

Combining contrastive and interlanguage analysis to apprehend transfer: detection, explanation, evaluation

Gaëtanelle Gilquin

Centre for English Corpus Linguistics
F.N.R.S. – Université catholique de Louvain

Abstract

This paper shows how contrastive analysis and interlanguage analysis can be combined into a model which aims to detect phenomena of transfer, explain them and evaluate them in terms of pedagogical relevance. The model, referred to as the Detection-Explanation-Evaluation (DEE) transfer model, relies on Granger's (1996) Integrated Contrastive Model and Jarvis's (2000) unified framework for transfer research, and uses corpus data as a basis for a range of comparisons between L₁ (mother tongue), L₂ (target language) and IL (interlanguage). As an illustration, the use of even if by French-speaking learners is investigated, and the possible influence of L₁ is examined in the light of the different comparisons involved in the model. Finally, some limitations of the DEE transfer model are discussed, all of which point to possible developments that could be implemented in the future.

1. Introduction

The importance of transfer in second language acquisition can now be regarded as an established fact. As observed by Kohn (1986: 21) two decades ago already, “there is no doubt that, despite its sometimes irritatingly elusive character, transfer is one of the major factors shaping the learner’s interlanguage competence and performance”. Two main sources of data have been used over the years to investigate the influence of the mother tongue (L₁) on the interlanguage (IL), namely contrastive analysis (CA) data and learner data. Selinker (1989: 287) notes that “CA is the best place to begin language transfer studies since structural congruence (or at the least, partial structural similarity) is most probably necessary, though not sufficient, for most types of language transfer to occur”. Indeed, Schachter’s (1974) analysis, which investigated the influence of L₁ on the use of English restrictive relative clauses by learners from several mother tongue backgrounds, was demonstrated to be flawed, for lack of reliable CA data (Zhao 1989, Kamimoto et al. 1992). Without samples of learner production to rely on, however, studies of transfer, as faultless as they may be from a contrastive point of view, could offer nothing but useless predictions. Both contrastive analysis and interlanguage analysis are therefore necessary for a sound and systematic assessment of the role of transfer in second language acquisition.¹

In this paper, two models of transfer which combine a CA and an IL approach, Granger (1996) and Jarvis (2000), are compared in an attempt to develop a refined method to apprehend transfer. Three varieties of language will be considered, viz. L_1 , L_2 (target language) and IL, and will lead to a number of comparisons which aim to detect transfer (when is it at work?), explain it (what are its origins?) and evaluate it from a pedagogical perspective (is it pedagogically relevant?). Reflecting this threefold purpose, the model will be referred to as the Detection-Explanation-Evaluation (DEE) transfer model. Because the interest here is in phenomena of transfer as they manifest themselves in naturally-occurring language, use will be made of corpora, whose value for linguistic descriptions has been underlined by numerous studies over the last few years.

The paper is organised as follows. First, the phenomenon of transfer is briefly described and recontextualised, with special emphasis on its links with contrastive analysis and interlanguage analysis. Granger's (1996) Integrated Contrastive Model and Jarvis's (2000) unified framework for the study of L_1 influence are then presented and compared. On the basis of this comparison, a new model of transfer is proposed, the DEE transfer model, which serves the threefold purpose of detecting, explaining and evaluating cases of transfer. The model is illustrated by means of a case study of the use of *even if* by French-speaking learners, after which some of its limitations are outlined. The article ends with some concluding remarks.

2. The transfer mystery

After a period in the 1950s and 60s when transfer was seen as all-pervasive and a period in the 1970s when its existence was denied, we have now reached a time when it is "generally accepted that transfer does occur, but is a far more complex phenomenon than hitherto believed" (Benson 2002: 68). Its exact nature, the circumstances in which it occurs or the processes it relies on in the learner's mind are still, to a large extent, a mystery to linguists. As Dechert & Raupach (1989: ix) put it in their introduction, "[i]n spite of three decades of intensive research [...] there is still no generally accepted agreement of what transfer in language actually is". Lado (1957: 2) refers to the transfer of "the forms and meanings and the distribution of forms and meanings of [individuals'] native language and culture to the foreign language and culture". Odlin's (1989: 27) definition is broader, as it covers the influence of "any other language that has been previously (and perhaps imperfectly) acquired". In other words, transfer is not limited to the influence of L_1 on L_2 , but could involve the influence of, say, L_2 on L_3 or L_3 on L_5 . Neither of these definitions, however, explicitly raises the possibility of a simultaneous influence of several languages. Nor do they allow for the fact that a language, most typically the native language, may be influenced by subsequently acquired languages. In order to take such phenomena into account, transfer will be defined here in very general terms, as the influence, within an individual's

linguistic system, of one or more languages over another.² Moreover, we will consider both cases of negative transfer, when the influence leads to erroneous usage,³ and positive transfer, when it leads to correct usage.

The notion of transfer is closely related to contrastive analysis. In fact, as pointed out by Ringbom (1995: 581), the original idea behind contrastive analysis was to bring to light the problems that learners from a specific mother tongue background might experience in learning a given foreign language. However, the study of learner language (so-called “interlanguage”, see Selinker 1972), initially investigated within the frame of error analysis, revealed that problems could occur in cases not predicted by contrastive analysis and, conversely, that not all areas predicted as difficult by contrastive analysis were actually problematic for learners (Ringbom 1995). Thus, while both contrastive analysis and interlanguage analysis appeared necessary to make discoveries about transfer, it also turned out that the link between the two was far from obvious and required further research.

The model of transfer presented here, the DEE transfer model, relies on two models combining contrastive and interlanguage analysis, namely Granger (1996) and Jarvis (2000), and aims to provide a systematic method for detecting transfer, explaining it and evaluating its pedagogical relevance. It involves three varieties of language (L_1 , L_2 and IL) and six types of comparison,⁴ as summarised in Table 1, which also shows the abbreviations that will be used throughout the article. Of these six comparisons, the first three are instances of contrastive analysis, whereas the last three involve so-called contrastive interlanguage analysis.⁵ Each of them serves a particular (primary) purpose, namely detection for NL_1/IL and $IL_a/IL_b \dots IL_z$ (see Section 6.1), explanation for OL_1/OL_2 and $SL_{1/2}/TL_{2/1}$ (see Section 6.2) and evaluation for NL_2/IL and IL_a/IL_a (see Section 6.3).

Table 1. Types of comparison and their abbreviations

Abbreviation	Explanation	Example for French learner of English
OL_1/OL_2	Comparison of original L_1 and original L_2	Original French vs. original English
$SL_{1/2}/TL_{2/1}$	Comparison of source L_1 and translated L_2 or vice versa	Source French vs. translated English or source English vs. translated French
NL_1/IL	Comparison of native L_1 and IL	Native French vs. learner English
NL_2/IL	Comparison of native L_2 and IL	Native English vs. learner English
IL_a/IL_a	Comparison of interlanguage by learners from the same L_1	French learner English vs. French learner English
$IL_a/IL_b \dots IL_z$	Comparison of interlanguage by learners from different L_1 s	French learner English vs. Spanish learner English vs. German learner English, etc

Although the model is mainly illustrated through cases of transfer from L₁ French to L₂ English, and most of the examples concern lexical (and some syntactic) phenomena, it is possible to apply the DEE transfer model to other languages, other types of transfer – e.g. from L₂ to L₃ (Hammarberg 2001: 23) or from L₂ to L₁ (Pavlenko & Jarvis 2002) – and other phenomena (phonological, pragmatic, etc).⁶ And while all the illustrations come from written corpus data, the model could also serve to study transfer in speech, provided the necessary material is available.

3. Granger's (1996) Integrated Contrastive Model

In Granger's (1996) Integrated Contrastive Model (ICM), two components are combined in an attempt to shed light on the phenomenon of transfer. The contrastive analysis (CA) component compares (i) original data from one language with original data from another language and (ii) source language and its translation in another language. The contrastive interlanguage analysis (CIA) component, similarly, consists of two types of comparison: (i) native language and a non-native variety of this language and (ii) different non-native varieties of one and the same language.

The peculiarity of the model is that it exclusively relies on authentic data coming from computerised corpora. This has not always been the case in contrastive and interlanguage studies. The first comparisons of two or more languages were often intuition-based and the traditional error analyses of learner language usually relied on very small collections of texts.⁷ Recently, however, with the computer revolution, large amounts of machine-readable authentic data have become available. Parallel corpora (that is, corpora of original texts and their translations into one or several languages) and comparable corpora (that is, corpora of original texts in two or more languages which are similar in terms of domain, genre, etc) have made it possible to reassess a number of contrastive claims. Tisthoud (1992), for instance, shows on the basis of bilingual corpus data that the morphological flexibility of English as compared to French has been largely overestimated, with a number of French derivations being translated by means of a paraphrase in English (e.g. *clochettes* = *little bells*, *improductivité* = *lack of productivity*). Large and balanced corpora of interlanguage have also been compiled over the last few years. One of the first corpora of this type was the *International Corpus of Learner English* (ICLE, see Granger et al. 2002), the latest version of which includes essays written by advanced learners of English from 16 different mother tongue backgrounds (Granger et al. forthcoming). Not only have such corpora allowed for the contextualisation of errors, but they have also enabled researchers to investigate what learners get right, what they overuse, i.e. use significantly more than native speakers, and what they underuse, i.e. use significantly less than native speakers. Corpora, therefore, have contributed to the

development of both CA and CIA, thus making the model resulting from their integration a powerful tool, as will be shown presently.

As appears from Figure 1, which represents the ICM, the notion of transfer occupies a central position, linking CA and CIA predictively and diagnostically. The predictive hypothesis (also called “CA a priori” or “strong CA hypothesis”) starts from the comparison of two languages (CA), either as original languages (OL vs. OL) or as source language and translated language (SL vs. TL), and on this basis makes predictions about the learner’s interlanguage. The assumption is that in the case of discrepancies between the learner’s mother tongue and the target language, the learner is likely to transfer the L_1 pattern to his/her interlanguage, hence producing an erroneous L_2 pattern (negative transfer). In the case of similarities between L_1 and L_2 , on the other hand, the learner is expected to produce a correct pattern in L_2 (positive transfer). By examining learner production data (CIA) and looking for traces of L_1 , one can test the accuracy of the predictions and thus establish the (potential) presence, or otherwise, of transfer. The diagram in Figure 1, by using a broken line to connect the CA and the CIA boxes, emphasises the hypothetical nature of the CA-based predictions: L_1/L_2 mismatches do not always lead to errors⁸ in interlanguage and, conversely, L_1/L_2 identity does not necessarily imply error-free use by learners (Gilquin 2000/2001: 101). The diagnostic hypothesis (also called “CA a posteriori” or “weak CA hypothesis”) follows the opposite path, starting from CIA and moving on to CA. By comparing the learner data with a reference native corpus (NL vs. IL) and with data produced by learners from other mother tongue backgrounds (IL vs. IL), one may notice L_1 -specific errors and look to contrastive analysis for an explanation. The diverging arrows in the diagram indicate that the explanation for an error will not always be found in the relation between the learner’s mother tongue and the target language. The errors observed may be due to other factors than L_1 influence, e.g. intralingual factors, developmental factors, teaching-induced factors or communication strategies in interlanguage (see James 1980: 146).

As pointed out by Granger (1996: 46), the ICM “involves constant to-ing and fro-ing between CA and CIA”. This can be illustrated by means of the analysis of the use of the passive by French-speaking learners of English. A quantitative study of the passive in native English and French-speaking learners’ English [CIA: NL vs. IL] brings out a significant underuse of the passive by the learners (Granger 1997). In order to test the role of L_1 in this underuse, one may compare the frequency of the passive in native French and native English. A comparable corpus of newspaper articles [CA: OL vs. OL] reveals that the passive is about twice as frequent in English as in French (Fivet 1995, Granger 1998), which seems to suggest that the French learners’ underuse is transfer-related. Before drawing any such conclusion, however, it is safe to consider other varieties of learner English. From this examination [CIA: IL vs. IL], it turns out that the underuse of the passive is not limited to French-speaking learners, but also concerns learners from other L_1 backgrounds (Granger 1997 highlights this tendency for Swedish- and Finnish-speaking learners as well), which appears to

challenge the transfer-related explanation. One last verification is necessary, however, namely checking whether the other native languages do not exhibit a lower frequency of the passive than in English, as French does. Using an English/French and an English/Swedish parallel corpus [CA: SL vs. TL], Granger (1998) shows that the proportion of English passives that are translated by means of a passive is higher in Swedish than in French. The frequency of the passive in Swedish is close to that of the passive in English, so that Swedish learners' underuse of the passive cannot be attributed to L_1 influence. The underuse of the passive exhibited by French-speaking learners of English therefore seems to be a universal feature of interlanguage, which may be related to learners' preference for the unmarked rather than marked option (Ellis 1994), or for a personal rather than impersonal style (see e.g. Petch-Tyson 1998). The influence of L_1 French, however, cannot be excluded, and it is probable that transfer reinforces the effect of other factors.

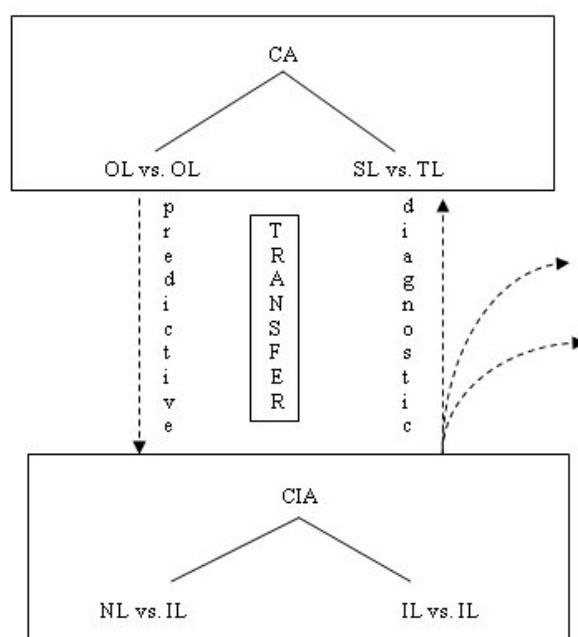


Figure 1. Integrated Contrastive Model (Gilquin 2000/2001: 100, based on Granger 1996: 47)

4. Jarvis's (2000) unified framework for transfer research

In his (2000) article, Jarvis, noting the amount of confusion that reigns in the field of transfer research, offers a possible solution in the form of a unified framework consisting of three components, viz. a theory-neutral definition of transfer, a list of the types of evidence to be considered and a set of outside variables to be controlled. Jarvis's proposed definition will not be discussed here (although it is important to stress that the types of evidence he advocates are derived from this definition), but the other two components will be examined, with particular emphasis on the evidence that should be produced to establish the existence of transfer-related phenomena with some confidence.

Jarvis recognises that the essence of transfer "lies beyond the reach of the researcher" (ibid. 253-254), but he proposes a list of three effects which are subject to direct investigation and which should be considered before one is justified in positing the existence (or at least likelihood) of L_1 influence. These three effects are:

- (i) intra- L_1 -group homogeneity in learners' IL performance;
- (ii) inter- L_1 -group heterogeneity in learners' IL performance;
- (iii) intra- L_1 -group congruity between learners' L_1 and IL performance.

The first effect of L_1 influence, the intra- L_1 -group homogeneity in learners' IL performance, refers to the fact that learners who share the same mother tongue should exhibit a similar behaviour when using the target language. In other words, a particular feature should not be the idiosyncrasy of one or two individuals, but should characterise the group as a whole. The second effect, the inter- L_1 -group heterogeneity in learners' IL performance, means that learners coming from different mother tongue backgrounds should perform differently in the target language. The behaviour of Italian learners of English, for example, should be specific to this learner population, and distinct from the behaviour displayed by, say, Finnish or Chinese learners of English. The third effect, the intra- L_1 -group congruity between learners' L_1 and IL performance, points to a parallelism between the learner's mother tongue and his/her interlanguage. If the interlanguage presents a particular feature, there should be a corresponding feature in L_1 .

While each of these effects has the potential to reveal the existence of L_1 influence, Jarvis demonstrates that none of them is necessary nor sufficient in itself, as transfer may take place despite the absence of one of the effects and, conversely, an effect may be present but not transfer-related. To give but one example, several L_1 s may share a particular feature (especially if they belong to the same language family) and therefore produce the same effect on IL. In such a case, inter- L_1 -group homogeneity would not exclude the possibility of transfer. This leads Jarvis to argue that at least two effects out of the three he mentions should be identified before concluding that transfer is at play.

In addition, Jarvis claims that any rigorous investigation of transfer should control for a number of outside variables, which may either promote or inhibit L_1 influence. On the basis of the literature, he draws the following list of variables (Jarvis 2000: 260-261):

- a) age;
- b) personality, motivation and language aptitude;
- c) social, educational and cultural background;
- d) language background (all previous L_1 s and L_2 s);
- e) type and amount of target language exposure;
- f) target language proficiency;
- g) language distance between the L_1 and target language;
- h) task type and area of language use;
- i) prototypicality and markedness of the linguistic feature.

Jarvis illustrates his methodological framework by means of a study of the referential lexical choices made in English by Finnish-speaking and Swedish-speaking Finns. Using three different elicitation tasks (film retell, lexical listing and selection of lexical items), he investigates how a number of objects and actions are referred to by the participants, a carefully selected subject pool consisting of several learner groups at various levels of age, L_2 exposure and L_2 proficiency, and their L_1 -control group counterparts. All the variables listed by Jarvis (see above) were controlled for, except the second one (personality, motivation and language aptitude). The analysis of the results shows that, generally speaking, the learners' lexical choices exhibit the three effects identified by Jarvis as typical of L_1 influence – although their presence may at times be quite subtle. First, it turns out that, despite differences in age and L_2 exposure, learners with the same mother tongue background tend to make similar lexical choices. Second, while the comparison of the lexical choices made by learners with distinct mother tongues does not reveal any significant differences, learners from different L_1 backgrounds and matched according to age and L_2 exposure still exhibit higher levels of heterogeneity than learners from the same L_1 background and with differences in age and L_2 exposure. Finally, although the comparison of learners' interlanguage and mother tongue provides only partial support for the third effect, the other two types of comparison (first and second effects) show evidence of L_1 -IL similarities. Jarvis also demonstrates that, while some outside variables appear to have an influence on the results (e.g. age, L_2 exposure, task type), none of them produces as consistent effects as the L_1 variable. His study therefore underlines the importance of the mother tongue background in the interlanguage lexicon of Finnish- and Swedish-speaking learners of English. More generally, it also demonstrates the usefulness of a rigorous methodological framework to investigate the role of transfer in interlanguage studies.

While Jarvis uses elicitation data to illustrate his model, one could envisage using authentic corpus data instead, in view of the advantages of corpora

outlined in Section 3. Paquot (2006) adopts a corpus-based approach to Jarvis's framework, testing the role of L₁ influence on French-speaking learners' overuse of the expression *on the contrary*. Her study relies on learner data from several ICLE components representing different mother tongue backgrounds, as well as comparable native data produced by French- and English-speaking students (CODIF and LOCNESS, see later). She tests the first and second effects (intra-L₁-group homogeneity and inter-L₁-group heterogeneity in learners' IL performance) by means of a simple analysis of variance (ANOVA), which indicates that the difference in frequency between the ICLE components is more significant than the difference between the distinct essays within each component. Computing pairwise t-tests with Bonferroni adjustment, she is then able to show that it is the French data which are mainly responsible for the statistically significant difference between the ICLE components. In other words, it turns out from the statistical analysis that the variability between the essays written by learners with different mother tongues is significant (inter-L₁-group heterogeneity) and is greater than the variability within the essays written by the French-speaking learners (intra-L₁-group homogeneity). In order to investigate the third effect, the intra-L₁-group congruity between learners' L₁ and IL performance, Paquot compares the relative frequency of *on the contrary* in French-speaking learners' interlanguage with its relative frequency in a corpus of native English on the one hand, and with the relative frequency of the French equivalent, *au contraire*, in a corpus of native French on the other hand. It appears that the results of the learner corpus are more similar to the results of the native French corpus than to those of the native English corpus, which suggests that French-speaking learners tend to transfer the frequency of *au contraire* when producing *on the contrary*. The presence of Jarvis's three effects leads Paquot to conclude that the use of *on the contrary* by French-speaking learners is (at least partly) L₁-induced.⁹

5. Comparison of Granger's (1996) ICM and Jarvis's (2000) framework

Granger's (1996) and Jarvis's (2000) models share a number of important features. The most obvious one is that they are both concerned with the phenomenon of transfer. In addition, although Jarvis does not use these specific terms, the two models approach transfer by combining contrastive analysis (comparison of two or more languages) and contrastive interlanguage analysis (comparison of two or more **varieties** of a language, including non-native language). While the specific types of comparison within CA and CIA are essentially different, as we will see below, there is one type which is shared by the two models, namely the comparison of several non-native varieties of the same language (IL_a/IL_b...IL_z, which corresponds to Jarvis's second effect, inter-L₁-group heterogeneity in learners' IL performance). Finally, Granger and Jarvis both acknowledge the existence of a number of outside variables that should be rigorously controlled. While these variables are not part of the ICM as such, Granger (1996: 44) underlines their importance, with particular emphasis on the

type of target language exposure (foreign or second language), the target language proficiency and the task type. In fact, the learner corpus she uses to test the ICM, the *International Corpus of Learner English*, has been carefully designed so as to keep these three variables constant (foreign language rather than second language, advanced proficiency level¹⁰ and essay writing) and has been encoded with additional variables which can easily be manipulated when selecting the data, e.g. amount of L₂ exposure, knowledge of other foreign languages, gender, task setting, etc (see Granger 2003).

Next to these similarities, the two models also present differences. Despite their common research object (transfer), the models have slightly different goals. The ICM, as we saw, has both a predictive and a diagnostic goal, whereas Jarvis's model simply aims to identify instances of transfer. Second, Granger and Jarvis use different data – corpus data and experimental data, respectively – although, as illustrated by Paquot's (2006) study, Jarvis's model can be operationalised by means of corpus data as well. The models also differ with respect to the types of CA and CIA analyses they advocate. For Granger, the contrastive analysis should consist in a comparison of the target language and the learner's mother tongue (either as original languages in a comparable corpus or as source and translated languages in a parallel corpus). Jarvis, on the other hand, proposes the comparison of the learner's mother tongue with his/her interlanguage (see endnote 5 on why this is considered a case of CA). In terms of contrastive interlanguage analysis, besides the common IL_a/IL_b...IL_z comparison, Granger includes as an individual (and essential) component of her model a comparison of the interlanguage and its native counterpart. Jarvis uses this type of comparison too, but only as part of the larger IL-IL comparative analysis, treating NL₂ as if it were another IL. Moreover, he regards this type of evidence as unreliable (Jarvis 2000: 300). Finally, Jarvis, unlike Granger, includes in his model a comparison of the interlanguage of several learners from the same mother tongue background.

The similarities and differences between the two models are summarised in Table 2. It will appear from the next section that all these elements have a role to play in a model of transfer, but that they serve different purposes.

Table 2. Comparison of the features of Granger's ICM and Jarvis's framework

Granger's (1996) ICM	Jarvis's (2000) framework
	Object of study = transfer
	Combination of CA and CIA
	IL _a /IL _b ...IL _z
	Control of outside variables
Predictive and diagnostic	Identification
Corpus data	Experimental data (but applicable to corpus data)
L ₁ vs. L ₂ (OL ₁ /OL ₂ or SL _{1/2} /TL _{2/1})	NL ₁ vs. IL
NL ₂ vs. IL as an individual component	IL _a vs. IL _a

6. Towards a comprehensive model of transfer: the DEE transfer model

The advantage of using authentic corpus data has been briefly demonstrated in Section 3, both for CA and CIA (see also Ellis & Barkhuizen 2005 on the value of naturally-occurring samples of learner language for investigating second language acquisition). The importance of considering outside variables has also been underlined by Granger and Jarvis, as well as other authors (e.g. Ellis 1994). This section will therefore focus on the six types of comparison borrowed from Granger's and Jarvis's models (see Table 1), and deal with them according to the primary purpose they serve, namely detection, explanation or evaluation.¹¹ The resulting model, the DEE transfer model, does not pretend to be a theory of transfer, but it proposes a methodology which seeks to apprehend transfer in a way that takes several of its facets into account, thus striving for a comprehensive view of the phenomenon.

6.1 Detection of transfer

Two types of comparison serve the primary purpose of detecting the (potential) presence of transfer, viz. NL_1 vs. IL and IL_a vs. IL_b (or more precisely IL_a vs. IL_b vs. ... vs. IL_z). Of these, the first one is an instance of CA and the second one, an instance of CIA, so that the detection of transfer relies on the combination of CA and CIA, as is also the case in Granger (1996) and Jarvis (2000). While $IL_a/IL_b...IL_z$ is common to both models, NL_1/IL is only found in Jarvis's model.

The comparison of the learner's interlanguage with his/her mother tongue (NL_1 vs. IL) is essential to detect transfer. Only if one can establish a similarity between the learner's behaviour in IL and NL_1 does the presence of transfer seem plausible enough.¹² The operationalisation of the NL_1/IL comparison may involve frequency, as proposed by Paquot (forthcoming), with the requirement that the frequency of the NL_1 item be as close as possible to that of the IL item (cf. frequency of *au contraire* in native French and of *on the contrary* in learner English in Paquot's [2006] study). The similarity may also be formal, in which case the method of reversed translation, which translates interlanguage back into the learner's mother tongue, may be helpful.¹³ Thus, disregarding frequency, it is clear that *in fact* may be considered as the formal equivalent of French *en fait* (*en* corresponds to the preposition *in* and *fait* is the French noun for *fact*; both expressions have the same function of adding information to emphasise what has just been said). Similarly, from a formal point of view, *on the contrary* would be the nearest equivalent of French *au contraire*, with *contraire* being related to *contrary* (more, for example, than to *contrast* – cf. *by contrast* – which has another equivalent in French, *contraste*). Using the method of reversed translation, Borgatti (2006) discovered that 57% of the occurrences of the verb *make* in the interlanguage of French-speaking advanced learners could be translated into French by means of the equivalent verb *faire* (e.g. *make sacrifices* = *faire des sacrifices*).¹⁴ Interestingly, he also found out that 83% of all the erroneous uses of *make* corresponded to French *faire* (e.g. **make a description*,

translatable into French by *faire une description*), which strongly suggests that these errors are due to negative transfer. It should be noted that the method of reversed translation involves authentic learner data and that contrastive corpus data may be helpful in identifying translational equivalence (see below). However, the comparison itself does not rely on corpus data, but on the researcher's capacity to translate learners' interlanguage back into their mother tongue. In that, it stands as an exception to the generally corpus-based approach of the DEE transfer model.

NL₁/IL congruity is not enough to establish transfer with some degree of confidence. This is because a learner's behaviour may be similar to his/her mother tongue, but also similar to the behaviour of learners from other mother tongue backgrounds, in which case the most likely explanation is not interlingual (i.e. due to L₁ influence), but intralingual (i.e. due to the inherent difficulties of learning L₂).¹⁵ The examination of the relationship between the learner's interlanguage and the interlanguage of learners from other mother tongue backgrounds (IL_a/IL_b...IL_z) should therefore come as a complement to the NL₁/IL comparison. On the basis of a simple NL₁/IL comparison, one may for instance come to the conclusion that the overuse of *more and more* by French-speaking learners is due to transfer, given that the frequency of this expression in their interlanguage is very similar to the frequency of its French counterpart, *de plus en plus*, in a comparable corpus of native French (CODIF, see Section 6.4). Examining data produced by learners from other mother tongue backgrounds, however, reveals that the overuse of *more and more* is a general characteristic of learner English, since out of the 16 L₁ backgrounds represented in ICLE 2 (Granger et al. forthcoming), 14 present such an overuse (statistically significant at the 0.001 level). While transfer cannot be totally excluded in the case of French-speaking learners, it seems as if intralingual factors prevail here. Since both Granger and Jarvis agree on the necessity of this type of comparison, we will not discuss it any longer.

One may wonder why the comparison of the two languages involved has been operationalised through an NL₁/IL analysis, rather than an OL₁/OL₂ or SL_{1/2}/TL_{2/1} analysis as is the case in the Integrated Contrastive Model. In contrast with the latter two types of comparison, the NL₁/IL analysis makes it possible to approach the learner's interlanguage as an independent linguistic system. Comparing the two native languages in their original (or source and translated) forms and directly applying these results to the question of transfer (i.e. the relationship between the learner's interlanguage and his/her mother tongue) would imply equating the interlanguage and the target language. However, several second language acquisition specialists have guarded against such an equation, however, underlining the dangers of what is called the "comparative fallacy", i.e. "the failure [...] to take the structure of the interlanguage on its own terms and a reliance on a comparison of the target language in order to study the structure of the interlanguage" (Bley-Vroman 1983: 3-4). It will be argued later that the criticism levelled by some against the native norm in interlanguage study is not necessarily justified, but it is still true that the learner's interlanguage, not

the target language, should be taken as a baseline for the comparison with NL₁, since equivalence (or lack of equivalence) between NL₁ and NL₂ does not necessarily imply a similar relationship between NL₁ and IL. Thus, there may be no correspondence in a parallel corpus between English *sit on the bus* and French *être assis sur le bus*, the latter describing someone sitting on the roof of a bus, but reversed translation would highlight the equivalence between the two expressions, since *sur le bus* is a literal translation of *on the bus*. It would also suggest that an English-speaking learner of French using the preposition *sur* in such cases is probably acting under the influence of English. The inappropriateness of bilingual corpora to deal with the relationship between the learner's interlanguage and his/her mother tongue is also particularly evident in the case of non-target forms in the learner's IL. The French component of ICLE (ICLE-FR) contains the expression *make abstraction of*, standing for *disregard*. An OL₁/OL₂ or SL_{1/2}/TL_{2/1} comparison would fail to establish a correspondence between these two items, since *make abstraction of* does not exist in the NL₂ system. Using the method of reversed translation, however, one may draw a parallel between *make abstraction of* and French *faire abstraction de*, thus showing that transfer is a likely explanation for the IL behaviour.

What precedes should not be taken to mean that OL₁/OL₂ or SL_{1/2}/TL_{2/1} comparisons have no role to play in the detection of transfer, or that NL₂ is completely out of the picture because of the comparative fallacy. Comparable or parallel corpora may be necessary before the NL₁/IL comparison proper, to identify the best candidate for translational equivalence. In the case of cognates (e.g. *fait/fact*, *description/description*), the equivalence is easily established. In some other cases, however, it may be more difficult, especially for non-native speakers of the learner's L₁, to determine the most probable equivalent. It is here that bilingual corpora may help, by showing in context what words fulfil the same function and express the same meaning in the learner's mother tongue and in the target language.¹⁶ As Danchev (1991: 89) puts it, translation obviates "the familiar difficulty of establishing comparability (sic)". (See, however, Section 7 for an important caveat.)

As for NL₂, it is often necessary as a baseline for comparing the frequencies of an item in IL and NL₁. The difference in frequency between *nevertheless* in French learners' IL and the equivalent *néanmoins* in a comparable corpus of native French may at first sight appear too great to be the result of transfer, with a relative frequency of 327 per million words for the former and 206 per million words for the latter. However, compared with the frequency of *nevertheless* in native English, which amounts to a mere 9 occurrences per million words, this difference between IL and NL₁ becomes insignificant, thus making the transfer hypothesis more likely.

Despite their possible participation in the detection of transfer, OL₁/OL₂, SL_{1/2}/TL_{2/1} and NL₂/IL comparisons have another, primary role to play in the DEE transfer model. The first two types of comparison make it possible to explain the presence, or otherwise, of L₁ influence, whereas the last type of comparison is especially helpful in a more applied context, when evaluating

whether an observed phenomenon of transfer should have consequences in foreign language teaching. These two roles, explanation and evaluation, are described in the next two sections.

6.2 Explanation of transfer

The existence of a formal equivalent in the other language system is not necessarily sufficient to lead to transfer. Thus, although both Swedish and French have structures corresponding to causative *make* + infinitive (with *få* in Swedish and *faire* in French), transfer seems to be at work only in the case of Swedish learners (compare Altenberg 2002 and Gilquin 2000/2001). This difference in IL behaviour may be explained by resorting to the concept of “language distance”, that is “the degree of similarity between two languages” (Odlin 1989: 32). As noted by Ellis (1994: 327), “[t]here is substantial evidence to indicate that the actual distance between the native and the target languages acts as a constraint on transfer.” Equivalents which are identical in all respects between the two language systems are therefore more likely to be transferred than equivalents which present differences.¹⁷

The degree of cross-linguistic equivalence can be established on the basis of parallel corpus data (SL_{1/2} vs. TL_{2/1}). A useful measure here is that of “mutual translatability” (or “mutual correspondence”), which refers to the frequency with which two items are translated as one another (on this concept, see Altenberg 1999). A value of 0% would indicate that the two items are never translated as one another, and a value of 100%, that the two items are always translated as one another. If we apply this measure to causative constructions in parallel corpus data of English and Swedish on the one hand (Altenberg 2002), and English and French on the other (Gilquin 2000/2001), a noticeable difference emerges. While causative *make* and its Swedish equivalent *få* exhibit a mutual translatability of 52%, in the case of causative *make* and its French equivalent *faire*, the value is under 14%.¹⁸ Considering all the English equivalents of causative *faire* in a sample of the parallel corpus PLECI¹⁹ (from source texts to translations and vice versa), for example, it turns out that a synthetic causative verb is the preferred option, with a total percentage of 56.7% (Gilquin 2000/2001: 105). This is illustrated by (1), where *faire disparaître* (‘make disappear’) corresponds to *dispel*.

- (1) Aussitôt qu’elle avait établi son bon droit, elle s’efforçait de *faire disparaître* toute trace de ressentiment chez son adversaire.
 = Her own way once insisted upon and secured she was swift to *dispel* any smallest hint of unpleasantness.
 (PLECI fiction TF-OE)

The degree of cross-linguistic equivalence between two items can also be established by means of a comparable corpus (OL₁ vs. OL₂), preferably as a supplement to the parallel corpus analysis. While this approach does not make it possible to identify the word or construction corresponding to the target item in a

specific context, since the texts are not translations of each other, it has the advantage of not being affected by any possible translation effects (“translationese”, see Gellerstam 1986). By carefully comparing the factors conditioning the use of two supposedly equivalent items in authentic original data, one may gain insight into the extent to which the two items overlap in terms of, say, lexical preferences, pragmatic functions or syntactic patterns. This, in turn, may help determine the “distance” between the two items and hence the likelihood of transfer. In the case of English and French causative constructions, the analysis of the comparable corpus data confirms the low degree of equivalence brought to light by the parallel corpus analysis (see Gilquin forthcoming). To give but one example, a sample of original texts from the PLECI corpus shows that causative constructions with *faire* are quite frequently causeless, i.e. with no mention of the entity changed or influenced by the subject (this represents over one quarter of all the occurrences of causative *faire*), cf. (2). In English, by contrast, causeless constructions with *make* are limited to the idiomatic expressions *make believe* and *make do*, which, incidentally, are not found in the PLECI corpus.

- (2) Fais vérifier ses papiers.
 ‘Make check his papers’
 Make someone check his papers/Have his papers checked.
 (PLECI fiction OF-TE)

In sum, OL_1/OL_2 and $SL_{1/2}/TL_{2/1}$ comparisons, by highlighting the degree of similarity between two items cross-linguistically, may provide an explanation for the presence or absence of transfer among learners. Two remarks are in order, here. The first one is that, although in what precedes the bilingual corpus data have been presented as a way of **explaining** transfer, thus corresponding to the diagnostic hypothesis of Granger’s model, they may also be used with a predictive goal, as a way of **predicting** where transfer is likely to occur or not. In other words, the OL_1/OL_2 and $SL_{1/2}/TL_{2/1}$ comparisons may precede or follow the detection of transfer, or indeed, as advocated by Granger, occur at different stages in the analysis. The second remark is that, in the same way as distance between two languages is a matter of degree (cf. degree of mutual translatability or degree of overlap between the conditioning factors of two items), the likelihood of transfer may be relatively high or low. One cannot determine a threshold below which transfer becomes impossible, so that both the explanatory (or diagnostic) power and the predictive power of the DEE transfer model are hypothetical, as in the Integrated Contrastive Model.

6.3 Evaluation of transfer

The two remaining types of comparison found in Granger (1996) and Jarvis (2000), NL_2/IL and IL_a/IL_a , also have a role to play in a comprehensive model of transfer. However, while Granger and Jarvis use these comparisons as (partial) evidence for transfer, they will be assumed here to be mainly (though by no

means only) useful for pedagogical purposes, as a way of evaluating whether the presence of transfer should affect foreign language teaching (FLT) or not.²⁰

Once a phenomenon of transfer has been detected with a certain degree of likelihood, a comparison of the interlanguage and the target native language (NL₂) is necessary to distinguish between positive and negative transfer. Consider the expressions *according to me* and *according to X* (X being another entity than the subject) as used by French-speaking learners of English. Applying the criteria established in Section 6.1, one can detect possible traces of transfer in both cases, from French *selon moi* and *selon X* respectively. However, it takes an examination (or at least knowledge) of NL₂ to reveal that *according to me*, unlike *according to X*, is not acceptable in native English. The presence of *according to me* in French-speaking learners' interlanguage may therefore be described as a case of negative transfer, that is, a case where learners produce a non-target form in the interlanguage as a result of negative influence of their mother tongue, whereas the presence of *according to X* could be a case of positive transfer, where learners produce a target form in the interlanguage as a result of positive influence of their mother tongue. Note that the comparison with native data is not only necessary to distinguish between cases of positive and negative transfer of form, but also of frequency.²¹ By transferring the frequency of an item from NL₁ to IL, a learner may "hit the right amount" (positive transfer), or s/he may overuse or underuse the item (negative transfer). The apparent positive transfer of form of *according to X* among French-speaking learners of English is accompanied by what looks like a positive transfer of frequency, with no statistically significant difference in frequency between non-native and native English. Negative transfer, on the other hand, seems to be at work in French-speaking learners' overuse of *on the contrary* (see above).

Given the time constraints that exist in the FLT classroom, it is best to focus on instances of negative transfer (e.g. the use of *according to me* or the overuse of *on the contrary*), showing learners how to avoid it (negative feedback), and let positive transfer (e.g. the use of *according to X*) make its way into the interlanguage. This, of course, implies that one takes native language as a target for foreign language learners, which may sound like a debatable issue in view of the "comparative fallacy" hypothesis (Bley-Vroman 1983) and the movement of English as a Lingua Franca (or English as an International Language, see e.g. Jenkins 2000, Seidlhofer 2001, Mauranen 2003), which both stigmatise the NL norm. However, I would like to argue that the native language still has a proper place (and an important one) in FLT contexts.²² In a survey conducted among some 200 students, Mukherjee (2005) found that 80% of the subjects chose a national variety of native English as the norm they wanted to approximate to (47% for British English, 30.5% for American English and 2.5% for other national varieties). By contrast, less than 17% of the subjects accepted as their target norm the English used by competent non-native speakers in international contexts. Although native English may not be the norm aimed at by all learners (many of the students who took part in Mukherjee's survey were future English teachers; the results might have been different with, say, business

students), it is still a target norm for **some** of them. In addition, especially in the early stages of FLT, students may not know what type of norm they will need in their future interactions (e.g. only comprehensible to non-native speakers, or also comprehensible to native speakers) and teachers may have to address mixed groups of learners, some of them aiming at a native norm and others happy to speak English as competent non-native speakers. Native language, therefore, appears as a reasonable common target in FLT education, and one which is relatively well-defined.²³

Like NL_2/IL comparisons, IL_a/IL_a comparisons will be taken here to be mainly useful for evaluative purposes. The fact that an item is used by only a few learners within an L_1 -group does not perforce exclude the possibility of transfer. In fact, it has been regularly pointed out in the literature on transfer that “the extent to which the L_1 influences L_2 development can vary strongly from individual to individual” (Kellerman 1995: 133).²⁴ Going back to the expression *make abstraction of* (see Section 6.1), there is no doubt that it is the result of transfer from French, although it occurs only once in ICLE-FR. Given its low frequency, however, it is less interesting in an FLT context than an expression such as *according to me*, which is found in ICLE-FR with a relative frequency of almost 10 occurrences per 100,000 words. By establishing the probability of a particular transfer-related phenomenon, IL_a/IL_a comparisons, therefore, provide relevant information to go through the inevitable process of material selection in the FLT curriculum.

The model of transfer proposed here, with its six types of comparison and three main stages, is illustrated in Figure 2. In what follows, the model will be applied to the use of *even if* by French-speaking learners of English.

6.4 Case study: the use of *even if* by French-speaking learners

This section presents a case study using the DEE transfer model. More precisely, we will go through the stages of detection, explanation and evaluation in order to study any phenomenon of transfer that might be attached to French-speaking learners' use of the expression *even if*. The learner corpus data on which the study is based come from ICLE 2 (Granger et al. forthcoming). Like all the data from this corpus (see Section 5), the selected data keep three variables constant, namely type of learner (foreign language learner rather than second language learner), stage of learning (advanced level) and text type (essay writing). A number of additional variables were controlled for when selecting the samples: mother tongue background (French), text type (argumentative essays) and conditions of writing (untimed and with reference tools). Finally, only those samples that contained at least 100,000 words after applying the above criteria were kept in the final analysis.²⁵ Table 3 presents the word counts of the remaining nine components. In addition, comparable corpora of native French and native English were used in the study – 208,419 words from the *Corpus de Dissertations Françaises* (CODIF) and 328,406 words from the *Louvain Corpus of Native English Essays* (LOCNESS) – as well as French-English corpus data from PLECI.²⁶

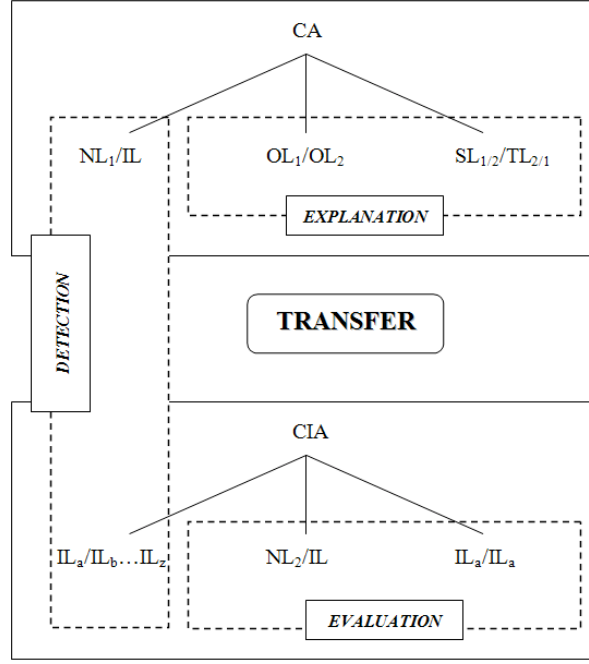


Figure 2. The DEE transfer model

Table 3. Word counts of ICLE samples

ICLE component	Number of words
Czech (ICLE-CZ)	131,435
Dutch (ICLE-DU)	163,908
Finnish (ICLE-FI)	126,599
French (ICLE-FR)	137,431
German (ICLE-GE)	110,890
Polish (ICLE-PO)	141,535
Russian (ICLE-RU)	168,046
Spanish (ICLE-SP)	99,905
Turkish (ICLE-TU)	105,006
TOTAL	1,184,755

Detection of transfer according to the DEE transfer model involves two types of comparison, viz. NL_1/IL and $IL_a/IL_b \dots IL_z$. A reversed translation of *even if* points to *même si* as a formal equivalent, both as a multi-word unit and as individual words (*even* = *même*; *if* = *si*). This equivalence is confirmed by the corpus data. The first type of comparison also reveals a striking similarity in

frequency between *even if* in ICLE-FR and *même si* in the corpus of native French essays, with a relative frequency of 458.41 and 489.40 occurrences per million words respectively.²⁷ The inter-L₁-group comparison (IL_a/IL_b...IL_z), the results of which are displayed in Table 4, shows that, despite the variability among the L₁ groups, French-speaking learners' behaviour differs markedly from that of the other learners. ICLE-FR emerges as the ICLE component which has the highest relative frequency of *even if*, and a Ryan procedure (REGWQ) indicates that French-speaking learners distinguish themselves significantly from all the other learner populations in their use of *even if*.²⁸ These two comparisons suggest that there might be a phenomenon of transfer from French *même si* to French-speaking learners' IL *even if*.²⁹

Table 4. Frequency of *even if* in ICLE samples

ICLE component	Absolute frequency	Relative frequency per million words
ICLE-FR	63	458.41
ICLE-FI	35	276.46
ICLE-GE	26	234.47
ICLE-PO	30	211.96
ICLE-CZ	26	197.82
ICLE-RU	26	154.72
ICLE-TU	12	114.28
ICLE-DU	17	103.72
ICLE-SP	7	70.07

Going on to the next stage, that of explanation, we can use bilingual corpus data to examine the degree of equivalence between English *even if* and French *même si*. Parallel corpus data from PLECI reveal a relatively high degree of correspondence between the two expressions. Out of 61 occurrences of *even if*, 40 correspond to *même si* in French, e.g. (3).

- (3) Dès lors, on comprend pourquoi la solution négociée arrangerait les affaires de tout le monde, la France et les Nations unies devant se targuer d'avoir provisoirement marqué un point, *même si* Washington, inquiet de la triple entente Kinshasa-Tripoli-Khartoum, n'a pas encore dit son dernier mot.
 = Clearly, a negotiated solution would be in everybody's interest. It would enable both France and the UN to score a point, at least temporarily, *even if* Washington, worried by the triple alliance between Kinshasa, Tripoli and Khartoum, may not have said its last word.
 (PLECI non-fiction OF-TE)

However, it also appears from the SL_{1/2}/TL_{2/1} and OL₁/OL₂ comparisons that *even if* and *même si* do not always fulfil the same function. In particular, while *even if*

is exclusively used to introduce a condition, *même si* can also introduce a concession, as in (4), where *même si* corresponds to English *even though*.

- (4) Or, dans les sociétés africaines traditionnelles, on demande toujours l'avis des femmes avant de prendre une décision, *même si* celles-ci ne parlent jamais en public.
 = But in traditional African societies the women are always asked for their opinion before a decision is taken, *even though* they never speak in public.
 (PLECI non-fiction OF-TE)

Although the equivalence between *even if* and *même si* is only partial, the high degree of correspondence between the two expressions and the inclusion of *even if* in the functional field of *même si* suggest that there is little distance between the expressions and thus serve to explain French-speaking learners' readiness to transfer the form, function and frequency of *même si* into their IL (see next paragraph on the transfer of function).

In terms of evaluation, finally, the comparison between NL₂ and IL underlines the negative nature of the transfer, and hence the relevance such material may have in an FLT context. This negative transfer manifests itself in two ways. First, it takes the form of an overuse of *even if* by French-speaking learners, as appears from Table 5 ($\chi^2 = 60.24$, statistically significant at the 0.001 level). Second, it results in a number of incorrect uses of the expression, in contexts where a concession, rather than a condition, is introduced. This type of error is illustrated by (5), where a native speaker would have used *even though* instead of *even if*.³⁰ As noted earlier, *même si* can fulfil a concessive function, which makes the influence of L₁ quite clear in such cases.³¹ These two types of negative transfer stress the importance of drawing French-speaking learners' attention to the native-like use of *even if*.

Table 5. Frequency of *even if* in learner English (ICLE-FR) and native English

Language variety	Absolute frequency	Relative frequency per million words
Learner English (ICLE-FR)	63	458.41
Native English	33	100.49

- (5) The problem is that *even if* such an international principle for equality has been established, some countries still keep on violating it (mostly in countries where religion prevails over international rights). (ICLE-FR)

This point is reinforced by the intra-L₁-group comparison (IL_a/IL_a), which shows that the use of *even if* is not limited to a few learners only. Querying each essay from ICLE-FR separately, one notices that the expression occurs in 22.37% of them (51 essays out of 228). While this proportion may seem relatively low, it

roughly corresponds to the proportion with which frequent words such as *difficult*, *look*, *feel* or *idea* occur, as emphasised by Paquot & Fairon (2006).

This brief case study has illustrated the operationalisation of the DEE transfer model, as described in Sections 6.1 to 6.3. While this model makes it possible to shed light on some important aspects of the phenomenon of transfer, it does not pretend to answer all questions related to the issue. In the next section, we examine some possible limitations of the model.

7. Some limitations of the DEE transfer model

In what follows, three possible limitations of the DEE transfer model are briefly discussed, one having to do with the CA component, the other with the CIA component and the last one with the model in general. The discussion, at the same time, opens up possibilities for further improving the model, in an attempt to achieve an even better understanding of transfer.

A first and, in my view, major limitation of the DEE transfer model as it stands is that, in its investigation of the relation between L_1 and L_2 , it only takes account of the objective distance, as evidenced in parallel or comparable corpora. Since transfer is a phenomenon that takes place first and foremost in the learner's cognitive system, however, what really counts is not the objective distance between two items, but the distance as it is perceived by the learner, that is, the learner's psychotypology (cf. Kellerman 1983).³² If a learner perceives two expressions as similar, this may lead him/her to transfer the characteristics of the L_1 expression to the L_2 expression, no matter whether the two expressions actually correspond to each other in authentic language data. On the other hand, if two expressions are not connected with each other in the learner's mental lexicon, their objective similarity is very unlikely to trigger transfer-related phenomena. While objective distance may be an indication of subjective distance (Gilquin 2000/2001), the ideal would be to probe directly into learners' perceptions. This could be done by means of elicitation tests, asking learners to provide the translational equivalent of a word or expression presented to them. The results of such tests might provide an explanation for some cases of transfer not accounted for by bilingual corpus data. Thus, the use of *indeed* by French-speaking learners presents all the signs of transfer. Its frequency in ICLE-FR is very different from its frequency in all the other ICLE components, but very similar to the frequency of its intuitive equivalent *en effet* in a comparable corpus of native French. It is also significantly overused in ICLE-FR (unlike the other ICLE components, which display an underuse or no significant difference) and appears in a large proportion of essays, which makes it particularly relevant to FLT. However, what seems to be lacking is an explanation for this transfer, since parallel corpus data from PLECI reveal a very low degree of correspondence, with only 4 out of 48 occurrences of *indeed* translated as *en effet*,³³ and hence, a great (objective) distance between the two expressions. A test carried out among 23 French-speaking advanced learners of English, however, shows that *en effet* is perceived

as the direct equivalent of *indeed* by most of them (20 learners).³⁴ The subjective similarity between *indeed* and *en effet*, therefore, explains why transfer is possible, although the expressions are, in reality, quite different. Inaccurate perception may also be at work in the case of so-called false friends. A learner may establish a link between, say, *actually* and *actuellement* (French for ‘at the moment’) because of the formal similarity between the words, but no bilingual corpus, or, indeed, competent “reversed” translator would suggest such a link. Elicitation data, because they offer more direct access to the learner’s mental lexicon, may explain phenomena unaccounted for by the corpus data and would, for this reason, deserve a place in the DEE transfer model.

Turning to the CIA component, it has been argued earlier that, despite recent criticism, a native norm is still useful in the study of transfer. However, what this norm should be is still very much an open question. In the case study of Section 6.4, the English native data were extracted from LOCNESS, a corpus of essays written by American and British students. American English and British English, however, display some differences, as is well known, and it might not be such a good idea to mix the two varieties. If only one variety is chosen, on the other hand, which one should it be? British English, American English or yet another national variety of English? Han (2004: 232) argues that the norm against which the interlanguage is examined should be “what the subjects have been exposed to”, but for foreign language learners, exposure might be limited to English as spoken by a non-native teacher. Some may also argue that a corpus of expert writing is what is needed, not a corpus of novice writing, since learners (or at least, some learners) ultimately aim to become expert writers. The problem is that the choice of the norm may affect the results of the analysis. An example of this is the word *chance*, which has a direct (though partly deceptive) equivalent in French, *chance*. Compared with data from LOCNESS, ICLE-FR reveals no significant difference in frequency ($\chi^2 = 1.51$). Compared with a corpus of expert academic writing, however, the difference becomes significant ($\chi^2 = 22.64$, $p < 0.001$), with the French-speaking learners overusing the word. Consequently, if transfer could be demonstrated to occur with *chance*, a comparison with a corpus of novice writing would suggest that it is unnecessary to draw learners’ attention to the word *chance* (or, at least, to its frequency), whereas a comparison with a corpus of expert writing would speak in favour of raising this issue in the FLT classroom. As long as the question of the norm has not been resolved, therefore, the DEE transfer model may make inappropriate recommendations in terms of FLT (although it remains valid for detecting and explaining transfer).

Another possible problem with the model proposed here is that it may give the impression that, once the influence of L_1 has been highlighted, no other explanation is to be sought to account for the observed IL behaviour. Transfer, however, does not exclude the possibility of other factors acting in conjunction with L_1 influence. Consider the overuse of *let’s/let us* by French-speaking learners, which Paquot & Fairon (2006) have shown to be transfer-related. An examination of the “Essay Writing” section of the French-English bilingual dictionary *Collins Robert* (Atkins et al. 1998) reveals the presence of a large

number of expressions with *let's/let us*, presented as translational equivalents of plural imperatives in French, e.g. *let us now consider...* (for French *considérons maintenant...*), *let us take ... as a starting point* (for French *prenons comme point de départ...*) or *let us not forget that...* (for French *n'oublions pas que...*). This indicates that in the case of the imperative, the influence of the mother tongue may be reinforced by teaching-induced effects, and more precisely by the unsuitability of some pedagogical materials and/or reference tools (Gilquin & Paquot 2008). Ideally, the DEE transfer model should therefore be combined with the analysis of other possible factors affecting the IL behaviour, so as to determine the extent to which transfer is responsible for this behaviour.

8. Conclusion

While the existence and the pervasiveness of transfer in second language acquisition is now widely acknowledged, one must recognise that transfer is still something of a mystery to linguists. In this article, the DEE transfer model has been proposed which, by integrating methods from Granger (1996) and Jarvis (2000), aims to detect cases of transfer, explain them and evaluate their pedagogical relevance, hence making it possible to come to a better understanding of the phenomenon. This model combines contrastive analysis and interlanguage analysis, both of which have been shown to be necessary for a sound investigation of transfer, and relies on corpus data, the guarantee for the authenticity of the observations.

It would be naïve, however, to think that the model of transfer presented here provides definite and definitive answers to all transfer-related questions. Transfer is a slippery phenomenon that does not lend itself easily to apprehension. It sometimes comes up when one does not anticipate it, and can fail to appear where one would expect it, thus escaping the attention of even the best of observers. But with enough patience and by putting in enough effort and rigour in the task, it is to be hoped that one day transfer will reveal all of its secrets. This model offers a key to unlock some of them.

Notes

- 1 I gratefully acknowledge the financial support of the Belgian National Fund for Scientific Research (F.N.R.S.). For helpful comments on an earlier version of this paper, I wish to thank (in alphabetical order) María Belén Díez-Bedmar, Sylviane Granger, Claire Hugon, Scott Jarvis, Szilvia Papp and Magali Paquot.
- 2 Sharwood Smith & Kellerman (1986: 1-2) prefer the term “crosslinguistic influence” to “transfer” (or “interference”). One of their arguments is that the term “transfer” has traditionally been reserved for the influence of L_1 on L_2 or L_2 on L_3 , not the other way round. However, following Pavlenko

& Jarvis (2002), “transfer” will be understood here in a bidirectional sense. The term “interference” will be avoided because of its negative connotations (see Sharwood Smith & Kellerman 1986: 1).

- 3 “Erroneous usage” is to be taken here both in a qualitative sense, referring to misuse (i.e. errors in the strict sense), and in a quantitative sense, covering over- and underuse (see later on these notions). The same is true of “correct usage”, which can be correct both from a qualitative and a quantitative perspective.
- 4 The three same varieties of language are examined by Borin & Prütz (2004) and Tono (2004), but none of them offers the range of comparisons proposed here.
- 5 See Granger (1996) and Section 3 of this paper on contrastive interlanguage analysis. Note that, while the third type of comparison, NL₁/IL, includes a non-native variety of language, it is considered a case of contrastive analysis because it involves distinct languages (e.g. native French and non-native English), and not two varieties of one and the same language as is the case in contrastive interlanguage analysis.
- 6 It should be borne in mind, however, that the claims of the model might be slightly different for, say, phonological or pragmatic phenomena, or for languages which are radically different from European languages in these aspects.
- 7 See Gilquin (2000/2001) for a brief history of CA and CIA, as well as some bibliographical references on the subject.
- 8 “Errors” in a broad sense, see endnote 3.
- 9 Other influences are not excluded. It seems plausible to assume, following Paquot (2006), that the overuse of *on the contrary* is a general characteristic of interlanguage, given that all ICLE components exhibit this tendency (although it is more marked in the French component). This overuse may also be teaching-induced since, as Lake (2004) notes, many textbooks make no distinction between *on the contrary* and *on the other hand*, listing both discourse markers under the functional heading of “contrast”, which leads to potential confusion in the learner’s mind.
- 10 It should be emphasised, however, that the stage of learning was determined on the basis of external criteria (students in their third or fourth year of university study) and that a rater’s assessment of some 200 texts from ICLE based on internal criteria revealed proficiency levels ranging from B2 to C2 in the Common European Framework (but with a majority of C1), see Granger & Thewissen (2005).

- 11 Detection roughly corresponds to Jarvis's (2000) idea of "identification". Explanation may be related to the "diagnostic" hypothesis of Granger's (1996) ICM. Evaluation is not explicitly present in either model.
- 12 Although Jarvis (2000: 258) cites a case where transfer could be present despite a lack of correspondence between NL₁ and IL, he acknowledges that NL₁-IL performance congruity "may, in fact, be the strongest type of evidence for L1 influence".
- 13 While reversed translation as such is not used by Jarvis (2000), his method of comparing NL₁ and IL is similar in that it examines congruence between the words used in NL₁ and in IL to refer to the same denotatum.
- 14 Borgatti (2006) uses the term "back-translation" to describe this technique. Here, however, this term will be avoided for its potential confusion with the concept of back-translation as defined by Ivir (1983), which refers to the observation, in a bilingual corpus, of translated words and their equivalents in the source language.
- 15 That is, unless all L₁s share the feature under study, in which case similarity in the learners' behaviour is to be expected in case of transfer.
- 16 This does not exclude the possibility that a word has several equivalents, all of which may result in transfer. Thus, English *perhaps* and *maybe* correspond to just one word in French, *peut-être*. While *maybe* is a closer equivalent from an etymological point of view (*peut* = *may* and *être* = *be*), *perhaps* is more similar phonologically. There is therefore no reason to favour one word over the other as an equivalent of *peut-être*, and hence one could envisage that an English-speaking learner of French is influenced both by *perhaps* and *maybe* when producing *peut-être* in his/her interlanguage.
- 17 The opposite view, namely that identity between L₁ and L₂ leads to avoidance, has also been expressed (Kellerman [2000] refers to this phenomenon as "homoiphobia", the fear of resemblance). However, this seems to be true of certain items only (typically marked forms such as idioms, cf. Kellerman 1978) and to be limited to intermediate learners (Sjöholm 1998). Since the illustrations provided here are based on advanced learner language and non-idiomatic items, it is reasonably safe to assume that language identity is more likely to lead to transfer than language distance.
- 18 This difference in mutual translatability may be argued to originate from the structural difference between English and Swedish on the one hand, and French on the other. In the former case, the normal position of the non-finite complement is after the causee (cf. *He made her laugh*, *Han fick henne att skratta*), whereas in French the causative verb and its non-finite complement are in principle inseparable (cf. *Je l'ai fait rire*).

However, the Swedish causative construction and its English counterpart also differ in two important respects, namely *få*, which is the verb used in the Swedish construction, corresponds to English ‘get’, not ‘make’ (the normal equivalent of ‘make’ in Swedish, *göra*, cannot be used as a causative verb followed by a non-finite verb; see Viberg 2006: 251-252), and the infinitival complement in Swedish has to be preceded by the marker *att*, ‘to’.

- 19 The PLECI (*Poitiers-Louvain Échange de Corpus Informatisés*) corpus is an English-French bilingual corpus of journalistic and fictional texts (see Gilquin 2000/2001).
- 20 The context here is that of FLT for learners from the same mother tongue background. In the case of teachers having to address students with different mother tongues or teaching materials designed for several learner populations, it is difficult to take transfer into account, except, perhaps, if the learners’ mother tongues belong to the same language family.
- 21 See Paquot (forthcoming) on the notions of transfer of frequency and transfer of form.
- 22 See also Tenfjord et al. (2006), who argue that the tagging of errors in a learner corpus, which entails a native norm, does not necessarily lead to comparative fallacy.
- 23 By contrast, the concept of English as a Lingua Franca is still relatively vague and to date no one has yet been able to compile a complete repertoire of “non-core” features of English which would represent “no threat to intelligibility for the (non-native) receiver” (Jenkins 1998: 123). Mollin (2006), in fact, shows that what is generally referred to as “Euro-English” is too heterogeneous to be considered a “variety of English”.
- 24 It may also vary in one and the same learner according to external circumstances. Nickel (1989: 298) notes that stress, for example, is a factor that tends to increase transfer.
- 25 The only exception to this is the Spanish component, which was included since its size is very close to the limit (99,905 words).
- 26 Use was made of the Poitiers web interface (<http://cabal.rezo.net>, last accessed on January 13, 2007), which contains about 350,000 words of original newspaper articles in English or French and their translations in the other language. While the use of a journalistic corpus may be questionable, given the nature of the other corpora (argumentative essays), it should be emphasised that bilingual corpora are not available for all genres, which severely limits the choice. Bilingual corpora of scientific articles do exist (cf. Kraif & Tutin 2006), but they are still relatively rare (and small), and not readily available. Newspaper articles, which are more widely available, have been chosen here as the next best solution.

- 27 To be compared with 100.49 in native English (see later).
- 28 I thank Magali Paquot and Yves Bestgen for their help with the statistical analysis.
- 29 While the Italian component of ICLE was excluded because of its small size, Prat Zagrebelsky (2001) has shown that Italian learners also have a tendency to overuse *even if*, a result which she explains by L₁ influence.
- 30 Interestingly, the overuse of *even if* by French-speaking learners goes hand in hand with an underuse of *even though* (significant at the 0.001 level), which is probably to be linked to the learners' common confusion between the two expressions.
- 31 The same phenomenon seems to be at work in Italian learner English, since the Italian conjunction *anche se* can be used to express the meanings of both *even if* and *even though* (Prat Zagrebelsky 2001).
- 32 For an example of a study showing the importance of psychotypology, see Singleton (1987).
- 33 An equivalent such as *même*, in fact, is more common than *en effet* in the corpus data (6 occurrences).
- 34 Two learners provided *en fait* ('in fact') as a translational equivalent, and one, the rather formal adverb *certes* ('admittedly'). Interestingly, both equivalents occur twice each in the parallel corpus data.

References

- Altenberg, B. (1999), 'Adverbial connectors in English and Swedish: semantic and lexical correspondences', in: H. Hasselgård and S. Oksefjell (eds) *Out of Corpora. Studies in Honour of Stig Johansson*. Amsterdam & Atlanta: Rodopi. 249-268.
- Altenberg, B. (2002), 'Causative constructions in English and Swedish: A corpus-based contrastive study', in: B. Altenberg and S. Granger (eds) *Lexis in Contrast. Corpus-based Approaches*. Amsterdam & Philadelphia: John Benjamins Publishing Company. 97-116.
- Atkins, B.T., A. Duval, R.C. Milne, P.-H. Cousin, H.M.A. Lewis, L.A. Sinclair, R.O. Birks and M.-N. Lamy (eds) (1998), *Collins Robert Unabridged. French-English, English-French Dictionary*. Fifth edition. New York: HarperCollins.
- Benson, C. (2002), 'Transfer/cross-linguistic influence', *ELT Journal*, 56(1): 68-70.
- Bley-Vroman, R. (1983), 'The comparative fallacy in interlanguage studies: the case of systematicity', *Language Learning. A Journal of Applied Linguistics*, 33(1): 1-17.

- Borgatti, E. (2006) *The use of the verbs 'make' and 'do' by French- and Dutch-speaking EFL learners: A corpus-based study*. Unpublished MA dissertation, Université catholique de Louvain.
- Borin, L. and K. Prütz (2004), 'New wine in old skins? A corpus investigation of L1 syntactic transfer in learner language', in: G. Aston, S. Bernardini and D. Stewart (eds) *Corpora and Language Learners*. Amsterdam & Philadelphia: John Benjamins Publishing Company. 67-87.
- Danchev, A. (1991), 'Some notes on a Bulgarian model of expanded contrastive analysis', in: V. Ivir and D. Kalogjera (eds) *Languages in Contact and Contrast: Essays in Contact Linguistics*. Berlin: Mouton de Gruyter. 87-101.
- Dechert, H. and M. Raupach (eds) (1989), *Transfer in Language Production*. Norwood, NJ: Ablex.
- Ellis, R. (1994), *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. and G. Barkhuizen (2005) *Analysing Learner Language*. Oxford: Oxford University Press.
- Fivet, O. (1995), *A contrastive study of the passive in English and French journalesse*. Unpublished MA dissertation, Université catholique de Louvain.
- Gellerstam, M. (1986), 'Translationese in Swedish novels translated from English', in: L. Wollin and H. Lindquist (eds) *Translation Studies in Scandinavia*. Lund: CWK Gleerup. 88-95.
- Gilquin, G. (2000/2001), 'The Integrated Contrastive Model. Spicing up your data', *Languages in Contrast*, 3(1): 95-123.
- Gilquin, G. (forthcoming), 'Causative *make* and *faire*: A case of mismatch', to appear in: M.L.A. Gómez González, J.L. Mackenzie and E.M. González Álvarez (eds) *Current Trends in Contrastive Linguistics*. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Gilquin, G. and M. Paquot (2008), 'Too chatty: Learner academic writing and register variation'. *English Text Construction* 1(1): 41-61.
- Granger, S. (1996), 'From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora', in: K. Aijmer, B. Altenberg and M. Johansson (eds) *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies. Lund 4-5 March 1994*. Lund: Lund University Press. 37-51.
- Granger, S. (1997), 'Automated retrieval of passives from native and learner corpora: precision and recall', *Journal of English Linguistics*, 25(4): 365-374.
- Granger, S. (1998), 'Interlanguages in contrast: translationese vs learnerese'. Paper presented at the conference *Languages in Contrast. Information structure in parallel texts*, University of Oslo, 30-31 October 1998.
- Granger, S. (2003), 'The International Corpus of Learner English. A new resource for foreign language learning and teaching and second language acquisition research', *TESOL Quarterly*, 37(3): 538-546.

- Granger, S. and J. Thewissen (2005), 'Towards a reconciliation of a "Can Do" and "Can't Do" approach to language assessment'. Paper presented at the *Second Annual Conference of EALTA (European Association of Language Testing and Assessment)*, Voss, Norway, 2-5 June 2005.
- Granger, S., E. Dagneaux and F. Meunier (eds) (2002), *The International Corpus of Learner English. Handbook and CD-ROM*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Granger, S., E. Dagneaux, F. Meunier and M. Paquot (eds) (forthcoming), *The International Corpus of Learner English. Handbook and CD-ROM. Version 2*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Hammarberg, B. (2001), 'Roles of L1 and L2 in L3 production and acquisition', in: J. Cenoz, B. Hufeisen and U. Jessner (eds) *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Clevedon: Multilingual Matters. 21-41.
- Han, Z. (2004), 'Fossilization: five central issues', *International Journal of Applied Linguistics*, 14(2): 212-242.
- Ivir, V. (1983), 'A translation-based model of contrastive analysis', *Jyväskylä Cross-Language Studies*, 9: 171-178.
- James, C. (1980), *Contrastive Analysis*. London: Longman.
- Jarvis, S. (2000), 'Methodological rigor in the study of transfer: identifying L1 influence in the interlanguage lexicon', *Language Learning*, 50(2): 245-309.
- Jenkins, J. (1998), 'Which pronunciation norms and models for English as an International Language?', *ELT Journal*, 52(2): 119-126.
- Jenkins, J. (2000), *The Phonology of English as an International Language*. Oxford: Oxford University Press.
- Kamimoto, T., A. Shimura and E. Kellerman (1992), 'A second language classic reconsidered – the case of Schachter's avoidance', *Second Language Research*, 8(3): 251-277.
- Kellerman, E. (1978), 'Giving learners a break: Native language intuitions as a source of predictions about transferability', *Working Papers in Bilingualism*, 15: 59-92.
- Kellerman, E. (1983), 'Now you see it, now you don't', in: S. Gass and L. Selinker (eds) *Language Transfer in Language Learning*. Rowley, MA: Newbury House. 112-134.
- Kellerman, E. (1995), 'Crosslinguistic influence: transfer to nowhere?', *Annual Review of Applied Linguistics*, 15: 125-150.
- Kellerman, E. (2000), 'Lo que la fruta puede decirnos acerca de la transferencia léxico-sémantica: una dimensión no estructural de las percepciones que tiene el aprendiz sobre las relaciones lingüísticas' [What fruit can tell us about lexicosemantic transfer: A non-structural dimension to learners' perceptions of linguistic relations], in: C. Muñoz (ed.) *Segundas lenguas. Adquisición en la aula*. Barcelona: Ariel. 21-37.

- Kohn, K. (1986), 'The analysis of transfer', in: E. Kellerman and M. Sharwood Smith (eds) *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon Press. 21-34.
- Kraif, O. and A. Tutin (2006), 'Looking for semi-frozen expressions using an aligned corpus: an application for academic writing for EFL learners'. Paper presented at the 7th *Conference of Teaching and Language Corpora (TaLC)*, Université Paris 7 – Denis Diderot, 1-4 July 2006.
- Lado, R. (1957), *Linguistics across Cultures*. Ann Arbor: University of Michigan Press.
- Lake, J. (2004), 'Using "on the contrary": the conceptual problems for EAP students', *ELT Journal*, 58(2): 137-144.
- Mauranen, A. (2003), 'The corpus of English as lingua franca in academic settings', *TESOL Quarterly*, 37(3): 513-527.
- Mollin, S. (2006), *Euro-English. Assessing Variety Status*. Tübingen: Gunter Narr Verlag.
- Mukherjee, J. (2005), 'The native speaker is alive and kicking: linguistic and language-pedagogical perspectives', *Anglistik*, 16(2): 7-23.
- Nickel, G. (1989), 'Some controversies in present-day error analysis: "contrastive" vs. "non-contrastive" errors', *International Review of Applied Linguistics in Language Teaching*, 27(4): 293-305.
- Odlin, T. (1989), *Language Transfer*. Cambridge: Cambridge University Press.
- Paquot, M. (2006), 'EAP phraseology in the International Corpus of Learner English: the role of transfer'. Paper presented at the *Congreso Internacional de Fraseología Y Paremiología*, Universidad Santiago de Compostela, 19-22 September 2006.
- Paquot, M. (forthcoming), 'Exemplification in learner writing: A cross-linguistic perspective', to appear in: S. Granger and F. Meunier (eds) *Phraseology: An Interdisciplinary Perspective*. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Paquot, M. and C. Fairon (2006), 'Investigating L1-induced learner variability: Using the Web as a source of L1 comparable data'. Paper presented at the *International Computer Archive of Modern and Medieval English (ICAME) Conference (Variation, Contacts and Change)*, University of Helsinki, 24-28 May 2006.
- Pavlenko, A. and S. Jarvis (2002), 'Bidirectional transfer', *Applied Linguistics*, 23(2): 190-214.
- Petch-Tyson, S. (1998), 'Writer/reader visibility in EFL written discourse', in: S. Granger (ed.) *Learner English on Computer*. New York: Addison-Wesley Longman. 107-118.
- Prat Zagrebelsky, M.T. (2001), 'Even if or even though? A corpus-based investigation of the Italian-ICLE subcorpus'. Poster presented at the *International Computer Archive of Modern and Medieval English (ICAME) Conference (Future Challenges in Corpus Linguistics)*, Université catholique de Louvain, 16-20 May 2001.

- Ringbom, H. (1995), 'Error analysis', in: J. Verschueren, J.-O. Östman and J. Blommaert (eds) *Handbook of Pragmatics*. Amsterdam & Philadelphia: John Benjamins Publishing Company. 581-583.
- Schachter, J. (1974), 'An error in error analysis', *Language Learning*, 24(2): 205-214.
- Seidlhofer, B. (2001), 'Closing a conceptual gap: the case for a description of English as a lingua franca', *International Journal of Applied Linguistics*, 11: 133-158.
- Selinker, L. (1972), 'Interlanguage', *International Review of Applied Linguistics in Language Teaching*, 10(3): 209-231.
- Selinker, L. (1989), 'CA/EA/IL: The earliest experimental record', *International Review of Applied Linguistics in Language Teaching*, 27(4): 267-291.
- Sharwood Smith, M. and E. Kellerman (1986), 'Crosslinguistic influence in Second Language Acquisition: An introduction', in: E. Kellerman and M. Sharwood Smith (eds) *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon Press. 1-9.
- Singleton, D. (1987), 'Mother and other tongue influence on learner French. A case study', *Studies in Second Language Acquisition*, 9: 327-345.
- Sjöholm, K. (1998), 'A reappraisal of the role of cross-linguistic and environmental factors in lexical L2 acquisition', in: K. Haastrup and Å. Viberg (eds) *Perspectives on Lexical Acquisition in a Second Language*. Lund: Lund University Press. 209-236.
- Tenfjord, K., J.E. Hagen and H. Johansen (2006), 'The "hows" and "whys" of coding categories in a learner corpus (or "How and why an error-tagged learner corpus is not 'ipso facto' one big comparative fallacy")', *Rivista di Psicolinguistica Applicata*, 6(3): 93-108.
- Tisthoud, N. (1992), *Conceptual and lexical gaps: A problem of translation*. Unpublished MA dissertation, Université catholique de Louvain.
- Tono, Y. (2004), 'Multiple comparisons of IL, L1 and TL corpora: The case of L2 acquisition of verb subcategorization patterns by Japanese learners of English', in: G. Aston, S. Bernardini and D. Stewart (eds) *Corpora and Language Learners*. Amsterdam & Philadelphia: John Benjamins Publishing Company. 45-66.
- Viberg, Å. (2006), 'What one verb can do: The Swedish verb *göra* in a crosslinguistic perspective', in: M. Suominen, A. Arppe, A. Airola, O. Heinämäki, M. Miestamo, U. Määttä, J. Niemi, K.K. Pitkänen and K. Sinnemäki (eds) *A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday*. Special Supplement to *SKY Journal of Linguistics*, 19: 243-257.
- Zhao, R. (1989), 'A discourse analysis of relative clauses in Chinese and English: An error in "An error in error analysis"', *IDEAL*, 4: 105-117.