers?
2) Is writing expertise an important variable in the use of epistemic stance adverbials?

3.3. Methodology

3.3.1. Description of the corpora

The data used for the present investigation have been taken from three different sources: the Italian subcorpus of the "International Corpus of Learner English" (ICLE), the "Louvain Corpus of Native English Essays" (LOCNESS) and a corpus of persuasive opinion articles from newspapers and magazines which is labelled "EMA"[6].

ICLE is a 3,7-million-word computer learner corpus. It collects 6,085 essays written by undergraduate students, who are

[6] "Editorial material" is the label which describes the "genre" of some newspaper files of the British National Corpus (BNC). Those files include letters-to-the editor, institutional editorials, personal editorials and some legal presentations (Lee, 2001: 54). The distinction between institutional editorials and personal editorials, which is often made in British journalism, corresponds to the distinction between editorials and columns that I made in my corpus. I chose the labels "editorials" and "columns" following the majority of the newspapers and magazines' on-line pages from where I took the material.

learners of English, in their third or fourth year of university. ICLE is the result of a project launched by Sylviane Granger at the University of Louvain in 1990. At that time, corpus linguistics had already provided researchers with new insights into the nature of language use. Corpora collecting different language varieties were available but non-native varieties were totally disregarded. The ICLE project, therefore, was an attempt to fulfil the need for non-native speakers' productions. At present the corpus is made up of 16 different non-native varieties of English. The language backgrounds are Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Turkish, and Tswana. Each national subcorpus contains c. 200,000 words and it has been compiled according to specific design criteria.

The learner profile includes information about sociolinguistic variables such as age and gender; the mother tongue; the language(s) used at home; the level of education; the language used as medium of instruction; the year of study; the proficiency level; the type of exposure to the English language (e.g. periods of time spent in an English-speaking country); and knowledge of other foreign languages[7].

The written material collected in ICLE is divided into two types of texts: argumentative essays and literary exam papers. Argumentative texts account for the largest part of the corpus. Each national subcorpus collects a maximum of 25% of literary exam papers. As regards argumentative texts, students were provided with a topic that stimulated argumentation. Suggested essay titles were, for instance:

- Crime does not pay.
- The prison system is outdated. No civilised society should punish its criminals: it should rehabilitate them.

---

[7] Information about the corpus collection guidelines and the learner profile can be obtained visiting the "Centre for English Corpus Linguistics" website (Université Catholique de Louvain): http://www.uclouvain.be/en-317607.html

- Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.
- A man/woman's financial reward should be commensurate with their contribution to the society they live in.

Students were asked to agree or disagree with a given point of view (expressed by the statement) and, consequently, to take a stance. ICLE, therefore, is suitable for research into the way non-native speakers express their opinions and point of view and interact with readers to persuade them.

The essays, both argumentative and literary texts, could be done as a home assignment that students could do in their own time but without seeking the help of native speakers or other sources, such as books; alternatively, essays could be done under exam conditions. Each text ranges between 500 and 1,000 words.

Table 3. Corpora for analysis

	ICLE-IT	LOCNESS	EMA
No. of words	227,085	175,087	208,183

The present investigation takes into consideration texts written in English by Italian university students. Henceforth, the Italian subcorpus will be referred to as ICLE-IT. ICLE-IT collects 392 essays and the total number of words is 227,085 (see Table 3 for a comparison of the size of the three corpora). The average length of the essays is of 579 words. The average age of learners is 24 and their proficiency level ranges from higher intermediate to advanced. Literary exam papers constitute a small percentage of the texts collected (15%). However, as my analysis deals with argumentation, I did not take literary exam papers into account.

The data provided by the Italian subcorpus are compared with data taken from LOCNESS, a 300,000-word corpus of essays by British and American students. This corpus comprises native English writing which is comparable to ICLE as will be explained shortly. LOCNESS has been compiled at Louvain University by the team of researchers working on the ICLE project. Their aim was to collect texts that were equivalent to ICLE texts. This need was felt for two main reasons (Granger et al. 2002a: 40-41). First, both learners and native speakers become proficient in writing only by degrees. Therefore, the results obtained through comparisons between non-native students' writing and professional or expert writing will be different from those obtained through comparisons between native and non-native students' essays. Second, no professional writing exactly corresponds to students' argumentative texts. The texts in LOCNESS are not all argumentative essays, but, as in ICLE, there are also literary exam papers. In my study these papers are not taken into consideration. LOCNESS is made up of three components: British pupils' A-level essays, British university students' essays and American university students' essays. Since I am only interested in the argumentative texts written by university students, the subcorpus that I use for my analysis is smaller than LOCNESS, and it contains 175,078 words.

The third source of data is a corpus of "editorial material" (EMA), which I compiled and computerised between May and September 2004. In order to ensure comparability with ICLE and LOCNESS, I collected 295 texts, whose length ranges from 500 to 1,000 words. The total number of words is 208,183 (Table 4).

Table 4. Sub-genres of EMA

Sub-genres	*Number of texts*	*Number of words*
Columns	119	102,358
Editorials (signed and unsigned)	162	105,825
TOTAL	281	208,183

Table 5. Geographic varieties of EMA

Geographic variety	*Number of texts*	*Number of words*
American English	158	113,333
British English	123	94,850
TOTAL	281	208,183

The articles were taken from web pages of newspapers and magazines[8]. The geographic varieties are American and British English (Table 5). Each text was saved as a plain ASCII text file so as to be processed by WordSmith Tools, i.e. the software that I use for the analysis of the three corpora. The texts of EMA are editorials (both signed and unsigned) and columns. Editorials are written by professional writers who comment on events or controversial topics in the news of the day. In general, editorials are not signed because they reflect the opinion of the entire editorial board rather than the writer's viewpoint. Some newspapers, however, have signed editorials. This kind of editorial differs from personal columns in that it expresses the opinion of both the author and the newspaper or magazine. Editorialists may also write personal columns; in that case the viewpoint expressed does not necessarily represent the stand of the editorial board. The newspaper or magazine, therefore, is not responsible for what is written. Columnists may also be freelance journalists who do not write for a living or they may be experts in a certain field and, consequently, be qualified to argue for or against a certain position. A typical editorial page contains both editorials and columns. In spite of this, however, some newspapers distinguish between an editorial page and an Op-Ed page. "Op-Ed" is a label that stands for "Opposite-Editorial", that is, the page facing the editorial page. When such a division is made, columns are placed in the Op-Ed page along with letters to the editor.

Both editorials and columns are characterised by comments and analysis rather than by objective reporting of facts. The authors express their opinion and provide information to support their evaluations. They take a stance and argue in favour of or

[8] See Appendix II for the complete list of articles.

against an idea or fact. The main purpose of editorialists and columnists is to persuade readers to adopt the same point of view or, at least, to provoke thought and discussion. Therefore, editorials and columns can be taken as a "norm" against which to compare ICLE and LOCNESS in the study of writer stance[9].

3.3.2. Analytical techniques

This section illustrates the analytical steps taken in the present study. The first task was to select the stance adverbs for analysis. This choice was based on the list of stance adverbs provided by Biber and Finegan (1988). The following step was to retrieve every instance of stance adverbs. As the adverbs for study were pre-selected, I extracted them by using the Concord Tool of WordSmith Tools. This software provides raw frequencies and sorts the selected item within its context. Each occurrence was displayed within five lines of context[10]. Contextualisation was fundamental because a major task of the study was that of disambiguation, which was done manually since the corpora are not tagged[11]. My aim was to identify adverbs functioning as adverbials. First of all, I distinguished between adverbs and adjectives. "Likely" is one of these cases. In example (31) "likely" functions as an adverbial, while in example (32) it functions as an adjective

(31) The problems will most *likely* be with the electronic voting. [...] [EMA]

(32) But nothing even close to that is *likely* to happen. [EMA].

---

[9] Appendix III includes four essay samples taken from the three corpora. One is taken from ICLE, one from LOCNESS and two from EMA (i.e. one column and one editorial). Stance adverbials are in italics.

[10] Appendix IV shows the concordance lines for the adverb "actually" in the three corpora.

[11] The present study is based on the first release of the ICLE corpus which was not tagged (Granger et al, 2002a). At present a new release is available (Granger et al., 2009) in which all learner essays are lemmatized and POS-tagged with CLAWS (Lancaster University (http://ucrel.lancs.ac.uk/claws/).

Subsequently, I distinguished between potential stance adverbs and manner adverbs. “Clearly”, for instance, may be glossed as “in a clear manner” or as “obviously”. Once I had identified the instances that might have stance meanings, I separated out adverbs functioning as adverbials (example 33) from modifiers of adjectives, adverbs and noun phrases. In example (34), “really” modifies the adjective “odd” and functions as an intensifier.

(33) […] it *really* couldn’t have gotten much worse than it had been. [LOCNESS]

(34) […] a *really* odd smelling cigarette [...] [LOCNESS]

The concordance lines that remained contained only adverbs functioning as adverbials. However, another step was necessary to complete the process of disambiguation. Some of the adverbials expressing certainty and actuality could be interpreted in two different ways when they occurred in a medial position with gradable verbs. They could be interpreted either as intensifiers of the verb, with the approximate meaning of “very much”, or as stance adverbials, with the meaning of “it is clear/obvious that” or “in reality”, “in actual fact” (example 35).

(35) Baseball *really* needed more fans who were interested in the game [...] [LOCNESS].

In most cases, however, there was ambiguity as to whether adverbials had one meaning or the other (example 36)

(36) Yet, they never *really* cared what anyone else thought. [LOCNESS].

As Biber et al. acknowledge, «even a wider context may not clarify which meaning the speaker/writer intended in such cases» (1999: 585). As a consequence, following Biber and Finegan (1988), I counted as stance adverbials those adverbials that could «be interpreted naturally as stance marker[s]» (*ibid*.:

9), including those that might also have had an emphasising function (example 36 above).

Once the process of disambiguation was completed, there followed the calculation of the two-variable chi-square test of statistical significance for each category of stance adverbials. The two-variable chi-square verifies whether the observed frequencies are sufficiently different from the expected frequencies to reject the hypothesis that the samples belong to the same population. This hypothesis is called the “null hypothesis”. In the present study the independent variable is the category of writers, and the dependent variable is the occurrence of stance adverbials. Before using any statistical test it is necessary to set the significance level of the experiment, that is, the probability of making an error in rejecting the null hypothesis. The choice of the level of significance depends on the research question and on the consequences of making an error[12]. In corpus linguistics differences may be considered statistically significant at the 0.05 level and highly significant at the 0.01 level (Virtanen, 1998: 106). Therefore the probability of error threshold in the present study was set at $p \geq 0.01$, which means that I took 1 chance in 100 that it might be wrong to generalise that the population from which the samples were taken shares the same characteristics as the samples (Oakes, 1998; Connor-Linton, 2006). The following section presents the frequencies and the results of the chi-square test.

3.4. Results and discussion

Table 6 shows that the three semantic categories of stance adverbials investigated in this study are represented in the three corpora by rather low frequencies.

[12] In the medical field, for instance, the consequences of being wrong may be extremely dangerous. Thus, researchers will probably take 1 chance in 100,000 or more that they are rejecting the null hypothesis when they should not.

Table 6. Certainty, doubt and actuality adverbials through corpora. Frequencies and normalised figures: (x/corpus size) 10,000[13]

	Doubt	*Certainty*	*Actuality*
ICLE-IT	280 (12.33)	344 (16.53)	115 (5.06)
LOCNESS	193 (11.02)	144 (8.22)	87 (4.97)
EMA	229 (11.00)	191 (8.42)	90 (4.32)

Relative to the size of the corpus, LOCNESS and EMA have approximately the same number of doubt and actuality adverbials. In both corpora doubt adverbials occur about eleven times out of 10,000, certainty adverbials occur slightly more than eight times out of 10,000, and actuality adverbials occur a little more than four times out of 10,000. ICLE-IT is characterised by a higher number of instances. There are more than twelve doubt adverbials, sixteen certainty adverbials and five actuality adverbials per 10,000 words.

Although the normalised figures show that there are some differences between ICLE-IT and the two native corpora, the chi-square test is necessary to determine whether those differences are statistically significant[14]. The results of the chi-square test demonstrate that in native and non-native productions the distribution of certainty, doubt and actuality adverbials is significantly different. Therefore, it is possible to affirm that Italian learners tend to overuse stance adverbials expressing certainty, doubt and actuality. As regards the proportion of one category to another, doubt adverbials are by far the most frequent items in both LOCNESS and EMA, followed by certainty and actuality adverbials. In ICLE-IT, the most frequent category is that of certainty adverbials, which is followed, respectively, by the categories of doubt and actuality adverbials.

---

[13] Displayed results are rounded to the second decimal. Since rounding implies an error between the approximation and the real mathematical value, the figures obtained by normalising the total number of occurrences of stance adverbials (Table 6) may be slightly different from those obtained by calculating the sum of single normalised data (Tables 5, 6 and 7).

[14] Chi-square ICLE-IT by LOCNESS = 17.84; d.f. = 2; $p < 0.01$; Chi-square LOCNESS by EMA = 1.81; d.f. = 2; not significant at 0.2 level; Chi-square ICLE-IT by EMA = 10.27; d.f. = 2; $p < 0.01$.

So far, it appears that expertise is not a significant variable in the expression of writer stance as the two native corpora do not seem to differ significantly. Since an extensive marking of "interactional evidentiality" (Biber and Finegan, 1989) is typical of conversation, it might be argued that the overuse of stance adverbials by Italian learners reflects the more informal style of their written productions (Altenberg and Trapper, 1998; Virtanen, 1998; Granger and Tayson, 1996). However, before rejecting the hypothesis that writing experience plays a role, individual items should be analysed, in particular paying attention to the functions that they perform.

3.4.1. Adverbials expressing doubt

Of the 24 items that were searched in the three corpora, only 10 were found in ICLE-IT. Some of the remaining items ("superficially", "reputedly", "purportedly", "outwardly" and "ostensibly") never occur in any corpus, either because they are rarely used as stance adverbials (e.g. "outwardly"), or because they are very formal (e.g. "purportedly"). Other adverbials occur in both LOCNESS and EMA, but they do not occur in ICLE-IT ("arguably", "ideally", "seemingly") (Table 7). However, their frequencies are extremely low; thus, larger databases would be necessary to verify whether Italian learners underuse them or not.

Table 7. Doubt adverbials. Frequencies and normalised figures: (x/corpus size) 10,000

	ICLE-IT		LOCNESS		EMA	
Allegedly	0	0	0	0	4	0.19
Apparently	5	0.22	6	0.34	30	1.44
Arguably	0	0	2	0.11	2	0.10
Conceivably	0	0	0	0	1	0.05
Doubtless	5	0.22	0	0	2	0.10
Formally	0	0	0	0	1	0.05
Ideally	0	0	1	0.06	3	0.14
Likely	2	0.09	11	0.63	7	0.34

Maybe	69	3.04	32	1.83	24	1.15
Officially	0	0	0	0	3	0.14
Ostensibly	0	0	0	0	0	0
Outwardly	0	0	0	0	0	0
Perhaps	70	3.08	59	3.37	73	3.51
Possibly	8	0.35	15	0.86	8	0.38
Presumably	1	0.04	1	0.06	12	0.58
Probably	118	5.20	54	3.08	43	2.06
Purportedly	0	0	0	0	0	0
Reportedly	0	0	1	0.06	8	0.38
Reputedly	0	0	0	0	0	0
Seemingly	0	0	1	0.06	3	0.14
Superficially	0	0	0	0	0	0
Supposedly	0	0	9	0.51	3	0.14
Technically	1	0.04	0	0	1	0.05
Theoretically	1	0.04	1	0.06	1	0.05
TOTAL	280	12.32	193	11.03	229	10.90

Only “supposedly” and “reportedly” can be considered as underused. Even in this case, however, a larger amount of corpus data is needed. Bigger corpora would help determine to what extent the underuse of those items is developmental in the learners’ pathway, as appears to be the case with “reportedly”, or whether it is due to other causes. “Allegedly”, “conceivably”, “formally” and “officially” are absent from both novice writers’ corpora, but they are extremely rare in EMA too.

The adverbials that occur in ICLE-IT are the following: “apparently”, “doubtless”, “likely”, “maybe”, “perhaps”, “possibly”, “presumably”, “probably”, “technically” and “theoretically”. “Presumably”, “technically” and “theoretically” occur only once. As regards “technically” and “theoretically”, their low frequency does not mean underuse, because these items occur only once even in the two native corpora. “Presumably”, instead, occurs once in both ICLE-IT and LOCNESS, but it is featured twelve times in EMA. The reason may be that novice writers prefer other terms to say that something is probably true. For instance, both native and non-native university students use “maybe” and “probably” more extensively than professional writers. Therefore, it seems that expertise implies greater lexical variety in the expression of probability and the ability to convey different shades of meanings.

Another adverb that is used to signal that a proposition may be true is "possibly". This adverb occurs almost as frequently in ICLE-IT as in EMA. Therefore, it would seem that Italian students do not differ from native speakers in the incidence of use of "possibly". However, by looking more closely at the way the adverb is being used by Italian learners, "possibly" turns out to be a "false friend". In the majority of the instances retrieved in ICLE-IT, "possibly" is used as if its meaning were "if possible", that is, as if "possibly" were equivalent to "possibilmente":

> (37) For example, the widespread antiabortion movement also in Italy declares that giving life must be the most natural act between two human beings, *possibly* married, so there is for any other solution. [ICLE-IT]

> (38) Of course money is not all which is required and, in my opinion, there should be strict controls on the mother and *possibly* psychological tests to understand whether she is apt or not to have a child [ICLE-IT]

It should be pointed out, however, that "possibly" occurs only eight times. Thus, more statistical evidence needs to be gathered in order to determine whether the misunderstanding of its meaning is typical of Italian learners or whether the present results are due to the idiosyncrasies of some learners.

One of the items that tends to be underused by Italian learners is "apparently", which is also underused by novice native writers. But particularly striking is the underuse of "likely", which occurs only twice in ICLE-IT in collocation with "very"[15]. The underuse of "likely", however, seems to be compensated by the above-mentioned overuse of "probably" and "maybe" which, along with "perhaps", are the most favourite

[15] In both native corpora "likely" is modified almost exclusively by "most" (e.g. "You will most likely find that it is a very similar situation" [LOCNESS]).

items of Italian university students. These three adverbials are responsible for the over-representation of the category of doubt adverbials in ICLE-IT.

"Probably", "maybe" and "perhaps" are extremely common in both written and spoken registers (Biber et al., 1999: 868). Nevertheless, in written registers "perhaps" is favoured, while "maybe" is felt to be informal (Quirk et al., 1985: 620). In conversation, on the other hand, "probably" is the most common adverb expressing doubt, which is followed, respectively, by "maybe" and "perhaps" (Biber et al., 1999: 869). In EMA and LOCNESS the frequencies of the three items reflect that register distribution, whereas in ICLE-IT the situation is different.

"Probably" is Italian learners' favourite adverbial. It is the most frequent item of all the items that were retrieved for the present investigation, including doubt and actuality adverbials. At least two reasons may be hypothesised for this result. One is that being similar to the Italian adverb "probabilmente", "probably" is more familiar to learners, and thus, it is used more confidently. The other reason is that the teaching methodology may have influenced the linguistic habits of learners. The communicative approach to language teaching places great emphasis on speech; as a consequence, learners are exposed in the main to informal conversations. Since one reason does not exclude the other, it is possible that they both account for the overuse of "probably".

"Perhaps" is not overused by Italian students. However, it occurs as frequently as "maybe". This is not the case in the two native corpora, where "maybe" is much less frequent than "perhaps". This would suggest that Italian learners are unaware of register restrictions and the result is that they use "maybe" and "perhaps" interchangeably.

The analysis of "probably", "maybe" and "perhaps" shows that Italian students tend to adopt a rather informal style. This tendency, however, is not limited to Italian learners. Agerström (2000) investigates hedging in the essays of advanced Swedish students and he finds that "probably", "maybe" and "perhaps" are among the most overused items (see also Aijmer, 2002).

However, it may be argued that the fact that foreign learners are more likely to be exposed to spoken rather than written register may not be the sole reason for this similarity. According to Granger and Rayson, the influence of the teaching methodology alone “cannot account for the learners’ more spoken style” (1998: 130). Conversational style often corresponds to a high degree of personal involvement, which seems to be an age-related feature that learners share with young and inexperienced native speakers (*ibid*.: 130). The data provided by ICLE-IT, LOCNESS and EMA corroborate this hypothesis. Although, on the whole, the two native corpora are not significantly different as regards the distribution of the three semantic categories, they differ in terms of the frequencies of single items. “Maybe” and “probably” are overused in LOCNESS as compared to EMA. Therefore, with respect to personal involvement, novice native writers are more similar to foreign learners than to professional writers.

The proportion of doubt adverbials in LOCNESS reflects the proportion of doubt adverbials in EMA. In both native corpora, “perhaps” is the most frequent item, followed respectively by “probably” and “maybe” (see Table 7). As mentioned above, this distribution is in accordance with the picture provided by the *LGSWE*. Therefore, although novice native writers tend to overuse adverbials expressing doubt, they are aware of register differences, which appears not to be the case with Italian learners.

To summarise, the frequencies of this category of stance adverbials seem to confirm the hypothesis that Italian learners produce texts that are characterised by features that are more typical of speech than writing. “Probably” is the most frequent adverbial, followed by “maybe” and “perhaps”, which have approximately the same frequency of occurrence. This suggests that Italian learners tend not to distinguish between formal and informal doubt adverbials. As regards the overuse of “probably” and “maybe”, it seems that the reason for this language behaviour is developmental, because novice native writers also overuse these adverbials as compared to professional writers. Contrary to what emerged by looking only at the frequencies of the three adverbial categories, expertise is an important factor, especially in terms of the degree of personal involvement in argumentation.

3.4.2. Adverbials expressing certainty

The picture of certainty adverbials is quite different from that of doubt adverbials. A greater number of items is absent from the three corpora ("decidedly", "unarguably", "incontestably", "incontrovertibly", "indisputably", "indubitably", "manifestly", "patently", "plainly"), whereas only one adverbial is absent exclusively from ICLE-IT: "admittedly". This adverbial, however, is not very frequent in either native corpora.

Table 8. Certainty adverbials. Frequencies and normalised figures: (x/corpus size) 10,000

	ICLE-IT		LOCNESS		EMA	
Admittedly	0	0	3	0.18	5	0.22
Assuredly	0	0	1	0.06	0	0
Certainly	79	3.79	25	1.48	31	1.37
Clearly	6	0.29	13	0.77	14	0.62
Decidedly	0	0	0	0	0	0
Definitely	6	0.29	15	0.89	2	0.09
Evidently	2	0.1	0	0	2	0.09
Incontestably	0	0	0	0	0	0
Incontrovertibly	0	0	0	0	0	0
Indeed	41	1.97	21	1.25	46	2.03
Indisputably	0	0	0	0	0	0
Indubitably	0	0	0	0	0	0
Manifestly	0	0	0	0	0	0
Obviously	43	2.07	19	0.89	8	0.35
Of course	88	4.23	38	2.25	56	2.47
Patently	0	0	0	0	0	0
Plainly	0	0	0	0	0	0
Surely	54	2.59	7	0.42	21	0.92
Unarguably	0	0	0	0	0	0
Undeniably	1	0.05	0	0	1	0.04
Undoubtedly	23	1.1	6	0.36	4	0.18

Unquestionably	1	0.05	0	0	1	0.04
TOTAL	344	16.53	144	8.55	191	8.42

Some adverbials which occur in ICLE-IT and EMA do not occur in LOCNESS ("evidently", "unquestionably", "undeniably") (see Table 8). But again, there is the limitation of extremely low frequencies of these items, which is an impediment to any plausible explanation.

An adverbial which learners use approximately to the same extent as novice native writers is "indeed". However, in the Italian sub-corpus of ICLE there are some instances of misuse. When "indeed" occurs in initial position it can be seen as a cohesive device with the function of corroborating the argument (Granger and Tyson, 1996: 20). Thus, the reader is led to expect both a logical connection to the previous sentence and some support for its truth-value, as in the following example:

> (39) The most plausible explanation is that the administration has focused so intensely on Iraq, which posed no nuclear threat, that it had little energy left for the real dangers. *Indeed*, the Harvard researchers said that if a tenth of the effort and resources devoted to Iraq in the last year was devoted to securing nuclear material wherever it might be, the job could be accomplished quickly. [EMA]

In many instances retrieved from ICLE-IT often there is no real corroboration and sometimes no obvious connection between the two sentences.

> (40) I live in Azzano San Paolo, a village of about 8,000 people near Bergamo. My house is a small detached one and it is set in the historic centre. *Indeed* it is placed in the backyard of a [...] of the XI century which was built to defend the village against the enemies but it was destroyed in a fire caused by the Guelphs in the 1409. [ICLE-IT]

> (41) I think that their role of satire can be a fitting examples to explain the different way of considering freedom of speech, depending on the different kinds of government. *Indeed* in a dictatorship an underground satire is frequently used and it is particularly cutting. [ICLE-IT]

In example (40), there is logical connection between the two sentences linked by "indeed", but the sentence framed by the adverbial does not corroborate a previous opinion. Thus "indeed" has no argumentative value. In example (41), it seems that the adverbial is used to provide an example of what has previously been said in an emphatic way. However, in this case too it does not perform an argumentative function.

An adverbial that Italian students apparently underuse is "clearly". Although it is difficult to provide a reason for this linguistic behaviour, it may be argued that Italian learners seem to favour other lexical items to express certainty such as "surely", "obviously", "of course" and "certainly". The data from ICLE-IT also suggest that "clearly" is felt by Italian learners more as a manner adverb than as a stance adverbial: it occurs eleven times as manner adverb (example 42) and only six as stance adverbial.

> (42) Firstly because the availability of guns would not miraculously disappear: this is *clearly* demonstrated by another social problem, the drug use. [ICLE-IT]

The items that Italian learners overuse are "certainly", "obviously", "of course", "surely" and "undoubtedly". "Obviously" indicates that knowledge has been obtained inductively (Chafe, 1986: 266). Induction, or inference, is a "mode of knowing in which evidence plays a central role" (*ibid*.: 266). In professional texts, the proposition framed by "obviously" is almost always preceded or followed by a justification for why something should be considered as obvious.

> (43) I have some sympathy, too, for those whose biological children have flown the nest and are understandably looking to compensate with something with an annoying bark that trails mud all over the carpet and doesn't know how to use a knife and fork. *Obviously*, pets have certain advantages over children. You don't have to take them on holiday […] though it is possible to put your kids into kennels for the summer - or, as they call it in America, 'camp'. [EMA]

In novice writing the basis for the assessment is rarely provided. The LOCNESS texts are midway between ICLE-IT and EMA. Sometimes novice native writers provide the basis for their evaluation, sometimes they simply signal that induction has taken place. This language behaviour is typical of speech. In conversation there is no time for careful argumentation, thus speakers exploit adverbials expressing certainty to convey a sense of conviction and to encourage solidarity (Biber and Finegan, 1988: 18). Although, on the whole, "obviously" is not used incorrectly by Italian learners, it tends to be overused in instances without indication of the nature of the evidence. Therefore, with respect to "obviously" Italian learners' production are characterised by a more spoken style.

(44) When new events occur, society always debates publicly, criticizes and tries to react against what is too innovative and hazardous to be accepted. The case of women who try to become mothers without having partners is *obviously*, one of the theme on which people creates more polemics. But judging on moral grounds and at the light of religious dogmas, it is easy to forget one of the oldest and most natural rights of a woman: to become a mother. [ICLE-IT]

In example (44), "obviously" is not used for the purpose of argumentation, but it is used to create common ground. However, the fact that artificial insemination is one of the issues that most stimulate discussion is not necessarily obvious. This is another characteristic of novice texts. Sometimes adverbials expressing certainty introduce propositions with the aim of establishing common ground, but those propositions cannot be taken for granted. "Of course" is often used in this way (example 45).

(45) Just think about our last-year textual analysis exam and about the great number of different interpretations which arose from the same short text, and what I mean saying that a novel or a poem is open to a wide range of interpretations, will be immediately clearer. *Of course* mental processes, which guide you to go deeper than the surface while reading a book, should be developed and the only way to do so is by trying to read in an intelligent way, since the more you read, the more

you become reading-conscious. [ICLE-IT]

"Certainly", "of course" and "surely" are used by Italian learners as native speakers do when their function is to prevent dispute. When these adverbials are employed to set up concessive clause relations, however, there are a number of differences between novice and professional writers. First of all, dissimilarities can be found in terms of the frequency with which novice and professional writers interact with readers. Novice native writers set up more concessive clause relations than professional writers, and novice non-native writers show an even higher number of instances of concession to the reader. The consequence of conceding too often points to the reader may be that the concessions lose their persuasive force[16]. Secondly, the concessive clause implies that the following sentence should be interpreted as unexpected by readers (Quirk et al., 1985: 1088). As Neff et al. point out (2003: 568), professional writers make additional claims that are careful second considerations of the opinion conceded. Novice writers, on the contrary, offer readers statements which are "so obvious in nature as to not constitute a claim at all" (*ibid*.: 569). Examples (46), (47) and (48) illustrate this situation.

(46) Puritans argue that where laws have been liberalised—in, for instance, the Netherlands, Germany and Australia—the new regimes have not lived up to claims that they would wipe out pimping and sever the links between prostitution and organised crime. *Certainly*, those links persist; but that's because, thanks to concessions to the opponents of liberalisation, the changes did not go far enough. [EMA]

(47) Using two leaders of France, it's interesting to note how Louis

[16] In their study of French-speaking learners of English, Granger and Tyson (1996: 127) underline the underuse of concessive subordinators in the French sub-corpus of ICLE as compared to LOCNESS. I have retrieved from the Italian sub-corpus of ICLE some of the subordinators listed by Granger and Tyson ("although", "while", "whilst", "though") to verify whether they are underused also by Italian learners. The data that I obtained corroborates Granger and Tyson's findings. One reason for this underuse may be that learners set up many concessive clause relations exploiting stance adverbials. This hypothesis, however, is merely speculative.

XIV- XVI and Napoleon desired to use the decorative arts to impress the globe with the glory of France. *Certainly*, these leaders hoped to communicate the opulence and prestige of their reigns. However, art also helps us communicate with each other in day-to-day living. Children's drawings are a good example of this. [LOCNESS]

(48) Probably saying that nowadays religion is replaced with television is going too far but whoever will think about the twentieth century will consider the television one of its mains features and elements. *Certainly* people don't follow the precepts of TV but it has an incisive role in our society the same. [ICLE-IT]

Although novice writers set up more concessive clause relations than professional writers, the overuse of adverbials expressing certainty by Italian learners cannot be explained exclusively in these terms. The overuse of “certainly”, “obviously”, “of course” and “surely” seems also to be due to the high number of instances in which these adverbials emphasise the truth value of the proposition.

(49) An adopted child might overcome this crisis making investigations about his parents but what could a child born from an artificial insemination do? How could his mother explain him that he is not the result of the union between two people loving each other but of that between his mother and an unknown man who gave her his semen? This child will *surely* be frightened about this news and probably he will start to think of being a sort of outcast, an artificial product who has no roots and who is very different from all of his friends. [ICLE-IT]

An extreme marking of “evidentiality” for emphasis is typical of an «involved, intense conversational style» (Biber and Finegan, 1989: 110). Therefore, it appears that even in the case of certainty adverbials Italian learners tend to adopt a more oral style. This conclusion seems to be supported by the fact that Italian students often place adverbials at the end of the clause, which is typical of a conversational style.

It should be pointed out, however, that interpersonal involvement may vary according to cultural background. There are cultures that tolerate a higher level of writer/reader visibility

(Petch-Tyson, 1998: 107). Therefore, the adoption of the terms "underuse" and "overuse" in the present study should be intended as descriptive, because the data is not adequate to exclude the hypothesis that the extensive use of certainty adverbials by Italian learners is attributable to different rhetorical strategies. In order to have a more complete picture of the situation it would be useful to carry out an analysis of stance adverbials in argumentative essays in Italian too.

3.4.3. Adverbials expressing actuality

This category of adverbials is characterised by three items: "actually", "factually", and "really". Only "actually" and "really" will be analysed because no instances of "factually" were found in any of the three corpora (Table 9).

Table 9. Actuality adverbials. Frequencies and normalised figures: (x/corpus size) 10,000

	ICLE-IT		LOCNESS		EMA	
Actually	35	1.54	32	1.83	37	1.78
Factually	0	0	0	0	0	0
Really	80	3.52	55	3.14	53	2.54
TOTAL	115	5.06	87	4.97	90	4.32

"Actually" is not overused by Italian learners. Yet differences exist between native and non-native speakers: Italian learners tend to overuse "actually" in initial position (see Appendix IV). This tendency is also shared by French and German learners (Granger and Tyson, 1996: 22). In the case of French learners, Granger and Tyson explain that the overuse of "actually" in initial position is to due to mother-tongue influence. French learners seem to translate English corroborative connectors with "en fait", which in general is placed at the beginning of the clause in French (*ibid*.: 22)[17]. As regards Italian stu-

[17] Granger and Tyson call "corroborative connectors" the stance adverbials "actually", "indeed", "in fact" and "of course". These adverbials either strengthen the argument or provide a new interpretative angle.

dents, however, there is no such mother tongue influence. Therefore, the explanation provided by Altenberg and Trapper appears more acceptable, in suggesting that «the tendency among learners to overuse corroborative adverbs is due to their argumentative 'style'» (1998: 90).

According to Oh (1999), "actually" favours different clause positions depending on the mode of discourse, i.e. speech or writing. Although "actually" favours medial position in both spoken and written language, in the written variety this tendency is stronger. Thus, written genres are characterised by high frequencies of "actually" in medial position and by relatively low frequencies in initial position. Spoken genres, instead, show higher frequencies in initial and also in final position. This would suggest that learners produce texts which are more similar to a spoken, involved argumentative style in terms of the sentence position of "actually".

When placed initially, "actually" appears to be employed as a cohesive device functioning as the contradiction of prior expectations. Thus, when "actually" is used at the beginning of the clause not only does it show actuality, it also links the proposition to a preceding sentence giving a new turn to the argument. Italian learners are not always aware that "actually" has a contrastive meaning.

> (50) Very little has been done to enlarge universities, to give them more autonomy and to encourage scientific research. *Actually* the Government has assigned to this Ministry less than one percent of its budget. [ICLE-IT]

> (51) Naturally there aren't only disadvantages in a theoretical degree. *Actually* students have often excellent cultural bases and this proves advantages when they decide to take a M.A. in another country. [ICLE-IT]

In example (50) "actually" is used to corroborate the previous statement and not to deny or contrast it. In example (51) the contrastive meaning of "actually" is minimised by "naturally" in the previous sentence, which shows a contrast itself. In addition, in this example "actually" does not contradict a previous expectation.

To sum up, the analysis of "actually" has provided data suggesting Italian learners are influenced by informal conversation in the placement of "actually" in the clause. This adverbial is neither underused nor overused by Italian students, but some instances of misuse have been found, especially when "actually" functions as both stance adverbial and connector. Italians appear to be unaware that "actually" is used to contradict expectations.

As regards "really", on the whole, there are no instances of misuse in the Italian texts. The major problem is that of overuse. Kaszubski (1996) has found that "really" is a word overused by many foreign learners of English, regardless of their mother tongue. His study involved the Chinese, Czech, Dutch, Finnish, French, German, Japanese, Spanish, Swedish and Polish subcorpora of ICLE. Apart from the Japanese and Swedish subcorpora, the overuse of "really" is highly significant in all the other subcorpora. In order to illustrate what happens in (part of) the Italian subcorpus it may be useful to give some examples of the patterns in which "really" occurs. The pattern "subject + operator + *really* + verb" occurs fourteen times in ICLE-IT, seven times in LOCNESS and nine times in EMA (example 52).

(52) They were our "paladines" and they have *really* given us down. [ICLE-IT].

The pattern "subject + *can*/*can't*/*cannot* + *really*" occurs nine times in ICLE-IT, three times in LOCNESS and once in EMA (example 53).

(53) […] they are always in a hurry also because they can *really* do much more things during a day than their grandparents. [ICLE-IT]

There are also patterns that never occur in either native corpora, such as "subject + modal verb + be + *really*", which was found seven times in ICLE-IT (example 54).

(54) However the explanation of a text may be *really* necessary if the reader is not cultivated enough to get the whole meaning [...] [ICLE-IT]

In the two native corpora one of the most common positions is immediately after the subject. In ICLE-IT, too, there are instances in which the adverbial is placed after the subject. However, many of those instances, are characterised by the pattern "*I* + *really* + verb", which occurs only once in both LOCNESS and EMA.

In the texts produced by Italian learners "really" tends to be placed in medial position between the operator, or a modal verb, and the main verb. When "really" occurs in initial position, or at the end of a clause, it indicates that the content of a proposition matches reality. When it is placed in medial position it keeps its stance meaning but it also intensifies the truth-value of the proposition expressed[18]. If learners overuse intensifiers, whether they convey a stance meaning or not, the impression is that of "overstatement", that is, of "too much communicative effort" (Lorenz, 1998: 61). In many of the instances taken from the Italian texts the proposition would have been impressive even without the adverb (e.g. "I *really* believe in the supremacy of the human person, and I think that moral principles contribute to create law […]" [ICLE-IT]). As Lorenz points out intensifiers "are not essential to the argument – on the contrary, they potentially even distract from the main points" (*ibid*.: 60).

[18] In medial position some problems arise as to whether "really" intensifies the entire clause in which it occurs (stance adverbial) or only part of it (emphasiser). The instances that could naturally be interpreted as having both functions were counted in the present study. The raw frequency of "really" was 164. After the disambiguation process only 80 instances remain. Most of them, however, are not straightforward examples of "really" as stance adverbial, but rather they are examples of ambiguous cases.

The data provided by the retrieval of adverbials expressing actuality has shown that Italian learners overuse this category of stance adverbials. This higher frequency, however, is due to the overuse of “really”. “Actually” is used almost to the same extent as native writers, though placed too frequently in initial position.

Conclusion

Now it is possible to answer the questions that I posed in Chapter 3, section 3.2.1. The first issue that I wished to investigate was whether there were differences in terms of the frequency of occurrence of the three categories of adverbials in the productions of professional writers, native and non-native university students. The present investigation has revealed that on the part of Italian learners there is a general overuse of all the categories of stance adverbials analysed: doubt, certainty and actuality. Thus, the initial hypothesis that Italian students would show a different argumentative style from professional writers may be accepted.

The argumentative style of Italian learners is characterised by an extensive use of adverbials marking certainty and doubt. Contrary to native speakers, Italian learners use more adverbials expressing certainty than adverbials expressing doubt. However, the difference between the frequencies of these two categories is smaller in ICLE-IT than in LOCNESS and EMA. In other words, adverbials expressing doubt and certainty are employed by Italian students to quite a similar extent. The significant use of these two categories may have an influence on the rhetorical effect of the argumentation (Aijmer, 2002: 73). The expression of doubt and the expression of certainty can be considered as opposite persuasive strategies, and the mixed use of «contradictory strategies [...] contributes to the impression sometimes given by a text that it has not been written by a native speaker» (*ibid*.: 73).

The analysis of single items has shown that the over-representation of adverbials expressing doubt and certainty depends on relatively few adverbials: “maybe”, “probably”, “certainly”, “obviously”, “of course”, “surely”, and “undoubtedly”. Apart from “undoubtedly”, all these adverbials tend to occur much more frequently in spoken contexts. Therefore, it seems that Italian students adopt a rather informal and speech-like style. This tendency is shared by a number of foreign learners (see Agerström, 2000; Aijmer, 2002; Kaszubski, 1996) and, in part, by inexperienced native speakers too.

With regard to adverbials expressing actuality there are no major differences in the frequencies of use of “actually”, but dissimilarities exist in the way Italian learners employ this adverbial. On the whole, Italian students seem to be unaware that “actually” implies an unexpected contrast. All too often “actually” is used to corroborate a previous statement. “Really”, instead, is significantly overused, and the impression is that learners adopt a hyperbolic style, which in many cases does not add to the force of their argumentation. The extensive use of “really” alone accounts for the high frequencies of the category of adverbials expressing actuality.

On the basis of the evidence that has emerged in this study, the first question, which was about whether there are differences in terms of frequency in the use of epistemic stance adverbials in the productions of professional writers, native and non-native university students can be answered affirmatively. Italian learners are more eager to convey their assessments of certainty, doubt and actuality than native speakers. There are no significant differences between professional and novice native writers in terms of the frequencies of the three categories.

The second question I asked was whether writing expertise is an important variable in the use of epistemic stance adverbials. The evidence suggests that expertise is an important variable. Novice native writers tend to use fewer types of doubt adverbials than professional writers, and the consequence is that they use a limited number of items more extensively, notably *probably* and *maybe*. This tendency, which reveals a conversational and involved style, makes novice native writers resemble foreign learners more than experienced writers.

Another interesting aspect which emerged from analysis is that professional writers almost always provide the basis for their evaluations of certainty. This does not happen when the proposition that writers frame as certain is a point that they concede to the reader. That point is an objection that may be raised by readers and that writers anticipate in order to neutralise criticism. In the productions of novice native and non-native writers the basis for the evaluation is rarely provided, and the sentence following the concession is not a careful second consideration of a concept or opinion, but an obvious logical consequence of what has been said before. As regards the function of certainty adverbials to create common ground, there are instances in which inexperienced writers take for granted a concept which is not necessarily shared by the majority of readers. This may result in pragmatic failure.

The findings of the present investigation have a number of pedagogical implications. As pointed out by Holmes (1988: 40), the study of epistemic modality has often concentrated on modal verbs, disregarding other syntactic and lexical devices, such as lexical verbs (e.g. "appear", "hope", "presume"), nouns (e.g. "assumption", "opinion", "belief"), adjectives (e.g. "certain", "plain", "sure") and adverbials. The consequence may be that learners make use of these alternative devices without knowing exactly how to do it and the contexts in which certain items should be used more sparingly. Hence, it would be advisable to present learners with a more realistic picture of modality, because «native speakers do not limit themselves to modal verbs» (*ibid.*: 40). This would increase the lexical knowledge of learners. Extending learners' vocabulary would perhaps have a positive effect on some situations of overuse, as in the case of "probably". In addition, along with the list of adverbials that some grammars provide (see Sinclair, 1990, *Collins COBUILD English Grammar*) there should be also indications of the relative frequency with which adverbials occur in different registers. In this way students would have at least some guidelines to help them avoid producing stylistically inadequate texts. For instance, it would be suitable to provide learners with information concerning some restrictions of use of "really" in terms of register distribution. The study of modal adverbials and adverbials expressing actuality, therefore, should be placed within a discourse perspective.

As regards the issue of personal involvement, I have pointed out that the evidence provided by the present study does not allow us to exclude the hypothesis that the extensive use of certainty adverbials by Italian learners is due to different rhetorical strategies. Even though novice native writers, too, appear to be more "visible" than professional writers, the cultural background of Italian students may have some influence over their productions. Therefore, it would be desirable for future investigations into the way Italian learners construct "writer stance" to take into account argumentative essays written in Italian too.

In conclusion, two methodological considerations are required. First, the present study has demonstrated that comparing foreign learners only to novice native writers may result in a partial picture of the nature of learners' interlanguage. Without a control corpus of professional writers developmental factors could hardly be identified. Second, the conclusions that I have drawn here would have been almost impossible to reach without adopting a corpus-based approach. The corpus-based approach has enabled me to carry out a systematic analysis of all the instances of stance adverbials in the three corpora in a relatively short time.

Appendix 1

Epistemic stance adverbs expressing certainty, doubt and actuality.

CERTAINTY	DOUBT	ACTUALITY
Admittedly	Allegedly	Actually
Assuredly	Apparently	Factually
Certainly	Arguably	Really
Clearly	Conceivably	
Decidedly	Doubtless	
Definitely	Formally	
Evidently	Ideally	
Incontestably	Likely	
Incontrovertibly	Maybe	
Indeed	Officially	
Indisputably	Ostensibly	
Indubitably	Outwardly	
Manifestly	Perhaps	
Obviously	Possibly	
Of course	Presumably	
Patently	Probably	

Plainly	Purportedly	
Surely	Reportedly	
Unarguably	Reputedly	
Undeniably	Seemingly	
Undoubtedly	Superficially	
Unquestionably	Supposedly	
	Technically	
	Theoretically	

This list is taken from Biber and Finegan (1988). Even though the adverb "probably" is not featured in the list provided by the two scholars, I included it in this study because I felt it was an important marker of doubt.

Appendix II

Editorials [Ed], Op-Ed Columns [Op-Ed] and Columns [C] from Newspapers and Magazines

US newspapers and magazines

[CST] Chicago Sun-Times: Daily newspaper. On-line edition: www.suntimes.com
[D] Dissent: Quarterly magazine of politics and culture of the left. On-line edition: www.dissentmagazine.org
[LAT] Los Angeles Times: Daily newspaper. On-line edition: www.latimes.com
[NYP] New York Post: Daily newspaper. On-line edition: www.nypost.com
[SFC] San Francisco Chronicle: Northern California daily newspaper. On-line edition: www.sfgate.com/chronicle
[HC] The Hartford Courant: Connecticut daily newspaper. On-line edition: www.courant.com
[N] The Nation: Weekly newsmagazine of the left. On-line edition: www.thenation.com
[NYT] The New York Times: Daily newspaper. On line edition: www.nytimes.com
[WP] The Washington Post: Daily newspaper. On-line edition: www.washingtonpost.com

British (and international) newspapers and magazines

[FT] Financial Times: Business and politics daily newspaper. On-line edition: www.news.ft.com
[I] Independent: Liberal daily newspaper. On-line edition: www.independent.co.uk
[DT] The Daily Telegraph: Conservative daily newspaper. On-line edition: www.telegraph.co.uk
[E] The Economist: Weekly magazine covering political and economic events. On-line edition: www.economist.com
[G] The Guardian: Daily newspaper of the left. On-line edition:www.guardisn.co.uk
[S] The Spectator: Conservative weekly magazine. On-line edition: www.spectator.co.uk

Newspaper or magazine	Title of the article	Date of publication	No. of words	Type of article
[CST]	Mammograms must be available for all women	Jun. 14, 2004	636	[Ed]
[CST]	Election rhetoric takes frightful turn	Sept. 9, 2004	660	[Ed]
[CST]	All parties must know that time is of the essence	Sept. 8, 2004	673	[Ed]
[CST]	Swabbing suspects for DNA could help prevent crime	May 30, 2004	685	[Ed]
[CST]	Preservation foresight paying off beautifully	Sept. 3, 2004	689	[Ed]
[CST]	Service sector can best fight outsourcing by facing reality	Sept. 6, 2004	690	[Ed]
[CST]	Quit playing politics with ban on assault weapons	Sept. 7, 2004	695	[Ed]
[CST]	One Zell of a blow to Kerry record that counts	Sept. 5, 2004	710	[Ed]
[CST]	Bush's chance to spell out why he's the one	Aug. 30, 2004	746	[Ed]
[D]	Response	Winter 2002	875	[C]
[D]	An appealto Dean supporters	Spring 2004	961	[C]
[FT]	No to champions	May 26, 2004	537	[Ed]
[FT]	Quality first for the EU Commission	May 27, 2004	561	[Ed]
[FT]	Co-operation is the best policy	Jun. 13, 2004	725	[C]

[FT]	Spain and Poland should stand firm on voting	May 26, 2004	733	[C]
[I]	There's bewilderment, and then there's obtuseness	Jun. 5, 2002	704	[C]
[I]	The politics of the playground is obstructing peace in Iraq	Apr. 25, 2004	722	[C]
[I]	He played good, he sang good	Jun. 13, 2004	728	[C]
[I]	Howard is going the wrong way about tackling UKIP	Jun. 13, 2004	747	[C]
[I]	The truth about torture and interrogation	May 12, 2004	757	[C]
[I]	Whispering in Bush's ear has got him nowhere. Blair must find his voice on Iraq	May 23, 2004	792	[C]
[I]	Political courage is needed to tackle our asylum myths	Jun. 5, 2004	797	[C]
[I]	It's insane - I've lost my right to vote	Jun. 5, 2004	801	[C]
[I]	What George Bush could learn from the Gipper	Jun. 7, 2004	829	[C]
[I]	The myth of Mr Kennedy's opposition to the war	Jun. 2, 2004	836	[C]
[I]	What's the point of giving 16-year-olds the vote?	May 5, 2004	878	[C]
[I]	They had the cheek to call me 'Charlie Chamberlain'	May 21, 2004	883	[C]
[I]	In the shadow of extreme nationalism	Jun. 8, 2004	886	[C]
[I]	This vote is really about being in or out of Europe	Apr. 20, 2004	887	[C]
[I]	America’s brutal culture of unseen oppression	May 25, 2004	902	[C]
[I]	Tony Blair must be more honest over Iraq	May 20, 2004	908	[C]
[I]	We need irreversible progress in tackling world poverty	Jun. 1, 2004	912	[C]
[I]	The electrodes’ switch is in Washington	May 2, 2004	964	[C]
[LAT]	Low-MPG Energy Policy	May 20, 2004	500	[Ed]
[LAT]	A continuing journey	May 17, 2004	512	[Ed]
[LAT]	Bow to global AIDS reality	May 19, 2004	519	[Ed]
[LAT]	A war there, a war here	May 31, 2004	522	[C]
[LAT]	Knock sense into taxes	Apr. 18, 2004	562	[C]
[LAT]	The schools go flabby	May 22, 2004	573	[Ed]

[LAT]	A win-win water deal	May 19, 2004	583	[C]
[LAT]	Putting a good face on the final adieu	Jun. 13, 2004	602	[C]
[LAT]	Stop winking at torture and codify It	Jun. 13, 2004	691	[C]
[LAT]	High, and at risk	Jun. 13, 2004	921	[C]
[LAT]	The danger of too much chumminess	Jun. 13, 2004	996	[C]
[NYP]	President Bush’s MemorialDay proclamation	May 31, 2004	506	[Ed]
[NYP]	Silver’s slippery slope	Jun. 13, 2004	571	[Ed]
[NYP]	UK’s Declaration of Independence	Jun. 15, 2004	571	[Op-Ed]
[NYP]	UN resolution riddle	Jun. 1, 2004	612	[Op-Ed]
[NYP]	A mile-long line of lives touched by true leader	Jun. 11, 2004	613	[Op-Ed]
[NYP]	Reagan myths	Jun. 15, 2004	650	[Op-Ed]
[NYP]	Rattling Riyadh	Jun. 15, 2004	693	[Op-Ed]
[SFC]	Grecian formula	Aug. 31, 2004	500	[Ed]
[SFC]	Roadblocks on driver's license bill	Sept. 2, 2004	515	[Ed]
[SFC]	An election about security	Sept. 3, 2004	543	[Ed]
[SFC]	Lessons of Sept. 11 -- charting a course for a vulnerable	Sept. 5, 2004	547	[Ed]
[SFC]	Crisis in the higher education: A leisurely college education	Jun. 6, 2004	652	[Ed]
[SFC]	Mourningin America	Jun. 6, 2004	714	[Ed]
[DT]	There's no word for 'standards' in Milibabble	Aug. 18, 2004	501	[Ed]
[DT]	Prisons without bars aren't prisons	Sept. 3, 2004	501	[Ed]
[DT]	Sistani is the key to peace in Najaf	Aug. 27, 2004	517	[Ed]
[DT]	PM's chance to be of lasting benefit to Britain	Sept. 8, 2004	529	[Ed]
[DT]	Reward sixth-formers by rationing A grades	Aug. 14, 2004	554	[Ed]
[DT]	Iran should be punished for nuclear cheating	Aug. 19, 2004	560	[Ed]
[DT]	Russia's children pay with their lives	Sept. 4, 2004	627	[Ed]
[DT]	The pips could hardly squeak any louder	Sept. 7, 2004	767	[Ed]
[DT]	It’s about freedom, not imperium	May 26, 2004	912	[C]
[DT]	Howard is honest but unbelievable	Jun. 16, 2004	937	[C]
[DT]	Imagine a country run by Heseltine	May 31, 2004	941	[C]

[DT]	Time for flippant, feminine modern man to stop hiding behind the masculine few	Jun. 5, 2004	942	[C]
[DT]	House prices have peaked – honestly	Jun. 14, 2004	945	[C]
[DT]	Where is the Tory alternative?	Jun. 2, 2004	963	[C]
[DT]	Freud had it right about uptown girls	Jun. 5, 2004	964	[C]
[DT]	Works of art are just as mortal as their creators	May 27, 2004	969	[C]
[DT]	The true spirit of independence	Jun. 7, 2004	978	[C]
[DT]	Edward Hopper shows us the joy of sadness	May 31, 2004	983	[C]
[DT]	Ministry of clowns threatens circus	Jun. 14, 2004	990	[C]
[DT]	What's the point of being posh when you can be American?	Jun. 16, 2004	994	[C]
[DT]	I didn't want the MMR – and now my baby has measles	Jun. 5, 2004	996	[C]
[DT]	McDonald's scores own goal with supersize spin	Jun. 14, 2004	997	[C]
[E]	Can Ariel Sharon really evacuate the Jewish settlements in Gaza?	Sept. 4, 2004	615	[Ed]
[E]	Garbage in, garbage out	May 27, 2004	659	[Ed]
[E]	The great fall of China?	May 13, 2004	828	[Ed]
[E]	Still at its mercy	May 20, 2004	900	[Ed]
[E]	Sex is their business	Sept. 2, 2004	968	[Ed]
[G]	Long division	Apr. 1, 2004	503	[Ed]
[G]	Why IT matters	May 20, 2004	508	[C]
[G]	It is as an outsider that I write	Sept. 24, 2003	548	[C]
[G]	Black hole blues	Apr. 8, 2004	606	[Ed]
[G]	Inside Europe	Apr. 26, 2004	606	[C]
[G]	A catastrophe too far	Jun. 1, 2004	610	[Ed]
[G]	Towards a British Islam	Apr. 1, 2004	612	[Ed]
[G]	Our business too	Sept. 7, 2004	616	[Ed]
[G]	Prison's revolving door	Apr. 13, 2004	624	[Ed]
[G]	Not in good shape	Feb. 26, 2004	631	[Ed]
[G]	Return verdict	Mar. 10, 2004	635	[Ed]
[G]	Unsustainable debt	Mar. 1, 2004	636	[Ed]
[G]	Curbing the courts	Jan. 5, 2004	644	[Ed]
[G]	A rose in winter	Jan. 6, 2004	645	[Ed]
[G]	Failure of will	Feb. 28, 2004	645	[Ed]
[G]	Europe's reality check	Jun. 17, 2004	652	[Ed]

[G]	On Labour's ground	Jun. 25, 2004	653	[Ed]
[G]	Peers versus people	Mar. 22, 2004	658	[Ed]
[G]	Eastern promise	Apr. 30, 2004	658	[Ed]
[G]	Kerry's chance to shine	Jul. 26, 2004	664	[Ed]
[G]	Settling nothing	May 4, 2004	672	[Ed]
[G]	Leading questions	May 22, 2004	673	[Ed]
[G]	Confidence trick	Sept. 9, 2004	674	[Ed]
[G]	Seven year verdict	May 1, 2004	700	[Ed]
[G]	Hot favourites, dark horses	Mar. 22, 2004	714	[C]
[G]	Is Gordon ready to be brutal?	Jul.13, 2004	746	[C]
[G]	Guide to age	Jun. 12, 2004	774	[C]
[G]	Decapitation	Jan. 30, 2004	779	[Ed]
[G]	Our kind of dictators	Jun. 9, 2004	783	[C]
[G]	Bush, the corporations' flag-carrier	Jan. 15, 2002	790	[C]
[G]	Towering he wasn't	Jun. 7, 2004	795	[C]
[G]	Change course or face defeat	Jun. 15, 2004	796	[C]
[G]	Adding new insult to old injury	Nov. 13, 2003	797	[C]
[G]	Death is the right price	Jul. 23, 2004	803	[C]
[G]	Nightmare in Neverland	Nov. 22, 2003	805	[C]
[G]	The Queen of chick lit	Jun. 15, 2004	812	[C]
[G]	The failing that could prove lethal	Jun. 14, 2004	818	[C]
[G]	Muslim schools don't cause riots	Jun. 10, 2004	825	[C]
[G]	We're patently going mad	Mar. 4, 2004	836	[C]
[G]	A Tale of Two Countries	Jun. 15, 2004	865	[C]
[G]	Too realistic to push	Mar. 26, 2004	899	[C]
[G]	Bush takes refuge in history	Jun. 3, 2004	955	[C]
[G]	An exquisite danger	Jun. 2, 2004	988	[C]
[G]	Terror as a weapon of occupation	Mar. 24, 2004	997	[C]
[HC]	Three reasons to hate electoral reform	Jun. 1, 2004	740	[C]
[HC]	Mistreatment of prisoners disturbing, not surprising	Jun. 1, 2004	545	[C]
[HC]	What brings young talent to the city	May 18, 2004	591	[C]
[HC]	Mr. Bush's five steps	May 26, 2004	598	[C]
[HC]	Memorial Roll grows longer	May 31, 2004	622	[C]
[HC]	The cubs of war	May 31, 2004	675	[C]
[HC]	Down the sewer to Abu Ghraib	May 31, 2004	676	[C]
[HC]	Giving the people of Hartford what they want	Jun. 1, 2004	678	[C]

[HC]	No one is clean	May 17, 2004	694	[C]
[N]	Sharon's bulldozers	May 6, 2002	512	[Ed]
[N]	The Big Money Election	May 27, 2004	512	[Ed]
[N]	Post-Feminism, R.I.P.	Apr. 29, 2004	512	[Ed]
[N]	Prick up your ears	Jan. 13, 2003	530	[Ed]
[N]	The vision thing	Jun. 16, 2003	535	[Ed]
[N]	Social security heist	Jul. 9, 2001	541	[Ed]
[N]	Scandal on scandal	May 26, 2003	543	[Ed]
[N]	War profiteering	May 12, 2003	559	[Ed]
[N]	The war is not over	Jul. 7, 2003	568	[Ed]
[N]	Bush, AIDS, Big Pharma	Apr. 8, 2004	572	[Ed]
[N]	The UN's relevance	Mar. 31, 2003	574	[Ed]
[N]	Iraq and beyond	Apr. 7, 2003	583	[Ed]
[N]	The horror of Abu Ghraib	May 6, 2004	584	[Ed]
[N]	White House dirty tricks	Oct. 20, 2004	603	[Ed]
[N]	The haunted archives	Apr. 15, 2004	612	[Ed]
[N]	Under the banner of the 'War on Terror'	Jun. 3, 2004	626	Signed [Ed]
[N]	Ulster says maybe	Dec. 20, 1999	647	[Ed]
[N]	Orders to torture	Jun. 7, 2004	649	[Ed]
[N]	Edward Said	Oct.20, 2003	652	[Ed]
[N]	Justice, not vengeance	Oct. 8, 2001	662	[Ed]
[N]	A great wound	Oct. 1, 2001	683	[Ed]
[N]	Bush's gulf of credibility	Feb. 17, 2003	706	[Ed]
[N]	Blindness in Gaza	Jun. 25, 2001	718	[Ed]
[N]	Grassroots Globalism	Feb. 18, 2002	755	[Ed]
[N]	Enron values	Feb. 25, 2002	774	[Ed]
[N]	Rumsfeld should go	Apr. 21, 2003	830	[Ed]
[N]	Take back our rights	Apr. 21, 2004	843	[Ed]
[N]	After Genoa	Aug. 6, 2001	957	[Ed]
[NYT]	Cleansing the Olympic team	Jun. 1, 2004	509	[Ed]
[NYT]	A real nuclear danger	May 28, 2004	521	[Ed]
[NYT]	China's soft landing	Jun. 11, 2004	547	[Ed]
[NYT]	The disability lobby and voting	Jun. 11, 2004	558	[Ed]
[NYT]	A hollow sovereignty for Iraq	May 29, 2004	575	[Ed]
[NYT]	Voting reform could backfire	May 9, 2004	586	[Ed]
[NYT]	When the umpires take sides	Mar. 29, 2004	610	[Ed]
[NYT]	The confusion over voter ID	Apr. 4, 2004	618	[Ed]

[NYT]	The UN go-ahead on Iraq	Jun. 10, 2004	624	[Ed]
[NYT]	Voting machines for New York	Jun. 11, 2004	628	[Ed]
[NYT]	Ronald Reagan	Jun. 7, 2004	629	[Ed]
[NYT]	Fiscal shenanigans	Jun. 3, 2004	633	[Ed]
[NYT]	The contrasts of Padre Island	Jun. 7, 2004	635	[Ed]
[NYT]	Lands in need of care	Jun. 1, 2004	649	[Ed]
[NYT]	A National ID	May 31, 2004	672	[Ed]
[NYT]	Bad new days for voting rights	Apr. 18, 2004	684	[Ed]
[NYT]	Forever the optimist	Jun. 7, 2004	685	[Ed]
[NYT]	Punishing the poor	Jun. 11, 2004	687	[Op-Ed]
[NYT]	Bitter at the top	Jun. 15, 2004	687	[Op-Ed]
[NYT]	A President who listened	Jun. 7, 2004	692	[Op-Ed]
[NYT]	Travesty of justice	Jun. 15. 2004	697	[Op-Ed]
[NYT]	Reagan’s next victory	Jun. 7, 2004	701	[Op-Ed]
[NYT]	Tear down this UN stonewall	Jun. 14, 2004	710	[Op-Ed]
[NYT]	A compromised voting system	Apr. 24, 2004	712	[Op-Ed]
[NYT]	Courting school troubles	Jun. 10, 2004	715	[Op-Ed]
[NYT]	An economic legend	Jun. 11, 2004	716	[Op-Ed]
[NYT]	Dooh Nibor Economics	Jun. 1, 2004	727	[Op-Ed]
[NYT]	Dare we call it genocide	Jun. 16, 2004	727	[Op-Ed]
[NYT]	Beating specialist baker	Jun. 5, 2004	729	[Op-Ed]
[NYT]	Reckonings; another useful crisis	Nov. 11, 2001	734	Signed [Ed]
[NYT]	Go negative on the allies	Jun. 15, 2004	750	[Op-Ed]
[NYT]	Level with Americans	Jun. 7, 2004	753	[Op-Ed]
[NYT]	Florida as the next Florida	Mar. 14, 2004	760	[Op-Ed]
[NYT]	Reading Thackeray's 'Vanity Fair' With the Illustrations Intact	Aug. 30, 2004	794	Signed [Ed]
[NYT]	George Tenet resigns	Jun. 4, 2004	802	[Ed]
[NYT]	Making votes count	Jun. 13, 2004	802	[Ed]
[NYT]	Nation builders and low bidders in Iraq	Jun. 15, 2004	814	[Op-Ed]
[NYT]	Today, Philadelphia's brotherly love belongs to Smarty Jones	Jun. 5, 2004	822	[Op-Ed]
[NYT]	D-Day in Iraq	Jun. 10, 2004	824	[Op-Ed]
[NYT]	What adolescents miss when we let them grow up in cyberspace	May 29, 2004	848	Signed [Ed]
[NYT]	When politics corrupts	Jun. 16, 2004	860	[Op-Ed]

[NYT]	Peak Performance in Athletics: Is That All We Can Expect?	Aug. 29, 2004	894	Signed [Ed]
[NYT]	Why are we hiding Bin Laden?	Nov. 11, 2001	1000	Signed [Ed]
[O]	Europe's help for Russia is now vital	Sept. 5, 2004	620	[Ed]
[O]	Still hoping Japan is land of rising sums	Jul. 18, 2004	667	[C]
[O]	Groan up	Apr. 4, 2004	712	[C]
[O]	I want my digital utopia	Apr. 11, 2004	715	[C]
[O]	Puppy love	Feb. 29, 2004	717	[C]
[O]	Make mine metrosexual	May 16, 2004	732	[C]
[O]	At least we had each other	May 23, 2004	732	[C]
[O]	Asexual harassment	Jul. 4, 2004	748	[C]
[O]	The morning after	Mar. 28, 2004	751	[C]
[O]	There are no such thing as nags	Jun. 20, 2004	775	[C]
[O]	Say it like you mean it	Jul. 11, 2004	788	[C]
[O]	Status complacency	Jun. 13, 2004	807	[C]
[O]	Near, far, wherever you are...	Jul. 18, 2004	815	[C]
[O]	The ugly truth	Jul. 25, 2004	841	[C]
[O]	Topped up with debt	Jan. 18, 2004	983	[C]
[S]	The flunking examiners	Jun. 19, 2004	829	[Ed]
[S]	It's about democracy	Apr. 24, 2004	836	[Ed]
[S]	A loss of respect	Apr. 17, 2004	853	[Ed]
[S]	Fat controllers	Jun. 26, 2004	857	[Ed]
[S]	The abuse of power	Aug. 28, 2004	861	[Ed]
[S]	Bring back the Sixties	Jul. 24, 2004	869	[Ed]
[S]	Gordon's great con	Mar. 6, 2004	869	[Ed]
[S]	Closed minds	Feb. 28, 2004	870	[Ed]
[S]	Listen hard and you can hear J. Bonington-Jagworth grumbling loudly.	Aug. 7, 2004	871	[Ed]
[S]	Rogue mail	May 1, 2004	873	[Ed]
[S]	Lock them up	Mar. 13, 2004	876	[Ed]
[S]	Victory for optimism	Jun. 12, 2004	878	[Ed]
[S]	First gold to Greece	Aug. 14, 2004	895	[Ed]
[S]	Democracy can wait	Apr. 10, 2004	899	[Ed]
[S]	Help the aged	Sept. 5, 2004	929	[Ed]
[WP]	Letters from prison	Apr. 15, 2004	511	[Ed]
[WP]	Starving science	May 29, 2004	518	[Ed]

[WP]	Double standards	May 14, 2004	521	[Ed]
[WP]	Fighting the wrong battles	Apr. 29, 2004	521	[Ed]
[WP]	No way to treat schools	May 29, 2004	525	[Ed]
[WP]	A start on democracy	Apr. 28, 2004	527	[Ed]
[WP]	Oil wars	Jul. 20, 2003	527	[Ed]
[WP]	Dealing drugs	Mar. 15, 2004	527	[Ed]
[WP]	Virginia’s reward	May 31, 2004	536	[Ed]
[WP]	The upset in India	May 14, 2004	538	[Ed]
[WP]	Arresting witnesses	May 22, 2004	538	[Ed]
[WP]	Abuse by outsourcing	May 26, 2004	545	[Ed]
[WP]	Talking to the World	Oct. 6, 2003	548	[Ed]
[WP]	Trade and Mr. Kerry	May 27, 2004	549	[Ed]
[WP]	Recall in Venezuela	Jun. 13, 2004	565	[Ed]
[WP]	Google vs. Wall Street	May 9, 2004	584	[Ed]
[WP]	The Arab Backlash	Mar. 10, 2004	592	[Ed]
[WP]	The SEC's challenge	Jun. 1, 2004	595	[Ed]
[WP]	Too slow on nukes	Jun. 11, 2004	605	[Ed]
[WP]	Finding a war in Iraq	May 9, 2004	611	[Ed]
[WP]	Squirrelly budgeting	Apr. 25, 2004	618	[Ed]
[WP]	Mr. Bush’s challenge	May 25, 2004	622	[Ed]
[WP]	No winks or nods	Jun. 14, 2004	623	[Ed]
[WP]	Legalizing torture	Jun. 9, 2004	633	[Ed]
[WP]	An inadequate response	May 8, 2004	632	[Ed]
[WP]	The security dilemma	May 19, 2004	648	[Ed]
[WP]	The war decision	Apr. 25, 2004	662	[Ed]
[WP]	Protecting the system	May 12, 2004	665	[Ed]
[WP]	Healing health care	May 15, 2004	670	[Ed]
[WP]	A corrupted culture	May 20, 2004	675	[Ed]
[WP]	Reveal the rules	May 23, 2004	679	[Ed]
[WP]	Mr. Kerry on security	May 30, 2004	681	[Ed]
[WP]	Rules of the System	May 11, 2004	682	[Ed]
[WP]	Iraq and the Conservative crack-up	Jun. 1, 2004	741	[C]
[WP]	Fix this bill	Jun. 13, 2004	745	[Ed]
[WP]	Winning the cold war	Jun. 14, 2004	533	[Ed]
[WP]	We can never take our guard down	Jun. 14, 2004	645	[C]
[WP]	September 11 and the economy	Sept. 8, 2004	758	[Ed]
[WP]	Valentines from a President	Jun. 14, 2004	761	[C]
[WP]	The ethics of antiterrorism	Jun. 14, 2004	772	[C]

[WP]	It's called victory	Jun. 14, 2004	820	[C]
[WP]	A liberal legacy	Jun. 14, 2004	851	[C]
[WP]	False memories and friends	Jun. 14, 2004	882	[C]
[WP]	Gov. Romney's trump cord	Jun. 14, 2004	894	[C]

Appendix III

ICLE

<ICLE-IT-TOR-0001.5> (No. of words: 597)

Probably the person who wrote this song had so much money that he did not know how to use it and, for this reason, money was perhaps more a source of anxiety than a fortune for him/her. Anyway, nobody never died of money and, especially nowadays, having at least a small amount of money helps people to survive. More and more here in Italy you hear people talking only about money and how to make it; the State does not know what else to invent in order to make people try their luck through different games: “scrape and win”, other national lotteries, horse-racing or football pools.

Another reason for trying to make as much money as possible is that the cost of living has gone and is still going up more and more. For a middle-class family (but far more for a lower-class one!) it is not enough just a single salary to live anymore. That is why people dream a win, for instance in the lottery, in order to improve their life conditions and, if possible, to change their life completely. Let us suppose that one morning, awakening, a person finds out he/she has won a lot of money; *probably* the first feeling would be of astonishment and fear, but *surely* then the first thing that this person will do, would be going to collect his/her money. Where is the evil in it, then? Is there anything wrong in money itself? Not at all. One may object that having money sometimes leads to anxiety, fear of being robbed or swindled and to moral corruption, in the sense that the more you have, the more you want, no matter who you trample on. This is partially true if you do not know how to make profit from your capital; it could also be objected that in that case too you could run the risk of being swindled by carrying out a wrong transaction. This is true as well, but "nothing venture, nothing have", as the saying goes. [...]

Apropos of songs, in the text of a song written by Alanis Morrisette, a Canadian singer, "An old man turned 98, he won the lottery and died the next day". *Surely* money does not make happiness nor health or immortality and fate in most cases decides for you, but in any case having money let you life better without thinking always of saving, of choosing cheap (and often of poor quality) goods. Having money in a certain sense helps you also with health, because if you have a serious disease you can pay for those drugs, hospital stays and treatments you eventually need. Moreover having money would be for poor people *really* a godsend. It is not necessary to remind in what conditions third-world people live. How could you say that for them "money would be the root of all evil"? Even if we do not look so far, we perceive that also in our country there are people living on the border-line of dignity; for all these people we ought to remember the importance and the civil progress brought by money circulation in order not to waste money but to appreciate it as a divine gift and to learn to be satisfied with what we have, without profitting by other people's misfortune and misery.

Governments should give much more money to the medial research and should provide a better chance of life, organizing all the public services that we need as best as they can. *Perhaps*, we would not need to use solutions against nature.

LOCNESS

<ICLE-BR-SUR-0017.3> (No. of words: 659)

Although Europe is a much discussed topic, with talks about subsidies and a single European currency, the concept of a united Europe is still difficult for the average person to visualise. *Perhaps* because they are worried and uncertain as to their role and Britain's in this new state, they prefer to believe that it is simply not possible or even probable.

Steps have already been taken, however, towards this goal. The 1972 European Communities Act and the Treaty of Rome have as their aims the eventual unification of the nation states. Hence the fact that there is no clause in the Treaty detailing how a member may leave. The treaty bound Britain to certain measures which have already curtailed our sovereignty. For example, the European Commission (an unelected body) can issue resolutions which do not even need to be incorporated into the domestic law of each member state – they are directly applicable. How has this affected our sovereignty? Until 1986 each member had the power of veto so that no new law could be passed without full agreement. In a way this did not affect our sovereignty. The Single European Act of 1986, however, introduced majority voting. It is now possible, therefore, that the UK opposes proposed legislation but that is voted through by the other states and becomes law here without our consent. This is *obviously* an infringement of our sovereignty already.

A unified Europe, though, could be expected to have a more representative legislative body, although there is no guarantee that we will not be in a minority then. *Perhaps* Europe will be organised on a federal system as a "United States of Europe", each state with limited legislative authority for its own affairs and one supreme house to co-ordinate foreign policy and the like. We have already begun to integrate our trade, and we have free movement of labour within the community.

The question of currency is a more vexed one for most people, who see the prospect of saying goodbye to sterling much harder to accept than the European Commission imposing laws on us. Although this feeling is perfectly natural, it will not long outlast the introduction of a new currency. Those of us familiar with sterling will feel a loss, but the next generation will be oblivious to it, just as I have no nostalgia for the sixpence or crown.

Certain other aspects of Britain would be more difficult to change though. Our legal system in this country is a common law one, while most countries on the continent have civil law systems. This brings about a different attitude to statutes and written law in Britain as well as a completely different approach to its treatment in court. Our law is not codified, and much has not yet even made the transition from being a common law wrong to a statutory offence. This confusing state of affairs horrifies our European partners, but it would be a lengthy process to change it, and one that would face much opposition.

There is also the question of where the Queen would fit in. She is our nominal??? (unreadable) head of state (although she is largely prevented from using her powers), but it is doubtful whether or not Europeans would be willing to accept her as their monarch as well - the Spanish already have a monarch, as do the Dutch. While she would *apparently* have no place in a single European state, would Britain be prepared to sacrifice her in order to join?

Many constitutional problems still block our road to Europe, as well as people's attitudes - we in Britain rather enjoy being an island and not attached to the continent - witness the opposition to the Channel Tunnel. It is impossible to say whether we will eventually be willing to sacrifice the rest of our sovereignty. *Maybe* if we move towards Europe slowly we will one day become accustomed to the idea.

EMA

Excerpt from: "The New York Times", *Travesty of justice*, June 15, 2004; Column (No. of words of complete article: 697)

For this column, let's just focus on Mr. Ashcroft's role in the fight against terror. Before 9/11 he was aggressively uninterested in the terrorist threat. He didn't even mention counterterrorism in a May 2001 memo outlining strategic priorities for the Justice Department. When the 9/11 commission asked him why, he responded by blaming the Clinton administration, with a personal attack on one of the commission members thrown in for good measure.

We can't tell directly whether Mr. Ashcroft's post-9/11 policies are protecting the United States from terrorist attacks. But a number of pieces of evidence suggest otherwise.

First, there's the absence of any major successful prosecutions. The one set of convictions that seemed fairly significant - that of the "Detroit 3" - appears to be collapsing over accusations of prosecutorial misconduct. (The lead prosecutor has filed a whistle-blower suit against Mr. Ashcroft, accusing him of botching the case. The Justice Department, in turn, has opened investigations against the prosecutor. Payback? I report; you decide.)

Then there is the lack of any major captures. Somewhere, the anthrax terrorist is laughing. But the Justice Department, you'll be happy to know, is trying to determine whether it can file bioterrorism charges against a Buffalo art professor whose work includes harmless bacteria in petri dishes.

Perhaps most telling is the way Mr. Ashcroft responds to criticism of his performance. His first move is always to withhold the evidence. Then he tries to change the subject by making a dramatic announcement of a terrorist threat.

For an example of how Mr. Ashcroft shuts down public examination, consider the case of Sibel Edmonds, a former F.B.I. translator who says that the agency's language division is riddled with incompetence and corruption, and that the bureau missed critical terrorist warnings. In 2002 she gave closed-door Congressional testimony; Senator Charles Grassley described her as "very credible . . . because people within the F.B.I. have corroborated a lot of her story." But the Justice Department has invoked the rarely used "state secrets privilege" to prevent Ms. Edmonds from providing evidence. And last month the department retroactively classified two-year-old testimony by F.B.I. officials, which was *presumably* what Mr. Grassley referred to […].

Excerpt from: “The Economist”, *Garbage in, garbage out*, May 27, 2004; Editorial (No. of words of complete article: 659)

Commonly used, but flawed, measures of economic output can lead to bad decision-making HOW big is the world economy? That sounds like a straightforward question. Simply to add up the size of all the world's national economies would seem to be the obvious way to answer it. But how that is done yields radically different results, and therein lies a tale. The most commonly used method is to convert national economic outputs to a single measure, namely the American dollar, using the market exchange rates of all the national currencies. That produces a figure of $36 trillion for 2003. But many professional economists think that it makes much more sense to use what they call purchasing-power parities (PPP), which take account of differences in prices of the same goods between countries, and so tries to measure the real purchasing power of inhabitants in each country, no matter what the world's fluctuating currency markets happen to be doing to exchange rates. Using this method, the world economy last year was worth $50 trillion.

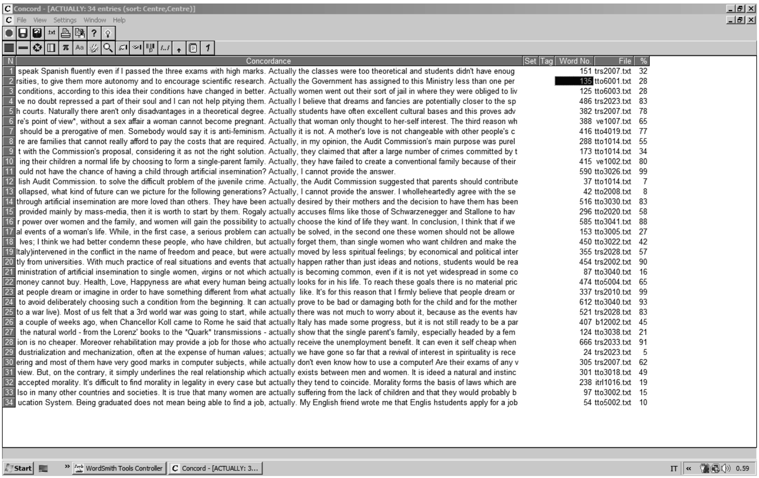
The precise size of the world economy may not matter much from a policy point of view, though a $14 trillion difference is hardly small change. However, which method of measurement is used also affects more important matters: the global rate of growth, the relative size of economies, and the extent of inequality between rich and poor. In these cases, using market exchange rates can produce misleading results and hence stimulate bad policies.

The best reason for not using market exchange rates is that prices tend to be lower in poorer countries, so a dollar of spending there is worth more. Market rates therefore understate their real level of development. *Indeed*, measured at market rates, developing economies' share of global income has fallen over the past two decades, to less than one quarter. This would back the claims of the anti-globalisation lobby that poor countries are being left behind. Yet measured at PPP, developing economies' share of world income has risen over the same period, to almost half of the total, which gives a more realistic impression.

Hot Air. PPPs are not always appropriate. Trade and capital flows which, unlike the bulk of GDP, are *actually* transacted at market exchange rates, should be converted at those rates into dollars. And businesses which trade internationally and have to convert revenues and profits to dollars are often more interested in the dollar purchasing power of various national markets than their real level of economic prosperity, so for businessmen the market-exchange-rate measure matters more […].

Appendix IV

Figure 1. Screenshot showing the concordance for "actually" in ICLE-IT.



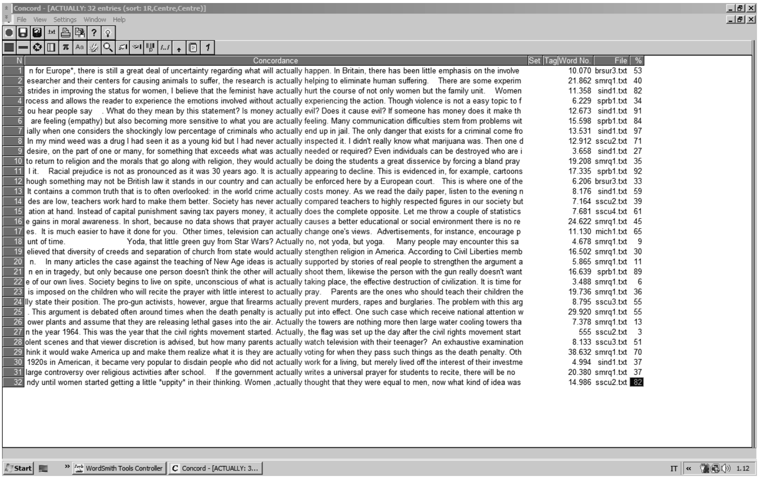
Concord - [ACTUALLY: 34 entries (sort: Centre,Centre)]

File View Settings Window Help

N	Concordance	Set	Tag	Word No.	File	%
1	speak Spanish fluently even if I passed the three exams with high marks. Actually the classes were too theoretical and students didn't have enoug			151	trs2007.txt	32
2	rsities, to give them more autonomy and to encourage scientific research. Actually the Government has assigned to this Ministry less than one per			135	tto6001.txt	28
3	conditions, according to this idea their conditions have changed in better. Actually women went out their sort of jail in where they were obliged to liv			125	tto6003.txt	28
4	ve no doubt repressed a part of their soul and I can not help pitying them. Actually I believe that dreams and fancies are potentially closer to the sp			486	trs2023.txt	83
5	h courts. Naturally there aren't only disadvantages in a theoretical degree. Actually students have often excellent cultural bases and this proves adv			382	trs2007.txt	78
6	re's point of view*, without a sex affair a woman cannot become pregnant. Actually that woman only thought to her-self interest. The third reason wh			388	ve1007.txt	65
7	should be a prerogative of men. Somebody would say it is anti-feminism. Actually it is not. A mother's love is not changeable with other people's c			416	tto4019.txt	77
8	re are families that cannot really afford to pay the costs that are required. Actually, in my opinion, the Audit Commission's main purpose was purel			288	tto1014.txt	55
9	t with the Commission's proposal, considering it as not the right solution. Actually, they claimed that after a large number of crimes committed by t			173	tto1014.txt	34
10	ing their children a normal life by choosing to form a single-parent family. Actually, they have failed to create a conventional family because of their			415	ve1002.txt	80
11	ould not have the chance of having a child through artificial insemination? Actually, I cannot provide the answer.			590	tto3026.txt	99
12	lish Audit Commission. to solve the difficult problem of the juvenile crime. Actually, the Audit Commission suggested that parents should contribute			37	tto1014.txt	7
13	ollapsed, what kind of future can we picture for the following generations? Actually, I cannot provide the answer. I wholleheartedly agree with the se			42	tto2008.txt	8
14	through artificial insemination are more loved than others. They have been actually desired by their mothers and the decision to have them has been			516	tto3030.txt	83
15	provided mainly by mass-media, then it is worth to start by them. Rogaly actually accuses films like those of Schwarzenegger and Stallone to hav			296	tto2020.txt	58
16	r power over women and the family, and women will gain the possibility to actually choose the kind of life they want. In conclusion, I think that if we			585	tto3041.txt	88
17	al events of a woman's life. While, in the first case, a serious problem can actually be solved, in the second one these women should not be allowe			153	tto3005.txt	27
18	lves; I think we had better condemn these people, who have children, but actually forget them, than single women who want children and make the			450	tto3022.txt	42
19	Italy)intervened in the conflict in the name of freedom and peace, but were actually moved by less spiritual feelings; by economical and political inter			355	trs2028.txt	57
20	tly from universities. With much practice of real situations and events that actually happen rather than just ideas and notions, students would be rea			454	trs2002.txt	90
21	ministration of artificial insemination to single women, virgins or not which actually is becoming common, even if it is not yet widespread in some co			87	tto3040.txt	16
22	money cannot buy. Health, Love, Happyness are what every human being actually looks for in his life. To reach these goals there is no material pric			474	tto5004.txt	65
23	at people dream or imagine in order to have something different from what actually like. It's for this reason that I firmly believe that people dream or			337	trs2010.txt	99
24	to avoid deliberately choosing such a condition from the beginning. It can actually prove to be bad or damaging both for the child and for the mother			612	tto3040.txt	93
25	to a war live). Most of us felt that a 3rd world war was going to start, while actually there was not much to worry about it, because as the events hav			521	trs2028.txt	83
26	a couple of weeks ago, when Chancellor Koll came to Rome he said that actually Italy has made some progress, but it is not still ready to be a par			407	b12002.txt	45
27	the natural world - from the Lorenz' books to the *Quark* transmissions - actually show that the single parent's family, especially headed by a fem			124	tto3038.txt	21
28	ion is no cheaper. Moreover rehabilitation may provide a job for those who actually receive the unemployment benefit. It can even it self cheap when			666	trs2033.txt	91
29	dustrialization and mechanization, often at the expense of human values; actually we have gone so far that a revival of interest in spirituality is rece			24	trs2023.txt	5
30	ering and most of them have very good marks in computer subjects, while actually don't even know how to use a computer! Are their exams of any v			305	trs2007.txt	62
31	view. But, on the contrary, it simply underlines the real relationship which actually exists between men and women. It is ideed a natural and instinc			301	tto3018.txt	49
32	accepted morality. It's difficult to find morality in legality in every case but actually they tend to coincide. Morality forms the basis of laws which are			238	itrl1016.txt	19
33	lso in many other countries and societies. It is true that many women are actually suffering from the lack of children and that they would probably b			97	tto3002.txt	15
34	ucation System. Being graduated does not mean being able to find a job, actually. My English friend wrote me that Englis hstudents apply for a job			54	tto5002.txt	10

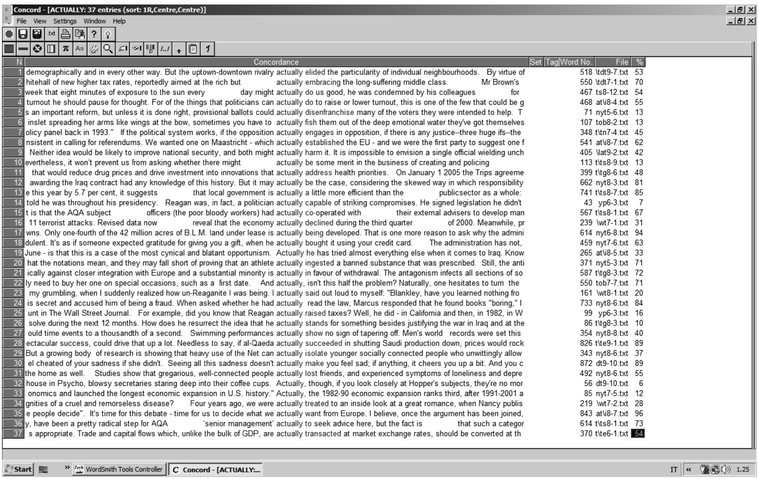
Start | WordSmith Tools Controller | Concord - [ACTUALLY: 3... | IT | 0.59

Figure 2. Screenshot showing the concordance for “actually” in LOCNESS



N	Concordance	Set	Tag	Word No.	File	%
1	n for Europe*, there is still a great deal of uncertainty regarding what will actually happen. In Britain, there has been little emphasis on the involve			10.070	brsur3.txt	53
2	esearcher and their centers for causing animals to suffer, the research is actually helping to eliminate human suffering. There are some experim			21.862	smrq1.txt	40
3	strides in improving the status for women, I believe that the feminist have actually hurt the course of not only women but the family unit. Women			11.358	sind1.txt	82
4	rocess and allows the reader to experience the emotions involved without actually experiencing the action. Though violence is not a easy topic to f			6.229	sprb1.txt	34
5	ou hear people say . What do they mean by this statement? Is money actually evil? Does it cause evil? If someone has money does it make th			12.673	sind1.txt	91
6	are feeling (empathy) but also becoming more sensitive to what you are actually feeling. Many communication difficulties stem from problems wit			15.598	sprb1.txt	84
7	ially when one considers the shockingly low percentage of criminals who actually end up in jail. The only danger that exists for a criminal come fro			13.531	sind1.txt	97
8	In my mind weed was a drug I had seen it as a young kid but I had never actually inspected it. I didn't really know what marijuana was. Then one d			12.912	sscu2.txt	71
9	desire, on the part of one or many, for something that exceeds what was actually needed or required? Even individuals can be destroyed who are i			3.658	sind1.txt	27
10	to return to religion and the morals that go along with religion, they would actually be doing the students a great disservice by forcing a bland pray			19.208	smrq1.txt	35
11	l it. Racial prejudice is not as pronounced as it was 30 years ago. It is actually appearing to decline. This is evidenced in, for example, cartoons			17.335	sprb1.txt	92
12	hough something may not be British law it stands in our country and can actually be enforced here by a European court. This is where one of the			6.206	brsur3.txt	33
13	It contains a common truth that is to often overlooked: in the world crime actually costs money. As we read the daily paper, listen to the evening n			8.176	sind1.txt	59
14	des are low, teachers work hard to make them better. Society has never actually compared teachers to highly respected figures in our society but			7.164	sscu2.txt	39
15	ation at hand. Instead of capital punishment saving tax payers money, it actually does the complete opposite. Let me throw a couple of statistics			7.681	sscu4.txt	61
16	e gains in moral awareness. In short, because no data shows that prayer actually causes a better educational or social environment there is no re			24.622	smrq1.txt	45
17	es. It is much easier to have it done for you. Other times, television can actually change one's views. Advertisements, for instance, encourage p			11.130	mich1.txt	65
18	unt of time. Yoda, that little green guy from Star Wars? Actually no, not yoda, but yoga. Many people may encounter this sa			4.678	smrq1.txt	9
19	elieved that diversity of creeds and separation of church from state would actually stengthen religion in America. According to Civil Liberties memb			16.502	smrq1.txt	30
20	n. In many articles the case against the teaching of New Age ideas is actually supported by stories of real people to strengthen the argument a			5.865	smrq1.txt	11
21	n en in tragedy, but only because one person doesn't think the other will actually shoot them, likewise the person with the gun really doesn't want			16.639	sprb1.txt	89
22	e of our own lives. Society begins to live on spite, unconscious of what is actually taking place, the effective destruction of civilization. It is time for			3.488	smrq1.txt	6
23	is imposed on the children who will recite the prayer with little interest to actually pray. Parents are the ones who should teach their children the			19.736	smrq1.txt	36
24	lly state their position. The pro-gun activists, however, argue that firearms actually prevent murders, rapes and burglaries. The problem with this arg			8.795	sscu3.txt	55
25	. This argument is debated often around times when the death penalty is actually put into effect. One such case which receive national attention w			29.920	smrq1.txt	55
26	ower plants and assume that they are releasing lethal gases into the air. Actually the towers are nothing more then large water cooling towers tha			7.378	smrq1.txt	13
27	n the year 1964. This was the year that the civil rights movement started. Actually, the flag was set up the day after the civil rights movement start			555	sscu2.txt	3
28	olent scenes and that viewer discretion is advised, but how many parents actually watch television with their teenager? An exhaustive examination			8.133	sscu3.txt	51
29	hink it would wake America up and make them realize what it is they are actually voting for when they pass such things as the death penalty. Oth			38.632	smrq1.txt	70
30	1920s in American, it became very popular to disdain people who did not actually work for a living, but merely lived off the interest of their investme			4.994	sind1.txt	37
31	large controversy over religious activities after school. If the government actually writes a universal prayer for students to recite, there will be no			20.380	smrq1.txt	37
32	ndy until women started getting a little *uppity* in their thinking. Women ,actually thought that they were equal to men, now what kind of idea was			14.986	sscu2.txt	82

Figure 3. Screenshot showing the concordance for “actually” in EMA



N	Concordance	Set	Tag	Word No.	File	%
1	demographically and in every other way. But the uptown-downtown rivalry actually elided the particularity of individual neighbourhoods. By virtue of			518	\tdt9-7.txt	53
2	hitehall of new higher tax rates, reportedly aimed at the rich but actually embracing the long-suffering middle class. Mr Brown's			550	\tdt7-1.txt	70
3	week that eight minutes of exposure to the sun every day might actually do us good, he was condemned by his colleagues for			467	ts8-12.txt	54
4	turnout he should pause for thought. For of the things that politicians can actually do to raise or lower turnout, this is one of the few that could be g			468	at\i8-4.txt	55
5	s an important reform, but unless it is done right, provisional ballots could actually disenfranchise many of the voters they were intended to help. T			71	nyt5-6.txt	13
6	inslet spreading her arms like wings at the bow, sometimes you have to actually fish them out of the deep emotional water they've got themselves			107	tob8-2.txt	13
7	olicy panel back in 1993." If the political system works, if the opposition actually engages in opposition, if there is any justice--three huge ifs--the			348	t\tn7-4.txt	45
8	nsistent in calling for referendums. We wanted one on Maastricht - which actually established the EU - and we were the first party to suggest one f			541	at\i8-7.txt	62
9	Neither idea would be likely to improve national security, and both might actually harm it. It is impossible to envision a single official wielding unch			405	\lat9-2.txt	42
10	evertheless, it won't prevent us from asking whether there might actually be some merit in the business of creating and policing			113	t\ts8-9.txt	13
11	that would reduce drug prices and drive investment into innovations that actually address health priorities. On January 1 2005 the Trips agreeme			399	t\tg8-6.txt	48
12	awarding the Iraq contract had any knowledge of this history. But it may actually be the case, considering the skewed way in which responsibility			662	nyt8-3.txt	81
13	e this year by 5.7 per cent, it suggests that local government is actually a little more efficient than the publicsector as a whole:			741	t\ts8-7.txt	85
14	told he was throughout his presidency. Reagan was, in fact, a politician actually capable of striking compromises. He signed legislation he didn't			43	yp6-3.txt	7
15	t is that the AQA subject officers (the poor bloody workers) had actually co-operated with their external advisers to develop man			567	t\ts8-1.txt	67
16	11 terrorist attacks. Revised data now reveal that the economy actually declined during the third quarter of 2000. Meanwhile, pr			239	\wt7-1.txt	31
17	wns. Only one-fourth of the 42 million acres of B.L.M. land under lease is actually being developed. That is one more reason to ask why the admini			614	nyt6-8.txt	94
18	dulent. It's as if someone expected gratitude for giving you a gift, when he actually bought it using your credit card. The administration has not,			459	nyt7-6.txt	63
19	June - is that this is a case of the most cynical and blatant opportunism. Actually he has tried almost everything else when it comes to Iraq. Know			265	at\i8-5.txt	33
20	hat the notations mean, and they may fall short of proving that an athlete actually ingested a banned substance that was prescribed. Still, the anti			371	nyt5-3.txt	71
21	ically against closer integration with Europe and a substantial minority is actually in favour of withdrawal. The antagonism infects all sections of so			587	t\tg8-3.txt	72
22	ly need to buy her one on special occasions, such as a first date. And actually, isn't this half the problem? Naturally, one hesitates to turn the			550	tob7-7.txt	71
23	my grumbling, when I suddenly realized how un-Reaganite I was being. I actually said out loud to myself: "Blankley, have you learned nothing fro			161	\wt8-1.txt	20
24	is secret and accused him of being a fraud. When asked whether he had actually read the law, Marcus responded that he found books "boring," I			733	nyt8-6.txt	84
25	unt in The Wall Street Journal. For example, did you know that Reagan actually raised taxes? Well, he did - in California and then, in 1982, in W			99	yp6-3.txt	16
26	solve during the next 12 months. How does he resurrect the idea that he actually stands for something besides justifying the war in Iraq and at the			86	t\tg8-3.txt	10
27	ould time events to a thousandth of a second. Swimming performances actually show no sign of tapering off. Men's world records were set this			354	nyt8-8.txt	40
28	ectacular success, could drive that up a lot. Needless to say, if al-Qaeda actually succeeded in shutting Saudi production down, prices would rock			826	t\te9-1.txt	89
29	But a growing body of research is showing that heavy use of the Net can actually isolate younger socially connected people who unwittingly allow			343	nyt8-6.txt	37
30	el cheated of your sadness if she didn't. Seeing all this sadness doesn't actually make you feel sad; if anything, it cheers you up a bit. And you c			872	dt9-10.txt	89
31	the home as well. Studies show that gregarious, well-connected people actually lost friends, and experienced symptoms of loneliness and depre			492	nyt8-6.txt	55
32	house in Psycho, blowsy secretaries staring deep into their coffee cups. Actually, though, if you look closely at Hopper's subjects, they're no mor			56	dt9-10.txt	6
33	onomics and launched the longest economic expansion in U.S. history." Actually, the 1982-90 economic expansion ranks third, after 1991-2001 a			85	nyt7-5.txt	12
34	gnities of a cruel and remorseless disease? Four years ago, we were actually treated to an inside look at a great romance, when Nancy publis			219	\wt7-2.txt	28
35	e people decide". It's time for this debate - time for us to decide what we actually want from Europe. I believe, once the argument has been joined,			843	at\i8-7.txt	96
36	y, have been a pretty radical step for AQA 'senior management' actually to seek advice here, but the fact is that such a categor			614	t\ts8-1.txt	73
37	s appropriate. Trade and capital flows which, unlike the bulk of GDP, are actually transacted at market exchange rates, should be converted at th			370	t\te6-1.txt	54

References

Agerström, J. (2000). Hedges in argumentative writing: A comparison of native and non-native speakers of English. In T. Virtanen and J. Agerström (eds.). *Three studies of learner discourse. Evidence from the international corpus of learner English* (6-42). Rapporter Från Växjö Universitet, Humaniora 10, Växjö Universitet.

Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, et al. (2002b), (55-76).

Altenberg, B. and Trapper M. (1998). The use of adverbial connectors in advanced Swedish learner's written English. In S. Granger (ed.), (1998), (80-93).

Aston, G. (1995). Corpora in language pedagogy: Matching theory and practice. In G. Cook and B. Seidlhofer (eds.). *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (257-270). Oxford: Oxford University Press.

Atkins, B.T.S., Clear J. and Ostler N. (1992). Corpus Design Criteria. *Literary and Linguistic Computing*, VII, 1-16.

Baker, P. (2006). *Using corpora in discourse analysis*. London: Continuum.

Barbera, M. (2009). *Schema e storia del «corpus taurinense». Linguistica dei corpora dell'italiano antico*. Alessandria: Edizioni dell'Orso.

Barbera, M., Onesti C. and Corino E. (2007). Cosa è un corpus? Per una definizione più rigorosa di corpus, token, markup. In M. Barbera, E. Corino e C. Onesti (eds.). *Corpora e linguistica in rete* (89-108). Perugia: Guerra Edizioni.

Bernardini, S. (2000). *Competence, capacity, corpora*. Bologna: CLUEB.

Besnier, N. (1993). Reported speech and affect on Nukulaelae Atoll. In J.H. Hill and J.T. Irvine (eds.). *Responsibility and evidence in oral discourse* (161-181). Cambidge: Cambridge University Press.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: CUP.

Biber, D. and Finegan E. (1988). Adverbial stance types in English. *Discourse Processes*, *11*, 1-34.

Biber, D. and Finegan E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, *9*, 93-124.

Biber, D., Conrad S. And Reppen R. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics*, *15*, 169-189.

Biber, D., Conrad S. and Reppen R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: CUP.

Biber, D., Johanasson S., Leech G., Conrad S. and Finegan E. (eds.). (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.

Bondi, M. and Mauranen A. (eds.). (2003). Evaluation in Academic Discourse. Special issue of *Journal of English for Academic Purposes*, *2*(4).

Bondi, M., Gavioli L. and Silver M. (2004). *Academic discourse genre and small corpora*. Roma: Officina Edizioni.

Brown, J.D. (2009). Foreign and Second Language Needs Analysis. In M.H. Long and C.J. Doughty (eds.). *The handbook of language teaching* (269-293). Singapore: Blackwell.

Brown P. and Levinson S.C. (1978). *Politeness. Some universals in language usage*. Cambridge: Cambridge University Press.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Carter, R. (1987). *Vocabulary: Applied linguistic perspectives*. London: Allen & Unwin.

Chafe, W. (1986). Evidentiality in English conversation and academic writing. In W. Chafe and J. Nichols (eds.). *Evidentiality: The linguistic coding of epistemology* (261-272). Norwood, NJ: Ablex.

Channell, J. (2000). Corpus-based analysis of evaluative lexis. In S. Hunston and G. Thompson (eds.) (2000), (38-55).

Charles, M. (2003). ‘This mystery. . .’: A corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, *2*, 313-326.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Connor-Linton, J. (2004). Chi-Square Tutorial. Available at http:// georgetown.edu/faculty/ball/webtools/web_chi_tut.html.

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

Crismore, A. (1989). *Talking with readers: Metadiscourse as rhetorical act*. New York: Peter Lang.

Dagneaux, E., Dennes S. and Granger S. (1998). Computer-aided error analysis. *System. An International Journal of educational Technology and Applied Linguistics*, *26* (2), 163-174.

Douthwaite, J. (2007). A stylistic view of modality. In G. Garzone and R. Salvi (eds.). *Linguistica, linguaggi specialistici, didattica delle lingue* (107-156). Rome: CISU.

Dressen, D. (2003). Geologists' implicit persuasive strategies and the construction of evaluative evidence. *Journal of English for Academic Purposes*, *2*, 273-290.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: OUP.

Firth, J.R. (1957). *Papers in Linguistics 1934-51*. Oxford: Oxford University Press.

Flowerdew, L. (1998). Integrating 'expert' and 'interlanguage' computer corpora findings on causality: Discoveries for teachers and students. *English for Specific Purposes*, *17*(4), 329–345.

Francis, G. (1986). *Anaphoric nouns*. Discourse Analysis Monographs, 11, University of Birmingham: English Language Research.

Francis W. and Kucera H. (1979). *Brown corpus manual*. Available at http://khnt.hit.uib.no/icame/manuals/brown/INDEX.HTM

Georgakopoulou, A. and Gustos D. (1997). *Discourse analysis: An introduction*. Edinburgh: Edinburgh University Press.

Granger, S. (ed.). (1998). *Learner English on computer*. London: Longman.

Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, Hung J. and Petch-Tyson S. (eds.) (2002b), (3-36).

Granger, S. and Rayson P. (1998). Automatic profiling of learner texts. In S. Granger (ed.) (1998), (119-131).

Granger, S. and Tyson S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, *15*(1), 17-27.

Granger, S. and Tribble C. (1998). Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven learning. In S. Granger (ed.) (1998), (199-209).

Granger, S., Dagneaux E. and Meunier F. (eds.). (2002a). *International corpus of learner English*. UCL: Presses Universitaires de Louvain (CD-ROM and Handbook).

Granger, S., Hung J. and Petch-Tyson S. (eds.). (2002b). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam: Benjamin.

Granger S., Lerot J. and Petch-Tyson S. (eds.). (2003). *Corpus-based approaches to contrastive linguistics and translation studies*. Amsterdam – New York: Rodopi.

Granger, S., Dagneaux E., Meunier F. and Paquot M. (2009). *International Corpus of Learner English v2*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Greenbaum, S. (1996). *The Oxford English Grammar*. Oxford: OUP.

Gries, S.T. and Stefanowitsch A. (eds.). (2006). *Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis*. Berlin – New York: Mouton de Gruyter.

Halliday, M.A.K. (1994). *An introduction to functional grammar*. London: Arnold, 2nd ed.

Hoey, M. (1979). *Signalling in Discourse*. Discourse Analysis Monographs No. 6, University of Birmingham: English Language Research.

Hoey, M. (1983). *On the Surface of Discourse*. London: Allen and Unwin.

Hoey, M. (2000). Persuasive rhetoric in linguistics: A stylistic study of some features of the language of Noam Chomsky. In S. Hunston and G. Thompson (eds.) (2000), (28-37).

Hoey, M. (2005). *Lexical Priming*. London and New York: Routledge.

Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied Linguistics*, *9*(1), 21-44.

Honigam, E.J.A. (ed.). (1997). *Othello*. Croatia: The Arden Shakespeare.

Hunston, S. (1994). Evaluation and organization in a sample of written academic discourse. In M. Coulthard (ed.). *Advances in written text analysis* (119-218). London: Routledge.

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

Hunston, S. (2001). Colligation, lexis, pattern, and text. In M. Scott and G. Thompson (eds.). *Patterns of text: In honour of Michael Hoey* (13-33). Amsterdam: Benjamins.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hunston, S. and Sinclair J.M. (2000). A Local Grammar of Evaluation. In S. Hunston and G. Thompson (eds.) (2000), (74-101).

Hunston, S. and G. Thompson. (eds.). (2000). *Evaluation in text. Authorial stance and the construction of discourse*. Oxford: OUP.

Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam, John Benjamins.

Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London: Longman.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, *20*, 207-226.

Hyland, K. (2005a). *Metadiscourse*. London: Continuum.

Hyland, K. (2005b). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, *7*(2), 173-192.

Johns, T. and Kings P. (eds.). (1991). Classroom concordancing. Special issue of *ELR Journal, 4*. University of Birmingham: Centre for English Language Studies.

Jordan, M.P. (1984). *Rhetoric of everyday English texts*. London: Allen and Unwin.

Kaszubski, P. (1996). A case for corpus-based EFL pedagogy. Available at http://main.amu.edu.pl/~przemka/przemek.html

Kenndy, D.G. (1998). *An introduction to corpus linguistics*. Harlow: Longman.

Kytö, M. (1996). *Manual to the diachronic part of the Helsinki corpus of English texts: Coding conventions and lists of source texts* (3rd ed.). Available at http://khnt.hit.uib.no/icame/manuals/HC/INDEX.HTM

Labov, W. (1972). *Language in the Inner City*. Philadelphia, University of Pennsylvania Press.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press: Ann Arbor.

Lee, D.Y.W. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology*, *5* (3), 37-72.

Leech, G. (1974). *Semantics*. Harmondsworth: Penguin Books.

Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (ed.). *Directions in corpus linguistics* (105-122). Berlin: Mouton de Gruyter.

Leech, G. (1998). Learner corpora: What they are and what can be done with them. In S. Granger (ed.) (1998), (XIV-XX).

Lorenz, G. (1998). Overstatement in advanced learners' writing: Stylistic aspects of adjective intensification. In S. Granger (ed.) (1998), (53- 66).

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis and E. Tognini-Bonelli (eds.). *Text and Technology* (157-176). Amsterdam: Benjamins.

Mark, K. L. (1998). The significance of learner corpus data in relation to the problems of language teaching. *Bulletin of General Education*, *312*, 77-90.

Martin, J.R. (2000). Beyond exchange: Appraisal system in English. In S. Hunston and G. Thompson (eds.) (2000), (142-175).

Martin, J.R. and White P.R.R. (2005). *The language of evaluation. Appraisal in English*. Houndmills/New York: Palgrave Macmillan.

McEnery, T. and Gabrielatos C. (2006). English corpus linguistics. In B. Aarts and A. McMahon (eds.). *The handbook of English linguistics* (33-71). Oxford: Blackwell.

McEnery, T. and Wilson A. (1996). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Meunier, F. (1998). Computer tools for the analysis of learner corpora. In S. Granger (ed.) (1998), (19-37).

Meyer, C.F. (2002). *English corpus linguistics. An introduction*. Cambridge: CUP.

Molino, A. (2010). Personal and impersonal authorial references: A contrastive study of English and Italian linguistics research articles. *Journal of English for Academic Purposes*, *9*(2), 86-101.

Mur Dueñas, P., Lorés Sanz R. and Lafuente Millán E. (eds.). (2010). Interpersonality in written academic discourse. Special issue of *Journal of English for Academic Purposes*, *9*(2).

Murison-Bowie, S. (1996). Linguistic corpora and language teaching. *Annual Review of Applied Linguistics*, *16*, 182-199.

Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, *10*(1), 1-35.

Neff, J.A., Ballesteros F., Dafouz E., Martinez F., Rica J.P. (2003). Formulating writer stance: A contrastive study of EFL learner corpora. In D. Archer, P. Rayson, A. Wilson and T. McEnery (eds.). *Proceedings of the corpus linguistics 2003 conference*. UCREL Technical Papers 16, Lancaster University: UCREL.

Oakes, M.P. (1998). *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.

Oh, S.Y. (1999). ‘Actually’ and ‘in fact’ in American English: A data-based analysis. *English Language and Linguistics*, *4*(2), 243-268.

Oostdijk, N. (1991). *Corpus linguistics and the automatic analysis of English*. Amsterdam – Atlanta: Rodopi.

Oxford, R.L. (ed.). (1996). *Language learning motivation: Pathways to the new century*. Mānoa: Second Language Teaching & Curriculum Center, University of Hawaiì at Mānoa.

Oxford, R.L. and Ehrman M.E. (1993). Second language research on individual differences. *Annual Review of Applied Linguistics*, *13*, 188-205.

Palmer, F.R. (1979). *Modality and the English modals*. London: Longman.

Petch-Tyson, S. (1998). Reader/writer visibility in EFL persuasive writing. In S. Granger (ed.) (1998), (107-118).

Pullum, G.K. and Sholz B.C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, *19*(9), 9-50.

Quirk, R. (1960). Towards a description of English usage. *Transactions of the Philological Society*, 40-61.

Quirk, R., Greenbaum S., Leech G. and Svartvik J. (1985). *A Comprehensive grammar of the English language*. London: Longman (2nd ed).

Ringbom, H. (1999). Contrastive and error analysis. In B. Spolsky (ed.). *Concise encyclopaedia of educational linguistics* (489-495). Amsterdam and Tokio: Elsevier.

Salager-Meyer, F. M., Ariza A.A. and Zambrano N. (2003). The scimitar, the dagger and the glove: Intercultural differences in the rhetoric of criticism in Spanish, French and English medical discourse (1930–1995). *English for Specific Purposes*, *22*(3), 223- 247.

Sampson, G. and McCarthy D. (eds.). (2004). *Corpus linguistics: Readings in a widening discipline*. London, New York: Continuum.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10*(3), 209-231.

Semino, E. and Short M. (2004). *Corpus stylistics: Speech, writing and thought presentation in a corpus of English writing*. London: Routledge.

Sinclair, J.M. (Editor-in-Chief). (1990). *Collins COBUILD English Grammar*. London: Collins.

Sinclair, J.M. (1991). *Corpus, concordance and collocation*. Oxford: OUP.

Sinclair, J.M. (1996). Preliminary Recommendations on Corpus Typology. EAGLES Document EAG—TCWG—CTYP/P, Version of May, 1996, available at http://www.ilc.cnr.it/EAGLES/corpustyp/corpustyp.html

Sinclair, J.M. (2004). Corpus and text: Basic principles. In M. Wynne (ed.). *Developing linguistic corpora: A guide to good practice* (1-16). Oxford: Oxbow Books. Available at http://ota.ahds.ac.uk/documents/creating/dlc/chapter1.htm.

Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell.

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.

Swales, J.M. (1990). *Genre analysis. English in academic and research settings*. Cambridge: CUP.

Tan, M. (ed.). (2002). *Corpus studies in language education*. Bangkok: IELE Press.

Tarone, E. (1999). Interlanguage. In B. Spolsky (ed.). *Concise encyclopaedia of educational linguistics* (507-512). Amsterdam and Tokio: Elsevier.

Thompson, G. (2001). Interaction in academic writing: learning to argue with the reader. *Applied Linguistics*, *22*(1), 58-78.

Thompson, G. and Hunston S. (2000). Evaluation: An introduction. In S. Hunston and G. Thompson (eds.) (2000), (1-27).

Thompson, G. and Ye Y. (1991). Evaluation in the reporting verbs used in academic papers. *Applied Linguistics*, *12*(4), pp. 365-382.

Tognini-Bonelli, E. (1993). Interpretative nodes in discourse. Actual and actually. In M. Baker, G. Francis and E. Tognini-Bonelli (eds.). *Text and Technology* (193-212). Amsterdam: Benjamins.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: Benjamins.

Van Leeuwen, T. and Wodak R. (1999). Legitimizing immigration control: A discourse-historical analysis. *Discourse Studies*, *1*(1), 83-118.

Virtanen, T. (1998). Direct questions in argumentative student writing. In S. Granger (ed.) (1998), (94-106).

White, P. (2003). Beyond Modality and Hedging: A Dialogic View of the Language of Intersubjective Stance. *Text* *23*(2), 2594–8.

Wichmann, A., Fligelstone S., McEnery A. and Knowles G. (eds.). (1997). *Teaching and language corpora*. Harlow: Longman.