

# Les corpus bilingues et la linguistique contrastive. Une étude de constructions causatives basée sur un corpus parallèle néerlandais-suédois

Gudrun Rawoens

Département des langues scandinaves – Université de Gand – Rozier 44  
B-9000 Gand – Belgique

Centre for English Corpus Linguistics – Université Catholique de Louvain  
Place Blaise Pascal 1 – B-1348 Louvain-la-Neuve – Belgique

## Abstract

The use of multilingual corpora in contrastive linguistics has been growing the last few decades and their value for various types of semantic and syntactic linguistic studies as well as a great number of terminological studies has been emphasized in a great number of publications (e.g. Johansson and Oksefjell 1998; Altenberg and Granger 2002). This paper will first discuss a number of theoretical issues concerning contrastive linguistic research and emphasize the value of multilingual corpora in the detailed study of one particular language. As shown in Aijmer and Altenberg (1996: 12), Aarts (1998: Introduction) and Johansson (2000: 4), the value of a contrastive linguistic analysis is not only that it allows for cross-linguistic comparison, but also in that it allows us to trace details about one specific language that might not have been discovered without the comparative angle in the study. The second part of this paper will discuss a selection of results from a study of analytical causative constructions in a Dutch-Swedish contrastive perspective (Rawoens 2007). The study is based on a Dutch-Swedish parallel corpus comprising a total of three million words. One result from this investigation is that the relation between the analytical causative constructions and their translations is asymmetrical. Whereas synthetic causatives appear to be more common in Swedish, Dutch uses more analytical causative constructions and more non-verbal causal constructions.

## Résumé

Récemment, l'utilisation de corpus multilingues dans les études contrastives a connu un succès grandissant. La valeur ajoutée des corpus multilingues pour un éventail d'études sémantiques et syntaxiques, ainsi que pour des études terminologiques, a été soulignée dans de nombreuses publications (e.g. Johansson et Oksefjell 1998, Altenberg et Granger 2002). L'étude dont il est question ici vise tout d'abord à discuter quelques questions théoriques concernant la recherche en linguistique contrastive et à souligner l'importance de l'emploi des corpus multilingues pour l'analyse détaillée de langues particulières. Aijmer et Altenberg (1996 : 12), Aarts (1998 : Introduction) et Johansson (2000 : 4) ont mis en avant que la valeur d'une analyse contrastive n'est pas uniquement liée au fait qu'elle permet de comparer plusieurs langues, mais aussi au fait qu'elle permet de déceler des détails d'une langue qui n'auraient pas été découverts sans l'aide d'une perspective contrastive. La deuxième partie de cet article portera sur les résultats d'une analyse contrastive de constructions causatives analytiques en néerlandais et en suédois (Rawoens 2007). L'analyse est basée sur un corpus parallèle bidirectionnel suédois-néerlandais comprenant trois millions de mots. Les résultats suggèrent que la relation entre les constructions causatives analytiques et leurs traductions est asymétrique. Il apparaît que les causatives synthétiques sont plus courantes en suédois, tandis que les constructions analytiques ainsi que les constructions non-verbales sont plus fréquentes en néerlandais.

**Mots-clés :** corpus multilingues, études contrastives.

## 1. Introduction

Selon Aijmer et Altenberg (1996 : 12), Aarts (1998 : Introduction) et Johansson (2000 : 4) l'intérêt d'une analyse contrastive ne réside pas uniquement dans la possibilité de comparer plusieurs langues : une perspective contrastive permet également de révéler certains détails d'une langue qui n'auraient pas été découverts autrement. En outre, une telle approche est particulièrement intéressante dans le cas d'une étude de deux langues affines telles que le néerlandais et le suédois, notamment parce que des différences subtiles entre ces deux langues ne peuvent être décelées que grâce à une comparaison basée sur un grand nombre de données.

Il est de notoriété publique que l'utilisation des données empiriques pour la recherche linguistique a connu un succès grandissant depuis que des grands corpus sont devenus plus accessibles (p.ex. McEnery et Wilson 2001). Les corpus peuvent être utilisés dans un grand nombre de domaines de recherche, ainsi que pour un éventail d'applications différentes. Ils peuvent servir comme des données de base qui permettent de formuler des hypothèses (l'approche qu'on appelle *corpus-driven*), ou ils peuvent être utilisés comme un moyen pour vérifier des hypothèses (l'approche qu'on appelle *corpus-based*) (Ooi, 1998 : 52).

Les avantages et désavantages de l'utilisation des corpus pour des études linguistiques ont été largement discutés. Un des avantages manifestes est que les corpus mettent à disposition du linguiste une grande quantité de données de langue authentique, ce qui lui permet d'effectuer la recherche linguistique d'une façon beaucoup plus objective que s'il en était réduit à sa propre compétence linguistique. Cependant, il est recommandé que le linguiste n'ait pas une confiance absolue dans les corpus. En d'autres termes, l'étude linguistique devrait être basée sur des corpus (*corpus-based*) et non déterminée par des corpus (*corpus-bound*) (Summers 1996 : 266). En fait, l'approche idéale consiste à mélanger plusieurs méthodes : la méthode basée sur des corpus en combinaison avec l'introspection et l'élicitation.

Depuis quelques décennies, l'utilisation de corpus multilingues dans les études contrastives a connu un succès croissant et leur valeur ajoutée pour un éventail d'études sémantiques et syntaxiques, ainsi que pour des études terminologiques, a été soulignée dans de nombreuses publications (e.g. Johansson et Oksefjell 1998 ; Altenberg et Granger 2002). Comme le font remarquer Altenberg et Granger (2002 : 14), les avantages tangibles des corpus multilingues pour des études contrastives sont, entre autres, qu'ils dévoilent des notions qui n'auraient pas été découvertes dans une étude basée sur des corpus monolingues. Les corpus multilingues favorisent aussi la découverte de différences entre les textes sources et les textes cibles. Par ailleurs, ce type de corpus peut être utilisé dans de nombreuses applications, entre autres la lexicographie, l'enseignement des langues et la traduction.

Plusieurs types de corpus multilingues peuvent être distingués. Il y a une certaine incohérence en ce qui concerne la terminologie (voir p.ex. Johansson 1998 : 4–5 à ce propos). Dans la typologie suivante, basée sur Altenberg et Granger (2002 : 8), on peut faire une différence entre les corpus de traduction et les corpus comparables (Figure 1). Les corpus de traduction comprennent des textes originaux et leurs traductions vers une ou plusieurs langues. Ces corpus peuvent être unidirectionnels ou bidirectionnels, selon la présence de traductions dans uniquement une direction ( $L1 > L2$ ) ou dans deux directions ( $L1 > L2$  et  $L2 > L1$ ). Les corpus comparables ne contiennent pas de traductions, mais des textes dans deux ou plusieurs langues qui sont comparables au niveau du genre ou du champ de spécialité.

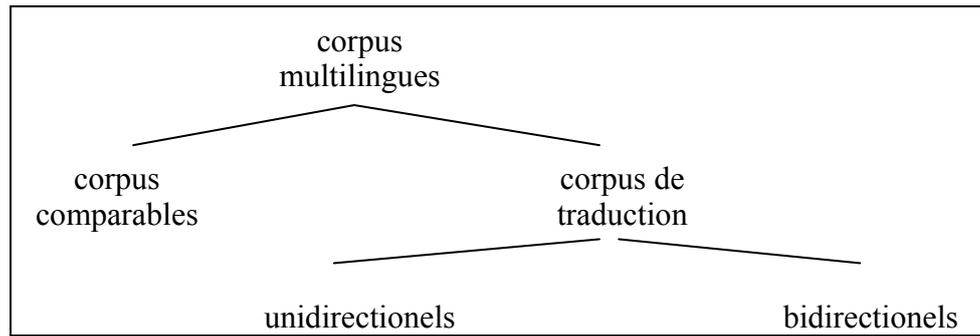


Figure 1. Une typologie de corpus multilingues (Altenberg et Granger 2002 : 7)

Il existe donc plusieurs types de corpus et il y a des avantages et des désavantages liés à chaque type. Cela implique que le choix du 'bon' corpus dépend du type de recherche visé. De plus, il est aussi important que les résultats de l'analyse soient liés au type du corpus.

L'avantage des corpus comparables est qu'ils contiennent uniquement des textes originaux, c'est-à-dire pas de traductions. Ce type de corpus est particulièrement approprié pour des études terminologiques (Lauridsen 1996). Cependant, de tels corpus sont difficiles à compiler, et par conséquent, ils sont rarement équilibrés dans leur contenu. Dans beaucoup de cas il est difficile de trouver, pour deux langues spécifiques, des corpus qui soient comparables en termes de fonction et de style (Aijmer et Altenberg 1996 : 13 ; Johansson 1998 : 5 ; Altenberg et Granger 2002 : 8-9, 13). En d'autres termes, il est très difficile, voire impossible, de trouver des équivalents cross-linguistiques dans un corpus comparable, ce qui explique pourquoi ce type de corpus n'est pas un moyen approprié à la comparaison de deux langues.

Si le linguiste veut étudier des équivalents de traduction, les corpus de traduction offrent un plus grand éventail de possibilités (Chesterman 1998 : 90 ; James 1980 : 67, 178 ; Johansson 2000 : 4). Dans les traductions il s'agit des similarités ou des équivalents et de la façon de transférer un élément particulier d'une langue à l'autre. C'est ce que Chesterman (1998 : 17) appelle « the translation identity assumption ». Cependant, les corpus de traduction connaissent aussi un nombre de désavantages. Un désavantage bien connu est que le texte cible peut être influencé par le texte source. Ces signes d'influence sont désignés par le terme *translationese* (Gellerstam 1996 : 53-54). En outre, le texte cible peut aussi montrer des signes qui sont typiques d'un texte traduit (James 1980 : 117-118 ; Aijmer et Altenberg 1996 : 13 ; Aarts 1998 : Introduction ; Johansson 1998 : 5). Altenberg et Granger (2002 : 19) soulignent également qu'il y a rarement une correspondance de cent pourcent entre les équivalents de traduction dans les corpus de traduction. Par ailleurs, il y a rarement une distribution égale entre les équivalents de traduction dans deux langues (Altenberg et Granger 2002 : 18). Pour apprendre s'il y a un haut degré de correspondance entre les langues en question, on pourrait calculer ce qu'on appelle la correspondance mutuelle dans le corpus de traduction. Un degré bas de correspondance mutuelle peut être expliqué par une série de facteurs comme la polysémie divergente ou des différences de système pragmatique (Altenberg et Granger 2002 : 19-21). Un autre désavantage des corpus de traduction est qu'ils sont rarement bien équilibrés à cause du fait que certains genres ou types de textes sont souvent surreprésentés (Johansson, 1998 : 6 ; Altenberg et Granger 2002 : 9). De plus, les corpus de traduction sont limités à la langue écrite (Johansson 1998 : 6 ; Altenberg et Granger 2002 : 9).

La conclusion d'Altenberg et Granger (2002 : 9) est qu'il faudrait considérer les deux types de corpus multilingues comme une ressource complémentaire dans les études contrastives. En

d'autres mots, on pourrait utiliser un corpus de traduction et un corpus comparable en même temps, en profitant des avantages respectifs de ces deux types de corpus.

## 2. Le SALT Dut-Swe corpus

Cette étude est basée sur un corpus parallèle néerlandais-suédois, le SALT Dut-Swe corpus. Ce corpus a été compilé en coopération avec des chercheurs de l'université de Gothenburg dans le cadre du projet SALT (Språkbankens Arkiv för Länkade Texter, 'archives de textes alignés de Språbanken'), un projet réalisé à l'initiative de Språbanken ('la banque du suédois'). Språbanken peut compter sur une longue tradition et expertise en ce qui concerne la compilation de corpus suédois monolingues, remontant aux années soixantes, quand le corpus suédois Press 65, d'un million de mots, a été compilé par Sture Allén. Språbanken a été établi en 1975 comme un centre national visant à rassembler des bases de données de corpus suédois. A l'heure actuelle, les corpus de Språbanken continuent de se développer et contiennent déjà plus de cent millions mots. Le projet SALT a été lancé en 2001 et trouve son origine dans la nécessité de créer un nombre de corpus bilingues avec le suédois comme langue centrale dans le but de les rendre accessibles aux chercheurs dans les domaines de la linguistique, pédagogie et applications technologiques comme CAT (*computer assisted translation*). Les corpus SALT contiennent des textes originaux en suédois et leurs traductions dans une autre langue (p.ex. le russe ou le néerlandais), ainsi que des textes originaux dans les langues étrangères et leurs traductions vers le suédois. De cette façon, les corpus SALT sont une combinaison de corpus de traduction et de corpus comparables. Toutefois, ils sont renseignés comme des corpus parallèles. Plusieurs styles et genres différents sont représentés dans le corpus néerlandais-suédois. Alors que les autres corpus SALT contiennent uniquement des textes de fiction, le corpus néerlandais-suédois contient également d'autres textes pour obtenir une plus grande variété dans la langue. Vingt auteurs différents et treize traducteurs sont représentés dans le corpus. Le corpus contient trois millions de mots et les textes sont alignés au niveau des phrases. La structure bidirectionnelle permet différentes sortes d'analyses. Non seulement est il possible de comparer les textes sources avec les textes cibles, mais il est aussi possible de comparer les textes sources dans les deux langues (comme dans un corpus comparable) ou de comparer les textes sources avec les textes cibles dans une des deux langues (Figure 2). Ce corpus bilingue peut être utilisé dans un éventail d'études contrastives lexicales, syntaxiques ou pragmatiques. De plus, il peut aussi être utilisé pour des objectifs pédagogiques ou pour des objectifs nettement quantitatifs.

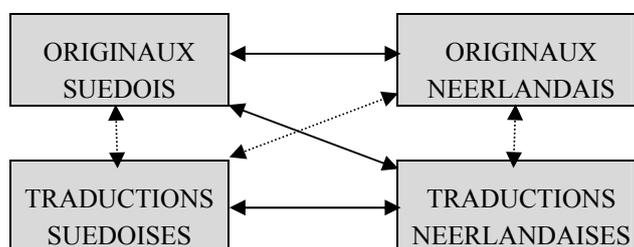


Figure 2. La structure du SALT Dut-Swe corpus

## 3. Les constructions causatives en néerlandais et suédois

La deuxième partie de cet article sera consacrée à la présentation d'une sélection de résultats d'une analyse contrastive de constructions causatives analytiques en néerlandais et suédois basée sur le SALT Dut-Swe corpus. L'étude qui est présentée ici fait partie d'une étude plus étendue des constructions causatives analytiques en néerlandais et suédois (Rawoens 2007) où

l'approche est bidirectionnelle. L'approche dans l'étude élaborée ici est limitée à une dimension, c'est-à-dire que l'étude vise à analyser les traductions en partant des constructions dans les textes originaux.

Les constructions causatives analytiques consistent en un prédicat causal et un prédicat d'effet (p.ex. Comrie 1985 ; Verhagen et Kemmer 1997). Les structures peuvent varier selon le type de prédicat causal et le type de complément. La présente étude traite des constructions dont le prédicat causal est un verbe grammaticalisé suivi d'une phrase infinitive. En néerlandais les prédicats causaux qui apparaissent dans de telles constructions sont *doen* ('faire') et *laten* ('laisser, faire'). Les structures des constructions sont comme suit, illustrées par des exemples du corpus parallèle.

- *doen* + NP + V<sub>inf</sub>  
Er klonk een aarzelend applaus dat hem verlegen naar zijn stoel terug *deed* lopen. (WoIN1)  
'Le son des applaudissements hésitants l'embarrassa et le fit retourner vers sa chaise.'
- *laten* + NP + V<sub>inf</sub>  
Bij 8-7 voor Frankrijk *liet* ze hen van speelhelft wisselen. (KraN1)  
'A 8-7 pour la France elle les fit échanger de côté.'

Une différence sémantique importante entre les deux verbes est que *doen* est utilisé pour exprimer des relations causales qui sont relativement simples et directes. *Laten* par contre exprime des relations causales plutôt indirectes (Verhagen et Kemmer 1997 : 70). Le sujet de *doen* peut être humain ou non-humain, tandis que le sujet de *laten* est normalement humain. Dans certains cas la différence entre les deux verbes est très claire, dans d'autres les différences sont plus subtiles et dépendent du style ou du registre. Dans quelques contextes limités les deux verbes sont interchangeable.

Les équivalents directs de *doen* et *laten* en suédois sont les verbes *få* (litt. 'recevoir', caus. 'faire') and *låta* (litt. 'laisser', caus. 'faire') respectivement. En outre, les verbes *komma* (litt. 'venir', caus. 'faire'), *ha* (litt. 'avoir', caus. 'faire') et *förmå* (litt. 'persuader', caus. 'faire') peuvent également figurer dans ces constructions. Leurs structures syntactiques sont les suivantes :

- *få* + NP + (*till*) + *att* + V<sub>inf</sub>  
Jag visste [...] att jag *fick* människor att göra som jag ville, [...]. (BerS1)  
'Je savais que je faisais faire aux gens ce que je voulais qu'ils fassent.'
- *låta* + NP + V<sub>inf</sub> Jag *lät* skriva ut materialet. (BerS1)  
'J'ai fait taper le matériel.'
- *komma* + NP + (*till*) + *att* + V<sub>inf</sub>  
Där fanns särskilt en bild som *kom* mig att yla av sorg. (BerS1)  
'Il y avait une image en particulier qui me fit pleurer de chagrin.'
- *ha* + NP + (*till*) + *att* + V<sub>inf</sub>  
Hon *hade* honom att bygga ett nytt garage. (SAG)  
'Elle le fit construire un nouveau garage.'

- *förmå* + NP + (*till*) + *att* + V<sub>inf</sub>  
*Jag förmådde* honom att söka psykiatrisk vård – ingenting hjälpte. (BerS1)  
 ‘Je le fis chercher de l’aide psychiatrique – rien n’aïda.’

La nature des différences entre les causatives en suédois diffère de celle des différences entre les causatives en néerlandais. Les verbes *få* et *låta* ne sont jamais interchangeables comme les verbes *doen* et *laten*. Par contre, les verbes *komma*, *ha* et *förmå* peuvent être utilisés comme des alternatives au verbe *få*. Ils sont pourtant beaucoup moins fréquents que *få* à cause d’une série de restrictions sémantiques et stylistiques (SAG 1999 ; Rawoens 2007) : ils peuvent exprimer différents degrés de contrainte et de modalité et ils sont souvent plus formels. Les différents verbes sont aussi typiquement combinés avec différents sujets. Le sujet de *få* peut être humain ou non-humain. *Komma* est combiné avec des sujets non-humains, tandis que les sujets de *låta*, *ha* et *förmå* sont normalement humains.

Nous ne discuterons pas plus en détail des différences sémantiques ou stylistiques entre les constructions en question. L’objectif principal de l’étude contrastive qui est présentée ici est d’analyser les traductions suédoises des constructions causatives analytiques en néerlandais et l’inverse. Le point de départ de l’étude sont les structures de traduction suivantes, qui sont représentés par trois catégories principales, CAT1, CAT2 et CAT3. CAT1 contient les cas où une construction causative analytique avec un complément infinitif est traduite par le même type de construction. CAT2 contient les cas où une construction causative analytique avec un complément infinitif est traduite par un autre type de construction causative analytique, p.ex. où le complément est une proposition subordonnée (CAT2ana), ou par un verbe lexical synthétique (CAT2syn). CAT3 contient également deux sous-catégories : soit une construction causative analytique est traduite par une construction causale non-verbal, p.ex. une conjonction ou une préposition (CAT3cau), soit il n’y a pas de traduction (CAT3Ø).

L’extraction des constructions causatives du corpus a été faite de façon semi-automatique, avec d’abord une extraction automatique basée sur les différentes formes morphologiques des verbes causatifs, suivie d’une phase d’édition manuelle pour supprimer les phrases non pertinentes. Pendant le processus d’analyse, le logiciel ParaConc (Barlow 1999) a été utilisé. Les résultats de l’analyse ont été testés statistiquement au moyen du Chi-carré. Les textes sources contiennent 797 phrases comprenant des constructions causatives analytiques, à savoir 374 phrases dans les textes néerlandais et 423 dans les textes suédois. Comme indiqué par les fréquences relatives dans le Tableau 1, les constructions causatives analytiques sont plus fréquentes dans les textes néerlandais que dans les textes suédois. Dans les textes néerlandais les constructions avec *laten* (301 occurrences) sont beaucoup plus fréquentes que les constructions avec *doen* (73). Dans les textes suédois les constructions avec *få* sont les plus fréquentes (215 occurrences), suivies par les constructions avec *låta* (165). Les verbes *komma* (31), *förmå* (10) et *ha* (2) sont beaucoup moins fréquents. Les proportions dans les textes néerlandais et suédois sont comparables aux proportions qui ont été relevées dans d’autres analyses monolingues de ces constructions en néerlandais et en suédois (Verhagen et Kemmer 1997 ; Rawoens 2007).

	Fréquences absolues	Fréquences relatives (par 100,000 mots)
Originaux suédois (1 026 003 mots)	423	41,23
Originaux néerlandais (429 410 mots)	374	87,10
Total (1 455 413 mots)	797	54,76

Tableau 1. Fréquences absolues et relatives des constructions causatives analytiques dans les textes originaux en suédois et en néerlandais

Les résultats de l'analyse des traductions suédoises (S) des constructions causatives analytiques en néerlandais (N) avec *doen* et *laten* (Tableau 2) révèlent que les constructions en néerlandais sont traduites par des constructions correspondantes, à savoir des constructions avec *få*, *komma*, *ha*, *förmå* ou *låta* dans quarante quatre pourcent des cas (CAT1). Nous retrouvons une proportion similaire dans la deuxième catégorie (CAT2), qui représente les traductions où sont utilisées des constructions alternatives verbales. Douze pourcent du total sont des traductions où des constructions non-verbales sont utilisées ou où il n'y a pas de traduction (CAT3).

CAT1	CAT2		CAT3		Total N > S
	CAT2ana	CAT2syn	CAT3cau	CAT3Ø	
165	5	159	3	42	
165	164		45		374
44,1%	43,9%		12%		100%

Tableau 2. Les traductions N > S

Les exemples (1) et (2) illustrent des traductions qui sont classifiées dans CAT1 : dans la première phrase une construction causative analytique en néerlandais avec *doen* est traduite en suédois par une construction causative analytique avec *få* et dans la deuxième phrase une construction avec *laten* est traduite par une construction avec *låta*.

- (1) Ik heb iets bij me dat uw hartje sneller zal *doen* kloppen. (ClaN1)  
 Jag har med mig något som ska *få* ditt hjärta att slå snabbare. (ClaS2)  
 'J'ai quelque chose avec moi qui va faire battre ton coeur plus vite.'
- (2) Om uit de voeten te kunnen met stok of looprek heeft zij alle overtollige stoelen en prullen door Oscar naar boven *laten* brengen. (EnqN1)  
 För att hon ska kunna ta sig fram med käpp eller rollator har hon *låtit* Oscar bära upp alla överflödiga stolar och prylar på övervåningen. (EnqS2)  
 'Afin de pouvoir se déplacer avec une canne ou une tribune, il fit porter en haut par Oscar toutes ses chaises superflues et ses babioles.'

Dans la deuxième catégorie CAT2 il y a une répartition déséquilibrée frappante entre les sous-catégories CAT2ana et CAT2syn (Tableau 2). CAT2ana contient uniquement cinq cas, tandis

que CAT2syn contient 159 cas, ce qui représente presque quarante trois pourcent du nombre total des traductions N > S.

Il est remarquable que la plupart des traductions dans la sous-catégorie CAT2syn sont des traductions des constructions avec le verbe *laten*, comme illustré dans l'exemple (3), où la construction néerlandaise *laat...opstellen* 'faire... établir' est traduite par le verbe causatif synthétique *sätter upp* 'établir'.

- (3) Ik ben blij als jullie er komen wonen, ik *laat* een huurcontract opstellen zodat je zekerheid hebt. (EnqN1)  
 Jag blir glad om ni flyttar in, jag sätter upp ett hyreskontrakt så att du har någon säkerhet. (EnqS2)  
 'Je serai content si vous emménagez, je ferai établir un contrat de location afin de vous garantir une certaine assurance.'

Le grand nombre d'occurrences dans cette sous-catégorie semble être lié au fait que les causatives synthétiques expriment souvent la signification causative d'une façon beaucoup plus tangible que si une construction causative analytique avait été employée. Ceci vaut surtout pour les constructions avec *laten*, ce qui peut être expliqué par le fait que les constructions avec *laten* peuvent aussi exprimer la permissivité. Si le traducteur choisit de traduire une construction avec *laten* par une causative synthétique, il souligne ainsi la signification causative et enlève en plus tout risque d'ambiguïté. Une traduction plus littérale avec une construction causative analytique, bien que possible, est souvent plus laborieuse en suédois. Ceci vaut par exemple pour l'exemple (4) où la construction néerlandaise *laten zien* 'faire voir, montrer' aurait pu être traduite en suédois par la construction *låta se*. Mais cette traduction aurait été soit considérée comme artificielle, soit mal interprétée comme permissive. A la place, le traducteur a choisi un verbe synthétique approprié, à savoir *visa* 'montrer', qui est l'expression la plus naturelle dans la langue cible.

- (4) Hij *laat* haar het kasteel van binnen zien, zij kijkt haar ogen uit. (ClaN1)  
 Han visar henne slottet från insidan, hon tror inte sina ögon. (ClaS2)  
 'Il lui montre le château de l'intérieur, elle ne peut pas en croire ses yeux.'

En ce qui concerne la troisième catégorie, CAT3, la sous-catégorie CAT3cau contient uniquement trois cas, tandis que la sous-catégorie CAT3Ø en contient quarante-deux. Parmi ceux-ci, neuf représentent des cas où la traduction est absente. Les autres sont des cas où les relations causales sont formulées d'une façon implicite.

Au premier abord, les traductions néerlandaises des constructions causatives analytiques suédoises manifestent une distribution comparable aux traductions suédoises des constructions néerlandaises dont nous avons discuté plus haut. Pourtant, si nous considérons aussi les sous-catégories, il apparaît qu'il y a des différences remarquables (Tableau 3).

CAT1	CAT2		CAT3		Total S > N
	CAT2ana	CAT2syn	CAT3cau	CAT3Ø	
	60	79	43	41	
200	139		84		423
47,2%	32,9%		19,9%		100%

Tableau 3. Les traductions S > N

Dans la première catégorie CAT1 il y a une majorité frappante de cas où une construction causative analytique avec *låta* est traduite par une construction causative analytique avec *laten*, comme dans l'exemple (5). Il est aussi intéressant de noter que ce groupe représente presque trente pourcent du total des traductions.

- (5) Jag *lät* skriva ut materialet. (BerS1)  
Ik liet het materiaal uittikken. (BerN2)  
'J'ai fait taper le matériel.'

Dans CAT2 il y a une distribution interne égale entre les deux sous-catégories : CAT2ana contient soixante occurrences et CAT2syn, soixante dix-neuf. En d'autres termes, le nombre de cas où une construction causative analytique est traduite par une causative synthétique est beaucoup plus bas que dans la catégorie équivalente dans les traductions néerlandaises.

La catégorie CAT3 représente environ un cinquième de toutes les traductions en néerlandais, ce qui rend cette catégorie plus étendue que la catégorie correspondante dans les traductions en suédois. La plupart des occurrences dans cette troisième catégorie appartiennent à la sous-catégorie CAT3cau comme dans l'exemple (6), où la préposition causale *van* 'de' est utilisée dans la traduction néerlandaise. La traduction littérale de la phrase néerlandaise est 'Il devenait malade de la lumière et du son.'

- (6) Ljus och ljud *fick* honom att må illa. (EkmS1)  
Van licht en geluid werd hij misselijk. (EkmN2)  
'La lumière et le son le rendaient malade.'

De plus, la sous-catégorie CAT3cau représente environ une dixième de toutes les traductions en néerlandais, ce qui est beaucoup plus que dans les traductions en suédois où cette catégorie constitue moins d'un pourcent de toutes les traductions vers le suédois. Une explication pour le grand nombre de constructions non-verbales dans les traductions néerlandais peut être que les conjonctions et adverbiaux causaux sont considérés comme les façons les plus communes pour exprimer des relations causales en néerlandais en général (Degand 2001).

#### 4. Conclusion

Dans cet article nous avons motivé le choix d'une approche contrastive et basée sur des corpus dans la recherche linguistique. Ensuite, nous avons présenté les résultats d'une analyse contrastive de constructions causatives analytiques en néerlandais et en suédois basée sur le SALT Dut-Swe corpus. Les résultats de l'étude ont montré que la relation entre les constructions causatives analytiques et leurs traductions est asymétrique. Alors que les causatives synthétiques sont plus courantes en suédois, en néerlandais ce sont les constructions causatives analytiques – surtout celles avec *laten* – ainsi que les constructions causales non-verbales qui sont les plus courantes. Les résultats ont également montré que les traducteurs choisissent souvent des traductions qui impliquent un haut degré d'explicitation. Ces résultats ont été confirmés par une analyse de corpus plus vaste, où les autres dimensions du corpus ont aussi été considérées (Rawoens 2007).

#### Références

- Aarts J. (1998). Introduction. In Johansson S. and Oksefjell S. editors, *Corpora and Cross-linguistic Research : Theory, Method, and Case Studies*. Amsterdam, Rodopi.

- Aijmer K. and Altenberg B. (1996). Introduction. In Aijmer K., Altenberg B. and Johansson M. editors, *Languages in Contrast. Papers from a symposium on text-based cross-linguistic studies in Lund*, 4-5 March 1994. Lund, Lund University Press, pages 11-16.
- Altenberg B. and Granger S. (2002). Recent trends in cross-linguistic lexical studies. In Altenberg B. and Granger S. editors, *Lexis in Contrast. Corpus-based approaches*. Amsterdam/Philadelphia, John Benjamins Publishing Company, pages 1-48.
- Barlow M. (1999). MonoConc 1.5 and ParaConc. *International Journal of Corpus Linguistics*, 4 (1). Amsterdam/Philadelphia, John Benjamins Publishing Company, pages 319-327.
- Chesterman A. (1998). *Contrastive functional analysis*. Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Comrie B. (1985). Causative verb formation and other verb-deriving morphology. In Shopen T. editor, *Language typology and syntactic description*. Cambridge, Cambridge University Press, pages 309-348.
- Degand L. (2001). *Form and Function of Causation. A Theoretical and Empirical Investigation of Causal Constructions in Dutch*. Leuven, Peeters.
- Gellerstam M. (1996). Translations as a source for cross-linguistic studies. In Aijmer K., Altenberg B. and Johansson M. editors, *Languages in Contrast*. Lund, Lund University Press, pages 53-62.
- James C. (1980). *Contrastive analysis*. London, Longman.
- Johansson S. (1998). On the role of corpora in cross-linguistic research. In Johansson S. and Oksefjell, S. editors, *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*. Amsterdam, Rodopi, pages 3-24.
- Johansson S. (2000). Contrastive Linguistics and Corpora no.3, *SPRIKreports*. Available on-line from <http://www.hf.uio.no/german/sprik/> (accessed: 28 October 2008).
- Johansson S. and Oksefjell S. editors (1998). *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*. Amsterdam, Rodopi.
- Lauridsen K. M. (1996). Text corpora and contrastive linguistics: Which type of corpus for which type of analysis? In Aijmer K., Altenberg B. and Johansson M. editors, *Languages in Contrast*. Lund, Lund University Press, pages 63-71.
- McEnery T. and Wilson A. (2001). *Corpus Linguistics*. Edinburgh, Edinburgh University Press.
- Ooi V. B.Y. (1998). *Computer Corpus Lexicography*. Edinburgh, Edinburgh University Press.
- Rawoens G. (2007). *Kausativa verbkonstruktioner i svenskan och nederländskan. En korpusbaserad syntaktisk-semantisk undersökning*. PhD dissertation, Ghent University.
- SAG I-IV = Teleman, U., S. Hellberg and E. Andersson (1999). *Svenska Akademiens grammatik*. Stockholm, Svenska Akademien.
- Summers D. (1996). Computer lexicography: the importance of representativeness in relation to frequency. In Thomas J. and Short M. editors, *Using corpora for language research: studies in the honour of Geoffrey Leech*. London, Longman, pages 260-266.
- Verhagen A. and Kemmer S. (1997). Interaction and causation: Causative constructions in modern standard Dutch. *Journal of Pragmatics* 27, pages 61-82.