



Advances in Corpus-based Contrastive Linguistics

*Studies in honour of
Stig Johansson*

Edited by

Karin Aijmer and Bengt Altenberg

Studies in Corpus Linguistics

54

JOHN BENJAMINS PUBLISHING COMPANY

Advances in Corpus-based Contrastive Linguistics

Studies in honour of Stig Johansson

Edited by

Karin Aijmer

University of Gothenburg

Bengt Altenberg

University of Lund



John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Cover design: Françoise Berserik
Cover illustration from original painting *Random Order*
by Lorenzo Pezzatini, Florence, 1996.

Library of Congress Cataloging-in-Publication Data

Advances In Corpus-based Contrastive Linguistics : Studies in Honour of Stig Johansson /

Edited by Karin Aijmer, Bengt Altenberg.

p. cm. (Studies in Corpus Linguistics, ISSN 1388-0373 ; v. 54)

Includes bibliographical references and index.

1. Contrastive linguistics. 2. Computational linguistics. 3. Corpora (Linguistics) I.

Aijmer, Karin, editor of compilation. II. Altenberg, Bengt, editor of compilation.

III. Johansson, Stig, 1939- honouree.

P134.A35 2013

410--dc23

2012043843

ISBN 978 90 272 0359 5 (Hb ; alk. paper)

ISBN 978 90 272 7232 4 (Eb)

© 2013 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

List of contributors

Karin Aijmer	University of Gothenburg, Sweden, karin.aijmer@eng.gu.se
Bengt Altenberg	University of Lund, Sweden, bengt.altenberg@englund.lu.se
Jorge Arús	The Complutense University of Madrid, Spain, jarus@filol.ucm.es
Kate Beeching	University of West England, Bristol, UK, Kate.Beeching@uwe.ac.uk
Sylvie De Cock	Université Catholique de Louvain, Belgium, sylvie.decock@uclouvain.be
Jarle Ebeling	University of Oslo, Norway, jarle.ebeling@usit.uio.no
Thomas Egan	Hedmark University College, Norway, Thomas.egan@hihm.no
Diane Goossens	Université Catholique de Louvain, Belgium, diane.goossens@uclouvain.be
Sylviane Granger	Université Catholique de Louvain, Belgium, Sylviane.Granger@uclouvain.be
Hilde Hasselgård	University of Oslo, Norway, hilde.hasselgard@ilos.uio.no
Jennifer Herriman	University of Gothenburg, Sweden, jennifer.herriman@eng.gu.se
Marlén Izquierdo	University of León, Spain, marlen.izquierdo@unileon.se
Kerstin Kunz	University of Saarland, Germany, k.kunz@mx.uni-saarland.de
Julia Lavid	The Complutense University of Madrid, Spain, lavid@filol.ucm.es
Marie-Aude Lefer	Université Catholique de Louvain, Belgium, marie-aude.lefer@uclouvain.be

Lara Moratón	The Complutense University of Madrid, Spain, laramoraton@gmail.com
Signe Oksefjell Ebeling	University of Oslo, Norway, s.o.ebeling@ilos.uio.no
Rosa Rabadán	University of León, Spain, rosa.rabadan@unileon.es
Anne-Marie Simon-Vandenberghe	University of Ghent, Belgium, AnneMarie.Vandenberghe@UGent.be
Erich Steiner	University of Saarland, Germany, e.steiner@mx.uni-saarland.de
Anna-Brita Stenström	University of Bergen, Norway, ab.stenstrom@telia.com
Åke Viberg	University of Uppsala, Sweden, Ake.Viberg@lingfil.uu.se

Table of contents

List of contributors	IX
Introduction	1
<i>Karin Aijmer and Bengt Altenberg</i>	
<i>Tertia comparationis</i> in multilingual corpora	7
<i>Thomas Egan</i>	
Seeing the lexical profile of Swedish through multilingual corpora: The case of Swedish <i>åka</i> and other vehicle verbs	25
<i>Åke Viberg</i>	
A corpus-based analysis of English affixal negation translated into Spanish	57
<i>Rosa Rabadán and Marlén Izquierdo</i>	
English adverbs of essence and their equivalents in Dutch and French	83
<i>Anne-Marie Simon-Vandenberg</i>	
A parallel corpus approach to investigating semantic change	103
<i>Kate Beeching</i>	
Youngspeak: Spanish <i>vale</i> and English <i>okay</i>	127
<i>Anna-Brita Stenström</i>	
Quantity approximation in English and French business news reporting: More or less the same?	139
<i>Sylvie De Cock and Diane Goossens</i>	
Enriching the phraseological coverage of high-frequency adverbs in English-French bilingual dictionaries	157
<i>Sylviane Granger and Marie-Aude Lefer</i>	
Using recurrent word-combinations to explore cross-linguistic differences	177
<i>Jarle Ebeling, Signe Okseffell Ebeling and Hilde Hasselgård</i>	
Cohesive substitution in English and German: A contrastive and corpus-based perspective	201
<i>Kerstin Kunz and Erich Steiner</i>	

The extraposition of clausal subjects in English and Swedish <i>Jennifer Herriman</i>	233
Thematic variation in English and Spanish newspaper genres: A contrastive corpus-based study <i>Julia Lavid, Jorge Arús and Lara Moratón</i>	261
Author index	287
Subject index	291

Introduction

Karin Aijmer and Bengt Altenberg

University of Gothenburg / University of Lund

This volume contains a selection of studies presented at a workshop on 'Corpus-based contrastive analysis' held in connection with the ICAME32 Conference in Oslo on 1–5 June 2011. In addition, a few specially invited contributions have been included.

The conference was organized in honour of Professor Stig Johansson, University of Oslo, who died in 2010. Stig Johansson was one of the founders of ICAME (the International Computer Archive of Modern and Medieval English) and an enthusiastic and inspiring pioneer of corpus-based language research.¹ In the 1970s he was a central member of the team that created the Lancaster-Oslo/Bergen Corpus (LOB), the British counterpart of the American Brown Corpus. Together the British and American corpora were to set the pattern for a number of computer corpora developed around the world in the following decades.

When Stig Johansson and his team of researchers at the universities of Oslo and Bergen created the English-Norwegian Parallel Corpus in the early 1990s (see Johansson & Hofland 1994) he initiated a new era in contrastive linguistics. Contrastive analysis had enjoyed a temporary peak of interest in the 1950s and 1960s, but the approach was then largely intuitive and limited to comparing abstract language systems (or subsystems) rather than exploring languages in use. Moreover, the practical applications, mainly in language learning and teaching, were over-optimistic and disappointing and interest in contrastive analysis as a linguistic discipline soon declined.

Basic to 'the new era' in contrastive linguistics is the comparison of different languages on the basis of computer corpora and the use of corpus-linguistic methods. Stig Johansson's corpus-based approach placed contrastive analysis on a sound empirical footing. By combining the methodological advantages of computer corpus linguistics and the possibility of contrasting 'parallel' texts in two (and later several) languages, he and his team made it possible to compare the actual

1. For information about ICAME, see <<http://icame.uib.no/>>

use of the languages involved at all levels of description, from lexis to discourse, with far greater accuracy and detail than had been possible before. Moreover, the approach proved to have fruitful applications in a number of areas, such as language teaching, lexicography, translation studies and computer-aided translation.

The English-Norwegian Parallel Corpus (ENPC) was a bidirectional translation corpus consisting of comparable English and Norwegian original texts, representing various (printed) genres, and their translations into the other language.² He called it a 'parallel' corpus, partly inspired by the Rosetta Stone with its 'interlinear' presentation of three languages (see Johansson 2007: 4). Another inspiration may have been the Anglo-Saxon translation of the Vulgate version of the Bible, with the Latin and English versions presented in parallel, interlinear fashion.

Contrastive research on bilingual and multilingual parallel corpora exploits translators' competence and ability to find the 'right' correspondence in context. The analyst can move back and forth between sources and translations in order to build up paradigms which show the correspondences between lexical elements or grammatical constructions and meanings in the compared languages. The paradigms provide a blueprint of the similarities and differences between the languages compared. They are raw material for a maximally rich representation of the meanings and functions of a linguistic item which is based on more objective data than the analyst's intuitions. This approach is particularly useful when we study elements which are multifunctional and have no clear core meaning.

The fact that the English-Norwegian Parallel Corpus was based on comparable original English and Norwegian texts (i.e. text types of similar character in terms of genre, register, period, etc.) was a recognition of the problems involved in relying on translations alone for contrastive analysis. Translations need to be used with care since they tend to be affected by various 'translation effects', i.e. influences from the source language or from general translation strategies. To eliminate or reduce this potential source of deviations in translated texts it is necessary to verify the results against language use in original texts. Moreover, translation corpora are seldom big enough to provide evidence of less common language features, nor do they cover all the text types or genres researchers may be interested in (e.g. business language or spontaneous speech). Comparable monolingual corpora are therefore a necessary complement to translation corpora, either as a starting-point or as a verification of results produced by translation corpora. Sometimes, in the absence of translations, comparable corpora are the only possible source of contrastive observations. However, comparable corpora have the drawback of lacking

2. For information about the ENPC, its composition and structure, see <<http://www.hf.uio.no/ilos/english/services/omc/enpc/index.html>>. On the coding and alignment of the corpus, see Johansson et al. (1996).

an obvious *tertium comparationis*, i.e. a “background of sameness against which differences can be viewed and described” (Johansson 2007: 39). As Stig Johansson put it (*ibid.* p. 3), “How do we know what to compare?”

Corpus-based contrastive analysis has experienced a dramatic development since the 1990s. It has been applied to an increasing number of language pairs and it has seen the creation of multilingual translation corpora, again inspired by Stig Johansson and his colleagues at Oslo and Bergen (the Oslo Multilingual Corpus).³ Useful software has been developed for linking source and target texts, for retrieving linguistic elements and for tagging and annotating the compared languages at different levels of analysis. Although computer corpus linguistics is primarily a methodology, the use of corpora for contrastive analysis has led to new insights into the languages compared. These insights, whether corpus-driven or corpus-informed, have resulted in more realistic, detailed and empirically sound comparisons of languages, both in terms of their structure and use.

Research in contrastive linguistics has recently ventured into new domains such as pragmatics, text linguistics and discourse. Contrastive analysis has for instance been an attractive approach to study elements which are multifunctional and have no clear meanings. In this case the translations can provide an answer to thorny questions about polysemy or multi functionality, core meaning and the distinction between meaning and function. The results from the contrastive analysis can also be used to trace diachronic changes of elements that have been grammaticalized, as shown in several contributions to this volume.

The rapid development of corpus-based contrastive analysis since the 1990s had led to great vitality and productivity in the field. This is reflected in an increasing number of languages compared and a growing variety of topics and methodological approaches. These developments are clearly demonstrated in the present volume. The studies compare linguistic phenomena in eight languages: English (the common hub), Dutch, Finnish, French, German, Norwegian, Spanish and Swedish. The topics range from comparisons of specific lexical categories, word combinations and collocations to syntactic constructions and discourse phenomena such as cohesion and thematic structure. The point of departure is typically a preselected linguistic form or category but the aim is generally to highlight similarities and differences in the structure, semantics or functions of the compared items across the languages boundaries, to reveal divergences in their use, or the emergence of new meanings and language change. The material used is either bilingual or multilingual parallel corpora or comparable monolingual corpora (often purpose-built), or a combination of these types. The emphasis varies from

3. On the Oslo Multilingual Corpus, see <<http://www.hf.uio.no/ilos/english/services/omc/>>

purely linguistic studies to those focusing on some practical application (e.g. in lexicography or translation studies).

Thomas Egan compares the coding of the concept of ‘betweenness’ in English and French on the basis of translations of the Norwegian preposition *mellom* in the Oslo Multilingual Corpus. Egan distinguishes seven senses encoded in *mellom* and by using these as *tertia comparationis* he explores the degree of resemblance between the English and French translations. The study shows that there is a considerable degree of similarity between English and French in their encoding of ‘betweenness’.

Using a typological perspective as a starting point, Åke Viberg examines Swedish verbs describing motion in a vehicle and their correspondences in a multilingual parallel corpus consisting of Swedish original texts and their translations into English, German, French and Finnish. He demonstrates that the languages differ not only with regard to their inventory of vehicle verbs, the degree to which a certain contrast is obligatory and the semantic extension of individual verbs, but also in terms of usage-based tendencies that favour certain perspectives or alternative ways of coding a certain type of situation.

Rosa Rabadán and Marlén Izquierdo examine how English affixal negation is translated into Spanish and the extent to which the use and distribution of the translations differ from those in non-translated Spanish texts. The study is carried out in two steps: the Spanish translations of negative English affixes are first determined on the basis of a parallel corpus; the Spanish translations are then matched against a monolingual corpus of original Spanish. The data show clear differences between translated Spanish and regular native usage. The findings are analyzed both quantitatively and qualitatively and interpreted in terms of various translation universals.

Anne-Marie Simon-Vandenberghe examines the French and Dutch correspondences of the English ‘adverbs of essence’, *basically*, *essentially* and *fundamentally*, on the basis of the British National Corpus and a bidirectional, trilingual translation corpus. The point of departure is that the French and Dutch translations can throw light on the semantics and pragmatics of the English adverbs. Although the three adverbs are semantically similar, the study reveals similarities as well as differences, both among the items and across the languages. For example, apart from their common core meaning, *basically* has developed an extension in the direction of a softener and *fundamentally* in the direction of an amplifier. The reason seems to be that different pragmatic implicatures have been foregrounded and conventionalized.

Kate Beeching uses translations as a means of establishing language change. More precisely, by looking at the English translations of the multifunctional French pragmatic marker *quand même* in three parallel corpora, including two

of spoken language, she can establish the emergence of a 'relational' (or interpersonal) function of the French expression.

Anna-Brita Stenström starts from an intuitively observed similarity between English *okay* and Spanish *vale* and examines their functional similarity as discourse markers as well as their social and gender-determined distribution in two comparable corpora of teenage speech. Her study shows that, although *vale* is slightly more common than *okay*, both perform the same functions. Both are used more frequently by teenagers than by adults.

Sylvie De Cock and Diane Goossens compare the range of approximating devices that appear with numbers in two comparable corpora of English and French business news. To establish comparable units they first use part-of-speech tagging to extract numbers in the two corpora and then a collocation program to retrieve recurrent approximators in their vicinity. The study demonstrates that, although the same semantic categories of approximation are represented in both corpora, some are preferred in either English or French. Most of the grammatical realizations are also shared by the two languages but some types are language-specific and some favoured in one of the corpora.

Sylviane Granger and Marie-Aude Lefer use a combination of monolingual and translation corpus data to check the coverage and treatment of phraseological units in three major English-French bilingual dictionaries. Their study focuses on two high-frequency adverbs, French *encore* and English *yet*. Using the n-gram method to extract 'lexical bundles' involving these words from monolingual corpora and adding concordance data from translation corpora, the authors match the results with the entries for *encore* and *yet* in bilingual dictionaries, revealing various shortcomings in the coverage, exemplification and authenticity of the phraseological units in the dictionaries. Although the translation corpus data are genre-restricted, the authors clearly demonstrate the usefulness of contrastive corpus research for applications in bilingual lexicography.

In a corpus-driven study Jarle Ebeling, Signe Oksefjell Ebeling and Hilde Hasselgård explore phraseological differences between English and Norwegian on the basis of the English-Norwegian Parallel Corpus. By extracting n-gram lists of recurrent word combinations in original and translated texts from the two languages they uncover cross-linguistic differences that are likely to elude investigations that start from preselected lexical or grammatical categories. In three case studies the authors reveal divergences that point to constructional, semantic and pragmatic differences between the languages.

Kerstin Kunz and Erich Steiner examine cohesive substitution in English and German on the basis of a combination of comparable and bidirectional translation corpora. Using the semantic/functional characteristics of substitution (as distinct from other types of cohesion) as a *tertium comparationis*, they outline the English

and German options for nominal, verbal and clausal substitution. By means of fine-grained extraction rules that allow multilevel queries, they find among other things that substitution is more common in English than in German, whereas German uses a greater variety of forms. The latter is partly related to the fact that, although verbal and clausal substitution hardly exist in a strict sense in German, nominal substitution is more finely differentiated. The English forms display a higher degree of grammaticalisation and lexical bleaching.

Jennifer Herriman uses the English-Swedish Parallel Corpus to compare the use of extraposed subject clauses in the two languages. Looking at their distribution in both source texts and translations she finds that, despite their formal similarity in the two languages, extraposition is on the whole more frequent in Swedish, partly due to grammatical differences between the two languages and partly to the greater English tolerance of placing new or 'weighty' information in preverbal position.

Julia Lavid, Jorge Arús and Lara Moratón explore the influence of genre-related and language-specific factors influencing the way English and Spanish journalists begin their clauses in comparable corpora of English and Spanish news reports and commentaries. The basis of the comparison is a language-independent systemic framework for analyzing the choices made in the thematic field of clauses in the two genres. The authors show that, although there are language-specific differences in the thematic structure at clause level, most of the clausal and discourse thematic choices are a reflection of the different communicative purposes which characterize news reports and commentaries. In other words, genre tends to play a greater role for the choice and structure of the theme than language-related differences.

References

- Johansson, S. 2007. *Seeing through Multilingual Corpora. On the Use of Corpora in Contrastive Studies* [Studies in Corpus Linguistics 26]. Amsterdam: John Benjamins.
- Johansson, S., Ebeling, J. & Hofland, K. 1996. Coding and alignment of the English-Norwegian Parallel Corpus. In *Languages in Contrast*, K. Aijmer & B. Altenberg (eds), 87–112. Lund: Lund University Press.
- Johansson, S. & Hofland, K. 1994. Towards an English-Norwegian parallel corpus. In *Creating and Using English Language Corpora*, U. Fries, G. Tottie & P. Schneider (eds), 25–37. Amsterdam: Rodopi.

Tertia comparationis in multilingual corpora

Thomas Egan

Hedmark University College

This paper compares coding of the concept of ‘betweenness’ in English and French on the basis of translations of expressions in the same Norwegian texts into both languages. It argues that the original Norwegian expressions comprise a viable *tertium comparationis* for the comparison of the other two languages. Seven different senses of ‘betweenness’ are distinguished. Data from the Oslo Multilingual Corpus show that English and French resemble one another closely in the means employed to code all seven of these senses.

Keywords: *Tertium comparationis*, 3-text corpora, semantic equivalence, prepositions, ‘betweenness’

1. Introduction

Johansson (2007: 39) touches on the status of *tertium comparationis* in contrastive studies. He writes: “Much discussion in contrastive analysis has revolved around the question of the *tertium comparationis*, i.e. the background of sameness against which differences can be viewed and described”. However, the status of various sorts of *tertia comparationis* would seem to have been more of a topic of discussion among pragmatists and sociolinguists than among corpus linguists (see references in Jaszczolt 2003). Nevertheless, any contrastive corpus linguist who takes translation equivalence as evidence of semantic equivalence is working on the overt or tacit assumption that there exists a viable *tertium comparationis* in the form of a meaning component common to both the source expression and its translation.

In this chapter I operationalise the notion of *tertium comparationis* in a study of how the notion of ‘betweenness’ is encoded in English and French, comparing translation equivalents in these two languages of the Norwegian preposition *mellom*, which encodes the ‘betweenness’ relationship. Section 2 contains a brief description of various types of linguistic equivalence and explains the reasons behind the approach taken in this study. Section 3 explores the semantic field of

'betweenness' and distinguishes seven main senses of the concept. In Section 4 I compare encodings of 'betweenness' in English and French. Finally, Section 5 contains a summary of the discussion and some suggestions for further research.

2. Types of equivalence and *tertium comparationis*

Expressions in two languages may resemble one another syntactically, semantically and/or pragmatically, or they may, of course, not resemble one another in any respect, in which case we have no grounds for comparing them. As Krzeszowski (1990) puts it:

All comparisons involve the basic assumption that the objects to be compared share something in common, against which differences can be stated. This common platform of reference is called *tertium comparationis*. Moreover, any two or more objects can be compared with respect to various features and, as a result, the compared objects may turn out to be similar in some respects but different in others. (Krzeszowski 1990: 15)

Two expressions may be syntactically or lexically equivalent, like the Norwegian predication in (1a) and the English one in (2a), both taken from the English Norwegian Parallel Corpus (ENPC).

- | | | | |
|-----|----|--|--------------------|
| (1) | a. | Hildegun <i>fulgte</i> henne ut på gangen. | (BV1) ¹ |
| | b. | Hildegun <i>accompanied</i> her onto the landing. | (BV1T) |
| (2) | a. | Mattie <i>followed</i> her up the stone steps, | (GN1) |
| | b. | Mattie <i>fulgte etter</i> henne opp steintrappen, | (GN1T) |

Although the predications in (1a) and (2a) are syntactically equivalent, they differ in their semantics. The translation equivalents in (1b) and (2b) point to the relevant differences in meaning between the English verb *follow* and its Norwegian cognate, which means 'accompany'. The Norwegian semantic equivalent of English *follow* in (2b) may be literally translated as 'follow after'. In this case the *tertium comparationis* is the formal similarity between the two verbs which share a common etymology but have later evolved into false friends.

The comparison of (1a) and (2a) is semasiologically motivated, based as it is on two similar word forms. The present study is onomasiologically motivated,

1. The code (BV1) refers to the text in the ENPC from which the example has been taken; the presence of a final 'T' indicates a translated text, the absence of a 'T' a source language text. In the examples from the Oslo Multilingual Corpus 'TE' means translated text in English, 'TF' translated text in French.

comparing expressions coding ‘betweenness’ in two languages. The *tertium comparationis* in this case is therefore semantic/pragmatic, rather than syntactic equivalence. This does not mean that semantically equivalent expressions may not also be syntactically equivalent. Consider in this respect the pairs of sentences in (3) and (4), taken from the Oslo Multilingual Corpus (OMC).

- (3) a. Markus Torjussen takes his hand *between* both of his and shakes it emphatically. (BHH1TE)
- b. Markus Torjussen prend sa main *entre* les deux siennes et là secoue vigoureusement. (BHH1TF)
- (4) a. And the lovely paths *among* the flower beds, strewn with crushed shells! (HW2TE)
- b. Et les jolis sentiers *entre* les plates-bandes, recouverts de sable! (HW2TF)

The pairs of sentences in (3) and (4) are both syntactically congruent, as well as semantically equivalent. They thus conform to the requirements proposed by Krzeszowski (1990) for primary data for contrastive analysis.

In an earlier work (Krzeszowski 1981: 123), I suggested that syntactic contrastive studies should be performed on data restricted in the following way: a contrastive grammar will take as its primary data (to be assigned the status of semantic-syntactic equivalence) the closest approximations to grammatical word-for-word translations and their synonymous paraphrases, if such forms exist. Such a constraining of primary data as the basis for syntactic contrastive studies bypasses the inherent difficulties of the proposals suggesting the use of unrestricted semantic equivalence as the basis for comparison. (Krzeszowski 1990: 19)

There is no doubt that the pairs of sentences in (3) and (4) comprise close “approximations to grammatical word-for-word translation”. Indeed they *are* grammatical word-for-word translations, though not of each other. Krzeszowski (1990:25) employs the term *2-text* to refer to texts in either parallel or translated corpora. The availability of multilingual corpora, such as the OMC, allows us to operate with the concept of the *3-text*, with expressions in a source language serving as potential *tertium comparationis* for their translations into two other languages. This means that we have another means of solving the problem of unrestricted semantic equivalence raised by Krzeszowski. The essence of this problem is the difficulty of identifying the exact semantic content common to a source item and its translation. What the two sets of translated items in tokens such as (3) and (4) have in common is simply the fact that they are both translations of the same source items.

The fact that we have a *tertium comparationis* in the form of the Norwegian original also means that word-for-word translations are not privileged in any way, as they are in Krzeszowski’s approach. What is of interest is how the other two

languages go about encoding ‘betweenness’ in similar semantic contexts. In some cases they may use similar word-for-word structures. In other cases they may employ quite dissimilar constructions. Thus the pairs of sentences in (5) and (6) are semantically, though not syntactically equivalent.

- (5) a. “You mean perhaps that I’m not capable of distinguishing *between* fiction and reality?” (BHH1TE)
 b. – Vous voulez peut-être dire que je ne suis pas capable de distinguer fiction et réalité? (BHH1TF)
- (6) a. What was the relationship *between* the brothers like? (BHH1TE)
 b. Et comment s’entendaient les deux frères? (BHH1TF)

Though (5a) and (5b) differ from one another syntactically, they resemble one another lexically, in so far as they contain cognate verbs, *distinguish* and *distinguer*. In (6), on the other hand, there is no syntactic resemblance between the structures. There is no doubt, however, that both code a ‘betweenness’ relationship between the people involved.

Finally, there may be some pairs of sentences in a 3-text corpus which are not obviously semantically equivalent. In the present case, these are pairs in which it is impossible to detect a coding of ‘betweenness’ in the English or the French predication, or indeed in both. The pair in (7) may serve as an example.

- (7) a. “Dad moved,” he says after a long silence *between* them. (BHH1TE)
 b. “Papa est parti tout seul”, dit-il, après un long moment de silence. (BHH1TF)

One might argue that (7a) and (7b) are pragmatically equivalent in that the silence in (7b) must necessarily pertain between those present. I have chosen, however, to omit (7) from the tokens examined in Section 4, as one could plausibly interpret the French version as referring to alternative sources of silence. Of a total of 423 tokens of *mellom*, there were 30 in which the ‘betweenness’ predication could be construed as absent in either the French or English text, including two in which it was absent in both.

Before proceeding to an examination of the data, two points should be aired regarding the suitability of the 3-text approach employed in this study. One of these is practical, the other theoretical. The practical question concerns the possibility that we are actually engaged in exploring what is either, in whole or in part, a 2-text rather than a 3-text corpus. In other words, did one or more of either the French or English translators make use of a prior translation into the other language? Signe Oksefjell Ebeling, the compiler of the corpus, thinks it unlikely that this was the case, at least with respect to the French and English versions (personal