

New Journal – Order Now and Save 20%

## ■ Corpus Linguistics and Linguistic Theory

Editors-in-Chief: Anatol Stefanowitsch and Stefan Th. Gries

### Subscription rates for Volume 1 (2005)

Print only or Online only:

Institutions/Libraries: € 148.00 / sFr 237.00 / approx. US\$ 178.00

\*\*Discount Price: € 118.00 / sFr 188.00 / approx. US\$ 142.00

ISSN 1613-7027 (Print)

ISSN 1613-7035 (Online)

*Corpus Linguistics and Linguistic Theory* is a newly founded, peer-reviewed journal publishing high-quality original corpus-based research focusing on theoretically relevant issues in all core areas of linguistic research (phonology, morphology, syntax, semantics, pragmatics), or other recognized topic areas. It will provide a forum for researchers from different theoretical backgrounds and different areas of interest that share a commitment to the systematic and exhaustive analysis of naturally occurring language.

The following kinds of articles will appear in the journal:

- papers that develop new corpus-linguistic methods or extensions of existing methods of interest in the context of linguistic theorizing;
- papers that test or evaluate theoretical claims using corpus data and corpus-linguistic methods;
- papers that offer systematic and detailed analyses of individual linguistic phenomena within a theoretical framework;
- papers that compare corpus data to other kinds of empirical data, such as experimental or questionnaire data;
- critical surveys of relevant areas of research;
- squibs (short notes on theoretical issues, short reports of interesting data, or short replies to or rebuttals of previously published articles);
- reviews of new books, corpora, or software packages.

Studies concerned with language description and studies with an applied focus (for example, in lexicography or language teaching) may be included if they are informed by or have clear relevance to current issues in linguistic theorizing.

**Editorial Board:** Harald Baayen, Holger Diessel, Michelle Gregory, Stefan Grondelaers, Jennifer Hay, Frank Keller, Adam Kilgarriff, Manfred Krug, Anke Lüdeling, Katja Markert, Laura Michaelis, Douglas Roland, Anna Theakston, and Joe Trotta.

*Corpus Linguistics and Linguistic Theory* invites authors to submit their manuscripts electronically to the editors at [clt.journal@degruyter.com](mailto:clt.journal@degruyter.com) in PDF, RTF or MS Word format (contact the editors beforehand if you plan to submit in a different format). A style sheet will be made available soon via the journal's web page. For the time being, please refer to the *Linguistics* style sheet available online at <http://www.degruyter.de/instructions/lingss.pdf>.

W  
DE  
G  
Mouton  
de Gruyter  
Berlin · New York

All prices: postage extra.  
\*\*For orders received by December 31st, 2005  
Prices are subject to change.

WWW.MOUTON-PUBLISHERS.COM

## Corpus Linguistics and Language Contrast: Cases for Compensation

Kjell Johan Sæbø (University of Oslo, Norway)

### Abstract

Some contrasts between two languages are asymmetric in the sense that for some resource in one language, the other has no counterpart and may have to compensate through alternative resources. I discuss what translation corpora can tell us about how such deficits are compensated for. On the basis of three case studies (Positive Polarity Items, Logophoric Mood, Progressive Aspect), I conclude that translation corpora are valuable sources as far as the scope of compensatory resources is concerned, but that they must be supplemented by other sources to assess whether these resources are in fact necessary and sufficient.

### 1. Introduction\*

It is generally agreed that (bilingual or multilingual) language corpora represent an invaluable source of information about language contrasts. Language comparison and cross-linguistic research are bound to benefit immensely from computerised methods of corpus comparison. Indeed, contrastive linguistics has been a causal and an instrumental factor in the development of parallel (translation and comparable) text systems.

At the same time, corpora are fallible sources of information about language contrasts. There are not only margins of error; in translation corpora, there is a human factor involved, insofar as,

in particular, the translator can be influenced by the source language or information can go missing; and generally, if something does not occur in one language, it is not decidable whether this is a necessity or a tendency, a matter of style or grammaticality – corpora can never provide negative evidence. Consequently, there is an acute awareness of methodological problems in corpus-based contrastive linguistics and of ways of overcoming them by supplementing one methodology by another (“validation”; cf. e.g. Johansson 1998, Fabricius-Hansen (forthcoming)).

There are a variety of types of language contrasts (Teich 2002), and different ways of studying corpora will be appropriate for different types of contrasts. I will in what follows focus on one type of contrast, namely, the phenomenon that may be referred to as compensation, and on how compensation is reflected in parallel translation corpora.

By compensation I mean the question whether and, in the event, how an asymmetric language contrast can be or is counterbalanced: When two (relatively similar) languages are such that a certain lexicogrammatical resource is available in one but unavailable in the other – i.e., when a choice between two expressions is neutralised – what, if any, alternative means are there to compensate, that is, to express the same ideational meaning? Prima facie, corresponding sentences in the language without the relevant resource can be assumed to be potentially ambiguous; but as any language strives to minimise ambiguity, we may expect there to be alternative means of realising the same semantics. To tentatively fix a general notion of compensation, we may formulate the following, very coarse and vague hypothesis:

#### Compensation in Principle

If a resource present in language L1 is absent in language L2, there will be resources in L2 that are sometimes necessary and sufficient to fill the function of the L1 resource.

\* The paper is based on research in the framework of the project SPRIK *Språk i kontrast* ‘Language(s) in Contrast’ at the University of Oslo. I am indebted to the other members of the group for valuable comments and discussions, and to an anonymous reviewer for bringing to my attention certain relevant considerations in translation studies and for suggesting ways to compensate for my initial unfamiliarity with them.

This is plausible but not completely trivial: On the one hand, it implies that there is no consistent redundancy in a language – for each resource, there will be a necessity that shows up somewhere; on the other hand, it presupposes that languages are consistently comparable in terms of functions, meaning (in the sense of ‘sense’, not ‘reference’) is cross-linguistically invariant – and this is not uncontroversial, cf. e.g. Santos (forthcoming); finally, it seems to imply that perfect translation is possible (to be sure, this must be relativised to meaning, excluding context in the sense of cultural (in)congruences).

The alternative resources may come from other subsystems of the grammar (morphology, syntax, lexicon, prosody, etc.), and there may be several different resources interacting to fill the function of the L1 resource. Examples are:

- How the lack of definiteness marking in e.g. Russian seems to be counterbalanced by, inter alia, word order and aspect;
- How the lack of perfectivity marking in e.g. German seems to be counterbalanced by less grammaticalised resources.

Compensation in Principle may be in need of qualifications, but the main focus will be on another hypothesis, “Compensation in Practice”, establishing the relevant connection between contrasts and corpora. This hypothesis says that parallel corpora can be relied on as a source of information about compensation – compensatory strategies can be observed in translations (S/T = the source/target language):<sup>1</sup>

#### Compensation in Practice

If a translator chooses to deviate from a direct translation (in a way that requires effort) and this can be traced to the absence in T of a resource in S, this deviation is necessary and sufficient to fill the function of the S resource.

Needless to say, this hypothesis, thus formulated, raises some serious questions. The term *direct*

*translation* is polysemous (cf. Shuttleworth and Cowie 1997, 40f.), and even in the sense intended here, something akin to ‘literal translation’, it represents a vague notion. Also, whether a deviation from a direct translation requires effort is a question which is very difficult to make precise. Regarding direct translation, I do not understand it as translation so close to the structures of the source language as not to observe target language norms. By the qualification that the deviation from a direct translation should require effort I mean that “solutions” to a translation problem consisting in evading it by simplifying the message should not be considered, as such translations are likely to lose information and be insufficient. Finally, whether a choice to translate indirectly can be traced to the absence in the target of some resource in the source is of course very difficult to ascertain. I intend to focus on relatively clear cases, using common sense and appealing to the good will of the reader; furthermore, the problematic notions will become more transparent in the course of the discussion.

Any case corroborating Compensation in Practice at once corroborates Compensation in Principle: A case testifying to the necessity of some indirect translation in response to the absence in T of some S resource simultaneously constitutes evidence that there are L2 resources that are sometimes necessary to fill the function of a resource specific to L1. At the same time, any case conforming to Compensation in Practice will give a partial answer to the question how compensation is provided – the question about the nature and range of alternative resources.<sup>2</sup>

The relevance of Corpus Linguistics for Language Contrast is in this perspective an indirect one: We do not need corpora to tell us that German does not have a grammatical aspect, but we need them to tell us what it entails for German not to have a grammatical aspect – and what German does instead. In short, we need parallel corpora to tell us more about a plus/minus contrast established through other sources.

<sup>1</sup> Note that my use of the term *compensation* differs slightly from the way it has been used in translation studies, in particular by Harvey (1995 or 1998), who understands it as a strategy triggered not so much by the lack of an equivalent lexicogrammatical resource in the target language but by the loss of a source text effect. Throughout this paper, the emphasis is on equivalence in a semanticist’s sense, primarily concerning truth conditions.

<sup>2</sup> On the other hand, a case invalidating Compensation in Practice does not at the same time falsify Compensation in Principle: If a translator has tried to circumvent an ambiguity by an insufficient or unnecessary means, it may still be that there are both necessary and sufficient means to circumvent that same type of ambiguity.

I will consider data that support Compensation in Practice as well as data that do not, both concerning whether a deviation is sufficient and whether it is necessary. In particular, I will focus on three phenomena: Polarity items in English in contrast with Norwegian (Section 2), the German *Konjunktiv* in reported speech in contrast with English and Norwegian (Section 3), and the English progressive aspect in contrast with German and Norwegian (Section 4). My conclusion will be that parallel corpora are valuable sources regarding the scope of compensatory resources, but that they must be supplemented by other sources to assess whether these resources are in fact necessary and sufficient.

## 2. Positive Polarity Items

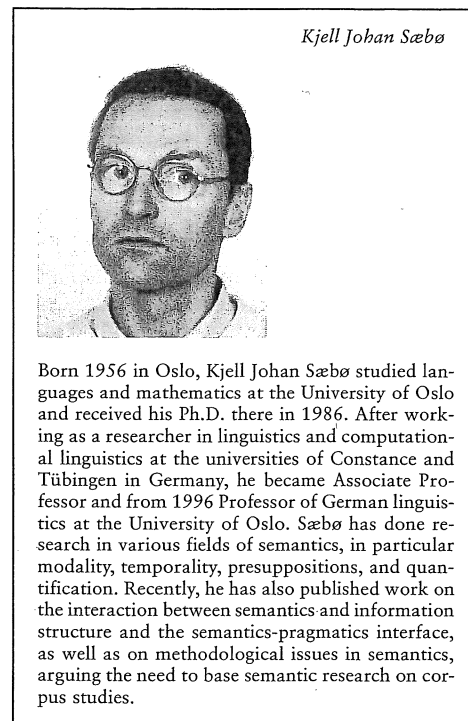
In English, a distinction is made between, on the one hand, *a, some* and some pronouns and adverbs based on *some* and, on the other hand, *any* and some pronouns and adverbs based on *any*. The *any(-)* items are either Free Choice Items or Negative Polarity Items (NPIs). As NPIs, they must occur in some downward entailing context, typically in the scope of some negative operator or quantifier (Ladusaw 1996).

A consensus on the proper definition of the contexts licensing NPIs or on the status of the Negative Polarity / Free Choice Item delimitation has not been reached (cf. Giannakidou 2001; 2002). In the following, I will only need to consider the *any* paradigm as NPIs in clearly negative contexts. My focus will be on the distinction between *any(-)* and *some(-)* as it manifests itself in such contexts. The items based on *some* are often called Positive Polarity Items (PPIs), so the focus will be on the opposition between NPIs and PPIs as witnessed by *any(-)* versus *some(-)*.

Mostly, the NPI / PPI distinction is informationally redundant, in the sense that there is no actual opposition between the two forms: When the context is clearly positive, the NPI is ruled out, as in (1).

- (1) \*Hermione’s trying to tell us anything.

When the context is negative, the PPI is not ruled out. But normally, it must then be interpreted outside the scope of the negative operator or quantifier, as in (2):



Kjell Johan Sæbø

Born 1956 in Oslo, Kjell Johan Sæbø studied languages and mathematics at the University of Oslo and received his Ph.D. there in 1986. After working as a researcher in linguistics and computational linguistics at the universities of Constance and Tübingen in Germany, he became Associate Professor and from 1996 Professor of German linguistics at the University of Oslo. Sæbø has done research in various fields of semantics, in particular modality, temporality, presuppositions, and quantification. Recently, he has also published work on the interaction between semantics and information structure and the semantics-pragmatics interface, as well as on methodological issues in semantics, arguing the need to base semantic research on corpus studies.

- (2) Hermione’s not telling us something.

Here the only possible reading is that where *something* outscopes *not*. So it is only superficially that the context is negative; the indefinite is not interpreted in a negative context.

Now the positive polarity (PP) properties do not seem to be as “strong” as the negative polarity (NP) properties; there are cases where a form based on *some* occurs felicitously in a clearly negative context (cf. Farkas 2002 for a recent discussion), like the authentic (3):

- (3) I wasn’t at all sure the Spedes themselves hadn’t done something.

This is a problem if *something* is a PPI in the sense that its PP property is a lexico-grammatical property. But according to Krifka (1995), forms based on *some* are not PPIs but get PP properties because forms based on *any* are NPIs:

I contend that NPs based on *some* are not polarity items at all. The observation about the scope dif-

ferences in cases like *Mary didn't see anyone* ( $\neg \exists$ ) and *Mary didn't see someone* ( $\exists \neg$ ) that have been adduced for the PPI status of *someone* rather should be explained as a paradigmatic effect induced by Grice's principle of ambiguity avoidance: In case a speaker wants to express the  $\neg \exists$  reading the unambiguous form containing *anyone* is preferred. It might very well be that this paradigmatic effect is so strong that it is virtually grammaticalized.

Since 1995, this pragmatic line of reasoning has become theoretically developed through Bidirectional Optimality Theory (Blutner 2000): A form-meaning pair is optimal if there is no better candidate form for the content and there is no better candidate content for the form (in terms of conditional informativity). Through the competition with *any(-)*, necessarily paired with the negative ( $\neg \exists$ ) interpretation, the pairing of *some(-)* with the positive ( $\exists \neg$ ) interpretation (if this interpretation is at all plausible) emerges as optimal. Thus NPI competitors will tend towards PP (Positive Polarity) properties: not to be interpreted in the scope of a negation in the same clause; if an interpretation of wide scope or specificity is plausible, this is the actual interpretation.

### 2.1. Translating PPIs I

In Norwegian, a corresponding distinction between NPI and PPI determiners or pronouns is not made generally (there is an NPI determiner in the singular, though). Thus both *something* and (NPI) *anything* map onto *noe*. This can be assumed to sooner or later cause an ambiguity in translation unless something else is called on for disambiguation. This prediction is borne out: (4b) is a maximally direct translation of (4a), and the inverse scope relation between the negation and the pronoun *some* is in practice reversed, so that (4b) really means the same as the back-translation (4c). (EO = English original; NT = Norwegian translation.<sup>3</sup>)

- (4) a. Or perhaps he deliberately didn't remember some of our names. (EO)

- b. Eller kanskje han ikke husket noen av navnene våre med vilje. (NT)  
c. Or perhaps he deliberately didn't remember any of our names.

Note that this piece of data is irrelevant for the hypothesis Compensation in Practice: Because the translator does not deviate from a direct translation, this hypothesis does not apply.

The next case to be considered seems to show that although some move is necessary and would in fact be sufficient, the move actually made by the translator is insufficient.

- (5) a. . . . , a fact which did not please some of the older veterans of the service. (EO)  
b. . . . , noe som ikke gledet de eldre veteranene i tjenesten. (NT)  
c. . . . , a fact which did not please the older veterans of the service.  
d. . . . , a fact which did not please any of the older veterans of the service.  
e. . . . , noe som ergrer noen av de eldre veteranene i tjenesten.  
f. . . . , a fact which annoyed some of the older veterans of the service.

The authentic translation (5b) expresses the same proposition as its direct English retranslation (5c), and we see that an essential distinction is lost: Although there is not an indefinite in the scope of the negation, as there would be in a direct translation, the interpretation is practically the same; (5c) is nearly synonymous to (5d). (5e), corresponding to the English (5f), shows a solution which would be sufficient.

However, a closer look at the example reveals that again, the hypothesis does not apply; this time because the deviation is not one that requires effort. The translator does deviate from a direct translation and it is reasonable to assume that this move is triggered by the wish to avoid the narrow scope interpretation of the indefinite *vis-à-vis* the negation, but the indirectness is a matter of simplification; the ambiguity is avoided by evading the problem. Note that (5e) would constitute an indirect translation requiring effort. (In a sense, of course, (5b) constitutes an indirect translation which does require an – intellectual – effort, entailing the recognition of the ambiguity; for the hypothesis to apply, though, this would

have to express itself in a positive deviation, as in (5e).)

So here, we have seen cases where the hypothesis Compensation in Practice does not apply, either because the translator does not deviate at all or because the deviation does not require effort. In both cases, the translation fails to preserve the truth conditions of the original.

### 2.2. Translating PPIs II

It is not difficult to find cases where the hypothesis Compensation in Practice does apply, though.

- (6) a. "D'you get the feeling Hermione's not telling us something?" Ron asked Harry. (EO)  
b. "Har du på følelsen at det er noe Hermine ikke forteller oss?" spurte Ron Harry. (NT)  
c. "D'you get the feeling there's something Hermione's not telling us?"  
d. Har du på følelsen at Hermine ikke forteller oss noe?

In (6b), the translator deviates from a direct translation in a way that requires effort and it is reasonable to trace this to the absence in Norwegian of the NPI / PPI distinction. A direct translation would be possible, (6d), but here the pronoun *noe* corresponding to *something* would preferably be interpreted in the scope of the negation. The translator has chosen to rephrase the embedded clause as a presentative construction, corresponding directly to the English (6c), with the effect that the pronoun is outside the overt scope of the negation. The result is truth-conditionally equivalent to the original. In sum, the deviation is both necessary and sufficient to fill the function of the English PPI. We have a case of compensation.

(7b) shows another strategy for dealing with a similar problem.

- (7) a. Neville almost always forgot to pack something. (EO)  
b. Nilus glemte nesten alltid noe når han skulle pakke. (NT)  
c. Neville almost always forgot something when packing.  
d. Nilus glemte nesten alltid å pakke noe.

Here the negation in the English original is implicit in the verb *forgot*, embedding an infinitival. The direct translation would be (7d), and this sentence has the preferred reading that Neville

almost always forgot to pack anything. The translator's choice is to make the pronoun corresponding to *something* the object of the verb corresponding to *forgot*, with the result that it escapes the scope of the implicit negation. Again, it is reasonable to assume that this deviation is motivated by the need to avoid the ambiguity potentially arising from the absence of the PPI in Norwegian. The move is necessary and sufficient to fill the PPI function. Again, we have a case of compensation. Both in (6b) and in (7b), the translation testifies to the presence in the target language of syntactic resources to compensate for the absence of a lexical resource. In (6b), one construction is preferred to another, in (7b), one valency frame for a verb is preferred to another (so this latter case can equally well be regarded as lexical). In the next case to be considered, the translator exploits a choice between two different word orders, more specifically, two different adjunction sites for an adverbial prepositional phrase:

- (8) a. it doesn't do in some countries to leave too many bills unpaid. (EO)  
b. i noen land nytter det ikke å la for mange regninger være ubetalt. (NT)  
c. in some countries it doesn't do to leave too many bills unpaid.  
d. det nytter ikke i noen land å la for mange regninger være ubetalt.

The direct translation of (8a) would be (8d), but here the indefinite Determiner Phrase *noen land* 'some/any countries' will tend to be interpreted in the scope of the negation. The actual translation (8b) deviates from this by left-adjoining instead of right-adjoining the adverbial PP *i noen land* 'in some/any countries', corresponding to (8c). In this way, the indefinite Determiner Phrase escapes the negative context, and the interpretation of the original (8a) is taken care of. Again, it is reasonable to assume that we have a case of compensation.

The cases considered so far would make us believe that indeed, when a translator deviates from a direct translation in a way that requires some effort and this can be assumed to be triggered by a need to compensate for something, then this move is necessary and sufficient as a compensation.

In the Oslo based English-Norwegian Parallel Corpus, it has not proved possible to find exam-

<sup>3</sup> The bulk of the translation data are from the Oslo based English-Norwegian Parallel Corpus (ENPC) or the Oslo Multilingual Corpus (OMC), consisting of fiction and non-fiction texts; all the data come from published originals and translations.

ples of translations of *some(-)* PPIs disproving Compensation in Practice. In the next section, we will look at a phenomenon where translations can cast doubt on the hypothesis.

### 3. The German Logophoric Mood

The central use of the present subjunctive (*Konjunktiv I*) in (written) German is to signal that the sentence has been uttered by somebody. This 'reportive subjunctive' occurs both in dependent and in independent clauses. When it occurs in dependent clauses, it is mostly redundant, since the superordinate verb is a verb of saying. When it occurs in independent clauses (*Berichtete Rede*, 'Reported Speech'), however, it is the only signal that the sentence has been uttered by somebody, and this can be assumed to cause ambiguities in translations into a language like English or Norwegian (cf. Pütz 1989). English and Norwegian can be expected to compensate for the lack of a reportive subjunctive in at least two ways:

1. through more frequent source citations,
2. through direct speech.

I will concentrate on 1 – the hypothesis that in English and Norwegian, source citations are sometimes necessary and sufficient to fill the function of the German reportive subjunctive mood. It is not difficult to find examples of translations adding source citations, evidently in order to keep the text unambiguous in the absence of a logophoric mood. The question is whether such additions are in fact necessary and sufficient. To answer this question, it is useful to consider translations into both English and Norwegian.

#### 3.1. Source Citations I

In the first case to be considered, both the English and the Norwegian translation contain an added source citation in the form of an apposition:

- (9) a. Der Generalsekretär der KPdSU Michail Gorbatschow hat verlautbart, es ginge der sowjetischen Führung beim 40. Jahrestag des Kriegsendes nicht darum, antideutsche Gefühle zu schüren. Die Sowjetunion trete für Freundschaft zwischen den Völkern ein.
- b. Mikhail Gorbachev, General Secretary of the Soviet Communist Party, has declared

that it is not the intention of the Soviet leaders to stir up anti-German feelings on the occasion of the 40th anniversary of the end of the war. The Soviet Union, he said, is committed to friendship between the nations.

- c. Generalsekretæren i Sovjetunionens kommunistiske parti Mikhail Gorbatsjov har tilkjennegitt at det ved førtiårsdagen for krigens slutt ikke for den sovjetiske ledelse gjaldt å hisse til antityske følelser. Sovjetunionen går, sa han, inn for vennskap mellom folkene.

In the German original, the present subjunctive form *trete* signals that the last sentence is a continuation of the announcement of General Secretary Gorbachev. It is reasonable to assume that the addition of *he said* and *sa han* is motivated by the need to signal the same. And the move is both necessary and sufficient: Without the added source citation, the interpretation that the last sentence is a conclusion drawn by the author of the text (in this case President von Weizsäcker) on the basis of the declaration of General Secretary Gorbachev is at least as readily accessible; cf. (9d):

- (9) d. Mikhail Gorbachev, General Secretary of the Soviet Communist Party, has declared that it is not the intention of the Soviet leaders to stir up anti-German feelings on the occasion of the 40th anniversary of the end of the war. The Soviet Union is committed to friendship between the nations.

When the two translations are as parallel as in (9a–c), the conclusion seems inevitable that the deviation from a direct translation is indeed necessary to compensate for the lack of something, here a logophoric mood. It may be, however, that the two translations both deviate, but that they do so in two different ways. In the next example, the Norwegian translation includes a source citation parallel to (9c), but the English translation is evidently designed to achieve the same effect by a more indirect route:

- (10) a. Ein schwarzäugiger, braunhäutiger Halbwüchsiger kam in Begleitung eines ihm ähnlichen Kindes zur Tür herein und tauschte an der Theke eine grosse leere Weinflasche gegen eine volle um; dabei stellte er das Kind als seinen Onkel vor. Er

gehe hier in die Volksschule, wo man eine besondere Ausländerklasse eingerichtet habe, welche "Bunte Klasse" heisse: nicht wegen der Farbstifte, die fast die einzigen Unterrichtswerkzeuge seien, sondern wegen der verschiedenen Hautfarben.

- b. A black-eyed, brown-skinned adolescent came in with a child who looked like him, and went to the bar, where he exchanged a large empty wine bottle for a full one. He introduced the child as his uncle and talked about himself. He went to the local public school; the special class that had been organized for foreigners was known as the "color class," not because of the crayons, which are virtually the only teaching aids in use, but because of the different skin colorations represented.
- c. En sortøyet, mørkhudet fremslenging kom inn i følge med en guttunge som var ganske lik ham; borte ved disken fikk de en stor flaske vin i bytte for tomflasken; den største gutten fortalte at den minste var hans onkel. Han gikk på folkeskolen her, sa han, og der var det satt opp en egen klasse for utlendingene, den ble kalt "broget klasse".

As the English deviation would be possible in the Norwegian translation as well, and vice versa, then if either deviation is sufficient to fill the function of the German mood, Compensation in Practice would seem to be invalidated: It is not necessary to add a source citation in just the way it is done in either translation.

Indeed, this case is not untypical, and it motivates a reformulation of Compensation in Practice:

#### Compensation in Practice (revised)

If a translator chooses to deviate from a direct translation (in a way that requires effort) and this can be traced to the absence in T of a resource in S, this or a similar deviation is necessary and sufficient to fill the function of the S resource.

#### 3.2. Source Citations II

So far, we have seen cases where both the English and the Norwegian translator feel the need to do something but where they may choose to do something different. There are cases, however, where only one of the two feels the need to do anything

at all. Consider the continuation of the text in (10a) and its translation into English and Norwegian:

- (11) a. Der Direktor sei stolz auf diese Klasse; sie habe sogar einen eigenen Eingang bekommen, und auch die Anfangszeiten seien andere als bei den österreichischen Schülern.
- b. The principal, said the boy, is proud of this class; he had even arranged for it to have a special entrance, and the hours are different from those of the Austrian classes.
- c. Rektor var stolt av denne klassen; den hadde egen inngang, og undervisningen begynte ikke samtidig med undervisningen for de østerrikske elevene.

The English translation (11b) contains an added source citation, while the Norwegian translation is a direct translation of the German sentence, modulo the tense: The German present subjunctive is rendered by the preterite, the same tense as in the surrounding text. This casts doubt on the necessity of the added source citation (or a similar addition) in the English translation. Probably, introspection or informant judgments will give the result that an English version corresponding to the Norwegian translation would in this case be sufficient for the interpretation that the sentence is a continuation of the boy's story:

- (11) d. The principal was proud of this class; he had even arranged for it to have a special entrance, and the hours were different from those of the Austrian classes.

In other cases, a closer investigation of a similar asymmetry between the two translations may give the opposite result: The added source citation is necessary to retain the report reading in a sufficient degree of clarity; cf. (12a–c):

- (12) a. Der Rat der Sachverständigen für Umweltfragen stellt in einem Gutachten fest, dass "in der Frauenmilch BetaHexachlorcyclohexan, Hexachlorbenzol und DDT oft in bedenklichen Konzentrationen gefunden werden". Diese Giftstoffe sind in Pflanzenschutzmitteln enthalten, die inzwischen aus dem Verkehr gezogen wurden. Ihre Herkunft sei ungeklärt.
- b. The Council of Experts on Environmental Issues determines in a report that "in mother's milk beta-hexachlorocyclohex-

ane, hexachlorobenzol and DDT are often found in significant concentrations". These toxic substances are contained in pesticides and herbicides that have by now been taken off the market. According to the report their origin is undetermined.

- c. Det sakkyndige råd for miljøspørsmål fastslår i sin betenkning at det "ofte blir funnet betaheksaklorcykloheksan, heksaklorbenzen og DDT i betenkelige konsentrasjoner i morsmelk". Disse giftstoffene finnes i plantevernmidler som i mellomtiden er blitt trukket tilbake fra markedet. Det er uklart hvor giftstoffene stammer fra.

The Norwegian translation is misleading; in fact, it seems incoherent: The third sentence (where the English translation has an added source citation) seems to contradict the second. Part of the reason is that in the German original, the first sentence, introducing the report, and the third sentence, continuing the report, are separated by a sentence in the indicative, interrupting the report. The English source citation *According to the report* is necessary to return to the report level. Along with this return goes the interpretation of *Ihre Herkunft, their origin* in the sense of 'how they entered the mother's milk', which is lost in the Norwegian translation, causing the sense of contradiction between the second and the third sentence.

The bottom line is that when there is a discrepancy between two translations such that *ceteris paribus*, only one translator has seen the need to add something, there is no principled way of knowing whether that addition is necessary. This may not be surprising. It is often hard to draw the line between a case of ambiguity and a case of a sufficiently preferred interpretation one way or the other, and these judgments may vary from hearer to hearer. In addition, when relations across larger segments of discourse are concerned, as in the cases considered in this section, there are subtle nuances, and sometimes interpretations that are in principle different may for practical purposes coincide.

What has not been falsified yet is that when it is necessary to add something, what is added is sufficient to retain the right interpretation. In the next section, we will have a look at a case where

this aspect of the Compensation in Practice hypothesis can be called into question.

#### 4. The English Progressive

As is well known, (Eastern and Southern) German does not have a grammaticalised means to match the English progressive; indeed, written German is not in possession of (viewpoint) aspect at all. There are various ways in which the English progressive aspect can be rendered in German (cf. Ebert 2000 for an overview); the more grammaticalised options consist in different prepositional locutions.

Quite often, the progressive is not rendered by anything without a loss of temporal information; other pieces of (lexical and non-lexical) information conspire to bring out the intended interpretation. Often enough, however, there is a need to bring in extra resources to express the temporal relations established by the progressive; in particular, when the lexical aspect (the situation type) is telic, the neutral forms will tend to convey a narrative progression which is halted by the progressive (cf. Tonne 2001 for corresponding findings as to Norwegian).

##### 4.1. Progressive Translation I

Thus in (13a), the progressive in the *when* clause depicts the dying as an extended process covering the waiting in the matrix clause; the simple past would favour the interpretation that the dying preceded the waiting. This will be the preferred interpretation of the German or Norwegian translation as well unless the sense of the progressive is conveyed by some means. In the German translation (13b), the *when* clause is rendered by a prepositional phrase, "at my father's deathbed", and in the Norwegian translation (13c), it is rendered by a temporal clause with a verb of posture, "lie".

- (13) a. I waited for two days with my father when he was dying, and wanted him to tell me that he loved me. (EO)  
 b. Ich habe zwei Tage am Sterbebett meines Vaters gegessen und mir gewünscht, dass er mir sagt, dass er mich liebt.  
 c. I to dager ventet jeg på å høre faren min si at han var glad i meg, da han lå for døden.

These or similar deviations from a direct translation (where a direct translation would be to simply render the past progressive by the past) are necessary and sufficient here. (14a–c) testifies to the same, supporting the revised version of Compensation in Practice. In (14a), the simple past instead of the progressive in the matrix clause would favour the interpretation that the dumping succeeded the cresting in the *as* clause, and to produce the progressive interpretation that the dumping surrounded the cresting, the German and the Norwegian translations must employ less grammaticalised locutions. The German translation (14b) includes the adverb *gerade* 'just', while the Norwegian translation (14c) makes use of the so-called pseudo-coordination of a verb of posture, here 'stand', and the translated verb:

- (14) a. As the Jaguar crested the headland Neil Pascoe was dumping rubbish into one of the two dustbins outside the caravan, . . .  
 b. Als der Jaguar die Hügelkuppe überquerte, stopfte Neil Pascoe gerade Abfall in eine der beiden Mülltonnen neben dem Wohnwagen, . . .  
 c. Da Jaguaren kom kjørende over toppen av neset, stod Neil Pascoe og la søppel i en av de to dunkene utenfor campingvognen, . . .

Again, these or similar moves are necessary and sufficient to convey the original interpretation, so again, Compensation in Practice is corroborated.

##### 4.2. Progressive Translation II

In the two cases to be considered next, however, there is in the German translation a deviation from a simple past tense translation almost certainly caused by the presence in English but absence in German of the progressive, but that deviation fails to preserve the meaning of the original in this connection. Some measure is necessary, but the course taken is not sufficient; in fact, one misinterpretation is avoided at the cost of producing another, or an infelicity is avoided only to produce a misinterpretation. Consider first (15a–c):

- (15) a. They were dragging me back when we discovered that the compound was burning.

- b. Sie zerrten mich zurück,  
 [they dragged me back  
 und dann entdeckten wir plötzlich, dass der  
 and then discovered we suddenly that the  
 ganze compound brannte.  
 whole compound burned]  
 c. Mens de halte meg tilbake,  
 [while they dragged me back  
 oppdaget vi at compounden brant.  
 discovered we that compound-the  
 burned]

In the English original (15a), the event of dragging me back starts before and is interrupted by the discovery that the compound was burning; with the simple past, it would be the other way around: The event of dragging me back would be triggered by the discovery. This interpretation would persist in a translation rendering the past progressive by the past. It is reasonable to assume that it is to avoid this reversal of the temporal relation between the two events that the German translator has chosen to present the discovery not in a temporal clause but in an independent clause introduced by 'and then'; the result, however, is an interpretation where the event of dragging me back (to the compound) is completed before the discovery, conflicting with the original interpretation that that event is interrupted by the discovery. The Norwegian translator solves the problem by presenting the event of dragging me back in a 'while' clause.

In (16a), the progressive in the *because* clause is important because the temporal adverb *already* makes the reference time punctual, and the telic verb phrase *open the door with a key* cannot be evaluated at a point of time; the atelic verb phrase *be opening the door with a key* can. In German and Norwegian, a translation of this clause using the present tense would be odd in a similar way as an English version with the simple present. The Norwegian translation (16c) uses a semi-grammaticalised progressive form *i ferd med* (cf. Tonne 2001), whereas in the German translation (16b), the present progressive is rendered by a present perfect, resulting in the interpretation that the event of opening the door is completed prior to the author's realisation that this is the right house.

- (16) a. At first I think there must be some mistake; but no, this is the house all right,

because my father is already opening the door with a key.

- b. Zuerst glaube ich, dass sie sich geirrt haben müssen; aber nein, es ist schon das richtige Haus, denn mein Vater hat bereits mit einem Schlüssel die Tür aufgeschlossen.
- c. Først tror jeg det må være en misforståelse; men nei, det er dette huset, for min far er alt i ferd med å låse opp døren.

While this misinterpretation may not be too grave, it does go to show the same point as the German translation (15b) above: An indirect translation requiring effort cannot be trusted to be sufficient to fill the function of a resource present in the source although it can be traced to the absence of this resource in the target.

## 5. Conclusions

When we speak of language contrasts, we often refer not just to facts about resources present in one language but absent in another language, but also to the means available in the other language to compensate. To shed light on what evidence for this type of contrast can be provided by translation corpus data, I have examined three different phenomena, each giving a different answer to the question to what extent translations that are evidently designed to offer compensation can in fact be assumed to constitute cases of a necessary and sufficient compensation.

As is to be expected, direct translations and indirect translations less complex than the original are found to result in misrepresentations or a loss of information. But as long as the translator has evidently tried to compensate, the first phenomenon considered – polarity items in English as compared to Norwegian (Section 2) – has not given counterevidence to the hypothesis that such translations can be trusted. The second phenomenon, however – the logophoric mood in German as compared to English and Norwegian (Section 3) – has called into question the claim that intended compensations are invariably necessary. And the third phenomenon – the progressive in English as compared to German and Norwegian (Section 4) – has shown that we cannot always assume an intended compensation to be sufficient to preserve the meaning of the original. In sum,

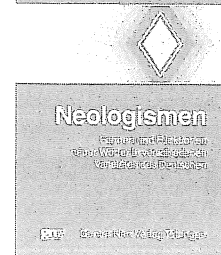
there is reason to treat parallel corpus data with caution even when the translator can be assumed to have made an effort. Translation corpora remain fallible as a source of data even outside the scope of “translationese”. The data that are *prima facie* relevant for compensation must be filtered through introspection or informant queries and supplemented by other methodologies.

This may not be surprising. After all, translations are not spontaneously produced texts. When we speak of something in one language filling the function of something in another language, we always understand a *ceteris paribus* premise – but in reality, other things are not as equal across languages as translations make them seem. It is to be expected, in particular as regards discourse-related phenomena like logophoricity or narrative progression, that two languages differ at a more global level than can be reflected in translation; that they simply structure the discourse differently (Doherty 1995, House 2002). Comparable corpora or language generation experiments can in principle provide a better picture of how the meaning of a lexicogrammatical resource in one language is produced in another language without that resource.

The main strength of parallel corpora in connection with compensation problems seems to lie in the inductive dimension, in the discovery procedure preliminary to in-depth investigations: There can be little doubt that translation corpora offer an invaluable source of information about the scope of compensatory strategies. To acquire a picture of the spectrum of different resources that can be applied to a compensation problem in a language, the examination of parallel corpora constitute an essential and efficient first step. What they seem to show may have to be validated through a filter of supplementary methods, like introspection and elicitation. But what is then established is certain to give a broader and sharper picture of a contrast than what might be and has traditionally been obtained through those other methodologies.

## References

- Blutner, Reinhard (2000), “Some Aspects of Optimality in Natural Language Interpretation”, in: *Journal of Semantics*, 17, 189-216.
- Doherty, Monika (1995), “Introduction”, in: M. Doherty, Ed., *Information Structure. A Key Concept for Translation Theory*. Special Issue of *Linguistics*, 34: 3, 441-457.
- Ebert, Karin (2000), “Progressive Markers in Germanic Languages”, in: Ö. Dahl, Ed., *Tense and Aspect in the Languages of Europe*, Berlin: Mouton de Gruyter, 605-653.
- Fabricius-Hansen, Cathrine (forthcoming), “Paralleltext und Übersetzung aus sprachwissenschaftlicher Sicht”, in: A. P. Frank, N. Greiner et al., Eds., *Übersetzung – Translation – Traduction. Ein internationales Handbuch zur Übersetzungsforschung/An International Encyclopedia of Translation Studies/Encyclopédie internationale des sciences de traduction*, Vol. 1, Berlin: Mouton de Gruyter.
- Farkas, Donka (2002), “Varieties of Indefinites”, in: *Proceedings of SALT 12*, 59-84.
- Giannakidou, Anastasia (2001), “The Meaning of Free Choice”, in: *Linguistics and Philosophy*, 24, 659-735.
- Giannakidou, Anastasia (2002), “Licensing and Sensitivity in Polarity Items: From Downward Entailment to (Non)Veridicality”, to appear in: M. Andronis, A. Pycha and K. Yoshimura, Eds., *CLS 38: Papers from the 38th Annual Meeting of the Chicago Linguistic Society, Parasession on Polarity and Negation*.
- Harvey, Keith (1995) “A Descriptive Framework for Compensation”, in: *The Translator*, 1, 65-86.
- Harvey, Keith (1998) “Compensation”, in: M. Baker, Ed., *Routledge Encyclopedia of Translation Studies*, London/New York: Routledge, 37-40.
- House, Juliane (2002), “Maintenance and convergence in covert translations English – German”, in: H. Hasselgård, S. Johansson, B. Behrens and C. Fabricius-Hansen, Eds., *Information Structure in a Cross-Linguistic Perspective*, Amsterdam: Rodopi, 199-211.
- Johansson, Stig (1998), “On the role of corpora in cross-linguistic research”, in: S. Johansson and S. Oksefjell, Eds., *Corpora and Cross-linguistic Research*, Amsterdam: Rodopi, 3-24.
- Krifka, Manfred (1995), “The Semantics and Pragmatics of Polarity Items”, in: *Linguistic Analysis*, 25, 209-257.
- Ladusaw, William (1980), *Polarity Sensitivity as Inherent Scope Relations*, New York: Garland.
- Ladusaw, William (1996), “Negation and Polarity Items”, in: S. Lappin, Ed., *The Handbook of Contemporary Semantic Theory*, Oxford: Blackwell, 321-341.
- Pütz, Herbert (1989), “Referat – vor allem berichtete Rede – im Deutschen und Norwegischen”, in: W. Abraham and T. Janssen, Eds., *Tempus – Aspekt – Modus. Die lexikalischen und grammatischen Formen in den germanischen Sprachen*, Tübingen: Niemeyer, 183-223.
- Santos, Diana (forthcoming), *Corpus-based contrastive semantics, with special reference to tense and aspect in Portuguese and English*, Amsterdam: Rodopi.
- Shuttleworth, Mark and Moira Cowie (1997), *Dictionary of Translation Studies*, Manchester: St. Jerome.
- Teich, Elke (2002), “System-oriented and text-oriented comparative linguistic research: Cross-linguistic variation in translation”, in: *Languages in Contrast*, 2, 187-210.
- Tonne, Ingebjørg (2001), *Progressives in Norwegian and the Theory of Aspectuality*, Ph.D. Dissertation, University of Oslo.



## Hilke Elsen Neologismen

Formen und Funktionen neuer Wörter in verschiedenen Varietäten des Deutschen

Tübinger Beiträge zur Linguistik 477, 2004,  
199 Seiten, div. Tab., € 48,-/Sfr 82,50  
ISBN 3-8233-6072-8

Wie und warum bilden wir neue Wörter? Wird das von einzelnen Sprechergruppen und -intentionen mitbedingt? Um diese Fragen zu beantworten, werden die Neologismen aus acht Varietäten des Deutschen, z.B. Werbe-, Zeitungs-, Fachsprachen, hinsichtlich Wortbildungstechnik und Gebrauchsfunktion analysiert. Die verschiedenen Sprachausprägungen unterscheiden sich dabei deutlich in der Wahl der bevorzugten Wortbildungsmöglichkeiten.

**gnv** Gunter Narr Verlag  
Postfach 2567 · D-72015 Tübingen