Cross-Disciplinary Approaches to the English Language

Cross-Disciplinary Approaches to the English Language: Theory and Practice

Edited by

Teodora Popescu, Rodica Pioariu and Crina Herțeg

CAMBRIDGE SCHOLARS

P U B L I S H I N G

Cross-Disciplinary Approaches to the English Language:
Theory and Practice,
Edited by Teodora Popescu, Rodica Pioariu and Crina Herțeg

This book first published 2011

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

TABLE OF CONTENTS

Chapter III: Language and Culture in the Global Village

PREFACE

Cross-disciplinary Approaches to the English Language: Theory and Practice provides an overview of a less tackled field of research, namely the main issues at stake when teaching English Language and Culture in Romania. The approach is an interdisciplinary and cross-cultural one, as the authors investigate problems, offer and probe solutions from a cross-curricular perspective. Interdisciplinarity has become a virtue of both modern research and applied disciplines, opening up multitudes of perspectives. The book is a collection of 10 contributions by teachers and researchers from Romania that draw on theoretical and applied methodological explorations into the challenges posed by teaching/learning English in a globalised context. Organised into three main chapters, the volume addresses the multifacetedness of language education as a cross-discipline.

Chapter I, *Cross-disciplinary Issues in the English Language* encompasses wide dimensions ranging from a corpus-based approach to translation studies and development of business students' mental lexicon, to an inventory of business English borrowings in the Romanian language, as well as the translatability of language humour. The first paper in the series, *On the Interdisciplinary Nature of Corpus-Based Translation Studies* discusses the evolution of corpus-based translation studies. The author asserts that the Romanian research community is taking baby-steps in this respect trying to acquire an overall image of the international endeavours. That is why, the introduction of this field of study in the Romanian research environment is crucial in order to add up new theories to the existing ones and to improve Romanian translators' performance, thus enhancing intercultural communication. The paper aims at presenting the status of online available corpora, along with considerations as regards the text genres, types, reliability, consistency, user-friendliness, authorship and purposes. The conclusion is a plea for the participation of Romanian linguistic research teams in such undertakings, as well as for the presence of the Romanian language as a language pair in contrastive corpus-based investigations. The second contribution, Teodora Popescu's *Developing Business Students' Mental Lexicon in English* tackles the teachability of the business lexis, and provides an insight into the characteristics of the business language as opposed to general English. The relevance of the

subchapter lies in the hands-on approach and methodological guidelines
for business English teachers. The main resources that lend themselves to
pedagogic exploitation are: concordancers, lexical databases, paper-based
and online dictionaries, thesauri, the business press, and/or other business
related documents (e.g. meetings transcripts). For each of these tools, the
author gives examples of practical activities. A student-centred approach is
advocated and teachers are called upon to empower students to take charge
of their own learning process. Crina Herțeg`s *Transgressing English
Language Boundaries. The Case of Business English Borrowings* deals
with the current stage of Romanian language which witnesses an ever
increasing process of enrichment. The author explains that this phenomenon
is in close connection with the needs of the community, with the evolution
of the society, with the social, economic, political, technologic and
historical factors influencing it. The author draws on the fact that,
irrespective of the field of activity, English words have become a prevalent
reality in Romanian language. The primary focus of the paper is on the
way of integrating and assimilating these borrowings, which means either
translating them into Romanian, with using an equivalent in the Romanian
language, or adjusting them according to the organic rules of the TL.
While scrutinizing specialized websites and online newspapers the author
provides us with an inventory of the newly introduced words in the field of
business and its components (banking, financing, marketing, etc.). The last
contribution, by Gina Măciucă, titled *Language Humour Interidiomatically
Viewed* explores the theory advanced by Coșeriu (1994) on linguistic
norms, and applies it to humour-generating devices. The author starts from
the premise that comic effects could indeed be traced back to the flouting
of each of the three norm types identified by Coșeriu: of congruence,
correctness and appropriateness. With respect to the translatability of such
comic effects, the scale put forward by Gina Măciucă shows that the three
types of humour rank quite differently from the corresponding linguistic
norms in Coșeriu's hierarchy (1994), namely: the congruence-flouting type
ranks highest, while the correctness-flouting kind is relegated to the lowest
position, with the appropriateness-flouting humour hovering somewhere in
between. In conclusion, the translator is compelled to reshape the
situational and linguistic context out of all recognition in order to
successfully put the message across to his readers.

Chapter II, *Teaching English as a Discipline at the Crossroads* addresses
the issue of language education to students of other disciplines, as viewed
from the perspective of two important genres, English for Specific
Purposes (ESP) and Content and Language Integrated Learning (CLIL).
The specific fields under scrutiny are teaching entrepreneurship to

humanities students and teaching English to students of Science and Technology. The paper *Employing Cooperative Strategies in Learning Professional English at University Level* by Alexandra Jacobsen sets forth a course in professional English, understood both as a teaching environment and a set of teaching materials having as main criterion its usefulness for the students including its relevance to their future careers. Taking into account the students' specific specialism (Applied Electronics), this course has a double focus. On the one hand, it attempts to employ learning methods and techniques already familiar to the students (e.g. project and team work). On the other hand, it attempts to address topics specific to their area of study. The point of departure for this exploration was the notion of 'soft' skills (Pulko and Parikh 2003), understood as general skills, such as making presentations and collaborative work. To achieve the course's aims, Alexandra Jacobsen resorts to theoretical insights from Cooperative Learning and CLIL methodology, respectively. Moise Achim and Laria Dragolea's paper, *Teaching Entrepreneurship to Humanities students*, sets out to present an overview of content-based education carried out through the medium of the English language. The authors give an account of their experience of teaching entrepreneurship to 31 MA students in Language and Communication for Business Administration, all with a non-business background. Students were randomly divided into an experimental and a control group. The pedagogic experiment revealed the double focus and efficiency of a course in entrepreneurship – the acquisition of business content (theoretical knowledge pertaining to setting up a business) and the acquisition of specialist business vocabulary in English as well as of specific pragmatic strategies. The research methods used were both quantitative (tests) and qualitative (questionnaires, interviews, etc.). The teaching methods were a combination of teacher mini-lectures, project work, pair-work, group work, simulations and role-plays. The results obtained were clearly indicative of the fact that Content and Language Integrated Learning (CLIL) education is efficient and motivating. Marinela Grănescu and Ema Adam's paper, *A Teachers and Students' Perspective on Their Engineering Under- and Post-Graduate English Syllabus in a Transylvanian Technical University* addresses the issue of ESP syllabus design. The authors start from the premise that designing a course syllabus represents a complex, time-consuming and difficult process, especially when the undergraduates and graduates aimed at display various levels of language command. Grănescu and Adam present an overview of possible syllabi types, as well as their suitability to various categories of students. They posit that the basis of product syllabuses remains fundamentally similar, while the

underlying assumptions about language and language learning from analytic approaches differ greatly: process type syllabuses assert that learning a language is transient and cannot be itemized; pedagogical procedure overweighing content. The authors consider that interdisciplinarity and transdisciplinarity bring a valuable contribution to both research-oriented and practice-based endeavours, especially in the case of ESP programmes, both under- and post-graduate.

Chapter III, *Language and Culture in the Global Village* reveals the fact that in the process of language learning and teaching the indispensable presence of culture is taken for granted by all educational factors. In other words there is a cultural component in language teaching. Teaching English in a genuinely intercultural context will prove helpful in real life situations when it is so important to act appropriately and adequately. Emilia Plăcintar, the author of the contribution *A Cultural Briefing on Romania: Insider and Outsider Perspectives* aspires to help students in business deal with intercultural communication during a cultural awareness programme. The researcher has in view a number of stages to go through with her students starting from familiarisation with key concepts in intercultural communication, organisation of cultural knowledge and adaptation of personal business cultural style to the culture of the new business community. A few illustrative activities for each stage and a set of teaching techniques and methods are included with a view to equipping the participants in the training course with practical solutions meant to avoid potential cross-cultural misunderstandings. This calls for important changes namely, a shift from a national to an international mindset allowing trainees to relate to other cultures more appropriately, understand how the system works in their clients' business environment and add these culture-based considerations to their task-based skills in order to proceed in the appropriate way and get the desired results. The second contribution, Rodica Pioariu`s *Cross-cultural Issues in Teaching English to Romanian Students* reiterates the importance of culture in language acquisition trying to provide an image of the evolution of the Romanian perspective and attitude to the study and teaching of English in Romania in a diachronic approach. The most significant moments in the development of English studies in Romania are honestly revealed without neglecting the inevitable ups and downs due to the competition with other widely-circulated languages more popular with certain intellectual circles at different historical epochs, be it the interwar period, the communist or the present-day one. It equally focuses on the obvious change of paradigm and teaching practices in Romanian schools and universities in the globalisation era and their beneficial effect on the younger generations

whose adherence to contemporary values and demands are undeniable. In the last study, *Culture and Collocations*, the authors Gabriela Mocan and Mariana Toma point to the reciprocal relationship between language and culture which can no longer be ignored since both culture and language combinations or collocations play an essential role in foreign language acquisition. They are more than likely to improve intercultural communication and avoid plenty of possible misunderstandings. The idea that there is great diversity in the way different language communities make use of phraseology is also stressed. Being in favour of "different collocational patterning" reflecting not only a particular system of knowledge and thinking, but also another cultural model foreign language students sometimes have difficulties in mastering idiomatic language and collocations. The practical section illustrates culturally-oriented approach for teaching and learning collocations with a view to helping students change their world view, their attitude to it and last, but not least, promote genuine cross-cultural communication.

To conclude with, the complexity and universality of the research enquiries and practical insights make the topics addressed valid across the globalising educational context nowadays. *Cross-disciplinary Approaches to the English Language: Theory and Practice* will be a useful tool to specialists and practitioners from ESP and CLIL domains alike, as well as graduate and postgraduate students in foreign language teaching.

ABBREVIATIONS

BNC (British National Corpus)

BOKR (The Russian Reference Corpus)

BULTREEBANK (Bulgarian Treebank)

CALP (Cognitive Academic Language Proficiency)

CEXI (English-Italian corpus)

CLIL (Content and Language Integrated Learning)

CIC (Cambridge International Corpus)

CED (Chambers English Dictionary)

CEFR (Common European Framework of Reference for Languages)

CILT (National Centre for Languages in the UK)

COCA (Corpus of Contemporary American English)

CTS (Corpus-based Translation Studies)

DEX (Dicționarul explicativ al limbii române)

DTS (Descriptive Translation Studies)

EAP (English for Academic Purposes)

ECC (English Comparable Corpus)

ECOLE (Expectations, Communication, Organisation, Leadership and Etiquette)

EFL (English as a Foreign Language)

ESL (English as a Second Language)

ENCP (English Norwegian Parallel Corpus)

ESP (English for Specific Purposes)

EST (English for Science and Technology)

EUROCLIC (European Network of Administrators, Researchers and Practitioners)

GEM (Global Entrepreneurship Monitor)

INCA (Intercultural Competences Assessment project)

L1 (mother tongue)

L2 (foreign/second language)

LDELC (Longman Dictionary of English Language and Culture)

LSP (Language for Specific Purposes)

LMR (Linear-active, Multi-active and Reactive)

NON-TEC (Non-Translational English Corpus)

NOS (National Occupational Standards)

PGDF (Pons Großwörterbuch Deutsch als Fremdsprache)

RADAR (Recognise, Analyse, Decide, Act and Review)

SL (Source Language)

TL (Target Language)

TS (Translation Studies)

TEC (Translational English Corpus)

UC (Uppsala Corpus)

UKCES (UK Commission for Employment and Skills)

CHAPTER I:

CROSS-DISCIPLINARY ISSUES IN THE ENGLISH LANGUAGE

INTRODUCTION

CRINA HERȚEG

The chapter *Cross-disciplinary Issues in the English Language* brings together various contributions to language research tackled from an interdisciplinary perspective. This type of approach means crossing of the boundaries of disciplinary research in language studies and opening new horizons in interpreting language phenomena.

The ever increasing number of research centres on interdisciplinary issues witnessed by the world research community as well as by the Romanian one, the conferences organized on this topic, adopting an interdisciplinary tenet within international research projects come to account for the introduction of these contributions in the chapter under discussion.

Interdisciplinarity represents a type of approach, under different embodiments: interaction (when two or more fields interact and aim at a common objective), circulation (when a discipline uses, borrows or assimilates the concepts of another discipline), convergence (when a new discipline emerges as a result of cooperation of scientists belonging to different disciplinary fields), divergence (when differing points of view address a certain issue), integration.

The interdisciplinary approach has advantages, as well as disadvantages. One of the first advantages refers to its openness, which at its turn entails overcoming the bridges which establish between different areas of research. A second implication of its open character could be the fact that it helps us integrate new knowledge, better interpret and understand language phenomena, innovate and bring novelty in language studies. It also enables the researcher to transgress the boundaries of language-related disciplines, as he is not confined any longer to pursue research within the limits of a discipline.

Another advantage would be the fact that it promotes collaborative research, which highly recommends this type of approach in research projects.

We believe that it would be honest to mention here some of the drawbacks entailed by the interdisciplinary perspective. Despite calling forth openness and innovation, bringing together researchers and theories,

this type of approach has limitations too. One of them refers to the fact that there are areas which totally ban collaborative research, yet, we believe that disciplines under the umbrella of language studies are not encompassed in these areas. Another difficulty lies in devising a set of rules to be put into practice when discussing a language issue. A distinct terminology, and a set of rules to be taken into account by researchers have not been devised so far. To this respect, the propagation of interdisciplinary-type of research should have in view the fact that the researcher is expected to acquire and eventually master the methods, tools and terminology at work in different cognate disciplines. Thus, the interdisciplinary approach is made possible only after specialization in different areas of study.

Traditional research promotes the use of certain values as well as devices necessary for measuring the final results. As far as interdisciplinarity is concerned the results of the research are difficult to quantify, which again could be considered a drawback of interdisciplinary research.

As this approach requires solid and specialised knowledge in the respective fields of investigations, we believe that the introduction of an interdisciplinary component should be a gradual one, starting with the academic curricula and continuing to an upper stage, that of research projects.

All the investigations under discussion in this chapter have in view several areas of language studies: translation studies (Gina Măciucă's *Language Humour Interidiomatically Viewed* and Mona Arhire's *On The Interdisciplinary Nature Of Corpus-Based Translation Studies*) and specialized language, the case of business English (Teodora Popescu's *Developing Business Students` Mental Lexicon in English* and Crina Herțeg's *Transgressing English Language Boundaries. The Case of Business English Borrowings*).

We could conclude by stating that interdisciplinarity facilitates a meeting point between the above-mentioned language related areas, yet another reason to consider that the topics gathered within this chapter embrace an interdisciplinary perspective. Further arguments calling forth the interdisciplinary character of this chapter are embodiments of interdisciplinarity such as convergence: Mona Arhire's *On The Interdisciplinary Nature of Corpus-Based Translation Studies*, divergence: approaching language humour from the perspective of language norms theory and the one of translation studies, interaction: penetration of specialized lexis from SL to TL.

References

Salter, Liora, and Alison Hearn. 1997. *Outside the Lines. Issues in Interdisciplinary Research*. Montreal:McGill-Queen's Press.

Thompson Klein, Julie. 1990. *Interdisciplinarity. History, Theory, and Practice*. Detroit: Wayne State University Press.

ON THE INTERDISCIPLINARITY
OF CORPUS-BASED TRANSLATION STUDIES

MONA ARHIRE

1 Introduction

Modern research relies heavily on corpora in a large variety of fields, fostering both humanistic and exact sciences, out of which disciplines appear to operate in unexpected combinations and often emerge into new *pair disciplines* or even *group disciplines*, sometimes made up of traditionally remote standing ones. Hence, the era of interdisciplinarity, as the last decades might be called, represents a revolutionary turn for the research community in plenty of fields since the concern for joint disciplines has been opening up new horizons.

2 The Interdisciplinary Character of Translation Studies

Casting a glance back towards the early history of translation studies, there is the prescriptive view that calls for recognition of its merits. After having thoroughly exploited the normative aspects of translations, scholars have opted for the challenging shift towards the descriptive perspective. This newly adopted, rather observant attitude has proved to be truly fruitful in findings, and has led to the age of what is known in the literature as Descriptive Translation Studies (DTS). The field has naturally determined the implication of researchers in various empirical studies. The cooperation with corpus linguistics and computational linguistics has soon been observed to increase the effectiveness and reliability of research, which has, in its turn, brought about the establishment of Corpus-based Translation Studies (CTS). Having emerged about two decades ago, CTS, at least partly, accounts for the interdisciplinary character of modern Translation Studies.

Looking at the very core of the problem, we have to count with the already well-established fact that Translation Studies makes up a field that is interdisciplinary by nature. By far not surprising anymore, this is

sustained by sundry of scholars, including Hatim, for instance, in asserting that *"translation studies has not remained a prisoner within one paradigm"*, while discussing *"linkages to other disciplines",* like contrastive analysis, sociolinguistics, corpus linguistics, text linguistics, psycholinguistics, cultural studies and deconstruction, gender studies, literature, etc. (Hatim and Mason 2001, 80).

The interdisciplinarity of Translation Studies is to be considered also from the viewpoint of translations interfering, in content, with any possible subject matter. As an intercultural means of communication between parties not sharing the same linguistic code, translation has always carried information, discoveries, inventions and findings across the world, having a decisive role in the progress of humanity within any possible field. This is one of the reasons why translation is considered both art and science, and speaks for the complexity of the activity, but also about its constraints. A translator is impossible to possess exhaustive knowledge in all the fields that he/she might have contact with, whether pertaining to the humanities or the exact sciences. The intermingling of translation studies with other disciplines provides it with an imprecise character in its delimitation from them, with direct consequences on the translator as a professional and social being. Roland Barthes' statement on interdisciplinarity sheds a doubt over the very identity of the discipline: *"Interdisciplinarity consists in creating a new object which does not belong to anybody"*. (Barthes 2007, 2). Wolf, resting on this assertion explains that interdisciplinarity is understood as a *"differentiated, multidimensional epistemological concept"* (Wolf 2007, 2).

Translation belongs to the area of intercultural communication, being, above all, a form of communication between two cultures, since there is no possible way to separate communication, language and culture in an act of translation (Tomalin and Stempleski 1993).

In the intention to provide a more orderly vision of interdisciplinarity within translation studies, we shall further on suggest a distinction between internal and external interdisciplinarity as far as the field of translation studies is concerned. *Internal interdisciplinarity* refers to the relation of translation studies with other disciplines pertaining to the wide area of letters, i.e. languages and literatures, for instance: applied linguistics, contrastive linguistics, intercultural communication, literary and cultural studies, discourse analysis, stylistics, pragmatics, lexicography, terminology, etc. By *external interdisciplinarity* we understand two distinct aspects: i) the contact of translation studies with so much and so diverse content knowledge, since it intrudes in any possible field, which, we daresay, makes of TS a genuine melting pot of disciplines. This

interdisciplinary type is inherent to the translational field and it has always existed since translation does not operate apart from it. It cannot help interfering with remote standing disciplines since the bulk of translation work does not lie within the field of letters. Secondly, ii) the necessary partnership with the field of computer science in translational research initiatives linked to domains like corpus linguistics, computational linguistics, and computer-based contrastive analysis.

The dawn of external interdisciplinarity of this latter type overlaps with translation studies evolving to corpus-based investigation. Accounting for a new type of interdisciplinarity it has by far not exhausted its resources and is still expected to unfold findings to contribute to the development of the translation field in all its aspects. It is this particular kind of interdisciplinarity in the domain of translational research which falls under the scope of this paper in its focus on Corpus-based Translation Studies (CTS), which depends by all means on computer processing throughout its research stages.

Obviously, the computational aid has become so invaluable and highly necessary in areas of language study, that it has turned to be integrated into fields traditionally standing at the opposite end of the range of sciences. It is no less the case of CTS, which does not operate beyond the involvement of information technology. On the contrary, the advantages soar up due the perspectives opened up by IT within linguistics generally, and TS particularly.

3 Electronic Corpora: General Overview

With a view to obtaining an overview of corpus-based research and the fields it pertains to, we have undertaken a survey of the existing electronic corpora, extracting information on the following criteria: i) *general criteria*, like: size, availability, initiation date, authorship, purpose(s), audience, user-friendliness, and ii) *specific criteria* referring to corpora types from various viewpoints: monolingual, bilingual, or multilingual; general or terminological; containing written or spoken language; synchronic or diachronic; full text, sample, or mixed; comparable versus parallel, but in terms of the text genres they include. Since an exhaustive incursion into the overwhelming amount of corpora available on the internet is impossible, we have considered a number of about 80 corpora, which we find fairly sufficient to provide relevant data. For this very article, we shall focus only on some of the aspects investigated, the ones in support of interdisciplinarity.

3.1 Purposes and audience of corpora

To start with, the purposes and target audience explicitly stated by corpora authors do not necessarily offer a clear image of the underlying intent. However, most of them mention in the guidelines that the corpus is designed for research, teaching and learning. Vagueness should not be surprising or interpreted as weakness. On the contrary, it stems for the large variety of options that unfold before several categories of professionals. Once a serious, reliable corpus has been created, it has tuned into food for scientific, academic, teaching and professional environments. This is a reason why is it worthwhile compiling a corpus: it is subject to subsequent uses in theoretical or applied areas, beyond the compiler's initial scope. Consequently, the potential audience is as varied as the corpora genres themselves, irrespective of whether stated as such in the corpus interface.

The generality of the aim statement of some of the corpora is counterbalanced by others stating overtly much more precise objectives and addressees. Purposes are expressed in terms of: knowledge extraction, text production, (machine) translation, lexicographic use, phonetics research, usage guides, natural language processing, comparative studies, for scholars investigating linguistic and literary history, social history, and other related fields. In contrast, exact sciences display a much weaker representation and indicate purposes like: computer application, technology development, or the acquisition of medical information. Most of the corpora stating specific purposes and audience will also consider the interest of the general public in simply getting informed across domains of knowledge.

3.2 Corpora genres

Text genres range from the ones containing everyday, accessible language to more specialised varieties. They vary so as to include journal and newspaper articles, literature, essays, biographies, autobiographies, personal letters, business correspondence, literature on medical ethics, college papers, and others. If we were to look at recordings of spoken language, we would notice a diversity of both professional and non-professional contexts where acts of oral communication have been selected for analysis.

3.3 Corpora types

Corpus typologies have constituted ground for debate among scholars especially since the emergence of Corpus-based Translation Studies. In this particular field, corpora are classified according to different criteria and on several levels, whose systematic presentation does not lie in the intention of the present paper. Still, the glance cast upon corpora types has provided evidence of the existence of parallel and comparable corpora in unbalanced numbers, the former being preferred to the latter. Researchers are obviously in favour of parallel corpora whether conceived as bilingual or multilingual text collections. Comparable corpora, even if reduced in numbers at present, are expected to raise more interest in future translational research due to the fact that they offer the opportunity of comparing the natural language with the language of translation, in the case of monolingual comparable corpora, for instance. Or, in another interesting approach, a bilingual comparable corpus can offer insights into two different natural languages. Obviously, the investigation entails similar text types in terms of genre, length, time stretch, etc. The few but fruitful investigations on comparable corpora have operated on the very essence of the translation universals, defining them on grounds of comparison between translational and non-translational language. A few languages are already represented in projects on comparable corpora, but it is not this type of corpus that the high amount of corpus-based projects is involved in. Probably, the most representative corpus of this category is the English Comparable Corpus (ECC), which is made up of two subcorpora, the Translational English Corpus (TEC) and the Non-Translational English Corpus (NON-TEC). The former includes a collection of computerized texts excerpted from different fields in translation into English from various languages, such as Italian, Portuguese, German, Spanish, French, Arabic, Jewish, Thai, and others. The corpus can be accessed free of charge and is available to the research community, together with a set of software tools which enable the investigation of the English language in translation.

Another important corpus, a mixed one - of spoken and written British English - is the British National Corpus, known as BNC, containing 100 million words, available on the internet as well. Other languages are also represented as far as research in translation theory is concerned, among which the following stand out: BOKR (the reference corpus of Russian), the UC (Uppsala Corpus, University of Tübingen, 2004), the English-Italian corpus CEXI (Aston, Bernardini & Zanettin, 2003), the English-Norwegian comparable corpus ENCP, and so many others for languages as Finnish, Swedish, Welsh or Irish. In North America, the most important

corpus is the COCA (Corpus of Contemporary American English), an open-ended, being permanently updatable. All these corpora are domain specific, and integrate within a determined time stretch.

The distinction by domain or content among corpora reveals that reference corpora are preferred to specialized ones. This means that the amount of heterogeneous corpora, possibly covering all varieties of a language (most commonly English) prevail over the ones limited to ESP research. Hence, the text collections have as sources more often pieces of writing addressing the general public, like newspapers, magazines, and literature, rather than specialized reviews, reference books, manuals, etc. One of the reasons must be the easy access to the former category as compared to the latter. If we were to take the compiler's stance, specialized material in a particular field requires specialist search, is often subject to copyright and confidentiality restrictions and is limited. However, research encounters in ESP can extend their impact outside the scientific environment and the academic use, serving the practitioners in a field or another as well. They would be of particular interest and use to translators, for example, whose voyage towards obtaining a coherent target language text might be thorny in highly specialized areas.

Written language corpora have been the initiating ones, the genres of which we have referred to previously. Even if written language collections were the first ones to be set under the scholars' magnifying glass, there is a rising amount of spoken language recording transcribed to make up distinct corpora or a subcorpus within a more extensive mixed one, consisting of both written and spoken language (e.g. BNC). Available genres of transcripts comprise political speeches, formal, semi-formal and informal meetings, presentations, telephone conversations, every-day conversations, private dialogue, conversations over lunch or monologues. Moreover, several varieties of the English language are sampled: American, Scottish, Irish, New Zeeland, South African, Australian English, etc. As far as the field of Corpus-based Translation Studies is concerned, it is obviously only the written texts that are of interest, translation pertaining to the written language.

Temporality is the criterion that accounts for the classification of corpora into synchronic and diachronic ones. Although diachronic corpora lie at the basis of interesting studies concerned with the contrastive analysis of the same language within distinct periods of time or different translations of the same classical literary works, the synchronic corpora type outnumbers the diachronic type. Most of the text compilations take into account a limited time stretch, within a maximum of one or two decades. Nevertheless, a number of corpora are open-ended, in that they

are updatable and constantly worked upon, the so-called monitor corpora. Interestingly, the majority of the corpora investigated were created in the 1990s, and the more comprising and reliable ones give evidence of regular updating, once in two or three years. After the year 2000 it seems that a number of languages caught up with the English language corpora and completed the list of languages represented as corpora in electronic format. One of the latest is probably the Bulgarian – English bilingual, parallel corpus, which is still under construction (BULTREEBANK GROUP), and, as far as our search has revealed, the first Bulgarian initiative of the kind.

3.4 Authorship and reliability of corpora

Among the features that a corpus needs to fulfil, a fore position is held by its reliability. This characteristic is comprisable enough to incorporate all the other ones: the quality, authorship, the considerable size, representativeness, good structure and user-friendliness.

As far as *authorship* is concerned, it is obvious that corpora have chiefly been created by researchers for researchers. Users of corpora might be guided in their search for material to work on looking for reputed authors and institutions, which usually guarantees for the reliability of a particular corpus. Evidently, corpora users, irrespective of their intentions and objectives, should critically assess a corpus before counting on it entirely. There are weakly administered corpora and sites in terms of temporality, mention of authorship, size, or other important data. Such items speak about the reliability and user-friendliness of a corpus, alongside clear instructions, good presentation, and internal structure. But generally, users are fortunate to find that we mainly owe the compilation of electronic corpora to universities or centres for linguistic studies, among which there are British ones, like Oxford and Cambridge Universities, the Centre for Speech Technology Research of the University of Edinburgh (BNC, CIC - Cambridge International Corpus being among the most representative ones). Northern American institutions have also contributed decisively from the University of Chicago, the Department of Linguistics; the English Language Institute of the University of Michigan (Michigan Corpus of Academic English), and others. But also non-English parts of the world have created corpora. For example in Italy: University of Bologna (CORIS/ CODIS), Göteborg University – the Department of Linguistics, Saarbrücken University in Germany (SCOSE), Karlovoy University of the Czech Republic (The Czech National Corpus), University of Wellington, New Zealand (Corpora at Victoria University of

Wellington), Japan (Tanaka Corpus). Sometimes corpora are the result of the work of joint research teams from around the world. A noteworthy amount of corpora are owed to information technology departments. Such cases are: the Department of Informatics of the University of Sussex, the School of Computing within the University of Leeds (The Quranic Arabic Corpus), the Computer Science Department of Leipzig University (WORTSCHATZ), and many others. Given the overwhelming representation of the English language among the existing corpora, there is only one conclusion that can be drawn: English it nowadays the language of highest interest throughout the world. Evidence thereof is the finding that not all the English corpora have origins in the English-speaking parts of the world. Most bilingual parallel corpora consider English as the pair language. Furthermore, it is definitely the *lingua franca* of corpus-based research since the interface, user's guides, instructions, metalanguage of the sites displaying corpora are in this very language.

Corpus size also speaks for reliability. The more comprising a corpus is, the more accurate the data extracted will be. Corpora vary considerably in size, but the most sizable ones range between hundreds of thousands and hundreds of millions of words. It is the merit of electronic devices that data can be collected nowadays from large volumes of computerized texts and therefore produce much more reliable and relevant empirical discoveries.

3.5 Accessibility

Most of the corpora are freely available on the internet, with special mention of the purposes as non-profit linguistic research and non-commercial ones. Some corpora require of the user, before allowing access, to fill in a statement in this respect. Others allow restricted use (to current students, faculty staff, authorised users) or offer access only to institutions. Some are open to be used by anybody but charge institutions differently from private users. The offer for a time-restricted free trial is also available at times. All this suggests that there is plenty of material available for research, but there is no exhaustive offer for any genre and text type ever, aspects that invite to further compilation, especially in specialized domains.

3.6 Preliminary conclusions

The diversity of the text genres and types stands for the considerable number of fields having interdisciplinary character simply by interfering

with computer science. This is due to the fact that text compilation, processing, and analysis rely on this indispensable and everywhere-intruding discipline. Computational Linguistics is one of the best examples, and its very name speaks for it. Obviously, the linguistic half of the discipline has its roots in corpus linguistics, which has been added value by the computational input, especially in the possibilities offered to natural language compilation and processing. Concretely, linguistics has benefited from corpora and the tools created by computational linguistics in a variety of ways: to discover unexpected linguistic patterns, to study workplace interaction, to observe collocational behaviour and semantic prosody, to enhance teaching foreign languages and translator training, but also to investigate speech acts, to perform genre or narrative analysis, to obtain lexical frequency lists, to compare variations of the same language, and many others. A special focus shall be given to the translational aspects of corpus-based research, which has displayed important progress once corpora started being taken as source of inspiration and linguistic discoveries. Moreover, the electronic tools set at the translation scholars' disposal have been crucial in providing new investigation methodology. It is precisely the meeting area between Translation Studies and IT that constitutes the departure point of Corpus-based Translation Studies.

4 Advantages of Corpus-based Research within Translation Studies

As indicated above, interdisciplinarity is a must in the field of corpus-based translation studies. The compilation of corpora depends on electronic resource to the same extent as the analysis of such corpora is determined by the possibilities of the software. Or, to put it differently, there is interdependency between analysis and IT tools, each stimulating the other. Here, we have touched upon the essence of interdisciplinarity, namely the fact that the domains at stake are subject to mutual influence, challenge and stimulation. This is the way in which new perspectives open up to researchers, while each field extends its reach over others.

The pure research has been enriched with outstanding means of investigation since the very dawn of corpus-based studies. Hence, it is also able to display a similarly invaluable harvest of findings. There are advantages to be traced throughout the stages of corpus-based translation studies as well: all the way from the very initiating, but crucial stage, namely that of corpus compilation, up to the highly productive linguistic analysis. It is precisely the interdisciplinary character of CTS that allows for in-depth scrutiny in both the process and the product dimension of

translations.

But in order to reach the point of investigation, we need to proceed through the activity of collecting texts or text samples to be included in the corpus. Nowadays, written text corpora are hardly imagined to be compiled, sorted and categorized outside the area of electronic devices. At this stage, the advantages that the cooperation with computer science offers are manifold. Even if traditionally collected texts are still scanned and electronically converted to the desired and unitary format, this activity can never compete with the amount of texts available on the internet in all the possible fields and of so diverse a genre. The electronic search tools definitely offer easy and time-saving identification, selection, ordering and preparation of texts for analysis, even if internet searching is far form being perfect. The constraints and limitations that the internet sets on corpora compilation are strictly linked to the features corpora need to fulfil in order to lead to reliable investigation results. Such conditions would refer to authorship, authenticity of texts and contexts, length, the author's writing skills, etc., some of which have been discussed above.

As for the analysis proper, tools have been constantly created to allow for new approaches. Within the translational field, we can benefit today from findings that couldn't have been possible with traditional means of language processing. The modern investigation technologies have led to the creation of research methodologies which have produced a turn of vision in the examination of translations and have determined shifts in the very essence of the translation universals. More precisely, the corpus-based methodology has favoured systematic investigation of the translational language and patterning, on sizable corpora, leading thus to more consistent results. Among others, the universals of *simplification, normalization and explicitation* have been subject to different perception changes due to the corpora-based approach to translations (Laviosa 2002, 58). Baker's investigation, in 1995, on a monolingual comparable corpora revealed that the translated language operated after different qualitative norms as compared to the language of original texts. Undoubtedly, the comparison rests on similar text genres. Ever since the change of view for performing contrastive analysis between original texts and translated ones into the same target language as the originals, the universals have been constantly enriched by researchers involving tens of languages in such encounters.

To be more specific, the list of electronic tools available for linguistic analysis, at different language levels, comprises concordancers, text aligners, word count instruments, as well as devices for the creation of lexical frequency lists and for identifying recurrent linguistic patterns.

Some sites providing corpora offer electronic tools for language processing as well, usually for free. In addition, the electronic tools are accompanied by instructions and there is also evidence to be found as concerns the research projects that used those tools for investigation. Further details on the topic have given birth to plenty of articles accounting for such undertakings. For example, in 2000, Wools' complex project on parallel corpora used computational programs, such as WordSmith, and became an important exploitation resource in a variety of ways, among which: to examine the translator's choices and strategies, to test the validity of claims made in translation theory, and to formulate pedagogic strategies for the training of translators. Furthermore, the study pinpoints one of the major advantages of using parallel and comparable corpora, namely the refinement of the contrastive knowledge of two languages that come into contact in the process of translation.

The outcome of the electronic-based analysis is of help to professionals, alongside other tools that they have at hand in the process of translation. The aid that they can take advantage of consists in obtaining concordances, fast and effective search for words, relevant search results due to contextual visualization of lexical items, illustrations of usage on the same contextual account, and others, all of them having as advantage the production of more accurate and fluent target language texts.

5 Conclusions

The collision between disciplines broadens also the social interaction of the professionals whether in the scientific, academic or practitioners' environment. That is why interdisciplinarity can be viewed as personal accomplishment, were we to consider the human involvement in other fields, constantly stimulating them and leading to the personal progress of the people involved, as professional and social beings. Translators cannot and should not be solitary in their work. They need the help of specialists or of specialized knowledge in the fields they interfere with, even if electronic corpora can offer fast and easily accessible information, being a valuable source of content knowledge. The translator is the social being operating in social and communicative events, always exceeding borders or meant of bridge disciplines. In recognition of this, the newly emerging area of study, the *sociology of translation* has been observed to acquire a well-deserved status over the last few years. The newly created discipline discusses the translator's social implication in, for instance, the translation of dialects or culture-related language.

A discussion about the overridden cultural transfer might seem

superfluous nowadays, but the analysis of translation errors has uncovered the fact that the cultural differences stem for most of the errors in translation rather than linguistic incompatibility between languages or insufficient linguistic competence on the translator's part. That is the reason why it is essential for translators to have *"bi-cultural vision"* (Hatim and Mason 1994, 223) or possess *"cultural fluency"* (Tomalin and Stempleski 1993), apart from the linguistic competences.

On account of its social and also interdisciplinary nature, translation studies extends its impact on a closely related discipline, one that is itself so generous and rewarding in terms of interdisciplinary, namely intercultural communication. Similar to translation studies, it pertains to whatever field can be imagined, envisaging increased quality of all kinds of international relations, as is the common and ultimate goal of the translational and the communicational fields.

The benefits of interdisciplinarity generally and within Corpus-based Translation Studies in particular are not debatable. But for its interdisciplinary character, the translational field could not have evolved into such a dynamic and constantly stimulating study area. The virtues of interdisciplinarity within Translation Studies are felt beyond the scientific environment, outside the pure discipline since it stretches its influence over the profession of translator via the academic and translator training programmes. The number of Centres for Corpus-based Research around the world, as well the multitude of national and international projects whose analyses are based on corpora, stand proof of the merits of corpus-based research, including CTS, merits that are to a considerable extent grounded on the feature of interdisciplinarity attributed to such research. In a nutshell, Translation Studies "has been enriched by dint of possessing such a multi-faceted nature" (Snell-Hornby 2007, VI).

In order to end in an optimistic key, we can only be hopeful as regards the Romanian translational research domain for fast development in the direction of corpus-based studies, and with a view to including the Romanian language on the long list of languages currently worked on.

References

Baker, Mona. 1996. "Corpus-based Translation Studies. The Challenges that Lie Ahead." In *Terminology, LSP and Translation*, edited by Harold Somers. Amsterdam: John Benjamins.

Chesterman, Andrew. 2007. "Bridge Concepts in Translation Sociology." In *Constructing a Sociology of Translation*, edited by Michaela Wolf, and Alexandra Fukari. Amsterdam/Philadelphia: John Benjamins

Publishing Company.

Corpas Pastor, Gloria, and Miriam Seghiri. 2007. "Specialized Corpora for Translators: A Quantitative Method to Determine Representativeness." *Translation Journal 11*(3). Accessed November 30, 2009. http://accurapid.com/journal/41corpus.htm.

Hatim, Basil, and Ian Mason. 1994. *Discourse and the Translator*. London/New York: Longman.

Hatim, Basil. 2001. *Teaching and Researching Translation*. U.K.: Longman.

Laviosa, Sara. 2002. *Corpus-based Translation Studies. Theory, Findings, Applications*. Amsterdam/New York: Rodopi.

Shuttleworth, Mark, and Moira Cowie. 1997. *Dictionary of Translation Studies*. U.K.: St. Jerome Publishing.

Tomalin, Barry, and Susan Stempleski. 1993. *Cultural Awareness*. U.K.: Oxford University Press.

Tymoczko, Maria. 1998. "Computerized Corpora and the Future of Translation Studies." *Meta: Translators' Journal 43*(4):652-660. Accessed January 10, 2010. http://www.erudit.org/revue/meta/1998/v43/n4/004515ar.pdf.

Wilkinson, Michael. 2006. "Compiling Corpora for Use as Translation Resources." *Translation Journal 10*(1). Accessed April 15, 2010. http://accurapid.com/journal/35corpus.

Wolf, Michaela. 2007. "The emergence of a sociology of translation." In *Constructing a Sociology of Translation*, edited by Michaela Wolf, and Alexandra Fukari. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Zanettin, Federico. 2000. "Parallel Corpora in Translation Studies: Issues in Corpus Design and Analysis." In *Intercultural Faultiness. Research Models in Translation Studies I: Textual and Cognitive Aspects*, edited by Maeve Olohan. Manchester: St. Jerome.

Annex 1. List of Corpora

BOKR (The Russian Reference Corpus), http://www.ruscorpora.ru/en/
BNC (British National Corpus), http://corpus.byu.edu/bnc/
BULTREEBANK (Bulgarian Treebank) GROUP,
    www.bultreebank.org/ESyntAC
COCA (The Corpus of Contemporary American English),
    http://www.americancorpus.org/
ENCP (English-Norwegian Parallel Corpus),
    http://www.hf.uio.no/ilos/tjenester/kunnskap/sprak/omc/enpc/ENPCma
    nual.html