

- Shuy, R. 1981. A holistic view of language. *Research in the Teaching of English* 15: 101-111.
- Sinclair, J. (ed). 1990. *Looking up*. London: Collins.
- Sinclair, J. 1991. *Corpus, concordance and collocation*. Oxford: Oxford University Press.
- Sperber, D. & D. Wilson. 1986. *Relevance: communication and cognition*. Oxford: Blackwell.
- Studdert-Kennedy, M. 1987. The phoneme as a perceptuomotor structure. *Language perception and production*, ed. by A. Allport, D. MacKay, W. Prinz & E. Scheerer, 67-84. London: Academic Press.
- Swales, J.M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Tannen, D. 1984. The pragmatics of cross-cultural communication. *Applied Linguistics* 5: 189-195.
- Tannen, D. 1989. *Talking voices: repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Taylor, T.J. 1992. *Mutual misunderstanding*. London: Routledge.
- Towell, R. & R. Hawkins. 1994. *Approaches to second language acquisition*. Clevedon, Avon: Multilingual Matters.
- Wesche, M.B. 1994. Input and interaction in second language acquisition. *Input and interaction in language acquisition*, ed. by C. Gallaway & B.J. Richards, 219-249. Cambridge: Cambridge University Press.
- Yip, V. 1994. Grammatical consciousness-raising and learnability. *Perspectives on pedagogical grammar*, ed. by T. Odlin, 123-139. Cambridge: Cambridge University Press.

From CA to CIA and back:  
An integrated approach to  
computerized bilingual and  
learner corpora

Sylviane Granger

### 1. Introduction

After dominating applied linguistics throughout the 1960s, contrastive analysis (CA) went through a period of disfavour which was to last for well over a decade. Today, however, it is reemerging as a major approach in translation and interlanguage studies. This revival can be attributed first and foremost to the new age of computerized bilingual and learner corpora, which have enabled researchers in translation and interlanguage studies to make use of the methods and tools first promoted in corpus linguistics and thus to adopt a scientific, empirically-based approach which was previously lacking.

### 2. CA: Computerized bilingual corpora

Indeed, one of the major weaknesses of contrastive studies in the past proved to be their slim empirical foundation. To Odlin (1989: 29) "the descriptive shortcomings of contrastive accounts are quite evident". Back in 1984, in his survey of the major contrastive projects of the 1970s, Filipovic had already stressed the necessity to base CA on representative bidirectional corpora. His survey showed that "all the projects [...] had some kind of corpus, but none of them had a standard corpus" (1984: 110). In those days, however, the difficulties of collecting and processing a computerized corpus were so great that Filipovic

himself, convinced though he was of the importance of bidirectionality, had to be satisfied with a unidirectional corpus.

It was recognition of that empirical weakness that led Wollin and Lindquist to organise the Lund Symposium on Translation Theory in 1985. They firmly believed that the time had come “to devote more energy to carrying out carefully planned, large-scale investigations of authentic translation with the aim to verify and falsify existing theories and theoretical proposals” (Wollin & Lindquist 1986: 10). They hoped to see translation studies “move away from their initial stage of impressionistic, non-professional or purely theoretical work to become a more empirically based research field” (p 11). Research based on computerized bilingual corpora is now making this possible.

According to Baker (1993: 235), we are at a turning point in the history of translation studies, a turning point which will come “as a direct consequence of access to large corpora of both original and translated texts, and of the development of specific methods and tools for interrogating such corpora”. She attests to the fact that “corpora in translation studies have so far been very modest affairs” (p 241) and expresses her dissatisfaction with the introspective methods commonly used by translation scholars. She is very optimistic about the contribution that bilingual corpora and corpus-driven methodology will make in the field of translation studies.

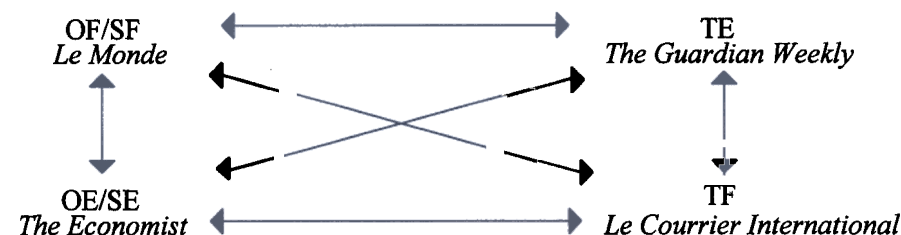
Computerized bilingual corpora may consist of original texts and their translations (type 1) or of parallel original texts in two or more languages (type 2). No fixed terminology has yet been established to describe these different types of corpus. Johansson & Hofland (1994) use the expression ‘parallel corpus’ to cover both types. Aijmer et al. (this volume) speak of ‘translation corpus’ for type 1 and restrict the use of the term ‘parallel corpus’ to type 2, while Baker (1993: 248) uses the terms ‘parallel corpus’ and ‘comparable corpus’ respectively. As the term ‘translation corpus’ has traditionally been used in CA to refer to corpora of original texts and their translations, it would seem sensible to adopt Aijmer et al.’s terminology.

Both types of bilingual corpora have their advantages and disadvantages. The major plus factor of the translation corpus is that it provides “a unique opportunity of showing how the same meaning is conveyed in languages” (Johansson & Hofland, 1994: 26). Its main weakness is that translated texts are regarded as “distorted versions of ‘real’ texts” (Baker 1993: 233). In the case of parallel corpora, there is no doubt about the authenticity of the data; the difficulty in this case is to ensure text type comparability.

The computerized corpus of English-French journalesse currently being compiled at Louvain is a combined parallel/translation corpus, as represented in Figure 1.

As such, the corpus allows for a 4-way comparison:

- (1) OF vs OE, ie original French journalesse (*Le Monde*) vs original English journalesse (*The Economist*);
- (2) SF vs TE / SE vs TF, ie original French journalesse (*Le Monde*) vs translated version in English (*The Guardian Weekly*) and original English journalesse (*The Economist*) vs translated version in French (*Le Courrier International*);
- (3) OE vs TE / OF vs TF, ie comparisons of original and translated French and English;
- (4) TE vs TF, ie comparisons of translated English and translated French.



OF/OE = Original French/English

SF/SE = French/English as Source Language

TF/TE = Translated French/English.

Figure Computerized corpus of English-French journalesse

The first two types of comparison aim to provide better contrastive descriptions of English and French, whilst the last two contribute to a better understanding of the nature of translated language, which, under the heading of ‘translationese’, has come to be viewed as a variety in its own right.

Even at its current modest size of 200,000 words, the corpus is realising its potential as a testbed for English-French contrastive statements such as those made in the influential English-French contrastive study by Vinay & Darbelnet, *Stylistique comparée du français et de l’anglais*. Though originally published as

early as 1958, it continues to have a deep influence on current contrastive studies, which draw on many of its basic tenets.

Several pilot studies carried out at Louvain demonstrate the usefulness of using corpus data to challenge what are often no more than carefully observed intuitions contained in these studies about English-French differences. They have shown that intuitions are rarely "wrong" but that they often give only a partial – and hence often misleading – view of reality.

For example, most contrastive studies on English and French contrast the flexibility of the English word formation system with the rigidity of the French one, giving convincing illustrations of the processes of compounding, derivation and conversion which can be used to create new words in English whilst French has to make do with paraphrases. This idea is found in Vinay & Darbelnet (1969: 69) and taken over in Van Hoof's (1989) more recent book on English-French translation. As appears from the following passage, Van Hoof (1989: 17-18) puts derivation and compounding on an equal footing, both processes being presented as more flexible in English than in French: "l'anglais accuse une dérivation plus complète que le français et offre des facilités de composition ignorées de celui-ci" ['English has a more complete system of derivation than French as well as facilities for compounding not available in French']. Tisthoud (1992) investigated all lexical gaps in the Louvain journalesse bilingual corpus in order to test the validity of this statement. The data confirmed that compounding was indeed much more prominent in English than in French, but showed that the situation concerning derivation was a lot more complex than the literature leads one to believe. In fact, Tisthoud found that both languages had an equivalent number of lexical gaps due to derivation, but whereas French practically always uses paraphrases to translate English derivations, English uses paraphrases *and* compounding, as the following examples demonstrate:

#### A. English to French

|                        |                                            |
|------------------------|--------------------------------------------|
| <u>unequivocally</u>   | sans ambiguïté aucune                      |
| <u>unwittingly</u>     | sans le savoir                             |
| <u>unthinkingly</u>    | à la légère                                |
| <u>exclusiveness</u>   | le caractère élitiste                      |
| <u>fellowship</u>      | le sentiment d'appartenir à une communauté |
| <u>homelessness</u>    | la crise du logement                       |
| <u>neighbourliness</u> | les relations de bon voisinage             |
| <u>breezy</u>          | dans un courant d'air                      |

#### B. French to English

##### B1. Paraphrases

|                          |                                              |
|--------------------------|----------------------------------------------|
| <u>improductivité</u>    | lack of productivity                         |
| <u>vingtaine</u>         | twenty or so                                 |
| <u>riverains</u>         | people living alongside (a river, road, etc) |
| <u>clochettes</u>        | little bells                                 |
| <u>ruelles</u>           | narrow streets                               |
| <u>approvisionnement</u> | obtaining supplies                           |
| <u>délabrement</u>       | state of disrepair                           |
| <u>comptabilités</u>     | book-keepers' records                        |
| <u>revanchisme</u>       | vengeful demands                             |

##### B2. Compounds

|                       |                    |
|-----------------------|--------------------|
| <u>tannage</u>        | leathermaking      |
| <u>contestataires</u> | troublemakers      |
| <u>contribuable</u>   | tax-payer          |
| <u>migraineux</u>     | migraine-sufferers |
| <u>festivalier</u>    | festival-goer      |
| <u>bétonneuse</u>     | cement mixer       |
| <u>glaçon</u>         | ice cube           |
| <u>vigneron</u>       | wine grower        |
| <u>attentiste</u>     | wait-and-see       |

Although relatively fragmentary, this study brings out two interesting phenomena: firstly, that French derivation is much richer than is generally assumed; and secondly that there is an interesting cross-linguistic link to be made between French derivation and English compounding. Interestingly, the only contrastive study that comes close to our conclusions is corpus-based. In Chuquet & Paillard's (1987) investigation of English-French translation difficulties, based on a reduced bidirectional journalesse corpus, it is stated that "l'anglais privilégie la composition là où le français privilégie la dérivation par affixation" (p 189) ['English favours compounding where French favours derivation by affixation'].

Other of our studies confirm intuitive contrastive statements. Frère (1993), for example, studied the process of transposition from noun to verb and from verb to noun in French-English/English-French translations. As French is usually presented as a nominal language and English as a verbal language (cf

Vinay & Darbelnet, p 102, Van Hoof, p 62, Chuquet & Paillard, p 135), one would expect to find much greater evidence of verb to noun transposition in French translation than in English translation. As shown in Table 1, this was overwhelmingly confirmed by Frère's research.

Table Noun><Verb transpositions

|                          | English > French | French > English  |
|--------------------------|------------------|-------------------|
| Verb > Noun              | <b>72% (66)</b>  | <b>32.5% (39)</b> |
| Noun > Verb              | 28% (26)         | 67.5% (81)        |
| Total N/V transpositions | 100% (92)        | 100% (120)        |

The research we have conducted so far shows that scholars who have adopted a largely impressionistic, non-statistical contrastive approach have a tendency to overgeneralize. The fact that English is particularly flexible in the use of *some* suffixes such as the noun forming suffix *-ness* or the adverb forming suffix *-ly* is taken to mean that derivation as a whole is much more flexible in English than in French. Ongoing research on connectors seems to point in the same direction. Because French makes greater use than English of *some* connectors (such as *ainsi* or *en effet*), it is concluded that French uses more explicit connectives to signal the logical relationships between sentences than English, which prefers to keep the links implicit. Preliminary results indicate that this too may well turn out to be an overstatement.

Even at this early stage in our investigations, it has become clear that contrastive Sprachgefühl needs to be set against hard contrastive data. Bilingual computer corpora are particularly useful in this respect because they can be submitted to a whole range of language software (concordancers, taggers, parsers, etc). However, the dual nature of bilingual corpora calls for the development of new computer tools, notably text alignment software (Johansson & Hofland 1994), bilingual concordancers and translation checkers (Isabelle 1992). These are unfortunately not yet freely available.

### 3. CIA: Computerized learner corpora

Whilst researchers at Louvain are actively investigating bilingual translation corpora, the major focus of current research is in the field of interlanguage (IL). In this field too, CA has always played an important role, first in its 'strong version', then later in its 'weak version' when it was relegated to a position of secondary importance, although never fully abandoned. Recently however, there has been widespread recognition of a severe limitation in the scope of traditional CA. Pery-Woodley (1990: 143) for example feels there is a need for CA research concerned with "comparing/contrasting what non-native and native speakers of a language do in a comparable situation". Selinker (1989: 285) equally sees great potential in this "new type of CA", whose object is to compare parallel NL (native language) and IL data.

Indeed, research into learner writing conducted at Louvain within the framework of the International Corpus of Learner English (ICLE) project has demonstrated the need for an integrated contrastive model, combining CA and a new type of CA, which I have called Contrastive Interlanguage Analysis or CIA. Unlike classical CA, CIA does not establish comparisons between two different languages but between native and learner varieties of one and the same language. Odlin (1989: 28) stresses the usefulness of this type of comparison to evaluate the occurrence of transfer, while James (1993: 14) sees it as a particularly apt basis for a "quantificational contrastive typology of a number of English ILs". According to Krzeszowski (1990: 212) this search for "individual style" in interlanguages involves massive statistical research since "each national interlanguage of English is probably characterized by a unique matrix of frequencies of occurrence of various linguistic forms". In his view, empirical verification of these hypotheses "will keep contrastive studies alive for some time to come".

The CIA approach is at the core of ICLE project (see Granger 1993), whose main aim is to uncover factors of 'foreign-soundingness' in learner writing.

The first step of the project was to gather and computerize a number of subcorpora of advanced non-native English essay writing, with each subcorpus representing a different language variety.<sup>1</sup>

Once computerized, the corpus is ready to be submitted to contrastive interlanguage analysis, as represented in Figure 2.

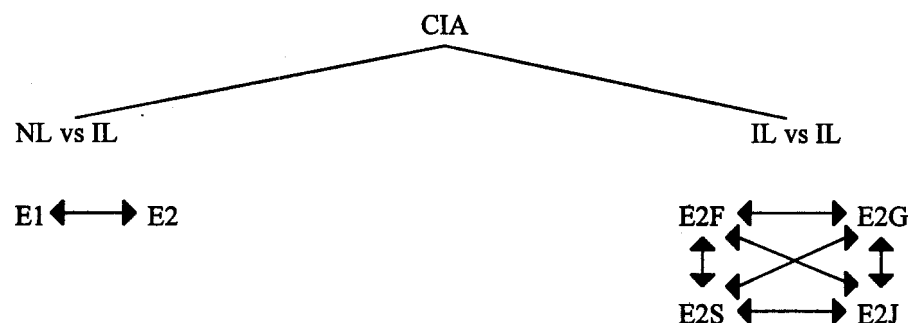


Figure 2. Contrastive interlanguage analysis

CIA involves two types of comparison:

- (1) NL vs IL, ie comparison of native and non-native varieties of one and the same language. Within ICLE, we contrast English as a first language (E1) and as a foreign language (E2).<sup>2</sup>
- (2) IL vs IL, ie comparison of different interlanguages of the same language: the English of French learners (E2F), German learners (E2G), Swedish learners (E2S), Japanese learners (E2J), etc.

It was clear from the start that the reliability of the investigation depended crucially on the quality of the data. Ellis (1994: 49) has underlined "the importance of collecting *well-defined samples* of learner language" (my italics). He notes that many EA (error analysis) studies are "difficult to interpret and almost impossible to replicate" because not enough attention is paid to the variety of factors that can influence learner language. In fact, one of the reasons why many important issues in applied linguistics have remained inconclusive is that researchers have not been comparing like with like. For this reason we established from the outset a number of variables which are rigorously controlled:

- o *type of learner*: foreign language learners, not second language learners
- o *stage of learning*: advanced
- o *text type*: argumentative essay writing

These variables are controlled through a learner profile questionnaire, which all learners are requested to fill in and which contains all their relevant biographical

information. This information is encoded in the form of standardised file headers and can be subsequently retrieved, enabling researchers to generate their own tailor-made subcorpora (a subcorpus of French-speaking learners of English with Dutch as their third language, for instance).

It is not enough, however, for the different ILs to be comparable. For the first type of CIA comparison (see Figure 2), we need to have access to a comparable corpus of NL English. The temptation is great to use existing computer corpora of English but in view of the genre sensitivity of many linguistic features, comparisons may be flawed.<sup>3</sup> A control corpus of comparable writing by native speakers of English has therefore also been assembled at Louvain.

The machine-readability of the data makes it possible to study the phenomena of over- and underrepresentation, which have largely been neglected in studies of learner language. These phenomena call for a quantitative contrastive approach, as Krzeszowski (1990: 206) points out: "In either case the learner deviates in plus or in minus from a certain statistical norm which characterizes native performance in a particular language. To ascertain such an error, one has to perform a quantitative contrastive study of texts written by native users of a particular language and by a non-native user of the same language and compare the frequencies of use of the investigated forms".

To be reliable this quantitative investigation has to be conducted on a large scale. This makes computerizing the data necessary rather than optional. As the ICLE corpus will eventually exist in three different versions – raw, tagged and parsed – it will be possible to go beyond the over- and underrepresentation of lexical words and phrases to investigate cases of over- and underuse of word categories and structures.

The contrastive investigation of raw frequencies, though undoubtedly the least sophisticated type of quantitative comparison, can in fact be a very fruitful undertaking. For instance, comparing the top 45 words in a 100,000 word corpus of French learners' writing and a control corpus of native speaker writing of similar size shows that learners overuse the preposition *of* and underuse *by*, and similarly that they overuse *but* and underuse *and* (see Granger et al. 1994). Qualitative research can then be carried out to analyse the reasons for this.

Concordancing allows researchers to go beyond this purely quantitative stage and find out more about the actual use of words in context. A mere glance at the concordances of the connector *so* in native and non-native writing suffices to see that there is a clear overuse of sentence initial *so* by the learners (see Granger & Tyson, forthcoming).

Other computer tools allow researchers to effect much more sophisticated searches. Collocation generators, for instance, enable analysts to retrieve sequences of words which recur in a text. This will help applied linguists form a clearer picture of learners' phrasal lexicon. One example of this comes from some research carried out in Louvain into learners' use of prefabs. While the unidiomaticity of their production is often attributed to a lack of prefabs, our own corpus-based research (see Granger, forthcoming) shows that it may equally well be assigned to an overuse of prefabs, notably topic introductory sequences such as *as far as X is concerned*, *we can X that* or *from the point of view of X*.

#### 4. Integrating CA and CIA

Far from seeing computerized bilingual corpora as the private ground of translation specialists and typologists and computerized learner corpora as the sole concern of applied linguists, we see the two types of corpora as closely interrelated. The integrated CA/CIA contrastive model we rely on can be represented as is Figure 3.

In keeping with the current state of applied linguistics research, the notion of transfer occupies a central position in this model. According to Selinker (1992: 171) the primacy of transfer can now be considered as "SLA fact". He views interlingual identifications as "a basic, if not *the* basic, SLA learning strategy". This should lead specialists to adopt "a richer language transfer perspective which would include translation phenomena" (p 259-260).

The model involves constant to-ing and fro-ing between CA and CIA.<sup>4</sup> CA data helps analysts to formulate predictions about interlanguage which can be checked against CIA data. So, for instance, the fact that the passive is twice as frequent in English as in French (see Granger 1976) may lead one to expect an underuse of the passive in the writing of French learners of English. Conversely, CIA results can only be reliably interpreted as being evidence of transfer if supported by clear CA descriptions. To take the example of the passive again, explaining cases of underuse or overuse in IL productions requires access to reliable CA data. It is important to note that the terms 'predictive' and 'diagnostic' refer to mere hypotheses, which can be confirmed or refuted by corpus investigation.

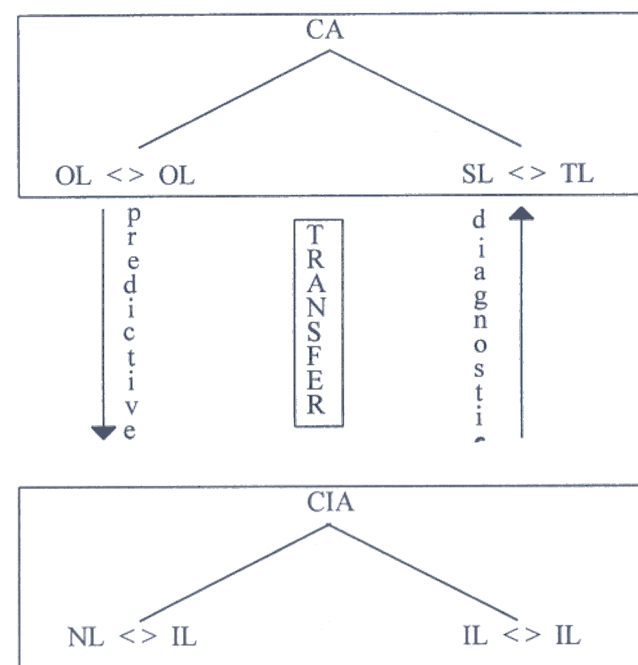


Figure 3. Integrated contrastive model

Lack of reliable CA data casts doubt on the validity of results in previous IL investigations. In a recent article Kamimoto et al (1992: 257) convincingly demonstrate that the lack of "L1 base-line data" deeply flaws the argumentation behind Schachter's influential paper on avoidance.<sup>5</sup> They reassess Schachter's conclusions using bilingual data and show that for avoidance to be firmly established "it is necessary to look at the first language form, distribution and function of the entity supposedly being avoided in the L2" (p 251).

Evidence of transfer does not only depend on CA data: it also crucially depends on the quality of the interlanguage data. The following example illustrates this very clearly. Faced with overuse of sentence initial *so* by French learners, we seemed to find a convincing transfer-related explanation in their also very frequent use of sentence initial *ainsi*, a translation equivalent for *so*. However, subsequent investigation into the use of *so* by learners from other language backgrounds, revealed that sentence initial position was also characteristic of their production. It is thus essential to conduct comprehensive IL comparisons in order to distinguish between L1-related features and cross-

linguistic invariants, ie those features which are common to all learners, irrespective of mother tongue.

A particularly promising type of comparison that the combined CA/CIA model allows is that between translated language and learner language. Both are situated somewhere between L1 and L2 and are likely to contain examples of transfer. The similarity between translationese and 'learnerese' is clearly brought out by the following two sentences, extracted from our journalesse English-French corpus and the French subcorpus of ICLE respectively.

- (1) While, under his leadership, the old Labour Party has shed some of the most controversial aspects of its manifesto and is admittedly more credible today, one more setback could be fatal to its leader.  
(*The Guardian Weekly*)
- (2) Although to achieve ultimate political integration of the member countries, a true monetary union and above all increased participation of the European countries represented in the Parliament, is required. (ICLE: E2F)

Both sentences give evidence of what Gellerstam (1986: 94) calls "syntactic fingerprints". The sentences both sound extremely stilted because the tendency in French to use particularly heavy themes has been transferred.<sup>6</sup>

Translated language does not only display L1-dependent features. Like interlanguage, it also displays universal features, ie "features which typically occur in translated text rather than original utterances and which are not the result of interference from specific linguistic systems" (Baker 1993: 243). The methodology Baker suggests using to set apart these two types of features is strikingly similar to CIA and could accordingly be called Contrastive Translation Analysis (CTA). A corpus of texts translated into English from a variety of languages helps to "isolate patterns which (i) occur across the corpus, irrespective of whether the source texts are French, Hebrew or Chinese, (ii) do not occur, or do not occur to the same degree/with the same frequency, in original texts" (p 245).

Several linguists restrict the use of the term 'translationese' to L1-dependent features. To Johansson & Hofland (1994: 26) the term refers to "deviance in translated texts induced by the source language". Because of the difficulty of distinguishing between L1-related and universal features, it would seem wise to use the term in the wider sense in which the term 'interlanguage' is

used, to cover all those features which distinguish translated language from original language, whether or not they are transfer-related.<sup>7</sup>

## 5. Conclusion

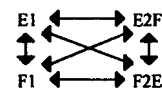
In 1970 Wardhaugh suggested that CA was due for "a period of quiescence". In the nineties, it is showing all the signs of effervescence and is once more a promising field of research. Equipped with the methods and tools of corpus linguistics, the new-look CA and its offshoot, CIA, represent an exciting new departure for both translation and interlanguage studies.

## Acknowledgements

We acknowledge the support provided by the Belgian National Scientific Research Council and the University of Louvain Research Fund in this research.

## Notes

1. Ten of these are currently covered by ICLE: Chinese, Czech, Dutch, Finnish, French, German, Japanese, Polish, Spanish and Swedish.
2. In view of our interest in English-French contrasts, we have also collected a similar corpus of French, ie essays written by native speakers of French and advanced English-speaking learners of French. This bidirectional interlanguage corpus can be represented as follows:



3. This danger is clearly demonstrated in a study of connectors in native and learner writing (Granger & Tyson, forthcoming). While a comparison of the frequency of some individual connectors in learner writing with the LOB corpus points to an overuse by the learners, a comparison with the Louvain English essay writing corpus shows a clear underuse.
4. The integrated contrastive model bears some resemblance to Danchev's (1991) "model of expanded contrastive analysis" which also advocates a combined CA/IL perspective. However, it differs from our model in several important respects, notably in its empirical foundation.
5. The authors also deplore the lack of methodological detail in Schachter's study, and in particular "the lack of controls on the writing task. We know nothing about topics selected, for instance, nor how much time was allocated to them, nor the mean lengths of the compositions by the various groups" (p 256). This statement lends support to the rigorous data collection procedures adopted within ICLE.
6. Following McCarthy & Carter (1994: 73) we use the term 'theme' to refer to "everything in first position that precedes the main verb of the main clause."
7. Here too no clear terminology has been fixed upon. Baker (1993: 249) restricts the use of the term 'translationese' to features which result from "the translator's inexperience or lack of competence", while Gellerstam (1986: 88) explicitly excludes "anecdotal instances of bad translations"

## References

- Baker, M. 1993. Corpus linguistics and translation studies: Implications and applications. *Text and technology. In honour of John Sinclair*, ed. by M. Baker, G. Francis & E. Tognini-Bonelli, 233-250. Philadelphia/Amsterdam: Benjamins.
- Chuquet, H. & M. Paillard. 1987. *Approche linguistique des problèmes de traduction anglais > français*. Paris: Ophrys.
- Danchev, A. 1991. Some notes on a Bulgarian model of expanded contrastive analysis. *Languages in contact and contrast: Essays in contact linguistics*, ed. by V. Ivir & D. Kalogjera, 87-101. Berlin: Mouton de Gruyter.
- Filipovic, R. 1984. What are the primary data for contrastive analysis? *Contrastive linguistics - Prospects and problems*, ed. by J. Fisiak, 107-117. Berlin/New York/Amsterdam: Mouton.
- Frère, L. 1993. *The verb-noun transposition in the English-French French-English translation of journalistic texts*. Unpublished MA dissertation. Université Catholique de Louvain.
- Gellerstam, M. 1986. Translationese in Swedish novels translated from English. *Translation studies in Scandinavia*, ed. by L. Wollin & H. Lindquist, 88-95. Lund: CWK Gleerup.
- Granger, S. 1976. Why the passive? *English-French contrastive analyses*, ed. by J. Van Roey, 23-57. Leuven: Acco.
- Granger, S. 1993. The International Corpus of Learner English. *English language corpora: Design, analysis and exploitation*, ed. by J. Aarts, P. de Haan & N. Oostdijk, 57-71. Amsterdam/Atlanta: Rodopi.
- Granger, S. Forthcoming. Prefabricated patterns of advanced EFL writing: collocations and lexical phrases. *Phraseology*, ed. by A. Cowie. Oxford: Oxford University Press.
- Granger, S., F. Meunier & S. Tyson. 1994. New insights into the learner lexicon: a preliminary report from The International Corpus of Learner English. *Entering text*, ed. by L. Flowerdew & A.K.K. Tong, 102-113. The Hong Kong University of Science and Technology.
- Granger, S. & S. Tyson. Forthcoming. Connector usage in the English essay writing of native and non-native EFL speakers of English. *World English* (1996).
- Isabelle, P. 1992. La bi-textualité: vers une nouvelle génération d'aides à la traduction et la terminologie. *Meta* 37: 721-737.
- James, C. 1994. Don't shoot my dodo: On the resilience of contrastive and error analysis. *IRAL* 32: 179-200.
- Johansson, S. & K. Hofland. 1994. Towards an English-Norwegian parallel corpus. *Creating and using English language corpora*, ed. by U. Fries, G. Tottie & P. Schneider, 25-37. Amsterdam/Atlanta: Rodopi.
- Kamimoto, T., A. Shimura & E. Kellerman. 1992. A second language classic reconsidered – the case of Schachter's avoidance. *Second Language Research* 8: 251-277.
- Krzeszowski, T. 1990. *Contrasting languages. The scope of contrastive linguistics*. Berlin/New York: Mouton de Gruyter.
- McCarthy, M. & R. Carter. 1994. *Language as discourse. Perspectives for language teaching*. London: Longman.
- × Odlin, T. 1989. *Language transfer — Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Pery-Woodley, M.P. 1990. Contrasting discourses: contrastive analysis and a discourse approach to writing. *Language Teaching* 23: 143-151.
- Schachter, J. 1974. An error in error analysis. *Language Learning* 24: 205-214.
- Selinker, L. 1989. CA/EA/IL: The earliest experimental record. *IRAL* 27: 267-291.
- Selinker, L. 1992. *Rediscovering interlanguage*. London/New York: Longman.
- Tisthoud, N. 1992. *Conceptual and lexical gaps: a problem of translation*. Unpublished MA Dissertation, Université Catholique de Louvain.
- Van Hoof, H. 1989. *Traduire l'anglais. Théorie et pratique*. Paris/Louvain-la-Neuve: Duculot.
- Vinay, J.P. & J. Darbelnet. 1969. *Stylistique comparée du français et de l'anglais*. Paris: Didier.
- Wardhaugh, R. 1970. The contrastive analysis hypothesis. *TESOL Quarterly* 4: 123-130.
- × Wollin, L. & H. Lindquist (eds). 1986. *Translation studies in Scandinavia*. Lund Studies in English 75. Lund: CWK Gleerup.