

Inhaltsverzeichnis

0.	Vorwort	9
1.	Kausalmarker und Diskursrelationen	13
1.1	Erklärungsmodelle für die Variation in der Kodierung von kausaler Information	13
1.2	Das Inventar an Kausalmarkern	20
1.3	Das Inventar an Diskursrelationen	22
1.4	Probleme bei der Analyse der Relationen in einem Korpus	28
1.5	Korrelationstendenzen bei der Kodierung kausaler Information.....	32
2.	Datenbasis und Methode	37
2.1	Das Deutsche Referenzkorpus DeReKo	37
2.2	Das Untersuchungsdesign	38
3.	Das Analyseraster: Vorstellung der Analyseparameter	47
3.1	Die formalen Merkmale	48
3.1.1	Konnexionsspezifische Merkmale	49
3.1.1.1	Das Merkmal POSITION DES KONNEKTORS	49
3.1.1.2	Das Merkmal LINEARISIERUNG DER KONNEKTE.....	54
3.1.1.3	Das Merkmal INTEGRATION	56
3.1.1.4	Das Merkmal KORRELATKONSTRUKTION	58
3.1.1.5	Das Merkmal MEHRFACHES AUFTRETEN VON MARKERN	59
3.1.2	Eigenschaften der Konnekte	61
3.1.2.1	Das Merkmal SATZMODUSTYP	61
3.1.2.2	Das Merkmal VERBSTELLUNGSTYP	63
3.1.2.3	Das Merkmal SATZFÖRMIGKEIT.....	64
3.1.2.4	Das Merkmal UMFANG DER KONNEKTE	67
3.1.2.5	Das Merkmal DEFINITHEIT DER KONNEKTE.....	68
3.1.2.6	Das Merkmal KONJUNKTIV	69
3.1.2.7	Das Merkmal AUSDRUCK DER 1. PERSON	69
3.1.2.8	Das Merkmal SPRECHEREINSTELLUNG.....	70

3.2	Die funktionalen Merkmale	71
3.2.1	Das Merkmal AGENS.....	74
3.2.2	Das Merkmal EXPERIENCER	75
3.2.3	Das Merkmal NONFAKT.....	77
3.2.4	Das Merkmal WERTUNG.....	79
3.2.5	Das Merkmal FREMDPERSPEKTIVE	80
3.3	Die Klassifikation und Bestimmung der Diskursrelationen	82
3.3.1	Die Diskursrelation PRAGMATIC CLAIM.....	82
3.3.2	Die Diskursrelation VOLITIONAL CAUSE	84
3.3.3	Die Diskursrelation NONVOLITIONAL CAUSE.....	86
3.3.4	Die Diskursrelation PURPOSE	87
4.	Auswertung der Ergebnisse: signifikante und nicht signifikante Korrelationen.....	89
4.1	Korrelationen bezogen auf die einzelnen Marker	89
4.1.1	Korrelationen zwischen Markern und Diskursrelationen	89
4.1.1.1	Globales Bild	89
4.1.1.2	Die „pragmatischen“ Diskursrelationen PRAGMATIC CLAIM I/II	94
4.1.1.3	Die „semantischen“ Diskursrelationen VOLITIONAL CAUSE und NONVOLITIONAL CAUSE	99
4.1.1.4	Die Diskursrelation PURPOSE	103
4.1.2	Korrelationen zwischen Markern und Formmerkmalen	105
4.1.2.1	Korrelationen mit der POSITION DES KONNEKTORS	105
4.1.2.2	Korrelationen mit der LINEARISIERUNG DER KONNEKTE	107
4.1.2.3	Korrelationen mit dem Merkmal INTEGRATION	108
4.1.2.4	Korrelationen mit dem Merkmal KORRELAT- KONSTRUKTION	109
4.1.2.5	Korrelationen mit dem Merkmal MEHRFACHES AUFTRETEN VON MARKERN.....	110
4.1.2.6	Korrelationen mit dem SATZMODUSTYP	111
4.1.2.7	Korrelationen mit dem VERBSTELLUNGSTYP.....	115

4.1.2.8	Korrelationen mit der SATZFÖRMIGKEIT	117
4.1.2.9	Korrelationen mit dem UMFANG DER KONNEKTE	120
4.1.2.10	Korrelationen mit der DEFINITHEIT DER KONNEKTE	120
4.1.2.11	Korrelationen mit dem Merkmal KONJUNKTIV	121
4.1.2.12	Korrelationen mit dem AUSDRUCK DER 1. PERSON	123
4.1.2.13	Korrelationen mit Ausdrücken der SPRECHER- EINSTELLUNG	124
4.1.3	Korrelationen zwischen Markern und funktionalen Merkmalen	125
4.1.3.1	Korrelationen mit den rollensemantischen Merkmalen AGENS im KONSEQUENS und EXPERIENCER im KONSEQUENS	125
4.1.3.2	Korrelationen mit dem Merkmal NONFAKT	127
4.1.3.3	Korrelationen mit dem Merkmal WERTUNG	129
4.1.3.4	Korrelationen mit dem Merkmal FREMDPERSPEKTIVE	131
4.2	Korrelationen bezogen auf die einzelnen Diskursrelationen	133
4.2.1	Globales Bild	133
4.2.2	Korrelationstabellen für Korrelationen zwischen Diskursrelationen und Formmerkmalen	142
4.2.2.1	Korrelationen mit der POSITION DES KONNEKTORS	143
4.2.2.2	Korrelationen mit der LINEARISIERUNG DER KONNEKTE	144
4.2.2.3	Korrelationen mit dem Merkmal INTEGRATION	146
4.2.2.4	Korrelationen mit KORRELATKONSTRUKTIONEN	146
4.2.2.5	Korrelationen mit dem SATZMODUSTYP	147
4.2.2.6	Korrelationen mit dem Merkmal SATZFÖRMIGKEIT	149
4.2.2.7	Korrelationen mit dem Merkmal KONJUNKTIV	152
4.2.2.8	Korrelationen mit dem AUSDRUCK DER 1. PERSON	153
4.2.2.9	Korrelationen mit Ausdrücken der SPRECHEREINSTELLUNG	154
4.2.3	Korrelationstabellen für die Korrelationen zwischen Diskursrelationen und funktionalen Merkmalen	155
4.2.3.1	Korrelationen mit dem Merkmal AGENS im KONSEQUENS	155

4.2.3.2	Korrelationen mit dem Merkmal EXPERIENCER im KONSEQUENS	156
4.2.3.3	Korrelationen mit dem Merkmal NONFAKT	156
4.2.3.4	Korrelationen mit dem Merkmal WERTUNG	158
4.2.3.5	Korrelationen mit dem Merkmal FREMDPERSPEKTIVE	160
5.	Von Gründen und Folgen: Zusammenfassung und Ausblick	165
6.	Literatur	173
7.	Anhang: Verzeichnis der verwendeten Quellen	183
7.1	Zeitungen und Zeitschriften	183
7.2	Belletristik, Sachliteratur und Sonstiges	184
8.	Verzeichnis der Abbildungen und Tabellen	187

0. Vorwort

Der vorliegende Band präsentiert Ergebnisse des in den Jahren 2007 und 2008 von der DFG geförderten Gemeinschaftsprojekts *Kausalitätsmarker als Kohärenzmittel und ihre Formalisierung für die automatische Textanalyse* (Kurztitel: *Kausale Kohärenz*), das in der Abteilung Grammatik des Instituts für Deutsche Sprache Mannheim und am Institut für Angewandte Computerlinguistik der Universität Potsdam durchgeführt wurde (Fördernr. BR 3463/1-1 und STE 733/7-1). Gegenstand des Projekts waren spezifisch grammatische Ausdrucksmittel für die Herstellung kausaler Kohärenz in Texten: konjunktionale Konnektoren (*weil, da, denn, sodass, weshalb* etc.), Adverbkonnektoren (*deshalb, daher, darum, nämlich, schließlich* etc.) sowie Präpositionen (*wegen, aufgrund*). Diese „Kausalmarker“ etablieren in Texten verschiedene Diskursrelationen zwischen Textsegmenten, in aller Regel zwischen benachbarten Einheiten, ihren „Konnekten“. Wiewohl unter dem herkömmlichen semantischen Relationsbegriff Kausalität subsumierbar, sind sie in konkreten Textumgebungen in aller Regel keineswegs äquivalent und wechselseitig ersetzbar.

Die Faktoren herauszufinden, die die Variation in der Kodierung kausaler Information steuern, war ein wesentliches Anliegen des Projekts. Konkretes Ziel war es, auf der Basis einer empirisch gestützten Analyse dieser Kausalmarker in ihren jeweiligen Kontextumgebungen sowohl den idiosynkratisch-lexikalischen Beitrag der einzelnen Marker als auch den Anteil bestimmter Kontextcharakteristika zur Spezifizierung von Diskursrelationen zu eruieren. Im besonderen Rahmen des Gemeinschaftsprojekts versteht sich dieses Vorhaben vor dem Hintergrund der Einsicht, dass oberflächennahe Verfahren der automatischen Textanalyse mittlerweile an ihre Grenzen gestoßen sind und infolgedessen nun angestrebt wird, für computerlinguistische Anwendungen wie etwa automatische Informationsextraktion oder Textzusammenfassungen wieder verstärkt Erkenntnisse der qualitativen linguistischen Analyse nutzbar zu machen. Hierzu liegen nun mit dem Handbuch der deutschen Konnektoren (Pasch et al. 2003) und weiteren Arbeiten aus dem Umfeld des Konnektorenprojekts am IDS (siehe etwa Blühdorn/Breindl/Waßner (Hg.) (2004); Blühdorn (2008, demn., i. Vorb.); Breindl/Waßner (i. Vorb.); Volodina (2007)) neuere detaillierte Erkenntnisse zur Syntax und Semantik eines Großteils dieser Marker vor. Ein zweiter für unsere Untersuchung einschlägiger Forschungsstrang ist an der Universität Utrecht angesiedelt, wo Diskursrelationen aus

einer psycholinguistischen und kognitiven Perspektive in ihren Ausprägungen in verschiedenen europäischen Sprachen untersucht werden. Neben der Frage nach der Systematik oder „Taxonomie“ von Diskursrelationen (Knott/Sanders 1998) beschäftigt sich die Utrechter Forschergruppe in den letzten Jahren intensiv mit Zusammenhängen zwischen Diskursrelationen und der sprachlichen Kodierung der Dimension „Subjektivität“ (Pit 2003, 2007; Stukker 2005; Mulder 2008).

Die Brücke zwischen der computerlinguistischen Implementierung und der feinkörnigen grammatischen Analyse bilden empirische, korpusbasierte Arbeiten, die in der Lage sind, den tatsächlichen Sprachgebrauch abzubilden und „typische“ und „weniger typische“ Verwendungen von Kausalkonnektoren zu erfassen. Zu diesem Zweck werden potentiell relevante formale und funktionale Charakteristika dieser Kontexte erhoben, die jeweiligen Ausprägungen an einem Korpus für jeden Marker separat analysiert und die Daten miteinander korreliert. Eine solche Studie liegt mit Frohning (2007) für ein Teilinventar unseres Gegenstands vor. Frohning ermittelte schriftsprachliche Korrelationen der Kausalmarker mit einer großen Reihe von Oberflächen-Kontextmerkmalen, die zur Abgrenzung der einzelnen Marker beitragen. Was dabei außer Acht blieb, ist der Zusammenhang mit den Diskursrelationen.

Nicht zuletzt darin geht die vorliegende Studie über die von Frohning hinaus. Sie beschreibt, ausgehend von „harten“ grammatischen Eigenschaften der Kausalmarker (wie z.B. der Zugehörigkeit zu einer bestimmten syntaktischen Klasse und den sich daraus ergebenden Kontextrestriktionen) deren spezifischen Variationsspielraum in Bezug auf eine Reihe von Merkmalen (z.B. hinsichtlich der topologischen Konfiguration der verknüpften Konnekte) und ermittelt dann auf der Basis eines umfangreichen schriftsprachlichen Korpus Korrelationen zwischen Markern, Umgebungsmerkmalen und Diskursrelationen. Die Ergebnisse der Studie können in qualifizierter und quantifizierter Form in eine lexikographische Darstellung des Gegenstands eingebracht werden. Im Rahmen des Projekts finden sie Eingang in das in Potsdam erarbeitete Diskursmarkerlexikon DimLex (siehe Stede 2008 und DimLex), das für die speziellen Bedürfnisse der automatischen Textverarbeitung entwickelt wurde.

Am Zustandekommen dieser Studie haben viele ihren Anteil. In erster Linie sind hier unsere Projektpartner zu nennen: Manfred Stede, der für den Potsdamer Anteil in leitender Funktion verantwortlich war; Michael Grabski, der zusammen mit Dagmar Frohning die hier zugrunde gelegte Taxonomie der Diskursrelationen erarbeitete; Kristin Irsig, die in Potsdam zusammen mit

Manfred Stede das Diskursmarkerlexikon aufbaute. In Mannheim profitierte das Projekt in der ersten Hälfte wesentlich von Dagmar Frohnings Expertise. Den Mannheimer Kollegen aus dem Konnektorenprojekt, Hardarik Blühdorn und Ulrich H. Waßner, verdanken wir nicht nur viele wertvolle Anregungen und nützliche Kritik, sondern auch direkte tatkräftige Zuarbeit. Gisela Zifonun danken wir für ihre Bereitschaft, das Manuskript kritisch zu lesen. In Berlin unterstützten uns Anke Lüdeling und Emil Kroymann. Die korpuslinguistische Knochenarbeit aber – 3.200 Belege zu kategorisieren und 2.410 von ihnen nach 18 Merkmalen und 5 Relationen zu analysieren – bewerkstelligten vor allem unsere Hilfskräfte Romina Barbera und Lisa Kolb. Auch im Dickicht der zahlreichen statistischen Berechnungen wie in dem der Manuskriptversionen haben sie die Übersicht nicht verloren.

Der Band ist wie folgt aufgebaut:

Kap. 1 führt in den Gegenstand ein, nennt die zentralen Forschungsfragen und gibt einen knappen Überblick über einige der neueren Zugänge zum Thema. Es versteht sich von selbst, dass hier nicht in extenso auf die reichhaltige Literatur zur Kausalität – etwa aus der philosophischen, der logischen, der rhetorischen, der illokutionsemantischen Tradition – eingegangen werden kann und wir uns auf die für uns zentralen grammatischen, korpuslinguistischen und diskurslinguistischen Ansätze konzentrieren. Ein zentrales Augenmerk liegt dabei auf dem, was in der Tradition von Sweetser (1990) als die drei „Domänen“ oder „Ebenen“ semantischer Relationen bezeichnet wird – eine Sachverhaltsebene der „real world“, eine epistemische Ebene der Sprecher-einstellungen und eine illokutive Ebene der Sprechakte – was sich in ähnlicher Form in der diskurslinguistischen Tradition in der Unterscheidung „semantischer“ von „pragmatischen“ Relationen (Knott/Sanders 1998) niederschlägt. Aus der Forschung werden die Ausgangshypothesen der Korpusuntersuchung abgeleitet, aus der sich auch die Auswahl der untersuchten Merkmale begründet.

Kap. 2 widmet sich der Datengrundlage und der Methode. Das benutzte Korpus wird in seiner Zusammensetzung beschrieben, und der Weg vom Gesamtkorpus (der Gesamtheit der schriftsprachlichen Korpora am IDS) zu einem Korpus von Kausalmarkern in einer begrenzten Kontextumgebung wird dargestellt. Dabei wird auch auf das Disambiguierungsproblem bei polyfunktionalen Markern eingegangen, bei denen eine nicht-kausale Verwendung (*da lässt sich nichts machen; es dreht sich darum; er redet ohne Wissen daher*) oder

eine zwar kausale aber nicht satzverknüpfende Verwendung (*es ist unklar, weshalb/weswegen diese Lösung gewählt wurde*) von einer kausalen geschieden werden muss. Ferner wird in Kap. 2 auf die Analysemethoden und die bei der Auswertung benutzten statistischen Testverfahren eingegangen.

Kap. 3 beschreibt im Detail die einzelnen erhobenen Merkmale, listet die möglichen Merkmalsausprägungen und belegt sie mit Beispielen aus dem Korpus. Problemfälle für die Merkmalszuweisung werden jeweils genannt und die hier getroffenen Entscheidungen begründet.

Die Ergebnisse der Korpusanalyse werden in **Kap. 4** präsentiert. Es werden die absoluten Merkmalsausprägungen und eine Vielzahl von Korrelationen dargestellt. Im ersten Teil geht es um das markerspezifische Auftreten bestimmter Form- und Funktionsmerkmale, und um die Frage der Spezifik der Marker in Bezug auf die Diskursrelationen. Im zweiten Teil wird von den einzelnen Markern abgesehen und es werden die Korrelationen zwischen Diskursrelationen und Merkmalen präsentiert. Wir geben zunächst einen Überblick über die signifikanten Korrelationen und bewerten die Ergebnisse, anschließend werden sämtliche Korrelationstabellen im Einzelnen angeführt.

Kap. 5 diskutiert die Ergebnisse vor dem Hintergrund der in Kap. 1 skizzierten Forschungslage, gibt einen zusammenfassenden Überblick und nennt die noch offenen Fragen.