

**Elemente
der
optimalitätstheoretischen Syntax**

Gereon Müller

gereon.mueller@ids-mannheim.de

STAUFFENBURG / VERLAG / TÜBINGEN
19. September 2000

Inhaltsverzeichnis

Vorbemerkung	1
1 Grundzüge der Theorie	5
1. Hintergrund	5
2. Zentrale Annahmen	8
3. Kandidaten	11
4. Kandidatenmengen	12
5. Struktur der Grammatik	13
6. Notation	16
7. Parametrisierung	17
8. Zahl vs. Qualität der Verletzungen	17
9. Eine erste Anwendung: ‘Do’-Einsetzung	18
9.1. Rekonstruktion bei Speas (1995)	19
9.2. Rekonstruktion bei Grimshaw (1997)	22
10. Optionalität	26
10.1. Mögliche Erklärungen	26
10.2. Ein Beispiel: Fehlende Komplementierer im Englischen	27
11. Typen von Beschränkungen	29
11.1. Treuebeschränkungen	29
11.2. Markiertheitsbeschränkungen	31
2 Evidenz und Gegenevidenz	33
1. Parametrische Variation	33
1.1. Einfache Fragesätze	33
1.2. Mehrfachfragen	35
2. Reparatur	39
2.1. ‘Do’-Einsetzung im Englischen	40
2.2. Der Ersatzinfinitiv im Deutschen	40
2.3. Pronominaladverbien im Deutschen	42
2.4. Expletives ‘es’ im Deutschen	48
2.5. Schluß	49
3. Wettbewerb und Blockade	51
3.1. Komplementierertilgung im Englischen	51
3.2. Partielle W-Bewegung im Deutschen	54
4. Das Konzept des unmarkierten Falles	57
4.1. Ökonomie der Bewegung	57
4.2. Default-Kasus	59
4.3. Parallelverschiebung	60
5. Beschränkungskonflikt	65

5.1.	Deutsche Wortstellung	66
5.2.	Objektbewegung und Quantorenskopos im Isländischen	68
5.3.	Ökonomieprinzipien	70
6.	Beziehungen unter Beschränkungen	72
6.1.	Pāninische Beschränkungsrelationen	72
6.2.	Stringenzrelationen	73
7.	Konzeptuelle Gegenevidenz	74
7.1.	Komplexität	74
7.2.	Natur der Kandidatenmenge	76
7.3.	Natur der Beschränkungen	76
7.4.	Parametrisierung	77
7.4.1.	Parametrisierung und faktorielle Typologie	77
7.4.2.	Parametrisierung und morphologische Fundierung	77
8.	Empirische Gegenevidenz	82
8.1.	Optionalität	82
8.2.	Grammatikalitätsgrade	82
8.3.	Absolute Ungrammatikalität	82
8.3.1.	Der Generator: Pesetsky (1997)	83
8.3.2.	Leere Kandidaten: Ackema & Neeleman (1998)	83
8.3.3.	Schlechte Gewinner: Grimshaw (1994), Müller (1997)	85
8.3.4.	Reparatur	86
8.3.5.	Neutralisierung: Legendre et al. (1995; 1998), Baković & Keer (1999)	88
3	Drei Analysen	93
1.	Grimshaws (1993; 1997) Theorie der Inversion	93
1.1.	Grundannahmen	93
1.2.	Inversion	94
1.3.	‘Do’-Einsetzung in Fragesätzen	97
1.4.	Blockierte Inversion	104
1.5.	Negative Operatoren	106
1.6.	Konditionale	108
1.7.	Komplementierer	110
1.7.1.	Optionales vs. obligatorisches ‘that’	110
1.7.2.	Bewegung	111
1.7.3.	Ein Problem: Doppel-Comp-Filter-Verletzungen	113
1.8.	Einordnung	115
1.8.1.	Kandidatenmengen	115
1.8.2.	Typen von Evidenz	116
1.8.3.	Ein empirisches Problem	117
1.8.4.	Wirkung	118
2.	Pesetskys (1997; 1998) Theorie der PF-Realisierung im C-System	118
2.1.	Grundannahmen	118

2.2.	Das C-System im Französischen	119
2.3.	Das C-System im Englischen	123
2.4.	Schwierigkeiten	125
2.5.	PF-Realisierung von Spuren	127
2.6.	Einordnung	129
3.	Legendre, Smolensky & Wilson (1998) über Lokalität	130
3.1.	Grundannahmen	130
3.2.	W-Extraktion im Chinesischen	131
3.3.	Extraktion und lokale Konjunktion	135
3.3.1.	Treuebeschränkungen	135
3.3.2.	Lokalitätsbeschränkungen	136
3.3.3.	W-Bewegung im Englischen	139
3.4.	Einordnung	143
4	Vorgänger und Alternativen	145
1.	Universalität	145
2.	Verletzbarkeit: Uszkoreits Wortstellungsanalyse	148
3.	Geordnetheit: Pafels Skopusanalyse	151
4.	Wettbewerb	156
4.1.	Transderivationelle Ökonomiebeschränkungen im Minimalis- mus	157
4.1.1.	Das Prinzip der wenigsten Schritte	158
4.1.2.	Das Prinzip der kürzesten Pfade	161
4.1.3.	Konklusion	166
4.2.	Blockade-Syntax	170
4.2.1.	Die Grundidee	170
4.2.2.	Komparativbildung im Englischen	170
4.2.3.	Possessivkonstruktionen im Englischen	172
4.2.4.	Optimalitätstheoretische Rekonstruktion	172
4.2.5.	Konklusion	174
4.3.	Das Prinzip der Vollständigen Inklusion	174
4.3.1.	Die Analyse	174
4.3.2.	Optimalitätstheoretische Rekonstruktion	176
5.	Versteckte Optimalität	178
5.1.	Vermeide Pronomina bei Chomsky (1981)	178
5.1.1.	Die Analyse	178
5.1.2.	Optimalitätstheoretische Rekonstruktion	179
5.2.	Kontrolle bei Stechow & Sternefeld (1988)	181
5.2.1.	Die Analyse	181
5.2.2.	Optimalitätstheoretische Rekonstruktion	182
5.3.	W-Bewegung und Komplementierer-Finalität bei Kayne (1994)	183
5.3.1.	Die Analyse	183
5.3.2.	Optimalitätstheoretische Rekonstruktion	185

5.4.	Phonologische Realisierung in Kopf-Ketten bei Roberts (1997)	185
5.4.1.	Die Analyse	185
5.4.2.	Optimalitätstheoretische Rekonstruktion	186
5.4.3.	Mehrfachfragen bei Grewendorf (1997)	187
5	Optionalität	189
1.	Das Phänomen und seine Konsequenzen für die Syntax	189
2.	Pseudo-Optionalität	193
3.	Echte Optionalität	196
4.	Kopplungen	200
4.1.	Global Hierarchische Kopplung	201
4.2.	Lokal Hierarchische Kopplung	206
4.3.	Lokal Konjunktive Kopplung	212
4.4.	Disjunktive Kopplungen	216
5.	Neutralisierung	219
6.	Konklusion	223
6	Wortstellung	225
1.	Einleitung	225
2.	Ein Pseudo-Optionalitäts-Ansatz: Choi (1996; 1999)	227
2.1.	Die Analyse	227
2.2.	Diskussion	231
3.	Ein anderer Pseudo-Optionalitäts-Ansatz: Buring (1996; 1997; 1999)	233
3.1.	Analyse	233
3.2.	Diskussion	236
4.	Ein erweiterter Kopplungs-Ansatz: Müller (1999c)	238
4.1.	Die Analyse	238
4.1.1.	Voraussetzungen	238
4.1.2.	Das Scrambling-Kriterium und die Aufspaltung des Optimalitätsbegriffs	241
4.1.3.	Empirie	245
4.2.	Diskussion	249
5.	Ein D-Struktur-Ansatz: Heck (2000)	251
5.1.	Die Analyse	251
5.2.	Diskussion	254
6.	Konklusion	256
7	Kasus	259
1.	Einleitung	259
2.	Inhärente vs. strukturelle Kasus: Woolford (1999)	259
3.	Nochmals inhärente vs. strukturelle Kasus: Fanselow (1999)	265
4.	Kasus in freien Relativsätzen: Vogel (1999)	272
5.	Lexikalische Dekompositionsgrammatik	277

5.1.	Der Ansatz	277
5.2.	Struktureller Kasus und Spezifität bei Stiebels (2000) . . .	281
5.3.	Inhärenter Kasus und Sichtbarkeit bei Wunderlich (2000) . .	285
6.	Konklusion	294
8	Neuere Entwicklungen	297
1.	Beschränkungen	297
2.	Harmonische Ausrichtung	301
2.1.	Hintergrund	301
2.2.	GF-Realisierung bei Aissen (1999)	303
3.	Mehrfache Optimierung	310
3.1.	Bindungstheorie bei Wilson (1998)	310
3.2.	Wortstellung und Quantorenskopos bei Heck (1999)	316
3.3.	Weitere Ansätze	319
4.	Möglichkeiten und Grenzen	323
	Literaturangaben	325
	Index	337

Vorbemerkung

Ziele

Zunächst eine Warnung: Was dieses Buch nicht leistet (sondern voraussetzt), ist eine Einführung in die Syntax. Was es dagegen leisten soll, ist eine systematische Darstellung der Grundkonzepte der optimalitätstheoretischen Syntax, einer neuen Syntaxtheorie, die erst seit Mitte der neunziger Jahre überhaupt existiert, aber in diesem relativ kurzen Zeitraum bereits einige Bedeutung erlangt hat. Das Neue an der optimalitätstheoretischen Syntax läßt sich gut an einem einzigen Punkt festmachen: Diese Theorie erlaubt es, bestimmte Ideen für die Analyse von syntaktischem Datenmaterial umzusetzen, die in Standardtheorien schlicht keinen Platz haben. Warum das so ist, zeigt der folgende Gedankengang.

Ist man bereit, von Besonderheiten des theoretischen Überbaus und der formalen Durchführung zu abstrahieren, dann stellt sich heraus, daß die grundsätzlichen Unterschiede zwischen den bestehenden Syntaxtheorien gar nicht so groß sind, und daß die weitaus meisten Analysen ohne entscheidenden Substanzverlust aus einer Syntaxtheorie in eine andere übersetzt werden können.¹ Der Fall liegt oft anders bei Analysen im Rahmen der optimalitätstheoretischen Syntax. Der Grund ist, daß diese Analysen wesentlich auf der Verbindung eines Konzepts des Wettbewerbs mit der Idee verletzbarer und geordneter Beschränkungen beruhen, und dies hat in Standardtheorien der Syntax aus prinzipiellen Gründen keinen Platz.² Daher erlaubt die Optimalitätstheorie einen neuen Blick auf wohlbekanntere syntaktische Phänomene, die bisher noch nicht adäquat erfaßt werden konnten. Solche Phänomene sind:

- Parametrisierung
- Reparatur

¹Vgl. etwa Stechow (1980, Abschnitt 3) über das Verhältnis von Chomskys GB-Theorie (Rektions- und Bindungs-Theorie, vgl. Chomsky (1981)) und der (kategorial-) syntaktischen Komponente einer Montague-Grammatik (siehe etwa Dowty, Wall & Peters (1981)); McCloskey (1988, 56) über das Verhältnis von GB-Theorie, GPSG (Generalisierte Phrasenstrukturgrammatik, vgl. Gazdar, Klein, Pullum & Sag (1985)) und LFG (Lexikalisch-Funktionale Grammatik, vgl. Bresnan (1982)); oder Engdahl (1996) über das Verhältnis von GB-Theorie und HPSG (Head-Driven Phrase Structure Grammar, vgl. Pollard & Sag (1994)).

²Dies soll nicht bedeuten, daß derartige Ideen in der syntaktischen Forschung bisher gar keine Rolle gespielt haben; wie Kapitel 4 zeigen wird, ist das Gegenteil der Fall. Allerdings zeigt sich, daß auf Beschränkungsverletzbarkeit, Beschränkungsgeordnetheit und Wettbewerb basierende Analysen immer Fremdkörper innerhalb der jeweiligen sie umgebenden syntaktischen Standardtheoriegebäude waren. Als solche waren sie bestenfalls ein zur Gesamtheorie eigentlich nicht passender Zusatz; schlimmstenfalls konnten sie zu theorieinterner Widersprüchlichkeit führen.

- Wettbewerb und Blockade
- Unmarkierter Fall
- Beschränkungskonflikt

In all diesen Fällen scheint eine optimalitätstheoretische Analyse zunächst einmal konkurrierenden Ansätzen überlegen. Andererseits wirft, wie sich zeigen wird, die optimalitätstheoretische Syntax auch viele neue Probleme auf, die es in Standardtheorien in dieser Form nicht gibt. Einwände, die gegen eine optimalitätstheoretische Syntax erhoben worden sind, kreisen u.a. um die folgenden Punkte:

- Komplexität
- Natur der Kandidatenmenge
- Natur der Beschränkungen
- Bestimmte Aspekte der Parametrisierung(!)
- Optionalität
- Grammatikalitätsgrade
- Absolute Ungrammatikalität

Angesichts dieser Situation versteht sich das Folgende als kritische Darstellung, die auf der Grundlage der größtmöglichen Offenheit der optimalitätstheoretischen Syntax gegenüber zwar die Vorteile dieses neuen Ansatzes herauszustellen versucht, mögliche Nachteile aber nicht verschweigen möchte.³

Voraussetzungen

Eine Bemerkung ist angebracht zu den Voraussetzungen, die dieses Buch macht. Da die Optimalitätstheorie im wesentlichen eine Theorie der verletzbaren und geordneten Beschränkungen ist, läßt sich im Prinzip jede syntaktische Theorie, die überhaupt mit Beschränkungen arbeitet, in optimalitätstheoretischer Weise weiterentwickeln und zu einer optimalitätstheoretischen Syntax transformieren. Ob Zufall oder nicht, es hat sich allerdings nun so ergeben, daß die weitaus meisten der gegenwärtig existierenden Arbeiten zur optimalitätstheoretischen Syntax die Begriffssprache und Beschränkungstypen der GB-Theorie und ihrer Weiterentwicklungen (Prinzipien-und-Parameter-Theorie, minimalistisches Programm; vgl. zu letzterem Chomsky (1995)) zum Ausgangspunkt nehmen. Die Zahl der LFG-basierten optimalitätstheoretischen Arbeiten zur Syntax ist demgegenüber vergleichsweise gering

³Da es sich um ein relativ neues und sehr aktives Forschungsfeld handelt, sind viele der in diesem Buch besprochenen Arbeiten derzeit nur als Manuskript zugänglich. Manche der noch unveröffentlichten Papiere sind vom *Rutgers Optimality Archive* (http://www.webslingerz.com/cgi-bin/oa_list.cgi) oder seinem Marburger Spiegel (<http://pc0880.germanistik-kunst.uni-marburg.de/~roa/>) abrufbar; in vielen Fällen muß man sich jedoch direkt an die Autoren wenden.

(vgl. aber etwa Choi (1996; 1999) und Bresnan (1997; 1998)); und auf GPSG oder HPSG fußende optimalitätstheoretische Arbeiten sind mir gar nicht bekannt. Im Einklang damit sind als Grundlage für dieses Buch wünschenswert Kenntnisse in der GB-Theorie und ihren Weiterentwicklungen, wie sie die meisten der auf dem Markt befindlichen einschlägigen Syntaxeinführungen vermitteln; vgl. im deutschsprachigen Bereich vor allem Fanselow & Felix (1987), Grewendorf (1988), Stechow & Sternefeld (1988), Sternefeld (1991) und Borsley (1997).

Aufbau

Kapitel 1 führt in die Grundzüge der Theorie ein. Kapitel 2 charakterisiert die Bereiche der Syntax, in denen eine optimalitätstheoretische Analyse, zumindest auf den ersten Blick, erfolgversprechend erscheint, und die Bereiche, die, wiederum zumindest auf den ersten Blick, Probleme aufwerfen. Kapitel 3 stellt drei wichtige Arbeiten der optimalitätstheoretischen Syntax vor (Grimshaw (1997), Pesetsky (1997; 1998), Legendre et al. (1995; 1998)); der Schwerpunkt liegt dabei durchweg auf der jeweiligen theoretischen Innovation. Kapitel 4 skizziert eine Reihe von prä-optimalitätstheoretischen Analysen, die wenigstens eine der grundlegenden optimalitätstheoretischen Annahmen teilen. Kapitel 5 widmet sich einem Bereich, der für die optimalitätstheoretische Syntax zunächst einmal schwierig erscheint: Es werden verschiedene Möglichkeiten diskutiert, Optionalität zu bewältigen. Kapitel 6 und 7 sind in erster Linie empirisch orientiert; sie beschäftigen sich mit der Wortstellung (vor allem mit den Stellungsregularitäten im deutschen Mittelfeld) und der Kasustheorie, respektive. Kapitel 8 schließlich stellt kurz einige neuere Entwicklungen vor und skizziert zusammenfassend Möglichkeiten und Grenzen der optimalitätstheoretischen Syntax.

Dank

Für Hilfe in der einen oder anderen Form bedanke ich mich bei Daniel Büring, Gisbert Fanselow, Caroline Féry, Silke Fischer, Jane Grimshaw, Fabian Heck, Tibor Kiss, Armin Mester, Tanja Schmid, Wolfgang Sternefeld, Barbara Stiebels, Sten Vikner, Ralf Vogel, Colin Wilson, Dieter Wunderlich und Antje Zajonz. Teile des in diesem Buch enthaltenen Materials sind verwendet worden in Lehrveranstaltungen an den Universitäten von Tübingen und Stuttgart, sowie in Kompaktkursen an den Universitäten von Leipzig (Dezember 1998) und Potsdam (Juli 1999). Auch den Teilnehmern dieser Kurse bin ich in vielerlei Hinsicht zu Dank verpflichtet. Eine englische Version von Kapitel 5 ist als Müller (1999b) erschienen. Die Arbeit ist gefördert worden durch die DFG-Zuweisungen MU 1444/1-1,2-1. Gewidmet ist das Buch Jason und Kirke.

Kapitel 1

Grundzüge der Theorie

1. Hintergrund

Die Optimalitätstheorie (OT) ist Anfang der neunziger Jahre von Alan Prince, Paul Smolensky, John McCarthy und anderen entwickelt worden, und zwar zunächst hauptsächlich für die Phonologie. Die ausführlichste und bis heute im wesentlichen kanonische Darstellung der Theorie findet sich in Prince & Smolensky (1993). Eine wichtige Weiterentwicklung ist McCarthy & Prince (1995). Eine umfassende Einführung in die optimalitätstheoretische Phonologie ist Kager (1999). Erste Anwendungen auf die Syntax sind Grimshaw (1997) und Pesetsky (1997; 1998) (bei beiden Autoren sind die Grundideen rückführbar auf die Zeit um 1993/1994). Ein Großteil der ansonsten im Rahmen der Optimalitätstheorie entstandenen syntaktischen Analysen nimmt direkt oder indirekt Bezug auf mindestens eine dieser beiden Arbeiten.

Die Optimalitätstheorie ist eine Theorie über Beschränkungskonflikte und ihre Auflösung durch Beschränkungsordnung bzw. -Gewichtung. Ein solcher Ansatz ist an sich, außerhalb der Sprachwissenschaft, nichts Ungewöhnliches: Unvereinbare Anforderungen, die nach ihrer Wichtigkeit geordnet werden müssen, um zu Entscheidungen zu kommen, finden sich in sehr vielen Regelsystemen, die uns tagtäglich umgeben. Ein einfaches Beispiel läßt sich der Straßenverkehrsordnung (StVO) entnehmen.¹

Angenommen, im Straßenverkehr entsteht die Situation, daß zwei oder mehr Verkehrsteilnehmer gleichzeitig an einer Kreuzung (bzw. Einmündung) ankommen. Dann gibt es zunächst einmal mehrere prinzipielle Lösungsstrategien. Welche davon die korrekte, also “optimale”, ist, wird durch das System der verletzbaren und zueinander gewichteten Vorfahrtsregeln entschieden. Diese Vorfahrtsregeln sind in (1) aufgelistet, wobei das Zeichen \gg die Beschränkungen gemäß ihrer Wichtigkeit ordnet.

¹Die folgende Darstellung richtet sich nach der StVO in Janiszewski (1998). Daß dieses System ganz wesentlich auf der Geordnetheit und Verletzbarkeit von Regeln beruht, hat bereits in etwas anderem Zusammenhang Savigny (1983, 120) ausgeführt. Savigny ging es jedoch in erster Linie um das unter Verkehrsteilnehmern in Deutschland konventionalisierte nicht-verbale Kommunikationssystem (er nennt es “Nicht-Verbale AUtofahrersprache”, kurz NIVEAU), das auf der Basis der gegebenen Verkehrsregeln mit Hilfe von Zeichen (wie Winkgeste, Richtungsgeste, Lichttuppe, Blinker, Kopfschütteln, etc.) funktioniert, allerdings interessanterweise selbst auch wieder die geltenden Verkehrsregeln unter bestimmten Umständen außer Kraft setzen kann.

- (1) a. V(ERKEHRS)-POL(IZIST) (StVO §36):
Die Zeichen und Weisungen von Polizeibeamten auf der Kreuzung sind zu befolgen. \gg (“Sie gehen allen anderen Anordnungen und sonstigen Regeln vor.”)
- b. BL(AULICHT)-EIN(SATZHORN) (StVO §§35, 38):
Fahrzeuge des Rettungsdienstes und der Polizei dürfen blaues Blinklicht zusammen mit einem Einsatzhorn verwenden; dies ordnet an: “Alle übrigen Verkehrsteilnehmer haben sofort freie Bahn zu schaffen.” \gg
- c. L(ICHT)-ZEI(CHEN) (StVO §37):
An Kreuzungen bzw. Einmündungen bedeuten Grün: “Der Verkehr ist freigegeben”; Gelb: “Vor der Kreuzung auf das nächste Zeichen warten”; Rot: “Halt vor der Kreuzung”. \gg (“Lichtzeichen gehen Vorrangregeln [und] vorrangregelnden Verkehrsschildern ... vor.”)
- d. V(ERKEHRS)-ZEI(CHEN)(A) (StVO §39):
Verkehrszeichen auf einem Fahrzeug ist Folge zu leisten. Sie gelten auch, wenn das Fahrzeug sich bewegt. \gg (“Sie gehen den Anordnungen der ortsfest angebrachten Verkehrszeichen vor.”)
- e. V(ERKEHRS)-ZEI(CHEN)(B) (StVO §39):
Ortsfesten Verkehrszeichen ist Folge zu leisten. \gg (“Regelungen durch Verkehrszeichen gehen den allgemeinen Verkehrsregeln vor.”)
- f. S(TRASSE) V(OR) F(ELDWEG) (StVO §8):
Fahrzeuge, die aus einem Feld- oder Waldweg auf eine andere Straße kommen, haben Vorfahrt zu gewähren. \gg (“[g] gilt nicht, wenn ...”)
- g. R(ECHTS) V(OR) L(INKS) (StVO §8a):
An Kreuzungen und Einmündungen hat die Vorfahrt, wer von rechts kommt.

Dieses Regelsystem ist nichts anderes als eine einfache optimalitätstheoretische Grammatik, nur eben nicht für die Sprache, sondern für den Straßenverkehr. Die wesentlichen Eigenschaften von OT-Grammatiken lassen sich hier aber genauso gut illustrieren, und die Straßenverkehrsordnung hat den Vorteil, daß sie der Leser bereits kennt.

Nehmen wir also einmal an, zwei Verkehrsteilnehmer A, B kommen gleichzeitig an eine Kreuzung, und A kommt von links. Dann hat A gemäß Regel RVL B die Vorfahrt zu gewähren. Befindet sich B jedoch auf einem Feldweg, dann wird die höher geordnete Regel SVF entscheidend: Die korrekte Auflösung der Situation ist die, daß A unter Verletzung der tiefer geordneten Regel RVL Vorfahrt hat, weil nur so die höher geordnete Regel SVF erfüllt werden kann. Natürlich muß sich B nicht unbedingt auf einem Feld- oder Waldweg befinden, um A trotz der Regel RVL Vorfahrt gewähren zu müssen. Es kann z.B. sein, daß an der Kreuzung auf der Straße von A ein Vorfahrtszeichen steht, und auf der Straße von B ein Vorfahrt-Gewähren-Zeichen. Auch dann ist die korrekte Auflösung die, daß A unter Verletzung von RVL, aber Respektierung der höher geordneten Regel V-ZEI Vorfahrt

vor B hat. An derselben Kreuzung können nun aber auch noch zusätzlich zu den Verkehrszeichen Ampeln stehen, und selbstverständlich können die Ampeln auch gerade rot für A und grün für B sein. In diesem Fall hat wiederum B die Vorfahrt vor A, unter Verletzung der Regel V-ZEI, denn nur so kann die höher geordnete Regel L-ZEI beachtet werden. An einer solchen Kreuzung sind die Verkehrszeichen natürlich erst dann überhaupt relevant, wenn die Ampel ausgeschaltet ist, z.B. in Zeiten, wo wenig Verkehr herrscht. Daß in der eben beschriebenen Situation die tiefstgeordnete Regel RvL wieder erfüllt werden kann, ist im übrigen zwar eine Tatsache, für das Funktionieren des Regelsystems aber offensichtlich irrelevant – wie gesehen ist ja im Zweifelsfall die Regel V-ZEI wichtiger als die Regel RvL.

Es ist klar, daß durch eine weitere Veränderung der Verkehrssituation das Vorfahrtsrecht an unserer Kreuzung trotz Ampel wieder wechseln kann. Nehmen wir zum Beispiel an, daß die Verkehrssituation im Prinzip so ist, wie zuletzt beschrieben (A kommt von links und hat zwar ein Vorfahrtszeichen, aber eine rote Ampel, und B kommt von rechts und hat neben einem Vorfahrt-Gewähren-Zeichen eine grüne Ampel); der einzige Unterschied sei, daß A kein gewöhnlicher Verkehrsteilnehmer ist, sondern ein Krankenwagen mit angeschaltetem Blaulicht und Martinshorn. Dann hat wieder A die Vorfahrt vor B, denn die Beachtung der Regel BL-EIN ist im Zweifelsfall wichtiger als die Beachtung der tiefer geordneten Regel L-ZEI. Alternativ könnte es sein, daß zwar kein Fahrzeug mit Blaulicht unterwegs ist, daß aber die Ampelanlage gestört ist und etwa auf der Fahrbahn von B auf Dauer-Grün steht, oder daß gerade ein wichtiger Verkehrsteilnehmer über die Kreuzung befördert werden soll, der es erforderlich macht, sich über das Signal der Ampelanlage hinwegzusetzen (Schwertransport, Staatskonvoi, etc.). In einem solchen Fall kann ein Verkehrspolizist die Vorfahrt regulieren: Was immer die Ampel anzeigt (und, natürlich, was immer das Verkehrszeichen und die Standardregelung “rechts vor links” sagen), die Anweisung des Verkehrspolizisten (V-POL) ist übergeordnet und setzt alle anderen Regeln außer Kraft. (Wie (1) zeigt, gilt dies im übrigen auch für nicht auszuschließende Unvereinbarkeiten der Regeln V-POL und BL-EIN: Im Zweifelsfall muß auch ein Krankenwagen mit Blaulicht der Anweisung eines Verkehrspolizisten Folge leisten.)

An diesem Beispiel lassen sich bereits einige attraktive Eigenschaften einer optimalitätstheoretischen Herangehensweise an komplexe Regelsysteme illustrieren, die unten noch ausführlich diskutiert werden: Dadurch, daß die Anweisungen der Regulierungsmittel als verletzbar und zueinander geordnet betrachtet werden, können sie allgemeiner und ohne Ausnahmeklausel formuliert sein (d.h., nichts an der durch das Vorfahrtsschild oder die Ampel gegebenen Anweisung nimmt Bezug auf die Möglichkeit der Außerkraftsetzung durch intervenierende Faktoren – die Anweisung ist “Freie Fahrt”, und nicht etwa “Freie Fahrt, es sei denn, ein Verkehrspolizist steht auf der Kreuzung, oder ein Krankenwagen mit Blaulicht überquert die Kreuzung, usw.”). Darüber hinaus ergibt sich auf eine sehr einfache Weise, daß, wenn nichts anderes gesagt ist (d.h., keine höher geordnete Regel eingreift), die tiefstgeordnete Beschränkung aktiv wird und korrekt den unmarkierten Fall (also den

Normalfall) vorhersagt (nämlich RVL). Und noch eine weitere Eigenschaft der Beschränkungsinteraktion in (1) ist der Erwähnung wert: Keine noch so große Anzahl an Verletzungen von niederrangigen Beschränkungen kann die Verletzung einer einzigen höherrangigen Beschränkung aufwiegen. D.h., auch wenn ein Verkehrsteilnehmer A von rechts kommt, ein Vorfahrtsschild sieht und die Ampel für ihn auf Grün steht, gilt im Zweifelsfall die entgegengesetzte Anweisung des Verkehrspolizisten – die kombinierte Verletzung von RVL, V-ZEI und L-ZEI wiegt nicht so schwer wie eine einzige Verletzung von V-POL.

Ähnliche Mechanismen lassen sich problemlos in vielen weiteren Regelsystemen finden.² Man beachte allerdings, daß die Straßenverkehrsordnung ebenso wie die in der letzten Fußnote genannten Regelsysteme ganz wesentlich auf einer *Konvention* beruhen; gerade dies ist bei der Grammatik als Kern der menschlichen Sprachfähigkeit, wie wir seit einiger Zeit zu wissen glauben, nicht der Fall. Der Zusammenhang ist also hier ein indirekter. Nichtdestoweniger scheint die Annahme plausibel, daß sich beim Zustandekommen dieser Konventionen sehr allgemeine, kognitiv verankerte Gesetze manifestieren, und unter dieser Perspektive ist die Analogie dann eben doch kein Zufall. Des weiteren sind auch aus der Verhaltensforschung Regelsysteme bekannt, die direkt als optimalitätstheoretische Grammatiken rekonstruierbar sind, und die nichts mit Konventionen zu tun haben. Dies gilt z.B. für die Hierarchie der Instinktaktivitäten beim Stichelingsmännchen (FORTPFLANZUNGSINSTINKT \gg NESTBAUINSTINKT, BRUTPFLEGEINSTINKT \gg VERPAARUNGSINSTINKT \gg DROHINSTINKT, BEISSINSTINKT, etc.). In diesem Bereich ergeben sich zweifelsohne interessante Perspektiven, insbesondere für die interdisziplinäre Beschäftigung mit der Frage, inwieweit man davon ausgehen kann, daß die einer optimalitätstheoretischen Grammatik für natürliche Sprachen eigene Funktionsweise letztlich Modul-unspezifisch ist. Für die gegenwärtigen Zwecke mögen diese Beispiele aber genügen; wir können uns jetzt der Linguistik zuwenden und mit den zentralen Annahmen der optimalitätstheoretischen Grammatik anfangen.

2. Zentrale Annahmen

Die vier zentralen Annahmen der Optimalitätstheorie sind in (2) zusammengestellt:

²Suggestive weitere Beispiele kann man u.a. der Rechtsprechung entnehmen (z.B.: Bundesrecht setzt Länderrecht außer Kraft, Urteile einer im Prinzip beliebigen Zahl niedrigrangiger Gerichte können durch das Urteil eines einzigen höherrangigen Gerichts zunichte gemacht werden, usw.); oder auch der Theologie (vgl. z.B. die Art und Weise, wie die Juden von Beterah mit Hilfe von Hillel dem Babylonier zu Beginn unserer Zeitrechnung durch Beschränkungsordnung den Konflikt aufzulösen trachteten, der sich zwischen den Anforderungen des Sabbats (“Keine Arbeit am Sabbat!”), des Passah-Opfers (“Beim Passahfest muß ein Lamm geopfert werden!”), und des gewöhnlichen Tamid-Opfers (“Zweimal täglich muß das Tamid-Opfer erbracht werden!”) ergibt; vgl. Drescher (1996, 8)).

- (2) a. *Universalität:*
Beschränkungen sind universell.
- b. *Verletzbarkeit:*
Beschränkungen können verletzt werden.
- c. *Geordnetheit:*
Beschränkungen sind geordnet.
- d. *Wettbewerb:*
Die Grammatikalität eines Kandidaten K ist nicht allein aufgrund interner Eigenschaften von K ermittelbar; vielmehr entscheiden externe Faktoren (der Wettbewerb von K mit anderen Kandidaten) über die Wohlgeformtheit von K.
(Zunächst einmal kann man sich als Kandidaten in der Syntax vereinfacht Sätze vorstellen, aber s.u.)

Diese vier Grundannahmen sind nicht alle logisch unabhängig. So setzt die Annahme der Geordnetheit von Beschränkungen (2-c) de facto die Annahme der Verletzbarkeit von Beschränkungen (2-b) voraus; umgekehrt gilt das nicht. Entscheidend ist, daß alle Annahmen in (2) üblichen Postulaten in syntaktischen Standardtheorien widersprechen. Dies sieht man sofort beim Vergleich von (2) mit entsprechenden Grundannahmen etwa der GB-Theorie:

- (3) *Grundannahmen in Standardtheorien (z.B. GB-Theorie) :*
- a. Nicht alle Beschränkungen sind universell (Parameter, sprachspezifische Filter, etc.).
- b. Beschränkungen dürfen nicht verletzt werden.
- c. Beschränkungen sind nicht geordnet (alle sind gleich wichtig).
- d. Die Grammatikalität eines Kandidaten K ist allein aufgrund interner Eigenschaften von K ermittelbar; externe Faktoren (die Eigenschaften anderer Kandidaten) sind irrelevant.

Ein wichtiger Unterschied zwischen der optimalitätstheoretischen Syntax und anderen Syntaxtheorien ist also, daß der fundamentale Begriff der Grammatikalität nicht nur unter Bezug auf interne Eigenschaften eines Kandidaten definiert ist. Die Optimalitätstheorie setzt vielmehr Grammatikalität mit Optimalität in der Kandidatenmenge gleich:

- (4) *Grammatikalität:*
Ein optimaler Kandidat aus der Kandidatenmenge ist grammatisch, alle nicht-optimalen Kandidaten sind ungrammatisch.

Zu beachten ist, daß alle nicht-optimalen (suboptimalen) Kandidaten gleichermaßen ungrammatisch sind. Dies bedeutet, daß das Konzept der graduellen Grammatikalität, wie es oft für Bewegungsphänomene (Subjazenzenz), Bindungsphänomene, Wortstellungsphänomene usw. angenommen wurde, ohne Modifikationen in der optimalitätstheoretischen Syntax nicht ausdrückbar ist.

Wie ist nun der Begriff der Optimalität selbst festgelegt? Optimal ist ein Kandidat mit dem besten Beschränkungsprofil. Im gewöhnlichen Verständnis heißt das, daß der optimale Kandidat ein besseres Beschränkungsprofil haben muß als alle anderen Kandidaten, und daß es somit nur einen optimalen Kandidaten geben kann. In der Optimalitätstheorie soll aber auch der Fall zugelassen werden, daß mehrere Kandidaten in einer Kandidatenmenge optimal sein können, wenn sie ein identisches Beschränkungsprofil haben und es keinen anderen Kandidaten gibt, der ein besseres Beschränkungsprofil hat (vgl. die Ausführungen zur Optimalität in Abschnitt 10). Die folgende Definition von Optimalität reflektiert dies:

(5) *Optimalität:*

Ein Kandidat K_i ist optimal hinsichtlich einer Beschränkungsordnung $\langle B_1 \gg B_2 \gg \dots \gg B_n \rangle$ gdw. es keinen anderen Kandidaten K_j in derselben Kandidatenmenge gibt, der ein besseres Beschränkungsprofil hat.

Der Hilfsbegriff des Beschränkungsprofils ist wie in (6) festgelegt.

(6) *Beschränkungsprofil:*

K_j hat ein besseres Beschränkungsprofil als K_i , wenn es eine Beschränkung B_k gibt, für die (a) und (b) gelten:

- a. K_j erfüllt B_k besser als K_i .
- b. Es gibt keine Beschränkung B_l , die höher als B_k geordnet ist und bei der sich K_i und K_j unterscheiden.

Ein Kandidat K_j erfüllt eine Beschränkung B besser als ein Kandidat K_i , wenn K_j B weniger oft verletzt als K_i . (Dies impliziert den Fall, daß K_j B gar nicht verletzt, K_i dagegen schon.) Jetzt ist gewährleistet, daß im Prinzip zwei Kandidaten in einer Kandidatenmenge optimal sein können: Optimal ist ein Kandidat genau dann, wenn es keinen besseren gibt, und nicht nur dann, wenn er besser ist als alle anderen.³

³Eine andere Definition von Optimalität, die unserem natürlich-sprachlichen Verständnis davon, "am besten" zu sein, näher kommt, wird in Grimshaw (1997, 373) gegeben; vgl.:

(i) *Optimalität:*

Ein Kandidat K_i ist optimal hinsichtlich einer Beschränkungsordnung $\langle B_1 \gg B_2 \gg \dots \gg B_n \rangle$ gdw. er beim Wettbewerb mit jedem anderen Kandidaten K_j in derselben Kandidatenmenge die höchst-geordnete Beschränkung besser erfüllt, wo K_i und K_j sich unterschiedlich verhalten.

Diese Festlegung präsupponiert, daß es immer eine Beschränkung gibt, bei der sich zwei Kandidaten unterscheiden, und es folgt, daß ein optimaler Kandidat notwendigerweise besser ist als alle Konkurrenten, daß es also nur einen geben kann. Tatsächlich sollte es bei zufällig identischem Beschränkungsprofil zweier Kandidaten aufgrund der nicht-erfüllten Präsupposition gar keinen optimalen Kandidaten geben. Damit (i) wie (5) mehrere Gewinner in einer Kandidatenmenge zuläßt, müßte die Präsupposition in einen *wenn*-Satz ausgegliedert werden ("Wenn es eine Beschränkung gibt, bei der sich K_i und K_j unterscheiden, dann muß K_i , um optimal zu sein, gegen K_j bei der höchst-geordneten Beschränkung gewinnen, wo sie sich unterscheiden."). Wie wir sehen

Zu klären ist jetzt als nächstes, welchen formalen Status die im Wettbewerb befindlichen Kandidaten haben und wie dieser Wettbewerb selbst (also das Konzept der Kandidatenmenge) definiert ist.

3. Kandidaten

Es sind verschiedene Festlegungen dafür vorgeschlagen worden, was die im Wettbewerb befindlichen Kandidaten für Objekte sind.

(7) *Kandidaten* (mögliche Festlegungen):

- a. Repräsentationen auf einer einzigen Strukturebene (z.B. S-Struktur-Repräsentationen)
- b. D-Struktur/S-Struktur-Paare
- c. N-Tupel (z.B. D-Struktur/S-Struktur/LF-Tupel)
- d. Vollständige Derivationen

Die Wahl eines Kandidatentyps für eine gegebene Analyse ist, läßt man einmal ausschließlich konzeptuell motivierte Präferenzen außer acht, eng daran gekoppelt, von welchem Typ die (verletzbaren und geordneten) Beschränkungen sind, die vorgeschlagen werden. Sind diese Beschränkungen z.B. sämtlich S-Struktur-Filter, ist es naheliegend, als Kandidaten wie in (7-a) nur S-Struktur-Repräsentationen anzunehmen; sind die postulierten Beschränkungen dagegen wenigstens zum Teil echt derivationell (d.h., nicht in offensichtlicher Weise als S-strukturelle Filter rekonstruierbar), dann wird man bei Kandidaten nicht mit S-Struktur-Repräsentationen auskommen, sondern braucht auch wie in (7-d) vollständige Derivationen.

Im Einklang mit der in der Literatur dokumentierten Variabilität bzgl. des Konzepts des syntaktischen Kandidaten werde ich im folgenden nicht eine einzige Definition voraussetzen. Wo eine Festlegung bei der Präsentation eines Arguments notwendig ist, wird sie an gegebener Stelle erfolgen, und wo eine Festlegung nicht notwendig scheint (weil die verschiedenen Konzepte keine unterschiedlichen Vorhersagen machen), wird sie unterbleiben.

Eine Bemerkung sei noch gemacht zur Entscheidung zwischen (7-a) und (7-d) in der optimalitätstheoretischen Syntax, und zum damit verbundenen Verhältnis von repräsentationeller und derivationeller Grammatikorganisation. Ursprünglich ist die Optimalitätstheorie in der Phonologie auch dadurch motiviert worden, daß durch eine Hinwendung zu (7-a) ein Verzicht auf (7-d) möglich ist; diese Position übernimmt etwa Grimshaw (1997) für die Syntax. In der klassischen GB-Theorie (Chomsky (1981)) sind die von einer Grammatik erzeugten Sätze Tripel, die aus

werden, geht Grimshaw (1997) davon aus, daß es mehrere optimale Kandidaten in einem Wettbewerb geben kann; implizit muß hier also das soeben skizzierte Verständnis von (i) gegeben sein. Vgl. in übrigen Sternefeld (1990) (und z.T. Lasnik & Uriagereka (1988, 53 & 161)) zu analogen Bemerkungen bzgl. der Ableitung des PRO-Theorems in Chomsky (1981).

D-struktureller, S-struktureller und LF-Repräsentation bestehen; will man diese Ansicht in der optimalitätstheoretischen Syntax beibehalten, dann ist der Kandidatenbegriff in (7-c) zu postulieren. (7-d) entspricht schließlich dem, was im minimalistischen Programm (Chomsky (1995)) angenommen wird, denn dort sind die von der Grammatik erzeugten Sätze vollständige Derivationen. (7-d) ist aber ebenfalls vollständig mit optimalitätstheoretischen Grundannahmen kompatibel. Eine derartige optimalitätstheoretische Syntax ist insofern dann näher am minimalistischen Programm, und die Unterschiede sind weniger frappierend als das vielfach behauptet wird. Generell gilt, daß die Frage der Kandidatendefinition orthogonal ist zu den sonstigen Unterschieden zwischen diesen Syntaxtheorien (vgl. etwa Prince & Smolensky (1993, 5) (Fußnote 3) oder Legendre, Smolensky & Wilson (1998, 285-286)), und daß eine optimalitätstheoretische Syntax im Prinzip sowohl repräsentationell als auch derivationell angelegt sein kann.

4. Kandidatenmengen

Ähnliche Variabilität besteht bei der Definition der Kandidatenmenge (= Referenzmenge bei Chomsky (1995); vgl. den Überblick bei Sternefeld (1997)). In (8) sind einige der möglichen Definitionen von Kandidatenmengen aufgelistet, wie sie sich in der Literatur finden:

- (8) *Kandidatenmenge* (mögliche Festlegungen):
 Zwei Kandidaten sind in derselben Kandidatenmenge gdw. sie
- a. dieselbe Numeration (dasselbe lexikalische Material) haben.
 - b. dieselbe LF haben.
 - c. dieselbe Bedeutung haben.
 - d. dieselbe D-Struktur haben.
 - e. dieselbe Oberflächenstruktur (bzw. S-Struktur) haben.
 - f. dieselbe Numeration und dieselbe LF haben.
 - g. dieselben Prädikat-/Argument-Strukturen und identische LFs haben.
 - h. dieselben Prädikat-/Argument-Strukturen haben und auf identische LFs zielen.

Die Definition der Kandidatenmenge ist die Definition des Wettbewerbs von Kandidaten; entsprechend hat die Festlegung auf ein Konzept ganz entscheidenden Einfluß auf die Grammatik. Generell gilt: Je größer die Kandidatenmenge, desto geringer ist die Chance, daß ein gegebener Kandidat optimal ist; je kleiner die Kandidatenmenge, desto größer ist die Chance der Optimalität eines gegebenen Kandidaten. Durch eine radikale Verkleinerung der Kandidatenmenge auf ein Element (z.B. durch die a priori ja mögliche Festlegung "Zwei Kandidaten sind in derselben Kandidatenmenge gdw. sie identisches lexikalisches Material und identische Derivationen aufweisen", was einen Kandidatenbegriff wie in (7-d) voraussetzen würde) kann man die Optimalitätstheorie vollkommen trivialisieren, weil bei nur einem Kandidaten im Wettbewerb dieser unweigerlich optimal sein muß; und

durch eine zu weite Fassung der Kandidatenmenge kann es leicht zu unerwünschten Blockaden von grammatischen Sätzen kommen. Hier gilt es also, sorgfältig abzuwägen zwischen dem möglichen explanativen Vorteil, den ein Wettbewerb über eine große Kandidatenmenge hinweg bietet (weil nur bei hinreichend großer Kandidatenmenge der OT-Mechanismus überzeugend arbeiten kann), und dem möglichen deskriptiven Vorteil, den ein kleiner Wettbewerb in Aussicht stellt (weil hier weniger die Gefahr besteht, fälschlicherweise wohlgeformte Kandidaten als suboptimal auszuschließen).

Es sei an dieser Stelle darauf hingewiesen, daß die Kandidatenmenge leicht sehr groß (und nicht-finit) werden kann, dies aber nicht notwendigerweise werden muß (vgl. z.B. Pesetskys (1997; 1998) Variante von (8-e), die in Abschnitt 2 von Kapitel 3 besprochen wird).

5. Struktur der Grammatik

Die Kandidaten werden gemäß (5) einem Wettbewerb ausgesetzt, der in einem Teil der Grammatik stattfindet, der Beschränkungen mit den Eigenschaften in (2) enthält. Dieser Teil der Grammatik wird auch ‘Harmonie-Evaluation’ oder ‘H-Eval’ genannt (vgl. Prince & Smolensky (1993)), weil nur ein maximal harmonischer (also optimaler) Kandidat wohlgeformt sein kann. Die Kandidaten – welche Definition in (7) wir auch immer annehmen – kommen nun aber nicht aus dem Nichts: Die verschiedenen konkurrierenden Strukturen müssen zunächst einmal erzeugt werden, bevor sie dem Wettbewerb ausgesetzt werden können. Die Erzeugung der Kandidaten leistet ein vorgelagerter Teil der Grammatik, der ‘Generator’ oder ‘Gen’ genannt wird. Diese Grammatikkomponente hat nur Beschränkungen mit (mehr oder weniger) den Eigenschaften in (3); insbesondere sind hier die Beschränkungen nicht verletzbar und nicht geordnet.⁴ Schematisch läßt sich der Aufbau einer optimalitätstheoretischen Syntax wie in Bild 1 darstellen.

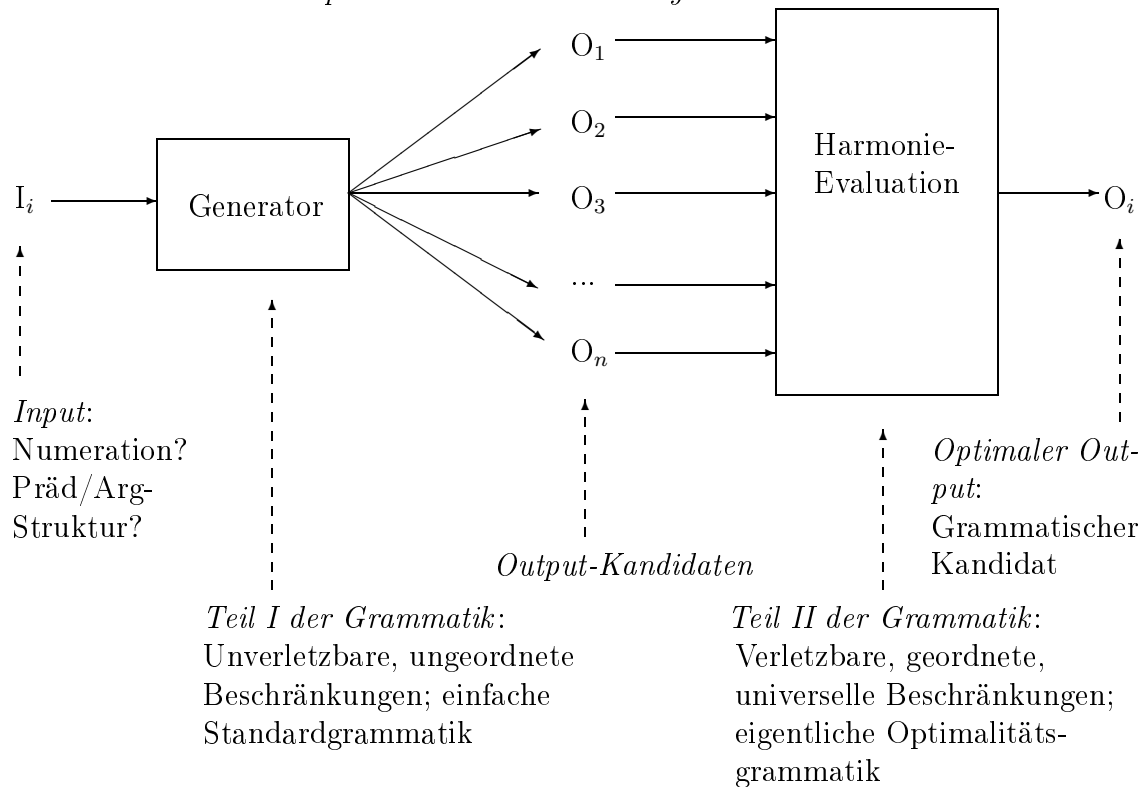
Die von Gen erzeugten und an den H-Eval-Teil der Grammatik weitergegebenen Kandidaten heißen auch Outputs. Das Bild weist zusätzlich zu dem schon Bekannten noch ein neues Konzept auf, das des Inputs. Die Idee ist hier, daß der Generator nicht einfach so Kandidaten (Outputs) erzeugt, sondern auf der Basis einer vorab gegebenen Menge von Informationen, eben des Inputs.⁵ Was genau als Input fungiert, darüber herrscht in der Syntax kaum Klarheit.⁶ Eine wesentliche Funktion

⁴Ob der Generator sprachspezifische Beschränkungen aufweisen kann, ist nicht ganz klar. Es liegt aber wohl in der Natur des optimalitätstheoretischen Ansatzes, hiervon nicht auszugehen.

⁵Der Rest dieses Abschnitts enthält konzeptuelle Erwägungen, die zum jetzigen Zeitpunkt vielleicht noch nicht so gut nachzuvollziehen sind, aber kaum an einem anderen Ort stehen können. Diese Erwägungen scheinen mir wichtig; sie sind jedoch für das weitere Verständnis nicht absolut essentiell.

⁶In der Phonologie ist der Fall etwas einfacher: Inputs sind die zugrundeliegenden Repräsen-

Bild 1: Struktur einer optimalitätstheoretischen Syntax



des Inputs in der klassischen Optimalitätstheorie ist es, den Wettbewerb festzulegen; d.h., Kandidatenmengen sind über Input-Identität definiert. Wenn wir nun die verschiedenen in der Literatur vorgeschlagenen Konzepte der Kandidatenmenge rekapitulieren (vgl. (8)), dann würde das für jede getroffene Festlegung der Kandidatenmenge bedeuten, daß die Festlegung dem Input äquivalent wäre. Falls man etwa wie Chomsky (1995) davon ausgehen will, daß zwei Kandidaten in derselben Kandidatenmenge sind genau dann, wenn sie dieselbe Numeration, d.h., dasselbe lexikalische Material, aufweisen (vgl. (8-a)), dann würde sich dies schlicht und ohne weitere Annahmen daraus ergeben, daß die Inputs Numerationen sind. Die Gleichsetzung von Kandidatenmengendefinition mit dem Input ist aber nicht ganz unproblematisch. Ein konzeptuelles Problem ergibt sich, wenn Kandidatenmengen unter Bezug auf syntaktische Struktur (z.B. unter Bezug auf identische D-Struktur, wie in (8-d)) definiert werden sollen; denn die syntaktische Struktur ist ja gerade das intendierte Ergebnis der Harmonie-Evaluation und nicht von vornherein gegeben (vgl. Archangeli & Langendoen (1997, 214)). Und wenn syntaktische Struktur bereits im Input vorhanden sein kann, dann stellt sich die Frage, mit welchem gram-

tationen, die im Lexikon aufgelistet sind.

matischen System diese Struktur erzeugt worden ist.⁷ Dieses Problem stellt sich auch bei den Kandidatenmengendefinitionen in (8-g) (Wettbewerb bei identischen Prädikat-/Argument-Strukturen und identischen LFs), die auf Grimshaw (1997) zurückgeht, und (8-h) (Wettbewerb bei identischen Prädikat-/Argument-Strukturen und identischen LF-Zielen), die Legendre et al. (1995; 1998) vorgeschlagen haben. Unter der Annahme, daß mit "LF" hier anders als in der Chomskyschen Tradition üblich keine genuin syntaktischen Strukturen gemeint sind, muß man im Input zwar keine syntaktischen Repräsentationen annehmen; allerdings kann kein Zweifel daran bestehen, daß auch rein semantische Repräsentationen von Sätzen hierarchisch gegliederte Strukturen sein müssen, so daß auch hier ein komplexer Aufbau des Inputs unvermeidbar scheint. Unklar ist dann wie vorher, mit Hilfe welchen Systems die bereits im Input vorhandene komplexe Struktur aufgebaut wird.⁸

Ein möglicher Ausweg wäre vielleicht, die Annahme aufzugeben, daß der Input vollständig die Kandidatenmenge definiert. Dann könnten wir z.B. sagen, daß Inputs Numerationen oder Prädikat-/Argument-Strukturen sind, daß aber die vom Generator erzeugten Output-Kandidaten nochmals durch einen Filter laufen, der nur solche Outputs in die Harmonie-Evaluation durchläßt, die identische syntaktische oder semantische Strukturen aufweisen; tatsächlich scheint Grimshaw (1997, 376) so etwas als Möglichkeit zu erwägen.⁹ Von hier ist es dann kein allzu großer Schritt mehr hin zu der radikalen Konklusion, daß man auf das Konzept des Inputs in der optimalitätstheoretischen Syntax vielleicht ganz verzichten könnte.¹⁰ Für die gegenwärtigen Zwecke sind derartige Überlegungen jedoch nicht entscheidend; ein eher intuitives Verständnis von Input (z.B. als Ansammlung von Inhaltswörtern, die in einem zu bildenden Satz verwendet werden sollen) mag zunächst einmal ausreichen. Im Einklang damit werde ich im folgenden den Begriff des Inputs auch ohne explizite Festlegung hin und wieder verwenden, wenn aus dem Kontext klar ist, was damit gemeint ist.

⁷Eine Lösung wäre hier die Mehrfachoptimierung. Vgl. Kapitel 8.

⁸Im übrigen sei vermerkt, daß die Annahme einer semantisch voll repräsentierten Struktur als Input einer Ableitung eines Satzes im wesentlichen dem entspricht, was von George Lakoff, James McCawley, John Ross, Paul Postal und anderen im Rahmen der Theorie der Generativen Semantik gegen Ende der sechziger Jahre vorgeschlagen wurde. Ein guter neuerer Überblick über diese Theorie findet sich in Huck & Goldsmith (1995).

⁹Ein solcher Ausweg bestünde allerdings nicht für den Kandidatenmengenbegriff in (8-h), denn hier ist ja eine abstrakte semantische Zielrepräsentation (die nicht notwendigerweise die echte semantische Repräsentation eines Satzes sein muß) der Referenzpunkt für miteinander im Wettbewerb befindliche Kandidaten; es ist nicht ohne weiteres zu sehen, wo eine solche abstrakte semantische Zielrepräsentation vorliegen kann, wenn nicht im Input. Ich werde in Kapitel 3 auf diesen Punkt zurückkommen.

¹⁰Dann müßten Beschränkungen, die sich explizit auf Input-Eigenschaften beziehen (Treuebeschränkungen, s.u.), entsprechend reformuliert werden.

6. Notation

Die Ordnung von zwei Beschränkungen A, B wird wie gesehen repräsentiert als $A \gg B$. Der Wettbewerb wird im allgemeinen veranschaulicht mit Hilfe von Tabellen (“Tableaux”). An dieser Stelle sind einige Notationskonventionen einzuführen: Ein optimaler Kandidat erhält im folgenden immer eine zeigende Hand, den sogenannten “pointing finger”: \Rightarrow . Eine Verletzung einer Beschränkung wird in einer Tabelle durch einen Stern $*$ markiert. Eine für einen Kandidaten fatale Verletzung einer Beschränkung (d.h., eine Verletzung, die den Kandidaten aus dem Wettbewerb ausscheiden läßt) schließlich wird mit einem Ausrufezeichen $!$ markiert. Oft (aber nicht in diesem Buch) werden noch andere Hilfsmittel benutzt, z.B. Schattierung von Feldern, in denen für den Wettbewerb irrelevante Beschränkungen erfüllt bzw. verletzt werden; oder umkringelte Sterne bei Beschränkungen, die von optimalen Kandidaten verletzt werden, etc. Diese Informationen sind redundant und dienen immer nur dem besseren Überblick.

Auf dieser Basis können wir nun ein erstes, noch vollkommen abstrakt gehaltenes Beispiel diskutieren. Angenommen, es gilt die Beschränkungsordnung $A \gg B \gg C$, die Kandidaten K_1 bis K_5 gehören derselben Kandidatenmenge an, und sie verletzen die Beschränkungen A, B, C wie in Tabelle $T_{1.1}$ angegeben. Dann erscheint K_1 als singulär optimaler Kandidat, und somit als einzig wohlgeformter.

$T_{1.1}$: Das Grundprinzip

Kandidaten	A	B	C
$\Rightarrow K_1$			*
K_2			**!
K_3		*!	
K_4	*!		
K_5		*!	*

Der Grund ist, daß kein Kandidat ein besseres Beschränkungsprofil hat als K_1 (tatsächlich hat hier auch kein anderer Kandidat ein ebenso gutes Beschränkungsprofil). Die höchst-geordnete Beschränkung, bei der sich K_1 und K_4 unterschiedlich verhalten, ist A. K_1 erfüllt A besser als K_4 . Die höchst-geordnete Beschränkung, bei der sich K_1 und K_5 unterscheiden, ist B. Hier gewinnt ebenfalls K_1 . Dasselbe gilt für K_1 und K_3 : Die höchst-geordnete Beschränkung, wo sich die beiden Kandidaten unterschiedlich verhalten, ist B, und K_1 erfüllt diese Beschränkung besser als K_3 . Im Falle von K_2 schließlich ist die höchst-geordnete Beschränkung, wo sich K_1 und K_2 unterscheiden, erst C; aber auch hier gewinnt K_1 , da dieser Kandidat C weniger oft verletzt als K_2 .

7. Parametrisierung

Aufgrund von Annahme (2-a) können syntaktische Unterschiede zwischen Einzelsprachen nicht durch parametrisierte Beschränkungen erfaßt werden; Beschränkungen im H-Eval-Teil der Grammatik sind universell gültig. Eine mögliche Alternative wäre Parametrisierung in Gen, aber auch hier wird oft angenommen, daß die Beschränkungen ganz allgemein (idealerweise universell) sein müssen. Die Optimalitätstheoretische Lösung ist verblüffend einfach und elegant:

(9) *Parametrisierung:*

Parametrisierung = Beschränkungsreihenfolge

Wenn z.B. eine Sprache nicht wie in $T_{1.1}$ die Ordnung $A \gg B \gg C$ zugrundelegt, sondern $A \gg C \gg B$, dann ist nicht K_1 der optimale Kandidat, sondern K_3 . Dies zeigt Tabelle $T_{1.2}$.

$T_{1.2}$: *Parametrisierung*

Kandidaten	A	C	B
K_1		*!	
K_2		**!	
☞ K_3			*
K_4	*!		
K_5		*!	*

8. Zahl vs. Qualität der Verletzungen

Die Zahl der von einem Kandidaten ausgelösten Verletzungen von Beschränkungen ist sekundär (vgl. auch das oben angeführte nicht-linguistische Beispiel der Beschränkungsinteraktion bei der Verkehrsregulierung). Das heißt, aufgrund von (5) kann keine auch noch so große Anzahl von Verletzungen einer untergeordneten Beschränkung auch nur eine einzige Verletzung einer übergeordneten Beschränkung aufwiegen: Qualität geht vor Quantität; das System ist in diesem Sinne nicht-kumulativ. Die Zahl der Verletzungen einer Beschränkung wird erst dann relevant, wenn zwischen Kandidaten hinsichtlich der Qualität der Verletzungen von Beschränkungen nicht mehr unterschieden werden kann. Dies veranschaulicht Tabelle $T_{1.3}$. In $T_{1.3}$ ist der optimale Kandidat (K_1) einer, der mehr Verletzungen akkumuliert als manche seiner suboptimalen Konkurrenten (K_3, K_4, K_5). Erst beim Vergleich der Kandidaten K_1 und K_2 , die sich hinsichtlich der hoch geordneten Beschränkungen A und B identisch verhalten, wird die Anzahl der Verletzungen von C relevant, und K_2 wird durch K_1 blockiert.

$T_{1,3}$: Irrelevanz der Zahl der Verletzungen an sich

Kandidaten	A	B	C
$\leftarrow K_1$			****
K_2			*****!*
K_3		*!	
K_4	*!		
K_5		*!	*

9. Eine erste Anwendung: ‘Do’-Einsetzung

An dieser Stelle ist es vielleicht nützlich, ein erstes Beispiel einer optimalitätstheoretischen Analyse anzuführen, das die Relevanz dieses Ansatzes für bestimmte Bereiche zeigt. Konkret soll es um das Phänomen der *do*-Einsetzung (sog. *do*-support) in Negationskontexten im Englischen gehen, das seit Chomsky (1957, 61-62) immer wieder im Blickpunkt des Interesses der generativen Syntax gestanden hat. Die in Chomsky (1991) entwickelte Analyse beruht ganz wesentlich auf der Grundidee, daß *do*-Einsetzung ein Reparaturprozeß ist, ein “letzter Ausweg” (“last resort”), der nur möglich ist, wenn alle anderen Optionen ausgeschlossen sind. Reparaturprozesse sind eines der klassischen Anwendungsgebiete der Optimalitätstheorie (s.u.), und so ist klar, daß diese Analyseidee für *do*-Einsetzung leicht optimalitätstheoretisch umgesetzt werden kann. Tatsächlich hat Speas (1995, 638-641) gezeigt, daß die Beschränkungen aus Chomskys (1991) Analyse unmittelbar in einer optimalitätstheoretischen Rekonstruktion übernommen werden können, mehr noch, daß Begriffe wie Verletzbarkeit und Geordnetheit von Beschränkungen eigentlich bereits in Chomskys (1991) Ansatz inhärent sind. Grimshaw (1997, 389-393) entwickelt eine optimalitätstheoretisch basierte Erklärung von *do*-Einsetzung, die letztlich ebenso auf Chomskys Grundidee zurückgeht, im Detail allerdings von seinem Ansatz abweicht. Diese beiden Analysen sollen hier kurz vorgestellt werden.

Zunächst die Fakten. *Do*-Einsetzung ist im Englischen obligatorisch, wenn eine Negation oberhalb eines Vollverbs auftritt und kein Auxiliärverb präsent ist:

- (10) a. *Mary not left
 b. Mary did not leave

Tritt keine Negation auf, ist *do*-Einsetzung nicht nur nicht obligatorisch; *do* ist dann unmöglich:¹¹

- (11) a. Mary left
 b. *Mary did leave

¹¹(11-b) wird akzeptabel bei Fokussierung von *did*. Diesen Umstand kann man für die gegenwärtigen Zwecke aber vernachlässigen. (Eine umfassende Analyse wird emphatisches *did* als Konsequenz einer intervenierenden, hoch geordneten Beschränkung auffassen.)

Die korrekte Generalisierung scheint also zu lauten: *do*-Einsetzung ist nur dann erlaubt, wenn sie notwendig ist, d.h., wenn es sonst keine Möglichkeit gibt, den Satz zu retten: Es handelt sich um ein Reparaturphänomen, einen letzten Ausweg.¹²

Nun zu den beiden erwähnten Analysen.

9.1. Rekonstruktion bei Speas (1995)

Speas setzt die drei Beschränkungen in (12) voraus, die alle auch in der Analyse von Chomsky (1991) wichtig sind (z.T. mit leicht anderer Formulierung, was darauf zurückzuführen ist, daß Chomsky nicht explizit von Beschränkungsverletzbarkeit und -Ordnung ausgeht).

- (12) a. ECP ("Empty Category Principle") (Chomsky (1981; 1991)):
Spuren müssen strikt registriert sein.
- b. LETZT-AUS ("Letzter Ausweg", "Last Resort") (Chomsky (1991)):
Sprachspezifische Sonderregeln dürfen nicht benutzt werden.
- c. ÖKON ("Ökonomie", "Economy", "Stay", "Don't Move") (Chomsky (1991; 1993; 1995)):
Bewegung ist verboten.

Die postulierte Beschränkungsordnung im Englischen ist die folgende:

- (13) *Beschränkungsordnung*:
ECP \gg LETZT-AUS \gg ÖKON

Angenommen wird von Speas (1995) wie von Chomsky (1991) zum einen, daß LETZT-AUS unter anderem *do*-Einsetzung blockiert (denn dies ist eine sprachspezifische Sonderregel des Englischen).¹³ Zum anderen ist wesentlich, daß das ECP verletzt ist, falls ein finites Hauptverb auf der S-Struktur unter einer Negation steht. Warum dies so sein soll, dafür gibt es verschiedene mögliche Erklärungen. Die von Chomsky (1991, 429) gegebene beruht darauf, daß LF-Bewegung eines auf der S-Struktur unter einer Negation in der VP stehenden Vollverbs in die I-Position obligatorisch ist, sofern kein Auxiliarverb in I steht. Die bei LF-Bewegung eines Vollverbs über eine intervenierende Negation hinweg entstehende Spur muß jedoch das

¹²Diese Intuition findet sich im wesentlichen so auch schon bei Chomsky (1957, 62). Gemäß dieser klassischen Analyse ist die für (11-a) postulierte Transformation "Affix Hopping", die die Flexion mit dem Verbstamm verbindet (vgl. die Regel (29-ii) in Chomsky (1957, 39)), bei Vorliegen einer intervenierenden Negation (genereller: bei fehlender Adjazenz der beiden Elemente) wie in (10-a) blockiert. In diesem Kontext, und nur hier, kann und muß dann die Transformation *do*-Einsetzung applizieren (vgl. die Regel (40) in Chomsky (1957, 62)).

¹³Dies würde dann natürlich doch präsupponieren, daß Gen nicht nur universelle Beschränkungen aufweist. Diesen Umstand können wir hier aber vernachlässigen, da es ja in erster Linie um die Rekonstruktion der Chomskyschen (1991) Analyse geht, und nicht so sehr um den bestmöglichen optimalitätstheoretischen Ansatz.

ECP verletzen. Unter diesen Voraussetzungen läßt sich nun die Chomskysche (1991) Analyse von *do*-Einsetzung ohne weiteres optimalitätstheoretisch formulieren. Tabelle T_{1.4} zeigt, daß in Negationskontexten der Kandidat mit *do*-Einsetzung trotz Verletzung von LETZT-AUS optimal ist; denn der Konkurrent ohne *do*-Einsetzung und mit V-in-situ verletzt das höher geordnete ECP. Der Einfachheit halber sei postuliert, daß die konkurrierenden Kandidaten hier S-Struktur-Repräsentationen sind (und daß das ECP schlicht finite Vollverben unter einer Negation verbietet), auch wenn dies nicht mit Chomskys (1991) Annahmen übereinstimmt, die eigentlich komplette Derivationen bis zur LF voraussetzen.¹⁴

T_{1.4}: *Negation und ‘do’-Einsetzung bei Speas*

Kandidaten	ECP	LETZT-AUS	ÖKON
☞ K ₁ : Mary did ₁ not t ₁ leave		*	*
K ₂ : Mary t ₁ not left ₁	*!		*

Demgegenüber zeigt Tabelle T_{1.5}, daß bei Fehlen einer intervenierenden Negation, und somit der Gefahr einer ECP-Verletzung, die durch *do*-Einsetzung ausgelöste LETZT-AUS-Verletzung fatal wird und der optimale Kandidat somit bloßes V-in-situ aufweist.¹⁵

T_{1.5}: *Verbot der ‘do’-Einsetzung ohne Negation bei Speas*

Kandidaten	ECP	LETZT-AUS	ÖKON
K ₁ : Mary did ₁ t ₁ leave		*!	*
☞ K ₂ : Mary t ₁ left ₁			*

Soweit der Kern der Analyse. Daß Speas (1995) hier tatsächlich wenig mehr getan hat, als ohnehin von Chomsky (1991) gemachte Annahmen explizit auf optimalitätstheoretische Art und Weise darzustellen, läßt sich leicht anhand des folgenden Zitats erkennen: “Um die korrekten Ergebnisse zu erzielen, muß das Prinzip des ‘letzten Auswegs’ so interpretiert werden, daß Prinzipien der UG angewendet

¹⁴Ein Merkmal dieser wie praktisch aller anderen Tabellen in diesem Buch (und in der Literatur zur optimalitätstheoretischen Syntax generell) ist, daß nur ein Ausschnitt des Wettbewerbs abgebildet wird und längst nicht alle Kandidaten betrachtet werden. Angesichts der Größe (im Normalfall sogar Nicht-Finitheit) der Kandidatenmenge ist dies unvermeidbar. Es tauchen daher hier und im folgenden immer nur einige der konkurrierenden Kandidaten in Tabellen auf, nämlich die, die am ehesten eine Chance haben, optimal zu werden, oder die auf andere Weise relevant scheinen.

¹⁵Wie man sieht, ist ÖKON für diese Wettbewerbe noch nicht entscheidend. Die bei Vollverben eingetragenen Verletzungen resultieren jeweils aus der von Chomsky (1991) postulierten S-strukturellen Absenkung der Flexionsmorphologie von I an das Hauptverb (Affix Hopping).

werden, wann immer das möglich ist, und daß sprachspezifische Regeln nur benutzt werden, um eine D-Struktur-Repräsentation zu “retten”, die ansonsten kein Ergebnis liefert” (Chomsky (1991, 427)).

Die Analyse ist natürlich in mehrerlei Hinsicht vereinfacht und unvollständig. Unterschlagen ist z.B. in den Tabellen $T_{1.4}$ und $T_{1.5}$ immer ein weiterer im Wettbewerb befindlicher Kandidat, der anstelle von V-in-situ oder *do*-Einsetzung das Vollverb nach I bewegt; vgl.:

- (14) a. *Mary left₁ not t₁
 b. *Mary left₁ t₁

Tatsächlich wäre (14-a) ein gefährlicher Wettbewerber in $T_{1.4}$: Nach Chomskys (1991) Annahmen würde dieser Kandidat weder LETZT-AUS verletzen (da er ohne *do*-Einsetzung auskommt), noch das ECP (das zwar LF-Bewegung eines Vollverbs über eine Negation verbietet, nicht jedoch S-Struktur-Bewegung im selben Kontext).¹⁶ Chomsky und Speas nehmen an, daß (14-ab) aus unabhängigen Gründen unmöglich sind: Englisch erlaubt anders als etwa Französisch prinzipiell niemals S-strukturelle Bewegung eines Vollverbs nach I; dies ist der sog. “V-I-Bewegungsparameter” (vgl. Pollock (1989) und die dort angegebene Literatur). Das Verbot S-struktureller V-I-Bewegung läßt sich als eine weitere Beschränkung fassen, die im Englischen hoch geordnet ist – höher jedenfalls als LETZT-AUS.

Abschließend sei noch auf einen Punkt hingewiesen, der wie bereits erwähnt entscheidenden Einfluß auf optimalitätstheoretische Analysen hat, nämlich die Definition des Wettbewerbs, d.h., die Festlegung der Kandidatenmenge. Unmittelbar einsichtig ist, daß die soeben vorgestellte Analyse nicht ohne weiteres mit sämtlichen oben erwähnten Möglichkeiten der Definition von Kandidatenmengen vereinbar ist. Nimmt man etwa an, daß Kandidatenmengen nur über identische Numerationen (identisches lexikalisches Material) festgelegt werden, dann können K_1 und K_2 in $T_{1.4}$ und $T_{1.5}$ nicht konkurrieren (weil sich die Kandidaten in der Präsenz bzw. dem Fehlen von *do* unterscheiden), und K_2 in $T_{1.4}$ und K_1 in $T_{1.5}$ können nicht als suboptimal blockiert werden. Dagegen entsteht z.B. unter der Annahme, daß die Kandidaten einer Kandidatenmenge Realisierungen derselben Prädikat-/Argument-Strukturen mit gleicher LF sind, der erwünschte Wettbewerb – vorausgesetzt, daß allein die Präsenz bzw. Absenz von expletivem *do* noch keinen Unterschied in der Prädikat-/Argument-Struktur herbeiführt, und daß die Kandidaten K_1 und K_2 in $T_{1.4}$ und $T_{1.5}$ auf LF jeweils tatsächlich identisch sind, was beides allerdings nicht unplausibel scheint. Dieses Konzept der Kandidatenmenge

¹⁶Man mag sich fragen, warum das so sein soll. Chomsky macht diese Annahme, um sicherzustellen, daß S-strukturelle Bewegung eines Vollverbs über die Negation im Französischen nicht zugunsten einer Operation wie der englischen *do*-Einsetzung blockiert wird. Die theoretische Umsetzung dieser Annahme ist einigermaßen komplex und im gegenwärtigen Zusammenhang auch nicht entscheidend; Chomsky (1991, 429) rekurriert dabei wesentlich auf die Frage der Tilgbarkeit von verbalen Kopf-Spuren.

wird bei Grimshaw (1997) zugrundegelegt, und dies führt zur zweiten optimalitätstheoretischen Analyse von *do*-Einsetzung im Englischen.

9.2. Rekonstruktion bei Grimshaw (1997)

Grimshaw (1997) setzt bei ihrer Analyse voraus, daß die im Wettbewerb befindlichen Kandidaten S-Struktur-Repräsentationen sind. Die folgenden Beschränkungen werden angenommen:

- (15) a. LEX-ÖKON (“Bewegungsökonomie für lexikalische Köpfe”, “No-Lex-Mvt”):
Bewegung von lexikalischen Köpfen ist verboten (X^0 -Spur_{lex} ist nicht erlaubt).
- b. KASUS (“Case”; Chomsky (1981)):
Der Kopf einer NP-Kette muß in einer Kasusposition sein.
- c. OB-KOPF (“Obligatorische Köpfe”, “Ob-Hd”):
Eine Projektion hat einen (nicht-leeren) Kopf.
- d. SUBJEKT (“Erweitertes Projektionsprinzip”, “Subj”; Chomsky (1982; 1995)):
Der höchste A-Spezifikator eines Satzes muß durch ein Argument gefüllt sein.
- e. VOLL-INT (“Vollständige Interpretation”, “Full Interpretation”; Chomsky (1986b)):
Expletiveinsetzung ist verboten.
- f. ÖKON (“Stay”, s.o.):
Bewegung ist verboten (Spur ist nicht erlaubt).

Die Ordnung entspricht der Reihenfolge in (15), also:

- (16) *Beschränkungsordnung*:
LEX-ÖKON \gg KASUS \gg OB-KOPF \gg SUBJEKT \gg VOLL-INT \gg ÖKON

Bevor wir uns mit diesem Rüstzeug den Daten zuwenden, sind noch einige Grundannahmen über expletives *do*, die Satzstruktur sowie die Negation zu klären. Grimshaw nimmt erstens an, daß expletives, also semantisch leeres *do* in I erzeugt wird und daß pro Satz nur ein finites Verb vorkommen kann (wenn also finites *do* in I steht, muß das Vollverb im Infinitiv auftreten). Der Gebrauch eines expletiven *do* ist dabei nicht kostenfrei; vielmehr wird hierbei die Beschränkung VOLL-INT verletzt, die Expletiveinsetzung verbietet.¹⁷

¹⁷Die Definition von VOLL-INT in (15) geht auf Chomsky (1986b) zurück. Bei Grimshaw (1997) lautet VOLL-INT tatsächlich etwas anders, nämlich wie in (i):

- (i) VOLL-INT:
Die lexikalisch-konzeptuelle Struktur muß respektiert werden.

Zweitens postuliert Grimshaw, daß die Satzstruktur nicht einheitlich ist: Sätze sind minimal VPs (denn das ist nach der von Grimshaw akzeptierten “Hypothese der prädikats-internen Subjekte” (“VP-internal subject hypothesis”) der minimale Bereich, in dem das Verb und seine Argumente vertreten sind); sie können darüber aber noch eine (im Prinzip unbeschränkte) Zahl von funktionalen Projektionen (sog. “erweiterten Projektionen” des Verbs; vgl. Grimshaw (1991)) haben, wenn dies durch das Harmonie-Evaluationssystem als optimal klassifiziert wird. Diese erweiterten V-Projektionen haben zunächst einmal keinen inhärent vorgegebenen, durch Merkmale bestimmten Status; aber gemäß ihrer hierarchischen Ordnung kann man sich wie in Standardtheorien darauf mit Namen wie IP, CP beziehen. Mit anderen Worten: Die Satzstruktur ist variabel, und ob ein optimaler Satz eine VP, IP, CP oder noch etwas Größeres ist, ist nicht von vornherein vorgegeben (also nicht durch den Generator bestimmt), sondern Resultat der Optimierung.

Drittens schließlich nimmt Grimshaw an, daß die englische Negation *not* der Kopf einer eigenen Projektion NegP ist, die oberhalb von VP steht. SpecNeg ist jedoch keine Kasusposition. Relevant werden jetzt die Beschränkungen KASUS und SUBJEKT in (15). KASUS verlangt, daß eine NP (bzw. der Kopf einer NP-Kette) in einer Kasusposition ist. Für die Zuweisung des Nominativs soll gelten, daß Subjekte diesen Kasus nur im Spezifikator eines Kopfes mit finiter Flexion erhalten können, d.h., im Spezifikator eines finiten Vollverbs (also in der VP) oder im Spezifikator eines Auxiliars (also in der IP). SUBJEKT dagegen erzwingt, daß eine Subjekt-NP im höchsten A-Spezifikator eines Satzes steht. Dies kann an sich SpecV sein, aber nur dann, wenn VP die höchste Schicht des Satzes ist; sobald oberhalb von VP eine NegP steht, kann ein Subjekt in SpecV SUBJEKT nicht mehr erfüllen, da nun SpecNeg der höchste A-Spezifikator ist (daß SpecNeg eine A-Position ist, ist hier natürlich vorausgesetzt.) Hieraus ergibt sich, daß bei Präsenz einer Negation (und also einer NegP) eine Subjekt-NP die Beschränkung SUBJEKT verletzt, wenn sie in ihrer Basisposition SpecV verharrt (vgl. (17-a)), und die Beschränkung KASUS, wenn sie in SpecNeg zu stehen kommt (vgl. (17-b)).¹⁸ Da diese beiden Beschränkungen im Englischen relativ hoch geordnet sind, bedeutet dies, daß im

Der Grund für diese abweichende Formulierung ist, daß Grimshaw mit VOLL-INT nicht nur die Kosten für Expletiveinsetzung messen möchte, sondern darüber hinaus auch noch erklären will, warum ausgerechnet *do* mit seiner extrem unspezifischen Semantik als expletives Verb im Englischen benutzt wird, und nicht etwa *sleep* oder irgendein anderes Verb. In Abschnitt 1 von Kapitel 3 werde ich auf diesen Punkt noch zurückkommen; für die gegenwärtigen Zwecke reicht die einfachere Definition von VOLL-INT in (15) aus.

¹⁸Warum kann in (17-b) nicht t_1 in SpecV KASUS erfüllen und $Mary_1$ in SpecNeg SUBJEKT? Die Antwort ergibt sich aus der Formulierung von KASUS: Der Kopf einer NP-Kette muß in einer Kasusposition sein. Der Kopf der NP-Kette $\langle Mary_1, t_1 \rangle$ ist die NP $Mary_1$, und diese NP ist nicht in einer Kasusposition. Man beachte, daß es hier ganz wesentlich ist, daß SpecNeg eine A-Position ist. Wäre SpecNeg eine A'-Position, dann wäre $\langle Mary_1, t_1 \rangle$ keine NP-Kette (sondern eine A'-Kette), und t_1 könnte KASUS erfüllen, in Analogie zu dem, was für W-Bewegung gilt.

optimalen Kandidaten eine zusätzliche funktionale Kategorie oberhalb der NegP projiziert werden muß, die IP. Ist der Kopf dieser IP leer, ist der Kandidat kettenidentisch zu dem in (17-b) (vgl. (17-c)) und also auch auszuschließen; tatsächlich wird hier zum einen OB-KOPF verletzt (und diese Verletzung wäre schon fatal), zum anderen jedoch auch noch die höher geordnete Beschränkung KASUS (denn für Nominativzuweisung ist ja per Annahme Spezifikator-Kopf-Kongruenz mit finiter Verbmorphologie notwendig). Der zusätzliche I-Kopf muß also gefüllt werden. Eine Möglichkeit wäre die, daß das finite Vollverb *left* sich von V nach I bewegt (vgl. (17-d) = (14-a)). Wir hatten jedoch bereits bei der Betrachtung von Speas' (1995) Chomsky-Rekonstruktion gesehen, daß dieser Fall ausgeschlossen ist, weil das Englische keine sichtbare V-I-Bewegung von Vollverben erlaubt. War bei Speas (1995) eine entsprechende hoch-geordnete Beschränkung allerdings noch implizit vorausgesetzt, so ist sie bei Grimshaw (1997) explizit: (17-d) verletzt fatal LEX-ÖKON. Somit bleibt, als letzte Möglichkeit, die Einsetzung eines expletiven Verbs *do* in I. Der resultierende Kandidat (17-e) erfüllt LEX-ÖKON, KASUS, OB-KOPF und SUBJEKT; er verletzt VOLL-INT, weil ein expletives Verb eingesetzt worden ist, aber da dies per Annahme die niedrigst-geordnete der hier relevanten Beschränkungen ist, resultiert Optimalität.

- (17) a. *_{[NegP Not [VP Mary left]]}
 b. *_{[NegP Mary₁ not [VP t₁ left]]}
 c. *_{[IP Mary₁ [I -] [NegP not [VP t₁ left]]]]}
 d. *_{[IP Mary₁ [I left₂] [NegP not [VP t₁ t₂]]]]}
 e. _{[IP Mary₁ [I did] [NegP not [VP t₁ leave]]]]}

Der eben skizzierte Wettbewerb ist in Tabelle T_{1.6} illustriert; die Reihenfolge der Kandidaten entspricht der Reihenfolge in (17).

T_{1.6}: Negation und 'do'-Einsetzung bei Grimshaw

Kandidaten	LEX- ÖKON	KA- SUS	OB- KOPF	SUB- JEKT	VOLL- INT	ÖKON
K ₁ : [Neg [VP NP ₁ V]]				*!		
K ₂ : [NP ₁ Neg [VP t ₁ V]]		*!				*
K ₃ : [NP ₁ - [t ₁ Neg [VP t ₁ V]]]		*!	*			**
K ₄ : [NP ₁ V ₂ [t ₁ Neg [VP t ₁ t ₂]]]	*!					***
☞ K ₅ : [NP ₁ did ₂ [t ₁ Neg [VP t ₁ V]]]					*	**

Da das Problem für den Kandidaten K₁ in T_{1.6} die Präsenz der Negation ist (nur deshalb kann die Subjekt-NP in SpecV nicht SUBJEKT erfüllen), sagt diese Analyse korrekt vorher, daß bei Fehlen einer Negation (somit einer NegP) der optimale Kandidat keine *do*-Einsetzung aufweisen darf, sondern das Subjekt in der VP verharren läßt (zumindest solange kein Auxiliar vorliegt, das eine weitere

Projektion erfordert):¹⁹ Können die höher geordneten Beschränkungen auch von einem Kandidaten erfüllt werden, ohne VOLL-INT zu verletzen, wird eine solche Verletzung sofort fatal. Die einschlägigen Daten sind in (18) aufgelistet ((18-ab) sind kettenidentisch):

- (18) a. [VP Mary left]
 b. *[IP Mary₁ [I –] [VP t₁ left]]
 c. *[IP Mary₁ [I left₂] [VP t₁ t₂]]
 d. *[IP Mary₁ [I did] [VP t₁ leave]]

Den zugrundeliegenden Wettbewerb veranschaulicht Tabelle T_{1.7}.

T_{1.7}: Verbot der ‘do’-Einsetzung ohne Negation bei Grimshaw

Kandidaten	LEX- ÖKON	KA- SUS	OB- KOPF	SUB- JEKT	VOLL- INT	ÖKON
☞K ₁ : [VP NP ₁ V]						
K ₂ : [NP ₁ – [VP t ₁ V]]		*!	*			*
K ₃ : [NP ₁ V ₂ [VP t ₁ t ₂]]	*!					**
K ₄ : [NP ₁ did ₂ [VP t ₁ V]]					*!	*

An dieser Stelle mag es nützlich sein, die Ansätze von Speas (1995) und Grimshaw (1997) zu vergleichen. Beidesmal wird die Chomskysche Grundintuition, daß *do*-Einsetzung in Negationskontexten ein Reparaturphänomen ist, optimalitätstheoretisch umgesetzt. Während jedoch bei Speas (wie bei Chomsky (1991)) der Reparatur-auslösende “Fehler” an dem Verhältnis von Negation und finitem Vollverb festgemacht wird (die Abfolge NP-Neg-V_{fin} verletzt fatal das ECP), sieht Grimshaw das Problem in dem Verhältnis von Negation und Subjekt-NP (die Abfolge NP-Neg-V_{fin} verletzt fatal KASUS). Man kann also sagen, daß die Rolle des ECP in Speas’ Analyse und die Rolle von KASUS in Grimshaws Analyse äquivalent sind als entscheidende Auslöser von *do*-Einsetzung in Negationskontexten. Ebenso äquivalent sind LETZT-AUS bei Speas und VOLL-INT bei Grimshaw: Beide Beschränkungen bestrafen die Verwendung eines expletiven Verbs *do*; sie sind in (Auxiliar-freien) Negationskontexten von optimalen Kandidaten verletzbar. LEX-ÖKON ist bei Speas implizit, bei Grimshaw explizit angenommen; dasselbe gilt auch für SUBJEKT (denn bei Annahme der Hypothese der prädikats-internen Subjekte muß auch Speas gewährleisten, daß ein Satz wie **Not Mary left* ausgeschlossen ist). Schließlich hat auch ÖKON in beiden Analysen dieselbe Funktion, nämlich interessanterweise – für die betrachteten Fälle – gar keine! Man kann sich leicht davon überzeugen, daß in den Tabellen T_{1.4} – T_{1.7} ÖKON nicht von einem einzigen Kandidaten *fatal* verletzt wird. ÖKON ist also nicht dafür verantwortlich, einen Kan-

¹⁹Tritt ein Auxiliar in I auf, z.B. in periphrastischen Tempora, muß aufgrund von KASUS eine Subjekt-NP von SpecV nach SpecI bewegt werden.

didaten aus dem Wettbewerb zu entfernen (es steht bei keiner ÖKON-Verletzung ein Ausrufezeichen). Dies bedeutet, daß ÖKON in den vorgestellten Analysen nicht *aktiv* ist. Das heißt natürlich nicht, daß ÖKON an sich überflüssig ist; bei anderen Konstruktionen, in anderen Wettbewerben mag diese Beschränkung, obschon sehr tief geordnet, Arbeit verrichten und für die Suboptimalität von Kandidaten verantwortlich sein. Dieselbe Folgerung ergibt sich bzgl. der Rolle von OB-KOPF, einer Beschränkung von Grimshaw, die kein Gegenstück bei Speas findet. OB-KOPF ist nicht aktiv in den vorliegenden Wettbewerben, aber es wird sich in Kapitel 3 zeigen, daß diese Beschränkung in anderen Kontexten sehr wohl wichtig wird.

Zusammenfassend kann man also sagen, daß die beiden Analysen in der Grundstruktur einander relativ ähnlich sind, daß jedoch Grimshaws Analyse zum einen sorgfältiger ausgearbeitet ist und zum anderen weniger ad hoc formulierte Beschränkungen bemüht (vgl. insbesondere die Beschränkung LETZT-AUS bei Speas, sowie auch die notwendige Unterscheidung zwischen S-struktureller und LF-Bewegung eines Vollverbs über die Negation, wie sie für das ECP notwendig war). Diesen Umstand kann man allerdings Speas nicht zum Vorwurf machen, denn ihr primäres Ziel ist der Nachweis impliziter Optimalität in Chomskys (1991) Ansatz, nicht die Entwicklung einer eigenen Analyse.

Nach dieser ersten Illustration optimalitätstheoretischer Analysen können wir nun zurückkehren zur Einführung in wichtige Konzepte der Theorie. Der nächste Begriff ist der der Optionalität.

10. Optionalität

10.1. Mögliche Erklärungen

Wenn es so aussieht, als könnten zwei konkurrierende Kandidaten K_i , K_j ein und desselben Wettbewerbes beide grammatisch sein, gibt es zunächst einmal drei grundsätzliche Möglichkeiten:²⁰

(19) *Optionalität zweier Kandidaten K_i , K_j :*

a. *Pseudo-Optionalität:*

K_i , K_j gehören bei genauem Hinsehen doch zu verschiedenen Kandidatenmengen.

b. *Echte Optionalität:*

K_i , K_j haben ein identisches Beschränkungsprofil.

c. *Kopplung* ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied"), d.h., zueinander nicht geordnet.

²⁰Zumindest zwei dieser drei Möglichkeiten (Pseudo-Optionalität und Kopplungen) lassen sich noch weiter untergliedern. Was das Konzept der Kopplung betrifft, so muß an dieser Stelle noch ein eher intuitives Verständnis genügen. Beide Punkte werden in Kapitel 5 ausführlich diskutiert.

In Tabellen werden Kopplungen von Beschränkungen durch benachbarte Anordnung und gestrichelte oder fehlende senkrechte Linien zwischen ihnen repräsentiert; hier werden gestrichelte Linien verwendet.²¹ In Tabelle T_{1.8} ist der einzige Unterschied zu Tabelle T_{1.1}, daß B und C gekoppelt sind; notiert wird eine Kopplung zweier Beschränkungen hier und im folgenden mit Hilfe des Symbols \circ (B \circ C).²² Gekoppelte Beschränkungen sind nicht zueinander geordnet; sie sind gleich wichtig. Daher ist nicht nur K₁ optimal (wie in T_{1.1}), und auch nicht nur K₃ (wie in der B und C vertauschenden Tabelle T_{1.2}); K₁ und K₃ sind vielmehr beide optimal. Es entsteht so eine systematische Alternation von Konstruktionen, also Optionalität.²³

T_{1.8}: *Kopplung von Beschränkungen*

Kandidaten	A	B C
☞ K ₁		*
K ₂		**!
☞ K ₃		*
K ₄	*!	
K ₅		*(!) *(!)

Die verschiedenen Herangehensweisen können gut am Phänomen des optionalen Komplementierers *that* in englischen Deklarativsätzen illustriert werden.

10.2. Ein Beispiel: Fehlende Komplementierer im Englischen

Im Englischen kann bei bestimmten Matrixverben wie *think* in eingebetteten finiten Deklarativsätzen optional der Komplementierer *that* fehlen:

- (20) a. I think – John will leave
 b. I think [_C that | John will leave

Gemäß (19) bieten sich verschiedene Erklärungen für diese Alternation an. Eine auf Pseudo-Optionalität (vgl. (19-a)) rekurrierende Analyse beruht auf der Annahme, daß (20-a) und (20-b) gar nicht im selben Wettbewerb sind; dann können die beiden Sätze jeweils in ihren verschiedenen Kandidatenmengen optimal sein. Für eine solche Analyse könnte man z.B. postulieren, daß Kandidatenmengen über identische

²¹Gestrichelte oder fehlende senkrechte Linien in OT-Tabellen zeigen in der Literatur manchmal auch einfach an, daß die Ordnung zweier Beschränkungen grundsätzlich irrelevant ist, oder daß man (noch) keine Evidenz hat, die eine Ordnung erzwingt.

²²Andere Notationen für Kopplungen, die man in der Literatur findet, sind “A \llgg B”, “A \approx B”, “A|B” und “A/B”.

²³Das eingeklammerte Ausrufezeichen (!) besagt bei Verletzungen von zwei gekoppelten Beschränkungen, daß eine der Verletzungen als zweite Verletzung der gekoppelten Beschränkungen fatal ist; zunächst einmal kann man sich aussuchen, welche.

Numerationen (identisches lexikalisches Material im Input) definiert sind. Dann führt das Fehlen bzw. die Präsenz des Komplementierers zu zwei unterschiedlichen Wettbewerben, und die beiden Kandidaten können nicht miteinander konkurrieren. Eine solche Analyse ist im Kern von Legendre, Wilson, Smolensky, Homer & Raymond (1995) und Baković & Keer (1999) vorgeschlagen worden.²⁴

Nehmen wir dagegen einmal an, daß Kandidatenmengen so definiert sind, daß (20-a) und (20-b) miteinander konkurrieren können (z.B., weil wir voraussetzen, daß Kandidaten miteinander im Wettbewerb sind, wenn sie Realisierungen von identischen Prädikat-/Argument-Strukturen mit gleicher Bedeutung bzw. LF sind).²⁵ Dann könnten (20-a) und (20-b) immer noch ein identisches Beschränkungsprofil haben, also genau dieselben Beschränkungen in gleichem Maße verletzen, und beide könnten im selben Wettbewerb als optimal erscheinen; wir hätten es also mit einer Instanz echter Optimalität (vgl. (19-b)) zu tun. Diese Lösung entspricht Grimshaws (1997) Vorschlag. Man beachte, daß dabei vorausgesetzt ist, daß der semantisch leere Komplementierer *that* anders als das semantisch leere Verb *do* nicht VOLL-INT verletzt; diese Annahme wird von Grimshaw (1997) tatsächlich gemacht (vgl. Abschnitt 1 von Kapitel 3).

Schließlich ließe sich auch eine Kopplungsanalyse gemäß (19-c) verfolgen, derzufolge die beiden Kandidaten zwar konkurrieren, ohne ein identisches Beschränkungsprofil aufzuweisen, sich aber lediglich bei zwei Beschränkungen unterscheiden, die miteinander gekoppelt sind. Die eine dieser beiden gekoppelten Beschränkungen würde dann z.B. eine Tilgung des Komplementierers (oder eine Leerlassung der C-Position) erfordern, und die andere würde eine leere Kopfposition verbieten. Dies ist die Erklärung von Pesetsky (1997; 1998) (vgl. Abschnitt 2 von Kapitel 3).

Welcher dieser Wege im vorliegenden Fall der beste ist, ist selbstverständlich nur in einem größeren Kontext zu entscheiden; aber prinzipiell scheinen zunächst einmal alle möglich. In Kapitel 5 werde ich auf dieses Thema noch genauer eingehen.

²⁴Ganz so einfach, wie es klingt, ist dieser Ansatz allerdings nicht. Was wäre etwa, wenn Gen einen im Input vorhandenen Komplementierer *that* durch Tilgung aus einem Output-Kandidaten entfernen könnte, oder wenn umgekehrt Gen die Fähigkeit besäße, einen Komplementierer einzusetzen, der im Input noch nicht vorhanden ist? Beides – Komplementierertilgung wie Komplementierereinsatz in der Syntax – ist in der Literatur bereits vorgeschlagen worden. Für die Pseudo-Optionalitäts-Analyse hätte das jedoch die Konsequenz, daß doch wieder Konkurrenzsituationen bei Optionalität entstehen. Tatsächlich betrachten Legendre et al. (1995) und Baković & Keer (1999) auch diesen Aspekt; sie versuchen jeweils zu zeigen, daß er die Analyse nicht unterminiert, sondern sogar positive Effekte zeitigt. Kapitel 5 wird sich dem Thema genauer widmen (unter dem Schlagwort "Neutralisierung", das allerdings schon vorher – in den Kapiteln 2 und 3 – im Zusammenhang mit der Ableitung absoluter Ungrammatikalität eine Rolle spielen wird).

²⁵Daß ein semantisch leerer Komplementierer *that* auf LF nicht vorhanden ist, haben u.a. Lasnik & Saito (1992) vorgeschlagen.

11. Typen von Beschränkungen

Neben anderen lassen sich zwei wichtige Typen von optimalitätstheoretischen Beschränkungen unterscheiden, die inhärent miteinander im Konflikt stehen, nämlich Treuebeschränkungen und Markiertheitsbeschränkungen.

11.1. Treuebeschränkungen

Bestimmte Beschränkungen sorgen dafür, daß sich bei einem Kandidaten Input und Output möglichst wenig unterscheiden. Diese Beschränkungen heißen *Treuebeschränkungen* (“faithfulness constraints”). Prince & Smolensky (1993) schlagen für die Phonologie an Treuebeschränkungen unter anderem PARSE und FILL vor. PARSE verbietet die Tilgung von im Input spezifizierten Segmenten, FILL umgekehrt die Einsetzung von Segmenten (Epenthese), die nicht im Input stehen.²⁶ Wie nun in Abschnitt 5 erläutert wurde, ist das Konzept des Inputs in der Syntax zwar derzeit noch nicht in einer allgemein akzeptierten Form definiert; der Begriff der Treuebeschränkung erweist sich nichtsdestoweniger auch hier als nützlich. Zur Illustration betrachte man nochmals die Beschränkung VOLL-INT, die in der Grimshawschen (1997) Analyse von *do*-Einsetzung eine Rolle gespielt hat (vgl. (15-e)).

(21) VOLL-INT:

Expletiveinsetzung ist verboten.

VOLL-INT läßt sich als eine Treuebeschränkung vom Typ FILL auffassen. Dies steht vollständig im Einklang mit Grimshaws (1997) Annahmen: In dieser Analyse ist der die Kandidatenmenge definierende Input eine Prädikat-/Argument-Struktur (mit vorgegebener LF), in der inhaltsleere Elemente wie funktionale Kategorien generell, aber speziell auch ein expletives *do*, fehlen. Wenn also ein expletives Verb *do* in einem Kandidaten auftaucht, dann kann es noch nicht Teil des Inputs gewesen sein; es ist vielmehr vom Generator eingesetzt worden. *Do*-Einsetzung ist unter dieser Sichtweise vollständig analog zur Epenthese in der Phonologie.

Wir könnten nun allerdings auch andere Annahmen über den Input machen, z.B. die, daß Inputs ausschließlich über Identität der LF bzw. der Bedeutung definiert sind, oder daß Kandidatenmengen gleich ganz ohne Bezug auf ein Konzept wie das des Inputs definiert werden, wiederum z.B. über Identität von LF bzw. Bedeutung (vgl. (8-b), (8-c)). In beiden Fällen könnte man ebenso eine Beschränkung wie VOLL-INT postulieren, nur daß nicht eine spätere Einsetzung eines Expletivs bestraft wird, sondern die Existenz eines Expletivs an sich:

²⁶Zwei Anmerkungen. Zum einen sind PARSE und FILL nicht tatsächlich singuläre Beschränkungen, sondern selbst wieder Klassen oder “Familien” von entsprechenden Beschränkungen. Zum anderen finden sich in McCarthy & Prince (1995) diese Beschränkungen leicht modifiziert und bei etwas geänderten Grundannahmen unter den Begriffen MAX (keine Tilgung) und DEP (keine Epenthese) wieder; diese Weiterentwicklung ist für die gegenwärtigen Zwecke aber unbedeutend.

- (22) VOLL-INT (versuchsweise modifiziert):
Expletiva sind verboten.

Wie man leicht nachprüfen kann, ändert sich bei einer solchen Variante an Grimshaws (1997) Analyse substantiell nichts. Insbesondere hat VOLL-INT dieselbe Funktion wie vorher: Das Auftauchen eines expletiven Verbs *do* wird bestraft und ist somit nur in Kontexten möglich, wo alle anderen konkurrierenden Kandidaten noch höher geordnete Beschränkungen erfüllen. Der einzige Unterschied ist ein konzeptueller: VOLL-INT ist nun im strengen Sinne keine Treuebeschränkung mehr, denn eine Verletzung der Beschränkung entsteht nicht aufgrund einer durch spätere Einsetzung eines Expletivs entstehenden Abweichung vom Input, sondern aufgrund der Präsenz des Expletivs an sich. Will man also in der einen Variante von einer Treuebeschränkung reden, in der anderen nicht? Möglicherweise will man das; eine Alternative könnte jedoch darin bestehen, das Konzept der Treuebeschränkung in der Syntax etwas liberaler zu fassen und im Zweifelsfall dabei an ein intuitives Verständnis zu appellieren.

Eine syntaktische Beschränkung, die dem phonologischen PARSE entspricht, ist das von Pesetsky (1997; 1998) vorgeschlagene Prinzip WIED, das sich vereinfacht wie in (23) formulieren läßt:

- (23) WIED (“Wiederauffindbarkeit”, “Recoverability”, “Rec”):
Bedeutungstragende Elemente dürfen nicht getilgt werden.

Als eine weitere Treuebeschränkung, die aber auf den ersten Blick weder PARSE noch FILL entspricht (weil sie weder Tilgung noch Einsetzung verbietet), kann man die oben bereits angeführte Ökonomiebeschränkung ÖKON betrachten, die in (24) wiederholt ist.²⁷

- (24) ÖKON:
Bewegung ist verboten (Spur ist nicht erlaubt).

ÖKON verbietet syntaktische Bewegung und trägt somit dazu bei, daß sich im Laufe einer syntaktischen Ableitung möglichst wenig ändert. Wiederum stellt sich die Frage, ob das für die Phonologie entwickelte Konzept von Treue als Input/Output-Identität hier ohne weiteres übernommen werden kann. Wenn ÖKON als “Spur ist nicht erlaubt” verstanden wird, ist die Antwort wohl positiv: Wie immer der Input definiert wird, Spuren syntaktischer Bewegung gibt es dort noch nicht, und somit ist die Existenz von Spuren eine Abweichung vom Input. Wenn dagegen ÖKON verstanden wird als “Bewegung ist verboten”, ist der Fall nicht mehr ganz

²⁷Man kann natürlich Bewegung so sehen, daß sie sowohl Einsetzung als auch Tilgung involviert (vgl. Chomskys (1993) Kopiertheorie der Bewegung): Zunächst wird von einem Element α eine Kopie hergestellt, eingesetzt und in die Zielposition bewegt (Einsetzung); dann wird die zurückgebliebene Konstituente α getilgt (Tilgung). Bei einer solchen Rekonstruktion von syntaktischer Bewegung ließe sich dann ÖKON doch wieder als Instanz von PARSE/FILL verstehen.

so klar. Nimmt man aber z.B. an, daß der Input bereits eine gewisse Struktur hat (also etwa nicht eine Numeration, sondern eine Prädikat-/Argument-Struktur ist) und daß der eigentliche Kern von ÖKON der ist, daß Abweichung von der in der Prädikat-/Argument-Struktur festgelegten Abfolge bestraft wird – denn genau das ist es, was Bewegung typischerweise bewirkt –²⁸ (vgl. insbesondere Williams (1999) und Haspelmath (2000)), dann kann auch unter diesem Verständnis von ÖKON gewährleistet werden, daß es sich um eine Treuebeschränkung im echten Sinne handelt.²⁹

11.2. Markiertheitsbeschränkungen

Wenn es nur Treuebeschränkungen dieser Art gäbe, dann hätte das zur Folge, daß D-Struktur, S-Struktur und LF niemals voneinander abwichen, und die Existenz von Expletiveinsetzungs-, Tilgungs- sowie Bewegungsoperationen in der Syntax natürlicher Sprachen wäre vollkommen unerklärt. Wesentlich ist daher, daß es auch eine Klasse von Beschränkungen gibt, die den Gegeneffekt von Treuebeschränkungen haben. Diese werden oft *Markiertheitsbeschränkungen* (“markedness constraints”) genannt, weil sie markierte Strukturen zu vermeiden trachten, die bei einer maximal treuen Repräsentation des Inputs im Output auftreten können. Diese Beschränkungen sind im wesentlichen solche, die die Erfüllung syntaktischer Wohlgeformtheitsanforderungen verlangen, wie z.B., daß sich bestimmte Operatoren auf der S-Struktur aus ihren D-strukturellen Positionen in speziell dafür vorgesehene Operatorenpositionen bewegen. Eine solche Beschränkung ist W-KRIT (“W-Kriterium”, “Wh-Criterion”), das die Bewegung einer W-Phrase in der Syntax verlangt:

(25) W-KRIT:

Eine W-Phrase muß auf der S-Struktur in SpecC stehen.

Wenn wir annehmen, daß im Deutschen die Ordnung W-KRIT \gg ÖKON gilt, dann folgt daraus, daß in einfachen Fragesätzen W-Bewegung immer obligatorisch sein muß, d.h., daß (26-a) (genauer, und ohne eine Entscheidung über die Natur von Kandidaten (s.o.) vorauszusetzen: der Kandidat, dem die S-Struktur (26-a) entspricht) (26-b) (bzw. den entsprechenden Kandidaten) als suboptimal blockiert:

- (26) a. (Ich weiß nicht) [_{CP} wen₁ C [_{IP} Fritz t₁ getroffen hat]]
 b. *(Ich weiß nicht) [_{CP} – C [_{IP} Fritz wen₁ getroffen hat]]

²⁸Allerdings nicht immer: Es gibt auch ketteninvariante Bewegung, die nicht Reihenfolgeverändernd ist. Vgl. Ross (1967).

²⁹Im übrigen ist ÖKON möglicherweise am besten einer dritten Klasse von Treuebeschränkungen subsumierbar, die in McCarthy & Prince (1995) vorgeschlagen wird: IDENT-F(EATURE). Dieser Typ von Beschränkungen bewirkt in der Phonologie, daß zwischen Input und Output möglichst wenig Merkmalsänderungen stattfinden.

T_{1.9}: S-strukturelle W-Bewegung im Deutschen

Kandidaten	W-KRIT	ÖKON
☞ K ₁ : ... wen ₁ ... Fritz t ₁ getroffen hat		*
K ₂ : ... Fritz wen ₁ getroffen hat	*!	

Andere Beschränkungen können eher indirekt die Verletzung von Treuebeschränkungen erzwingen; dies gilt z.B. für das oben bereits angeführte ECP, das bei der Ordnung ECP \gg LETZT-AUS (VOLL-INT) wie gesehen in der Analyse von Chomsky (1991) bzw. Speas (1995) im Englischen in bestimmten Kontexten die Einsetzung eines expletiven Verbs *do* zur Folge hat. Allgemein gilt, daß, je nach der relativen Gewichtung einer Treuebeschränkung und einer ihr entgegengesetzten (bzw. einen entgegengesetzten Effekt hervorrufenden) Markiertheitsbeschränkung in einer Sprache, der optimale Kandidat in einem Wettbewerb den Erfordernissen der Treue- oder der Markiertheitsbeschränkung nachkommt.

Die so entstehende Taxonomie von syntaktischen Beschränkungstypen ist natürlich noch sehr allgemein, und es mag gute Gründe dafür geben, die beiden großen Gruppen (Treue- und Markiertheitsbeschränkungen) weiter zu unterteilen.³⁰ Für die gegenwärtigen Zwecke soll die hier vorgenommene Grobklassifikation von Beschränkungstypen (in Beschränkungen, die syntaktische Operationen motivieren, und Beschränkungen, die Treue zum Input verlangen) aber genügen.

Damit sind die Grundzüge der optimalitätstheoretischen Syntax dargelegt. Das nächste Kapitel widmet sich der Frage, inwieweit eine derartige Syntax empirisch gerechtfertigt werden kann, und inwieweit sie Probleme aufwirft.

³⁰Grimshaw (1998) schlägt z.B. vor, zwischen Treuebeschränkungen im engeren Sinne und Ökonomiebeschränkungen wie ÖKON zu unterscheiden, sowie zwischen Markiertheitsbeschränkungen im engeren Sinne, "Strukturbeschränkungen", "Abbildungsbeschränkungen" (W-KRIT wäre je nach Interpretation dann einer der letzten beiden Klassen einzugruppieren) und "Ausrichtungsbeschränkungen" (alignment constraints). Vgl. auch Abschnitt 1 von Kapitel 8.

Kapitel 2

Evidenz und Gegenevidenz

Zunächst sollen in diesem Kapitel in den Abschnitten 1–5 fünf Typen von empirischer Evidenz angeführt werden, die für einen optimalitätstheoretischen Ansatz sprechen, nämlich Parametrisierung, Reparatur und letzter Ausweg, Wettbewerb und Blockade, das Konzept des unmarkierten Falls sowie Beschränkungskonflikte. Danach widmet sich der eher theoretische Abschnitt 6 den Beziehungen zwischen Beschränkungen. Im Anschluß daran wird in den Abschnitten 7–8 potentielle Gegenevidenz gegen eine optimalitätstheoretische Organisation der Syntax diskutiert, die sowohl konzeptueller wie empirischer Natur ist.

1. Parametrische Variation

1.1. Einfache Fragesätze

Durch schlichtes Umordnen von Beschränkungen ergibt sich einzelsprachliche Variation in unmittelbarer Weise. Ein Beispiel hierfür ist die optimalitätstheoretische Rekonstruktion des W-Bewegungsparameters, der festlegt, in welchen Sprachen S-strukturelle W-Bewegung möglich und in welchen Sprachen diese Bewegung unmöglich ist.¹ Im Deutschen, Englischen, Niederländischen etc. wird wie im letzten Abschnitt des ersten Kapitels gesehen in einfachen Fragesätzen die W-Phrase auf der Oberfläche bewegt, im Koreanischen, Japanischen, Chinesischen dagegen nicht. So zeigt (1-a) im Koreanischen, daß W-Phrasen in situ stehen können, und aus (1-b) ist ersichtlich, daß W-Phrasen nicht auf der S-Struktur nach SpecC geschoben werden dürfen. Die Evidenz für die Unmöglichkeit von (1-b) muß allerdings indirekt sein, denn die Wortkette an sich ist durchaus wohlgeformt. Es gibt aber gute Gründe dafür, anzunehmen, daß diese Kette dann immer die Struktur von (1-c) hat und nicht die von (1-b), d.h., daß optionale W-Voranstellung in diesen Sprachen immer eine Instanz von Adjunktion an IP oder VP, also Scrambling, ist, und nicht eine

¹Optimalitätstheoretische Analysen des W-Bewegungsparameters, die prinzipiell auf der hier vorgestellten Argumentation basieren, aber im Detail dann doch erheblich voneinander abweichen, finden sich in Legendre et al. (1995, 606-613; 1998, 265-268), Billings & Rudin (1996), Müller (1997, 268-274) (woran sich auch die folgende Darstellung orientiert) und Ackema & Neeleman (1998, 451-470). Grimshaw (1997, 404) weist ebenfalls auf die Ableitbarkeit des W-Bewegungsparameters durch Beschränkungsumordnung hin; sie hält es jedoch für möglich, daß andere (morphologische) Prinzipien für das W-in-situ-Phänomen verantwortlich sind, dergestalt, daß etwa das, was im Japanischen wie eine W-Phrase aussieht, in dem für eine Beschränkung wie W-KRIT (ihr OP-SPEZ) relevanten Sinn gar keine echte W-Phrase sei, weil das W-Merkmal fehle.

Instanz von Substitution in SpecC.

- (1) a. [CP – [IP Nö-nün muös-ül₁ sa-ss-]-ni] ?
 du_{top} was_{akk} kaufen-PRÄT-Q
- b. *[CP [SpecC Muös-ül₁] [IP nö-nün t₁ sa-ss-]-ni] ?
 was_{akk} du_{top} kaufen-PRÄT-Q
- c. [IP Muös-ül₁ [IP nö-nün t₁ sa-ss-]]-ni ?
 was_{akk} du_{top} kaufen-PRÄT-Q

Ein Grund dafür, optionale S-strukturelle W-Bewegung im Koreanischen nicht anzunehmen und die entsprechenden Daten per W-Scrambling zu reanalysieren, ergibt sich bei Betrachtung von längeren Abhängigkeiten. Es ist bekannt, daß W-Bewegung bei Argumenten wie bei Adjunkten im allgemeinen sukzessiv-zyklisch applizieren und so zu längeren Abhängigkeiten führen kann. Scrambling ist jedoch in vielen Sprachen (wie z.B. dem Deutschen, vgl. Fanselow (1990), Müller & Sternefeld (1993; 1994), Grewendorf & Sabel (1994) und die dort angegebene Literatur) ein Satz-gebundener Prozeß. Im Koreanischen sieht es nun so aus, daß Scrambling von Argumenten nicht Satz-gebunden ist, Scrambling von Adjunkten dagegen schon. Dies bedeutet, daß bei langen Adjunkt-Abhängigkeiten die zu (1-c) analoge Variante aus unabhängigen Gründen nicht möglich ist. Somit ergibt sich eine klare Vorhersage für den Fall der optionalen W-Voranstellung: Wenn dieser Prozeß wie in (1-b) echte W-Bewegung nach SpecC involvieren kann, dann sollten lange Adjunkt-Abhängigkeiten mit W-Voranstellung im Koreanischen möglich sein. Wenn dieser Prozeß dagegen nur wie in (1-c) das hier von vornherein ausgeschlossene Scrambling involvieren kann, dann sollten lange Adjunkt-Abhängigkeiten mit W-Voranstellung im Koreanischen nicht möglich sein. Die Datenlage zeigt nun, daß die zweite Prognose korrekt ist: In diesen Fällen ist W-in-situ die einzig mögliche Strategie, und W-Voranstellung ist ausgeschlossen:

- (2) a. [CP Nö-nün [CP Ch'ölsu-ka wae₁ o-ass-ta-ko]
 du_{top} Ch'ölsu_{nom} warum kommen-PRÄT-DEKL-C
 saengkakha-ni] ?
 denken-Q
 “Warum₁ denkst du, daß Ch'ölsu gekommen ist t₁?”
- b. *[CP Wae₁ nö-nün [CP t'₁ Ch'ölsu-ka t₁ o-ass-ta-ko]
 warum du_{top} Ch'ölsu_{nom} kommen-PRÄT-DEKL-C
 saengkakha-ni] ?
 denken-Q
 “Warum₁ denkst du, daß Ch'ölsu gekommen ist t₁?”

Hieraus läßt sich schließen, daß der W-Bewegungsparameter so formuliert sein sollte, daß er für W-in-situ-Sprachen wie Koreanisch W-Bewegung nicht nur nicht erfordert, sondern tatsächlich verbietet.

Der Unterschied zwischen Sprachen wie Deutsch und Koreanisch folgt, wenn die beiden Beschränkungen W-KRIT und ÖKON im Koreanischen gegenüber dem

Deutschen umgedreht werden: ÖKON \gg W-KRIT. Für (1) zeigt dies Tabelle T_{2.1} (vgl. damit Tabelle T_{1.9} aus dem letzten Kapitel):

T_{2.1}: Keine S-strukturelle W-Bewegung im Koreanischen

Kandidaten	ÖKON	W-KRIT
K ₁ : [<i>SpecC</i> muős-ül ₁ nǒ-nǔn t ₁ sa-ss-ni ?	*!	
☞K ₂ : [<i>SpecC</i> - nǒ-nǔn muős-ül ₁ sa-ss-ni ?		*

1.2. Mehrfachfragen

Die bisher angeführte Variation bzgl. syntaktischer W-Bewegung ist zwar gut optimalitätstheoretisch erfaßbar, aber per se noch kein zwingendes Argument für eine optimalitätstheoretische Analyse. Denn den Unterschied zwischen Sprachen mit S-struktureller W-Bewegung und Sprachen ohne S-strukturelle W-Bewegung kann man auf die eine oder andere Weise auch in Standardtheorien erfassen (vgl. z.B. die Vorschläge bei Lasnik & Saito (1992) und Chomsky (1995)). Interessanterweise ergibt sich nun aber bei der Betrachtung von Mehrfachfragen eine weitere Möglichkeit, wie sich Sprachen unterscheiden können. In Sprachen wie dem Deutschen, Englischen und Niederländischen wird bei Präsenz von mehreren W-Elementen in einem Satz immer nur genau eines bewegt; vgl.:²

- (3) a. *[_{CP} (Es) hat Fritz wann₁ [_{NP} welches Buch]₂ gelesen] ?
 b. [_{CP} Wann₁ hat Fritz t₁ [_{NP} welches Buch]₂ gelesen] ?
 c. *[_{CP} Wann₁ [_{NP} welches Buch]₂ hat Fritz t₁ t₂ gelesen] ?

Im Koreanischen (Japanischen, Chinesischen) ist auch in Mehrfachfragen wie zu erwarten keine W-Bewegung nach *SpecC* erlaubt:³

- (4) a. [_{CP} - [_{IP} Nǒ-nǔn muős-ül₁ wae₂ sa-ss-]-ni] ?
 du_{top} was_{akk} warum kaufen-PRÄT-Q
 b. *[_{CP} Muős-ül₁ [_{IP} nǒ-nǔn t₁ wae₂ sa-ss-]-ni] ?
 was_{akk} du_{top} warum kaufen-PRÄT-Q
 c. *[_{CP} Muős-ül₁ wae₂ [_{IP} nǒ-nǔn t₁ t₂ sa-ss-]-ni] ?
 was_{akk} warum du_{top} kaufen-PRÄT-Q

Im Bulgarischen (und Rumänischen) schließlich müssen, wie Rudin (1985; 1988) beobachtet hat, alle W-Phrasen in Mehrfachfragen auf der S-Struktur in den Bereich

²Die Variante mit mehrfachem W-in-situ in (3-a) ist mit spezieller Betonung marginal möglich als Echo-W-Frage. Davon wird hier abstrahiert.

³Die *-Markierung in den Sätzen (4-b) und (4-c) betrifft wiederum jeweils nur die Struktur (W-Bewegung in den Bereich von C), nicht die lineare Kette an sich (die per Scrambling erzeugt werden kann und dann wohlgeformt ist).

von C W-bewegt werden:⁴

- (5) a. $*[_{CP} - [_{IP} \text{Koj}_1 \text{ vižda kogo}_2]] ?$
 wer_{nom} sieht was_{akk}
- b. $*[_{CP} \text{Koj}_1 [_{IP} t_1 \text{ vižda kogo}_2]] ?$
 wer_{nom} sieht was_{akk}
- c. $[_{CP} \text{Koj}_1 \text{ kogo}_2 [_{IP} t_1 \text{ vižda } t_2]] ?$
 wer_{nom} was_{akk} sieht

Welche Positionen nehmen die vorangestellten W-Phrasen in (5-c) ein? Rudin (1988) setzt voraus, daß es sich bei der Position der jeweils links-peripheren W-Phrase in der C-Domäne um eine SpecC-Substitutionsposition handelt, und bei den Positionen von weiter rechts in der C-Domäne auftauchenden W-Phrasen um SpecC-Adjunktionspositionen; vgl. (6-a). Wenn dem so ist und wenn eine an SpecC adjungierte W-Phrase dadurch dem Gebot nachkommt, "in SpecC" zu stehen, dann können wir weiterhin annehmen, daß alle W-Bewegung, auch mehrfache, von der Beschränkung W-KRIT in ihrer Formulierung in (25) aus Kapitel 1 ausgelöst wird. Allerdings weisen Koster (1987) und Stechow & Sternefeld (1988) zurecht darauf hin, daß Strukturen wie (6-a) die Frage aufwerfen, wie denn die Spur t_2 von ihrem Antezedens c-kommandiert werden kann, so daß sie das elementare Prinzip erfüllt, daß Spuren gebunden sein müssen.⁵ Aus diesem Grund (sowie aus einem weiteren, der sich später noch als wichtig erweisen wird), werde ich im folgenden annehmen, daß in Konstruktionen mit mehrfacher S-struktureller W-Bewegung tatsächlich die rechts-periphere W-Phrase in der SpecC-Substitutionsposition steht, und daß alle vorangehenden W-Phrasen an CP adjungiert sind; vgl. (6-b).

- (6) a. $[_{CP} [_{SpecC} [_{SpecC} W_1] W_2] [_{C'} C [_{IP} \dots t_1 \dots t_2 \dots]]]$
 b. $[_{CP} W_1 [_{CP} W_2 [_{C'} C [_{IP} \dots t_1 \dots t_2 \dots]]]]$

Die Annahme von (6-b) statt (6-a) zwingt nun dazu, W-KRIT leicht zu modifizieren. Eine Möglichkeit ist (7):

- (7) W-KRIT (modifiziert gegenüber (25) aus Kapitel 1):
 Eine W-Phrase muß auf der S-Struktur in der Domäne von C stehen.

Hierbei soll gelten, daß die Domäne eines Kopfes X (i) den Spezifikator von X, (ii) Adjunktionspositionen von X und (iii) Adjunktionspositionen von XP umfaßt; dies ist eine etwas einfachere Variante der entsprechenden Definition bei Chomsky

⁴Die unmöglichen Strukturen in (5-a) und (5-b) sind im vorliegenden Fall Ketten-identisch; sie können aber in bestimmten Kontexten (z.B. bei Präsenz eines klitischen Elements in C) auseinander gehalten werden.

⁵Ebenso sei darauf hingewiesen, daß (6-a) die Bedingung des Strikten Zyklus, wie sie z.B. in Chomsky (1995) formuliert ist, verletzen würde.

(1993).⁶

Nun wird noch eine dritte Beschränkung (neben ÖKON und W-KRIT) benötigt, die zwischen Bulgarisch und Deutsch zu unterscheiden gestattet. Da W-KRIT Bewegung aller W-Phrasen auslöst und ÖKON die Bewegung aller W-Phrasen verbietet, muß die gesuchte Beschränkung gerade so beschaffen sein, daß sie nur mehrfache, nicht einfache W-Bewegung verbietet. In Frage kommt hier eine Beschränkung, die in Grimshaw (1997) aus ganz anderen Gründen (und für einen ganz anderen Datensatz) eingeführt worden ist (und die weiter unten noch besprochen werden wird), nämlich PUR-EP, was hier in einer gegenüber Grimshaw (1997) etwas vereinfachten Form erscheint.⁷

(8) PUR-EP (“Reinheit der Erweiterten Projektion”, “Purity of Extended Projection”):

Adjunktion an CP und C⁰ ist verboten.

Per Annahme macht ja mehrfache W-Bewegung Adjunktion an CP unter Verletzung von PUR-EP notwendig, während einfache W-Bewegung in SpecC per Substitution landen und somit PUR-EP erfüllen kann.⁸

Nehmen wir jetzt noch für die Sprachen Deutsch, Koreanisch und Bulgarisch die Ordnungen in (9) an, dann ergeben sich die Effekte des W-Bewegungsparameters auch für Mehrfachfragen auf relativ einfache Art und Weise, und ohne sprachspezifische Sonderannahmen außer der, daß die relative Ordnung von ÖKON, W-KRIT und PUR-EP entscheidend ist: Deutsch und Koreanisch unterscheiden sich in diesem Bereich nur darin, daß W-KRIT und ÖKON vertauscht sind, Deutsch und Bulgarisch nur darin, daß W-KRIT und PUR-EP vertauscht sind, und Koreanisch und Bulgarisch nur darin, daß W-KRIT einmal hoch und einmal tief geordnet ist.

- (9) a. *Ordnung im Deutschen:*
 PUR-EP \gg W-KRIT \gg ÖKON
 b. *Ordnung im Koreanischen:*
 PUR-EP \gg ÖKON \gg W-KRIT
 c. *Ordnung im Bulgarischen:*
 W-KRIT \gg PUR-EP \gg ÖKON

Den (3) zugrundeliegenden Wettbewerb im Deutschen veranschaulicht Tabelle T 2.2.

⁶Eine minimal von (6-b) unterschiedene Version könnte hier gemäß Chomsky (1995) mehrere SpecC-Positionen für W₁, W₂ vorsehen. Dann müßte W-KRIT nicht reformuliert werden.

⁷Wie sich zeigen wird, macht Grimshaws (1997) Original- PUR-EP auch nicht exakt die gleichen Vorhersagen.

⁸Wenn wie in Fußnote 6 erwogen anstatt einer Kombination von Substitution in SpecC und Adjunktion an CP mehrfache SpecC-Positionen angenommen würden, müßte die PUR-EP entsprechende Beschränkung die Bildung mehrerer Spezifikatoren verbieten.

T_{2.2}: Mehrfachfragen im Deutschen

Kandidaten	PUR-EP	W-KRIT	ÖKON
K ₁ : [CP – [IP ... wann ₁ welches Buch ₂]]		**!	
☞ K ₂ : [CP wann ₁ [IP ... t ₁ welches Buch ₂]]		*	*
K ₃ : [CP wann ₁ welches Buch ₂ [IP ... t ₁ t ₂]]	*!		**

Die die koreanischen Daten in (4) und die bulgarischen Daten in (5) erklärenden Wettbewerbe sind in Tabelle T_{2.3} und Tabelle T_{2.4} dargestellt.

T_{2.3}: Mehrfachfragen im Koreanischen

Kandidaten	PUR-EP	ÖKON	W-KRIT
☞ K ₁ : [CP – [IP ... muös-ül ₁ wae ₂ ...]]			**
K ₂ : [CP muös-ül ₁ [IP ... t ₁ wae ₂ ...]]		*!	*
K ₃ : [CP muös-ül ₁ wae ₂ [IP ... t ₁ t ₂ ...]]	*!	**	

T_{2.4}: Mehrfachfragen im Bulgarischen

Kandidaten	W-KRIT	PUR-EP	ÖKON
K ₁ : [CP – [IP koj ₁ vižda kogo ₂]]	*!*		
K ₂ : [CP koj ₁ [IP t ₁ vižda kogo ₂]]	*!		*
☞ K ₃ : [CP koj ₁ kogo ₂ [IP t ₁ vižda t ₂]]		*	**

Somit ist die in Mehrfachfragen sichtbare typologische Dreiteilung von Sprachen hinsichtlich S-struktureller W-Bewegung auf sehr einfache Weise optimalitätstheoretisch zu erfassen. Diese komplexere Variante des W-Bewegungsparameters ist in Standardansätzen, die keine verletzbaren und geordneten Beschränkungen kennen, sehr viel schwieriger zu formulieren, und das Ergebnis sieht aus prinzipiellen Gründen nicht so einheitlich aus.⁹

Abschließend ist noch eine offensichtliche Frage zu klären, die sich an das hier vorgestellte System der Umordnung von drei Beschränkungen zur Erfassung des

⁹Im minimalistischen System von Chomsky (1995) etwa kann zwar der Unterschied Deutsch/Koreanisch leicht erfaßt werden (starke vs. schwache W-Merkmale auf C, zusammen mit der Festlegung, daß nur zur Überprüfung starker Merkmale S-strukturelle Bewegung erfolgen darf); der Unterschied Deutsch/Bulgarisch bleibt aber ein Problem. Die nächstliegende Erklärungsmöglichkeit wäre wohl die, daß im Bulgarischen starke W-Merkmale auf W-Phrasen mehrfache W-Bewegung erzwingen, die im Deutschen alle schwach sind – aber dies ist einerseits mit der minimalistischen Idee inkompatibel, daß alle Bewegung Attraktion (also exklusiv eine von Eigenschaften der Zielposition ausgelöste Operation) ist, und andererseits konzeptuell schon deshalb problematisch, weil als Auslöser für ein und denselben Prozeß der W-Bewegung in zwei Sprachen unterschiedliche Faktoren identifiziert werden: im Deutschen ein starkes W-Merkmal auf C, im Bulgarischen starke W-Merkmale auf W-Phrasen.

W-Bewegungsparameters anschließt: Was ist mit den anderen drei möglichen Ordnungen von W-KRIT, PUR-EP und ÖKON, die in (10) aufgelistet sind?

- (10) a. W-KRIT \gg ÖKON \gg PUR-EP
 b. ÖKON \gg PUR-EP \gg W-KRIT
 c. ÖKON \gg W-KRIT \gg PUR-EP

Die Ordnung (10-a) weicht von der oben für das Bulgarische postulierten Ordnung (9-c) nur insofern ab, als die untergeordneten Beschränkungen PUR-EP und ÖKON vertauscht sind. Man kann sich anhand von Tabelle T_{2.4} leicht davon überzeugen, daß unter dieser Ordnung immer noch der Kandidat mit mehrfacher W-Bewegung optimal ist. Zumindest für die hier betrachteten Sätze können wir also genauso gut (10-a) wie (9-c) als im Bulgarischen gültige Ordnung annehmen. Welche der beiden Ordnungen sich dann letztlich als korrekt erweist, oder ob die Wahl vielleicht in überhaupt keinem empirischen Bereich einen Unterschied macht, ließe sich nur anhand von weiteren Daten ermitteln. Analoges gilt für (10-b) und (10-c) im Koreanischen: Solange ÖKON W-KRIT dominiert, wird der optimale Kandidat in Fragesätzen immer der mit W-in-situ sein, egal ob PUR-EP über ÖKON geordnet ist, wie oben in (9-b) angenommen, zwischen ÖKON und W-KRIT, wie in (10-b), oder aber unter W-KRIT, wie in (10-c). Wiederum heißt dies, daß wir zunächst einmal jede dieser drei Ordnungen für das Koreanische postulieren können; welche Ordnung(en) sich letztlich als korrekt erweist bzw. erweisen, das kann man nur bei Einbeziehung weiterer Daten entscheiden.

Damit sind die Möglichkeiten der relativen Ordnung der drei Beschränkungen W-KRIT, PUR-EP und ÖKON abgehandelt, und es hat sich gezeigt, daß tatsächlich jeder logisch möglichen Ordnung zunächst einmal auch eine linguistisch mögliche Ordnung entspricht. Die Hypothese, daß jeder möglichen Beschränkungsordnung auch eine mögliche natürliche Sprache entsprechen muß, taucht in der Literatur zur optimalitätstheoretischen Syntax hin und wieder auf; vgl. etwa Ackema & Neeleman (1998, 445). Der einschlägige Fachbegriff heißt hier “faktorielle Typologie” (“factorial typology”). Realistisch betrachtet ist es aber bei Einbeziehung einer größeren Zahl von geordneten Beschränkungen nicht klar, ob man erwarten darf, daß jede Ordnung eine mögliche Grammatik ist, und es scheint daher gewagt, eine entsprechende Forderung explizit als syntaktische OT-Analysen leitend zu formulieren.

2. Reparatur

Ein zweiter Typ von Evidenz, der zunächst einmal eine optimalitätstheoretische Analyse nahelegen kann, betrifft Reparaturphänomene. Ein Reparaturphänomen liegt immer dann vor, wenn die empirische Evidenz suggeriert, daß eine wohlgeformte Konstruktion Eigenschaften aufweist, die normalerweise (d.h., in anderen Konstruktionen) von der Grammatik nicht toleriert werden. Die Analyseidee ist dann, daß die Konstruktion in der Tat gegen eine unabhängig motivierte Beschränkung verstößt, aber nichtsdestoweniger der optimale Kandidat in ihrem Wettbe-

werb ist. M.a.W.: Wenn alle anderen konkurrierenden Kandidaten noch wichtigere Beschränkungen verletzen, kann eine Konstruktion als letzten Ausweg, bzw. als Reparaturmaßnahme, die minimale Verletzung einer Beschränkung wählen, deren Verletzung normalerweise, in anderen Kontexten, fatal ist. Im folgenden werden vier Beispiele angeführt.

2.1. ‘Do’-Einsetzung im Englischen

Das erste Beispiel ist oben bereits diskutiert worden: Die optimalitätstheoretische Rekonstruktion der Chomskyschen (1991) Analyse von *do*-Einsetzungsstrukturen in Negationskontexten im Englischen durch Speas (1995) und Grimshaw (1997) macht deutlich, daß die Einsetzung von *do* zwar nicht kostenfrei ist (d.h., eine Beschränkung der Grammatik wie LETZT-AUS oder VOLL-INT verletzt), aber nichtsdestoweniger unter bestimmten Umständen die einzige Möglichkeit, d.h., optimal, ist (wenn alle alternativen Kandidaten gegen höherrangige Beschränkungen verstoßen): Expletives *do* ist eine Reparaturform.

2.2. Der Ersatzinfinitiv im Deutschen

Im Deutschen tritt, wenn ein Modalverb wie *wollen* unter einem Perfekt-Auxiliar eingebettet ist und selbst wiederum einen Infinitiv einbettet, das Modalverb anstatt in der eigentlich zu erwartenden Partizipialform obligatorisch im Infinitiv auf:¹⁰

- (11) a. *daß sie das Lied singen *gewollt* hat
 b. daß sie das Lied hat singen *wollen*

Bettet das Modalverb nicht wiederum einen Infinitiv ein, dann erscheint es obligatorisch in der Partizipform; vgl. (12). Dies spricht zunächst einmal dagegen, einfach anzunehmen, daß das Auxiliar *haben* im Deutschen einen Infinitiv selektieren kann.

- (12) a. daß sie das *gewollt* hat
 b. *daß sie das hat *wollen*

Die traditionelle Sicht der Dinge ist daher die, daß die Infinitivform des Modalverbs in (11-b) eine Reparaturform ist, die als letzter Ausweg gewählt wird, weil hier die an sich zu erwartende Partizipialform unabhängig blockiert ist. Im Einklang damit heißt das Phänomen üblicherweise auch “Ersatzinfinitiv”.

Diese Grundintuition ist optimalitätstheoretisch umgesetzt worden in Schmid (1999).¹¹ Die zwei im gegenwärtigen Kontext wesentlichen Beschränkungen sind in

¹⁰Ähnliche Phänomene gibt es auch in anderen westgermanischen Sprachen.

¹¹Das Folgende ist eine starke Vereinfachung sowohl der Datensituation, als auch der in Schmid (1999) entwickelten optimalitätstheoretischen Analyse. Ausgespart sind hier z.B. der Stellungunterschied (präverbal bei Partizip vs. postverbal bei Ersatzinfinitiv), das unterschiedliche Verhalten anderer Verbklassen (bei Perzeptionsverben ist der Ersatzinfinitiv im Deutschen optional, bei An-

(13) aufgelistet:

- (13) a. *_{GE-MV} (“*ge-Modalverben”):
 Modalverben sind keine Partizipien. (Das Präfix *ge-* darf nicht auf einem Modalverb auftreten.)
- b. PROJ-P (“Projektionsprinzip”) (Chomsky (1981)):
 Die Subkategorisierungseigenschaften lexikalischer Elemente müssen in der Syntax respektiert werden.

*_{GE-MV} ist eine Markiertheitsbeschränkung, die nach Schmid Teil einer festen Hierarchie – einer sog. “Subhierarchie” – von Beschränkungen ist, die jeweils das Präfix *ge-* bei einer Verbklasse ausschließen und zueinander zwar eine fixe Ordnung aufweisen, relativ zu anderen Beschränkungen aber unterschiedlich angeordnet sein können. (Eine Subhierarchie im technischen Sinne versammelt Beschränkungen in fester Ordnung, die jeweils unterschiedliche Instantiierungen eines allgemeinen Schemas sind.) *_{GE-MV} als Teil der *_{GE-V}-Subhierarchie nun fordert das Weglassen der Partizipialmorphologie bei Modalverben. Umgekehrt ist PROJ-P eine Treuebeschränkung, die hier verlangt, daß die im Lexikon bei einem Auxiliar spezifizierte selegierte Verbform auch auf der syntaktischen Oberfläche erscheint; im vorliegenden Fall heißt das, daß das Auxiliar *hat* in (11) das unmittelbar eingebettete Verb in der Partizipialform verlangt. Bei einer Ordnung *_{GE-MV} \gg PROJ-P bedeutet dies nun, daß die Input-untreue Form (11-b) der Input-treuen, aber die Markiertheitsbeschränkung *_{GE-MV} fatal verletzenden Form (11-a) vorgezogen wird. Das heißt: Um eine Verletzung von *_{GE-MV} zu vermeiden, muß anstelle der (nach lexikalischer Selektion zu erwartenden) Partizipialform eine Infinitivform auftreten, auch wenn dies die untergeordnete Beschränkung PROJ-P verletzt. Dies zeigt Tabelle T_{2.5}.

T_{2.5}: *Ersatzinfinitive bei Modalverben im Deutschen*

Kandidaten	* _{GE-MV}	PROJ-P
K ₁ : ... singen gewollt hat	*!	
☞ K ₂ : ... hat singen wollen		*

Die hier ersichtliche Treueverletzung kann nur in Kauf genommen werden, um noch schlimmeres Unheil zu vermeiden, und gerade dies ist der Aspekt von Schmid's (1999) Analyse, der den Ersatzinfinitiv als ein Reparaturphänomen erweist.

Eine sich hier unmittelbar anschließende Frage ist, wieso keine Ersatzinfinitivform auf einem Modalverb auftaucht, wenn letzteres nicht wiederum einen Infinitiv einbettet; vgl. (12). Diese Frage kann man auf zwei Arten beantworten: Zum einen

hebungsverben dagegen unmöglich), sowie die bzgl. des Phänomens feststellbare Variation in den westgermanischen Sprachen.

könnte es eine weitere, noch höher als *GE-MV geordnete Treuebeschränkung geben, die zwar nicht bei einer Ersatzinfinitivform in (11) verletzt ist, aber bei einer Ersatzinfinitivform in (12); in diesem Fall kann der Input-treue Kandidat trotz einer durch die Partizipialform bedingten Verletzung von *GE-MV optimal sein. Zum anderen, und das ist Schmid's Erklärung, könnte man die Markiertheitsbeschränkung *GE-MV sensitiv machen für den Kontext, in dem das Modalverb auftritt, so daß in (12) ein Input-treuer Kandidat optimal sein kann, ohne PROJ-P zu verletzen, um *GE-MV zu respektieren. Konkret schlägt Schmid vor, *GE-MV so zu modifizieren, daß das Präfix *ge-* nicht generell auf einem Modalverb verboten ist, sondern nur dann, wenn es nicht am Ende einer erweiterten V-Kette steht. Der Begriff "erweiterte V-Kette" ist aus Chomsky (1986a) übernommen. Alle V-Köpfe eines Satzes (einer CP) sind demzufolge Teil einer durch Koindizierung (nicht durch S-strukturelle Bewegung) zustande gekommenen erweiterten Kette; gemeint ist hier im wesentlichen das klassische Konzept des Kohärenzfeldes. Dies läßt sich wie in (14) formulieren:

- (14) *GE-MV ("*ge-Modalverben", revidiert):
 Modalverben sind in nicht-finalen Positionen erweiterter V-Ketten keine Partizipien.

Somit ist gewährleistet, daß in "normalen" Kontexten wie in (12) die gemäß lexikalischer Subkategorisierung zu erwartende Form erscheint, und nicht die Reparaturform: Die PROJ-P-Verletzung, die in T_{2.5} in Kauf genommen werden kann, weil nur so *GE-MV erfüllbar ist, wird in T_{2.6} fatal, weil hier auch der Input-treue Kandidat *GE-MV respektiert.

T_{2.6}: Verbot des Ersatzinfinitivs bei Modalverben im Deutschen

Kandidaten	*GE-MV	PROJ-P
☞ K ₁ : ... das gewollt hat		
K ₂ : ... das hat wollen		*!

2.3. Pronominaladverbien im Deutschen

Eine ganz ähnliche Situation liegt vor bei den Pronominaladverbien im Deutschen. Unter bestimmten Umständen erscheint innerhalb einer PP im Deutschen anstatt des eigentlich zu erwartenden Personalpronomens ein R-Pronomen *da* (interrogativ *wo*; "R-Pronomen" deshalb, weil bei vokalischem Anlaut von P ein epenthetisches *r* erscheint). Die gesamte PP heißt dann Pronominaladverb. (Dieser Begriff hat sich eingebürgert, ist aber nicht ganz glücklich gewählt, da die PP sowohl als Argument, als auch als Adverb auftreten kann.) Insbesondere erfolgt Pronominaladverbbildung in den Fällen, wo eigentlich ein Personalpronomen *es* erscheinen sollte. Weil *da* zunächst einmal ein adverbiales Pronomen ist und allein, in anderen Kontexten, niemals die Funktion eines Personalpronomens haben kann, liegt es nahe,

Pronominaladverbien in Analogie zu Ersatzinfinitiven als Reparaturformen zu analysieren. Die obligatorische Ersetzung von PP-internem *es* durch die Reparaturform *da* illustriert das Satzpaar in (15):

- (15) a. *Fritz hat [_{PP} an [_{NP} es]] gedacht
 b. Fritz hat [_{PP} da-r-an] gedacht

Umgekehrt zeigen die Beispiele in (16), daß ein Personalpronomen wie *sie*, das auf einen Menschen referiert, zum einen PP-intern möglich ist, und zum anderen dann auch Pronominaladverbbildung mit *da* blockiert.¹²

- (16) a. Fritz hat [_{PP} an [_{NP} sie]] gedacht (Interpretation: [+menschlich])
 b. *Fritz hat [_{PP} da-r-an] gedacht (Interpretation: [+menschlich])

Als Objekte von V sind natürlich alle Typen von Personalpronomina erlaubt; und im Einklang damit ist hier ein *da* vollkommen unmöglich:

- (17) a. Fritz hat [_{VP} es zurückgegeben]
 b. *Fritz hat [_{VP} da zurückgegeben]

Es sieht also so aus, als sei das R-Pronomen *da* normalerweise nicht als Personalpronomen zu benutzen; möglich wird es in dieser Funktion nur als letzter Ausweg, wenn in einem Kontext das Personalpronomen unmöglich ist.¹³ Eine optimalitätstheoretische Umsetzung dieser Analyse kann mit Hilfe der folgenden drei Beschränkungen erfolgen:

- (18) a. PP-INSEL:
 PP ist eine Barriere für XP, falls XP Kasus benötigt.
 b. PRON-KRIT:
 Ein schwaches NP-Pronomen steht in der Wackernagelposition.
 c. PROJ-P:
 Die Subkategorisierungseigenschaften lexikalischer Elemente müssen in der Syntax respektiert werden.

¹²Zumindest gilt dies für das Standarddeutsche. Einige norddeutsche Varietäten erlauben interessanterweise in (16-b) auch Bezug auf einen Menschen. Um diese Varietäten soll es im folgenden nicht gehen.

¹³Wie beim Ersatzinfinitiv wird hier die Datenlage vereinfacht und nur ein Ausschnitt des Gesamtbildes betrachtet (z.B. ist Pronominaladverbbildung bei bestimmten Personalpronomina wie unbelebtem *ihm*, *sie* PP-intern nicht strikt obligatorisch oder unmöglich, sondern optional). Ebenso wird hier wie beim Ersatzinfinitiv davon abstrahiert, daß mit der Veränderung der Form bei der Pronominaladverbbildung auch eine Veränderung der linearen Abfolge einhergeht – Präposition bei Personalpronomen vs. Postposition bei R-Pronomen. Die Grundidee – Pronominaladverbien als Reparaturformen – läßt sich jedoch auch bei einer Erweiterung und Verfeinerung der Analyse aufrechterhalten; vgl. Müller (2000a), woran sich dieser Abschnitt orientiert.

Die ersten beiden Beschränkungen bedürfen einer kurzen Erläuterung. PP-INSEL ist hoch geordnet; diese Beschränkung sorgt dafür, daß kasusbedürftige NPs nicht aus PPs herausbewegt werden können, d.h., daß das Deutsche (anders als z.B. das Englische) kein Präpositionsstranden aufweist; vgl. (19-a). Das Deutsche hat demgegenüber Postpositionsstranden (vgl. (19-b)), und dies steht im Einklang mit PP-INSEL, vorausgesetzt, daß R-Pronomina anders als reguläre NPs keinen Kasus brauchen; vgl. Ross (1967), Fanselow (1983; 1991) und Müller (1995).¹⁴

- (19) a. *_{[NP Diesem Problem]₁} hat Fritz _[PP mit t₁] gerechnet
 b. Da₁ hat Fritz _[PP mit t₁] gerechnet

PRON-KRIT ist eine W-KRIT analoge Markiertheitsbeschränkung, die die syntaktische Bewegung von schwachen Pronomina in die Wackernagelposition hinein verlangt. Zwei Fragen sind zu klären: Was sind schwache Pronomina, und was ist die Wackernagelposition? Obwohl ein genauere Blick schnell zeigt, daß die Sachlage letztendlich in beiden Fällen etwas komplizierter ist, mag es für den gegenwärtigen Zweck ausreichen, anzunehmen, daß zu den schwachen Pronomina im Deutschen *es* gehört, zu den starken dagegen belebt (oder menschlich) interpretiertes *sie, ihn*;¹⁵ und daß die Wackernagelposition eine Domäne am linken Rand des Mittelfeldes bildet, die wir im Sinne von PRON-KRIT mit dem Spezifikator sowie Adjunktpositionen einer eigenständigen funktionalen Projektion π P gleichsetzen können. (Ich werde dies der Übersichtlichkeit halber im folgenden tun, aber alternativ könnte man die Wackernagelposition anstatt strukturell auch rein linear interpretieren – unabhängige synchrone Evidenz für eine Kopfposition π gibt es meines Wissens nicht.) PROJ-P schließlich ist in Formulierung und Funktion identisch zu dem, was wir im letzten Abschnitt gesehen haben: Diese Beschränkung verbietet den Austausch einer subkategorisierten Form (z.B. eines von einer Präposition subkategorisierten NP-Personalpronomens) durch eine nicht-subkategorisierte (z.B. ein eigentlich adverbiales R-Pronomen). Die vorausgesetzte Ordnung ist die folgende:

- (20) PP-INSEL \gg PRON-KRIT \gg PROJ-P

Hieraus ergibt sich das Auftreten von Pronominaladverbien im Deutschen als optimalitätstheoretisch erfaßbares Reparaturphänomen. Der einzige Unterschied zur Argumentation bzgl. des Ersatzinfinitivs im letzten Abschnitt ist, daß an die Stelle

¹⁴Daß PP-INSEL hoch geordnet ist, ist zwar eine Voraussetzung für die Ableitung der Ungrammatikalität von (19-a); noch nicht beantwortet ist damit jedoch die Frage, wie der mit (19-a) konkurrierende und diesen Satz blockierende optimale Kandidat dieser Kandidatenmenge aussieht. Dies ist eine Instanz des Problems der absoluten Ungrammatikalität in der optimalitätstheoretischen Syntax, dem ich mich in Abschnitt 8.3 dieses Kapitels zuwende. Die dort diskutierten Lösungen lassen sich auf den vorliegenden Fall übertragen.

¹⁵Worunter *sie* und *ihn* als unbelebte Pronomina fallen, ist schon weniger klar; vgl. Fußnote 13. Derartige Komplikationen können aber für die gegenwärtige Argumentation vernachlässigt werden.

der einen höher geordneten Markiertheitsbeschränkung *GE-MV zwei Beschränkungen PP-INSEL, PRON-KRIT getreten sind. Die Rolle der tiefer geordneten Treuebeschränkung PROJ-P bleibt gleich; hierdurch ist die Reparaturform in gewöhnlichen Kontexten blockiert. Was PP-INSEL und PRON-KRIT zusammen leisten, ist, schwache Pronomina in PPs in ein Dilemma zu führen: Einerseits muß sich ein schwaches Pronomen gemäß PRON-KRIT aus der PP herausbewegen, andererseits verletzt eine solche Bewegung aber PP-INSEL.¹⁶ In genau diesem Kontext erweist sich als optimaler Kandidat der, in dem das das Dilemma verursachende Element, also das schwache Pronomen, ersetzt worden ist durch eine Reparaturform: das nicht subkategorisierte R-Pronomen adverbialen Ursprungs. Denn R-Pronomina könnten zwar ohne PP-INSEL-Verletzung aus einer PP in andere Mittelfeldpositionen bewegt werden, sie unterliegen jedoch von vornherein nicht der Beschränkung PRON-KRIT:

- (21) a. daß da₁ der Fritz dem Karl einen Schlüssel [_{PP} t₁ für] gegeben hat
 b. daß der Fritz dem Karl da₁ einen Schlüssel [_{PP} t₁ für] gegeben hat
 c. daß der Fritz dem Karl einen Schlüssel [_{PP} da₁-für] gegeben hat

Den soeben skizzierten Wettbewerb, der die Existenz von Pronominaladverbien im Deutschen begründet, veranschaulicht Tabelle T_{2.7} (vgl. (15)).

T_{2.7}: *Pronominaladverbien als Reparaturformen*

Kandidaten	PP-INSEL	PRON-KRIT	PROJ-P
K ₁ : Fritz hat – π an es ₁ gedacht		*!	
K ₂ : Fritz hat es ₁ π an t ₁ gedacht	*!		
☞K ₃ : Fritz hat – π da ₁ -ran gedacht			*

Kandidat K₁ (= (15-a)) verletzt fatal PRON-KRIT, weil das schwache Pronomen *es* nicht in die Wackernagelposition Specπ bewegt wurde, sondern in der PP verharrt; Kandidat K₂ verletzt fatal PP-INSEL, weil das schwache Pronomen *es* als kasusbedürftige NP aus einer PP extrahiert worden ist; und Kandidat K₃ (= (15-b)) gelingt es als einzigem Wettbewerber, diese zwei hoch geordneten Beschränkungen durch die Wahl der Reparaturform, des Pronominaladverbs, zu respektieren. Der zu zahlende Preis ist eine PROJ-P-Verletzung, die jedoch in diesem Kontext in Kauf genommen werden kann. Besteht für einen Kandidaten allerdings die Möglichkeit, die zwei übergeordneten Beschränkungen ohne PROJ-P-Verletzung zu erfüllen, so

¹⁶Zwei Bemerkungen. Vorausgesetzt ist hier zum einen, daß Rattenfang (pied piping), also Bewegung der gesamten PP in die Wackernagelposition, nicht ausreicht, um PRON-KRIT zu erfüllen; diese Konklusion scheint durch die Daten gut begründet. Zum anderen hätte man anstelle der zwei Beschränkungen PP-INSEL und PRON-KRIT auch einfach eine Beschränkung postulieren können, die in Analogie zu *GE-MV qua Stipulation schwache Pronomina PP-intern verbietet; auf ein solches Vorgehen sollte aber dort verzichtet werden, wo es möglich ist, unabhängig etablierte Beschränkungen zu bemühen.

wird diese Verletzung sofort fatal. Dies ist der Fall bei den starken Pronomina, die ja per Annahme PRON-KRIT nicht verletzen, wenn sie in situ stehen bleiben; vgl. (16). Den Wettbewerb hier zeigt Tabelle T_{2.8}.

T_{2.8}: Verbot der Pronominaladverbbildung bei belebtem Pronomen

Kandidaten	PP-INSEL	PRON-KRIT	PROJ-P
☞K ₁ : Fritz hat – π an sie ₁ gedacht			
K ₂ : Fritz hat sie ₁ π an t ₁ gedacht	*!		
K ₃ : Fritz hat – π da ₁ -ran gedacht			*!

Der optimale Kandidat ist hier nicht der, der die Reparaturform verwendet (K₃), und auch nicht der, der das Pronomen in die Wackernagelposition bewegt (K₂). Vielmehr erweist sich als Kandidat mit dem besten Beschränkungsprofil K₁, weil hier gar keine der betrachteten Beschränkungen verletzt wird.

Ein analoges Verbot der Reparaturform finden wir auch bei schwachen Pronomina in all jenen Kontexten, wo eine Erfüllung von PRON-KRIT durch das Pronomen nicht automatisch eine Verletzung von PP-INSEL mit sich bringt, also z.B. dann, wenn das Pronomen ein Objekt von V ist; vgl. (17). Weil in diesen Sätzen allerdings kein lexikalisches Material zwischen der Objektposition und dem Beginn des Mittelfelds interveniert, kann man nicht unmittelbar sehen, ob Wackernagelbewegung von *es* erfolgt ist. Der Fall liegt anders bei den Beispielen in (22). Hier zeigt sich deutlich, daß der optimale Kandidat das schwache Pronomen in die Wackernagelposition bewegen muß.

- (22) a. Der Fritz hat es₁ π gestern der Maria t₁ geschenkt
 b. *Der Fritz hat – π gestern der Maria es₁ geschenkt
 c. *Der Fritz hat – π gestern der Maria da₁ geschenkt

Der zugrundeliegende Wettbewerb wird durch Tabelle T_{2.9} verdeutlicht: Die Kandidaten K₁ (Pronomen in situ) und K₃ (Reparaturform) sind suboptimal und also ungrammatisch, weil sie in fataler Weise PRON-KRIT respektive PROJ-P verletzen.

T_{2.9}: Verbot der Pronominaladverbbildung bei Objektpronomen

Kandidaten	PP-INSEL	PRON-KRIT	PROJ-P
K ₁ : Der Fritz hat – π gestern der Maria es ₁ V		*!	
☞K ₂ : Der Fritz hat es ₁ π gestern der Maria t ₁ V			
K ₃ : Der Fritz hat – π gestern der Maria da ₁ V			*!

Soweit die Evidenz für eine optimalitätstheoretische Behandlung der Pronominaladverbbildung als Reparaturstrategie. Die Analyse erklärt, warum in bestimmten Kontexten statt eines Personalpronomens eine Reparaturform auftritt. Man mag sich nun noch fragen, warum es ausgerechnet das Adverbialpronomen *da* (bzw., in Fragesätzen, das entsprechende Interrogativpronomen *wo*) ist, das als

Reparaturform gewählt wird, und nicht z.B. *dort*, *dann* oder ein noch komplexerer Ausdruck.¹⁷ Eine einfache Antwort wäre, daß *da* eben zufällig im Lexikon neben seinen anderen Funktionen auch als Expletiv gekennzeichnet ist. Eine etwas komplexere, letztlich aber eher befriedigende Antwort könnte auf die Generalisierung Bezug nehmen, daß Sprachen als Expletiva immer die Formen mit der am wenigsten spezifischen, allgemeinsten Bedeutung wählen. Dies ist so beim englischen V-Expletiv *do*, es ist so beim R-Pronomen *da* (das als adverbiales Pronomen eine ‘‘Allerweltsproform’’ (Altmann (1981)) mit extrem reduzierter Semantik ist), und wir werden weiter unten noch sehen, daß die Situation bei W-Expletiven in den verschiedensten Sprachen ähnlich ist. Wie in Fußnote 17 aus Kapitel 1 bereits erwähnt wurde, hat Grimshaw (1997) einen Weg ersonnen, VOLL-INT so zu formulieren, daß die Beschränkung nicht nur den Preis für das Auftreten eines Expletivs festlegt, sondern darüber hinaus auch noch dafür Sorge trägt, daß immer nur die Form mit der allgemeinst-möglichen Semantik als Expletiv Verwendung finden kann. Diese Idee wird in Kapitel 3 noch eingehender besprochen werden.

Abschließend sei auf ein Datum verwiesen, das unter der gegenwärtigen Analyse noch unerklärt ist. Für das Ausbleiben der Pronominaladverbbildung bei belebtem Personalpronomen war die Annahme unerlässlich, daß PRON-KRIT nur schwache Pronomina in die Wackernagelposition treibt. Wie aber kommt es dann, daß sich bei Ersetzung eines schwachen Pronomens *es* durch ein starkes Pronomen wie belebtes *sie* im Kontext (22) exakt dieselbe Grammatikalitätsverteilung ergibt, daß also auch hier Wackernagelbewegung obligatorisch ist?

- (23) a. Der Fritz hat sie₁ π gestern der Maria t₁ vorgestellt
 b. *Der Fritz hat – π gestern der Maria sie₁ vorgestellt
 c. *Der Fritz hat – π gestern der Maria da₁ vorgestellt

Es sieht hier so aus, als müßte man belebte Personalpronomina gleichzeitig PRON-KRIT unterwerfen und sie davon ausnehmen. Dieses Dilemma kann dadurch aufgelöst werden, daß die Beschränkung PRON-KRIT in eine Subhierarchie von PRON-KRIT-Teilbeschränkungen mit fixer interner Ordnung aufgelöst wird, die neben anderen (vgl. Müller (2000a)) die folgenden beiden Beschränkungen enthält (dieses Vorgehen ist wiederum vergleichbar dem von Schmid (1999) im Falle der *GE-V-Verbotsbeschränkungen).

- (24) a. PRON-KRIT^{schwach}:
 Ein schwaches NP-Pronomen steht in der Wackernagelposition.
 >> ... >>
 b. PRON-KRIT^{unbetont}:
 Ein unbetontes NP-Pronomen steht in der Wackernagelposition.

¹⁷ Allerdings existiert als etwas stärker restringierte Variante zum R-Pronomen *da* noch *hier*, wie in *hiermit*, *hieran*.

PRON-KRIT^{schwach} gilt dabei für *es*, während PRON-KRIT^{unbetont} für z.B. (unbetontes) belebtes *sie* gilt. Beide Beschränkungen sind über ÖKON geordnet und lösen so obligatorisch Wackernagelbewegung aus.¹⁸ Die feste Ordnung hat zur Folge, daß schwache NP-Pronomina wie *es* stärker in die Wackernagelposition streben als lediglich unbetonte NP-Pronomina wie belebtes *sie*. In den bisherigen Tabellen ist PRON-KRIT somit durch PRON-KRIT^{schwach} zu ersetzen; PRON-KRIT^{unbetont} ist dagegen unterhalb von PROJ-P geordnet, so daß sich am Wettbewerb T_{2.8} substantiell nichts ändert: Der optimale Kandidat läßt das NP-Pronomen in der NP, unter Verletzung der tief geordneten Beschränkung PRON-KRIT^{unbetont} und Erfüllung der höher geordneten Beschränkungen PP-INSEL, PRON-KRIT^{schwach} und PROJ-P.

2.4. Expletives ‘es’ im Deutschen

Ein viertes Beispiel involviert das deutsche Expletivpronomen *es*. Ein Standardproblem der deutschen Syntax ist es, zu erklären, warum ein echt expletives Pronomen *es* nur im Vorfeld, nicht aber im Mittelfeld erlaubt ist (vgl. z.B. Koster (1986), Grewendorf (1989), Cardinaletti (1990)). Dies zeigt der Kontrast in (25):¹⁹

- (25) a. Es haben viele Leute das Buch gelesen
 b. *daß es viele Leute das Buch gelesen haben

Wesentlich scheint hier die Beobachtung, daß (im Deklarativsatz) die Einsetzung von *es* im Spezifikator eines V/2-Kopfes in (25-a) erfolgen muß, um eine ansonsten ungrammatische Struktur zu “reparieren” (es sei denn, Topikalisierung eines Mittelfeld-internen Elements findet statt); vgl. (26-a). Gerade dies ist bei (25-b) anders: Die Struktur ist ohne *es*-Einsetzung wohlgeformt; vgl. (26-b):

- (26) a. *- haben viele Leute das Buch gelesen
 b. daß viele Leute das Buch gelesen haben

Die richtige Generalisierung könnte also diese sein: *Es*-Einsetzung im Deutschen ist nur als letzter Ausweg möglich, um eine Position zu füllen, die obligatorisch besetzt werden muß (nämlich SpecV/2 im Deklarativsatz), und ansonsten blockiert. Dies legt eine optimalitätstheoretische Analyse nahe: Wie die Einsetzung anderer Expletiva auch (z.B. des semantisch leeren Verbs *do* im Englischen, s.o.) verstößt *es*-Einsetzung gegen VOLL-INT (vgl. (15-e) aus Kapitel 1). Daher ist diese Opera-

¹⁸Betonte NP-Pronomina verhalten sich im Prinzip wie nicht-pronominale NPs; sie unterliegen einer PRON-KRIT-Teilbeschränkung, die unter ÖKON geordnet ist. Damit ist Wackernagelbewegung hier unmöglich.

¹⁹Angenommen wird hier, daß das *es* in (i) gar kein Expletiv, sondern ein Argument ist.

(i) Ich habe es nicht bedauert [CP daß ich das Buch gelesen habe]

Dementsprechend ist die CP hier auch appositiv und kein Argumentsatz. Vgl. Ross (1967), Bennis (1986) und Vikner (1995).

tion normalerweise blockiert, wie in (25-b)/(26-b). Allerdings gibt es eine mit dem Verbot der Expletiv-Einsetzung potentiell konfligierende Beschränkung, derzufolge das Vorfeld bei V/2 nicht radikal leer bleiben darf, sondern durch eine phonetisch realisierte Kategorie oder eine Spur davon sichtbar gemacht werden muß:

(27) SPECV/2:

Der Spezifikator von V/2 muß gefüllt sein.

Wie nun (25-a) vs. (26-a) zeigt, ist letztere Beschränkung im Deutschen offenbar ersterer übergeordnet: SPEC-V/2 \gg VOLL-INT (in Sprachen ohne obligatorische Vorfeldbesetzung bei V/2 würde das Gegenteil gelten).

Nimmt man nun weiterhin an, daß die Beschränkung VOLL-INT mit ÖKON im Deutschen gleich geordnet bzw. gekoppelt ist (VOLL-INT \circ ÖKON), dann ergibt sich, daß (28) (Vorfeldbesetzung durch Bewegung) zwar mit (25-a) (Vorfeldbesetzung durch *es*-Einsetzung) konkurriert, daß aber beide Kandidaten anders als (26-a) als optimal klassifiziert werden.

(28) Viele Leute₁ haben t₁ das Buch gelesen

Den optimalitätstheoretischen Wettbewerb bei der Vorfeldbesetzung zeigt Tabelle T_{2.10}.

T_{2.10}: Vorfeldbesetzung bei V/2

Kandidaten	SPECV/2	ÖKON VOLL-INT
☞K ₁ : Es haben viele Leute das Buch gelesen		*
☞K ₂ : Viele Leute ₁ haben t ₁ das Buch gelesen		*
K ₃ : – haben viele Leute das Buch gelesen	*!	

Daß die Einsetzung eines expletiven *es* im Mittelfeld nicht möglich ist (vgl. (25-b), (26-b)), folgt dann aus der aus Tabelle T_{2.10} ersichtlichen Beschränkungsordnung ohne Zusatzannahmen. Was ebenso folgt, ist, daß Spec *daß* niemals durch Expletiveinsetzung oder Topikalisierung gefüllt werden kann:

- (29) a. *(Ich glaube) [_{CP} es daß viele Leute das Buch gelesen haben]
 b. *(Ich glaube) [_{CP} viele Leute₁ daß t₁ das Buch gelesen haben]

All dies veranschaulicht Tabelle T_{2.11}: Sobald ein Kandidat die höher geordnete Beschränkung auch ohne die Verletzung einer tiefer geordneten Beschränkung erfüllen kann, wird die Verletzung der tiefer geordneten Beschränkung (in diesem Fall: von ÖKON oder VOLL-INT) sofort fatal.

2.5. Schluß

Die vorangegangenen Abschnitte haben plausibel zu machen versucht, daß eine optimalitätstheoretische Analyse gut mit dem Auftreten von syntaktischen Reparaturphänomenen zurecht kommt. Standardtheorien haben in diesem Bereich

T_{2.11}: Blockiertes ‘es’ und ‘daß’

Kandidaten	SPECV/2	ÖKON VOLL-INT
☞ K ₁ : daß viele Leute das Buch gelesen haben		
K ₂ : viele Leute ₁ daß t ₁ das Buch gelesen haben		*!
K ₃ : es daß viele Leute das Buch gelesen haben		*!
K ₄ : daß es viele Leute das Buch gelesen haben		*!

klassischerweise Probleme, denn die Idee, daß eine als exzeptionell und “markiert” ausgezeichnete Eigenschaft nur in bestimmten Kontexten auftreten kann, und zwar gerade dann, wenn die eigentlich zu erwartende Eigenschaft nicht auftreten kann, setzt voraus, daß Beschränkungen verletzbar und unterschiedlich wichtig sind, und daß ein Wettbewerb von Kandidaten besteht. Mit “exzeptioneller Eigenschaft” war in den hier behandelten Fällen immer eine bestimmte morphologische Form gemeint: das expletive Verb *do*, der Ersatzinfinitiv, das R-Pronomen *da*, das expletive Vorfeld-*es*. Tatsächlich sind viele Reparaturphänome in der Syntax von diesem Typ; unter die Rubrik “Reparaturformen” subsumierbar sind z.B. auch die resumptiven Pronomina, wie sie bei Legendre, Smolensky & Wilson (1998), Pesetsky (1998) und Vogel (1999) auftauchen (vgl. die Abschnitte 2 und 3 von Kapitel 3 und Abschnitt 4 von Kapitel 7). Zur Zeit noch weniger gut erforscht, aber prinzipiell ebenso einer optimalitätstheoretischen Analyse zugänglich ist ein anderer Typ von Reparaturphänomenen, die man als “Reparaturbewegungen” bezeichnen kann: Ein Bewegungsprozeß, der normalerweise in einer gegebenen Sprache auf ein gegebenes Element nicht angewandt werden kann, wird exzeptionell doch möglich, um eine höher geordnete Beschränkung zu erfüllen. Nur kurz sei ein Fall skizziert, der sich einer optimalitätstheoretischen Analyse als Reparaturphänomen in besonderer Weise anzubieten scheint, nämlich Quantorenanhebung als Reparaturbewegung in VP-Ellipsen-Konstruktionen im Englischen, wie sie von Fox (1995) vorgeschlagen worden ist.

Fox zeigt zunächst, daß auf der Ebene der Logischen Form (LF) Quantorenanhebung nur dann möglich ist, wenn als Resultat dieser Bewegung der bewegte Quantor Skopus über einen anderen Quantor gewinnt; d.h., semantisch leere Quantorenanhebung ist verboten. Diese Beschränkung können wir als Q-ÖKON (“Quantorenökonomie”) fassen. Darüber hinaus gibt es jedoch in VP-Ellipsen-Konstruktionen im Englischen eine sehr hoch geordnete Beschränkung PARALLELISMUS, die verlangt, daß die beiden Konjunkte parallel aufgebaut sind, d.h., wenn in einem Konjunkt Quantorenanhebung erfolgt, muß dies auch im anderen Konjunkt geschehen. (Die Grundannahme ist hier, daß die phonologisch fehlende Struktur im zweiten Konjunkt syntaktisch vorhanden und für LF-Bewegungen zugänglich ist.) So ist (30-a) nur zweifach (anstatt vierfach) mehrdeutig. Der Satz kann bedeuten, daß es (mindestens) einen Jungen gibt, der jeden Lehrer bewundert, und (mindestens) ein Mädchen, das auch jeden Lehrer bewundert (keine Quantorenanhebung); oder daß es für jeden Lehrer einen Jungen gibt, der ihn bewundert, und auch ein Mädchen,

das ihn bewundert (Quantorenanhebung von *every teacher* in beiden Konjunkten). Der Satz kann dagegen nicht etwa bedeuten, daß es für jeden Lehrer einen Jungen gibt, der ihn bewundert (aber z.B. keinen, der alle Lehrer bewundert), und daß es ein Mädchen gibt, das jeden Lehrer bewundert (Quantorenanhebung von *every teacher* im ersten, aber nicht im zweiten Konjunkt). Interessant ist nun, daß auch (30-b) zweifach ambig ist. Der Satz kann zum einen bedeuten: Mary bewundert jeden Lehrer, und es gibt darüber hinaus noch (mindestens) einen Jungen, der auch jeden Lehrer bewundert. Wichtiger ist im gegenwärtigen Zusammenhang aber die zweite, ebenfalls mögliche Lesart: Für jeden Lehrer gilt, daß Mary ihn bewundert, und daß ein Junge ihn bewundert (es muß nicht immer derselbe sein, wie bei der ersten Lesart). Hier muß zur Herstellung des weiten Skopus für *every teacher* im ersten Konjunkt semantisch leere Quantorenanhebung erfolgen, um PARALLELISMUS zu erfüllen. Diese Bewegung verletzt jedoch die tiefer geordnete Beschränkung Q-ÖKON. Das bedeutet: Hier ist ein Bewegungsschritt erfolgt, der normalerweise nicht erfolgen darf (weil er eine fatale Verletzung von Q-ÖKON involviert), der aber dann ausnahmsweise erlaubt ist, wenn dies die einzige Möglichkeit ist, eine noch höher geordnete Beschränkung (PARALLELISMUS) zu respektieren.

- (30) a. [_{CP₁} Some boy admires every teacher], and [_{CP₂} some girl does [_{VP} -]
too]
b. [_{CP₁} Mary admires every teacher], and [_{CP₂} some boy does [_{VP} -] too]

Zu diesem Fall wäre selbstverständlich noch weit mehr zu sagen. Die kurze Skizze des Kerns der Analyse von Fox (1995) mag aber ausreichen, um zu zeigen, daß in der optimalitätstheoretischen Syntax nicht nur solche Phänomene gut behandelt werden können, die Reparaturformen betreffen, sondern auch solche, die andere syntaktische Reparaturprozesse wie z.B. Bewegung involvieren.²⁰

3. Wettbewerb und Blockade

Ein dritter Typ von empirischer Evidenz legt den Schluß nahe, daß zwei Konstruktionen in ein und derselben Sprache miteinander konkurrieren, die an sich beide nur Eigenschaften aufweisen, von denen man weiß, daß sie im Prinzip (d.h., in anderen Konstruktionen) von der Grammatik erlaubt werden. Die Analyseidee ist dann, daß die eine Konstruktion die andere als suboptimal blockiert.

3.1. Komplementierertilgung im Englischen

Ein einfaches Beispiel ist die in Kapitel 1 schon erwähnte Alternation bzgl. Komplementierertilgung im Englischen (vgl. (20) aus Kapitel 1):

²⁰Eine ausführliche Behandlung dieses Falles sowie weiterer Instanzen von Reparaturbewegung findet sich in Heck & Müller (2000ab); vgl. auch Kapitel 8.

- (31) a. I think – John will leave
 b. I think that John will leave

Es sei für die gegenwärtigen Zwecke vorausgesetzt, daß die beiden Kandidaten Teil derselben Kandidatenmenge sind und daß die Alternation auf ein identisches Beschränkungsprofil zurückgeht. Dies ist Grimshaws (1997) Erklärung, und die soll hier zugrundegelegt werden (ganz ähnlich kann man aber auch im Rahmen von Pesetskys (1997; 1998) Analyse argumentieren, die die Alternation in (31) auf eine Kopplung von zwei Beschränkungen zurückführt).

Wie in Abschnitt 9.2 aus Kapitel 1 dargelegt, nimmt Grimshaw (1997) an, daß Sätze keine einheitliche Größe haben – sie sind minimal VPs, können aber auch IP, CPs oder noch größere Projektionen funktionaler Kategorien sein, wenn dies zu einem optimalen Beschränkungsprofil führt. Die Beschränkungen, die eine Rolle spielen, kennen wir zumindest vom Namen her alle schon:

- (32) a. PUR-EP:
 Adjunktion an die höchste XP einer eingebetteten erweiterten Projektion und Bewegung in deren Kopf sind verboten.
 b. OB-KOPF:
 Eine Projektion hat einen (nicht-leeren) Kopf.
 c. ÖKON:
 Bewegung ist verboten (Spur ist nicht erlaubt).

PUR-EP (“Purity of Extended Projection”) erscheint hier allerdings in der Originalformulierung von Grimshaw (1997) und nicht in der Form, die sich oben für die Erfassung parametrischer Variation bei W-Bewegung als nützlich erwiesen hat (vgl. (8)).²¹ Die Beschränkungsordnung für das Englische entspricht der Reihenfolge in (32). Tabelle T_{2.12} zeigt, warum die Sätze in (31) im Englischen trotz Mitgliedschaft in derselben Kandidatenmenge beide optimal sein können.

Weil das englische Futur die Präsenz eines Auxiliars bedingt, muß ein Kandidat hier mindestens eine IP (oder, neutraler gesagt, eine zusätzliche Projektion oberhalb der VP des Hauptverbs) sein, um optimal zu werden. Wie in Kapitel 1 gezeigt wurde, verlangen dann die Beschränkungen KASUS und SUBJEKT, daß die Subjekt-NP *John* in den Spezifikator des Auxiliars angehoben wird; diese Beschränkungen sind der Übersichtlichkeit wegen hier ausgelassen. Wesentlich ist nun,

²¹Die wesentlichen Unterschiede zwischen den beiden Versionen von PUR-EP sind (a), daß Grimshaws Beschränkung Adjunktion nicht invariant an CP und C⁰ verbietet, sondern an die höchste erweiterte Projektion von V, was immer dies ist (es kann sich aufgrund der variablen Satzstruktur zwar um eine CP handeln, aber, je nach Kontext, auch um eine VP, IP, etc.); und (b), daß Grimshaws Beschränkung derartige Adjunktionen nur im eingebetteten Fall verbietet, nicht im Wurzelsatz. Tatsächlich ließen sich die beiden Versionen von PUR-EP mit Hilfe einiger Zusatzannahmen vereinheitlichen. Da hier jedoch die Anhäufung von Evidenz für optimalitätstheoretische Analysen im Vordergrund steht, werde ich von derartigen Vereinheitlichungsversuchen absehen.

T_{2.12}: Komplementierertilgung im Englischen

Kandidaten	PUR-EP	OB-KOPF	ÖKON
☞ K ₁ : think [_{CP} that [_{IP} John ₁ will [_{VP} t ₁ leave]]]			*
☞ K ₂ : think [_{IP} John ₁ will [_{VP} t ₁ leave]]			*
K ₃ : think [_{CP} e [_{IP} John ₁ will [_{VP} t ₁ leave]]]		*!	*

daß der eingebettete Satz als IP ohne lexikalischen Komplementierer (K₂) dasselbe Beschränkungsprofil hat wie als CP mit lexikalischem Komplementierer (K₁) – beidesmal resultiert eine auf Subjektanhebung zurückzuführende Verletzung von ÖKON, aber keine Verletzung von PUR-EP, OB-KOPF oder sonst einer von Grimshaw (1997) angenommen Beschränkung.²² Anders liegt der Fall bei K₃, wo eine CP-Schale errichtet worden ist, die dann nicht durch einen Komplementierer *that* gefüllt wurde; hier ist fatal OB-KOPF verletzt.

Somit ergibt sich zunächst einmal, daß Konstruktionen sowohl mit Komplementierertilgung, als auch mit lexikalischem Komplementierer im Englischen möglich sind.²³ Dies bedeutet jedoch nicht, daß nicht die eine Option unter gewissen Umständen die andere als suboptimal blockieren kann, sobald weitere Beschränkungen ins Spiel kommen. In der Tat gibt es einige Kontexte, in denen die Optionalität des Komplementierers verloren geht. Eine solche Situation liegt z.B. vor bei bestimmten Matrixverben wie in (33), die nicht zur Klasse der Brückenprädikate gehören:

- (33) a. *He resented – John did it
 b. He resented that John did it

Ein zweiter Kontext, der Komplementierertilgung verbietet, wird durch Subjektivsätze gebildet (vgl. Stowell (1981) und die dort angegebene Literatur):

- (34) a. *[_{IP} He left so early] shows that he was tired
 b. [_{CP} That he left so early] shows that he was tired

Drittens ist Komplementierertilgung unmöglich bei eingebetteter Topikalisierung:

²²Dies bedeutet natürlich, daß der traditionelle Begriff der Komplementierer- *Tilgung* in Grimshaws Analyse streng genommen nicht mehr einschlägig ist – wo der Komplementierer fehlt, war er von vornherein nicht da. Als bloßen Namen für die Konstruktion werde ich diesen Begriff aber auch bei Grimshaws Analyse weiterhin benutzen.

²³Das Geheimnis hinter diesem auf den ersten Blick vielleicht erstaunlichen Resultat ist, daß die Präsenz eines Komplementierers *that* an sich vollkommen kostenfrei ist – es wird dadurch gegenüber der bloßen IP-Einbettung keine weitere Beschränkung verletzt. Dies präsupponiert (wie in Kapitel 1 bereits erwähnt) u.a., daß *that* anders als expletives *do* nicht VOLL-INT verletzt. Nichtsdestoweniger werden wir in Kapitel 3 und in Kapitel 5 noch sehen, daß die Annahme, daß eine Präsenz von *that* gar keine Beschränkungen verletzt, in Grimshaws eigener Analyse kaum aufrechterhalten werden kann, und daß folglich für die Daten in (31) ein Problem entsteht.

- (35) a. *I think – to John₁ Mary gave a book t₁
 b. I think [_C that] to John₁ Mary gave a book t₁

Dasselbe gilt für die Präsenz von eingebetteten Satzadjunkten (vorausgesetzt sei hier, daß diese Adjunkte nicht durch Bewegung in ihre Position gelangen – ansonsten handelte es sich um eine Variante des dritten Kontexts):²⁴

- (36) a. I think [_{CP} that [_{IP} most of the time [_{IP} John could accept this solution]]]
 b. *I think [_{IP} most of the time [_{IP} John could accept this solution]]

In allen diesen Fällen kann Komplementierertilgung dadurch blockiert werden, daß es eine im Wettbewerb befindliche Variante mit Komplementierer gibt, die ein besseres Beschränkungsprofil aufweist. Dies sei exemplarisch anhand von Grimshaws Erklärung der fehlenden Optionalität in (36) gezeigt; vgl. Tabelle T_{2.13}.

T_{2.13}: Komplementierertilgung und Satzadjunkte

Kandidaten	PUR-EP	OB-KOPF	ÖKON
☞ K ₁ : V [_{CP} that [_{IP} Adjunkt [_{IP} NP ₁ I t ₁ V NP]]]			*
K ₂ : V [_{IP} Adjunkt [_{IP} NP ₁ I t ₁ V NP]]	*!		*
K ₃ : V [_{CP} e [_{IP} Adjunkt [_{IP} NP ₁ I V t ₁ NP]]]		*!	*

Hier führt Präsenz vs. Fehlen von *that* zu einem unterschiedlichen Beschränkungsprofil: Da *most of the time* per Annahme an IP adjungiert ist (das Adverbial geht dem Subjekt voran und folgt einem Komplementierer), muß ein Kandidat mit bloßer IP-Einbettung hier fatal PUR-EP verletzen; diese Verletzung ist vermieden, wenn es oberhalb der IP noch eine weitere Projektion (CP) gibt, die dann die höchste erweiterte Projektion des Verbs ist und auf diese Weise die tiefere IP vor einer Verletzung von PUR-EP schützt.

3.2. Partielle W-Bewegung im Deutschen

Konstruktionen mit partieller W-Bewegung (W-W-Konstruktionen) im Deutschen (vgl. (37-a)) koexistieren normalerweise mit Konstruktionen mit sukzessiv-zyklischer W-Bewegung (vgl. (37-b)).

- (37) a. Was₁ glaubst du [_{CP} wann₁ sie t₁ gekommen ist] ?
 b. Wann₁ glaubst du [_{CP} t'₁ daß sie t₁ gekommen ist] ?

Bei partieller W-Bewegung wird die W-Phrase in eine [-w] markierte SpecC-Position verschoben; in der [+w] markierten Zielposition (bei längeren Abhängigkeiten

²⁴Bei (35) und (36) gibt es in manchen Varietäten des Englischen Variation dergestalt, daß auch die Komplementierer-freien Versionen akzeptabel scheinen; vgl. z.B. Rochemont (1989). Hiervon wird im folgenden abgesehen, auch wenn sich dieser Sachverhalt unter Umständen für die Grimshawsche Analyse als problematisch erweisen könnte.

optional auch in eingebetteten weiteren [-w] markierten SpecC-Positionen) wird ein sog. Skopusmarker *was* in SpecC eingesetzt, der als W-Expletiv klassifiziert werden kann (vgl. Riemsdijk (1982), Stechow & Sternefeld (1988), McDaniel (1989) sowie die Aufsätze in Lutz, Müller & Stechow (2000), die auch andere Analysemöglichkeiten für das hier auftretende *was* diskutieren).

Erneut kann die Optionalität durch die Annahme einer Kopplung zweier Beschränkungen gewährleistet werden. Tatsächlich sieht es so aus, als könne die relevante Kopplung gerade die oben schon für das Deutsche vorgeschlagene sein, nämlich ÖKON \circ VOLL-INT (vgl. Müller (1997)): Ein W-Bewegungsschritt verletzt ÖKON, die Einsetzung eines – per Annahme expletiven – Skopusmarkers verletzt VOLL-INT, und somit sind beide Operationen als gleich teuer identifiziert. Des weiteren müssen wir annehmen, daß beide Konstruktionen gleich gut die Beschränkung erfüllen können, die W-Bewegung auslöst. Diese Rolle hat oben W-KRIT gespielt (vgl. (7)); aber diese Beschränkung hat bisher nur gefordert, daß eine W-Phrase auf der S-Struktur im Bereich von C steht – offen gelassen war damit, ob C ein [+w]-Merkmal oder ein [-w]-Merkmal hat. Was wir nun brauchen, ist eine Erweiterung des W-Kriteriums, die dafür sorgt, daß zum einen W-Phrasen in (irgendeiner) C-Domäne stehen, daß zum anderen aber auch der eigentliche C_[+w]-Kopf ein geeignetes lexikalisches Element in seiner Domäne fordert (vgl. neben anderen May (1985) und Lasnik & Saito (1992), wo das W-Kriterium in zwei separate, konjunktiv verbundene Beschränkungen geteilt wird). Lange W-Bewegung erfüllt beide Forderungen simultan durch die Lokalisierung des W-Elements in der SpecC_[+w]-Position; partielle W-Bewegung erfüllt aber ebenso beide Forderungen – die erste durch partielle W-Bewegung in eine SpecC_[-w]-Position, die zweite durch die Einsetzung eines Skopusmarkers *was* in SpecC_[+w]. Die letzte Beschränkung, die eine Rolle spielen wird, ist schließlich die Barrierenbeschränkung BAR, derzufolge Kettenbildung nicht über eine Barriere hinweg erfolgen darf (vgl. Chomsky (1986a)). Die vier Beschränkungen sind im folgenden aufgelistet; die Reihenfolge entspricht ihrer Ordnung im Deutschen:

- (38) a. W-KRIT (modifiziert gegenüber (7)):
- (i) Eine W-Phrase muß auf der S-Struktur in der Domäne von C stehen.
 - (ii) C_[+w] verlangt ein geeignetes lexikalisches Element in seiner Domäne.
- b. BAR:
Kettenbildung darf keine Barriere überschreiten.
- c. ÖKON:
Bewegung ist verboten (Spur ist nicht erlaubt).
- d. VOLL-INT:
Expletiveinsetzung ist verboten.

Hieraus ergibt sich, daß (37-a) und (37-b) im selben Wettbewerb optimal sein können. Ebenso folgt, daß bei Präsenz eines Skopusmarkers trotzdem noch partielle W-Bewegung notwendig ist; vgl. Tabelle T_{2.14}. Versucht ein Kandidat, die

beiden bei der Kopplung von ÖKON und VOLL-INT bei K_1 und K_2 entstehenden Verletzungen durch Weglassen des Skopusmarkers und/oder der W-Bewegung zu reduzieren bzw. zu vermeiden, ergibt sich automatisch eine fatale Verletzung der höher geordneten Beschränkung W-KRIT.

T_{2.14}: Partielle und lange W-Bewegung im Deutschen

Kandidaten	W-KRIT	BAR	ÖKON VOLL-INT
$\Rightarrow K_1$: $[_{CP}$ was ... $[_{CP}$ wann ₁ ... t ₁			* *
$\Rightarrow K_2$: $[_{CP}$ wann ₁ ... $[_{CP}$ t' ₁ ... t ₁			**
K_3 : $[_{CP}$ was ... $[_{CP}$ - ... wann ₁	*!		*
K_4 : $[_{CP}$ - ... $[_{CP}$ - ... wann ₁	*!*		
K_5 : $[_{CP}$ - ... $[_{CP}$ wann ₁ ... t ₁	*!		*

Somit ist abgeleitet, daß partielle W-Bewegung (mit Skopusmarkereinsatz) ebenso wie sukzessiv-zyklische W-Bewegung einer W-Phrase zunächst einmal im Deutschen koexistierende, reguläre Möglichkeiten sind, lange W-Abhängigkeiten auszudrücken. Interessant ist nun, daß trotz der auf die Kopplung zurückführbaren grundsätzlichen Optionalität der beiden Konstruktionen die eine die andere in bestimmten Kontexten als suboptimal blockieren kann. Dies gilt etwa für Subjekt-Insel-Kontexte wie in (39-a) (unmögliche partielle W-Bewegung) gegenüber (39-b) (halbwegs akzeptable sukzessiv-zyklische Bewegung); vgl. McDaniel (1989):

- (39) a. *Was₁ ist es schade $[_{CP}$ $[_{PP}$ mit wem]₁ Hans t₁ gesprochen hat] ?
 b. ? $[_{PP}$ Mit wem]₁ ist es schade $[_{CP}$ t'₁ daß Hans t₁ gesprochen hat] ?

In Müller (1997) wird argumentiert, daß (39-a) gegenüber (39-b) mehr Verletzungen der Lokalitätsbedingung BAR aufweist. Konkret wird angenommen, daß in (39-b) nur eine BAR-Verletzung durch S-strukturelle Bewegung erfolgt, während in (39-a) zwei BAR-Verletzungen resultieren, und zwar durch S-strukturelle Kettenbildung sowie LF-Anhebung der W-Phrase in ihre Zielposition. Dies zeigt Tabelle T_{2.15}.²⁵

T_{2.15}: Verbot partieller W-Bewegung in Subjektsätzen im Deutschen

Kandidaten	W-KRIT	BAR	ÖKON VOLL-INT
K_1 : $[_{CP}$ was ... $[_{CP}$ mit wem ₁ ... t ₁		**!	* *
$\Rightarrow K_2$: $[_{CP}$ mit wem ₁ ... $[_{CP}$ t' ₁ ... t ₁		*	**
K_3 : $[_{CP}$ was ... $[_{CP}$ - ... mit wem ₁	*!	**	*
K_4 : $[_{CP}$ - ... $[_{CP}$ - ... mit wem ₁	*!*	*	
K_5 : $[_{CP}$ - ... $[_{CP}$ mit wem ₁ ... t ₁	*!	*	*

²⁵In diesem Ansatz wird davon ausgegangen, daß vollständige Derivationen miteinander im Wettbewerb sind, nicht nur z.B. S-Struktur-Repräsentationen. Entsprechend stehen die Repräsentationen in T_{2.15} auch nur für einen Ausschnitt aus den Derivationen K_1 , K_2 etc.

Damit bricht in Inselkontexten das identische Beschränkungsprofil von partieller und langer W-Bewegung zusammen, und erstere Option wird durch letztere als suboptimal ausgeschlossen. Diese Erklärung beruht ganz wesentlich auf dem Konzept des Wettbewerbs: Stünden partielle und lange W-Bewegung nicht miteinander in Konkurrenz, dann bestünde keine Möglichkeit, erstere in Inselkontexten durch letztere zu blockieren.²⁶

Die Annahme eines Wettbewerbs von langer und partieller W-Bewegung erlaubt auch eine einfache Erklärung des Faktums, daß in vielen norddeutschen Varietäten partielle W-Bewegung tatsächlich mehr oder weniger die einzige Möglichkeit zum Ausdruck langer W-Abhängigkeiten zu sein scheint (hier ist die Kopplung in eine eindeutige Ordnung ÖKON \gg VOLL-INT aufgelöst).²⁷ Riemsdijk (1982, 13) stellt in diesem Zusammenhang fest, daß "die Existenz des Skopusmarkiers *was* etwas zu tun hat mit der Unmöglichkeit langer Bewegung" in diesen Varietäten; diese Sicht geht natürlich wesentlich von einer Wettbewerbssituation der beiden Strategien aus, wie sie durch die Optimalitätstheorie explizit gemacht werden kann.

4. Das Konzept des unmarkierten Falles

Beim vierten Typ von empirischer Evidenz lassen sich die Daten durch eine Generalisierung des folgenden Typs beschreiben: Wenn nicht von anderen Beschränkungen der Grammatik etwas Gegenteiliges erfordert wird, zeichnet sich eine Konstruktion durch die Eigenschaft *P* aus; m.a.W.: Die Eigenschaft *P* zu besitzen ist der unmarkierte Fall. Die Analyseidee ist hier, daß die Eigenschaft *P* einer Konstruktion durch eine sehr tief geordnete Beschränkung erzwungen wird. Die Effekte dieser Beschränkung werden häufig durch die Effekte von konfligierenden, höher geordneten Beschränkungen überlagert. Nur wenn alle potentiell mit ihr im Konflikt befindlichen übergeordneten Beschränkungen bei einem gegebenen Kandidaten keine Rolle spielen, tritt die tief geordnete Beschränkung in Kraft, was den Eindruck des "unmarkierten" Falles hervorruft (vgl. hierzu insbesondere McCarthy & Prince (1994), die von "emergence of the unmarked" reden).

4.1. Ökonomie der Bewegung

Eine Beschränkung wie ÖKON ist typischerweise ziemlich niedrig eingeordnet und wird so ständig aufgrund der Erfordernisse von höher geordneten Beschränkungen

²⁶Warum der optimale Kandidat (39-b) nicht vollkommen akzeptabel ist, muß unter Standard-OT-Annahmen unklar bleiben; s.u.

²⁷Mehr wäre dann allerdings zur Vorfeldbesetzung zu sagen, denn auch in norddeutschen Varietäten koexistieren Vorfeld-*es* und Topikalisation. Eine mögliche Lösung wäre es, von zwei unterschiedlichen VOLL-INT-Beschränkungen auszugehen, von denen in der einen Grammatik nur eine mit ÖKON gekoppelt ist, während die Kopplung mit ÖKON in der anderen Grammatik beide VOLL-INT-Beschränkungen betrifft.

verletzt (W-KRIT, SPECV/2 im Deutschen etwa). Wenn aber zwei Kandidaten ein ansonsten identisches Beschränkungsprofil haben, dann wird ÖKON wirksam und prognostiziert korrekt den unmarkierten Fall. Eine solche Situation ist möglicherweise bei Fällen von ketteninvariantem Scrambling im Deutschen gegeben, die es seit Ross (1967) auszuschließen gilt. Das Problem ist, daß unter der üblichen Annahme, Scrambling sei iterierbar (z.B., weil es Adjunktion an VP ist), von der Theorie zunächst einmal Fälle von ständig wiederholtem, an der eigentlichen PF-Kette nichts änderndem Scrambling erlaubt sein sollten, daß also für jede gegebene Argumentabfolge eine nicht-finite Menge an wohlgeformten Derivationen zur Verfügung stehen sollte. Eine solche leere Ambiguität erscheint höchst unplausibel, und deshalb ist man stets darum bemüht gewesen, ketteninvariantes Scrambling zu verbieten. Das Problem wird veranschaulicht durch die verschiedenen Analysen ein und desselben Satzes in (40).²⁸

- (40) a. daß der Fritz₁ der Maria₂ das Buch₃ überlassen hat
 b. *daß der Fritz₁ der Maria₂ das Buch₃ t₁ t₂ t₃ überlassen hat
 c. *daß der Fritz₁ der Maria₂ das Buch₃ t₁ t₂ t₃ t₁ t₂ t₃ überlassen hat
 d. *daß der Fritz₁ der Maria₂ das Buch₃ t₂ t₃ t₁ t₂ t₃ überlassen hat
 e. *daß der Fritz₁ der Maria₂ das Buch₃ t₁ t₃ t₁ t₂ t₃ überlassen hat
 f. *daß der Fritz₁ der Maria₂ das Buch₃ t₁ t₂ t₁ t₂ t₃ überlassen hat
 g. *daß der Fritz₁ der Maria₂ das Buch₃ t₁ t₁ t₂ t₃ t₁ t₂ t₃ überlassen hat
 h. *daß der Fritz₁ t₁ der Maria₂ das Buch₃ überlassen hat
 i. *daß der Fritz₁ der Maria₂ t₁ t₂ das Buch₃ überlassen hat

Wie Tabelle T_{2.16} zeigt, folgt aus einer zwar tief geordneten (und somit durch Beschränkungen wie W-KRIT oder SPECV/2 außer Kraft zu setzenden) Beschränkung ÖKON ohne weitere Annahmen das Verbot von ketteninvariantem Scrambling: Wenn das restliche Beschränkungsprofil aller Kandidaten in (40) gleich ist (was hier exemplarisch durch die abstrakten Beschränkungen A, B und C illustriert ist), dann wird eine ansonsten leicht in Kauf zu nehmende ÖKON-Verletzung fatal; mit anderen Worten, im unmarkierten Fall findet keine Bewegung statt.

Tatsächlich erweist sich, daß in T_{2.16} K₁ gegen seine Wettbewerber unter jeder beliebigen Beschränkungsordnung gewinnt. Es läßt sich somit an dieser Stelle ein weiterer wichtiger optimalitätstheoretischer Begriff einführen, der ein bestimmtes Verhältnis zweier Kandidaten beschreibt, nämlich der Begriff der "harmonischen Begrenzung" ("harmonic bounding").

²⁸Eine hier im Einklang mit dem Großteil der einschlägigen Literatur gemachte Annahme ist, daß die Grundabfolge bei Verben wie *überlassen* im Deutschen Dativ- vor Akkusativobjekt ist. In Kapitel 6 wird argumentiert werden, daß diese Annahme nicht aufrechtzuerhalten ist, und daß vielmehr Akkusativ- vor Dativobjekt die uniforme D-strukturelle Abfolge bei Doppel-Objekt-Verben im Deutschen ist. Für die gegenwärtigen Zwecke ist dies aber irrelevant.

T_{2.16}: Verbot von ketteninvariantem Scrambling

Kandidaten	A	B	C	ÖKON
☞ K ₁ : daß der Fritz ₁ der Maria ₂ das Buch ₃ V		*	**	
K ₂ : daß der Fritz ₁ der Maria ₂ das Buch ₃ t ₁ t ₂ t ₃ V		*	**	*! **
K ₃ : daß der Fritz ₁ der Maria ₂ das Buch ₃ t ₁ t ₂ t ₃ t ₁ t ₂ t ₃ V		*	**	*!*****
K ₄ : daß der Fritz ₁ der Maria ₂ das Buch ₃ t ₂ t ₃ t ₁ t ₂ t ₃ V		*	**	*!*****
K ₅ : daß der Fritz ₁ t ₁ der Maria ₂ das Buch ₃ V		*	**	*!
K ₆ : daß der Fritz ₁ der Maria ₂ t ₁ t ₂ das Buch ₃ V		*	**	*!*

(41) *Harmonische Begrenzung:*

Ein Kandidat K₁ begrenzt einen anderen Kandidaten K₂ in derselben Kandidatenmenge harmonisch gdw. gilt:

- Es gibt keine Beschränkung, die K₂ besser erfüllt als K₁.
- Es gibt wenigstens eine Beschränkung, die K₁ besser erfüllt als K₂.

Somit begrenzt K₁ in T_{2.16} K₂, K₃, K₄, K₅ und K₆ harmonisch. Aus dem Begriff der harmonischen Begrenzung läßt sich zusammen mit dem Begriff des Beschränkungsprofils folgendes Theorem ableiten:

(42) *Theorem der harmonischen Begrenzung:*

Wenn K₁ K₂ harmonisch begrenzt, dann hat K₁ unter jeder Beschränkungsordnung ein besseres Beschränkungsprofil als K₂.

Dies ist eine Möglichkeit, im Rahmen der Optimalitätstheorie syntaktische Universalien abzuleiten. Harmonisch begrenzte Kandidaten können unter keiner Beschränkungsordnung optimal werden, und da Beschränkungsordnung das optimalitätstheoretische Mittel zur Erfassung einzelsprachlicher Variation ist, bedeutet dies: Harmonisch begrenzte Kandidaten können in keiner Sprache grammatisch sein.

4.2. Default-Kasus

Im unmarkierten Fall, d.h., wenn nicht von den Prinzipien der Kasuszuweisung (oder der Kasuskongruenz) etwas anderes festgelegt wird, erhalten NPs im Deutschen den Nominativ: Das ist der Default-Kasus (vgl. z.B. Fanselow (1991)). Der Default-Nominativ manifestiert sich unter anderem in *als*-PPs (sofern keine Kasuskongruenz vorliegt), in Infinitiven bei Kasuskongruenz mit PRO, bei Linksversetzung (wenn keine Kasuskongruenz vorliegt), oder bei prädikativen *und*-Konstruktionen (vgl. Kiss (1994), Reis (1995)). Dies zeigen der Reihe nach die Beispiele in (43).

- die Ehrung des Kanzlers als großer Politiker/*großen Politiker
- Wir baten die Männer [_{CP} PRO einer nach dem anderen/*einen nach dem anderen durch die Sperre zu gehen]

- c. Der Kaiser/*Den Kaiser, dem verdanken wir nichts
- d. Der/*Den und ein Buch lesen? (Daß ich nicht lache!)

In typischen Kasuszuweisungskontexten kann der Nominativ natürlich nicht optional erscheinen; dort ist er zugunsten des zugewiesenen Kasus blockiert. Willkürlich herausgegriffene Beispiele sind hierfür etwa die Akkusativzuweisung an ein direktes Objekt (vgl. (44-a)), die lexikalisch markierte Genitivzuweisung an ein obliquus Objekt (vgl. (44-b)), oder die Akkusativzuweisung an eingebettete Infinitivsubjekte in A.c.I.-Konstruktionen (sog. “Exceptional Case Marking” (ECM), vgl. (44-c)).

- (44) a. *daß ich er getroffen habe
 b. *daß man der Mann gedachte
 c. *daß wir der Mann das Buch lesen sehen

Daß im unmarkierten Fall im Deutschen auf einer NP der Nominativ erscheint, folgt aus einer tief geordneten Beschränkung wie (45):

- (45) DEFAULT-NOM:
 Eine NP trägt Nominativ.

Diese Beschränkung wird im Normalfall durch höher geordnete Kasuszuweisungsbeschränkungen überlagert (etwa der Art “In SpecV wird Akkusativ zugewiesen”), aber wenn nichts anderes gesagt wird, tritt DEFAULT-NOM in Kraft und sagt korrekt den unmarkierten Fall vorher.²⁹

4.3. Parallelverschiebung

Als ein drittes Beispiel dafür, wie eine tief geordnete Beschränkung in einer Optimalitätstheoretischen Syntax zwar in vielen Konstruktionen durch die Anforderungen höherrangiger Beschränkungen überlagert werden, aber dennoch bei ansonsten identischem Beschränkungsprofil den unmarkierten Fall vorhersagen kann, sei die Beschränkung PAR-VER (“Parallel-Verschiebung”) diskutiert, die in Müller (1999a) ausführlich motiviert wird.

- (46) PAR-VER:
 Wenn α β auf der Ebene L_n c-kommandiert, dann c-kommandiert α β auch auf der Ebene L_{n+1} (wobei α , β Argumente sind).

Die Idee hinter PAR-VER ist die, daß die relative Ordnung der Argumente in einem Satz im unmarkierten Fall in der Derivation bewahrt werden muß. Als für (46) relevante Ebenen mögen gelten: D-Struktur, S-Struktur, LF.³⁰ PAR-VER hat

²⁹Das Verhältnis von Kasus und Optimalitätstheorie wird genauer diskutiert in Kapitel 7.

³⁰Formal gesehen ist PAR-VER somit eine sog. “globale Regel” im Sinne von Lakoff (1971), d.h., eine Beschränkung, die weder als reiner Oberflächenfilter fungiert (der z.B. auf der S-Struktur kon-

in der syntaktischen Literatur einige Vorgänger. So schlagen Lakoff (1971), Kroch (1974), Huang (1982) und Reinhart (1983) vor, daß die Determination von relativem Quantorenskopos auf der Ebene der semantischen Interpretation sich im unmarkierten Fall nach den auf der syntaktischen Oberfläche bestehenden relativen C-Kommando-Verhältnissen richtet, d.h., ordnungsbewahrend ist. Vergleichbare Beschränkungen sind von Watanabe (1992) zur Behandlung von Superiorität und sog. Anti-Superiorität im Englischen und Japanischen vorgeschlagen worden, sowie von Meinunger (1995) und Haegeman (1995) zur Behandlung von mehrfacher A-Bewegung in Kasuspositionen hinein. All diesen Vorgängerbeschränkungen gemeinsam ist jedoch, daß sie konstruktionsspezifisch formuliert werden mußten (also z.B. nur Quantoren betrafen, oder nur Kasusbewegung); der Grund dafür ist, daß sie ansonsten als verletzbar hätten betrachtet werden müssen, und dieses Konzept stand in den jeweils angenommenen Theorien nicht zur Verfügung. In einem Optimalitätstheoretischen Rahmen kann jedoch PAR-VER in ganz allgemeiner Form formuliert werden und so eine einheitliche Erklärung von Ordnungsbewahrungssphänomenen gewährleisten, die bei den unterschiedlichsten Bewegungen auftreten.

Konkret ist PAR-VER eine Treuebeschränkung, die normalerweise ziemlich tief geordnet sein muß. Daher können ihre Effekte nur unter "günstigen Umständen" beobachtet werden; in der Regel wird PAR-VER durch höher geordnete Markiertheitsbeschränkungen überlagert, gerade so wie ÖKON.

Ein relevanter Anwendungskontext von PAR-VER ist die mehrfache W-Bewegung im Bulgarischen. Man erinnere sich aus Abschnitt 1.2 (Tabelle T_{2.4}), daß im Bulgarischen unter der Ordnung W-KRIT \gg PUR-EP \gg ÖKON in Mehrfachfragen derjenige Kandidat optimal ist, der alle W-Phrasen in die Domäne von C bewegt, um so maximal W-KRIT zu erfüllen, auch wenn das mit (im Deutschen fatalen) Verletzungen von PUR-EP einhergeht (wegen CP-Adjunktion). Aus der bisherigen Analyse folgt jedoch noch nichts über die Abfolge der vorangestellten W-Phrasen im Bulgarischen, und so würden wir zunächst einmal erwarten, daß jede Abfolge möglich ist. Dies ist jedoch, wie Rudin (1985; 1988) festgestellt hat, nicht der Fall. Vielmehr muß normalerweise die Abfolge der W-Phrasen in der C-Domäne im Bulgarischen gerade derjenigen Abfolge entsprechen, die unabhängig als die D-strukturelle Grundabfolge in der IP motiviert werden kann. Bei drei Argumenten eines Verbs wie *fragen* ist die Grundabfolge *wer* vor *wen* vor *was*, und genau diese

trolliert werden kann), noch als rein derivationale Beschränkung (die etwa die Anwendung einer Bewegungsregel in der Derivation lizenziert oder verbietet), sondern nicht-adjazente Repräsentationen (im vorliegenden Fall sind dies Repräsentationsebenen) in einer Derivation miteinander in Beziehung setzt; die Kandidaten müssen folglich hinreichend komplexe Objekte sein. In dieser Hinsicht hat PAR-VER somit denselben Status wie Chomskys (1981, 38) Projektionsprinzip, demzufolge lexikalisch determinierte Subkategorisierungseigenschaften von einer Ebene zur nächsten bewahrt werden müssen. Wie gesehen ist es tatsächlich auch möglich, Chomskys Projektionsprinzip optimalitätstheoretisch als verletzbare und geordnete Beschränkung zu erfassen; vgl. (13-b), (18-c).

Abfolge muß, wie die Daten in (47) zeigen, auch in der C-Domäne vorliegen – alle anderen Abfolgen sind ungrammatisch.

- (47) a. $[_{CP} \text{Koj}_1 \text{kogo}_2 \text{kakvo}_3 \text{C} [_{IP} \text{t}_1 \text{e pital } \text{t}_2 \text{t}_3]] ?$
 wer wen was fragte
 b. $*[_{CP} \text{Koj}_1 \text{kakvo}_3 \text{kogo}_2 \text{C} [_{IP} \text{t}_1 \text{e pital } \text{t}_2 \text{t}_3]] ?$
 c. $*[_{CP} \text{Kakvo}_3 \text{koj}_1 \text{kogo}_2 \text{C} [_{IP} \text{t}_1 \text{e pital } \text{t}_2 \text{t}_3]] ?$
 d. $*[_{CP} \text{Kakvo}_3 \text{kogo}_2 \text{koj}_1 \text{C} [_{IP} \text{t}_1 \text{e pital } \text{t}_2 \text{t}_3]] ?$
 e. $*[_{CP} \text{Kogo}_2 \text{koj}_1 \text{kakvo}_3 \text{C} [_{IP} \text{t}_1 \text{e pital } \text{t}_2 \text{t}_3]] ?$
 f. $*[_{CP} \text{Kogo}_2 \text{kakvo}_3 \text{koj}_1 \text{C} [_{IP} \text{t}_1 \text{e pital } \text{t}_2 \text{t}_3]] ?$

Nehmen wir nun an, daß die bisher postulierte Beschränkungsordnung im Bulgarischen (vgl. (9-c)) wie in (48) durch eine tief geordnete Beschränkung PAR-VER ergänzt wird.

- (48) *Ordnung im Bulgarischen* (erweitert):
 W-KRIT \gg PUR-EP \gg ÖKON \gg PAR-VER

Dann ergibt sich wie erwünscht als Konsequenz, daß bei Mehrfachfragen im Bulgarischen der optimale Kandidat der ist, in dem alle W-Phrasen auf der S-Struktur in die C-Domäne bewegt werden und sich dort wieder in der D-strukturellen Abfolge versammeln, wie durch PAR-VER gefordert; also hier K_1 in $T_{2.17}$. Die Kandidaten K_5 , K_6 usw. erfüllen W-KRIT zwar ebenso gut wie K_1 , und sie weisen auch nicht mehr Verletzungen von PUR-EP und ÖKON auf, aber sie verletzen die tief geordnete Treuebeschränkung PAR-VER in fataler Weise.³¹

$T_{2.17}$: Mehrfache W-Bewegung im Bulgarischen mit Ordnungsbewahrung

Kandidaten	W-KRIT	PUR-EP	ÖKON	PAR-VER
$\Rightarrow K_1: [_{CP} \text{koj}_1 \text{kogo}_2 \text{kakvo}_3 \dots \text{t}_1 \text{t}_2 \text{t}_3]$		**	***	
$K_2: [_{CP} \text{koj}_1 \dots \text{t}_1 \text{kogo}_2 \text{kakvo}_3]$	*!*		*	
$K_3: [_{CP} \text{koj}_1 \text{kogo}_2 \dots \text{t}_1 \text{t}_2 \text{kakvo}_3]$	*!	*	**	
$K_4: [_{CP} - \dots \text{koj}_1 \text{kogo}_2 \text{kakvo}_3]$	*!**			
$K_5: [_{CP} \text{koj}_1 \text{kakvo}_3 \text{kogo}_2 \dots \text{t}_1 \text{t}_2 \text{t}_3]$		**	***	*!
$K_6: [_{CP} \text{kakvo}_3 \text{koj}_1 \text{kogo}_2 \dots \text{t}_1 \text{t}_2 \text{t}_3]$		**	***	*!*

Zwei Bemerkungen noch zu diesem Wettbewerb. Erstens begrenzt K_1 die suboptimalen Kandidaten K_5 und K_6 in $T_{2.17}$ harmonisch. Wenn es also keine anderen Beschränkungen in der Grammatik gibt, die diese Kandidaten besser abschneiden lassen als K_1 , dann ergibt sich die Vorhersage, daß Abfolgen wie in K_5 und K_6 in

³¹Tabelle $T_{2.4}$ wäre dementsprechend so zu ergänzen, daß ein vom dortigen K_3 sich nur durch die W-Phrasen-Abfolge unterscheidender Kandidat (Objekt vor Subjekt statt Subjekt vor Objekt) ebenfalls durch eine fatale PAR-VER-Verletzung als suboptimal ausgeschlossen wird.

keiner Sprache optimal sein können.³²

Zweitens wird K_3 zwar von keinem Mitbewerber allein harmonisch begrenzt, aber man kann leicht sehen, daß dieser Kandidat trotzdem unter keiner Ordnung der vier involvierten Beschränkungen optimal sein kann. Wenn man die Redeweise von der harmonischen Begrenzung beibehalten will, dann kann man sagen, daß K_3 nicht von einem einzelnen Kandidaten, sondern von einer *Menge* von Kandidaten harmonisch begrenzt wird; vgl. Samek-Lodovici & Prince (1999). Diese Menge ist $\{K_1, K_2, K_4\}$, umfaßt also die möglichen Gewinner unter einer Beschränkungsordnung. Somit wird vorausgesagt, daß es zwar Sprachen gibt, die in Mehrfachfragen nur eine W-Phrase bewegen, und Sprachen, die alle W-Phrasen bewegen, aber keine Sprachen, die genau zwei W-Phrasen bewegen bzw. genau eine W-Phrase in situ verharren lassen – Kandidaten, die dies tun, sind kollektiv harmonisch begrenzt und daher “komplexe Verlierer” (‘complex losers’), wie Samek-Lodovici und Prince das nennen. (Zumindest gilt dies wiederum, solange sichergestellt ist, daß keine weiteren Beschränkungen, z.B. des Typs “Genau zwei W-Phrasen müssen bewegt werden”, doch noch K_3 favorisieren können.)

Wie erwähnt bedarf es günstiger Umstände, um PAR-VER bei der Arbeit zu sehen. Diese sind bei Mehrfachfragen im Bulgarischen gegeben, denn die konkurrierenden Kandidaten mit multipler W-Bewegung in die C-Domäne haben per Annahme ein ansonsten identisches Beschränkungsprofil, und so kann die durch PAR-VER präferierte Struktur als unmarkierter Fall erscheinen. In der Regel sind die Umstände jedoch nicht so günstig; d.h., PAR-VER kann problemlos verletzt werden. Dies gilt z.B. schon bei einfacher W-Bewegung eines Objekts über ein Subjekt im Bulgarischen. Ein Beispiel ist in (49) angeführt.

(49) [_{CP} Kakvo₂ pravi Ivan₁ t₂] ?
 was macht Ivan

Es ist klar, daß gemäß PAR-VER hier das Objekt eigentlich nicht über das Subjekt bewegt werden dürfte (wenn überhaupt, so sollte sich nur das Subjekt bewegen können). Daß Objektbewegung in diesem Kontext dennoch optimal ist, liegt daran, daß nur auf diese Weise die übergeordnete Beschränkung W-KRIT erfüllt werden kann; der unmarkierte Fall kann somit nicht eintreten. Den Wettbewerb illustriert Tabelle T_{2.18}.

Zunächst einmal könnte man vielleicht denken, daß PAR-VER nur in solchen Sprachen W-Bewegung aktiv restringiert, deren sonstige Beschränkungsordnung mehrfache Voranstellung von W-Phrasen erzwingt. Dem ist jedoch nicht so: PAR-

³²Rudin (1988) beobachtet, daß Sprachen wie Polnisch, Tschechisch und Russisch zwar mehrfache W-Voranstellung kennen, aber nicht die im Bulgarischen vorliegende feste Reihenfolge der bewegten W-Phrasen aufweisen. Sie argumentiert jedoch, daß in diesen Sprachen die mehrfache W-Voranstellung gar keine echte W-Bewegung in die C-Domäne ist, sondern so etwas wie obligatorisches W-Scrambling an die IP. Was immer diesen Prozeß auslöst, die soeben gemachte Generalisierung ist dadurch nicht gefährdet.

T_{2.18}: Einfache W-Bewegung im Bulgarischen

Kandidaten	W-KRIT	PUR-EP	ÖKON	PAR-VER
☞ K ₁ : [_{CP} kakvo ₂ pravi Ivan ₁ t ₂]			*	*
K ₂ : [_{CP} – Ivan ₁ pravi kakvo ₂]	*!			

VER wird immer dann für W-Bewegung relevant, wenn das System der sonstigen Beschränkungen an sich Ambiguität bei der Regelanwendung zeigt. Gerade dieser Fall liegt vor bei den Superioritätseffekten in Mehrfachfragen im Englischen. Typische Beispiele, die den Superioritätseffekt des Englischen veranschaulichen, sind die in (50).

- (50) a. I wonder [_{CP} who₁ C [_{IP} t₁ bought what₂]]
 b. *I wonder [_{CP} what₂ C [_{IP} who₁ bought t₂]]

Die ursprünglich für Fälle wie diese von Chomsky (1973) vorgebrachte Erklärung beruhte auf einer speziellen Superioritätsbedingung, derzufolge bei an sich bestehender Ambiguität in der Anwendung der W-Bewegungsregel immer die höhere (bzw., bei mehr als zwei W-Phrasen, die höchste) W-Phrase W-bewegt werden muß, und die andere(n) W-Phrase(n) in situ verharren muß bzw. müssen. PAR-VER hat für die betrachteten Fälle dieselben Konsequenzen: In (50-a) c-kommandiert das Subjekt das Objekt in Übereinstimmung mit PAR-VER auf D- wie S-Struktur; in (50-b) dagegen werden die D-strukturellen C-Kommando-Verhältnisse zur S-Struktur hin unter Verletzung von PAR-VER umgekehrt. Nehmen wir an, daß PAR-VER im Englischen wie im Bulgarischen relativ tief geordnet ist (vgl. (51)), dann folgen Superioritätseffekte wie in Tabelle T_{2.19} gezeigt.

- (51) *Ordnung im Englischen* (erweitert):
 PUR-EP ≫ W-KRIT ≫ ÖKON ≫ PAR-VER

T_{2.19}: Mehrfachfragen im Englischen

Kandidaten	PUR-EP	W-KRIT	ÖKON	PAR-VER
☞ K ₁ : [_{CP} who ₁ C t ₁ ... what ₂]		*	*	
K ₂ : [_{CP} – C who ₁ ... what ₂]		*!*		
K ₃ : [_{CP} who ₁ what ₂ C t ₁ ... t ₂]	*!		**	
K ₄ : [_{CP} what ₂ C who ₁ ... t ₂]		*	*	*!
K ₄ : [_{CP} what ₂ who ₁ C t ₁ ... t ₂]	*!		**	*

Wiederum ist durch die tiefe Einordnung von PAR-VER sichergestellt, daß einfache W-Bewegung eines W-Objekts über ein [-w] markiertes Subjekt hinweg unproblematisch ist. Ein einschlägiges Beispiel ist (52); der in Tabelle T_{2.20} dokumentierte Wettbewerb ist in allen relevanten Punkten zu dem in Tabelle T_{2.18} für das Bulgarische dargestellten identisch.

(52) I wonder [_{CP} what₂ C [_{IP} John₁ bought t₂]]

$T_{2,20}$: Einfache W-Bewegung im Englischen

Kandidaten	PUR-EP	W-KRIT	ÖKON	PAR-VER
☞ K ₁ : [_{CP} what ₂ C John ₁ ... t ₂]			*	*
K ₂ : [_{CP} - C John ₁ ... what ₂]		*!		

Durch die Annahme einer tief geordneten Beschränkung PAR-VER läßt sich somit ableiten, daß S-strukturelle W-Bewegung im unmarkierten Fall, d.h., bei ansonsten bestehender Ambiguität in der Regelanwendung, die D-strukturelle Abfolge der NP-Argumente beibehält, und dies im Bulgarischen wie im Englischen. Demgegenüber hat sich eine einheitliche Analyse der Reihenfolgeeffekte in Mehrfachfragen im Bulgarischen und Englischen in einer Standardsyntax, die ohne verletzbare Beschränkungen arbeitet, als schwierig erwiesen; die Superioritätsbedingung (ebenso wie die im wesentlichen äquivalente Minimal Link Condition aus Chomsky (1995)) jedenfalls kann zwar die Daten im Englischen erklären, ist aber nicht ohne weiteres auf die bulgarischen Daten anwendbar.³³

Abschließend sei noch vermerkt, daß PAR-VER keineswegs nur für W-Bewegung den unmarkierten Fall vorhersagt, sondern durch die allgemeine Formulierung natürlich im Prinzip jeden Bewegungstyp beschränkt. Tatsächlich werden in Müller (1999a) noch einige weitere Arten von Bewegung betrachtet und als für PAR-VER sensitiv zu erweisen versucht (z.B. mehrfache Wackernagelbewegung im Deutschen).

5. Beschränkungskonflikt

Angenommen, durch die empirische Evidenz sind unabhängig voneinander zwei Beschränkungen B₁ und B₂ gut motiviert; es kann aber in bestimmten Fällen zum

³³Man fragt sich hier vielleicht, wieso das Deutsche unter der bisher eingeführten, dem Englischen bis auf die Kopplung von ÖKON und VOLL-INT in diesem Bereich doch relativ ähnlichen Beschränkungsordnung anders als das Englische keinen typischen Superioritätseffekt aufweist (vgl. Haider (1983; 1993), neben vielen anderen):

- (i) a. Ich weiß nicht [_{CP} wer₁ C t₁ was₂ gekauft hat]
 b. Ich weiß nicht [_{CP} was₂ C wer₁ t₂ gekauft hat]

Zunächst einmal würden wir erwarten, daß (i-b) mit (i-a) im selben Wettbewerb ist und eine gegenüber (i-a) fatale Verletzung von PAR-VER aufweist. Es muß also noch einen anderen, intervenierenden Faktor geben, der (i-b) ein optimales Beschränkungsprofil zu haben gestattet. Ich möchte an dieser Stelle nicht weiter auf dieses Thema eingehen; es sei hier lediglich vermerkt, daß eine Lösung des vorliegenden Problems vielleicht darauf Bezug nehmen könnte, daß die Wortstellung im Deutschen IP-intern sehr viel freier ist als im Englischen; vgl. Fanselow (1997). Siehe auch Kapitel 6.

Konflikt von B_1 und B_2 kommen, wobei die eine Beschränkung selektiv verletzbar erscheint. In Standardtheorien gibt es angesichts einer solchen Situation nur die Möglichkeit, auf die eine Beschränkung entweder ganz zu verzichten, oder aber sie durch eine Ausnahmeklausel (oder durch Verlagerung ihrer Anwendung auf eine abstraktere Repräsentationsebene wie LF) in ihrer allgemeinen Gültigkeit einzuschränken, so daß ein Beschränkungskonflikt vermieden wird. In einem optimalitätstheoretischen Ansatz können B_1 und B_2 in einer solchen Situation beide in einer maximal generellen Form aufrechterhalten werden; die Konfliktresolution erfolgt durch Beschränkungsordnung. Drei Beispiele (deutsche Wortstellung, skandinavische Objektbewegung, Ökonomieprinzipien im minimalistischen Programm) sollen dies illustrieren.

5.1. Deutsche Wortstellung

Es ist häufig beobachtet worden, daß die Faktoren, die die Wortstellung im deutschen Mittelfeld determinieren, in vielen Fällen Unvereinbares verlangen. Diese Faktoren können wie in (53) als konfligierende Beschränkungen kodiert werden.³⁴

- (53) a. DEF (“Definitheitsbedingung”):
Eine definite NP geht einer indefiniten NP voran.
b. BEL (“Belebtheitsbedingung”):
Eine belebte NP geht einer unbelebten NP voran.
c. FOK (“Fokusbedingung”):
Ein nicht-fokussiertes Element geht einem fokussierten Element voran.
d. O-KASUS (“Objektkasusbedingung”):
Eine Dativ-NP geht einer Akkusativ-NP voran.

Es läßt sich für all diese Beschränkungen bestätigende Evidenz anführen. Nichtsdestoweniger ist klar, daß die Beschränkungen oft miteinander im Konflikt stehen. Wenn z.B. im Mittelfeld eine unbelebte Dativ-NP und eine belebte Akkusativ-NP zusammen auftreten, ergibt sich ein Beschränkungskonflikt zwischen BEL und O-KASUS; wenn eine definite Akkusativ-NP und eine indefinite Dativ-NP vorliegen, ergibt sich ein Konflikt zwischen DEF und O-KASUS; usw. Man braucht hier, wenn man auf keine der Beschränkungen in (53) verzichten möchte, eine Gewichtung der Beschränkungen, die den Konflikt auflöst und somit eine Abfolge der anderen vorzuziehen gestattet. Um nur ein Beispiel zu geben (und somit der ausführlichen Analyse in Kapitel 6 vorzugreifen): Es sieht so aus, als sei BEL O-KASUS übergeordnet. In (54) ist die optimale Abfolge Akkusativ-NP vor Dativ-NP.

³⁴Es sei darauf hingewiesen, daß die folgenden Bemerkungen nur sehr vorläufig sind und lediglich dem Zweck dienen, das Vorhandensein von Beschränkungskonflikten bei der deutschen Wortstellung plausibel zu machen. Die Thematik wird etwas genauer in Kapitel 4 behandelt werden, und sie steht dann im Mittelpunkt von Kapitel 6.

- (54) a. daß der Fritz die Kinder diesem Einfluß entzogen hat
 b. ?daß der Fritz diesem Einfluß die Kinder entzogen hat

Der Wettbewerb ist in Tabelle T_{2.21} skizziert.

T_{2.21}: *Belebtheit vs. Kasus bei der Wortstellung im deutschen Mittelfeld*

Kandidaten	BEL	O-KASUS
☞K ₁ : ... die Kinder diesem Einfluß V ...		*
K ₂ : ... diesem Einfluß die Kinder V ...	*!	

Man sieht hier, daß K₁ (= (54-a)) zwar die Beschränkung O-KASUS verletzt (weil die Akkusativ-NP der Dativ-NP vorangeht), aber dafür die höher geordnete Beschränkung BEL erfüllt und somit optimal wird. K₂ (= (54-b)) verletzt demgegenüber fatal BEL, bei Erfüllung von O-KASUS. Die Vorhersage ist nun die, daß dann, wenn ein Einfluß der höher geordneten Beschränkung BEL (sowie anderer höher geordneter Beschränkungen) in (53) ausgeschlossen wird (weil die konkurrierenden Kandidaten sich dort gleich verhalten), die tiefer geordnete Beschränkung O-KASUS entscheidend wird und so die optimale Abfolge vorhersagt. Dies ist, wie (55) zeigt, der Fall – die optimale Abfolge ist hier Dativ-NP vor Akkusativ-NP.

- (55) a. daß die Maria dem Fritz die Kinder entzogen hat
 b. ?daß die Maria die Kinder dem Fritz entzogen hat

Den Wettbewerb veranschaulicht Tabelle T_{2.22}.

T_{2.22}: *Der Einfluß von Kasus auf die Wortstellung im deutschen Mittelfeld*

Kandidaten	BEL	O-KASUS
K ₁ : ... die Kinder dem Fritz V ...		*!
☞K ₂ : ... dem Fritz die Kinder V ...		

Dies ist wiederum eine Instanz dessen, was im letzten Abschnitt behandelt wurde: Im unmarkierten Fall, d.h., wenn alle höher geordneten Beschränkungen nicht zwischen den Kandidaten unterscheiden, ist im Deutschen die Abfolge Dativ-NP vor Akkusativ-NP. Mit der Optimalitätstheorie hat man also ein geeignetes Mittel, um sowohl das offensichtliche Auftreten von Beschränkungskonflikten bei der variablen Abfolge im deutschen Mittelfeld zu erfassen, als auch korrekt vorherzusagen, daß im unmarkierten Fall die Abfolge Dativ-NP vor Akkusativ-NP ist. Allerdings sei nicht verschwiegen, daß sich in diesem Bereich auch ein schwerwiegendes Problem zeigt: (54-a) und (55-a) sind tatsächlich vollkommen wohlgeformt, aber (54-b) und (55-b) scheinen nicht so sehr ungrammatisch, als vielmehr “weniger normal” (oder “weniger grammatisch”) zu sein. Wie bereits erwähnt wurde, erlaubt die Optimalitätstheorie jedoch nicht die systematische Erfassung des Konzepts der Grammatikalitätsgrade; vgl. die Definition von Grammatikalität in (4) in Kapitel 1. Versuche, dieses Problem zu lösen, werden in Kapitel 6 besprochen.

5.2. Objektbewegung und Quantorenskopus im Isländischen

Ein zweites Beispiel dafür, daß die empirische Evidenz stark das Vorliegen eines Beschränkungskonfliktes nahelegt, betrifft die Interaktion von Quantorenskopus und Objektbewegung (sog. “Object Shift”) im Isländischen. Die empirische Beobachtung geht auf Diesing (1997) zurück; die im folgenden skizzierte optimalitätstheoretische Umsetzung ist in Vikner (1997b) geleistet worden.

Es gibt zwei Generalisierungen, die an und für sich für die Syntax des Isländischen beide gut bestätigt sind. Die erste Generalisierung lautet, daß der relative Skopus von quantifizierten Ausdrücken, z.B. von einer indefiniten NP und einem temporalen Adverbial wie *immer*, mit der syntaktischen Oberflächenabfolge Hand in Hand geht: Wenn ein quantifizierter Ausdruck Q_1 Skopus über einen anderen quantifizierten Ausdruck Q_2 hat, dann c-kommandiert Q_1 Q_2 auch asymmetrisch auf der S-Struktur. Diese Generalisierung läßt sich durch eine Beschränkung SKOPUS (“Scoping Condition”) erfassen.³⁵

Die zweite Generalisierung besagt, daß Objektbewegung einer NP über ein an VP adjungiertes Adverbial im Isländischen nur dann möglich ist, wenn das Hauptverb die VP verlassen hat (also im Isländischen nach I oder C bewegt worden ist); in diesem Kontext ist die Bewegung (bei nicht-pronominalen NPs) optional. Diese Generalisierung läßt sich für die gegenwärtigen Zwecke als OBJEKT-LIZ (“Objekt-Lizensierung”) festhalten.

Oft gibt es keinerlei Konflikt zwischen den beiden Beschränkungen. Wo Objektvoranstellung lizensiert ist, weil das finite Hauptverb die VP verlassen hat und nach I bzw. C gewandert ist, entspricht der relative Skopus von indefiniter NP und temporalem Adverbial stets der Oberflächenabfolge. So hat in (56-a) obligatorisch das Adverbial weiten Skopus über die indefinite NP (“Wenn Auslandsnachrichten laufen, gibt es immer ein Interview mit Clinton”), und in (56-b) liegt der umgekehrte Fall vor (“Wenn es ein Interview mit Clinton gibt, dann läuft das immer in den Auslandsnachrichten”).³⁶

- (56) a. Auk þess sýna₂ þau [_{Adv} alltaf] t₂ [_{NP₁} viðtal við Clinton]
 übrigens zeigen sie immer Interview mit Clinton

³⁵Diese Beschränkung ist möglicherweise als Instanz von PAR-VER rekonstruierbar, zumal SKOPUS dem entspricht, was Lakoff (1971) im Rahmen der Generativen Semantik vorgeschlagen hat (s.o.). Hierzu müßten zum einen bestimmte theoretische Voraussetzungen gegeben sein (z.B. die, daß quantifizierte Ausdrücke obligatorisch LF-Anhebung unterliegen, und daß relativer Skopus auf LF abgelesen wird), von denen Diesing und Vikner nicht notwendigerweise ausgehen; und zum anderen wäre PAR-VER so zu modifizieren, daß nicht nur Argumente, sondern – im Fall von quantifizierten Ausdrücken – auch Adverbiale darunter fallen. Für die gegenwärtigen Zwecke ist aber die Annahme einer Beschränkung wie SKOPUS ausreichend, und wir können von der Frage absehen, ob SKOPUS zugunsten einer noch allgemeineren Beschränkung aufgegeben werden kann.

³⁶Für manche Sprecher ist allerdings (56-a) offenbar auch ambig, was dann Zusatzannahmen erfordert.

- í erlendu fréttunum
in den Auslandsnachrichten
- b. Auk þess sýna₂ þau [_{NP₁} viðtal við Clinton] [_{Adv} alltaf] t₂ t₁
übrigens zeigen sie Interview mit Clinton immer
í erlendu fréttunum
in den Auslandsnachrichten

Dieses Bild ändert sich, sobald der syntaktische Kontext so modifiziert wird, daß keine Bewegung des Hauptverbs aus der VP mehr möglich ist. Letzteres ist bei periphrastischen Tempora der Fall, also bei Präsenz eines finiten Auxiliars; hier muß das Hauptverb-Partizip in situ, in der VP bleiben. Interessanterweise muß genau dann, wenn die relativen Skopusverhältnisse nicht durch Objektbewegung transparent gemacht werden können, eine solche Desambiguierung durch syntaktische Bewegung auch nicht erfolgen, und der relative Skopus muß dann nicht der Oberflächenabfolge entsprechen. Dies zeigen die Daten in (57). In (57-b) ist Objektbewegung vor das Adverbial trotz fehlender Hauptverbbewegung aus der VP erfolgt, und das Resultat ist Ungrammatikalität. In (57-a) dagegen ergibt sich bei unterbleibender Objektbewegung nicht nur Grammatikalität (wie in (56-a)), sondern tatsächlich Skopusambiguität.

- (57) a. Auk þess hafa þau [_{Adv} alltaf] sýnt [_{NP₁} viðtal við Clinton]
übrigens haben sie immer gezeigt Interview mit Clinton
í erlendu fréttunum
in den Auslandsnachrichten
- b. *Auk þess hafa þau [_{NP₁} viðtal við Clinton] [_{Adv} alltaf] sýnt t₁
übrigens haben sie Interview mit Clinton immer gezeigt
í erlendu fréttunum
in den Auslandsnachrichten

In dem Kontext, wo keine Hauptverbbewegung aus der VP heraus vorliegt, ergibt sich somit ein Widerspruch zwischen den beiden ansonsten im Isländischen gut bestätigten Beschränkungen SKOPUS und OBJEKT-LIZ. Wie der Unterschied in (57) zeigt, ist dieser Beschränkungskonflikt durch Beschränkungsordnung aufzulösen: OBJEKT-LIZ ist die wichtigere Beschränkung; die maximale Erfüllung dieser Beschränkung erlaubt es, die untergeordnete Beschränkung SKOPUS minimal zu verletzen.³⁷

³⁷Vorausgesetzt ist hier, daß semantische Information in die Definition der Kandidatenmenge eingeht; im selben Wettbewerb sind nur Kandidaten mit identischer LF (bzw. identischer Bedeutung). Das heißt, die S-Struktur in (57-a) ist Gewinner zweier verschiedener Wettbewerbe, je nachdem, ob Skopus der NP über das Adverbial oder Skopus des Adverbials über die NP vorliegt; aber während in letzterem Wettbewerb keine SKOPUS-Verletzung durch die optimale Form erfolgt, muß im ersteren Wettbewerb der optimale Kandidat SKOPUS verletzen.

5.3. Ökonomieprinzipien

Als drittes und letztes Beispiel für das Vorliegen von Konflikten unter unabhängig motivierten Beschränkungen seien einige Ökonomieprinzipien angeführt, wie sie im minimalistischen Programm vorgeschlagen worden sind (vgl. Chomsky (1991; 1993; 1995)). Neben der bereits bekannten Beschränkung ÖKON sind dies insbesondere das ZAUDERPRINZIP (“Procrastinate”), das insofern ein Spezialfall von ÖKON ist, als es S-strukturelle Bewegung gegenüber LF-Bewegung noch einmal zusätzlich verbietet; die Beschränkung PF-KONVERGENZ, derzufolge starke Merkmale vor der S-Struktur (bzw. vor der Operation Spell-Out, die in Chomsky (1993; 1995) die traditionelle S-Struktur ersetzt) per Bewegung überprüft und somit aus der Derivation entfernt werden müssen; sowie die Beschränkung KONVERGENZ, derzufolge alle Merkmale vor Erreichen der LF-Repräsentation durch Bewegung überprüft und entfernt werden müssen.

- (58) ÖKON:
Bewegung ist verboten.
- (59) ZAUDERPRINZIP:
S-strukturelle Bewegung ist verboten.
- (60) KONVERGENZ:
Merkmale müssen durch Bewegung überprüft werden.
- (61) PF-KONVERGENZ:
Starke Merkmale müssen vor der S-Struktur durch Bewegung überprüft werden.

Diese vier Beschränkungen sind recht allgemein formuliert, und daher stehen sie zum Teil miteinander im Konflikt – ÖKON und ZAUDERPRINZIP verbieten Bewegung, KONVERGENZ und PF-KONVERGENZ erzwingen Bewegung. In einer Standardsyntax ohne verletzbare und geordnete Beschränkungen gibt es angesichts dieser Situation nur die Möglichkeit, die Beschränkungen auf andere Weise zu formulieren, so daß kein direkter Beschränkungskonflikt entsteht; tatsächlich wird genau das im minimalistischen Programm auch getan. Der Preis, der hier zu zahlen ist, ist aber, daß die Definitionen dann um einiges komplexer werden und in die jeweilige Formulierung zum Teil explizit Ausnahmeklauseln eingebaut werden müssen. Konkret schlägt Chomsky (1991) vor, ÖKON als komplexe transderivationelle Beschränkung zu fassen, derzufolge von zwei im Wettbewerb befindlichen Derivationen immer die mit der geringeren Anzahl von Bewegungsoperationen zu wählen ist (das *Prinzip der wenigsten Schritte* (“Fewest Steps”), auf das in Abschnitt 4.1 von Kapitel 4 noch eingegangen werden wird).³⁸ In Chomsky (1993; 1995) findet

³⁸Eine transderivationelle Beschränkung ist nicht wie andere, lokal überprüfbare Beschränkungen erfüllt oder verletzt, sondern als Auswahlanweisung zu verstehen, derzufolge gemäß einem

sich demgegenüber ÖKON als mehr oder weniger impliziter Teil der Beschränkungen *Gier* (“Greed”) bzw. *Attraktion* (“Attract”), denenzufolge Bewegung nur dann erlaubt ist, wenn sie in Merkmalsüberprüfung endet. Das ZAUDERPRINZIP ist in der Standardformulierung ebenfalls keine lokale Beschränkung wie (59), sondern transderivationell. Es besagt bei Chomsky (1993) ungefähr folgendes: Wenn sich zwei ansonsten identische und konvergierende Derivationen dadurch unterscheiden, daß im einen Fall Bewegung auf der S-Struktur erfolgt und im anderen Fall auf LF, dann muß die zweite Derivation gewählt werden (vgl. Marantz (1995)). Daß ein Verständnis des ZAUDERPRINZIPS wie in (59) aber möglicherweise intuitiv plausibler und einfacher ist, das scheint mir die Redeweise von “forced (unforced) *violations of Procrastinate*” zu suggerieren (Hervorhebung von mir), die sich in Chomsky (1995) findet. Verletzungen des ZAUDERPRINZIPS kann es klarerweise nur geben, wenn diese Beschränkung verletzbar ist; eine transderivationelle Beschränkung kann aus Prinzip niemals verletzt werden.³⁹

Die Grundidee hinter ÖKON und ZAUDERPRINZIP im minimalistischen Programm ist jeweils, daß diese Beschränkungen immer dann greifen, wenn sich die Frage der Konvergenz nicht stellt. Dies bedeutet aber, daß das System der Chomskyschen Ökonomieprinzipien auf einfache Weise optimalitätstheoretisch rekonstruiert werden kann, und zwar als Dominanz der Ökonomiebeschränkungen durch die Konvergenzbeschränkungen. Konkret gelten bei Annahme der Formulierungen von ÖKON und ZAUDERPRINZIP in (58) und (59) im minimalistischen Programm die folgenden Ordnungen:⁴⁰

- (62) a. PF-KONVERGENZ \gg ZAUDERPRINZIP, ÖKON
 b. KONVERGENZ \gg ÖKON

Generell sieht es also so aus, als sei es bei Annahme der Verletzbarkeit von Beschränkungen oft möglich, diese einfacher zu formulieren, weil Beschränkungskonflikte nicht mehr umständlich durch Ausnahmeklauseln oder Relegation der Beschränkung auf eine andere Ebene umgangen werden müssen.⁴¹

bestimmten Maß ein Kandidat aus einer Kandidatenmenge ausgewählt wird.

³⁹Vgl. im übrigen Abschnitt 4.1 von Kapitel 4 zu genaueren Ausführungen, sowie zu einer Klärung des Zusammenhangs von transderivationellen Beschränkungen und dem Begriff der Optimalität.

⁴⁰Eine explizite Ordnung von drei sehr allgemeinen und daher konfligierenden Ökonomieprinzipien auf der Basis des minimalistischen Programms schlägt auch Sauerland (1995) vor.

⁴¹Mit “andere Ebene” ist hier zunächst einmal “andere Komplexitätsebene” gemeint, also der gerade besprochene Fall, daß eine lokal formulierbare Beschränkung transderivationell gefaßt werden muß. Verlagerung auf eine andere Ebene liegt aber auch dann vor, wenn eine Beschränkung lediglich auf einer abstrakten Ebene wie LF gelten kann. Diese Situation ergibt sich z.B. bei Chomskys (1986b) Prinzip VOLL-INT, das oben bereits mehrfach erwähnt worden ist. Bei Chomsky (1986b) wie bei Grimshaw (1997) ist die Intuition hinter VOLL-INT, daß Expletiva verboten sind, weil sie nicht zur Interpretation beitragen. Aber bei Chomsky (1986b) muß die Beschränkung VOLL-INT,

Damit ist die Skizzierung derjenigen Typen von empirischer Evidenz, die Argumente für eine optimalitätstheoretische Analyse liefern können, abgeschlossen. Natürlich sind die identifizierten fünf Konzepte (Parametrisierung, Reparatur, Wettbewerb und Blockade, unmarkierter Fall, Beschränkungskonflikt) im Einzelfall nicht immer klar voneinander zu trennen, und es gibt systematische Überlappungen. Eine saubere Taxonomie dieser eher informellen Kategorien wird hier auch nicht angestrebt. In den folgenden Kapiteln wird sich aber zeigen, daß viele optimalitätstheoretischen Analysen um Phänomene zentriert sind, die sich wenigstens einem der vorgestellten Konzepte subsumieren lassen.

Bevor ich mich nun der potentiellen Gegenevidenz gegen eine optimalitätstheoretische Syntax zuwende, sollen noch in einem Exkurs Klärungen von zwei Begriffen vorgenommen werden, die Beziehungen unter Beschränkungen betreffen und für das optimalitätstheoretische Arbeiten nützlich sind.

6. Beziehungen unter Beschränkungen

6.1. Pāninische Beschränkungsrelationen

Für jede Beschränkung B können wir feststellen, ob sie eine gegebene Kandidatenmenge KM unterteilt in solche Kandidaten, die B erfüllen, und solche Kandidaten, die B verletzen. Angenommen, eine Kandidatenmenge KM_1 hat ein W -Element im Input, eine andere Kandidatenmenge KM_2 nicht. Dann unterteilt W -KRIT KM_1 in eine Gruppe von Kandidaten, die die Beschränkung erfüllen (und W -Bewegung vornehmen), und eine andere Gruppe von Kandidaten, die die Beschränkung verletzen (und W -in-situ aufweisen). Die Kandidatenmenge KM_2 dagegen wird von W -KRIT nicht unterteilt, weil alle Kandidaten diese Beschränkung trivial erfüllen. Im ersteren Fall appliziert W -KRIT nicht leer beim Input von KM_1 ; im letzteren Fall appliziert W -KRIT leer beim Input von KM_2 .

Man betrachte nun die Beziehung der Beschränkungen W -KRIT und $ÖKON$. Wenn W -KRIT bei einem Input nicht leer appliziert, verlangen die beiden Beschränkungen Gegenteiliges. Dabei ist $ÖKON$ die allgemeinere Beschränkung. Daher gilt, daß W -KRIT zu $ÖKON$ als speziell zu generell in einer sog. Pāninischen Relation steht (“in honour of the first known investigator in the area”; Prince & Smolensky (1993, 81)):

(63) *Pāninische Beschränkungsrelation* (Prince & Smolensky (1993, 82)):

Es seien S und G zwei Beschränkungen. S steht zu G als speziell zu generell in einer Pāninischen Relation, wenn für jeden Input I , bei dem S nicht leer

weil unverletzbar, auf die abstrakte Repräsentationsebene LF beschränkt sein (wo sie dann in Chomskys (1986b; 1991) System den Prozeß der “Expletiversetzung” (“expletive replacement”) auslöst); erst die Annahme der Verletzbarkeit von $VOLL$ -INT bei Grimshaw (1997) macht es möglich, diese Beschränkung auch als S -strukturell gültig anzunehmen.

appliziert, gilt, daß jede S-Erfüllung in einem Output $O(I)$ eine G-Verletzung impliziert.

Es steht zwar W-KRIT zu ÖKON in einer Pāninischen Relation, aber nicht ÖKON zu W-KRIT. Der Grund ist, daß es natürlich Inputs geben kann, bei denen ÖKON nicht leer appliziert (weil auch andere Bewegungsoperationen, die durch andere Beschränkungen ausgelöst werden, durch ÖKON restringiert sind), ohne daß automatisch durch ÖKON-Erfüllung W-KRIT verletzt wird.

Hat man erst einmal herausgefunden, daß zwei Beschränkungen in einer Pāninischen Relation stehen, so ergibt sich eine interessante Konsequenz bzgl. ihrer Ordnung (vgl. Prince & Smolensky (1993, 82 & 221)). Angenommen, die generelle Beschränkung ist in der Beschränkungsordnung einer Sprache bei einem Input *aktiv* in dem Sinne, daß sie bei diesem Input Kandidaten aus dem Wettbewerb eliminiert (d.h., daß sie bei einem Input von mindestens einem Kandidaten fatal verletzt wird). Dann folgt, daß die spezielle Beschränkung über der generellen Beschränkung angeordnet sein muß, um selbst auch beim vorliegenden Input aktiv zu sein; ansonsten verrichtet die spezielle Beschränkung keine Arbeit. Dies ist Pānini's Theorem zur Beschränkungsordnung:⁴²

(64) *Pānini's Theorem zur Beschränkungsordnung* :

Es seien S und G zwei Beschränkungen einer Beschränkungsordnung, die als speziell zu generell in einer Pāninischen Relation stehen. Dann gilt: Wenn G bei einem Input I aktiv ist und $G \gg S$ gilt, dann ist S bei I nicht aktiv.

Es ergibt sich daraus eine praktische Konsequenz für das optimalitätstheoretische Arbeiten: Wenn eine spezielle Beschränkung, die in einer Pāninischen Relation zu einer generellen Beschränkung steht, in einer Grammatik Arbeit verrichten, also nicht irrelevant sein soll, dann muß sie die entgegengesetzte generelle Beschränkung dominieren. Für das konkrete Beispiel mit W-KRIT und ÖKON bedeutet dies: Wenn in einer Sprache $\text{ÖKON} \gg \text{W-KRIT}$ gilt, dann ist bei jedem Input, wo ÖKON aktiv ist, W-KRIT inaktiv; m.a.W.: W-KRIT hat hier keine Effekte. Gerade dies ist ja für Sprachen wie das Koreanische auch intendiert.

6.2. Stringenzrelationen

Unter der Voraussetzung, daß PUR-EP eine Beschränkung für durch Bewegung erzeugte Adjunktion ist,⁴³ ergibt sich, daß auch PUR-EP und ÖKON als speziell zu

⁴²Die Formulierung des Theorems in Prince & Smolensky (1993, 82) läßt versehentlich die entscheidende Klausel "wenn $G \gg S$ gilt" aus. Daß dies gemeint ist, wird jedoch zum einen aus dem Kontext klar; zum anderen ist der Fehler an einer späteren Stelle im Buchmanuskript (Prince & Smolensky (1993, 221)) korrigiert.

⁴³Bei Grimshaw (1997) gilt PUR-EP allerdings auch für basisgenerierte Adjunktion; siehe oben und Kapitel 3. Unter dieser Annahme gilt das Folgende nicht.

generell in einer Beziehung stehen. Dies kann jedoch nicht die Pāninische Relation sein. Der Grund ist, daß selbstverständlich nicht jede *Erfüllung* von PUR-EP eine Verletzung von ÖKON impliziert. Dies zeigen etwa die W-in-situ-Kandidaten K_4 aus T_{2.17}, K_2 aus T_{2.18}, K_2 aus T_{2.19}, sowie K_2 aus T_{2.20}. Es ist vielmehr so, daß offenbar jede *Verletzung* von PUR-EP auch eine Verletzung von ÖKON bewirkt (aber nicht umgekehrt, weil es ja Bewegungen gibt, die nicht Adjunktion an CP bzw. C^0 involvieren). Diese Beziehung nennt Baković (1995) nach einem Vorschlag von Alan Prince Stringenzrelation.

(65) *Stringenzrelation* (Baković (1995, 28)):

Es seien S und G zwei Beschränkungen. S steht zu G als speziell zu generell in einer Stringenzrelation, wenn für jeden Input I, bei dem S nicht leer appliziert, gilt, daß jede S-Verletzung in einem Output O(I) eine G-Verletzung impliziert.

Für zwei Beschränkungen, die in einer Stringenzrelation stehen, gilt nicht ein so einfaches Theorem wie (64). Eine spezielle Beschränkung kann im Prinzip in einer Grammatik Arbeit verrichten, wenn sie von einer generellen Beschränkung dominiert wird, zu der sie in einer Stringenzrelation steht. Die Frage, wo die spezielle Beschränkung in diesem Fall eingeordnet werden muß, um Arbeit zu verrichten, hängt vor allem von der Existenz und relativen Gewichtung von weiteren Beschränkungen ab, die mit den in einer Stringenzrelation stehenden Beschränkungen konfligieren.

Nach dieser Begriffsklärung können wir uns nun solcher Evidenz zuwenden, die für eine optimalitätstheoretische Analyse zumindest auf den ersten Blick problematisch ist. Die Gegenevidenz wird hier untergliedert in konzeptuelle und empirische Aspekte.

7. Konzeptuelle Gegenevidenz

7.1. Komplexität

Zur Ermittlung der Grammatikalität bzw. Ungrammatikalität eines Kandidaten ist in der Optimalitätstheorie ein globaler Vergleich, der Wettbewerb, notwendig. Somit ist es intuitiv plausibel, daß ein höheres Maß an Komplexität entsteht als bei anderen Theorien, die ohne globale Vergleiche auskommen und über die Wohlgeformtheit eines Kandidaten allein auf der Grundlage von Eigenschaften zu entscheiden erlauben, die dem Kandidaten inhärent sind. Diese größere Komplexität von OT-Grammatiken wird häufig als Nachteil betrachtet. Prince & Smolensky (1993) diskutieren diesen Einwand ausführlich. Der Kern ihrer Haltung dazu ist im folgenden Zitat zusammengefaßt (vgl. Prince & Smolensky (1993, 197)):

“Diese Bedenken sind in einem Mißverständnis darüber begründet, was Grammatiken sind. Es obliegt einer Grammatik nicht, komputationelle Operationen durchzuführen, worauf Chomsky über die Jahre wiederholt hingewiesen hat. Eine Grammatik ist vielmehr eine Funktion, die Sätzen Strukturbeschreibungen zuordnet; was formal relevant ist, ist, daß die Funktion wohldefiniert ist. Die Anforderungen der Erklärungsadäquatheit (an Grammatiktheorien) und der Beschreibungsadäquatheit (an Grammatiken) beschränken und evaluieren den Raum für Hypothesen. Grammatiktheoretikern steht es bei der Verfolgung dieser Ziele frei, ein beliebiges formales Mittel zu ersinnen; tatsächlich *müssen* sie sich gestatten, hier ohne Restriktionen vorzugehen, wenn gute Theorien entdeckt werden sollen. Ebenso steht es einem nicht frei, willkürliche zusätzliche Metabeschränkungen (wie z.B. “komputationelle Plausibilität”) einzuführen, die mit den wohldefinierten grundsätzlichen Zielen des gesamten Unternehmens im Konflikt stehen könnten. In der Praxis haben sich Computer-Linguisten immer als findig erwiesen. Sämtliche existierenden Komplexitätsresultate für bekannte Theorien sind frappierend weit entfernt von der Kapazität des menschlichen Prozessors ... doch nichtsdestoweniger sind alle Arten von Grammatiktheorien erfolgreich in Parsern implementiert worden; die Implementierung ist in unterschiedlich starkem Maße erfolgt, aber die Effizienz war vergleichbar. ... Es gibt weder prinzipielle noch praktische Gründe für die Annahme, daß Überlegungen bzgl. Komplexität, die direkt auf Grammatikformalismen angewandt werden, informativ sind.”

In eine ähnliche Richtung zielen auch die Untersuchung von Frank & Satta (1998) und ihre Weiterentwicklung durch Wartena (2000), wo gezeigt wird, daß eine optimalitätstheoretische Grammatik unter bestimmten Voraussetzungen (z.B. der, daß die geordneten Beschränkungen nicht selbst transderivationell sind), so formalisierbar ist, daß wie bei Standardgrammatiken keine besonderen Komplexitätsprobleme auftreten, und daß sie somit auch ohne weiteres implementierbar ist.

Nichtsdestoweniger bleibt die Komplexität vermutlich der meistgeäußerte Vorbehalt gegen eine optimalitätstheoretische Organisation der Syntax (oder der Grammatik insgesamt). Intuitiv, so die Annahme, ist es nicht einleuchtend, eine komplexere Syntax für einen gegebenen Phänomenbereich zu entwickeln, wenn es eine weniger komplexe Standard-Syntax ebenso gut tut. Möglicherweise ist diese Annahme an sich nicht vollkommen unbegründet; sie setzt allerdings voraus, daß man plausible und einfache Analyseideen für jeden empirischen Phänomenbereich in der Standard-Syntax ebenso gut umsetzen kann wie in der optimalitätstheoretischen Syntax. Gerade dies ist jedoch, wie ich oben zu zeigen versucht habe, und wie sich in den folgenden Kapiteln noch weiter manifestieren wird, nicht der Fall. Die Konklusion aus alledem sollte vielleicht sein, daß man Komplexitätsfragen ernst, aber nicht zu ernst nehmen muß: Eine optimalitätstheoretische Syntax muß sich rechtfertigen dadurch, daß sie bei der Analyse von empirischen Phänomenen größere Erklärungsadäquatheit erreicht als Standard-Syntaxen; kann sie das nicht, gibt es zunächst einmal keinen guten Grund, nicht eine weniger komplexe Syntaxtheorie anzunehmen.

7.2. Natur der Kandidatenmenge

Die Kandidatenmenge wird, wie oben schon erwähnt, leicht sehr groß und unüberschaubar. Darüber hinaus sieht es so aus, als sei ein einheitliches Konzept, das für alle verschiedenen Anwendungen gleichermaßen gut geeignet ist, schwer zu formulieren. (Zum Beispiel legen manche empirischen Phänomene nahe, daß nur Kandidaten mit derselben Bedeutung konkurrieren können, andere Daten suggerieren, daß trotz unterschiedlicher Bedeutung Wettbewerb vorliegt.) Die Frage der Größe von Kandidatenmengen wird bei der Diskussion von Pesetskys (1997; 1998) Theorie in Abschnitt 2 von Kapitel 3 eine Rolle spielen; das Problem widerstreitender Evidenz für Kandidatenmengen taucht in den Kapiteln 4 und 8 auf.

7.3. Natur der Beschränkungen

Innertheoretische Probleme können oftmals nur allzu leicht durch Hinzufügen hoch geordneter Ad-hoc-Beschränkungen gelöst werden, die zum Teil auch noch erheblich überlappen mit anderen, umfassenderen Beschränkungen. Das Problem scheint hier die potentiell fehlende Restriktivität: Was ist eine mögliche syntaktische Beschränkung in der Optimalitätstheorie? Eine naheliegende Antwort hierauf ist, daß andere Syntaxtheorien in der Regel auch keine prinzipiellen Restriktionen für mögliche Beschränkungen formulieren, abgesehen von solchen, die durch Grundsatzentscheidungen zwangsläufig sind (wenn etwa eine Syntaxtheorie strikt repräsentationell ist, können Beschränkungen nicht derivationell sein). Das Hauptkriterium für eine "gute" syntaktische Beschränkung ist hier wie dort Eleganz (dies beinhaltet u.a. möglichst große Einfachheit in der Formulierung und möglichst große Allgemeinheit in der Anwendung).⁴⁴

Es scheint daher durchaus vertretbar, bei Beachtung der Kriterien der Kompatibilität mit Grundvorgaben der Theorie (wie erwähnt können optimalitätstheoretische Syntaxen repräsentationell, derivationell, als eine Mischform mit D-Struktur, S-Struktur und LF wie in Chomsky (1981), oder noch ganz anders organisiert sein) und der Analyse-internen Eleganz bei der Formulierung von Beschränkungen die Frage von weiteren externen Restriktionen für optimalitätstheoretische Beschrän-

⁴⁴Chomskys (1995) minimalistisches Programm ist hier auf den ersten Blick eine Ausnahme, weil im Prinzip nur noch wenige Beschränkungstypen zugelassen sind – solche, die "konzeptuell notwendig" sind. Aber abgesehen davon, daß "konzeptuelle Notwendigkeit" ein vager Begriff ist, zeigt die Praxis minimalistischer Analyse, daß die Hauptrestriktionen für Beschränkungen zwar durch die Grundvorgaben der Theorie gegeben sind (Beschränkungen müssen aufgrund des Fehlens von S-Struktur und D-Struktur entweder auf den Schnittstellenebenen LF und PF applizieren, oder sie sind strikt derivationell), daß aber ansonsten wie vorher auch das entscheidende Kriterium das der Eleganz ist. Man vergleiche hierzu etwa die umfassende Studie von Epstein, Groat, Kawashima & Kitahara (1998), die eine minimalistische Analyse einer Reihe von verschiedenen Phänomenen vornimmt und dabei auch keineswegs nur auf "konzeptuell notwendige" Beschränkungen zurückgreift.

kungen hintan zu stellen. Ich werde aber in Kapitel 8 auf diesen Punkt zurückkommen.⁴⁵

7.4. Parametrisierung

Wir haben oben gesehen, daß parametrische Variation zunächst einmal auf sehr einfache Weise in der optimalitätstheoretischen Syntax erfaßt werden kann. Das Konzept der Parametrisierung als Beschränkungsumordnung bringt jedoch auch zwei konzeptuelle Probleme mit sich. Das erste Problem hängt mit dem oben bereits erwähnten Begriff der faktoriellen Typologie zusammen.

7.4.1. Parametrisierung und faktorielle Typologie

Entspricht jede Umordnung von Beschränkungen einer möglichen einzelsprachlichen Grammatik, wie es das Postulat der faktoriellen Typologie vorsieht? Dies war so bei dem Fragment, das in Abschnitt 1 eingeführt wurde und nur auf drei geordneten Beschränkungen beruhte (PUR-EP, ÖKON, W-KRIT). Werden jedoch größere Fragmente betrachtet, oder stellt man sich sogar eine gesamte optimalitätstheoretische Syntax einer Sprache vor, so scheint es unwahrscheinlich, daß jede Ordnung einer möglichen Grammatik entspricht. Gibt es dann Prinzipien, die festlegen, was eine mögliche und was eine unmögliche Beschränkungsumordnung (resultierend in parametrischer Variation) ist? Feste Ordnungen sind oben bereits angenommen worden für die Mitglieder von Subhierarchien von Beschränkungen. Ebenfalls in festen Ordnungen resultieren zwei Mittel der Erzeugung von Beschränkungen, die wir noch kennenlernen werden: lokale Konjunktion (Kapitel 3) und harmonische Ausrichtung (Kapitel 8). Vgl. auch Smolensky (1995) zu diesen drei Konzepten.

7.4.2. Parametrisierung und morphologische Fundierung

Der V-I-Bewegungsparameter Wenn Parametrisierung nur auf Beschränkungsumordnung basieren kann, dann kann es nicht in offensichtlicher Weise eine morpho-

⁴⁵Ein Einwand, der von Vertretern funktionaler Ansätze immer wieder gegen im Rahmen der generativen Grammatik vorgeschlagene Beschränkungen vorgebracht worden ist, beruht darauf, daß es keinen unabhängigen Grund gebe, warum die jeweilige Beschränkung gelten sollte. Die Antwort der generativen Grammatiker ist hier normalerweise, daß dies eben Ausdruck der Tatsache ist, daß das grammatische Wissen des Menschen nicht von der allgemeinen kognitiven Fähigkeit gesteuert, sondern modulspezifisch in separaten Hirndomänen verankert ist; mit anderen Worten: Die Beschränkungen sind so, wie sie sind, weil sie angeboren sind. Diese Auseinandersetzung läßt sich im Rahmen der Optimalitätstheorie auf genau dieselbe Weise führen. So meint Haspelmath (2000), daß optimalitätstheoretische Beschränkungen funktional, durch unabhängig gegebene Restriktionen für Sprecher und Hörer, gerechtfertigt werden müssen, und daß dies bei einer Reihe von wichtigen in der Literatur vorgeschlagenen Beschränkungen auch möglich ist. Als generativer Grammatiker hat man folglich zu zeigen, daß diese Position nicht haltbar ist; vgl. Müller (2000b). Eine Darstellung dieser Auseinandersetzung würde hier jedoch zu weit führen.

logische Fundierung der Parameterfixierung geben. Letzteres ist jedoch oft in der Literatur behauptet worden. Als erstes Beispiel sei der sog. “V-I-Bewegungsparameter” betrachtet, demzufolge nur Sprachen mit einem hinreichend elaborierten Konjugationsparadigma S-strukturelle Bewegung des finiten Voll-Verbs nach I erlauben. Wie Emonds (1978), Pollock (1989) und Chomsky (1991) festgestellt haben, läßt sich bei Annahme einer fixen Position (z.B. einer VP-Adjunktionsposition) für Adverbiale wie *oft* der in (66) ersichtliche Unterschied in der Wortstellung zwischen dem Französischen und dem Englischen darauf zurückführen, daß das Französische obligatorische S-strukturelle Bewegung des finiten Vollverbs über das Adverbial hinweg nach I aufweist, während diese Bewegung im Englischen auf der Oberfläche nicht stattfinden kann.

- (66) a. Jean embrasse₁ souvent [VP t₁ Marie]
 b. *Jean souvent [VP embrasse₁ Marie]
 c. *John kisses₁ often [VP t₁ Mary]
 d. John often [VP kisses₁ Mary]

Dieser V-I-Bewegungsparameter wurde häufig damit korreliert, daß die Verbalflexion im Französischen etwas reichhaltiger ist als im Englischen, d.h., er wurde an ein (wenn auch relativ abstraktes) Konzept der morphologischen Stärke gekoppelt. Diese Sichtweise ist überprüft worden durch Detailuntersuchungen an romanischen wie germanischen Sprachen und Dialekten. Insbesondere bei letzteren hat sich der erwartete Kontrast zwischen solchen Sprachen bzw. Varietäten ergeben, die reichhaltige Verbalflexion aufweisen, und solchen, die das nicht tun (z.B. verhalten sich Dänisch und Schwedisch wie Englisch, Isländisch und Jiddisch dagegen wie Französisch); vgl. etwa Roberts (1993), Rohrbacher (1994; 1999) und Vikner (1995), sowie die dort angegebene Literatur. Wie genau “morphologische Reichhaltigkeit” für die Zwecke des V-I-Bewegungsparameters zu definieren ist, ist dabei unter den Experten aber genauso sehr umstritten geblieben, wie die Idee der morphologischen Fundierung an sich weithin akzeptiert worden ist.

Das Problem für eine optimalitätstheoretische Analyse ist nun, daß zunächst einmal der Unterschied zwischen Französisch und Englisch in (66) auf syntaktische Beschränkungsumordnung zurückzuführen sein sollte und somit für morphologische Fundierung kein Platz zu sein scheint. Tatsächlich schlägt Vikner (1999) vor, daß der V-I-Bewegungsparameter aus der relativen Ordnung der Beschränkungen in (67) resultiert.

- (67) a. LEX-ÖKON (“Bewegungsökonomie für lexikalische Köpfe”, “No-Lex-Mvt”):
 Bewegung von lexikalischen Köpfen ist verboten (X^0 -Spur_{lex} ist nicht erlaubt).
 b. STR-BD (“Strikte Bindung”, “Proper Binding Condition”, “Pr-Bd”):
 Spuren müssen auf der S-Struktur gebunden sein.

LEX-ÖKON haben wir schon in Kapitel 1 bei der Darstellung von Grimshaws (1997) Analyse von *do*-Einsetzung kennengelernt; diese Beschränkung bestraft S-strukturelle Bewegung des Vollverbs nach I. Den umgekehrten Effekt hat STR-BD: Diese Beschränkung verbietet Absenkung der per Annahme in I erzeugten Flexionsmorphologie nach V ("Affix Hopping" nach Chomsky (1957), "Regel R" nach Chomsky (1981)). Vorausgesetzt sei nun, daß V und I auf die eine oder andere Weise in der Syntax zusammenkommen müssen. Dann ist auch ohne Illustration durch Tabellen klar, daß unter der Ordnung in (68-a) Sprachen des französischen Typs resultieren, in denen V-I-Bewegung stattfinden muß, während unter der Ordnung in (68-b) Sprachen des englischen Typs erscheinen, in denen V-I-Bewegung auf der S-Struktur verboten ist und stattdessen I abgesenkt wird.

- (68) a. Französisch: STR-BD \gg LEX-ÖKON
 b. Englisch: LEX-ÖKON \gg STR-BD

Eine morphologische Fundierung des Unterschieds in (66) ist bei dieser Analyse nicht mehr gegeben. Es bieten sich die folgenden Konklusionen an. Zum einen könnte man die vermeintliche Schwäche der Analyse in eine Stärke zu wenden versuchen, indem man argumentiert, daß die behauptete morphologische Fundierung des V-I-Bewegungsparameters ohnehin nicht unproblematisch ist. Angesichts dessen, daß der hier relevante Begriff der morphologischen Stärke von I im Detail umstritten ist und auf jeden Fall einigermaßen abstrakt sein muß, erscheint diese Strategie nicht von vornherein vergeblich.

Eine zweite Möglichkeit könnte darin bestehen, die fehlende morphologische Fundierung explizit als Metabeschränkung für Beschränkungsordnungen zu stipulieren, also etwa auf Beschränkungsordnungen einen Filter applizieren zu lassen der Art (69):

- (69) STR-BD \gg LEX-ÖKON nur bei morphologisch reichhaltigem I.

Ein solcher Ausweg wäre aber ad hoc, und er wäre auch konzeptuell fragwürdig, da er einen neuen Beschränkungstyp (Metabeschränkungen für Beschränkungsordnungen) in die Grammatik einführen würde.

Der dritte, radikalste, aber an sich wohl eleganteste Weg aus dem Dilemma ist der, die beobachteten morphologischen Korrelationen nicht als Ursache des Parameters zu sehen, sondern als Konsequenz seiner Setzung, d.h., als Konsequenz einer gegebenen Beschränkungsordnung. Letztlich haben wir genau dies bereits bei Grimshaws (1997) *do*-Einsetzungsanalyse gesehen: Ob eine Sprache in Negationskontexten expletives *do* aufweist, hängt nicht von unabhängigen morpho-lexikalischen Gegebenheiten ab, sondern einzig und allein von einer bestimmten Beschränkungsordnung. Dasselbe gilt für die Verwendung von W-Skopusmarkern in der oben vorgestellten Analyse aus Müller (1997). Legendre, Smolensky & Wilson (1998, 266) haben für derartige Fälle einen griffigen Slogan geprägt: "Das funktionale Lexikon ist Sklave der Syntax". Für die Flexionskategorie I als Teil des funktionalen Lexikons wäre dann zu zeigen, daß sie unter der Ordnung STR-BD \gg LEX-ÖKON in einer

Sprache als morphologisch reichhaltig erscheint, unter der Ordnung LEX-ÖKON \gg STR-BD dagegen als schwach; der Grad der morphologischen Reichhaltigkeit von I wäre somit nichts als ein Reflex des V-I-Bewegungsparameters. Ein solches Vorgehen scheint bei aller Verwegenheit nicht aussichtslos. In der Tat sind verschiedene Schritte in dieser Richtung unternommen worden, aber die konkrete Umsetzung hat sich bislang immer als relativ problematisch erwiesen.⁴⁶

Der Pro-Drop-Parameter Eine vergleichbare Situation ergibt sich beim Pro-Drop-Parameter, der bestimmt, ob eine Sprache pronominale Argumente lexikalisch unrealisiert lassen kann. Häufig wird angenommen, daß die Möglichkeit der Weglassung pronominaler Subjekte wiederum mit der Reichhaltigkeit von I korreliert (daß hier jedoch ein anderes Konzept von Reichhaltigkeit als beim V-I-Bewegungsparameter relevant sein muß, sieht man schon daran, daß im Französischen der eine Parameter negativ gesetzt ist, der andere positiv). Klassische Beispiele, die die unterschiedliche Setzung des Pro-Drop-Parameters anhand des Italienischen und Englischen zeigen, finden sich in (70):

- (70) a. – Ha cantato
 b. *– Has sung

Wiederum kann in einer genuin optimalitätstheoretischen Analyse dieser Parameter nicht direkt vom morphologischen Reichtum der Verbalflexion einer Sprache abhängen, sondern muß auf Beschränkungsumordnung zurückgehen. Grimshaw & Samek-Lodovici (1995; 1998) und Samek-Lodovici (1996) argumentieren, daß Pro-Drop wie in (70) nur möglich ist, wenn das Pronomen semantisch als Topik interpretiert wird. Ihre Analyse basiert u.a. auf den folgenden zwei Beschränkungen (die Darstellung ist vereinfacht; aber die weiteren Beschränkungen in der Analyse von Grimshaw und Samek-Lodovici ändern nichts am hier zu erläuternden Kern):

⁴⁶Tatsächlich ist die Analyse in Vikner (2000) zwar noch optimalitätstheoretisch formuliert; sie kehrt aber im wesentlichen zu einer prä-optimalitätstheoretischen Argumentation zurück insofern, als die Reichhaltigkeit der Verbalflexion als unabhängig von der Syntax festgelegt betrachtet wird. Konkret ist die Annahme, daß in der morphologischen Komponente entschieden wird, ob eine Sprache reichhaltige Verbflexion hat oder nicht. In der Syntax interagieren dann eine Beschränkung, die zwecks Merkmalsüberprüfung Kopfbewegung hin zu reicher Verbflexion fordert, und eine Ökonomiebeschränkung, die dies verbietet. Da in diesem System keine Beschränkung existiert, die Kopfbewegung hin zu schwacher Verbflexion fordert, läßt sich ableiten, daß keine Sprache V-I-Bewegung hat, die schwache Verbflexion aufweist (umgekehrt sind aber solche Sprachen als möglich vorgesehen, die trotz starker Verbflexion keine V-I-Bewegung zeigen, wofür es – zumindest im Bereich der SVO-Sprachen – kaum Evidenz gibt). Wenn man so will, ist dies also eine vierte Möglichkeit, morphologische Fundierung von syntaktischer Parametrisierung in eine optimalitätstheoretische Analyse zu integrieren. Man kann sich aber leicht klar machen, daß dieser Ansatz für die betrachteten Fälle nicht sehr weit von z.B. einer minimalistischen Analyse entfernt ist, die die Überprüfung starker Merkmale durch syntaktische Bewegung verlangt und aufgrund des Zauderprinzips Analoges für schwache Merkmale ausschließt.

- (71) a. TOP-TIL (“Topik-Tilgung”, “Drop Topic”):
Argumente, die mit dem Topik koreferent sind, dürfen strukturell nicht realisiert werden.
- b. REAL (“Realisierung”, “Parse”):
Input-Elemente müssen realisiert werden.

Im Italienischen gilt die Ordnung TOP-TIL \gg REAL. Dies bedeutet, daß im Input vorhandene, als Topik interpretierte Pronomina in dieser Sprache unter Verletzung von REAL weggelassen werden müssen, um TOP-TIL zu erfüllen. Sätze wie *Lui ha cantato*, mit realisiertem Pronomen, sind daher nur möglich, wenn das Pronomen nicht als Topik interpretiert wird. Umgekehrt ist die Situation im Englischen: Unter der dort angenommenen Ordnung REAL \gg TOP-TIL muß ein als Topik interpretiertes Pronomen auch im Output erscheinen (ebenso wie nicht-topikalisch interpretierte Pronomina). Von den drei oben erwähnten Wegen, die man angesichts einer weithin behaupteten Korrelation eines Parameters mit morphologischen Gegebenheiten einschlagen kann, gehen Grimshaw und Samek-Lodovici den ersten: Sie argumentieren, daß ihre Analyse den Standardanalysen, die auf morphologischer Fundierung des Pro-Drop-Parameters beruhen, überlegen ist, und daß eine solche Fundierung nur scheinbar ist. So erklärt ihre Analyse anders als auf morphologischer Fundierung basierende Pro-Drop-Ansätze, warum die Weglassung von Subjektpronomina generell an eine Topikinterpretation gekoppelt zu sein scheint.⁴⁷ Darüber hinaus wird ein notorisches Problem für die Idee der morphologischen Fundierung vermieden: Sprachen wie Chinesisch, Japanisch und Koreanisch weisen Pro-Drop auf, ohne über Verbalflexion zu verfügen (abgesehen einmal vom Phänomen der Höflichkeitskongruenz, wie es z.B. das Koreanische zeigt). Im vorgestellten optimalitätstheoretischen Ansatz heißt das, daß diese Sprachen die Ordnung TOP-TIL \gg REAL aufweisen; in den Standardanalysen sieht es dagegen so aus, als müsse auf der Basis dieser Evidenz der Begriff der morphologischen Reichhaltigkeit stark modifiziert oder letztlich sogar ganz aufgegeben werden.⁴⁸

⁴⁷Artstein (1998) entwickelt einen vergleichbaren optimalitätstheoretischen Ansatz, der vor allem mit dem Einfluß der grammatischen Kategorie Person (1./2. Person vs. 3. Person) auf die Möglichkeit von Pro-Drop im Modernen Hebräischen befaßt ist und ebenfalls auf morphologische Fundierung verzichtet.

⁴⁸Vgl. etwa Jaeggli & Safir (1989), die das Vorhandensein von “morphologisch uniformen Flexionsparadigmen” in einer Sprache als entscheidend für die Lizenzierung von Pro-Drop ansehen (und nicht morphologische Reichhaltigkeit) – ein leeres Flexionsparadigma kommt dabei der Forderung nach morphologischer Uniformität per se nach.

8. Empirische Gegenevidenz

8.1. Optionalität

Echte Optionalität (die auf ein identisches Beschränkungsprofil oder auf eine Koppelung zurückgeht) erwartet man in einer Theorie der Optimalität eigentlich nicht. (Normalerweise kann nur eine(r) der/die beste sein.) Tatsächlich stellt sich bei genauerer Betrachtung heraus, daß keines der in (19) in Kapitel 1 erwähnten Mittel, Optionalität zu behandeln, vollkommen unproblematisch ist (inkl. der Pseudo-Optionalität). Dieses Thema steht im Vordergrund von Kapitel 5.

8.2. Grammatikalitätsgrade

Dieses Konzept ist, wie erwähnt, in einer optimalitätstheoretischen Syntax zunächst einmal nicht erfaßbar. Man kann nun versuchen, diesen Mangel durch geeignete, möglichst konservative Modifikationen der Optimalitätstheorie zu beheben (vgl. Kapitel 6). Dabei sollte man aber nicht übersehen, daß das Konzept der Grammatikalitätsgrade zwar in manchen Standardtheorien erfaßt wird, aber nicht in prinzipieller Weise, sondern schlicht durch Stipulation: Ob die Verletzung einer Beschränkung z.B. in der GB-Theorie zu starker (ECP) oder schwacher (Subjazenz) Ungrammatikalität führt, folgt aus keiner inhärenten Eigenschaft der Beschränkungen; der Unterschied ist schlicht und einfach so behauptet. Dasselbe gilt für die weit verbreitete Annahme, daß ein Satz desto stärker ungrammatisch wird, je mehr Beschränkungen er verletzt. Im Prinzip könnte man derartige ad-hoc-Annahmen wohl auch in der optimalitätstheoretischen Syntax machen (abhängig von der Art, wie man Ungrammatikalität an sich ableitet, s.u.). Der zu zahlende Preis wäre wiederum eine willkürliche Erweiterung des theoretischen Rahmens. Es scheint daher gerechtfertigt, diesem Vorbehalt gegen eine optimalitätstheoretische Syntax nicht zu viel Gewicht beizumessen: Andere Theorien stehen hier auch nicht besser da, und falls es gelänge, eine systematische Behandlung der Grammatikalitätsgrade im Rahmen der Optimalitätstheorie zu leisten, ohne dabei deren Grundannahmen aufzugeben, wäre dies letztlich sogar ein möglicher Vorteil.

8.3. Absolute Ungrammatikalität

Zunächst einmal sieht es so aus, als müsse es, damit man einen Satz als ungrammatisch ausschließen kann, einen anderen Kandidaten in derselben Kandidatenmenge geben, der ihn als suboptimal blockiert (der beste Kandidat muß an sich gut genug sein). Dies ist nicht immer offensichtlich. Man betrachte ein ungrammatisches Beispiel aus dem Deutschen:

(72) *Was₁ ist Fritz eingeschlafen [CP nachdem er t₁ gelesen hat] ?

Hier ist unzulässigerweise eine W-Phrase aus einer Adjunkt-Insel extrahiert worden. Im folgenden sind einige der in der Literatur vorgeschlagenen Wege aufgeführt, die

absolute Ungrammatikalität in Fällen wie (72) in einem optimalitätstheoretischen Rahmen abzuleiten. (Dieses Problem ist in der Literatur auch unter dem Schlagwort “Unaussprechbarkeit” (‘ineffability’) bekannt.)

8.3.1. *Der Generator: Pesetsky (1997)*

Wenn auch der Schwerpunkt optimalitätstheoretischer Analysen auf der Grammatikkomponente der Harmonie-Evaluation (mit verletzbaren und geordneten Beschränkungen) liegt, so darf man doch nicht vergessen, daß dieser Komponente der Generator vorgelagert ist, der mit Hilfe von unverletzbaeren und nicht geordneten Beschränkungen die Kandidaten erzeugt (vgl. die Abbildung auf Seite 14). Eine naheliegende Möglichkeit, Sätze wie (72) auszuschließen, ist daher die, sie gar nicht erst durch den Generator erzeugen zu lassen. Dann kommt (72) nicht in den Wettbewerb, und das Problem, einen besseren konkurrierenden Kandidaten zu finden, der (72) blockiert, stellt sich nicht. Hierzu könnte man annehmen, daß eine Beschränkung wie ADJUNKT-INSEL, die W-Bewegung über eine Adjunkt-Insel hinweg verbietet, schon Teil des Generators ist.

(73) ADJUNKT-INSEL:

Bewegung darf kein Adjunkt überqueren.

Ein derartiger Vorschlag wird für vergleichbare Konstruktionen von Pesetsky (1997, 150f.) gemacht. Der Generator kann dann aber keine ganz reduzierte Standardgrammatik sein; er muß größere Aufgaben übernehmen können, als manche Vertreter der optimalitätstheoretischen Syntax ihm zuzugestehen bereit sind. Eine weitere Konsequenz ergibt sich, wenn wir annehmen, daß die Beschränkungen des Generators universell sind. Dann würde bei dieser Erklärung der Ungrammatikalität von (72) vorhergesagt werden, daß es keine Sprache geben kann, in der ein solcher Satz wohlgeformt ist. Falls sich diese Prognose als zu stark erweisen sollte, würde dies dafür sprechen, die Ungrammatikalität von (72) doch im Rahmen der Harmonie-Evaluation abzuleiten; falls sie sich jedoch bestätigen sollte, könnte dies umgekehrt für eine Lokalisierung im Generator sprechen.

Die im folgenden zu besprechenden vier Möglichkeiten beruhen alle auf der Annahme, daß ein Satz wie (72) vom Generator erzeugt und dem optimalitätstheoretischen Wettbewerb ausgesetzt wird. Sie unterscheiden sich jedoch in der Identifikation des optimalen Kandidaten, der den Wettbewerb gewinnt.

8.3.2. *Leere Kandidaten: Ackema & Neeleman (1998)*

Eine auf Prince & Smolensky (1993) zurückgehende und von Ackema & Neeleman (1998, 480) (anhand eines etwas anderen Falles, der Unmöglichkeit von Mehrfachfragen im Irischen) für die Syntax ausgearbeitete Erklärung beruht auf der Annahme, daß Kandidatenmengen so definiert sind, daß in jedem Wettbewerb auch ein Kandidat teilnimmt, der den Input vollständig unrealisiert läßt. Dies ist der “leere Kandidat” (“Null Parse”) \emptyset . Normalerweise gewinnt \emptyset in einem Wettbewerb nicht

gegen die konkurrierenden Kandidaten, denn es gibt eine Beschränkung, die dieser Kandidat verletzt und die sehr hoch geordnet ist:⁴⁹

(74) * \emptyset (“Vermeide leere Kandidaten”, “Avoid Null Parse”):

Der Input darf nicht vollständig unrealisiert sein.

Der ungrammatische Satz (72) kann nun durch einen leeren Kandidaten als Gewinner des Wettbewerbs blockiert werden, wenn ADJUNKT-INSEL eine verletzbarere Beschränkung ist, die noch höher geordnet ist als * \emptyset ; denn der leere Kandidat verletzt ADJUNKT-INSEL trivialerweise nicht. Ebenfalls höher geordnet als * \emptyset muß die Beschränkung W-KRIT sein; denn ansonsten würden wir erwarten, daß der optimale Kandidat die Verletzung von ADJUNKT-INSEL nicht durch Nicht-Realisierung aller Input-Elemente, sondern durch eine W-in-situ-Strategie vermeidet. Dies zeigt Tabelle T_{2.23}.

T_{2.23}: Ungrammatikalität und leere Kandidaten

Kandidaten	ADJUNKT-INSEL	W-KRIT	* \emptyset
K ₁ : was ₁ ... [<i>CP</i> nachdem er t ₁ V]	*!		
K ₂ : – ... [<i>CP</i> nachdem er was ₁ V]		*!	
☞K ₃ : \emptyset			*

Die zugrundeliegende Idee ist hier also quasi, daß es in manchen Fällen besser ist, zu schweigen, als etwas sehr Falsches zu sagen. Dieser Weg, Ungrammatikalität abzuleiten, ist vielleicht für manch einen intuitiv plausibel; er ist aber nicht ganz unproblematisch, denn er impliziert, daß alle Beschränkungen, die oberhalb von * \emptyset geordnet sind, de facto unverletzbar werden. Mit der Analyse von W-Bewegung in Mehrfachfragen im Deutschen, die in Abschnitt 1 entwickelt wurde, ist die Analyse in Tabelle T_{2.23} z.B. nicht vereinbar, denn dort war es wesentlich, daß W-KRIT von wohlgeformten Sätzen verletzt werden konnte; vgl. etwa Tabelle T_{2.2}. Eine Unterordnung von W-KRIT unter * \emptyset , wie sie Mehrfachfragen notwendig machen würden, würde aber in T_{2.23} zu einem nicht intendierten Ergebnis führen, daß nämlich optimal nicht der leere Kandidat K₃, sondern der W-in-situ-Kandidat K₂ wäre. Dieses Ergebnis entspricht tatsächlich einem dritten Weg, Ungrammatikalität in der optimalitätstheoretischen Syntax abzuleiten.

⁴⁹Ackema & Neeleman nennen diese Beschränkung PARSE. Intendiert ist jedoch nicht das, was die oben anhand der Analyse von Grimshaw & Samek-Lodovici (1995; 1998) besprochene Beschränkung REAL (“Parse”) (vgl. (71-b)) leistet – nämlich für jedes einzelne Input-Element Nicht-Realisierung zu verbieten –; vielmehr soll nur die Nicht-Realisierung des gesamten Inputs verboten werden. Um terminologische Konfusion zu vermeiden, verwende ich für letzteres Verständnis einen anderen Namen.

8.3.3. *Schlechte Gewinner: Grimshaw (1994), Müller (1997)*

Nehmen wir einmal an, daß es entweder leere Kandidaten in der Syntax nicht gibt, oder daß es sie zwar gibt, daß aber die Beschränkung, die sie verbietet (* \emptyset), die hier zur Debatte stehenden anderen Beschränkungen dominiert. Somit kann im Wettbewerb der W-in-situ-Kandidat gewinnen. Dann kann es sein, daß dieser Kandidat trotz syntaktischer Optimalität Eigenschaften besitzt, die ihn für andere linguistische Komponenten (z.B. Phonologie oder Semantik) unzugänglich machen, d.h., daß der Gewinner zwar im technischen Sinne optimal, aber nichtsdestoweniger in gewisser Weise "schlecht" ist. Somit wäre die Ungrammatikalität von (72) abgeleitet. Derartige Vorschläge sind in Grimshaw (1994, 15) und Müller (1997, 285) gemacht worden.⁵⁰

Konkret könnten wir für den vorliegenden Fall annehmen, daß die Beschränkung PUR-EP in der Variante, die mehrfache W-Bewegung blockiert (vgl. (8)), nur auf der S-Struktur gilt, W-KRIT dagegen auch auf LF. Somit wäre (75-a) zwar die optimale S-Struktur, aber auf LF wäre der optimale Kandidat dann gemäß Tabelle T_{2.2} das als S-Struktur-Repräsentation wegen PUR-EP noch suboptimale (75-b) (vgl. (3)).

- (75) a. [_{CP} Wann₁ hat Fritz t₁ [_{NP} welches Buch]₂ gelesen] ? S-Struktur
 b. [_{CP} Wann₁ welches Buch₂ hat Fritz t₁ t₂ gelesen] ? LF

Weiterhin sei einmal vorausgesetzt, daß Fragesätze nur von der semantischen Komponente interpretiert werden können, wenn alle W-Elemente auf LF in der Domäne von C_[+w] stehen.⁵¹ Für unseren ungrammatischen Satz (72) ergeben sich dann die erwünschten Konsequenzen: In Tabelle T_{2.24} blockiert K₂ K₁ als suboptimal; weil ADJUNKT-INSEL S-strukturelle und LF-Bewegung gleichermaßen blockiert, verharrt in K₂ die W-Phrase auch auf LF noch in situ. Somit ist K₂ zwar syntaktisch wohlgeformt, aber von der semantischen Komponente nicht interpretierbar (signalisiert wird das in (76) durch das Doppelkreuz #).

T_{2.24}: *Ungrammatikalität und schlechte Gewinner*

Kandidaten	ADJUNKT-INSEL	W-KRIT
K ₁ : was ₁ ... [_{CP} nachdem er t ₁ V]	*!	
☞K ₂ : - ... [_{CP} nachdem er was ₁ V]		*

- (76) #Fritz ist eingeschlafen [_{CP} nachdem er was₁ gelesen hat] ?

⁵⁰Grimshaw redet hier von einer "no good output situation" – dies ist im übrigen durchaus vergleichbar dem, was Chomsky (1995) meint, wenn er davon spricht, daß bestimmte Derivationen "als Kauderwelsch konvergieren" ('converge as gibberish').

⁵¹Dies ist semantisch nicht abwegig und tatsächlich oft explizit so vorgeschlagen worden; vgl. etwa Stechow (1993) und Beck (1996).

In gewisser Weise ist dieser Lösungsvorschlag dem von Pesetsky verfolgten Ansatz vergleichbar: In beiden Fällen erfolgt ein Bezug auf eine zusätzliche grammatische Domäne mit unverletzbaren Beschränkungen; nur ist diese Domäne dem optimalitätstheoretischen Harmonie-Evaluationssystem im einen Fall vorgelagert (Generator), im anderen Fall nachgelagert (semantische Interpretation). In beiden Fällen kann man jedoch sagen, daß die Erklärung für die Ungrammatikalität eines Satzes keine genuin optimalitätstheoretische ist. Anders als bei der Generator-basierten Erklärung der Ungrammatikalität von (72) besteht in der vorliegenden Analyse aber Spielraum für sprachspezifische Variation: Falls W-KRIT über ADJUNKT-INSEL geordnet ist, können Adjunktinseln von W-Bewegung verletzt werden.

8.3.4. Reparatur

Die zunächst einmal wohl attraktivste Erklärung von Fällen absoluter Ungrammatikalität würde in dem Nachweis bestehen, daß ein ungrammatischer Satz immer durch einen Reparaturkandidaten blockiert wird, der eine Verletzung einer Treuebeschränkung involviert, die nur in diesem Kontext in Kauf genommen werden kann. Diese Treuebeschränkung könnte sich auf semantische Eigenschaften von lexikalischen Einheiten beziehen; im vorliegenden Fall könnte sie z.B. die Reinterpretation des Fragepronomens *was* als Indefinitpronomen *was* (vgl. *etwas*) verbieten. Die Reparaturform, die diese Treuebeschränkung verletzt, aber (72) als suboptimal blockiert, weil nunmehr ADJUNKT-INSEL (es liegt keine Bewegung vor) und W-KRIT (es liegt kein W-Element vor, das bewegt werden müßte) beide leer erfüllt werden können, wäre dann (77-a). Alternativ könnte sich die fragliche Treuebeschränkung auf die Einsetzung zusätzlichen lexikalischen Materials beziehen; im vorliegenden Fall könnte eine solche Einsetzung z.B. ein resumptives Pronomen und eine Matrixpräposition für das W-Element betreffen. Auch in der entsprechenden Reparaturform (77-b) wären somit ADJUNKT-INSEL (weil keine Bewegung mehr vorliegt) und W-KRIT beide erfüllt, und (72) könnte als suboptimal blockiert werden.

- (77) a. Fritz ist eingeschlafen [_{CP} nachdem er *was*₁ gelesen hat] (= etwas)
 b. Bei *was*₁ ist Fritz eingeschlafen [_{CP} nachdem er *es*₁ gelesen hat] ?

Zur Illustration sei hier die erste Option kurz ausgeführt. Es sei also vorausgesetzt, daß der Generator in der Lage ist, ein [+w] markiertes Element des Inputs in ein [-w] markiertes Element umzuwandeln und so aus einem W-Element *was* ein Indefinitum *was* zu machen.⁵² Diese Reinterpretation verletzt die folgende Treue-

⁵²Diese Umwandlung ist natürlich nur bei einer Teilklasse von W-Elementen ohne lexikalische Manipulationen möglich – das Beispiel in (72) ist insofern bereits günstig gewählt worden. Bei *welches* oder *warum* funktioniert dies bereits nicht mehr so einfach – hier wären die entsprechenden [-w] markierten Formen vielleicht *ein* bzw. ein noch komplexerer Ausdruck wie *aus einem Grund*. Zur semantischen Treueverletzung käme also in jedem Fall noch eine lexikalische. Dies

beschränkung, die unterhalb von ADJUNKT-INSEL und W-KRIT geordnet ist.

(78) W-TREUE:

Eine im Input vorhandene W-Merkmalsspezifikation muß im Output realisiert werden.

Warum unter diesen Annahmen (77-a) (72) als Reparaturstrategie blockieren kann, zeigt Tabelle T_{2.25}.⁵³

T_{2.25}: *Ungrammatikalität und Reparatur*

Kandidaten	ADJUNKT-INSEL	W-KRIT	W-TREUE
K ₁ : was _{1[+w]} ... [CP nachdem er t ₁ V]	*!		
K ₂ : - ... [CP nachdem er was _{1[+w]} V]		*!	
☞ K ₃ : - ... [CP nachdem er was _{1[-w]} V]			*

K₁ ist die zu blockierende Variante; dieser Kandidat verletzt fatal ADJUNKT-INSEL. K₂ ist der Kandidat mit W-in-situ; er verletzt fatal W-KRIT. K₃ dagegen verletzt lediglich die tiefstgeordnete Beschränkung W-TREUE und ist somit optimal. Wie generell bei optimalitätstheoretischen Reparaturanalysen, so würde sich auch hier ergeben, daß die Treueverletzung in T_{2.25} wirklich nur als letzter Ausweg erlaubt ist: Wenn es möglich ist, W-KRIT ohne eine Treueverletzung zu erfüllen, ist eine solche Verletzung immer fatal.

Diese Lösung scheint somit auf den ersten Blick ganz gut zu funktionieren. Tatsächlich ist sie als generelle Erklärung aber leider nicht haltbar (und das ist auch der Grund, warum dieser Abschnitt keinen entsprechenden Literaturverweis im Titel enthält). Der Grund ist, daß die Konstruktionen in (77) eben keine offensichtlichen Reparaturstrategien sind, denn sie sind auch in Fällen möglich, wo lange W-Bewegung ebenfalls erlaubt ist, und wo wir nach dem eben Gesagten eigentlich erwarten sollten, daß sie blockiert sind: In (79-a) ist lange W-Bewegung möglich, und nichtsdestoweniger kann sowohl ein W-Indefinitum anstatt eines echten Fragewortes auftreten (vgl. (79-b)), als auch im Matrixsatz ein optionales W-Argument stehen, das im eingebetteten Satz ein resumptives Pronomen bindet (vgl. (79-c)):

- (79) a. Was₁ glaubt Fritz [CP t'₁ daß er t₁ lesen sollte] ?
 b. Fritz glaubt [CP daß er was lesen sollte]
 c. Von was₁ glaubt Fritz [CP daß er es₁ lesen sollte] ?

verkompliziert dann zwar die Analyse etwas, ändert aber substantiell nichts.

⁵³Wenn W-KRIT so formuliert ist, daß auch C_[+w] Lexikalisierung in seiner Domäne verlangt (s.o.), dann muß auch hier die [$\pm w$]-Spezifikation umgestellt werden, um W-KRIT ohne Bewegung zu erfüllen. Somit ergäbe sich für K₃ eine weitere W-TREUE-Verletzung. Da diese zusätzliche Verletzung keine weiteren Konsequenzen hätte, werde ich sie hier und im folgenden ignorieren.

Für (79-b) vs. (79-a) ist diese fatale Konsequenz in Tabelle T_{2.26} gezeigt: K₃ wird hier immer durch K₁ blockiert werden.

T_{2.26}: *Eine falsche Vorhersage*

Kandidaten	ADJUNKT-INSEL	W-KRIT	W-TREUE
☞ K ₁ : was _{1[+w]} ... [_{CP} t' ₁ daß er t ₁ V]			
K ₂ : - ... [_{CP} - daß er was _{1[+w]} V]		*!	
K ₃ : - ... [_{CP} daß er was _{1[-w]} V]			*!

Der letzte hier zu diskutierende Vorschlag versucht, den Kern des Reparaturansatzes zu bewahren, aber dabei die fatale Konsequenz der Untergenerierung zu vermeiden.

8.3.5. Neutralisierung: Legendre et al. (1995; 1998), Baković & Keer (1999)

Was die in Abschnitt 2 und danach behandelten Reparaturformen (wie expletives *do*, *es*, *was*, den Ersatzinfinitiv und das R-Pronomen) von den gerade diskutierten Formen (wie indefinit interpretiertem *was* oder A-Quer-Ketten der Art $\langle \textit{von was, es} \rangle$) wesentlich unterscheidet, ist, daß man gut motivieren kann, daß erstere beim Auftreten in einem beliebigen Output per se eine Treuebeschränkung verletzen müssen (wie VOLL-INT oder PROJ-P), während letztere dies nicht notwendigerweise tun. Wir haben gesehen, daß ein indefinites *was* im Output, das durch Umwandlung eines [+w]-Merkmals in ein [-w]-Merkmal entsteht, eine Beschränkung wie W-TREUE verletzt; aber nichts spricht dagegen, anzunehmen, daß es ein indefinites *was* im Output geben kann, das bereits im Input mit einem [-w]-Merkmal markiert war und somit nicht W-TREUE verletzt. Läßt man nun prinzipiell beides zu, so zeigt sich, daß man einerseits sicherstellen kann, daß in bestimmten Kontexten sowohl W-Bewegung als auch eine W-Indefinit-Strategie möglich ist (weil beide Konstruktionen Gewinner ihrer jeweiligen Wettbewerbe sein können, die sich im Fall von *was* nur bzgl. der [$\pm w$]-Spezifikation im Input unterscheiden), und andererseits ableiten kann, daß in anderen Kontexten (z.B. bei Adjunktinseln) W-Bewegung unter Verletzung einer Beschränkung wie W-TREUE zugunsten der W-Indefinit-Strategie als letztem Ausweg blockiert werden kann (weil in beiden Wettbewerben, bei [+w] wie bei [-w] im Input, der Gewinner die W-Indefinit-Strategie ist). Was in letzterem Fall passiert, ist also eine Neutralisierung eines Input-Unterschiedes, der zwei verschiedene Kandidatenmengen erzeugt, im Output. Derartige Analysen von Fällen absoluter Ungrammatikalität sind von Legendre et al. (1995, 615ff.; 1998, 274f.), Baković (1997), Keer & Baković (1997), Baković & Keer (1999), Vogel (1999) und Wilson (1999) entwickelt worden. Abschließend sei nun dargelegt, wie eine Neutralisierungsanalyse in Inselkontexten W-Bewegung zugunsten einer W-Indefinit-Strategie blockieren kann, ohne die fatale Konsequenz

des Reparaturansatzes nach sich zu ziehen.⁵⁴

Beginnen wir mit dem bewegungstransparenten Kontext in (79), und betrachten wir zunächst den Fall, wo das Pronomen *was* im Input als [+w] markiert ist. Tabelle T_{2.27} ist identisch mit Tabelle T_{2.26}. Hier gewinnt wie erwünscht Kandidat K₁ (= (79-a)), der die Merkmalspezifizierung beibehält und das W-Element in die Zielposition bewegt. Kandidat K₂ verletzt demgegenüber fatal W-KRIT, und Kandidat K₃, der das W-Element in ein Indefinitpronomen umwandelt, verletzt fatal W-TREUE.

T_{2.27}: *Bewegungstransparente Kontexte ohne Neutralisierung: ‘was* [+w] *’ im Input*

Kandidaten	ADJUNKT-INSEL	W-KRIT	W-TREUE
☞ K ₁ : was _{1[+w]} ... [CP t' ₁ daß er t ₁ V]			
K ₂ : - ... [CP - daß er was _{1[+w]} V]		*!	
K ₃ : - ... [CP daß er was _{1[-w]} V]			*!

Dies bedeutet nun jedoch nicht mehr, daß K₃ (= (79-b)) fälschlich als ungrammatisch vorhergesagt wird. K₃ verliert tatsächlich den Wettbewerb, in dem im Input [+w] steht, gewinnt aber den minimal unterschiedlichen Wettbewerb, in dem das Pronomen *was* im Input als [-w] spezifiziert worden ist. Hier ist umgekehrt K₁ aufgrund einer fatalen Verletzung von W-TREUE (Umwandlung von [-w] in [+w]) ausgeschlossen (dasselbe gilt natürlich auch für K₂); vgl. Tabelle T_{2.28}.

T_{2.28}: *Bewegungstransparente Kontexte ohne Neutralisierung: ‘was* [-w] *’ im Input*

Kandidaten	ADJUNKT-INSEL	W-KRIT	W-TREUE
K ₁ : was _{1[+w]} ... [CP t' ₁ daß er t ₁ V]			*!
K ₂ : - ... [CP - daß er was _{1[+w]} V]		*!	*
☞ K ₃ : - ... [CP daß er was _{1[-w]} V]			

In bewegungstransparenten Kontexten ist somit eine durch Umwandlung des [w]-Merkmals von Input zu Output verursachte Verletzung von W-TREUE immer fatal. Dies liegt daran, daß zwar die W-Bewegungsstrategie und die W-Indefinitstrategie miteinander konkurrieren, aber nicht mit derselben Ausgangsposition (“not on a level playing field”, wie Legendre, Smolensky & Wilson (1998, 257) sagen) – der “untreue” Kandidat spielt jeweils im Wettbewerb des “treuen” Kandidaten mit, hat aber qua W-TREUE-Verletzung einen inhärenten Nachteil, der ihm im Normalfall keine Chance läßt, optimal zu werden. Dieser Normalfall ist jedoch nicht mehr gegeben, wenn strikte Inselbeschränkungen wie ADJUNKT-INSEL

⁵⁴Wiederum könnte man alternativ eine Analyse entwickeln, die nicht semantische Treueverletzungen involviert, sondern ausschließlich lexikalische Treueverletzungen (Präposition plus resumptives Pronomen).

ins Spiel kommen: In diesem Fall wird neutralisiert. Die beiden relevanten Daten sind noch einmal wiederholt:

- (80) a. *Was₁ ist Fritz eingeschlafen [_{CP} nachdem er t₁ gelesen hat] ?
 b. Fritz ist eingeschlafen [_{CP} nachdem er was₁ gelesen hat] (= etwas)

Betrachten wir wiederum zunächst den Wettbewerb, in dem *was* im Input [+w] markiert ist. Tabelle T_{2.29} ist identisch mit Tabelle T_{2.25}. Hier ist der optimale Kandidat K₃ (= (80-b)); K₃ blockiert trotz W-TREUE-Verletzung K₁ (= (80-a)) und K₂, weil diese Konkurrenten die höher geordneten Beschränkungen ADJUNKT-INSEL bzw. W-KRIT verletzen.

T_{2.29}: Inselkontexte mit Neutralisierung: 'was [+w]' im Input

Kandidaten	ADJUNKT-INSEL	W-KRIT	W-TREUE
K ₁ : was ₁ [+w] ... [_{CP} nachdem er t ₁ V]	*!		
K ₂ : - ... [_{CP} nachdem er was ₁ [+w] V]		*!	
☞ K ₃ : - ... [_{CP} nachdem er was ₁ [-w] V]			*

Diesmal ist das Ergebnis dasselbe wie beim minimal variierenden Wettbewerb, in dem *was* [-w] markiert ist. Auch hier gewinnt K₃ gegen K₁ und K₂, und diesmal noch klarer, weil ohne W-TREUE-Verletzung. Dies zeigt Tabelle T_{2.30}.

T_{2.30}: Inselkontexte mit Neutralisierung: 'was [-w]' im Input

Kandidaten	ADJUNKT-INSEL	W-KRIT	W-TREUE
K ₁ : was ₁ [+w] ... [_{CP} nachdem er t ₁ V]	*!		*
K ₂ : - ... [_{CP} nachdem er was ₁ [+w] V]		*!	*
☞ K ₃ : - ... [_{CP} nachdem er was ₁ [-w] V]			

Die Neutralisierungsanalyse ist fraglos eine elegante Art, um Ungrammatikalität im Rahmen der Optimalitätstheorie abzuleiten, und sie benutzt anders als etwa Erklärungen, die die Lösung im Generator oder in einer Syntax-externen Grammatikkomponente suchen, auch nur Mittel, die dem optimalitätstheoretischen System immanent sind.⁵⁵ Dazu kommt, daß Neutralisierung auch das Phänomen der der Optionalität behandeln kann (vgl. Kapitel 5).

Allerdings ist auch dieses Konzept nicht vollkommen unproblematisch, weil in verschiedenen Bereichen die Komplexität ansteigt. Erstens entstehen systematisch leere derivationelle Ambiguitäten, in dem Sinne, daß ein Satz wie (80-b) zwei unterschiedliche, aber an sich beide wohlgeformte Entstehungsgeschichten haben kann,

⁵⁵Wie beim Ansatz mit leeren Kandidaten ergibt sich allerdings für den vorliegenden Fall eine Inkompatibilität mit der Analyse von W-Bewegung in Abschnitt 1: In Mehrfachfragen im Deutschen sollten alle W-in-situ-Elemente in Indefinita umgewandelt werden.

nämlich den Wettbewerb in $T_{2.29}$ und den Wettbewerb in $T_{2.30}$. Es ist nicht ganz klar, ob derartige leere Ambiguitäten nicht für den Spracherwerb und die Sprachverarbeitung problematisch sind; aber es ist erwähnenswert, daß Prince & Smolensky (1993) eine zusätzliche Annahme machen, die solche Ambiguitäten vermeiden soll: Sie schlagen ein Meta-Optimierungsverfahren der "Input-Optimierung" vor (in der Phonologie bekannt als "Lexikon-Optimierung"), das de facto ganze Wettbewerbe einem optimalitätstheoretischen Wettbewerb zweiter Stufe aussetzt: Im konkreten Fall gewänne der Wettbewerb in $T_{2.30}$ gegen den Wettbewerb in $T_{2.29}$ aufgrund des besseren Profils des optimalen Kandidaten, und (80-b) wäre so nicht mehr derivationell ambig. Ich werde in den Kapiteln 3 und 5 nochmals auf die Input-Optimierung zurückkommen; aber es sollte bereits klar sein, daß es sich hierbei um eine nicht-triviale Erweiterung des eigentlichen optimalitätstheoretischen Systems handelt.

Darüber hinaus setzen Neutralisierungsanalysen sehr große Kandidatenmengen voraus: Der Generator muß in der Lage sein, beliebige Input-Elemente im Output zu ignorieren, Input-Elemente in andere Output-Elemente umzuwandeln, und Output-Elemente einzufügen, die nicht im Input vorlagen. All dies ist in der Phonologie Standard; es ist aber nicht klar, ob sich derartiges in der Syntax, die mit komplexeren Strukturen arbeitet, aufrechterhalten läßt. Die vorliegende Analyse würde dabei sowohl in vielen Fällen das Einsetzen von nicht im Input vorhandenem lexikalischem Material erfordern (vgl. die Bemerkungen oben zu W-Elementen wie *welches*, die keine bloße Indefinitesart haben), als auch systematisch die Abänderung von semantischer Information (Umwandlung von einem Fragewort in ein Indefinitum). Dies bedeutet aber nicht nur, daß es völlig ausgeschlossen ist, Kandidatenmengen unter Bezug auf identische LF-Strukturen (bzw. semantische Interpretationen) der Kandidaten zu formulieren (vgl. Legendre, Smolensky & Wilson (1998) und Abschnitt 3 von Kapitel 3). Letzten Endes ist die Konsequenz, daß in jeder Kandidatenmenge jeder vom Generator erzeugbare Satz mit jedem anderen Satz konkurriert – auch wenn die Kandidaten, die dem jeweiligen Input in hinreichendem Maße untreu sind, niemals eine Chance haben, optimal zu werden. Damit ergibt sich aber nicht nur ein potentiell Problem der übermäßigen Komplexität; auch aus rein linguistischer Perspektive stellt sich die Frage, inwieweit die Annahme Plausibilität beanspruchen kann, daß über die Grammatikalität eines Satzes nicht nur der Vergleich mit (in einem durch geeignete Definition der Kandidatenmenge zu präzisierenden Sinne) "ähnlichen" Formen entscheidet, sondern der Wettbewerb mit allen überhaupt nur denkbaren satzähnlichen Formen. Will man eine solche Konsequenz vermeiden, muß man sicherstellen, daß der Generator nicht mit der Fähigkeit der willkürlichen Input-Manipulation ausgestattet ist. Dann werden jedoch Neutralisierungsanalysen schnell unmöglich.

Zusammenfassend ergibt sich, daß keiner der existierenden Ansätze, die absolute Ungrammatikalität von Sätzen in prinzipieller Weise abzuleiten, ganz ohne Probleme ist; aber abgesehen von einem reinen Reparaturansatz ist auch keiner von vornherein zum Scheitern verurteilt. Schließlich ist auch nicht ausgeschlossen, daß

mehrere Ansätze koexistieren können. Denkbar wäre z.B., gleichzeitig für bestimmte, als universell unmöglich postulierte Konstruktionen die Ungrammatikalität im Generator zu lokalisieren, für andere Konstruktionen dagegen mit einer Neutralisierungsanalyse aufzukommen (die idealerweise die Fähigkeiten des Generators zur massiven Input-Manipulation in Grenzen hält).

Kapitel 3

Drei Analysen

In diesem Kapitel sollen drei wichtige optimalitätstheoretische Analysen etwas genauer besprochen und eingeordnet werden, vor allem unter dem Aspekt, inwieweit sich die in Kapitel 2 herausgearbeiteten Typen von Evidenz für eine optimalitätstheoretische Herangehensweise an syntaktische Phänomene wiederfinden. Die Analysen sind auch mit der Maßgabe ausgewählt worden, daß sie jeweils neue Konzepte einführen. Konkret handelt es sich um Grimshaws Theorie der Inversion, Pesetskys Theorie der PF-Realisierung im C-System und die Lokalitätstheorie von Legendre, Smolensky & Wilson.

1. Grimshaws (1993; 1997) Theorie der Inversion

In diesem Abschnitt wird die in Grimshaw (1997) entwickelte Theorie behandelt; ein Vorläufer dieser Arbeit ist Grimshaw (1993).¹ Grimshaws Hauptziel ist es, den Nachweis zu erbringen, daß sich die subtilen Interaktionen von Operatorbewegungen, Kopfbewegungen und Füllungen von Kopfpositionen durch semantisch leere Köpfe im Englischen elegant optimalitätstheoretisch herleiten lassen. Die wichtigsten Konzepte sind Verbinversion und *do*-Einsetzung. Eine Instanz von letzterem haben wir bereits in Abschnitt 9.2 von Kapitel 1 kennengelernt. Zum Teil sind dort auch schon wichtige Grundannahmen eingeführt worden; im folgenden werden aber noch einmal alle wesentlichen Voraussetzungen erwähnt.

1.1. Grundannahmen

Grimshaw geht davon aus, daß die konkurrierenden Outputkandidaten S-Struktur-Repräsentationen sind. Kandidatenmengen werden wie in (1) definiert:

- (1) *Kandidatenmenge* (Grimshaw (1997, 375-376)):
Zwei Kandidaten K_i , K_j sind in derselben Kandidatenmenge gdw. gilt:
- a. K_i und K_j sind Realisierungen von identischen Prädikat-/Argument-Strukturen.
 - b. K_i und K_j haben identische LFs.

¹Grimshaw (1993) ist wohl das erste genuin optimalitätstheoretische Syntaxpapier, das größeren Einfluß erlangt hat. Ebenso früh, aber in der Wirkung nicht vergleichbar ist die Studie über Kasus und GF-Realisierung von Legendre, Raymond & Smolensky (1993).

Die Annahme ist hier, daß als Input jeweils eine Prädikat-/Argument-Struktur vorliegt (mit einer Spezifikation von Tempus und Aspekt), und daß zwei konkurrierende Kandidaten zusätzlich immer noch identische (genauer sagt Grimshaw: nicht-distinkte) LFs besitzen müssen. Somit ergibt sich strenggenommen eine Dissoziation von Input und Kandidatenmengendefinition; vgl. die Bemerkungen dazu in Abschnitt 5 von Kapitel 1.²

Eine gegebene Prädikat-/Argument-Struktur kann nun auf unterschiedliche Art und Weise in der Syntax realisiert werden. Diese unterschiedlichen Realisierungen haben immer denselben lexikalischen VP-Kern, sie können darüber hinaus aber beliebig komplexe erweiterte Projektionen (im Sinne von Grimshaw (1991)) dieser VP sein. Erweiterte Projektionen über der VP sind funktionale XPs, und die Satzstruktur ist variabel, je nachdem, ob und, falls ja, wie viele funktionale XPs über der VP sind. Wie bereits in Kapitel 1 dargelegt wurde, hat eine funktionale XP keine unabhängig definierbaren Merkmale; der funktionale Kopf X ist auch noch nicht im Input vorhanden, sondern wird durch Gen erst erzeugt. Dies gilt für Komplementierer wie *that* ebenso wie für expletives *do* im Englischen. Nichtsdestoweniger ist es für die gegenwärtigen Zwecke nützlich, sich auf diese funktionalen XPs oberhalb der durch die Prädikat-/Argument-Struktur festgelegten Kernprojektion mit den klassischen Termini IP, CP usw. zu beziehen: IP ist dann (normalerweise, wie in Kapitel 1 gesehen jedoch nicht bei Auftreten einer Negation) die erste funktionale Projektion über der VP, CP die zweite, und wo noch andere erweiterte Projektionen ins Spiel kommen, kann man diese mit weiteren mnemotechnisch günstigen Termini benennen. Entsprechend dieser Vorgabe läßt sich *do* als I-Kategorie betrachten und *that* als C-Kategorie.

Im Einklang mit der Annahme, daß die Outputkandidaten S-Struktur-Repräsentationen sind, sind alle verwendeten Beschränkungen S-Struktur-Filter. In Kapitel 2 haben wir bereits eine Reihe von Beschränkungen kennengelernt. Diese Beschränkungen werden wie die neu dazukommenden im folgenden an der Stelle eingeführt, wo sie relevant werden.

1.2. Inversion

Mit "Inversion" ist hier das gemeint, was in der klassischen Transformationsgrammatik "Subjekt-Auxiliar-Inversion" hieß, und was heute als Kopfbewegung von I

²Grimshaw (1997, 376) erwägt allerdings auch die Möglichkeit, daß die LF-Struktur Teil des Inputs ist, wie das Legendre, Smolensky & Wilson (1998) vorschlagen; s.u. Wenn jedoch der Input eine LF-Struktur beinhaltet, können semantisch untreue Kandidaten davon abweichen, und die Konklusion, daß konkurrierende Kandidaten nicht-distinkte LFs haben müssen, kann nicht mehr gelten. Wie bei der Diskussion von Ungrammatikalität in Kapitel 2 deutlich geworden ist und wie wir unten noch sehen werden, ist eben dies bei Legendre, Smolensky & Wilson (1998) intendiert. Da Grimshaw diese Konsequenz nicht vorsieht und im Gegenteil voraussetzt, daß konkurrierende Kandidaten notwendigerweise nicht-distinkte LFs haben, können wir schließen, daß in diesem Ansatz die LF-Struktur nicht Teil des Inputs ist.

nach C interpretiert wird. Wie (2) zeigt, ist diese Inversion im Deklarativsatz im Englischen unmöglich:

- (2) a. They will read some books
 b. *Will₂ they t₂ read some books

In nicht eingebetteten Fragesätzen ist dagegen Inversion obligatorisch:

- (3) a. Which books₁ will₂ they t₂ read t₁ ?
 b. *Which books₁ – they will read t₁ ?

Dies bedeutet, daß die Bewegung des Auxiliars von I nach C im Englischen möglich ist nur dann, wenn sie notwendig ist. Mit anderen Worten: Subjekt-Auxiliar-Inversion ist ein Reparaturphänomen und bietet sich daher für eine optimalitätstheoretische Analyse an. Die drei Beschränkungen, die benötigt werden, um den elementaren Fall abzuleiten, sind OB-KOPF und ÖKON, die bereits in Kapitel 2 eingeführt wurden, sowie die Beschränkung OP-SPEZ, die eine verallgemeinerte Variante des W-Kriteriums (bzw. der bereits diskutierten Beschränkung W-KRIT) ist insofern, als sie nicht nur über W-Phrasen redet, sondern ganz allgemein über Operatoren.³

- (4) a. OP-SPEZ (“Operator im Spezifikator”, “Op-Spec”):
 Syntaktische Operatoren müssen in einer Spezifikatorposition sein.
 (i) W-Operatoren müssen alle anderen Elemente der erweiterten V-Projektion c-kommandieren.
 (ii) Andere Operatoren (negative Quantoren z.B.) müssen nur die minimale Proposition c-kommandieren.
 b. OB-KOPF (“Obligatorische Köpfe”, “Ob-Hd”):
 Eine Projektion hat einen (nicht-leeren) Kopf.
 c. ÖKON (“Bewegungsökonomie”, “Stay”):
 Bewegung ist verboten (Spur ist nicht erlaubt).

Die Ordnung dieser drei Beschränkungen im Englischen ist OP-SPEZ \gg OB-KOPF \gg ÖKON. Die Erklärung für obligatorische Inversion in Matrixfragesätzen wie (3) ergibt sich nun wie folgt: Aufgrund von hoch geordnetem OP-SPEZ muß das W-Objekt in eine externe Spezifikatorposition SpecX (SpecC) bewegt werden. Die Existenz dieser Position verlangt ihrerseits nicht nur aus offensichtlichen Gründen die Existenz einer funktionalen XP-Schale (CP), sondern aufgrund von OB-KOPF auch die Existenz eines gefüllten Kopfes X (C). Diese Füllung erfolgt durch Kopfbewegung des in I befindlichen Auxiliars nach C. Insgesamt erfolgen hier also zwei Bewegungsschritte. Dadurch ist ÖKON zweifach verletzt. Nichtsdestoweniger ist die-

³Wie man jedoch anhand der Fallunterscheidung in (4-a) sehen kann, hat diese Verallgemeinerung einen Preis, denn W-Phrasen und andere Operatoren verhalten sich eben nicht hundertprozentig gleich. Die Evidenz dafür wird unten noch diskutiert.

ser Kandidat optimal, denn die Konkurrenten verletzen die höher geordneten Beschränkungen OP-SPEZ oder OB-KOPF, weil sie W-Bewegung oder Kopfbewegung unterlassen (oder beides). Dies zeigt Tabelle T_{3.1}.⁴

T_{3.1}: *Matrixfragesätze*

Kandidaten	OP-SPEZ	OB-KOPF	ÖKON
K ₁ : [<i>IP</i> NP will [<i>VP</i> read which books ₁]]	*!		
K ₂ : [<i>CP</i> e [<i>IP</i> NP will [<i>VP</i> read which books ₁]]	*!	*	
K ₃ : [<i>CP</i> which books ₁ e [<i>IP</i> NP will [<i>VP</i> read t ₁]]]		*!	*
☞K ₄ : [<i>CP</i> which books ₁ will ₂ [<i>IP</i> NP t ₂ [<i>VP</i> read t ₁]]]]			**
K ₅ : [<i>CP</i> will ₂ [<i>IP</i> NP t ₂ [<i>VP</i> read which books ₁]]]]	*!		*

Der Auslöser für die Verbbewegung von I nach C ist hier in letzter Instanz OP-SPEZ – nur wegen dieser Beschränkung muß eine funktionale Projektion oberhalb der IP geschaffen werden, deren Kopf dann wegen OB-KOPF zu füllen ist. Wir erwarten also, daß in Deklarativsätzen, wo OP-SPEZ leer erfüllt ist, weil keine W-Phrase (und, per Annahme, kein anderer Operator) vorhanden ist, Inversion unmöglich wird: Die Verletzung von tief geordnetem ÖKON ist fatal, wenn sie nicht zur Erfüllung höher geordneter Beschränkungen beiträgt. Diese Vorhersage ist wie in (2) gesehen korrekt. Den zugrundeliegenden Wettbewerb veranschaulicht Tabelle T_{3.2}.

T_{3.2}: *Matrixdeklarativsätze*

Kandidaten	OP-SPEZ	OB-KOPF	ÖKON
☞K ₁ : [<i>IP</i> NP will [<i>VP</i> read some books]]			
K ₂ : [<i>CP</i> e [<i>IP</i> NP will [<i>VP</i> read some books]]		*!	
K ₃ : [<i>CP</i> will ₂ [<i>IP</i> NP t ₂ [<i>VP</i> read some books]]]]			*!

In Kapitel 2 haben wir gesehen, wie man mit Hilfe einer Beschränkung wie W-KRIT ableiten kann, daß sich in Mehrfachfragen im Bulgarischen alle W-Phrasen in die C-Domäne bewegen, während sich in Mehrfachfragen im Deutschen oder Englischen nur genau eine W-Phrase nach SpecC bewegt und die anderen in situ verbleiben. Im Prinzip stünde eine derartige Erklärung auch Grimshaw (1997) zur Verfügung (insbesondere, da sie selbst die Beschränkung einführt, von der eine Variante für gerade diesen Zweck in Kapitel 2 benutzt worden ist, nämlich PUR-EP;

⁴Unterschlagen ist hier (wie in der folgenden Tabelle) die Bewegung des Subjekts von SpecV nach SpecI, die durch die Präsenz des Tempusauxiliars *will* erzwungen wird. Die Erklärung ist hier komplett analog zu der in Abschnitt 9.2 von Kapitel 1 gegebenen: Um KASUS und SUBJEKT zu erfüllen, muß sich der Kopf einer Subjekt-NP-Kette in der höchsten A-Spezifikatorposition befinden, die mit der finiten Verbmorphologie kongruiert. Bei Präsenz eines Auxiliars in I ist das SpecI.

s.u.). Tatsächlich wählt Grimshaw jedoch diesen Weg nicht. Ihre Formulierung von OP-SPEZ ist so, daß in Mehrfachfragen mehrfache W-Bewegung diese Beschränkung auch nicht besser erfüllen kann als eine Strategie mit einfacher W-Bewegung und W-in-situ. Denn gemäß (4-a) müssen W-Operatoren alle anderen Elemente der erweiterten V-Projektion c-kommandieren, und bei mehr als einer W-Phrase in einem Satz kann naturgemäß nur eine einzige die höchste sein, die dann alle anderen Elemente (inkl. aller anderen W-Phrasen) c-kommandiert.⁵ Somit ergibt sich, daß sich (5-a) hinsichtlich von OP-SPEZ genauso verhält wie (5-b) und (5-c) (beidesmal liegt entweder Adjunktion oder Bewegung in einen weiteren Spezifikator vor – letzteres impliziert dann jeweils noch einen leeren Kopf).

- (5) a. What will they put where ?
 b. *What will where they put ?
 c. *What where will they put ?

Die Sätze in (5-bc) akkumulieren allerdings gegenüber (5-a) ÖKON-Verletzungen, die fatal sind und zu Suboptimalität führen (bei Spezifikatorbewegung kommen noch zusätzliche OB-KOPF-Verletzungen hinzu). Dies zeigt Tabelle T_{3.3}.

T_{3.3}: Mehrfachfragen

Kandidaten	OP-SPEZ	OB-KOPF	ÖKON
\hookrightarrow K ₁ : [CP W ₁ Aux ₂ [IP NP t ₂ [VP V t ₁ W ₃]]]	*		**
K ₂ : [CP W ₁ Aux ₂ [W ₃ (e) [IP NP t ₂ [VP V t ₁ t ₃]]]]	*	(*)	***!
K ₃ : [W ₁ (e) [CP W ₃ Aux ₂ [IP NP t ₂ [VP V t ₁ t ₃]]]]	*	(*)	***!

Man beachte, daß diese Sichtweise aufgrund der harmonischen Begrenzung von K₂ und K₃ durch K₁ ganz unabhängig von Fragen der Beschränkungsordnung dazu zwingt, mehrfache Voranstellung von W-Phrasen im Bulgarischen nicht als einheitliches Phänomen zu sehen. Vielmehr muß sich hier eine W-Phrase bewegen, um OP-SPEZ zu erfüllen, und für die anderen vorangestellten W-Phrasen muß eine andere Beschränkung als Auslöser gefunden werden.⁶

1.3. ‘Do’-Einsetzung in Fragesätzen

Es gibt drei hauptsächliche syntaktische Kontexte, in denen im Englischen das expletive Verb *do* eingesetzt werden muß: (a) Negationskontexte, (b) Fragekontexte,

⁵Als Erläuterung dieser speziellen Einschränkung für W-Elemente, von der andere Operatoren nicht betroffen sind, sagt Grimshaw, daß W-Elemente anders als andere Operatoren den logischen Typ eines Satzes verändern (aus einer Proposition wird eine Menge von Propositionen).

⁶Vgl. auch Bošković (1999), der explizit dafür argumentiert, daß mehrfache W-Voranstellung im Bulgarischen kein homogenes Phänomen ist, und daß nur die höchste W-Phrase tatsächlich W-bewegt worden ist.

sowie (c) Fokus- bzw. Emphasekontexte. Die Interaktion von Negation und *do*-Einsetzung bei Grimshaw (1997) ist bereits in Kapitel 1 behandelt worden, zu Emphasekontexten wird hier nichts gesagt. Es geht im folgenden daher um *do*-Einsetzung in Fragesätzen. Die zugrundeliegende Generalisierung ist im Prinzip genau die gleiche wie bei der Negation: *Do* als expletives Verb ist möglich nur dann, wenn es notwendig ist (vgl. Chomsky (1957)). Grundannahme ist, daß expletives *do* nicht im Input auftaucht, sondern erst von Gen in I (also in einer funktionalen X-Position oberhalb der VP) eingesetzt wird, und daß es nur ein finites Verb pro erweiterter Projektion geben kann (d.h., sobald finites *do* auftritt, muß V infinit sein). Die Beschränkung, gegen die expletives *do* verstößt, ist VOLL-INT, das hier anders als in Kapitel 1 und 2 nicht mehr einfach als Beschränkung erscheint, die Expletiveinsetzung verbietet (vgl. (15-e) aus Kapitel 1), sondern in der "offiziellen" von Grimshaw vorgeschlagenen Form.

(6) VOLL-INT ("Vollständige Interpretation", "Full-Int"):

Die lexikalisch-konzeptuelle Struktur muß respektiert werden.

Diese Variante von VOLL-INT soll nicht nur die Einsetzung von Expletiva durch Gen verbieten, sie soll darüber hinaus auch noch erklären, welche Elemente als Expletiva verwendet werden können. Die Grundidee ist, daß Expletiva im Prinzip ganz gewöhnliche lexikalische Elemente sind, die für den expletiven Gebrauch in gewisser Weise "zweckentfremdet" wurden. Diese Zweckentfremdung erfolgt dadurch, daß alle spezifischen Bedeutungselemente des lexikalischen Elements entfernt werden. Wenn wir nun annehmen, daß die Bedeutung eines lexikalischen Elements in einer ihm zugeordneten lexikalisch-konzeptuellen Struktur festgelegt ist, dann impliziert die Entfernung semantischer Information der lexikalisch-konzeptuellen Struktur immer Verletzungen von VOLL-INT in (6). Je mehr semantische Merkmale entfernt werden müssen, desto mehr VOLL-INT-Verletzungen ergeben sich. Die wenigsten VOLL-INT-Verletzungen bei einer Verwendung als Expletiv wird folglich immer dasjenige lexikalische Element hervorrufen, das von vornherein schon die allgemeinste, am wenigsten durch spezifische Merkmale verengte Bedeutung hat. Bei Verben ist das lexikalische Element mit der allgemeinsten Bedeutung *do* (*sleep* z.B. würde VOLL-INT als Expletiv noch stärker verletzen); bei als Skopusmarker zu verwendenden W-Elementen (vgl. Kapitel 2) ist es im Deutschen *was* (im Hindi, im Ungarischen und in anderen Sprachen ist es ein dem deutschen *was* entsprechendes Element). Tatsächlich scheint im großen und ganzen die Annahme gut bestätigt, daß als Expletiva für eine Funktion X immer diejenigen Vertreter der Klasse X verwendet werden, die die allgemeinste Bedeutung haben, und dies folgt nun aus VOLL-INT in der Formulierung in (6). Was ebenfalls daraus (aber nicht aus der in Kapitel 1 vorausgesetzten einfacheren Variante) folgt, ist, daß ein Element, das von vornherein keine Bedeutung trägt, auch keine VOLL-INT-Verletzung hervorrufen kann. Dies gilt per Annahme bei Grimshaw (1997) für den Komplementierer *that*, und es erweist sich als essentiell, wenn es darum geht, Fälle der Optionalität von Komplementierern zu erfassen.

Nachdem geklärt ist, warum expletives *do* VOLL-INT verletzt und warum ein expletives *sleep* dieselbe Beschränkung noch häufiger verletzen würde (und daher im Englischen auch nicht als expletives Verb auftreten kann), bleibt noch die relative Ordnung von VOLL-INT zu den anderen bisher eingeführten Beschränkungen zu bestimmen. Bei der Diskussion von *do*-Einsetzung in Negationskontexten in Kapitel 2 hatte es sich gezeigt, daß VOLL-INT relativ niedrig geordnet sein muß, so daß ein optimaler Kandidat nicht, anstatt *do*-Einsetzung vorzunehmen, z.B. KASUS oder SUBJEKT verletzen kann. Ähnlich ist auch bei der W-Bewegung die Konsequenz: VOLL-INT muß unterhalb von OB-KOPF stehen, und, da OB-KOPF unterhalb von OP-SPEZ sein muß, per Transitivität der Ordnungsrelation auch unterhalb von OP-SPEZ. Nun können wir uns den Daten zuwenden.

(7) (vgl. (18) aus Kapitel 1) zeigt, daß *do*-Einsetzung im einfachen (nicht-negierten, nicht-emphatischen) Deklarativsatz unmöglich ist.

- (7) a. [VP John left]
 b. *[IP John₁ [I did] [VP t₁ leave]]

Da *do*-Einsetzung eine Reparaturstrategie ist und die inhärente VOLL-INT-Verletzung durch die Aktivität einer höher geordneten Beschränkung erzwungen werden muß, ergibt sich wie in Tabelle T_{3.4} gezeigt, daß der Kandidat K₂ mit *do*-Einsetzung zugunsten des Kandidaten K₁ blockiert ist. Man beachte, daß K₁ eine einfache VP ist; eine IP mit leerem Kopf wie in K₃ verletzt fatal OB-KOPF, ohne bei irgendeiner höher geordneten Beschränkung zu einem besseren Verhalten zu führen.

T_{3.4}: Einfache Deklarativsätze und 'do'

Kandidaten	OP-SPEZ	OB-KOPF	VOLL-INT	ÖKON
☞ K ₁ : [VP NP V]				
K ₂ : [IP NP do [VP V]]			*!	
K ₃ : [IP e [VP NP V]]		*!		

Die Situation ist anders in Matrixfragesätzen, die ein W-Objekt enthalten. Angenommen, es liegt kein Auxiliar vor, also kein Modalverb und auch kein Tempusauxiliar wie im Perfekt oder im Futur. Dann ergibt sich folgende Situation: Um OP-SPEZ zu erfüllen, muß die W-Phrase in einen externen Spezifikator (SpecC) wandern; vgl. (8-a) (das als echte Frage nicht wohlgeformt ist). Damit verlangt OB-KOPF die Füllung des nunmehr geschaffenen Kopfes (C); dies blockiert (8-b) sowie (8-c), wo *do*-Einsetzung ohne Inversion erfolgt ist. Der C-Kopf ist auf drei Arten zu füllen. Die erste Möglichkeit ist die Bewegung des finiten Hauptverbs; vgl. (8-d). Die zweite Möglichkeit ist die direkte Einsetzung von *do* in C; vgl. (8-e). Die dritte und allein erlaubte Möglichkeit ist schließlich die Bewegung eines in I erzeugten *do* nach C; vgl. (8-f).

- (8) a. *[VP She said what₁] ?

- b. $*[_{CP} \text{What}_1 - [_{VP} \text{she said } t_1]] ?$
 c. $*[_{CP} \text{What}_1 - [_{IP} \text{she did } [_{VP} t \text{ say } t_1]]] ?$
 d. $*[_{CP} \text{What}_1 \text{ said}_2 [_{VP} \text{she } t_2 t_1]] ?$
 e. $*[_{CP} \text{What}_1 \text{ did}_2 [_{VP} \text{she say } t_1]] ?$
 f. $[_{CP} \text{What}_1 \text{ did}_2 [_{IP} \text{she } t_2 [_{VP} t \text{ say } t_1]]] ?$

(8-d) verletzt fatal die Beschränkung LEX-ÖKON, die bereits in Kapitel 1 im Zusammenhang mit *do*-Einsetzung bei Negation eine Rolle gespielt hat, und von der auch dort schon geschlossen werden konnte, daß sie über VOLL-INT geordnet ist (denn ansonsten würde man in Negationskontexten im Englischen Hauptverbbe-
 wegung statt *do*-Einsetzung erwarten).

- (9) LEX-ÖKON (“Bewegungsökonomie für lexikalische Köpfe”, “No-Lex-Mvt”):
 Bewegung von lexikalischen Köpfen ist verboten (X^0 -Spur_{lex} ist nicht erlaubt).

Direkte Einsetzung von *do* in C wie in (8-e) führt dazu, daß das Subjekt nicht mehr mit finiter Verbmorphologie kongruiert und somit fatal die Beschränkung KASUS verletzt, die hier ebenfalls aus Kapitel 1 wiederholt ist und sich dort bereits als im Englischen über VOLL-INT geordnet erwiesen hat.⁷

- (10) KASUS (“Case”):

Der Kopf einer NP-Kette muß in einer Kasusposition sein.

Do-Einsetzung in I plus Bewegung nach C wie in (8-f) verletzt dagegen nur die tief geordneten Beschränkungen VOLL-INT und ÖKON. Daher ist dieser Kandidat optimal. Der soeben skizzierte Wettbewerb ist graphisch aufbereitet in Tabelle T_{3.5}.⁸

T_{3.5}: Matrixfragesätze und ‘do’

Kandidaten	OP-SPEZ	LEX-ÖKON	KASUS	OB-KOPF	VOLL-INT	Ö-KON
K ₁ : $[_{VP} \text{NP V } W_1]$	*!					
K ₂ : $[_{CP} W_1 e [_{VP} \text{NP V } t_1]]$				*!		*
K ₃ : $[_{CP} W_1 e [_{IP} \text{NP } do_2 [_{VP} t \text{ V } t_1]]]]$				*!	*	**
K ₄ : $[_{CP} W_1 V_3 [_{VP} \text{NP } t_3 t_1]]$		*!				**
K ₅ : $[_{CP} W_1 do_2 [_{VP} \text{NP V } t_1]]]$			*!		*	*
☞K ₆ : $[_{CP} W_1 do_2 [_{IP} \text{NP } t_2 [_{VP} t \text{ V } t_1]]]]$					*	***

⁷(8-e) liefert immerhin die richtige Kette. Daß dies aber keine mögliche Struktur ist, wird zusätzlich zur eben vorgestellten Argumentation bzgl. KASUS auch die Argumentation im nächsten Abschnitt zeigen. Dort wird es wesentlich sein, daß *do* zur Erfüllung von OP-SPEZ und OB-KOPF bewegt wird.

⁸Wie üblich wären noch mehr Kandidaten zu betrachten, aber keiner davon kann ein besseres Beschränkungsprofil haben als der optimale Kandidat.

Diese Argumentation beruht darauf, daß zur Erfüllung von OB-KOPF (und letztlich von OP-SPEZ) in Fragesätzen ein Kopf nach C bewegt werden muß.⁹ Dieses Element kann wegen LEX-ÖKON kein Hauptverb sein, und wenn aufgrund von Inputspezifikationen nicht anderweitig ein Auxiliar vorhanden ist, das ohne Verletzung von LEX-ÖKON von I nach C bewegt werden kann, muß *do* in I eingesetzt und dann bewegt werden. Das bedeutet aber, daß bei Präsenz eines Auxiliars (Modalverb oder Temporalauxiliar) *do*-Einsetzung blockiert ist. Dies zeigen die folgenden Beispiele:

- (11) a. What will she say ?
 b. *What will she do say ?
 c. *What does she will say ?

Aus Tabelle T_{3.1} ist ja ersichtlich, daß ein Kandidat wie (11-a) OP-SPEZ und OB-KOPF erfüllt und lediglich ÖKON verletzt. Da in (11-a) genauso wenig LEX-ÖKON, KASUS und VOLL-INT verletzt werden, ist die in (11-bc) unvermeidbare VOLL-INT-Verletzung automatisch fatal. Aus demselben Grund ist auch mehrfache *do*-Einsetzung blockiert: Hinzufügen von weiteren *dos* führt zu keiner Verbesserung des Beschränkungsprofils, es kostet nur weitere VOLL-INT-Verletzungen:

- (12) a. What did she say ?
 b. *What did she do say ?

Bislang haben wir nur W-Phrasen betrachtet, die als Objekte in Komplementpositionen erzeugt werden und sich daher zur Erfüllung von OP-SPEZ bewegen müssen. W-Subjekte sind nun bereits an sich in einer Spezifikatorposition: entweder in SpecV oder, wenn ein Auxiliar vorliegt, in SpecI (die Anhebung von SpecV nach SpecI im letzteren Fall erfolgt, um KASUS und SUBJEKT zu erfüllen). Die Vorhersage ist damit, daß ein W-Subjekt sich nicht mehr extra bewegen muß, um OP-SPEZ zu erfüllen, wenn die SpecV- bzw. SpecI-Position, die es einnimmt, bereits die höchste gefüllte Position im Fragesatz ist (hier unterscheidet sich OP-SPEZ wesentlich von W-KRIT, das auch W-Subjekte in eine designierte Position – SpecC – zwingt). Die durch *do*-Einsetzung hervorgerufene Verletzung der tief geordneten Beschränkung VOLL-INT sollte in diesem Fall also fatal sein. Die Beispiele in (13) zeigen, daß diese Vorhersage korrekt ist: Nicht-emphatisches *do* ist hier unmöglich.¹⁰

- (13) a. [_{VP} Who₁ left]?
 b. * [_{IP} Who₁ did₂ [_{VP} t₁ leave]] ?

⁹Warum kann nicht einfach im Matrixfragesatz ein Komplementierer *that* eingesetzt werden, um OB-KOPF zu erfüllen? Die Antwort wird weiter unten gegeben, bei der Betrachtung von Komplementierern.

¹⁰Eine notwendige Annahme ist dabei, daß SpecV in (13-a) gleichzeitig A-Position und A'-Position sein kann.

c. * $[_{CP} \text{Who}_1 \text{did}_2 [_{IP} t'_1 t_2 [_{VP} t_1 \text{leave}]]]$?

Den Wettbewerb illustriert Tabelle T_{3.6}. Der Übersichtlichkeit halber sind hier nur die Kandidaten in (13) betrachtet; alternative Kandidaten, die z.B. W-Bewegung nach SpecI oder SpecC ohne *do*-Einsetzung vornehmen, oder die anstelle von *do*-Einsetzung das Hauptverb bewegen, haben ein noch schlechteres Beschränkungsprofil als K₂ und K₃ – tatsächlich verletzt ja K₁ überhaupt keine der hier betrachteten Beschränkungen.

T_{3.6}: *W-Subjekt-Fragen*

Kandidaten	OP-SPEZ	LEX-ÖKON	KA-SUS	OB-KOPF	VOLL-INT	Ö-KON
☞ K ₁ : $[_{VP} W_1 \text{left}]$						
K ₂ : $[_{IP} W_1 \text{do}_2 [_{VP} t_1 \text{leave}]]$					*!	*
K ₃ : $[_{CP} W_1 \text{do}_2 [_{IP} t'_1 t_2 [_{VP} t_1 \text{leave}]]]$					*!	***

Eine wichtige Eigenschaft dieser Analyse ist, daß die Frage, ob *do*-Einsetzung bei Fragesätzen in einer Sprache vorkommt, nicht von morphologischen bzw. lexikalischen Gegebenheiten abhängt, sondern einzig und allein von einer bestimmten Beschränkungsordnung. Dies ist, wie schon in Kapitel 2 im Zusammenhang mit *do* in Negationskontexten gesagt wurde, eine gute Illustration des von Legendre, Smolensky & Wilson (1998) formulierten Slogans, daß das funktionale Lexikon Sklave der Syntax ist. Man kann also nicht z.B. sagen, daß das Russische keine *do*-Einsetzung hat (vgl. (14-a)), weil dieser Sprache ein geeignetes lexikalisches V-Expletiv fehlt. Vielmehr muß das Fehlen von *do*-Einsetzung (sowie von Hauptverbbewegung in diesem Kontext) eine Konsequenz einer anderen Beschränkungsordnung sein. Dieselbe Konsequenz ergibt sich unter dieser Perspektive auch für das Deutsche. Daß in Matrixfragesätzen keine *do*-Einsetzung erfolgt, sondern durchweg Verbbewegung (vgl. (14-b)), kann nicht am Fehlen eines geeigneten Expletivs liegen, sondern muß auf eine andere Beschränkungsordnung zurückgeführt werden.¹¹

- (14) a. $[_{CP} \check{C}t_{O1} - [_{VP} \text{on delaet } t_1]]$?
 was er macht
 b. $[_{CP} \text{Was}_1 \text{macht}_2 [_{VP} \text{er } t_1 t_2]]$?

¹¹ Abgesehen ist hier von der Möglichkeit der *tun*-Konstruktion in Substandard-Varietäten.

- (i) $[_{CP} \text{Was}_1 \text{tut}_2 [_{IP} \text{er } [_{VP} t t_1 \text{machen}] t_2]]$?

Die Existenz von (i) ist unter Umständen ein Problem für Grimshaws Analyse, da die Konstruktion nicht den Status eines Reparaturphänomens zu haben scheint, sondern in den Varietäten, wo sie überhaupt möglich ist, immer mit der Variante ohne *tun* koexistiert – und zwar nicht nur in speziellen Kontexten (Negation, Fragesatz, Emphase), sondern auch in normalen Deklarativsätzen.

Die Ordnung, die in englischen Matrixfragesätzen (mit W-Objekt und ohne Auxiliar) für *do*-Einsetzung verantwortlich ist, ist die in (15-a): OP-SPEZ erzwingt die Präsenz einer erweiterten Projektion CP, und deren Kopf darf dann wegen LEX-ÖKON und OB-KOPF weder durch das Hauptverb gefüllt, noch leer gelassen werden, sondern wird unter Verletzung von VOLL-INT durch *do* gefüllt. In einer minimal variierenden Beschränkungsordnung, in der LEX-ÖKON und VOLL-INT über OB-KOPF geordnet sind, wird der Kandidat optimal, der bei W-Bewegung den C-Kopf leer läßt (vgl. (15-b)); und in einer wiederum nur minimal abweichenden Beschränkungsordnung, in der OB-KOPF und VOLL-INT über LEX-ÖKON geordnet sind, ergibt sich Optimalität für den Kandidaten, der bei W-Bewegung im Matrixsatz das Hauptverb bewegt (vgl. (15-c)). Somit wären die Daten in (14) im Prinzip erfassbar.¹²

- (15) a. OP-SPEZ \gg LEX-ÖKON, OB-KOPF \gg VOLL-INT
 b. OP-SPEZ \gg LEX-ÖKON, VOLL-INT \gg OB-KOPF
 c. OP-SPEZ \gg OB-KOPF, VOLL-INT \gg LEX-ÖKON

Man vergleiche Tabelle T_{3.5}, in der gemäß Ordnung (15-a) *do*-Einsetzung gewinnt, mit den Tabellen T_{3.7} und T_{3.8}, in denen gemäß (15-bc) ein leerer C-Kopf bzw. Hauptverbanhebung optimal sind.

T_{3.7}: Matrixfragesätze und leere C-Köpfe

Kandidaten	OP-SPEZ	LEX-ÖKON	KA-SUS	VOLL-INT	OB-KOPF	Ö-KON
K ₁ : [VP NP V W ₁]	*!					
☞K ₂ : [CP W ₁ e [VP NP V t ₁]]					*	*
K ₃ : [CP W ₁ e [IP NP do ₂ [VP V t ₁]]]				*!	*	**
K ₄ : [CP W ₁ V ₃ [VP NP t ₃ t ₁]]		*!				**
K ₅ : [CP W ₁ do ₂ [VP NP V t ₁]]]			*!	*		*
K ₆ : [CP W ₁ do ₂ [IP NP t ₂ [VP t V t ₁]]]				*!		***

Zum Abschluß dieses Abschnitts sei noch auf eine interessante Konsequenz von Grimshaws Analyse hingewiesen: Bei Negationskontexten wie bei Fragekontexten erfolgt im Englischen *do*-Einsetzung unter Verletzung von tief geordnetem VOLL-INT. Der Grund für die Optimalität von Kandidaten mit dieser Verletzung ist jedoch nicht in beiden Fällen der gleiche: In Negationskontexten zwingen die hoch

¹²Eine genauere Analyse macht aber schnell klar, daß sowohl für das Russische, als auch für das Deutsche die Sachlage noch komplizierter ist. Um etwa systematisch alle Verbbewegung in russischen Matrixfragesätzen auszuschließen (d.h., auch von Auxiliaren), müßte noch zusätzlich ÖKON über OB-KOPF geordnet sein. Die gegenwärtigen Überlegungen dienen also in erster Linie der Illustration, wie englische *do*-Einsetzung in einem nur geringfügig unterschiedenen System "abgeschaltet" werden kann, und sind nicht als ernstgemeinte Vorschläge für Beschränkungsordnungen des Russischen oder Deutschen zu betrachten.

T_{3,8}: Matrixfragesätze und Hauptverbbewegung

Kandidaten	OP-SPEZ	VOLL-INT	KA-SUS	OB-KOPF	LEX-ÖKON	Ö-KON
K ₁ : [_{VP} NP V W ₁]	*!					
K ₂ : [_{CP} W ₁ e [_{VP} NP V t ₁]]				*!		*
K ₃ : [_{CP} W ₁ e [_{IP} NP do ₂ [_{VP} V t ₁]]]		*!		*		**
☞ K ₄ : [_{CP} W ₁ V ₃ [_{VP} NP t ₃ t ₁]]					*	**
K ₅ : [_{CP} W ₁ do ₂ [_{VP} NP V t ₁]]]		*!	*			*
K ₆ : [_{CP} W ₁ do ₂ [_{IP} NP t ₂ [_{VP} t V t ₁]]]]		*!				***

geordneten Beschränkungen KASUS und SUBJEKT eine Subjekt-NP in den Spezifikator einer funktionalen Projektion, deren Kopf dann durch *do*-Einsetzung gefüllt werden muß (aufgrund von KASUS und OB-KOPF). In Matrixfragesatzkontexten ist es dagegen die hoch geordnete Beschränkung OP-SPEZ, die eine W-Phrase in den Spezifikator einer funktionalen Projektion zwingt, und der Kopf muß dann allein wegen OB-KOPF durch *do*-Einsetzung gefüllt werden. Damit ist *do*-Einsetzung als inhomogenes Phänomen erwiesen: Nichts spricht z.B. dagegen, daß in einer Sprache *do*-Einsetzung in Fragesätzen obligatorisch sein könnte, in negierten Deklarativsätzen dagegen verboten.

1.4. Blockierte Inversion

Bisher ist abgeleitet worden, daß Inversion von Aux oder *do* in Matrixfragesätzen bei Nicht-W-Subjekten obligatorisch ist. Jetzt ist zu erklären, warum Inversion in eingebetteten Fragesätzen unterbleiben muß; vgl. (16).

- (16) a. I wonder [_{CP} what₁ – [_{IP} she₂ will [_{VP} t₂ say t₁]]]
 b. *I wonder [_{CP} what₁ will₃ [_{IP} she₂ t₃ [_{VP} t₂ say t₁]]]]

Relevant wird nun eine Beschränkung, die schon in Kapitel 2 eine Rolle gespielt hat, nämlich PUR-EP.¹³

- (17) PUR-EP (“Purity of Extended Projection”, “Pure-EP”):
 Adjunktion an die höchste XP einer eingebetteten erweiterten Projektion und Bewegung in deren Kopf sind verboten.

PUR-EP verbietet I-nach-C-Bewegung in (16-b). Eine derartige Bewegung wird aber, wie wir gesehen haben, im Matrixfragesätzen im Englischen durch OB-KOPF erzwungen; OB-KOPF ist also in (16-a) verletzt. Grimshaws Lösung ist nun offen-

¹³Dort wurde sowohl mit einer gegenüber Grimshaw (1997) modifizierten Beschränkung gearbeitet (bei der Behandlung des W-Bewegungsparameters; vgl. (8) aus Kapitel 2)), als auch mit der Originalbeschränkung (bei der Illustration des Konzepts “Wettbewerb und Blockade” durch das Phänomen der Komplementierertilgung im Englischen; vgl. (32-a) aus Kapitel 2).

sichtlich: Im Englischen ist PUR-EP über OB-KOPF geordnet. Tatsächlich spricht nichts dagegen, PUR-EP unter den bislang eingeführten Beschränkungen als un-dominiert zu betrachten, und Grimshaw tut dies auch. Der (16) zugrundeliegende Wettbewerb ist in Tabelle T_{3.9} veranschaulicht (die Beschränkungen SUBJEKT und VOLL-INT sind von keinem der betrachteten Kandidaten verletzt und daher aus Platzgründen weggelassen).

T_{3.9}: Blockierte Aux-Inversion in eingebetteten Fragesätzen

Kandidaten	PUR-EP	OP-SPEZ	KA-SUS	OB-KOPF	Ö-KON
☞K ₁ : [_{CP} W ₁ e [_{IP} NP ₂ Aux ₃ [_{VP} t ₂ V t ₁]]]				*	**
K ₂ : [_{CP} W ₁ Aux ₃ [_{IP} NP ₂ t ₃ [_{VP} t ₂ V t ₁]]]	*!				***
K ₃ : [_{CP} W ₁ Aux ₃ [_{VP} NP ₂ V t ₁]]]			*!		*
K ₄ : [_{IP} NP ₂ Aux ₃ [_{VP} t ₂ V W ₁]]]		*!			*

Man beachte, daß in K₃ *will* in C basisgeneriert ist; daher wird PUR-EP hier nicht verletzt. Was verletzt wird, ist dagegen KASUS, und zwar bei NP₂. Wichtig ist hier also eine Ordnung KASUS ≫ OB-KOPF.¹⁴

Diese Argumentation überträgt sich unmittelbar auf diejenigen Kontexte, die im Matrixfragesatz *do*-Einsetzung verlangen. Im eingebetteten Fragesatz ist *do*-Einsetzung blockiert, weil die OB-KOPF-Verletzung, die *do*-Einsetzung im Matrixfragesatz unter Inkaufnahme einer VOLL-INT-Verletzung verhindert, im eingebetteten Fragesatz aufgrund von höher geordnetem PUR-EP ohnehin unvermeidbar ist; vgl. (18) und Tabelle T_{3.10} (diesmal ist lediglich SUBJEKT weggelassen).

- (18) a. I wonder what she said
 b. *I wonder what did she say
 c. *I wonder what she did say
 d. *I wonder she did say what

T_{3.10}: Blockierte 'do'-Einsetzung in eingebetteten Fragesätzen

Kandidaten	PUR-EP	OP-SPEZ	KA-SUS	OB-KOPF	VOLL-INT	Ö-KON
☞K ₁ : [_{CP} W ₁ e [_{VP} NP ₂ V t ₁]]]				*		*
K ₂ : [_{CP} W ₁ do ₃ [_{IP} NP ₂ t ₃ [_{VP} t ₂ V t ₁]]]	*!				*	***
K ₃ : [_{CP} W ₁ do ₃ [_{VP} NP ₂ V t ₁]]]			*!		*	*
K ₄ : [_{CP} W ₁ e [_{IP} NP ₂ do ₃ [_{VP} t ₂ V t ₁]]]				*	*!	**
K ₅ : [_{IP} NP ₂ do ₃ [_{VP} t ₂ V W ₁]]]		*!			*	*

¹⁴Auf Seite 375 von Grimshaw (1997) (im Text über Abbildung 1, nicht in Abbildung 1 selbst) findet sich versehentlich die umgekehrte, falsche Ordnung.

1.5. Negative Operatoren

Bislang haben wir gesehen, daß Inversion (also I-nach-C-Bewegung) durch OB-KOPF und OP-SPEZ erzwungen, im eingebetteten Satz aber durch PUR-EP blockiert wird. Man beachte jedoch, daß PUR-EP lediglich Adjunktion bzw. Kopfbewegung in die höchste eingebettete erweiterte Projektion verbietet. Angenommen nun, OP-SPEZ wird durch Bewegung in einen Spezifikator eines Satzes erfüllt, der nicht der höchste der erweiterten Projektion ist. Dann verletzt Kopfbewegung in den Bereich dieses Spezifikators nicht PUR-EP, und wir erwarten, daß OB-KOPF durch Inversion erfüllt werden muß. Grimshaw argumentiert, daß sich dieses Phänomen bei der Voranstellung von negativen Operatoren im Englischen manifestiert. Wie zunächst (19) zeigt, werden negative Operatoren im Englischen in Matrixsätzen vorangestellt; dabei ist Inversion obligatorisch.

- (19) a. $[_{PP}$ Under no circumstances] will₂ she t₂ work this hard again
 b. * $[_{PP}$ Under no circumstances] she will₂ work this hard again
 c. *She will₂ work this hard again $[_{PP}$ under no circumstances]

Hier ist die Situation identisch zur W-Bewegung: OP-SPEZ erzwingt Anhebung des Neg-Operators in einen über IP liegenden Spezifikator, und OB-KOPF sorgt dann für Aux-Inversion; vgl. Tabelle T_{3.11}.¹⁵

T_{3.11}: Anhebung negativer Operatoren im Matrixsatz

Kandidaten	PUR-EP	OP-SPEZ	OB-KOPF	Ö-KON
☞ K ₁ : $[_{TP}$ Neg-Op ₁ Aux ₂ $[_{IP}$ NP ₃ t ₂ $[_{VP}$ t ₃ V t ₁]]]				***
K ₂ : $[_{TP}$ Neg-Op ₁ e $[_{IP}$ NP ₃ Aux ₂ $[_{VP}$ t ₃ V t ₁]]]			*!	**
K ₃ : $[_{IP}$ NP ₃ Aux ₂ $[_{VP}$ t ₃ V Neg-Op ₁]]		*!		*
K ₄ : $[_{TP}$ Aux ₂ $[_{IP}$ NP ₃ t ₂ $[_{VP}$ t ₃ V Neg-Op ₁]]]		*!		**
K ₅ : $[_{TP}$ e $[_{IP}$ NP ₃ Aux ₂ $[_{VP}$ t ₃ V Neg-Op ₁]]]		*!	*	*

Wie die Beispiele in (20) zeigen, ist bei negativen Operatoren Inversion auch im eingebetteten Satz erlaubt und notwendig.

- (20) a. She said $[_{CP}$ that $[_{PP}$ under no circumstances] would₂ she t₂ work this hard again]
 b. *She said $[_{CP}$ that $[_{PP}$ under no circumstances] she would₂ work this hard again]
 c. *She said $[_{CP}$ that she would₂ work this hard again $[_{PP}$ under no circumstances]]

¹⁵Warum hier die höchste Projektion statt als CP als TP notiert wird, wird bei der nächsten Tabelle erklärt. Darüber hinaus werden nur die unmittelbar relevanten Beschränkungen betrachtet. Wenn kein Auxiliar vorhanden ist, erfolgt auch hier wie oben gezeigt unter Verletzung von VOLL-INT *do*-Einsetzung.

Was die erlaubte OP-SPEZ-induzierte Inversion bei Voranstellung negativer Operatoren im eingebetteten Satz im Englischen von den oben betrachteten unmöglichen Fällen OP-SPEZ-induzierter Inversion bei Voranstellung von W-Operatoren unterscheidet, ist die Präsenz von *that* als höchstem Kopf des eingebetteten Satzes in (20). Tatsächlich ist bei Fehlen von *that* eingebettete Neg-Op-Voranstellung im Englischen blockiert:

(21) *She said [_{PP} under no circumstances] would₂ she t₂ work this hard again

Tabelle T_{3.12} zeigt, daß bei Präsenz von *that* als höchstem Kopf im eingebetteten Satz eine Kombination von durch OP-SPEZ und OB-KOPF bedingter Operatorbewegung und Kopfbewegung in eine tiefere Domäne (die zwischen der CP mit dem Komplementierer und der IP liegt und hier TP (“Topik-Phrase”) genannt wird) optimal ist, denn nunmehr wird dadurch PUR-EP respektiert.

T_{3.12}: Anhebung negativer Operatoren im eingebetteten Satz

Kandidaten	PUR-EP	OP-SPEZ	OB-KOPF	Ö-KON
☞K ₁ : [_{CP} that [_{TP} Neg-Op ₁ will ₂ [_{IP} NP ₃ t ₂ [_{VP} t ₃ V t ₁]]]]]				***
K ₂ : [_{CP} that [_{TP} Neg-Op ₁ e [_{IP} NP ₃ will ₂ [_{VP} t ₃ V t ₁]]]]]			*!	**
K ₃ : [_{CP} that [_{IP} NP ₃ will ₂ [_{VP} t ₃ V Neg-Op ₁]]]		*!		*
K ₄ : [_{CP} that [_{TP} will ₂ [_{IP} NP ₃ t ₂ [_{VP} t ₃ V Neg-Op ₁]]]]]		*!		**
K ₅ : [_{CP} that [_{TP} e [_{IP} NP ₃ will ₂ [_{VP} t ₃ V Neg-Op ₁]]]]]		*!	*	*
K ₆ : [_{TP} Neg-Op ₁ will ₂ [_{IP} NP ₃ t ₂ [_{VP} t ₃ V t ₁]]]	*!			***
K ₇ : [_{CP} e [_{TP} Neg-Op ₁ will ₂ [_{IP} NP ₃ t ₂ [_{VP} t ₃ V t ₁]]]]]			*!	***

Relevant ist der Vergleich des optimalen Kandidaten K₁ mit den suboptimalen K₆, K₇ (K₂ – K₅ sind bis auf *that* identisch zur Matrixsatz-Tabelle). In K₆ fehlt die CP-Schale über TP. Damit ist TP die höchste XP der eingebetteten erweiterten Projektion im Sinne von PUR-EP, und diese Beschränkung wird durch Inversion fatal verletzt. Die Situation ist hier analog zu dem, was über das Verbot von Verb-anhebung in eingebetteten W-Fragen gesagt wurde. In K₇ ist die CP-Schale zwar präsent, aber zur Vermeidung der PUR-EP-Verletzung ist C leer geblieben. Dies verstößt dann in fataler Weise gegen OB-KOPF; das Profil ist damit im wesentlichen wie bei K₂, nur daß noch eine weitere Bewegung (und somit ÖKON-Verletzung) hinzugekommen ist (nämlich die von *will* nach T).

Die Frage, die sich nun stellt, ist, warum dieser Ausweg aus dem PUR-EP/OB-KOPF-Dilemma nicht für W-Phrasen zur Verfügung steht – denn daß er nicht zur Verfügung steht, zeigt ja das Faktum, daß nicht nur (22-b), sondern auch (22-c) gegenüber (22-a) suboptimal ist ((22-ab) sind aus (16) wiederholt).

- (22) a. I wonder [_{CP} what₁ – she₂ will t₂ say t₁]
 b. *I wonder [_{CP} what₁ will₃ she₂ t₃ t₂ say t₁]
 c. *I wonder [_{CP} that what₁ will₃ she₂ t₃ t₂ say t₁]

Grimshaws (1997, 413) Antwort ist vielleicht nicht vollkommen befriedigend, wird aber durch die empirische Evidenz nahegelegt: W-Operatoren und andere Operatoren unterliegen doch nicht exakt denselben Bedingungen. Während W-Operatoren alle anderen Elemente im Satz c-kommandieren müssen, müssen etwa negative Operatoren nur die "minimale Proposition" c-kommandieren – und wie immer letzterer Begriff auch genau zu definieren ist, es ist klar, daß ein semantisch leerer Komplementierer *that* nicht Teil der minimalen Proposition ist. Die Formulierung von OP-SPEZ ist aus (4-a) hier wiederholt:

- (23) OP-SPEZ ("Operator im Spezifikator", "Op-Spec"):
 Syntaktische Operatoren müssen in einer Spezifikatorposition sein.
- W-Operatoren müssen alle anderen Elemente der erweiterten V-Projektion c-kommandieren.
 - Andere Operatoren (negative Quantoren z.B.) müssen nur die minimale Proposition c-kommandieren.

Somit ergibt sich neben der Betrachtung von Mehrfachfragen (vgl. Tabelle T_{3.3}) noch ein zweiter theorieinterner Grund für die strikten Forderungen an W-Operatoren, die (23-a) verkörpert. Die Erweiterung von Tabelle T_{3.9} in T_{3.13} (um K₅) veranschaulicht, daß *that*-Einsetzung bei eingebetteten W-Fragesätzen immer suboptimal sein muß.

T_{3.13}: *Blockierte Aux-Inversion in eingebetteten Fragesätzen, erweitert*

Kandidaten	PUR-EP	OP-SPEZ	KA-SUS	OB-KOPF	Ö-KON
☞ K ₁ : [CP W ₁ e [IP NP ₂ Aux ₃ [VP t ₂ V t ₁]]]				*	**
K ₂ : [CP W ₁ Aux ₃ [IP NP ₂ t ₃ [VP t ₂ V t ₁]]]	*!				***
K ₃ : [CP W ₁ Aux ₃ [VP NP ₂ V t ₁]]]			*!		*
K ₄ : [IP NP ₂ Aux ₃ [VP t ₂ V W ₁]]]		*!			*
K ₅ : [CP that [TP W ₁ Aux ₃ [IP NP ₂ t ₃ [VP t ₂ V t ₁]]]]]		*!			***

1.6. Konditionale

Im letzten Unterabschnitt haben wir eine Möglichkeit kennengelernt, wie trotz PUR-EP im eingebetteten Satz Inversion erfolgen kann: Die Zielposition für Verbbewegung ist nicht der Kopf der höchsten XP der eingebetteten erweiterten Projektion. Grundsätzlich sieht das optimalitätstheoretische System aber noch eine zweite Möglichkeit vor: Es muß lediglich eine Beschränkung geben, die noch höher als PUR-EP geordnet ist, die wie OB-KOPF Verbbewegung erzwingt, und die so eine Verletzung von PUR-EP erlaubt. Grimshaw nimmt an, daß diese Situation bei englischen Konditionalsätzen vorliegt: Im wohlgeformten Satz (24-a) erfolgt Inversion, im ungrammatischen Satz (24-b) unterbleibt sie.

- (24) a. [_{CP} Were₁ he t₁ to be asked] he would probably say no
 b. *_{[IP} He were₁ to be asked] he would probably say no

Man beachte, daß die eingebetteten (topikalisierten) Sätze keine funktionale XP über der CP aufweisen, daß also eine PUR-EP-Verletzung unvermeidlich ist. Die von Grimshaw ins Spiel gebrachte höher geordnete Beschränkung ist KOND.

- (25) KOND (“Konditionalbedingung”, “Cond”):
 Ein abhängiger Kopf *were* c-kommandiert die erweiterte Projektion, zu der er gehört.

Per Annahme ist der Kopf *were* in (24) kein unabhängiges Element; er “erfordert eine Verbindung zum Nachsatz”. Also muß *were* gemäß KOND seine erweiterte Projektion c-kommandieren, d.h., sich nach C bewegen. Diese Bewegung verletzt PUR-EP, und aufgrund der Wohlgeformtheit von (24-a) gilt daher: KOND ≫ PUR-EP. Das Ergebnis illustriert Tabelle T_{3.14}.¹⁶

T_{3.14}: *Konditionalinversion*

Kandidaten	KOND	PUR-EP	ÖKON
☞K ₁ : [_{CP} were ₁ [_{IP} he ₂ t ₁ [_{VP} t ₂ ...]]]]		*	**
K ₂ : [_{IP} he ₂ were ₁ [_{VP} t ₂ ...]]]]	*!		*
K ₃ : [_{CP} that [_{TP} were ₁ [_{IP} he ₂ t ₁ [_{VP} t ₂ ...]]]]]]	*!		**

Eine Voraussetzung für das Funktionieren dieser Analyse ist, daß das *if*-Konditional in (26) nicht im selben Wettbewerb wie (24-a) ist; denn (26-a) vermeidet ja eine PUR-EP-Verletzung.

- (26) [_{CP} If he were to be asked] he would probably say no

Zu diesem Zweck nimmt Grimshaw an, daß (24-a) und (27) nicht identische LFs haben. Dies mag ebenso wie die Redeweise von den abhängigen Köpfen in KOND nicht vollkommen unproblematisch sein. Worum es im gegenwärtigen Zusammenhang jedoch zuvorderst geht, ist die Logik der Argumentation: Normalerweise blockiert PUR-EP Inversion in die höchste Schale eines eingebetteten Satzes, aber gegen dieses Verbot kann zugunsten der Erfüllung einer höher geordneten Beschränkung verstoßen werden.

¹⁶K₃ entspricht dem, was bei eingebetteter Voranstellung negativer Operatoren optimal ist. Dieser Kandidat erfüllt zwar PUR-EP, ist aber aufgrund einer fatalen KOND-Verletzung dennoch suboptimal – der Kopf *were* c-kommandiert hier nicht seine gesamte erweiterte Projektion. Dies ist vollständig analog zur Situation bei W-Operatoren.

1.7. Komplementierer

1.7.1. Optionales vs. obligatorisches ‘that’

In Abschnitt 3.1 von Kapitel 2 sind bereits einige Merkmale von Grimshaws Analyse des englischen Komplementierers *that* behandelt worden. Zunächst einmal gilt, daß *that* in Objektsätzen häufig optional ist:

- (27) a. I think [_{IP} John will leave]
 b. I think [_{CP} that [_{IP} John will leave]]

Dies folgt, weil *that* zunächst einmal kostenfrei ist, d.h., es verletzt nicht automatisch eine Beschränkung, um die die komplementiererlose Variante herkommen würde. Insbesondere verletzt dieser Komplementierer wie oben gesehen nicht VOLL-INT – *that* trägt zwar keine Bedeutung, ist aber, so zumindest Grimshaws Annahme, kein Expletiv, das durch das Wegstreichen semantischer Merkmale aus einer bedeutungstragenden Form abgeleitet worden ist. Die Optionalität von “Komplementierertilgung” im Englischen ergibt sich aus Tabelle T_{3.15} (dies ist eine geringfügig modifizierte Variante von T_{2.12} aus Kapitel 2).

T_{3.15}: Komplementierertilgung im Englischen

Kandidaten	PUR-EP	OB-KOPF	VOLL-INT	Ö-KON
☞K ₁ : think [_{CP} that [_{IP} John ₁ will [_{VP} t ₁ leave]]]				*
☞K ₂ : think [_{IP} John ₁ will [_{VP} t ₁ leave]]				*
K ₃ : think [_{CP} e [_{IP} John ₁ will [_{VP} t ₁ leave]]]		*!		*

In manchen Kontexten kann *that* allerdings obligatorisch werden. Ein erster solcher Kontext ist wie gesehen die eingebettete Voranstellung negativer Operatoren, die ihrerseits Verbanhebung induziert; hier verletzt die *that*-lose Variante fatal PUR-EP. Ein zweiter Kontext mit obligatorischem Komplementierer im Englischen wird durch Objektsätze gebildet, die durch Nicht-Brückenverben eingebettet sind; und ein dritter durch Subjektsätze (vgl. (28-ab), (28-cd) respektive). Zu letzteren beiden Fällen sagt Grimshaw nichts.

- (28) a. *He resented [_{IP} John did it]
 b. He resented [_{CP} that John did it]
 c. * [_{IP} He left so early] shows that he was tired
 d. [_{CP} That he left so early] shows that he was tired

Ein vierter Kontext ist die eingebettete Topikalisierung wie in (29-ab), und ein fünfter Adjunktion im eingebetteten Satz wie in (29-cd).

- (29) a. *I think [_{IP} to John₁ Mary gave a book t₁]
 b. I think [_{CP} that to John₁ Mary gave a book]
 c. *I think [_{IP} most of the time [_{IP} John could accept this solution]]

d. I think [_{CP} that [_{IP} most of the time [_{IP} John could accept this solution]]]

Unter der Annahme, daß in diesen beiden Fällen unvermeidbar IP-Adjunktion vorliegt (d.h., daß der Unterschied zwischen beiden Konstruktion lediglich einer zwischen durch Bewegung erzeugter und basisgenerierter Adjunktion ist), ergibt sich die Obligatorizität des Komplementierers *that* auf im Prinzip die gleiche Weise wie bei der Voranstellung negativer Operatoren; nur ist es diesmal nicht die Verbinversion, die eine lexikalisch gefüllte Schale mit *that* verlangt, sondern die Adjunktion an die IP – denn dieser Prozeß verletzt PUR-EP, falls IP die höchste Projektion des eingebetteten Satzes ist. Dies ist in Tabelle T_{3.16} gezeigt, die im wesentlichen Tabelle T_{2.13} aus Kapitel 2 entspricht.

T_{3.16}: Komplementierertilgung und Satzadjunkte

Kandidaten	PUR-EP	OB-KOPF	VOLL-INT	Ö-KON
☞K ₁ : V [_{CP} that [_{IP} XP [_{IP} NP ₁ I t ₁ V NP]]]				*
K ₂ : V [_{IP} XP [_{IP} NP ₁ I t ₁ V NP]]	*!			*
K ₃ : V [_{CP} e [_{IP} XP [_{IP} NP ₁ I V t ₁ NP]]]		*!		*

Ein sechster Kontext für obligatorisches *that* ist schließlich die Subjekt-Relativierung im Englischen. All diesen Phänomenen gegenüber steht schließlich der Komplementierer-Spur-Effekt, der die Optionalität von CP- und IP-Einbettung in die andere Richtung auflöst und obligatorisch Komplementierertilgung verlangt. Die letzteren beiden Punkte werden im folgenden Abschnitt abgehandelt.

1.7.2. Bewegung

Der Komplementierer-Spur-Effekt im Englischen, d.h., die Unmöglichkeit der Subjektbewegung über einen benachbarten Komplementierer *that* (vgl. Chomsky & Lasnik (1977)) wird abgeleitet in Anlehnung an Déprez (1991). Die Grundidee ist, daß Subjektbewegung über Komplementierer gerade durch die Möglichkeit als suboptimal blockiert wird, daß man auch eine Variante ohne Komplementierer haben kann. An neuen Beschränkungen kommen S-REK und S-LEX-REK ins Spiel.

- (30) a. S-REK (“Spur-Rektion”, “T-Gov”):
Eine Spur ist regiert.
- b. S-LEX-REK (“Lexikalische Spur-Rektion”, “T-Lex-Gov”):
Eine Spur ist lexikalisch regiert.

S-REK und S-LEX-REK konfigrieren nicht. Sie stehen vielmehr als speziell zu generell in einer Stringenzrelation, denn jede S-REK-Verletzung bei einem Kandidaten impliziert auch eine S-LEX-REK-Verletzung: Jede Spur, die nicht regiert ist, ist ja naturgemäß auch nicht lexikalisch regiert (umgekehrt gilt das nicht notwendigerweise). Für in Stringenzrelation befindliche Beschränkungen gelten wie in Kapitel 2 bemerkt nicht per se fixe Ordnungen, um Aktivität zu erreichen; tatsächlich ist für

die hier zu betrachtenden Fälle die relative Ordnung der beiden Beschränkungen unwesentlich. Wichtig ist, daß S-REK und S-LEX-REK unter OP-SPEZ geordnet sind; Grimshaw nimmt an, daß diese Beschränkungen von allen bisher betrachteten Beschränkungen dominiert werden.

Zunächst sei der Fall der Objektbewegung über *that* betrachtet. Hier ist der Komplementierer optional:

- (31) a. Who₁ do you think [_{CP} that she likes t₁] ?
 b. Who₁ do you think [_{IP} she likes t₁] ?

Dies folgt wie in Tabelle T_{3.17} angegeben: Beide Extraktionen, in K₁ und K₂, verletzen keine der hier relevanten Beschränkungen; Objektspuren sind (von V) sowohl regiert, als auch lexikalisch regiert. W-in-situ-Strategien wie in K₃ und K₄ verletzen dagegen fatal OP-SPEZ.

T_{3.17}: Objektbewegung

Kandidaten	OP-SPEZ	S-REK	S-LEX-REK
☞K ₁ : who ₁ ... V [_{CP} that [_{VP} she likes t ₁]]			
☞K ₂ : who ₁ ... V [_{VP} she likes t ₁]			
K ₃ : ... V [_{CP} that [_{VP} she likes who ₁]]	*!		
K ₄ : ... V [_{VP} she likes who ₁]	*!		

Ebenso liegt bei Adjunktbewegung Optionalität von *that* vor:

- (32) a. When₁ do you think [_{CP} that they will see them t₁] ?
 b. When₁ do you think [_{IP} they will see them t₁] ?

Hier ist die Erklärung etwas anders: Denn Adjunktspuren sind per Annahme weder lexikalisch regiert, noch überhaupt von irgendeiner Kategorie regiert. Daher verletzen (32-a) und (32-b) (= K₁ und K₂ in Tabelle T_{3.18}) gleichermaßen S-REK und S-LEX-REK. Da diese Beschränkungen jedoch unter OP-SPEZ geordnet sind, ist auch hier eine in-situ-Strategie zum Scheitern verurteilt, selbst wenn die entsprechenden Kandidaten K₃ und K₄ S-REK und S-LEX-REK erfüllen, weil sie keine W-Spuren aufweisen.

T_{3.18}: Adjunktbewegung

Kandidaten	OP-SPEZ	S-REK	S-LEX-REK
☞K ₁ : when ₁ ... V [_{CP} that [_{IP} they will see them t ₁]]		*	*
☞K ₂ : when ₁ ... V [_{IP} they will see them t ₁]		*	*
K ₃ : ... V [_{CP} that [_{IP} they will see them when ₁]]	*!		
K ₄ : ... V [_{IP} they will see them when ₁]	*!		

Der interessante Fall ist nun die Subjektbewegung. Hier ist *that* blockiert.

- (33) a. *Who₁ do you think [_{CP} that t₁ will see them] ?

b. Who₁ do you think [_{IP} t₁ will see them] ?

Der Grund dafür ist, daß in diesem Kontext das identische Verhalten von t₁ bzgl. S-REK und S-LEX-REK aufhört: Beidesmal ist t₁ regiert (durch *that* bzw. das Matrixverb). Aber nur in (34-b) ist t₁ auch lexikalisch regiert (durch das Matrixverb); in (34-a) kann keine lexikalische Rektion von t₁ erfolgen, weil der Komplementierer *that* keine lexikalische Kategorie ist.¹⁷

T_{3.19}: *Subjektbewegung*

Kandidaten	OP-SPEZ	S-REK	S-LEX-REK
K ₁ : who ₁ ... V [_{CP} that [_{IP} t ₁ will see them]]			*!
☞ K ₂ : who ₁ ... V [_{IP} t ₁ will see them]			
K ₃ : ... V [_{CP} that [_{IP} who ₁ will see them]]	*!		
K ₄ : ... V [_{IP} who ₁ will see them]	*!		

Somit ist der Fall abgehandelt, wo auf den Komplementierer *that* zugunsten von IP-Einbettung (also “Komplementiererertilgung”) verzichtet werden muß. Es steht noch aus die Beschreibung des sechsten oben erwähnten Kontexts, in dem *that* im Englischen obligatorisch wird, nämlich die Subjekt-Relativierung im Englischen. Wie (34) zeigt, gibt es hier eine Art “Anti-Komplementierer-Spur-Effekt”:

- (34) a. The people [_{CP} that t₁ will see them]
 b. *The people [_{IP} t₁ will see them]

Die Strukturen hier sind vielleicht etwas ungewöhnlich, denn Grimshaw nimmt keinen leeren Operator an, der die Subjektspuren bindet; das Antezedens ist unmittelbar das Matrix-N *people*. Mit diesen Strukturen ergibt sich die Obligatorizität von *that* aber ohne weitere Stipulationen. Die wesentliche Beobachtung ist, daß Relativsätze keine Argumente sind; sie sind syntaktisch in einer Adjunktionsposition. Daher ist von außen überhaupt keine Rektion von t₁ möglich, weder in (34-a) noch in (34-b) (vgl. Chomsky (1986a)). Rektion durch *that* ist somit noch das Beste, was für eine Subjektspur in einem Relativsatz zu kriegen ist, und daher ist diese Variante optimal. Vgl. Tabelle T_{3.20}.

1.7.3. Ein Problem: Doppel-Comp-Filter-Verletzungen

Eine Komplikation ist noch zu erwähnen, die bereits in Fußnote 9 angeklungen ist. T_{3.9} hat gezeigt, daß der optimale Kandidat bei eingebetteten W-Fragen einer ist,

¹⁷Diese Argumentation präsupponiert einen Rektionsbegriff, demzufolge ein Kopf zwar sein Komplement und den Spezifikator seines Komplementes regieren kann, aber nicht seinen Spezifikator. Ansonsten würde ja bei VP-internen Subjekten (in englischen Sätzen ohne Auxiliar oder *do*) V SpecV lexikalisch regieren, und S-LEX-REK wäre fälschlicherweise durch das eingebettete Verb erfüllt, unabhängig von der Präsenz oder dem Fehlen eines Komplementierers.

T_{3,20}: Subjektrelativsätze

Kandidaten	S-REK	S-LEX-REK
☞ K ₁ : people [_{CP} that [_{IP} t ₁ will see them]]		*
K ₂ : people [_{IP} t ₁ will see them]	*	*

bei dem aufgrund der Ordnung PUR-EP \gg OB-KOPF Inversion unterbleiben muß; vgl. (35) (= (16)):

- (35) a. I wonder [_{CP} what₁ – [_{IP} she₂ will [_{VP} t₂ say t₁]]]
 b. *I wonder [_{CP} what₁ will₃ [_{IP} she₂ t₃ [_{VP} t₂ say t₁]]]

Der optimale Kandidat (35-a), so war die Konklusion, verletzt OB-KOPF, um PUR-EP zu erfüllen. Es gibt nun einen gefährlichen Konkurrenten, der hier in C statt des Verbs einen Komplementierer *that* einfügt und so eine OB-KOPF-Verletzung vermeidet:

- (36) *I wonder [_{CP} what₁ that [_{IP} she₂ will [_{VP} t₂ say t₁]]]

Der Kandidat (35-a) involviert eine Verletzung von OB-KOPF, und zunächst einmal sieht es so aus, als würde (36) diese Verletzung ohne Mehrkosten vermeiden. Es gilt also, eine Beschränkung zu finden, die (36) verletzt, und die OB-KOPF dominiert, so daß (35-a) immer noch optimal ist. Verschiedene Möglichkeiten bieten sich an. Eine Beschränkung, derzufolge Komplementierer phonologisch am linken Rand von CP stehen müssen (also bei Präsenz eines Komplementierers in SpecC nur Spuren oder phonetisch nicht realisierte XPs erlaubt sind), würde diese Aufgabe erfüllen, falls sie höher geordnet ist als OB-KOPF.¹⁸

Grimshaw geht jedoch einen anderen, etwas komplizierteren Weg: Die Beschränkungen KOPF-LINKS, KOPF-RECHTS werden notwendig.

- (37) a. KOPF-LINKS (“Hd-Lft”):
 Ein Kopf (X⁰ oder X′) steht am linken Rand seiner Projektion.
 b. KOPF-RECHTS (“Hd-Rt”):
 Ein Kopf (X⁰ oder X′) steht am rechten Rand seiner Projektion.

Die vorgeschlagene Ordnung für das Englische ist KOPF-LINKS \gg KOPF-RECHTS (dies wird bei Spezifikatoren im Englischen durch eine weitere Beschränkung SPEZ-LINKS überlagert), woraus folgt, daß im Englischen Köpfe links von ihren Komplementen stehen. Wichtig im gegenwärtigen Zusammenhang ist noch die Ordnung

¹⁸Wie wir im nächsten Abschnitt sehen werden, schlägt Pesetsky (1997; 1998) eine Beschränkung vor (vgl. sein LR(CP)), die in diese Richtung geht. Allerdings entspricht LR(CP) nicht exakt der im Text ins Auge gefaßten Beschränkung: Bei Pesetsky würden (35-a) und (36) gleichermaßen LR(CP) verletzen, weil LR(CP) nicht sagt: “Ein Komplementierer steht am linken Rand der CP”, sondern: “Eine CP hat am linken Rand einen Komplementierer”.

KOPF-LINKS \gg OB-KOPF. Dies bedeutet: Bevor *that* in nicht-links-peripherer Position wie in (36) auftritt, ist es besser, C erscheint überhaupt nicht, wie in (35-a). Hiermit ist gewährleistet, daß Komplementierereinsatz kein Weg sein kann, um einen durch OP-SPEZ und OB-KOPF geforderten Kopf zu füllen (und zwar weder im eingebetteten Satz noch im Matrixsatz).

Diese Analyse erzwingt jedoch einige Zusatzannahmen, die man als weniger schön ansehen mag. Beispielsweise darf I-nach-C-Bewegung im Hauptsatz nicht KOPF-LINKS verletzen, also muß noch weiter zwischen “perfekten” Köpfen (wie *that*) und “bewegten” Köpfen (wie V/2) unterschieden werden. Darüber hinaus folgt, daß leere Köpfe (wie in (35-a)) entweder nicht perfekt sein müssen (so daß sie KOPF-LINKS nicht verletzen), oder daß sie überhaupt nicht in der Struktur da sind, und C' IP ohne Verzweigung in eingebetteten W-Fragen dominiert. Schließlich merkt Grimshaw selbst an (p. 411), daß die Analyse nicht kompatibel ist mit der freien Alternation von CPs und IPs/VPs, wie in T_{3.15} dargestellt. Der Grund ist, daß die CP-Variante immer eine zusätzliche links-köpfige XP involviert, und somit automatisch einmal mehr KOPF-RECHTS verletzt als die konkurrierende IP/VP-Variante. Diese Verletzung, obschon sie eine sehr niedrig geordnete Beschränkung betrifft, sollte dann fatal sein: *That*-Einsetzung ist nicht mehr vollkommen kostenlos. Diese Überlegungen könnte man zum Anlaß nehmen, Doppel-Comp-Filter-Konstruktionen nicht durch eine Beschränkung auszuschließen, die ganz allgemein nicht-links-periphere Köpfe verbietet, sondern durch eine über OB-KOPF geordnete Beschränkung, die speziell auf Komplementierer Bezug nimmt, wie oben erwogen.

1.8. Einordnung

An dieser Stelle ist es vielleicht nützlich, einige wesentliche Eigenschaften der Grimshawschen Analyse Revue passieren zu lassen und im Hinblick auf allgemeine Merkmale von optimalitätstheoretischen Analysen zu überprüfen.

1.8.1. Kandidatenmengen

Ein erster Punkt betrifft die Größe von Kandidatenmengen. Wie gesehen sind solche Kandidaten miteinander im Wettbewerb, die Realisierungen derselben Prädikat-/Argument-Struktur sind, und die darüber hinaus auch noch eine identische LF besitzen. Die auf dieser Basis zustandekommenden Kandidatenmengen sind notwendigerweise nicht-finit. Dies kann man gut an zwei generellen Eigenschaften des Systems sehen. Die erste Eigenschaft ist die variable Projektion: Es ist ja ein Ziel, optimalitätstheoretisch abzuleiten, daß erweiterte Projektionen zwar an sich beliebige Größe haben können, daß die optimalen Projektionen (u.a. wegen OB-KOPF) aber immer den kleinstmöglichen Umfang haben. Dies heißt aber, daß zu jeder erweiterten Projektion XP auch noch eine Projektion [_{YP} XP] im Wettbewerb ist; eine obere Grenze gibt es nicht. Eine zweite Eigenschaft, die zu nicht-finiten Kandidatenmengen führt, ist die freie Adjunktion: Prinzipiell erlaubt das System beliebig iterierbare Adjunktion an eine XP. Dies wird optimalitätstheoretisch zwar

ab einem bestimmten Punkt immer durch ÖKON ausgeschlossen; entscheidend ist aber, daß dieser Ausschluß durch den Wettbewerb erreicht wird, nicht durch Gen.

1.8.2. Typen von Evidenz

Ein weiterer Punkt betrifft die Frage, wie die in Kapitel 2 behandelten Typen von Evidenz, die eine Optimalitätstheoretische Analyse stützen, sich in Grimshaws Ansatz wiederfinden: (i) parametrische Variation, (ii) Reparatur, (iii) Wettbewerb und Blockade, (iv) das Konzept des unmarkierten Falles sowie (v) inhärente Beschränkungskonflikte.

Parametrische Variation Grimshaw (1997, 404) betrachtet die faktorielle Typologie, die sich aus den möglichen Ordnungen der Beschränkungen OP-SPEZ, OB-KOPF und ÖKON ergibt. Ableitbar sind daraus neben Sprachen des englischen Typs solche Sprachen, die überhaupt keine W-Bewegung haben (wobei wie in Kapitel 2 erwähnt erwogen wird, daß Sprachen wie Japanisch oder Koreanisch dieser Ordnung gar nicht bedürfen, weil dort die W-Phrasen vielleicht gar keine Operatoren im Sinne von OP-SPEZ sind), Sprachen mit W-Bewegung, aber ohne Inversion (vgl. die Bemerkungen zum Russischen oben), sowie Sprachen, die in Matrix- wie eingebetteten Sätzen W-Bewegung plus Inversion haben (ein Beispiel hierfür wird nicht gegeben). Darüber hinaus erlaubt das System wie gesehen eine einfache "Abschaltung" von *do*-Einsetzung durch Beschränkungsumordnung. Insgesamt kann man nichtsdestoweniger sagen, daß parametrische Variation kein Schwerpunkt dieser Analyse ist.

Reparatur Dies ist dagegen ein zentraler Punkt. Zum einen ist natürlich *do*-Einsetzung ein Kardinalbeispiel für ein syntaktisches Reparaturphänomen; zum anderen läßt sich, da das Englische keine generelle V/2-Sprache ist, die durch OP-SPEZ und OB-KOPF bedingte Inversion als Instanz Reparatur-getriebener Bewegung auffassen: Etwas, das "normalerweise" nicht möglich ist, wird in bestimmten Kontexten möglich, um höher geordnete Beschränkungen zu erfüllen.

Wettbewerb und Blockade Dieser Fall liegt vor bei der Analyse der Distribution von *that*. Normalerweise ergibt sich hier freie Alternation zweier konkurrierender Kandidaten (mit und ohne *that*), aber unter bestimmten Umständen blockiert die eine Variante die andere Variante als suboptimal.

Das Konzept des unmarkierten Falles ÖKON ist tief geordnet, was zur Folge hat, daß im unmarkierten Fall Bewegung nicht erfolgt.

Beschränkungskonflikt Auch diese Evidenz findet sich in der vorliegenden Analyse. Grimshaw argumentiert, daß OP-SPEZ und OB-KOPF nur dann maximal

generell und elegant formuliert werden können, wenn man die beiden Beschränkungen verletzbar macht; vgl. die Diskussion von Mehrfachfragen im Englischen. Standardanalysen müssen die eine Bedingung als Ausnahmeregelung in die andere hineinschreiben. Allerdings ist hier einschränkend festzuhalten, daß OP-SPEZ bei genauerem Hinsehen doch nicht so elegant und einfach formuliert ist, wie man das gerne hätte (vgl. die Fallunterscheidung in (4-a)); und für OB-KOPF gilt (zumindest bei Annahme von KOPF-LINKS, KOPF-RECHTS) letztlich wohl dasselbe.

1.8.3. Ein empirisches Problem

Versucht man, die Analyse von Grimshaw auf das Deutsche zu übertragen, stellt sich sofort ein empirisches Problem: In vielen Fällen (etwas vereinfacht: bei Einbettung durch Brückenverben) sind im deklarativen Nebensatz im Deutschen sowohl V/2-Stellungen, als auch *daß*-Einsetzungen möglich.

- (38) a. Ich glaube [_{CP} den Fritz₁ hat₂ [_{IP} keiner t₁ gesehen t₂]]
 b. Ich glaube [_{CP} daß [_{IP} den Fritz₁ keiner gesehen hat]]

Zunächst einmal ist unklar, wie man diese Alternation überhaupt ableiten kann. Da die Sätze in (38) auf dieselben Prädikat-/Argument-Strukturen zurückgehen und auch noch eine identische Bedeutung zu haben scheinen, sieht es so aus, als müßten sie im selben Wettbewerb sein. Dann stellt sich aber die Frage, wie sie dasselbe Beschränkungsprofil haben können, da ja (38-a) gegenüber (38-b) zwei Bewegungen (und somit zwei ÖKON-Verletzungen) mehr aufweist.¹⁹

Ganz abgesehen vom Wettbewerb mit (38-b) ist aber (38-a) auch für sich genommen schon ein Problem. V/2-Bewegung in eingebetteten Deklarativsätzen im Deutschen verletzt PUR-EP; denn anders als z.B. bei der Voranstellung negativer Operatoren oder bei der Topikalisierung im Englischen kann ja hier keine Schale mit einem Komplementierer *daß* vorliegen, die V/2-Bewegung vor einer PUR-EP-Verletzung schützt. Man könnte nun meinen, daß das Problem durch Unterordnung von PUR-EP unter OB-KOPF lösbar sei. Dem ist jedoch nicht so, denn eingebettete W-Sätze blockieren im Deutschen wie im Englischen V/2, und dies wird man auf dieselbe Weise wie im Englischen ableiten wollen, als Resultat einer Ordnung PUR-EP ≫ OB-KOPF.

- (39) a. Er sagte mir [_{CP} wen – [_{IP} er getroffen hat]]

¹⁹Man kann versuchen, dieses Problem dadurch zu vermeiden, daß man argumentiert, daß (38-a) und (38-b) doch nicht dieselbe Bedeutung bzw. LF haben, weil z.B. die NP *den Fritz₁* in (38-a) als Topik interpretiert werde; vgl. etwa Vikner (1997a). Ob ein solcher Schritt der Datenlage angemessen ist, ist jedoch unklar. Denn einerseits ist oft festgestellt worden (vgl. z.B. Grewendorf (1989) und die dort zitierte Literatur), daß topikalisierte NPs im Deutschen alternativ als Topik, als Fokus, oder gar nicht in besonderer Weise (d.h., durch Rekonstruktion) interpretiert werden können; und andererseits kann *den Fritz₁* auch in (38-b) als Topik interpretiert werden, wenn entsprechende intonatorische Verhältnisse (vgl. die sog. "Hut-Kontur") gegeben sind.

b. *Er sagte mir [_{CP} wen hat [_{IP} er getroffen]]

Um deklarative V/2-Einbettungen im Deutschen abzuleiten, sind also auf jeden Fall noch weitere Annahmen notwendig.

1.8.4. Wirkung

Die Inversionsanalyse von Grimshaw (1993; 1997) hat großen Einfluß auf die optimalitätstheoretische Syntaxforschung gehabt. Drei Beispiele seien stellvertretend für viele andere erwähnt.

Baković (1995; 1998) und Vikner (1997a; 1999) entwickeln jeweils das Grimshawsche Modell zur Behandlung ähnlicher Daten aus anderen Sprachen in unterschiedliche Richtungen weiter – Baković zur Erfassung der Interaktion von Operatorenbewegungen unterschiedlichen Typs (z.B. Argumente vs. Adjunkte) und Verb-Anhebung in verschiedenen Varietäten des Spanischen, und Vikner zur Integration der Parametrisierung von V-nach-I- und I-nach-C-Bewegungen (also V/2) in den germanischen Sprachen (sowie im Französischen).

Dickey (1995) versucht, durch eine geringfügige Modifikation der im Erwachsenenenglisch bestehenden Beschränkungsordnung die Optionalität von Inversion in W-Matrixfragen im Kinderenglischen zu erklären. Er schlägt vor, daß für Englisch lernende Kinder OB-KOPF und ÖKON bis zu einem bestimmten Punkt gekoppelt, also gleich wichtig sind. Daher ist im Matrixsatz sowohl Inversion bei W-Bewegung optimal (wie in der Zielsprache), als auch W-Bewegung ohne Inversion.

2. Pesetskys (1997; 1998) Theorie der PF-Realisierung im C-System

Pesetsky stellt sich die Aufgabe, eine optimalitätstheoretische Analyse der phonologischen Realisierung von S-strukturell im C-System vorhandenem Material zu entwickeln. Diese Analyse wird dann ausgedehnt auf die PF-Realisierung in Bewegungsketten. Es wird sich zeigen, daß trotz teilweise identischen Datenbereichen und trotz beidesmal strikter Akzeptanz von OT-Grundprinzipien Pesetskys Ansatz und Grimshaws Ansatz in ihrer theoretischen Ausrichtung weiter voneinander kaum entfernt sein könnten.

2.1. Grundannahmen

Anders als Grimshaw setzt Pesetsky voraus, daß die im Wettbewerb befindlichen Kandidaten nicht verschiedene S-Struktur-Repräsentationen sind, sondern verschiedene PF-Realisierungen ein und derselben S-Struktur-Repräsentation. Der Wettbewerb ist also wie in (40) festgelegt:

(40) *Kandidatenmenge:*

Zwei Kandidaten K_i , K_j sind in derselben Kandidatenmenge gdw. gilt:
 K_i und K_j sind PF-Realisierungen derselben S-Struktur-Repräsentation.

Mit der Beschränkung, daß alle im Wettbewerb befindlichen Kandidaten PF-Varianten derselben S-Struktur-Repräsentation sind, verkleinert Pesetsky radikal den H-Eval-Bereich, in dem OT-Syntax stattfindet, zugunsten einer Stärkung von Gen. In diesem System können somit OT-Beschränkungen nichts darüber sagen, wie S-Strukturen überhaupt entstehen, und welchen nicht-phonologisch ausgerichteten Beschränkungen sie unterliegen.

Eine unmittelbare und interessante Konsequenz dieser Schmälerung der Rolle von H-Eval ist, daß anders als bei Grimshaw (1997) und der überwiegenden Zahl existierender optimalitätstheoretischer Analysen bei Pesetsky die Kandidatenmengen finit sind. Dies zeigt die folgende Überlegung: Angenommen, eine S-Struktur-Repräsentation weist genau n terminale Knoten auf. Vorausgesetzt, daß jeder terminale Knoten nur entweder PF-realisiert (ausgesprochen) werden kann oder nicht, ergibt sich eine finite Kandidatenmenge, die 2^n Elemente hat. Allerdings sind hierzu zwei Einschränkungen zu machen. Zum einen kommt man auch auf diese Art und Weise sehr rasch in Bereiche hinein, die aller Voraussicht nach vom menschlichen Gehirn nicht mehr komplett überblickt werden können. Hat z.B. ein Satz 20 terminale Knoten auf der S-Struktur, besteht die Kandidatenmenge, soweit sie nicht noch durch Zusatzrestriktionen beschränkt wird, bereits aus 1.048.576 Kandidaten. Zum anderen sieht Pesetsky (1998, 366) zumindest in einigen Fällen noch die Möglichkeit vor, daß es neben vollständiger PF-Realisierung und fehlender PF-Realisierung (“Tilgung”) noch eine dritte Variante gibt: Partielle PF-Realisierung einer XP durch ein Pronomen. Wenn somit statt zwei PF-Realisierungswerten (Ausprache: ja oder nein) drei PF-Realisierungswerte (Ausprache: ja, nein, oder partiell, als Pronomen) zur Verfügung stehen, bleibt die Kandidatenmenge zwar immer noch finit, sie wächst aber weiter ins Unüberschaubare.

2.2. Das C-System im Französischen

Der erste zu besprechende Datenkomplex involviert das C-System in französischen Deklarativ-, Relativ- sowie Fragesätzen. (41) zeigt, daß im Französischen im eingebetteten Deklarativsatz der Komplementierer *que* nicht fehlen darf.²⁰

- (41) a. Je crois [CP [C que] Pierre a faim]
 ich glaube daß Pierre hat Hunger
- b. *Je crois [CP [C –] Pierre a faim]
 ich glaube – Pierre hat Hunger

²⁰Man beachte, daß Pesetsky anders als Grimshaw davon ausgeht, daß eingebettete finite Sätze nicht in der Projektionsgröße variabel sind, sondern immer CP-Status haben (und einen Komplementierer besitzen) müssen. Eine Variante von (41-b) mit bloßer IP-Einbettung ist bei Pesetsky per Annahme nicht durch Gen erzeugbar – die einzige Option für Konstruktionen mit “Komplementierertilgung” ist daher eben das: die Nicht-Realisierung eines S-strukturell vorhandenen Komplementierers.

Im Einklang mit der Definition von Kandidaten und Kandidatenmengen sind die von Pesetsky postulierten Beschränkungen alle von der Art, daß sie phonologische Realisierung von S-Struktur-Konstituenten erzwingen oder verbieten. Zunächst sind drei Beschränkungen von Bedeutung:

- (42) a. WIED (“Wiederauffindbarkeit“, “Recoverability“, “Rec”):
Ein bedeutungstragendes syntaktisches Element muß PF-realisiert werden, es sei denn, es hat ein hinreichend lokales Antezedens.
b. LR(CP) (“Linker Rand(CP)“, “Left Edge(CP)“, “LE(CP)“):
Das erste PF-realisierte Wort in einer CP ist der Komplementierer.
c. TEL (“Telegraph“):
Funktionswörter dürfen nicht PF-realisiert werden.

WIED verbietet bestimmte Tilgungsoperationen, TEL verlangt bestimmte Tilgungsoperationen. Die Rolle von LR(CP) ist etwas komplexer: Diese Beschränkung verbietet tendentiell die Tilgung des Komplementierers, aber sie verlangt tendentiell die Tilgung von Elementen in SpecC. Die für das Französische angenommene Beschränkungsordnung ist (43):

- (43) *Ordnung im Französischen:*
WIED \gg LR(CP) \gg TEL

Für die Frage, ob im Beispielpaar (41) Komplementierertilgung erfolgt oder nicht, ist WIED noch nicht wichtig. Der Konflikt tritt hier auf zwischen der Beschränkung TEL, die Komplementierertilgung verlangt, und der Beschränkung LR(CP), die diese Tilgung verbietet, damit der Komplementierer *que* das erste PF-realisierte Wort in der eingebetteten CP ist. Wie Tabelle T_{3.21} zeigt, löst die Ordnung in (43) den Konflikt zugunsten von LR(CP) auf: Der Komplementierer muß PF-realisiert werden.²¹

T_{3.21}: *Komplementierer im Französischen*

Kandidaten	WIED	LR(CP)	TEL
☞ K ₁ : ... [CP [C que] Pierre a faim]			*
K ₂ : ... [CP [C -] Pierre a faim]		*!	

Wenn in Relativsätzen ein Objektpronomen nach SpecC bewegt worden ist, gibt es im Prinzip vier Möglichkeiten der PF-Realisierung: SpecC und C werden beide realisiert; nur SpecC wird realisiert; nur C wird realisiert; weder SpecC noch C wird realisiert. (44) zeigt, daß nur die dritte Möglichkeit grammatisch ist:

²¹Ein Kandidat mit bloßer IP-Einbettung (vgl. die letzte Fußnote) würde beide Beschränkungen trivial erfüllen. Die Annahme, daß Gen einen derartigen Kandidaten nicht erzeugen kann, ist also bei Pesetsky theorieintern unabdingbar.

- (44) a. *l'homme [CP qui₁ [C que] je connais t₁]
 der Mensch den wo ich kenne
 b. *l'homme [CP qui₁ [C -] je connais t₁]
 der Mensch den - ich kenne
 c. l'homme [CP -₁ [C que] je connais t₁]
 der Mensch - wo ich kenne
 d. *l'homme [CP -₁ [C -] je connais t₁]
 der Mensch - - ich kenne

Wenn dagegen in einem Relativsatz eine PP mit einem Relativpronomen darin bewegt wird, also sog. Rattenfang (“pied piping”) stattfindet, ist diejenige Variante optimal, die den Komplementierer tilgt und die PP in SpecC PF-realisiert:

- (45) a. *l'homme [CP [PP avec qui]₁ [C que] j'ai dansé t₁]
 der Mensch mit dem wo ich habe getanzt
 b. l'homme [CP [PP avec qui]₁ [C -] j'ai dansé t₁]
 der Mensch mit dem - ich habe getanzt
 c. *l'homme [CP [PP - -]₁ [C que] j'ai dansé t₁]
 der Mensch - - wo ich habe getanzt
 d. *l'homme [CP [PP - -]₁ [C -] j'ai dansé t₁]
 der Mensch - - - ich habe getanzt

An dieser Stelle wird die undominierte Beschränkung WIED relevant: Solange WIED nicht verletzt ist (d.h., solange keine wichtige Information verlorengeht, weil ein hinreichend lokales Antezedens bei Tilgung die Information wiederaufzufinden gestattet), wird im französischen Relativsatz aufgrund der Ordnung LR(CP) ≫ TEL der Spezifikator von C PF-getilgt; nur der Komplementierer wird PF-realisiert. Dies zeigt Tabelle T_{3.22} für den Fall der Objektbewegung: Die Tilgung von *qui* verletzt nicht WIED, weil dadurch keine Information verlorengeht, die nicht durch das Antezedens *l'homme* wiederauffindbar wäre.

T_{3.22}: Relativsätze mit Objektbewegung im Französischen

Kandidaten	WIED	LR(CP)	TEL
K ₁ : ... [CP qui ₁ [C que] je connais t ₁]		*!	*
K ₂ : ... [CP qui ₁ [C -] je connais t ₁]		*!	
☞ K ₃ : ... [CP - ₁ [C que] je connais t ₁]			*
K ₄ : ... [CP - ₁ [C -] je connais t ₁]		*!	

Wenn jedoch WIED bei Tilgung von SpecC verletzt würde, dann muß Spezifikator-Tilgung unter Verstoß gegen LR(CP) unterbleiben; und da LR(CP) dann ohnehin nicht mehr erfüllbar ist, tritt die tiefstgeordnete Beschränkung TEL in Kraft und löst ihrerseits C-Tilgung aus. Diese Situation liegt vor bei PP-Bewegung: Bei Tilgung von *avec qui* ist die Bedeutung des Elements *avec* nicht mehr auffindbar. Vgl. Tabelle T_{3.23}, insbesondere den Wettbewerb von K₃ und K₂ (der die Relevanz

von WIED zeigt) und den Wettbewerb von K_1 und K_2 (der den Einfluß von TEL verdeutlicht).

$T_{3.23}$: *Relativsätze mit Rattenfang im Französischen*

Kandidaten	WIED	LR(CP)	TEL
K_1 : ... $[_{CP} [_{PP} \text{ avec qui }]_1 [_{C} \text{ que }] \text{ j'ai dansé } t_1]$		*	*!
$\rightarrow K_2$: ... $[_{CP} [_{PP} \text{ avec qui }]_1 [_{C} -] \text{ j'ai dansé } t_1]$		*	
K_3 : ... $[_{CP} [_{PP} - -]_1 [_{C} \text{ que }] \text{ j'ai dansé } t_1]$	*!		*
K_4 : ... $[_{CP} [_{PP} - -]_1 [_{C} -] \text{ j'ai dansé } t_1]$	*!	*	

Problematisch für K_3 ist die Tilgung von *avec*, nicht die Tilgung von *qui* (die an sich wäre, wie $T_{3.22}$ zeigt, unproblematisch für WIED). Damit ergibt sich aber als gefährlicher Konkurrent für K_2 in $T_{3.23}$ (46): Hier wird nur das PP-interne Pronomen getilgt. Die Präposition bleibt erhalten, wodurch WIED respektiert wird:

(46) *l'homme $[_{CP} [_{PP} \text{ avec } -]_1 [_{C} -] \text{ j'ai dansé } t_1]$
 der Mensch mit - - ich habe getanzt

(46) hat zunächst einmal dasselbe Beschränkungsprofil wie K_2 . Die Konklusion hieraus muß sein, daß es noch eine höher geordnete Beschränkung gibt, die hier Tilgung innerhalb der PP bestraft. Wie diese Beschränkung aussieht, ist allerdings nicht klar.²²

Aus den bisher eingeführten Annahmen läßt sich die Vorhersage ableiten, daß sich eingebettete W-Fragesätze im Französischen genau wie Relativsätze mit Rattenfang verhalten, auch wenn sie selbst keinen Rattenfang aufweisen. Der Grund ist, daß ein W-Pronomen immer ein bedeutungstragendes syntaktisches Element ist, das kein lokales Antezedens aufweist, das die Information wiederaufzufinden gestattet. Diese Prognose wird durch (47) bestätigt: In eingebetteten Fragesätzen wird das W-Element in SpecC realisiert und C obligatorisch getilgt.

- (47) a. *Je me demande $[_{CP} \text{ quand}_1 [_{C} \text{ que }] \text{ Pierre arrivera } t_1]$
 ich mich frage wann daß Pierre wird ankommen
 b. Je me demande $[_{CP} \text{ quand}_1 [_{C} -] \text{ Pierre arrivera } t_1]$
 ich mich frage wann - Pierre wird ankommen
 c. *Je me demande $[_{CP} -_1 [_{C} \text{ que }] \text{ Pierre arrivera } t_1]$
 ich mich frage - daß Pierre wird ankommen

²²Pesetsky (1998, 342) erwägt, daß der Ausschluß von (46) geleistet wird durch "was immer auch generell Präpositionsstranden im Französischen verbietet". Dies ist aber nicht ganz unproblematisch, weil der optimale Kandidat K_2 in der im nächsten Unterabschnitt befindlichen Tabelle $T_{3.27}$ für englische Relativsätze mit Rattenfang genau denselben gefährlichen Konkurrenten (mit Tilgung des Pronomens und Realisierung der Präposition) hat, das Englische aber Präpositionsstranden erlaubt.

- d. *Je me demande [_{CP} \neg_1 [_C -] Pierre arrivera t₁]
 ich mich frage - - Pierre wird ankommen

Den Wettbewerb illustriert Tabelle T_{3.24}.

T_{3.24}: *Eingebettete Fragesätze im Französischen*

Kandidaten	WIED	LR(CP)	TEL
K ₁ : ... [_{CP} quand ₁ [_C que] Pierre arrivera t ₁]		*	*!
☞ K ₂ : ... [_{CP} quand ₁ [_C -] Pierre arrivera t ₁]		*	
K ₃ : ... [_{CP} \neg_1 [_C que] Pierre arrivera t ₁]	*!		*
K ₄ : ... [_{CP} \neg_1 [_C -] Pierre arrivera t ₁]	*!	*	

Soweit zur PF-Realisierung im C-System des Französischen. Das C-System im Englischen unterscheidet sich davon u.a. dadurch, daß mehr Optionen bestehen.

2.3. Das C-System im Englischen

Im Englischen ist wie bei der Diskussion von Grimshaw (1997) gesehen ein Komplementierer im eingebetteten Objektdeklarativsatz optional; vgl. das Beispielpaar in (27), das hier mit der bei Pesetsky angenommenen Struktur für die Komplementierertilgungsvariante (also PF-getilgtes C statt IP-Einbettung) wiederholt ist.

- (48) a. I think [_{CP} [_C -] [_{IP} John will leave]]
 b. I think [_{CP} [_C that] [_{IP} John will leave]]

(48-a) verletzt LR(CP); (48-b) verletzt TEL. Da beide Kandidaten optimal sind, schließt Pesetsky, daß im Englischen LR(CP) und TEL gekoppelt, also gleich geordnet sind.²³ Die im Englischen resultierende Beschränkungsordnung ist (49):

- (49) *Ordnung im Englischen:*
 WIED \gg LR(CP) \circ TEL

Tabelle T_{3.25} verdeutlicht, wie die Optionalität von Komplementierertilgung in (48) zustande kommt.

Folgerichtig gibt es im Englischen auch in Relativsätzen mit Objektbewegung mehr PF-Realisierungsmöglichkeiten: Tatsächlich ist außer gleichzeitiger Beibehaltung von SpecC und C jede Variante grammatisch.

- (50) a. *the person [_{CP} whom₁ [_C that] I met t₁]
 b. the person [_{CP} whom₁ [_C -] I met t₁]

²³Pesetskys Kopplungsbegriff entspricht keinem der bei Prince & Smolensky (1993) gemachten Vorschläge für die Interpretation von Kopplungen und macht für andere Daten einige interessante Prognosen. Darauf wird in Kapitel 5 genauer eingegangen. Für die gegenwärtigen Zwecke und die hier betrachteten Beispiele reicht ein intuitives Verständnis von Kopplungen, wie es in Kapitel 1 dargelegt wurde, vollkommen aus.

T_{3.25}: Komplementierer in englischen Objektsätzen

Kandidaten	WIED	LR(CP) TEL
☞K ₁ : ... [CP [C that] John will leave]		*
☞K ₂ : ... [CP [C -] John will leave]		*

- c. the person [CP -₁ [C that] I met t₁]
 d. the person [CP -₁ [C -] I met t₁]

Warum dies so ist, zeigt Tabelle T_{3.26}. Bei Tilgung von *whom* in K₃ kann LR(CP) erfüllt werden, bei Tilgung von *that* in K₂ dagegen TEL. Interessant ist, daß K₂ und K₄ exakt dasselbe Profil haben und daher auch K₄ optimal ist: TEL wird in K₄ durch C-Tilgung erfüllt. Da zum einen auch K₂ LR(CP) nicht besser erfüllen kann (weil weder in K₂ noch in K₄ das erste PF-realisierte Wort der CP ein Komplementierer ist – auch wenn das erste PF-realisierte Wort in K₂ *whom* ist, in K₄ aber *I*), und zum anderen SpecC-Tilgung hier nicht WIED verletzt, unterscheiden sich die beiden Kandidaten hinsichtlich der betrachteten Beschränkungen (sowie, dies ist vorausgesetzt, aller weiterer Beschränkungen der Grammatik) nicht. Hier haben wir es daher strenggenommen nicht mit einer reinen Kopplungsanalyse zu tun, sondern mit einem Ansatz, der zur Behandlung von freier Variation zweier konkurrierender Formen auf echte Optionalität ebenso zurückgreift wie auf Beschränkungskopplung (vgl. (19-b), (19-c) aus Kapitel 1).

T_{3.26}: Relativsätze mit Objektbewegung im Englischen

Kandidaten	WIED	LR(CP) TEL
K ₁ : ... [CP whom ₁ [C that] I met t ₁]		*(!) *(!)
☞K ₂ : ... [CP whom ₁ [C -] I met t ₁]		*
☞K ₃ : ... [CP - ₁ [C that] I met t ₁]		*
☞K ₄ : ... [CP - ₁ [C -] I met t ₁]		*

Wenn jedoch PP-Bewegung mit Rattenfang ins Spiel kommt, so zeigt sich, daß die englischen PF-Realisierungsmöglichkeiten in Relativsätzen trotz der Kopplung von LR(CP) und TEL mit denen des Französischen identisch sind:

- (51) a. *the person [CP [PP to whom]₁ [C that] Mary spoke t₁ at the party]
 b. the person [CP [PP to whom]₁ [C -] Mary spoke t₁ at the party]
 c. *the person [CP [PP - -]₁ [C that] Mary spoke t₁ at the party]
 d. *the person [CP [PP - -]₁ [C -] Mary spoke t₁ at the party]

Das liegt daran, daß in diesem Fall SpecC-Tilgung eine fatale WIED-Verletzung impliziert. So wird die Optionalität aufgehoben; vgl. Tabelle T_{3.27}.

Wie der Vergleich von T_{3.23} und T_{3.24} gezeigt hat, verhalten sich aufgrund der Tatsache, daß WIED beidesmal nicht-leer appliziert, eingebettete Fragesätze im Französischen genau wie Relativsätze mit Rattenfang. Da im Englischen Relativ-

T_{3,27}: Relativsätze mit Rattenfang im Englischen

Kandidaten	WIED	LR(CP)	TEL
K ₁ : ... [CP [PP to whom] ₁ [C that] Mary spoke t ₁]		*(!)	*(!)
☞ K ₂ : ... [CP [PP to whom] ₁ [C -] Mary spoke t ₁]		*	
K ₃ : ... [CP [PP - -] ₁ [C that] Mary spoke t ₁]	*!		*
K ₄ : ... [CP [PP - -] ₁ [C -] Mary spoke t ₁]	*!	*	

sätze mit Rattenfang trotz Kopplung analog zu ihren französischen Gegenstücken analysiert werden, ist auch klar, daß englische eingebettete Fragesätze nichts Neues bringen; vgl. (52). Die dazugehörige Tabelle sieht genau wie T_{3,27} aus und wird nicht noch einmal extra angeführt.

- (52) a. *I don't know [CP who₁ [C that] she met t₁]
 b. I don't know [CP who₁ [C -] she met t₁]
 c. *I don't know [CP -₁ [C that] she met t₁]
 d. *I don't know [CP -₁ [C -] she met t₁]

2.4. Schwierigkeiten

Zunächst einmal stellt sich die Frage, worin eine weitere Umordnung der Beschränkungen resultieren würde (vgl. das Konzept der faktoriellen Typologie). WIED mag aus unabhängigen Erwägungen immer sehr hoch geordnet (oder sogar undominiert, bzw. zu Gen gehörend) sein.²⁴ Ist aber (53) eine mögliche Ordnung einer Sprache? Unter dieser Ordnung würden anders als im Französischen in T_{3,21} bis T_{3,24} immer Kandidaten ohne PF-Realisierung des Komplementierers gewinnen.

- (53) WIED ≫ TEL ≫ LR(CP)

Pesetskys (1998, 347) spekuliert hier, daß diese Situation möglicherweise im umgangssprachlichen Haitianischen vorliegt, verfolgt die Sache aber nicht weiter.

Das zweite Problem ist entgegengesetzt, weil das Pesetskysche System zu Untergenerierung zu führen scheint: Unter keiner relativen Ordnung bzw. Kopplung von LR(CP) und TEL gewinnt der Kandidat mit dem Doppel-Comp-Effekt, also gleichzeitiger Füllung von SpecC und C. Der Grund ist, daß dieser Kandidat durch einen konkurrierenden Kandidaten, der bei Füllung von SpecC C tilgt, harmonisch begrenzt ist: Keiner der beiden Kandidaten kann LR(CP) erfüllen, aber der zweite Kandidat respektiert wenigstens TEL. Wie kann das System dann Gegenbeispiele wie (54-a) (in Substandardvarietäten des Deutschen) und (54-b) (in Substandardvarietäten des Französischen, aus Pesetsky (1997, 169)) erfassen?

²⁴Chomsky (1965, 255) sagt über die Wiederauffindbarkeitsbedingung, daß sie "niemals in der Grammatik ausgedrückt werden muß, weil sie eine Voraussetzung für das Funktionieren von Grammatiken ist", was sicherlich Unverletzbarkeit garantiert.

- (54) a. eine Frau [_{CP} die₁ [_C wo] ich t₁ mag]
 b. Je me demande [_{CP} quand₁ [_C que] Pierre arrivera t₁]
 ich mich frage wann daß Pierre wird ankommen

Pesetsky (1997, 169; 1998, 377) schlägt als Ausweg für den französischen Fall in (54-b) vor, daß in diesen Varietäten zwei CPs da sind. Dies setzt dann weitere Beschränkungen voraus, die Entsprechendes im Standardfranzösischen verhindern. Darüber hinaus ist nicht klar, wieso dann das W-Element immer in der höheren SpecC-Position stehen muß und der Komplementierer in der tieferen C-Position (und wieso nicht z.B. in den beiden CPs nur zwei W-Elemente oder nur zwei Komplementierer realisiert werden können). Eine (rein stipulative) Alternative kann man in Anlehnung an Heck (1997) formulieren: Angenommen, es gibt eine zusätzliche Beschränkung SPRICH-KOPF, die z.B. in den Varietäten des Deutschen, die (54-a) erlauben, mit TEL und LR(CP) gekoppelt ist. Diese Beschränkung würde für C praktisch das Gegenteil von LR(CP) fordern und kann ad hoc wie folgt formuliert werden: "Ein Kopf muß ausgesprochen werden, aber nicht am linken Rand seiner Projektion." Dann ergibt sich korrekt die Vorhersage, daß die Kandidaten K₁, K₂ und K₃ in T_{3,28} optimal sind (ein identisches Beschränkungsprofil für diese Kandidaten folgt aus sämtlichen Standardinterpretationen für Kopplungen; s.u., Kapitel 5).

T_{3,28}: Relativsätze mit Objektbewegung in Varietäten des Deutschen

Kandidaten	?	LR(CP)	TEL	SPRICH-KOPF
☞ K ₁ : eine Frau [_{CP} die ₁ [_C wo] ich t ₁ mag]		*	*	
☞ K ₂ : eine Frau [_{CP} die ₁ [_C -] ich t ₁ mag]		*		*
☞ K ₃ : eine Frau [_{CP} - ₁ [_C wo] ich t ₁ mag]			*	*
K ₄ : eine Frau [_{CP} - ₁ [_C -] ich t ₁ mag]	*!	*		*

Im Englischen müßte demgegenüber SPRICH-KOPF tief geordnet sein, so daß K₁ nicht optimal sein kann. Ein zusätzliches Problem ist aber noch zu lösen: K₄ ist im Deutschen (anders als im Englischen) unmöglich, obwohl der Kandidat auf den ersten Blick ein mit K₂ identisches Beschränkungsprofil hat. Es muß folgerichtig dann also noch eine weitere Beschränkung geben, die K₄ verletzt und K₁, K₂ und K₃ erfüllen, und die nicht mit LR(CP), TEL und SPRICH-KOPF gekoppelt ist. Eine Möglichkeit könnte hier eine Beschränkung sein, die Füllung mindestens einer Position im C-System erzwingt (wobei man dann aber Sorge tragen muß, daß diese Beschränkung trotz niedriger Einordnung nicht unerwünschterweise im Englischen aktiv werden kann). Diese Überlegungen mögen jedoch genügen, um zum einen zu zeigen, wie man im Pesetskyschen System widerstreitende Evidenz erfassen kann, und um zum anderen zu veranschaulichen, daß die resultierende Theorie dann

schnell Gefahr läuft, mechanistische Züge zu erhalten.²⁵

2.5. PF-Realisierung von Spuren

Eine Analyse auf der Grundlage des Konzepts der optimalen PF-Realisierung schlägt Pesetsky auch für das Phänomen vor, daß in Basispositionen von Ketten meist Spuren stehen, manchmal aber auch resumptive Pronomina. Die wesentliche Generalisierung ist, daß resumptive Pronomina oft nur dann möglich sind, wenn Spuren unmöglich sind (vgl. etwa Shlonsky (1992)). Pesetskys Annahme ist nun, daß Spuren im klassischen Sinne und resumptive Pronomina auf der S-Struktur gar nicht unterschieden sind. Er adaptiert Chomskys (1993; 1995) Kopiertheorie der Bewegung, derzufolge syntaktische Bewegung eine vollständige Kopie hinterläßt; nicht-initiale Kopien in S-Struktur-Ketten werden dann "Spuren" genannt, unabhängig von ihrer PF-Realisierung. Ob eine Kopie-Spur dann gar nicht PF-realisiert wird (wie Spuren im klassischen Sinne), als resumptives Pronomen PF-realisiert wird, oder sogar als vollständige Kopie PF-realisiert wird, hängt einzig und allein von der Ordnung verletzbarer und gewichteter Beschränkungen ab. Die hier relevanten Beschränkungen sind ST-SP und INSEL.²⁶

- (55) a. ST-SP ("Stille Spur", "Silent t", "Sil-Trc"):
Spuren (= nicht-initiale Kopien in Ketten) dürfen nicht PF-realisiert werden.
- b. INSEL ("Island Constraints", "Island Cond"):
* α ... [_{Insel} ... β ...], wobei β eine nicht PF-realisierte Spur von α ist.

In Inselkontexten konfigrieren die beiden Beschränkungen: ST-SP verlangt die vollständige PF-Tilgung von Kopie-Spuren, aber INSEL verbietet genau dies, wenn zwischen der Kopie-Spur und ihrem Antezedens eine Insel interveniert. Angenommen nun, es gilt die Ordnung INSEL \gg ST-SP. Dann kann die Spur nicht PF-getilgt werden. Die Idee ist nun, daß bei Verletzung von ST-SP deshalb aber nicht etwa eine vollständige Kopie erscheint. Um die Verletzung von ST-SP minimal zu halten, wird vielmehr die Spur als resumptives Pronomen PF-realisiert, das nur die Φ -Merkmale (Person, Numerus, Genus) einer XP aufweist, nicht jedoch deren sonstige PF-Merkmale (was ST-SP per Annahme stärker verletzen würde). Resumptive Pronomina sind also eine Reparaturstrategie mit inhärenter Verletzung von ST-SP, die die Grammatik wählt, wenn ST-SP aufgrund einer höher geordneten Beschränkung (INSEL) verletzt werden muß, weil so die Verletzung minimal gehalten werden

²⁵Im übrigen wäre zu den Doppel-Comp-Effekten in Varietäten des Deutschen gerade in einem optimalitätstheoretischen Rahmen noch weit mehr zu sagen (u.a. bzgl. der Kasuskongruenz von Relativpronomen und modifizierter NP); vgl. Bayer (1984).

²⁶Pesetsky möchte nicht behaupten, daß INSEL tatsächlich eine universelle Beschränkung ist. INSEL soll sich idealerweise als Theorem aus einer geordneten Menge von allgemeineren Lokalisierungsbeschränkungen ergeben.

kann. Diesen Effekt illustriert Pesetsky anhand von W-Inseln im Englischen. (56) zeigt, daß hier von den drei a priori möglichen PF-Realisierungen der Kopie-Spur die Variante mit resumptivem Pronomen die einzig grammatische ist.

- (56) a. * $[_{NP}$ Which picture of John $]_1$ were you wondering $[_{CP}$ whether $_{-1}$ was going to win a prize at the exposition $] ?$
 b. * $[_{NP}$ Which picture of John $]_1$ were you wondering $[_{CP}$ whether $[_{NP}$ which picture of John $]_1$ was going to win a prize at the exposition $] ?$
 c. $[_{NP}$ Which picture of John $]_1$ were you wondering $[_{CP}$ whether it_1 was going to win a prize at the exposition $] ?$

Der Wettbewerb ist in Tabelle T_{3.29} dargestellt.

T_{3.29}: *W-Inseln im Englischen*

Kandidaten	INSEL	ST-SP
K ₁ : $[_{NP}$ which picture of John $]_1 \dots [_{W-Insel} \dots [_{NP} \text{ }_{-1}] \dots]$	*!	
K ₂ : $[_{NP}$ which picture of John $]_1 \dots$ $[_{W-Insel} \dots [_{NP}$ which picture of John $]_1 \dots]$		**!
\Rightarrow K ₃ : $[_{NP}$ which picture of John $]_1 \dots [_{W-Insel} \dots [_{NP} it_1] \dots]$		*

Wenn es sich hier um ein Reparaturphänomen handelt, sollten resumptive Pronomina in bewegungstransparenten Kontexten im Englischen blockiert sein. Daß dem so ist, zeigen die Beispiele in (57); warum dem so ist, zeigt Tabelle T_{3.30}.²⁷

- (57) a. Who₁ do you like $_{-1} ?$
 b. *Who₁ do you like who₁ ?
 c. *Who₁ do you like him₁ ?

T_{3.30}: *Bewegungstransparente Kontexte im Englischen*

Kandidaten	INSEL	ST-SP
\Rightarrow K ₁ : W ₁ ... t ₁		
K ₂ : W ₁ ... who ₁		*!*
K ₃ : W ₁ ... him ₁		*!

Die so entstehende Theorie der PF-Realisierung in Bewegungsketten wird von Pesetsky ausgeweitet auf eine Reihe von Sprachen, wo sich die Effekte zum Teil

²⁷Man erinnere sich in diesem Zusammenhang an die Diskussion des Konzepts absoluter Ungrammatikalität bei Adjunkt-Inselverletzungen im Deutschen in Abschnitt 8.3 von Kapitel 2. Dort war bereits die Möglichkeit diskutiert worden, daß im Idealfall die Blockade durch eine Reparaturform absolute Ungrammatikalität erfassen kann. Es hat sich aber gezeigt, daß eine solche Lösung für viele Fälle nicht funktioniert, weil die "Reparaturformen" bei genauerem Hinsehen keine sind (sondern auch in bewegungstransparenten Kontexten vorkommen können); vgl. (79) in Kapitel 2. Die von Pesetsky betrachteten Inseldaten kommen dem Idealfall also näher.

noch besser als im Englischen zeigen lassen (Modernes Hebräisch, Polnisch, Russisch); um das Grundprinzip zu illustrieren, reichen die betrachteten Daten aus dem Englischen aber aus. Auf die Situation im Polnischen werde ich in Kapitel 5 noch zurückkommen, denn die postulierte Kopplung von ST-SP mit zwei weiteren Beschränkungen motiviert ganz wesentlich Pesetskys (1997; 1998) besondere Interpretation von Kopplungen.

2.6. Einordnung

Pesetskys (1997; 1998) Ansatz hat einige Merkmale, die ihn von anderen optimalitätstheoretischen Syntax-Analysen unterscheiden. Zum einen ist die Kandidatenmenge finit; da aber immer noch zur Ermittlung von Optimalität eine gewaltige Anzahl von Kandidaten simultan betrachtet werden muß, ist nicht ganz klar, wie groß der Vorteil hier ist. Zum anderen ist die Rolle von Gen gegenüber dem H-Eval-Teil der Grammatik gestärkt. Auf den ersten Blick sieht es so aus, als ginge es hier im H-Eval-Teil nur um in gewisser Weise periphere Fragen der Phonologie-Syntax-Schnittstelle, und als entspräche Gen mehr oder weniger einer elaborierten Standardgrammatik (z.B. einer minimalistischen oder einer GB-Syntax), und nicht einer reduzierten Variante davon. Daß dieser Eindruck aber nicht ganz korrekt ist, zeigt die Tatsache, daß Pesetskys System der PF-Realisierung Arbeit tut, die ansonsten auch Standardsyntaxen übernehmen, und daß die verwendeten Beschränkungen explizit auf elaborierte syntaktische Konzepte Bezug nehmen (dies gilt z.B. für die Beschränkung INSEL, aber auch für die Beschränkung TCP ("Tilgung in CP"), die in Kapitel 5 noch behandelt wird und die zwischen Komplement-CPs und Nicht-Komplement-CPs unterscheidet). Die Aufgaben, die das H-Eval-System der PF-Realisierung bei Pesetsky erfüllt, sind damit naturgemäß Gen entzogen, und letztere Komponente ist daher doch eine reduzierte Variante der Standardgrammatik. Zum Beispiel ergibt sich für Pesetskys Gen-Syntax, daß sie ohne einen Teil der Lokalisierungsbeschränkungen für Bewegung auskommt – denn diese sind ja wie gesehen im H-Eval-Teil lokalisiert (auch wenn sie dort als Aussprachebeschränkungen kodiert sind).

Bei aller grundsätzlichen Verschiedenheit von anderen OT-Analysen (wie z.B. der von Grimshaw (1997)) finden sich auch bei Pesetsky die Typen von Evidenz, die in Kapitel 2 als eine optimalitätstheoretische Herangehensweise nahelegend identifiziert wurden. Das Konzept der *parametrischen Variation* per Beschränkungsanordnung ist bei der Behandlung von PF-Realisierungsoptionen im C-System sehr wichtig (und Pesetsky nimmt das Konzept der faktoriellen Typologie so ernst, daß er wie gesehen auch die zunächst einmal nicht in offensichtlicher Weise dokumentierte Ordnung TEL \gg LR(CP) als möglich annehmen will); und ebenso zeigt Pesetsky, daß sich sprachspezifische Unterschiede bei der PF-Realisierung in Bewegungsketten gut durch Beschränkungsanordnung erfassen lassen. *Reparaturphänomene* spielen wie gesehen eine wesentliche Rolle (vgl. die Diskussion resumptiver Pronomina). Dasselbe gilt für das Konzept des *Wettbewerbs*: Wie bei Grimshaw

sind komplementiererlose und komplementiererhaltige Sätze miteinander im Wettbewerb, und während in englischen Objektsätzen durch die Kopplung von TEL und LR(CP) beide Varianten optimal sein können, sieht das Ganze im Französischen anders aus (und auch im Englischen sorgt die bereits erwähnte Beschränkung TCP dafür, daß etwa in Subjektsätzen die eine Variante die andere als suboptimal blockieren kann; vgl. Kapitel 5). Aus Pesetskys Ausführungen folgt weiterhin, daß *im unmarkierten Fall* Spuren nicht PF-realisiert werden (denn dies fordert die tiefgeordnete Beschränkung ST-SP); und schließlich beruht das System auf jeweils gut etablierten Beschränkungen, deren maximal generelle Formulierung automatisch zu *Beschränkungskonflikten* führt.

Insgesamt gilt, daß die bei Pesetsky vorgenommene Stärkung von Gen ebenso wie der Versuch, H-Eval-Beschränkungen auf eine bestimmte formale Domäne (PF-Realisierung) zu konzentrieren, für die Diskussion des Aufbaus einer Optimalitätstheoretischen Syntax wichtig waren und sind. Im Einklang damit ist Pesetskys (1997; 1998) Untersuchung für die weitere Forschung einflußreich gewesen. Ein Beispiel sind die Arbeiten von Broekhuis & Dekkers (1997) und Dekkers (1999); die Autoren versuchen hier u.a., das von Pesetsky entwickelte System auf PF-Realisierungsoptionen in Relativsätzen im Niederländischen auszudehnen. Eine weitere wesentlich von Pesetsky (1997; 1998) beeinflusste Untersuchung ist Fanselow & Ćavar (2000), wo verschiedene Typen von W-Abhängigkeiten (lange W-Bewegung, partielle W-Bewegung, W-Kopierbewegung) und NP-Aufspaltungskonstruktionen im Deutschen und in anderen Sprachen bei Annahme der Kopiertheorie der Bewegung durch Optimierung der PF-Realisierung erfaßt werden.

3. Legendre, Smolensky & Wilson (1998) über Lokalität

Hauptziel von Legendre, Smolensky & Wilson (1998) ist die Formulierung einer Lokalitätstheorie, die auf der einen Seite hinreichend flexibel ist, die dokumentierte Sprachvariation bei Extraktionsmöglichkeiten zu erfassen, und die auf der anderen Seite dennoch stark genug ist, um weitreichende Vorhersagen zu machen. Eine Variante dieser Analyse ist Legendre, Wilson, Smolensky, Homer, & Raymond (1995). Wichtige theoretische Konzepte, die Verwendung finden, sind die Neutralisierung (vgl. Kapitel 2) und die lokale Konjunktion von Beschränkungen.

3.1. Grundannahmen

So wie bei Grimshaw (1997) sind die im Wettbewerb befindlichen Output-Kandidaten S-Struktur-Repräsentationen. Kandidatenmengen sind jedoch etwas anders definiert, nämlich wie in (58).

(58) *Kandidatenmenge* (Legendre, Smolensky & Wilson (1998, 257)):

Zwei Kandidaten K_i , K_j sind in derselben Kandidatenmenge gdw. gilt:

- a. K_i und K_j sind Realisierungen von identischen Prädikat-/Argument-Strukturen.

- b. K_i und K_j zielen auf identische LFs.

Der entscheidende Unterschied zu Grimshaws Definition ist, daß zwei Kandidaten nicht notwendigerweise identische LFs haben müssen, um zu konkurrieren; es reicht, wenn sie auf identische LFs abzielen. Legendre, Smolensky & Wilson (1998) nehmen weiterhin an, daß der Wettbewerb rein über Input-Identität festgelegt werden kann, d.h., daß im Input jeweils eine Prädikat-/Argument-Struktur vorliegt, für die eindeutig eine LF vorgegeben ist. Die (von den Basispositionen abweichenden) Zielpositionen für LF-Interpretation sind dabei durch abstrakte Skopusmarker festgelegt.²⁸ Der syntaktische Input ist in dieser Theorie also bereits ein äußerst komplexes, hierarchisch geordnetes Gebilde (wie in Abschnitt 5 von Kapitel 1 bemerkt, ist diese Sichtweise nahe an der der Generativen Semantik).

Als technischen Begriff für ihr spezielles Input-Konzept in der Syntax führen Legendre et al. (1995; 1998) den Begriff des *Index* ein. Zwei Kandidaten sind also im selben Wettbewerb, wenn sie auf denselben Index zurückgehen. Dies setzt voraus, daß sie im Index (Input) dieselbe LF haben; man beachte aber, daß nicht notwendigerweise impliziert ist, daß die konkurrierenden Kandidaten auch dieselbe LF im Output haben müssen. Es kann sein, daß der Gewinner in einem Wettbewerb der im Index angezielten LF untreu sein muß, weil die entsprechende Treuebeschränkung tiefer geordnet ist als die Beschränkung, die ein LF-treuer Kandidat verletzen muß. Damit ergibt sich Neutralisierung, eine elegante Möglichkeit, um das Phänomen der absoluten Ungrammatikalität in den Griff zu bekommen (vgl. Kapitel 2): Es gewinnt ein semantisch untreuer "partieller Parse", der die ansonsten fatalen Merkmale des Kandidaten unrealisiert läßt; und der Gewinner dieses Wettbewerbs ist oberflächenidentisch mit dem semantisch treuen Gewinner eines anderen Wettbewerbs, in dem keine Abweichung von der angezielten LF stattfinden muß. Somit wird ein in zwei minimal unterschiedlichen Indizes vorhandener Kontrast (z.B. weiter vs. enger Skopus) im Output, also auf der syntaktischen Oberfläche, neutralisiert. Der nächste Unterabschnitt illustriert eine erste, einfache Anwendung dieses Mechanismus.

3.2. W-Extraktion im Chinesischen

Betrachtet werden Fälle von W-Extraktion im Chinesischen aus Deklarativsätzen und aus Fragesätzen. Das Chinesische wird klassischerweise als W-in-situ-Sprache betrachtet, in der W-Bewegung nicht auf der S-Struktur stattfindet, sondern auf LF. Demgegenüber nehmen Legendre, Smolensky & Wilson (1998, 269) an, daß man im Chinesischen S-strukturelle W-Extraktion zwar nicht unmittelbar sehen kann, daß sie aber dennoch in gewisser Weise erfolgt ist: Es gibt eine S-strukturelle

²⁸Dies ist bei Grimshaw (1997) anders, denn dort ist LF-Information wie gesehen nicht im Input vermerkt, obschon das Konzept der LF-Identität in die Definition der Kandidatenmenge eingeht.

Kette zwischen einem leeren Operator in A'-Position und einer damit koindizierten W-Phrase in situ. Diese Kette ist im Prinzip denselben Lokalitätsbeschränkungen unterworfen wie andere S-strukturelle Ketten, in denen die W-Phrase in einer A'-Position steht und eine koindizierte Spur bindet.²⁹

Unter diesen Prämissen kann man nun zunächst festhalten, daß eine Adjunktextraktion aus eingebetteten Deklarativsätzen im Chinesischen möglich ist; d.h., (59-a) kann die Bedeutung in (59-b) haben.

- (59) a. Ni renwei [_{CP} Lisi yinggai zenmeyang chuli zhe-jian
 du denkst Lisi sollte wie behandeln diese-CL
 shi] ?
 Angelegenheit
 b. "Wie₁ denkst du, daß Lisi diese Angelegenheit t₁ behandeln sollte?"

Unmöglich ist dagegen im Chinesischen eine Adjunktextraktion aus eingebetteten Fragesätzen; W-Inseln sind für Adjunktketten strikt. Dies zeigt das Beispiel (60-a), in dem das W-Adjunkt *zenmeyang* wie durch die Übersetzungen in (60-bc) angedeutet nicht im Matrixsatz interpretiert werden kann, sondern nur im eingebetteten Satz.

- (60) a. Ni xiang-zhidao [_{CP} shei zenmeyang chuli zhe-jian shi] ?
 du fragst dich wer wie behandeln diese-CL Angelegenheit
 b. "*Wie₁ fragst du dich, wer diese Angelegenheit t₁ behandelt hat?"
 c. "Du fragst dich, wer diese Angelegenheit wie behandelt hat."

Drei Beschränkungen spielen bei der Erklärung dieses Unterschieds eine Rolle:

- (61) a. SEL ("Selektion", "Select"):
 Lexikalisch markierte Selektion muß respektiert werden.
 b. BAR^{2[-ref]} ("Zwei-Barrieren-Bedingung für Adjunktketten"):
 Ein Kettenglied einer nicht-referentiellen (Adjunkt-) Kette darf nicht zwei Barrieren überkreuzen.
 c. SKOP-TREUE ("Skopus-Treue", "Parse-Scope"):
 Ein im Index markierter Skopus muß durch S-strukturelle Kettenbildung in diese Position hinein realisiert werden.

Die Treuebeschränkung SEL entspricht Chomskys (1981) Projektionsprinzip und ist

²⁹Die Argumentation ist hier ganz ähnlich wie bei Pesetsky (1998, 361-362), der davon ausgeht, daß das, was klassischerweise als LF-Bewegung gedeutet wurde, tatsächlich S-strukturelle Bewegung ist, bei der die Beschränkung ST-SPUR (in (55-a)) maximal verletzt werden muß. Vgl. auch die Rekonstruktion von Grewendorf (1997) in Kapitel 4. Konkret nehmen Legendre, Smolensky & Wilson (1998, 269) an, daß im Chinesischen die Beschränkung, die Spuren verbietet (ihr *t) der Beschränkung, die leere Operatoren verbietet (ihr *Q) übergeordnet ist. Wie bereits in Kapitel 2 erwähnt, legen Legendre et al. (1995; 1998) letztlich eine vollständig ausgearbeitete Variante des W-Bewegungsparameters vor. Darum soll es im folgenden aber nicht gehen.

im wesentlichen auch der Beschränkung PROJ-P aus Kapitel 2 äquivalent (vgl. dort (13-b), (18-c)). Sie besagt, daß lexikalisch determinierte Selektionseigenschaften in der Syntax respektiert werden müssen. Für Prädikate wie *sich fragen*, die lexikalisch einen [+w]-Komplementsatz selektieren, gilt somit, daß SEL einen [+w]-Satz in der Syntax verlangt; für Prädikate wie *denken* erfordert SEL in der Syntax einen [-w]-Komplementsatz. $\text{BAR}^{2[-ref]}$ ist ein Spezialfall einer umfassenden, viel allgemeineren Lokalitätsbedingung, die unten noch eingeführt wird; für die gegenwärtigen Daten ist jedoch nur der Spezialfall von Bedeutung. Wichtig ist die Annahme, daß IP und VP immer Barrieren sind.³⁰ Die Treuebeschränkung SKOP-TREUE schließlich verlangt, daß der im Index vorgegebene Skopus einer W-Phrase in der Syntax durch Kettenbildung realisiert wird. Eine nicht-fatale Verletzung eben dieser Beschränkung erweist sich als Kern der Neutralisierungsanalyse der Unmöglichkeit langer Adjunktbewegung aus W-Inseln im Chinesischen.

Die drei Beschränkungen sind im Chinesischen wie in (62) geordnet.

- (62) *Ordnung im Chinesischen:*
 SEL \gg $\text{BAR}^{2[-ref]}$ \gg SKOP-TREUE

Die Ungrammatikalität von (60-a) unter der Lesart (60-b), also mit langer Adjunktkettenbildung, ergibt sich nun wie in den Tabellen T_{3.31} und T_{3.32} gezeigt. Hier soll Q₁ jeweils für die tatsächlich durch syntaktische Kettenbildung erreichte Position stehen, [_S] dagegen für die im Index vorgegebene Skopus-Zielposition (letzteres weicht ein bißchen von der Notation in Legendre et al. (1995; 1998) ab, stellt aber keine substantielle Änderung dar). In T_{3.31} ist im Index weiter Skopus für das W-Adjunkt vermerkt. Beim Versuch von K₁, dieser Vorgabe durch syntaktische Kettenbildung zu entsprechen, wird jedoch fatal $\text{BAR}^{2[-ref]}$ verletzt, weil die Barrieren VP und IP im Matrixsatz überquert werden müssen. Daher ist hier K₂ optimal, wo der Skopus des W-Adjunkts gegenüber dem Index vom Matrix- auf den eingebetteten Satz umgestellt worden ist. Dies verletzt zwar SKOP-TREUE, aber aufgrund der Ordnung $\text{BAR}^{2[-ref]}$ \gg SKOP-TREUE in nicht-fataler Weise.

T_{3.31}: W-Inseln für weiten Adjunktskopus im Chinesischen

Kandidaten	SEL	$\text{BAR}^{2[-ref]}$	SKOP-TREUE
K ₁ : [_S Q ₁] ... V _[+w] [_{CP} t' ₁ ... wie ₁ ...]		*!	
☞ K ₂ : [_S -] ... V _[+w] [_{CP} Q ₁ ... wie ₁ ...]			*

Ist dagegen wie beim durch Tabelle T_{3.32} dokumentierten Wettbewerb im Index von vornherein schon ein eingebetteter Skopus für das W-Adjunkt vorgegeben,

³⁰Dies folgt automatisch, wenn man mit Chomsky (1986a) voraussetzt, daß alle XPs Barrieren sind, die nicht L-markiert (also, etwas vereinfacht, nicht Schwestern lexikalischer Köpfe) sind, und keine Zusatzstipulationen etwa über den "defektiven" Charakter von IP oder L-Markierung von VP durch ein nach I gewandertes V selbst machen möchte; vgl. Sternefeld (1991).

kann der optimale Kandidat (K_1) sowohl $\text{BAR}^{2[-ref]}$, als auch SKOP-TREUE erfüllen. Lange Kettenbildung wie bei K_2 involviert hier nicht nur eine Treueverletzung gegenüber dem Index, sondern auch noch eine Barrierenverletzung; beide Verletzungen reichen bereits allein für die Klassifikation von K_2 als suboptimal aus.

$T_{3.32}$: Kurzer Adjunktskopos im Chinesischen

Kandidaten	SEL	$\text{BAR}^{2[-ref]}$	SKOP-TREUE
$\text{K}_1: \dots V_{[+w]} [CP [S Q_1] \dots \text{wie}_1 \dots]$			
$\text{K}_2: Q_1 \dots V_{[+w]} [CP [S -] \dots \text{wie}_1 \dots]$		*!	*

Die optimalen Kandidaten K_2 in $T_{3.31}$ und K_1 in $T_{3.32}$ sind oberflächenidentische Formen; der Unterschied zwischen beiden Inputs hinsichtlich der W-Skopus-Zielposition ist im W-Insel-Kontext neutralisiert. Die so entstehende derivationelle Ambiguität wird von Legendre, Smolensky & Wilson (1998) durch das von Prince & Smolensky (1993) entwickelte Meta-Optimierungsverfahren der Input-Optimierung vermieden, das bereits in Abschnitt 8.3.5 von Kapitel 2 kurz eingeführt worden ist: Der auf "kompliziertere" Art und Weise entstandene, Index-untreue optimale Kandidat K_2 in $T_{3.31}$ ist gegenüber dem Index-treuen Kandidaten K_1 in $T_{3.32}$ blockiert, weil er eine höher-rangige Beschränkung verletzt.

Konkret stellt sich das Verfahren der Input-Optimierung von Prince & Smolensky (1993) wie folgt dar.³¹

(63) *Input-Optimierung* (nach Prince & Smolensky (1993, 192)):

Man nehme an, daß verschiedene unterschiedliche Inputs I_1, I_2, \dots, I_n in einer Grammatik zu korrespondierenden Outputs O_1, O_2, \dots, O_n führen, die alle als dieselbe Form Φ realisiert werden. Es muß einer dieser Outputs der harmonischste sein, weil er die am wenigsten signifikanten Verletzungen hervorruft; dieser Output sei O_k . Dann sollte der Lerner als zugrundeliegende Form von Φ den Input I_k wählen.

So werden durch Neutralisierung entstehende derivationelle Ambiguitäten durch sekundäre Optimierung ausgeschlossen.

Damit ist die Unmöglichkeit von langer Adjunktbewegung aus $[+w]$ -Sätzen erklärt. Was noch fehlt, ist die Analyse von langer Adjunktbewegung aus $[-w]$ -Sätzen wie in (59). An dieser Stelle wird die höchst-geordnete Beschränkung SEL wichtig,

³¹Prince und Smolensky entwickeln dieses Verfahren vor allem anhand der Phonologie; sie nennen es auch nicht "Input-Optimierung", sondern "Lexikon-Optimierung". Allerdings stellen sie selbst fest (p. 192, Fußnote 73), daß der Begriff "Lexikon" hier letztlich irreführend ist, denn es geht grundsätzlich um zugrundeliegende Formen, nicht notwendigerweise um lexikalische Einheiten, wie in der Phonologie. Im vorliegenden Fall etwa betrifft die Input-Optimierung nicht etwa ein im weiteren Sinne lexikalisches Element wie den abstrakten Skopusmarker, sondern nur bestimmte strukturelle Konfigurationen, in denen der Skopusmarker auftritt.

denn abgesehen von dieser Beschränkung verhalten sich die Extraktion aus einem [+w]-Satz und die Extraktion aus einem [-w]-Satz unter den bisherigen Annahmen zunächst einmal gleich: Auch in (59) ist eine Verletzung von $\text{BAR}^{2[-ref]}$ bei weitem W-Skopus unumgänglich. (Die beiden Extraktionsstrategien verhalten sich entweder deshalb gleich, weil eine Zwischenlandung in der eingebetteten SpecC-Position in beiden Fällen möglich ist, oder weil in (59) von vornherein nur eine von V L-markierte IP eingebettet wird – die Lokalitätsverletzung ergibt sich beidesmal durch Überschreitung der Barrieren VP und IP im Matrixsatz.) Dies bedeutet, daß anders als in den meisten klassischen Extraktionstheorien (vgl. z.B. Chomsky (1986a) oder Rizzi (1990)) W-Inseln nicht darauf zurückgeführt werden, daß in SpecC ein W-Element steht, das eine Fluchtstelle ('escape hatch') blockiert; die Lokalitätsbedingung $\text{BAR}^{2[-ref]}$ ist an sich nicht sensitiv für die Gefülltheit oder Leere von SpecC-Positionen. Der W-Insel-Effekt ergibt sich vielmehr daraus, daß eine durch die Ordnung $\text{BAR}^{2[-ref]} \gg \text{SKOP-TREUE}$ ausgelöste Reduktion des W-Skopus vom Matrixsatz auf den eingebetteten Satz nur dann die noch höher geordnete Treuebeschränkung SEL respektieren kann, wenn der eingebettete Satz ein Fragesatz ist. Ist der eingebettete Satz dagegen ein Deklarativsatz, verletzt eine Skopusreduktion fatal SEL, weil ein lexikalisch als [-w] markierter Satz als [+w] reinterpretiert werden muß. Mit anderen Worten: Rein unter Lokalitätsgesichtspunkten sind eingebettete Deklarativsätze ebenso sehr Inseln wie eingebettete Fragesätze; daß erstere dennoch bewegungstransparent sind, haben sie einzig der Tatsache zu verdanken, daß eine Neutralisierung durch Skopusreduktion die undominierte Beschränkung SEL verletzen würde. Diesen Gedankengang illustriert Tabelle T_{3.33}.

T_{3.33}: Deklarativsätze und weiter Adjunktskopus im Chinesischen

Kandidaten	SEL	$\text{BAR}^{2[-ref]}$	SKOP-TREUE
$\text{K}_1: [S \text{ Q}_1] \dots V_{[-w]} [CP \text{ t}'_1 \dots \text{wie}_1 \dots]$		*	
$\text{K}_2: [S -] \dots V_{[-w]} [CP \text{ Q}_1 \dots \text{wie}_1 \dots]$	*!		*

Nach dieser einführenden Illustration sollen im folgenden die extraktionsrelevanten Beschränkungen etwas genauer betrachtet und das Konzept der lokalen Konjunktion eingeführt werden.

3.3. Extraktion und lokale Konjunktion

3.3.1. Treuebeschränkungen

Neben SKOP-TREUE nehmen Legendre, Smolensky & Wilson (1998, 260) noch zwei weitere Beschränkungen an:

- (64) a. W-TREUE ("Parse(wh)"): Ein im Index auf XP auftretendes W-Merkmal muß im Output durch eine Operator-Variablenkette realisiert werden.

- b. TOP-TREUE (“Topik-Treue”, “Parse(top)”):
Ein im Index auf XP auftretendes Top-Merkmal muß im Output durch eine Operator-Variablenkette realisiert werden.

W-TREUE ist (in etwas unterschiedlicher Formulierung, aber essentiell identisch) bereits bei der Illustration von Neutralisierungsanalysen von absoluter Ungrammatikalität in Kapitel 2 aufgetaucht (vgl. das dortige (78)). Wie bei SKOP-TREUE führt eine Verletzung von W-TREUE in einem optimalen Kandidaten zu Neutralisierung. Im ersteren Fall bleibt eine im Index vorhandene Skopusmarkierung zwar per Skopusreduktion unrealisiert, aber das W-Element wird nach wie vor als W-Element gedeutet. Im letzteren Fall dagegen wird das W-Element gar nicht mehr als solches interpretiert (sondern z.B. als Indefinitum), was sich als noch drastischerer Schritt erweist. TOP-TREUE hat denselben Status wie W-TREUE; eine nicht-fatale Verletzung dieser Beschränkung bei langen Abhängigkeiten erklärt die Ungrammatikalität von zu langen Extraktionen, die statt W-Elementen Topik-Elemente involvieren.³²

Die Treuebeschränkungen interagieren systematisch mit den Lokalitätsbeschränkungen.

3.3.2. Lokalitätsbeschränkungen

Bislang haben wir mit $\text{BAR}^{2[-ref]}$ (hier wiederholt als (65-a)) den Spezialfall einer allgemeinen Lokalitätsbeschränkung BAR kennengelernt. BAR selbst ist wie in (65-b) definiert (vgl. auch (38-b) aus Kapitel 2).

- (65) a. $\text{BAR}^{2[-ref]}$ (“Zwei-Barrieren-Bedingung für Adjunktketten”):
Ein Kettenglied einer nicht-referentiellen Kette darf nicht zwei Barrieren überkreuzen.
- b. BAR (“Barrierenbedingung”):
Ein Kettenglied darf keine Barriere überkreuzen.

Zu klären ist, wie man von der allgemeinen Beschränkung BAR zu speziellen Ausformungen wie $\text{BAR}^{2[-ref]}$ kommt, und was überhaupt die Distinktion erfordert. Beginnen wir mit dem zweiten Punkt, und nehmen wir versuchsweise einmal an, daß nur BAR als Lokalitätsbeschränkung Verwendung findet. Dann ergibt sich sofort eine falsche Prognose für die beiden Kandidaten in Tabelle T_{3.34}. (Hier steht jeweils β für eine Barriere, also eine nicht L-markierte Kategorie wie z.B. VP oder IP, α_1 für das Erstglied einer Kette und t_1 für das Endglied derselben Kette.)

Das Problem ist hier, daß K_1 und K_2 dasselbe Beschränkungsprofil haben und so (bei Suboptimalität anderer Kandidaten) beide optimal sein sollten, obwohl bei K_1 sukzessiv-zyklisch und bei K_2 in einem Rutsch bewegt wird und K_2 deswegen

³²Analoge Treuebeschränkungen wären für alle Bewegungstypen zu formulieren, die zwar lang applizieren können, aber dabei Insel-sensitiv sind.

$T_{3.34}$: Eine falsche Prognose unter BAR

Kandidaten	BAR
$\Rightarrow K_1: \alpha_1 \dots \beta \dots \beta \dots t'_1 \dots \beta \dots t_1$	** *
$\Rightarrow K_2: \alpha_1 \dots \beta \dots \beta \dots \beta \dots t_1$	***

schlechter als K_1 sein sollte. Wenn es nun aber neben BAR auch noch eine Beschränkung gibt, die die simultane Überquerung von zwei Barrieren bestraft, und dann noch eine Beschränkung, die die simultane Überquerung von drei Barrieren bestraft, usw., dann kann abgeleitet werden, daß K_1 ein besseres Beschränkungsprofil hat als K_2 . Natürlich möchte man diese zusätzlichen Barrieren-Bedingungen nicht einzeln stipulieren, sondern durch allgemeine Prinzipien rekursiv erzeugen. Das hierfür verwendete Mittel ist die Gewinnung von Beschränkungshierarchien durch lokale Konjunktion (“local conjunction”; Smolensky (1995)). Lokale Konjunktion nimmt “primitive” (nicht-konjugierte) Beschränkungen als Basis (hier: BAR) und erzeugt rekursiv immer komplexere Beschränkungen, die zueinander in fixer Ordnung stehen. Das Konzept der lokalen Konjunktion ist in (66) definiert:

(66) *Lokale Konjunktion*:

- a. Die lokale Konjunktion zweier Beschränkungen B_1, B_2 bzgl. eines Bereichstyps D ergibt eine neue Beschränkung $B_1 \&_l B_2$, die verletzt wird, falls zwei unterschiedliche Verletzungen von B_1 und B_2 in einem einzigen Bereich vom Typ D auftreten.
- b. Universell gilt: $B_1 \&_l B_2 \gg \{B_1, B_2\}$

Aus (66-b) folgt, daß die Verletzung einer konjugierten Beschränkung immer schwerer wiegt als die Verletzung der Beschränkungen, aus denen sie entstanden ist.³³ Wichtig für den vorliegenden Fall ist die Annahme, daß lokale Konjunktion reflexiv sein kann, daß man also eine Beschränkung wie BAR mit sich selbst konjugieren kann. Der Bereich D , von dem (66-a) redet, ist für BAR jeweils ein einziges Kettenglied.³⁴ Rekursive lokale Konjunktion mit BAR produziert die folgende Subhierarchie von Beschränkungen mit fixer Ordnung; diese Beschränkungssubhierarchie nennen Legendre, Smolensky & Wilson (1998) die “Subhierarchie der minimalen Kettenglieder” (“MinLink subhierarchy”).

³³Hiermit hat man in der Optimalitätstheorie ein Mittel, um Kumulativitätseffekte zu integrieren: Bei einer Ordnung $B_1 \&_l B_2 \gg B_3 \gg B_1, B_2$ kann zwar direkt keine noch so häufige Verletzung von B_1 und B_2 eine einzige Verletzung von B_3 aufwiegen, indirekt jedoch schon, weil eine gleichzeitige Verletzung von B_1 und B_2 immer auch eine Verletzung der konjugierten Beschränkung $B_1 \&_l B_2$ impliziert. Dies wird bei Fischer (1999) für eine OT-Rekonstruktion der Skopustheorie von Pafel (1993; 1998) ausgenutzt; vgl. Kapitel 4.

³⁴Wäre der Bereich die gesamte Kette, wäre lokale Konjunktion nutzlos, da in diesem großen Bereich K_1 und K_2 aus $T_{3.34}$ wieder denselben Status hätten.

(67) *Subhierarchie der minimalen Kettenglieder:*

- a. $\text{BAR}\&_l\text{BAR} = \text{BAR}^2$:
Ein Kettenglied darf nicht zwei Barrieren überkreuzen.
- b. $\text{BAR}^2\&_l\text{BAR} = \text{BAR}^3$:
Ein Kettenglied darf nicht drei Barrieren überkreuzen.
- c. BAR^n :
Ein Kettenglied darf nicht n Barrieren überkreuzen.
- d. *Fixe Ordnung*:
 $\dots \gg \text{BAR}^3 \gg \text{BAR}^2 \gg \text{BAR}^1$

Wie Tabelle $T_{3.35}$ zeigt, verschwindet nun das Problem, sukzessiv-zyklische Bewegung gegenüber Bewegung in einem Rutsch als besser zu klassifizieren.

$T_{3.35}$: *Eine korrekte Prognose unter der BAR-Subhierarchie*

Kandidaten	BAR^3	BAR^2	BAR^1
$\Leftarrow K_1: \alpha_1 \dots \beta \dots \beta \dots t'_1 \dots \beta \dots t_1$		*	*
$K_2: \alpha_1 \dots \beta \dots \beta \dots \beta \dots t_1$	*!		

Zu dieser Tabelle ist zu bemerken, daß selbstverständlich auch K_2 BAR^2 und BAR^1 verletzt, denn es folgt unmittelbar aus dem Begriff der lokalen Konjunktion, daß BAR^n zu BAR^{n-1} als speziell zu generell in einer Stringenzrelation steht (vgl. (65) aus Kapitel 2). Dies kann jedoch nie eine Rolle spielen, und daher muß man die entsprechenden Verletzungen nicht unbedingt mehr in Tabellen berücksichtigen. Man kann also, um die Tabellen übersichtlicher zu gestalten, annehmen, daß BAR^n nur von Kettengliedern verletzt werden kann, die exakt n Barrieren überschreiten.

Eine weitere Beschränkung, die eine Rolle spielt, ist diese:

(68) REF (“Referentialität”):
Ketten sind referentiell.

REF wird immer von nicht-referentiellen Ketten verletzt. Nicht-referentielle Ketten umfassen Adjunkt-Ketten sowie Ketten von Argumenten, die als nicht-referentiell gelten (vgl. u.a. Rizzi (1990) und Cinque (1990)); von letzterem Fall kann hier abgesehen werden. Die Idee ist nun, daß REF mit den Beschränkungen der Subhierarchie der minimalen Kettenglieder BAR^n , BAR^{n-1} , ..., BAR^1 jeweils durch lokale Konjunktion verbunden werden kann. Es resultiert dann die Subhierarchie der minimalen nicht-referentiellen Kettenglieder in (69):³⁵

³⁵Intuitiv ist einigermaßen klar, wie man z.B. von den durch BAR^1 und REF jeweils ausgedrückten Forderungen zu dem kommt, was die durch lokale Konjunktion entstandene komplexe Beschränkung besagt: Aus “Ein Kettenglied darf keine Barriere überschreiten” und “Eine Kette ist referentiell” wird “Ein Kettenglied einer nicht-referentiellen Kette darf keine Barriere überschreiten”. Es ist jedoch nicht vollkommen offensichtlich, wie sich die Bedeutung der komplexen

(69) *Subhierarchie der minimalen nicht-referentiellen Kettenglieder :*

a. $\text{BAR}^n \&_I \text{REF} = \text{BAR}^{n[-ref]}$:

Ein Kettenglied einer nicht-referentiellen (Adjunkt-) Kette darf nicht n Barrieren überkreuzen.

b. *Fixe Ordnung:*

$\dots \gg \text{BAR}^{3[-ref]} \gg \text{BAR}^{2[-ref]} \gg \text{BAR}^{1[-ref]}$

Jetzt ist klar, wie die oben für das Chinesische diskutierte Beschränkung $\text{BAR}^{2[-ref]}$ durch lokale Konjunktion entsteht, und zwar durch eine Kombination von (a) reflexiver lokaler Konjunktion von BAR mit sich selbst, und dann (b) lokaler Konjunktion von BAR^2 mit REF . Da $\text{BAR}^{n[-ref]}$ aus BAR^n per lokaler Konjunktion abgeleitet ist, gilt schließlich aufgrund von Annahme (66-b) noch universell die folgende fixe Ordnung:

(70) $\text{BAR}^{n[-ref]} \gg \text{BAR}^n$

Der nächste Unterabschnitt illustriert einige Konsequenzen für die Lokalitätstheorie des Englischen.

3.3.3. *W-Bewegung im Englischen*

Legendre, Smolensky & Wilson (1998, 275) schlagen für das Englische die folgende Beschränkungsordnung vor.

Beschränkung in systematischer, kompositioneller Weise aus den Bedeutungen der Teilbeschränkungen ergibt. Nehmen wir zunächst einmal die naheliegende Lösung an und interpretieren die lokale Konjunktion $\text{BAR}^{1[-ref]}$ als logische Konjunktion. Dann resultiert (i-a), mit \wedge als logischem “und”. In der Aussagenlogik ist nun ein Satz der Art $A \wedge B$ wahr genau dann, wenn A und B beide wahr sind; ist einer der beiden Teilsätze falsch, ist schon der Gesamtsatz falsch. Dies bedeutet jedoch, daß logische Konjunktion nicht den intendierten Effekt für die lokale Konjunktion $\text{BAR}^{1[-ref]}$ hat: Wir wollen ja, daß $\text{BAR}^{1[-ref]}$ nur verletzt ist, wenn beide Teilbeschränkungen verletzt sind. Um dieses Resultat zu erreichen, müssen wir lokale Konjunktion tatsächlich wie in (i-b) als logische Disjunktion interpretieren, mit \vee als logischem “oder”.

(i) *Kompositionelle Interpretation von $\text{BAR}^{1[-ref]}$:*

a. [Ein Kettenglied darf keine Barriere überschreiten] \wedge [Eine Kette ist referentiell].

b. [Ein Kettenglied darf keine Barriere überschreiten] \vee [Eine Kette ist referentiell].

Ein Satz der Art $A \vee B$ ist wahr genau dann, wenn wenigstens einer der beiden Teilsätze wahr ist, und falsch folglich nur dann, wenn beide Teilsätze falsch sind. Das ist die intendierte Interpretation für $\text{BAR}^{1[-ref]}$: Diese Beschränkung ist verletzt nur dann, wenn sowohl BAR^1 , als auch REF verletzt ist. Als Ergebnis können wir somit festhalten, daß zumindest im vorliegenden Fall (de facto bei den meisten syntaktischen Anwendungen, die mir bekannt sind) die lokale Konjunktion logisch als Disjunktion zu interpretieren ist. Darüber hinaus ist in jedem Fall eine Spezialannahme für reflexive lokale Konjunktion zu machen, um den Effekt der kumulativen Verletzung ein und derselben Beschränkung zu erfassen.

(71) *Ordnung im Englischen:*

$$\begin{array}{l} \text{SEL} \gg \\ \text{W-TREUE} \gg \\ \text{SKOP-TREUE} \gg \end{array} \quad \begin{array}{l} \text{BAR}^{4[-ref]} \gg \text{BAR}^{3[-ref]} \gg \text{BAR}^4 \gg \text{BAR}^3 \gg \\ \text{BAR}^{2[-ref]} \gg \\ \text{BAR}^2 \gg \text{BAR}^{1[-ref]} \gg \text{BAR}^1 \end{array}$$

Entscheidend ist die relative Ordnung von Treuebeschränkungen (SEL, W-TREUE, SKOP-TREUE) und Lokalitätsbeschränkungen (entweder aus der Subhierarchie der minimalen Kettenglieder oder aus der Subhierarchie der minimalen nicht-referentiellen Kettenglieder).³⁶ Aus (71) folgt, daß kein Kettenglied drei Barrieren überqueren darf, egal ob es sich um eine Argument- oder eine Adjunkt-kette handelt; muß es das doch tun, wird unweigerlich neutralisiert, z.B. dadurch, daß im optimalen Kandidaten W-TREUE verletzt wird, wodurch das W-Merkmal im Index im Output ignoriert wird, das W-Element als Indefinitum erscheint und so die relevante Lokalitätsbeschränkung (BAR³ oder höher) leer erfüllt wird. Weiterhin folgt aus (71), daß Adjunkt-kettenglieder relativ strikte Lokalität erfüllen müssen: Sie dürfen höchstens eine Barriere überqueren. Falls sie mehr Barrieren überqueren müßten, wird neutralisiert, solange dies mit SEL kompatibel ist. Schließlich sind bei Argumentkettengliedern Überquerungen von zwei Barrieren ohne Neutralisierung im Prinzip (aber nicht immer; vgl. Fußnote 36) möglich, da alle zu Neutralisierung führenden Treuebeschränkungen höher geordnet sind.

Inseln haben üblicherweise die Eigenschaft, mehr Struktur bereitzustellen, und damit auch mehr Barrieren. Somit wird in diesen Fällen neutralisiert zugunsten eines Index-untreuen Kandidaten, der entweder W-TREUE verletzt und sein W-Merkmal unrealisiert läßt, oder aber SKOP-TREUE verletzt und engeren Skopus unterhalb der Insel nimmt. Auf dieser Grundlage seien nun einige Beispiele betrachtet.³⁷ (72) zeigt kurze, satzgebundene Bewegung eines Adjunkts und eines Arguments.

- (72) a. How₁ did [_{IP} she [_{VP} do it] t₁] ?
 b. What₁ did [_{IP} she [_{VP} do t₁]] ?

³⁶Man beachte, daß die Ordnung $\text{BAR}^2 \gg \text{BAR}^{1[-ref]}$ kompatibel ist mit den oben gemachten Annahmen über lokale Konjunktionen. Legendre, Smolensky & Wilson (1998, 276) schlagen diese Ordnung vor, um bestimmte superioritätsartige Effekte im Englischen abzuleiten. Der relevante Datenkontrast ist der in (i).

- (i) a. How₁ did [_{IP} he [_{VP} fix what₂] t₁] ?
 b. *What₂ did [_{IP} he [_{VP} fix t₂] how₁] ?

Die Erklärung ist hier, daß (i-a) zwar $\text{BAR}^{1[-ref]}$ verletzt (weil das Adjunkt eine IP-Barriere überschreitet), daß allerdings (i-b) eine fatale Verletzung von BAR^2 aufweist (weil das Argument eine VP- und eine IP-Barriere überschreitet). Dies setzt voraus, daß BAR^2 gegenüber $\text{BAR}^{1[-ref]}$ die höher geordnete Beschränkung ist.

³⁷Diese Fälle werden von Legendre, Smolensky & Wilson (1998) nicht im Detail besprochen, aber ihre Analyse macht klare Vorhersagen.

Unter der Annahme, daß das Adjunkt *how* in einer VP-Adjunktionsposition basiserzeugt ist, verletzt (72-a) $\text{BAR}^{1[-ref]}$ (mit IP als einzig intervenierender Barriere). (72-b) verletzt dagegen BAR^2 . Beide Beschränkungen sind tiefer als alle Treuebeschränkungen geordnet, und daher sind beide Kandidaten in ihren Kandidatenmengen optimal.

In (73) sind ein Adjunkt und ein Argument sukzessiv-zyklisch aus einem eingebetteten Deklarativsatz herausbewegt worden.

- (73) a. How_1 do [_{IP} you [_{VP} think [_{CP} t'_1 that [_{IP} she [_{VP} did it] t_1]]]] ?
 b. What_1 do [_{IP} you [_{VP} think [_{CP} t'_1 that [_{IP} she [_{VP} did t_1]]]]] ?

In (73-a) überquert das Kettenglied $\langle t'_1, t_1 \rangle$ eine Barriere (IP); das Kettenglied $\langle \text{how}_1, t'_1 \rangle$ überquert zwei Barrieren (VP und IP im Matrixsatz). Dies bedeutet, daß die höchst-geordnete Lokalitätsbeschränkung, die (73-a) verletzt, $\text{BAR}^{2[-ref]}$ ist. Zunächst einmal würden wir daher erwarten, daß statt langer Bewegung der optimale Kandidat Neutralisierung per Skopusreduktion aufweist, denn SKOP-TREUE ist ja unterhalb von $\text{BAR}^{2[-ref]}$ geordnet. Ein solcher Ausweg steht jedoch nicht zur Verfügung, denn bei Skopusreduktion wird die undominierte Beschränkung SEL verletzt, wie oben anhand des analogen chinesischen Falles gezeigt wurde. Neutralisierung per W-Merkmalstilgung verstößt gegen W-TREUE; diese Verletzung ist aufgrund der Ordnung $\text{W-TREUE} \gg \text{BAR}^{2[-ref]}$ fatal. Daher ist (73-a) trotz $\text{BAR}^{2[-ref]}$ -Verletzung optimal. (73-b) verletzt demgegenüber bei beiden Kettengliedern ($\langle \text{what}_1, t'_1 \rangle$, $\langle t'_1, t_1 \rangle$) BAR^2 und ist daher von vornherein unproblematisch.

Der Fall liegt anders bei W-Inseln. Bereits eine infinite W-Insel wie in (74-a) führt bei Adjunktextraktion zu Ungrammatikalität (vgl. Chomsky (1986a)); für Argumentextraktion wie in (74-b) sind jedoch infinite W-Inseln transparent.

- (74) a. * How_1 do [_{IP} you [_{VP} wonder [_{CP} t'_1 what [_{IP} PRO to fix t_1]]]] ?
 b. What_1 do [_{IP} you [_{VP} wonder [_{CP} t'_1 when [_{IP} PRO [_{VP} to fix t_1]]]]] ?

Unter der Annahme, daß hier eine Zwischenspur in SpecC etablierbar ist, verletzt die Argumentextraktion in (74-b) im ersten wie im zweiten Schritt BAR^2 , so daß Neutralisierung unterbleibt. Adjunktextraktion verletzt dagegen im ersten Schritt zwar nur $\text{BAR}^{1[-ref]}$, im zweiten Schritt jedoch $\text{BAR}^{2[-ref]}$. Wegen $\text{BAR}^{2[-ref]} \gg \text{SKOP-TREUE}$ wird folglich hin zu (75-a) neutralisiert; der Skopus wird unter Respektierung von SEL vom Matrix- auf den eingebetteten Satz umgestellt. Ein wie in (75-b) neutralisierender weiterer Wettbewerber verletzt fatal W-TREUE.³⁸

³⁸Der Neutralisierungskandidat (75-a) weist eine eingebettete Mehrfachfrage und daher gemäß der für das Englische geltenden Beschränkungsordnung W-in-situ auf. – Ob im übrigen (74-a) zu (75-a) oder (75-b) neutralisiert wird, ist zwar theorieintern relevant, spielt aber für das Ergebnis an sich natürlich keine Rolle: In beiden Fällen resultiert Ungrammatikalität für (74-a), und in beiden Fällen ist der blockierende Kandidat auch Gewinner eines anderen Wettbewerbs, wo er

- (75) a. You wonder [_{CP} how₁ [_{IP} PRO [_{VP} to fix what] t₁]]
 b. You wonder [_{CP} what [_{IP} PRO [_{VP} to fix t] in some way₁]]

Abschließend seien noch Subjekt- und Adjunkt-Inseln betrachtet. Letztere sind für Adjunktketten wie für Argumentketten strikt; vgl. (76).

- (76) a. *How₁ was [_{IP} he [_{VP} fired [_{CP} after behaving t₁]]] ?
 b. *What₁ was [_{IP} he [_{VP} fired [_{CP} after reading t₁]]] ?

Selbst wenn der Adjunktkomplementierer *after* eine besetzbare Spezifikatorposition aufweisen sollte, wird, weil die Adjunkt-CP eine Barriere ist, in (76-a) mindestens $\text{BAR}^{3[-ref]}$ verletzt, in (76-b) mindestens BAR^3 . Neutralisiert wird in diesem Fall zu Kandidaten wie (77-ab), die W-TREUE verletzen und das W-Merkmal unrealisiert lassen; denn eine Neutralisierung per Skopusreduktion verletzt neben tief geordneter SKOP-TREUE auch noch die undominierte Beschränkung SEL.

- (77) a. He was [_{VP} fired [_{CP} after behaving in some way₁]]
 b. He was [_{VP} fired [_{CP} after reading something₁]]

Subjektinseln werden schließlich durch (78-ab) illustriert. Sie sind im Englischen für Adjunkte wie für Argumente strikt (vgl. Cinque (1990)):

- (78) a. *How₁ would [_{IP} [_{CP} t'₁ PRO to behave t₁] be inappropriate] ?
 b. *Who₁ would [_{IP} [_{CP} t'₁ PRO to kiss t₁] be inappropriate] ?

Man möchte hier wohl dieselbe Erklärung geben wie bei Adjunkt-Inseln: (78-ab) überqueren im zweiten Schritt zu viele Barrieren und werden daher zu (79-ab) neutralisiert, unter Inkaufnahme einer nicht-fatalen W-TREUE-Verletzung.

- (79) a. [_{IP} [_{CP} PRO to behave in some way₁] would be inappropriate]
 b. [_{IP} [_{CP} PRO to kiss someone₁] would be inappropriate]

Das Problem ist jedoch, daß es so aussieht, als sei die höchste Lokalitätsbeschränkung, die die Beispiele in (78) verletzen, $\text{BAR}^{2[-ref]}$ bzw. BAR^2 (mit der nicht L-markierten Subjekt-CP sowie der Matrix-IP als Barrieren). Das sollte für Argumentextraktion dann gar keine Konsequenzen haben, und auch Adjunktextraktion sollte möglich sein, weil eine durch die Ordnung $\text{BAR}^{2[-ref]} \gg \text{SKOP-TREUE}$ induzierte Skopusreduktion SEL verletzen würde. Diese Überlegungen legen nahe, daß in (78) mehr als zwei Barrieren durch Kettenbildung überquert werden, d.h., daß Subjektinseln mehr Struktur haben als hier angezeigt. Eine Möglichkeit wäre die Postulierung einer zusätzlichen NP-Schale zwischen CP und IP, wie sie seit Ross (1967) immer wieder einmal vorgeschlagen worden ist. Diese leere NP liefert dann die dritte Barriere. Somit verletzen die Sätze in (78) $\text{BAR}^{3[-ref]}$ bzw. BAR^3 , und es erfolgt beidesmal Neutralisierung durch W-TREUE-Verletzung.

ohne Treueverletzung auskommt.

3.4. Einordnung

Das hier entwickelte System von Lokalitätsbeschränkungen mit durch allgemeinen Annahmen über lokale Konjunktion begründeten fixen internen Ordnungen scheint in Interaktion mit den in der Hierarchie der Lokalitätsbeschränkungen an verschiedenen Stellen eingefügten Treubeschränkungen gut in der Lage, universelle Tendenzen in der Bewegungstheorie zu erfassen. Man kann z.B. ableiten, daß sich Argumente immer leichter bewegen lassen als Adjunkte. Ebenso folgt, daß, je mehr Barrieren überschritten werden, die Wahrscheinlichkeit zunimmt, daß Ungrammatikalität resultiert. Darüber hinaus ist das System flexibel genug, um Variation in den Griff zu bekommen: Die Ordnung der Lokalitätsbeschränkungen zueinander ist zwar (relativ, vgl. Fußnote 36) fix, aber wo in die Hierarchie Treubeschränkungen eingefügt werden, kann von Sprache zu Sprache variieren. Das Konzept der Neutralisierung als Mittel, absolute Ungrammatikalität in OT abzuleiten auch in Fällen, wo keine offensichtliche Gewinnerform im selben Wettbewerb aufzutreten scheint (die nicht auch Gewinner eines anderen, "einfacheren" Wettbewerbes ist), ist zumindest theorieintern überzeugend.

Potentielle Probleme wirft die Analyse dagegen in den folgenden vier Bereichen auf. Erstens hat sich gezeigt, daß es im Einzelfall bei der Lokalitätstheorie stark auf spezielle Annahmen über die Satzstruktur und Ähnliches ankommt. Zweitens muß wirklich jede überschrittene Barriere gezählt werden, was unter Umständen in einer kognitiv ausgerichteten Theorie als problematisch angesehen werden kann; dies geht darauf zurück, daß das Konzept der Kumulativität in die Optimalitätstheorie integriert worden ist.³⁹ Drittens ist festzuhalten, daß gerade die Lokalitätstheorie ein Bereich ist, in dem dafür argumentiert wurde, daß Grammatikalitätsgrade eine Rolle spielen sollten; vgl. etwa die Idee, daß Subjazenverletzungen "schwach" und ECP-Verletzungen "stark" sind. Wie in Kapitel 2 festgestellt wurde, ist es aber eine allgemeine Eigenschaft der optimalitätstheoretischen Syntax, graduelle (Un-)Grammatikalität nicht zu erfassen. (Allerdings erfolgt die Behandlung dieses Phänomens in anderen Syntaxtheorien, wie ebenfalls in Kapitel 2 gesehen, auch nur rein stipulativ.) Viertens schließlich sei auf eine konzeptuelle Merkwürdigkeit hingewiesen. Der Mechanismus der lokalen Konjunktion von Beschränkungen ist aufgrund seiner Rekursivität zwar elegant, aber auch sehr mächtig. Ohne weitere Restriktionen gibt man damit eine elementare Annahme über Grammatiken auf, nämlich die, daß die Menge der Beschränkungen in einer Grammatik finit ist. Nicht-finite Beschränkungsmengen sind nun aufgrund der für lokale Konjunktionen geltenden fixen Ordnungen vielleicht empirisch kein unüberwindliches Problem. Sie lassen aber die Komplexität des Gesamtsystems weiter ansteigen. Wie wir in Kapi-

³⁹Wie Prince & Smolensky (1993) ausführen, gilt dieser Einwand nicht für die Optimalitätstheorie an sich: Unterschieden werden muß bei zwei Kandidaten jeweils immer nur, welcher mehr bzw. weniger Verletzungen einer Beschränkung aufweist; die Kenntnis der absoluten Zahl der Verletzungen einer Beschränkung spielt de facto nie eine Rolle.

tel 2 gesehen haben, gilt dasselbe für den Mechanismus der Neutralisierung. Somit erweist sich, daß die beiden Kernstücke des Ansatzes von Legendre, Smolensky & Wilson (1998) – lokale Konjunktion und Neutralisierung – zwar zu einer überzeugenden Theorie der Lokalität von Ketten führen, dies aber nicht umsonst: Der Preis ist ein Komplexitätsanstieg.

Kapitel 4

Vorgänger und Alternativen

Die zentralen Annahmen der Optimalitätstheorie sind in (1) wiederholt.

- (1) a. *Universalität:*
Beschränkungen sind universell.
- b. *Verletzbarkeit:*
Beschränkungen können verletzt werden.
- c. *Geordnetheit:*
Beschränkungen sind geordnet.
- d. *Wettbewerb:*
Die Grammatikalität eines Kandidaten K ist nicht allein aufgrund interner Eigenschaften von K ermittelbar; vielmehr entscheiden externe Faktoren (der Wettbewerb von K mit anderen Kandidaten) über die Wohlgeformtheit von K.

Diese Annahmen sind für sich genommen nicht neu: Jede ist in der ein oder anderen Form in der syntaktischen Literatur bereits vor dem Aufkommen der Optimalitätstheorie schon einmal gemacht worden. Das Spezielle und Neue an der Optimalitätstheorie ist, daß diese vier Annahmen zusammengebracht werden. Im folgenden sollen diese Annahmen anhand von prä-optimalitätstheoretischen Analysen untersucht werden (und zwar mit unterschiedlicher Ausführlichkeit, die in (1) von oben nach unten zunimmt); es soll gezeigt werden, wieso diese Arbeiten noch keine genuine optimalitätstheoretischen Analysen darstellen; und es soll zum Teil auch gezeigt werden, wie die jeweiligen Analysen optimalitätstheoretisch erfaßt werden können.

1. Universalität

Das Postulat der Universalität syntaktischer Beschränkungen ist zum einen unter den vier Konzepten in (1) das am wenigsten wichtige (im Prinzip funktioniert eine optimalitätstheoretische Analyse in einer Einzelsprache auch bei Annahme der Sprachspezifität der geordneten und verletzbaren Beschränkungen). Zum anderen ist es aber zugleich auch das am wenigsten kontroverse. So ist etwa in der GB- bzw. Prinzipien-und-Parameter-Theorie von Anfang an versucht worden, die Zahl sprachspezifischer Filter zu minimieren (bzw. auf solche Filter ganz zu verzichten). Eine andere Quelle der Nicht-Universalität syntaktischer Beschränkungen ist das klassische Konzept der Parametrisierung. Hier ist die Standardannahme in der GB-Theorie immer gewesen, daß die Beschränkungen zwar an sich universell sind, daß

sie jedoch offene Stellen enthalten, deren unterschiedliche Füllung bzw. Belegung Sprachvariation abzuleiten gestattet. Dies bedeutet jedoch letztlich, daß unterschieden werden muß zwischen einer universellen "Proto"-Beschränkung und ihren durch Parameterfixierung gewonnenen sprachspezifischen Realisierungen. Letztere sind die eigentlichen Beschränkungen einer einzelsprachlichen Syntax, und in diesem Sinne sind dann die syntaktischen Beschränkungen nicht universell, auch wenn sie Ausprägungen einer universellen Proto-Beschränkung mit offenen Parametern sind.

Es hat sich aber über die Jahre innerhalb der Prinzipien-und-Parameter-Theorie immer mehr die Auffassung durchgesetzt, daß auch eigentliche Beschränkungen im obigen Sinne universell gültig und nicht durch von Sprache zu Sprache variierende Parameterfixierung unterschieden sind; vielmehr beziehen sich "Parameter" eher auf Klassen von sprachlichen Ausdrücken. Diese Tendenz läßt sich gut anhand der Erklärung der Variation bei der Lizenzierung von anaphorischen Elementen (Reflexiv- bzw. Rezipropronomen) illustrieren.

Betrachtet man Reflexivpronomen im Deutschen und Isländischen im Vergleich, so sieht es zunächst einmal so aus, als müßte das für Prinzip A der Bindungstheorie wichtige Konzept der bindenden Kategorie (bzw. regierenden Kategorie) unterschiedlich festgelegt, also parametrisiert, werden, und zwar dergestalt, daß bindende Kategorien im Isländischen größer sind als im Deutschen. Denn wie (2-a) im Deutschen gegenüber (2-b) im Isländischen zeigt, sind Kontrollinfiniteive zwar bindende Kategorien für die Anapher *sich* im Deutschen, aber nicht für die Anapher *sig* im Isländischen.

- (2) a. Maria₁ befahl Fritz₂ [_{CP} PRO₂ sich_{*1/2} umzuziehen]
 b. Maria₁ skipaði Harald₂ [_{CP} að PRO₂ raka sig_{1/2}]
 Maria befahl Harald zu waschen sich

Das Matrixprädikat *befehlen* ist ein Objektkontrollverb, und somit kann im Deutschen wie im Isländischen ein Reflexivum innerhalb des Kontrollinfiniteivs an PRO gebunden werden, was wiederum mit dem Matrixobjekt koindiziert ist. Koindizierung der Anapher mit dem Matrixsubjekt ist dagegen zwar im Isländischen möglich, nicht jedoch im Deutschen.

Es gibt nun aber gute Gründe, die dagegen sprechen, diesen Unterschied durch sprachspezifische Parametrisierung einer syntaktischen Beschränkung zu erfassen. Wie Manzini & Wexler (1987) beobachten, können vergleichbare Unterschiede zwischen zwei Typen von Anaphern auch innerhalb einer einzigen Sprache auftreten, so daß eine Parametrisierung von Sprache zu Sprache nicht weiterhilft. Diese Situation liegt im Italienischen vor, wo es die beiden Reflexiv-Typen *sè* und *se stesso* gibt. Beide Reflexivpronomen müssen zwar innerhalb des nächsten finiten Satzes gebunden sein, und (anders als im Isländischen) auch innerhalb des nächsten Kontrollinfiniteivs; während allerdings auch A.c.I.-Infiniteive (hier konkret Infiniteive, die durch Perzeptionsverben eingebettet werden) für *se stesso* eine bindende Kategorie errichten, innerhalb derer diese Anapher Prinzip A erfüllen muß, sind A.c.I.-Infiniteive für lange Bindung von *sè* transparent. Dies zeigen die Daten in (3)

(vgl. Manzini & Wexler (1987, 416 & 423)).

- (3) a. Mario₁ vide [_{IP} Gianni₂ colpire se stesso_{*1/2}]
 Mario saw Gianni hit himself
 b. Alice₁ vide [_{IP} Mario₂ guardare sè_{1/2} nello specchio]
 Alice saw Mario look at himself in the mirror

Manzini und Wexler schließen hieraus, daß man für die Bindungstheorie keine Parametrisierung von Grammatik zu Grammatik annehmen sollte, sondern nur von lexikalischem Element zu lexikalischem Element; diese Annahme geht ursprünglich auf Borer (1984) zurück und ist auch als “Hypothese der Lexikalischen Parametrisierung” bekannt. Dieser Hypothese zufolge betrifft also die Variation lediglich unterschiedliche Subklassen von lexikalischen Elementen, und das einzige, was ein Kind im Erstspracherwerb lernen muß, ist, zu welcher Subklasse von anaphorischen Elementen ein gegebenes Reflexivpronomen zählt: *Sich* im Deutschen und *se stesso* im Italienischen zählen zu der Subklasse von Anaphern, die die kleinste Bindungsdomäne haben, *sè* im Italienischen hat eine etwas größere Bindungsdomäne, und *sig* im Isländischen eine noch größere.

Es ist des weiteren der Versuch unternommen worden, aus der morphologischen Form des Reflexivpronomens die Zugehörigkeit zu einer der Subklassen unmittelbar abzuleiten (ein wesentlicher Faktor scheint z.B. zu sein, ob das Reflexivum komplex ist, also noch einen *selbst*-Teil aufweist, oder nicht). Dies soll uns an dieser Stelle jedoch nicht interessieren. Wichtig ist allein die folgende Konsequenz: Unter der Voraussetzung, daß Evidenz gegen eine Parametrisierung syntaktischer Beschränkungen von Sprache zu Sprache nicht nur auf den Bereich der Anaphern beschränkt ist, sondern auch in anderen Bereichen der Syntax angeführt werden kann, läßt sich schließen, daß kein Grund vorliegt, nicht durchweg von ganz allgemeinen, universon Beschränkungen auszugehen.¹

¹Eine Diskussion aller einschlägigen empirischen Bereiche würde hier zu weit führen. Verwiesen sei lediglich noch auf einen besonders offensichtlichen Fall, den der Subjazenbeschränkung. Rizzi (1982) hatte argumentiert, daß Chomskys (1977) Subjazenbeschränkung parametrisiert ist: Die Beschränkung besagt, daß Bewegung nicht zwei Grenzknoten überschreiten darf, und als Grenzknoten zählen im Englischen (in modernerer Terminologie) NP und IP, im Italienischen dagegen NP und CP. Dies erklärt nach Rizzi, warum W-Inseln im Italienischen nicht strikt Extraktion blockieren. Die Parametrisierung der Subjazenbeschränkung galt eine zeitlang als der Inbegriff einer erfolgreichen Anwendung des Konzepts der Parametrisierung syntaktischer Beschränkungen, spielt jedoch zum gegenwärtigen Zeitpunkt in der Forschung kaum noch eine Rolle.

Zum einen wird spätestens seit Chomsky (1986a) versucht, unabhängige Eigenschaften einer XP dafür verantwortlich zu machen, ob sie ein Grenzknoten (bzw. eine Barriere) für Bewegung ist oder nicht (eine Grenzknoten-auslösende Eigenschaft einer XP scheint z.B. die zu sein, nicht L-markiert zu werden – vereinfacht: nicht Schwester eines lexikalischen Kopfes zu sein). Unter dieser Perspektive gibt es keinen Platz für eine Parametrisierung des Grenzknotenbegriffs, und im Einklang damit stellt Chomsky (1986a, 37) fest, daß eine parametrisierte Subjazenbeschränkung mit “zusätzlichen [d.h., nicht von der Theorie vorausgesagten] Grenzknoten” arbeiten müßte, “over and above the system just outlined”.

Syntaktische Beschränkungen gelten unter dieser Perspektive in der Prinzipien- und-Parameter-Theorie generell, sie können jedoch empirisch ohne Konsequenzen bleiben, wenn in einer Sprache kein geeignetes lexikalisches Element vorhanden ist, das durch sie beschränkt wird.

Der Schluß aus diesen Überlegungen kann somit nur lauten, daß die Annahme der Universalität syntaktischer Beschränkungen im H-Eval-System einer Optimalitätstheoretischen Grammatik keineswegs einen radikalen Bruch mit der GB-orientierten Forschung darstellt; vielmehr handelt es sich hier um eine naheliegende, konsequente Weiterentwicklung.

2. Verletzbarkeit: Uszkoreits Wortstellungsanalyse

In Abschnitt 5.1 von Kapitel 2 ist bereits bemerkt worden, daß die Beschränkungen, die die Abfolge von Nominalen im deutschen Mittelfeld steuern, miteinander im Konflikt stehen, weil sie manchmal Unvereinbares von einem Satz verlangen. Um diesem Sachverhalt gerecht werden zu können, entwirft Uszkoreit (1984; 1986) ein Modell der Wortstellung im deutschen Mittelfeld, das systematisch mit verletzbaaren Beschränkungen arbeitet. Konkret nimmt Uszkoreit die folgenden Linearisierungsregeln an:

- (4) a. [+Nominativ] vor [+Dativ]
 b. [+Nominativ] vor [+Akkusativ]
 c. [+Dativ] vor [+Akkusativ]
 d. [-Fokus] vor [+Fokus]
 e. [+Pronomen] vor [-Pronomen]

Diese Linearisierungsbeschränkungen müssen im Prinzip verletzbar sein. Wenn z.B. in einem Satz mit einem transitiven Verb das Akkusativargument pronominal ist, das Nominativargument dagegen nicht pronominal, ergibt sich sofort ein Beschränkungskonflikt, und da ein entsprechender Satz nicht notwendigerweise ungrammatisch ist, muß wenigstens eine der beiden involvierten Linearisierungsregeln verletz-

Zum anderen hat sich herausgestellt, daß Englischsprecher zu den von Rizzi (1982) ins Feld geführten (relativ) wohlgeformten italienischen Daten vollständig analoge Beispiele oft ebenso gut bewerten (vgl. Chomsky (1986a, 37) und Grimshaw (1986, 365)); umgekehrt erweisen sich viele ungrammatische Daten, die ursprünglich die Subjanzbeschränkung im Englischen motiviert haben, bei direkter Umsetzung ins Italienische als ebenso schlecht (dies beobachtet z.T. schon Rizzi (1982) selbst). In Müller & Sternefeld (1993, 494) wird dementsprechend erwogen, daß das, was zunächst einmal wie ein Unterschied zwischen Italienisch und Englisch (bzw. Deutsch) aussieht, bei genauerem Hinsehen tatsächlich ein Unterschied zwischen verschiedenen Konstruktionstypen (z.B. Relativierung und Topikalisierung vs. W-Bewegung) sein könnte. Wenn dem so ist, ergeben sich für die Variation bei Extraktionsoptionen ganz ähnliche Konsequenzen wie im Text für die Variation bei Anapherbindung dargestellt: Die Beschränkungen für Extraktion sind in allen Sprachen dieselben; die relevanten Unterschiede betreffen lediglich die Natur des extrahierten Elements und seine Landestelle (z.B.: Topik und Topikposition vs. W-Phrase und W-Position).

bar sein. Uszkoreit (1984) macht nun die folgenden drei Annahmen:

- (5) a. Ein Satz ist (bzgl. der Linearisierungsbeschränkungen in (4)) wohlgeformt, wenn jeder Verstoß gegen eine der Linearisierungsbeschränkungen bei einem NP-Paar $\langle NP_1, NP_2 \rangle$ sanktioniert wird durch die Erfüllung einer anderen Linearisierungsbeschränkung beim selben NP-Paar $\langle NP_1, NP_2 \rangle$. Ist dies nicht der Fall, resultiert Ungrammatikalität.
- b. Es gibt Grade der Akzeptabilität: Ein gemäß (5-a) wohlgeformter Satz wird in dem Maße besser, in dem er mehr der Beschränkungen in (4) erfüllt.
- c. Die Beschränkungen sind untereinander nicht gewichtet; relevant ist nur die Anzahl der Verletzungen dieser Beschränkungen, nicht die Qualität.

Auf der Basis dieser Annahmen betrachte man nun die vier Sätze in (6), die sich nur hinsichtlich (a) der Realisierung des Fokus (genauer: des Fokusexponenten) auf direktem bzw. indirektem Objekt und (b) der linearen Abfolge der beiden Objekte unterscheiden (Fokussierung wird durch Kapitälchen angezeigt).

- (6) a. Dann will der Doktor dem Mann die PILLE geben
 [+Nom,-Fokus] [+Dat,-Fokus] [+Akk,+Fokus]
- b. Dann will der Doktor dem MANN die Pille geben
 [+Nom,-Fokus] [+Dat,+Fokus] [+Akk,-Fokus]
- c. Dann will der Doktor die Pille dem MANN geben
 [+Nom,-Fokus] [+Akk,-Fokus] [+Dat,+Fokus]
- d.??Dann will der Doktor die PILLE dem Mann geben
 [+Nom,-Fokus] [+Akk,+Fokus] [+Dat,-Fokus]

Zunächst einmal ist in (6-a) keine der Linearisierungsbeschränkungen in (4) verletzt; insbesondere entspricht die Abfolge Dativobjekt vor Akkusativobjekt sowohl der Dativregel (4-c) als auch der Fokusregel (4-d). Da hier auch nicht gegen irgendeine andere Grammatikregel verstoßen wird, ist (6-a) zunächst einmal gemäß Annahme (5-a) grammatisch wohlgeformt; und da es bzgl. der Erfüllung der Linearisierungsbeschränkungen keinen besseren Kandidaten geben kann, ist (6-a) somit gemäß Annahme (5-b) auch maximal unmarkiert.² Im Gegensatz zu (6-a) enthalten (6-b) und (6-c) jeweils einen Verstoß gegen eine Linearisierungsbeschränkung, der lokal sanktioniert wird: Die Abfolge Dativobjekt vor Akkusativobjekt in (6-b) verletzt zwar die Fokusregel, wird aber durch die Erfüllung der Kasusregel lizenziert, und die Abfolge Akkusativobjekt vor Dativobjekt in (6-c) verletzt zwar die Kasusregel, wird aber durch die Erfüllung der Fokusregel erlaubt. Beide Sätze werden vom Uszkoreitschen System somit als grammatisch eingestuft; sie bekommen

²An dieser Stelle mag ein intuitives Verständnis der Grade von Markiertheit bzw. Akzeptabilität von an sich grammatischen Sätzen ausreichen; aber das Konzept der (Un-) Markiertheit von Sätzen ist nicht zu verwechseln mit dem Konzept des unmarkierten Falles. Vgl. Kapitel 6.

jedoch gemäß Annahme (5-b) einen intermediären Status bzgl. der Akzeptabilität bzw. Markiertheit zugewiesen. Die Abfolge Akkusativobjekt vor Dativobjekt in (6-d) schließlich wird durch keine der Linearisierungsbeschränkungen sanktioniert, weil hier sowohl die Fokus-, als auch die Dativregel verletzt wird (und von den beiden Objekte auch keine andere Linearisierungsbeschränkung zum Ausgleich erfüllt wird). Gemäß (5-a), (5-b) wird (6-d) daher nicht nur als noch stärker markiert, sondern sogar als ungrammatisch klassifiziert (vgl. Uszkoreit (1984, 183)).

Dieses Ergebnis ist nun allerdings nicht ganz unproblematisch, denn im Vergleich mit echt ungrammatischen Sätzen im Deutschen schneidet (6-d) fraglos noch ganz gut ab. Uszkoreit (1986, 896-899) (vgl. z.T. auch schon Uszkoreit (1984, 189)) nimmt dieses Problem zur Kenntnis. Er schlägt vor, daß auf Annahme (5-c) doch verzichtet werden sollte, d.h., daß die Beschränkungen für die lineare Abfolge im deutschen Mittelfeld doch gewichtet sind. Konkret sei die Fokusregel die schwächste Linearisierungsbeschränkung in (4), die Pronomenregel die stärkste, und die Kasusregeln seien dazwischen angesiedelt. Zusammen mit der Zusatzannahme, daß "eine Verletzung aller anwendbaren Linearisierungsregeln ... marginal akzeptabel ist, wenn die Verletzung einer der schwächer gültigen Linearisierungsregeln attribuiert werden kann" (Uszkoreit (1986, 897)), läßt sich so die eigentlich von der Theorie vorausgesagte vollständige Ungrammatikalität eines Beispiels wie (6-d) vermeiden.³

Ein anderes Beispiel sei noch betrachtet. In (6-a) wird keine Linearisierungsregel verletzt, und das Ergebnis ist maximale Unmarkiertheit. Ein bißchen anders liegt der Fall beim ebenfalls als perfekt eingestuftem (7):

- (7) Dann hatte ihm ein Freund den BALL gegeben
 [+Dat,+Prn] [+Nom,-Prn] [+Akk,-Prn]

Die Abfolge Dativobjekt vor Nominativobjekt verstößt gegen die Kasusregel, wird aber erlaubt durch die Erfüllung der Pronomenregel. Mit einer vollkommen analogen Argumentation läßt sich nun aber auch der Satz in (8) als vom System in Uszkoreit (1984) als vollkommen grammatisch klassifiziert erweisen:

- (8):*Dann hatte ein Freund dem Fritz ihn gegeben
 [+Nom,-Prn] [+Dat,-Prn] [+Akk,+Prn]

Die Abfolge Dativobjekt vor Akkusativobjekt verstößt gegen die Pronomenregel, wird aber durch die Erfüllung der Kasusregel lokal sanktioniert. Das Problem ist dabei, daß (8) gegenüber (7) stark markiert, wenn nicht gar ungrammatisch ist (zumindest scheint das der Fall zu sein, wenn *ihn* unbetont ist). Wiederum suggeriert dies, daß entgegen der Analyse in Uszkoreit (1984) auch die Qualität von Beschrän-

³Die genuin optimalitätstheoretischen Analysen von Choi (1996; 1999) und Büring (1997; 1999) machen ebenfalls zunächst einmal die Vorhersage, daß (6-d) ungrammatisch ist. Die Versuche von Choi und Büring, dieses Problem zu lösen, werden in Kapitel 6 besprochen.

kungsverletzungen eine Rolle spielt, derart, daß im vorliegenden Fall die Erfüllung der Pronomenregel wichtiger ist als die Erfüllung der Kasusregel. Die Aufgabe der Annahme (5-c) in Uszkoreit (1986) bei gleichzeitiger stärkerer Gewichtung der Pronomenregel gegenüber den Kasusregeln liefert dann auch für diesen Fall empirisch bessere Ergebnisse.

Zusammenfassend läßt sich festhalten, daß die Uszkoreitsche Theorie der freien Wortstellung im Deutschen ganz wesentlich auf der Idee basiert, daß die Linearisierungsbeschränkungen, die die Abfolge determinieren, verletzbar sind: Ein Satz kann grammatisch sein, obwohl er gegen eine oder mehrere Beschränkungen verstößt. Dies ist ganz im Sinne der Optimalitätstheorie. Im Unterschied zur Optimalitätstheorie nimmt Uszkoreit jedoch auch noch an, daß die Zahl von Beschränkungsverletzungen ein wesentlicher Faktor ist, und (zumindest in der ursprünglichen Version dieser Theorie) daß die Qualität von Beschränkungsverletzungen keine Rolle spielt. Letztere Annahme wird dann jedoch tentativ (Uszkoreit (1984)) bzw. explizit (Uszkoreit (1986)) im Lichte empirischer Gegenevidenz verworfen, und dies ist dann bereits ein weiterer Schritt hin zu einer genuin optimalitätstheoretischen Analyse. Obwohl noch weit mehr zum Zusammenhang von deutscher Wortstellung und verletzbaren Beschränkungen zu sagen wäre, mag das an dieser Stelle genügen. Ich werde in Kapitel 6 auf dieses Thema zurückkommen und mich jetzt einer Analyse des relativen Skopus im Deutschen zuwenden, die nicht nur eine Verletzbarkeit von Beschränkungen postuliert, sondern auch deren relative Gewichtung.

3. Geordnetheit: Pafels Skopusanalyse

Wie schon erwähnt, präsupponiert die Annahme, daß Beschränkungen geordnet sind, de facto die Annahme, daß Beschränkungen verletzbar sind: Ist nämlich eine Beschränkung nicht verletzbar, so resultiert ein Verstoß gegen diese Beschränkung bereits in Ungrammatikalität, und eine Ordnung unter den Beschränkungen bleibt empirisch leer.⁴

Eine Theorie, die auf verletzbare und zueinander gewichtete Beschränkungen

⁴In der GB-Theorie wurde im allgemeinen angenommen, daß Verstöße gegen bestimmte Beschränkungen (wie z.B. in manchen Fällen die Subjazenbedingung, vgl. (i-ab)) weniger starke Ungrammatikalität zur Folge haben als Verstöße gegen bestimmte andere Beschränkungen (wie z.B. das ECP oder die Adjunkt-Insel-Beschränkung, vgl. (i-c) und (i-d) = (72) aus Kapitel 2).

- (i) a. ??Radios₁ weiß ich nicht [_{CP} wie₂ man t₂ t₁ repariert]
 b. ??What₁ do you wonder [_{CP} whether John likes t₁] ?
 c. *Who₁ do you think [_{CP} t'₁ that [_{IP} t₁ left]] ?
 d. *Was₁ ist Fritz eingeschlafen [_{CP} nachdem er t₁ gelesen hat] ?

Es ergäbe jedoch wenig Sinn, von einer Ordnung der Art ECP ≫ Subjazen zu sprechen, denn die beiden Beschränkungen interagieren (konfigurieren) ja nicht. Vielmehr sind hier per Stipulation mit Verletzungen von bestimmten Beschränkungen bestimmte Grade der Ungrammatikalität assoziiert; vgl. Abschnitt 8.2 von Kapitel 2.

Bezug nimmt, ist die in Pafel (1993; 1998) und d’Avis & Pafel (1992) entwickelte Theorie des relativen Skopus von quantifizierten Elementen (die folgende Darstellung orientiert sich an Pafel (1993)). Die Grunderkenntnis ist, daß verschiedene Faktoren den relativen Skopus determinieren, und zwar u.a.:⁵

- (9) a. Syntaktische Konstellation
- b. Grammatische Funktion
- c. Grad der Distributivität
- d. Grad der Spezifität

Das Verhalten eines quantifizierten Ausdrucks (QA) zu jedem dieser Faktoren wird durch einen Skopuswert gemessen. QA erhält bei (9-a) je nach syntaktischer Konstellation einen hohen oder niedrigen Skopuswert; mit “syntaktischer Konstellation” ist im wesentlichen die lineare Abfolge gemeint.⁶ Bei (9-b) richtet sich die Höhe des Skopuswertes nach der grammatischen Funktion, und zwar gemäß einer Hierarchie *Subjekt* \succ *Indirektes Objekt* \succ *Präpositionales Objekt* \succ *Direktes Objekt*, mit nach unten absteigendem Skopuswert. Je nach Grad der Distributivität erhält ein QA gemäß (9-c) einen hohen oder niedrigen Skopuswert bei diesem Faktor. Analoges gilt für den Grad der Spezifität gemäß (9-d). Die von einem QA angehäuften Skopuswerte werden nunmehr addiert und – zur Ermittlung des relativen Skopus – mit dem Gesamtwert eines anderen im selben Satz auftauchenden QA verglichen: Wer von zwei gegebenen QAs mehr Punkte gesammelt hat, gewinnt gegen den anderen. Ist der Abstand vom Gewinner QA_i auf den Verlierer QA_j hinreichend groß, hat QA_i weiten Skopus über QA_j; ist der Abstand nur gering, resultiert Ambiguität, allerdings mit Präferenz für den Gewinner. Einige Beispiele mögen die Arbeitsweise dieses Systems verdeutlichen. Man betrachte zunächst (10):

- (10) Ein Mann₁ liebt jede Frau₂

Es gibt hier einen Gesamtskopuswertvorsprung von QA₁ gegenüber QA₂ (der darauf beruht, daß QA₁ bei (9-a) und (9-b) einen Skopuswert anhäuft, der größer ist als der von QA₂ bei (9-c) aufgrund der inhärent gegebenen Distributivität von *jede* zusammengekommene); dieser Vorsprung ist jedoch nicht so groß, daß er entscheidend ist und zu Skopuseindeutigkeit führt. Somit resultiert hier Ambiguität, allerdings mit leichter Präferenz für weiten Skopus von QA₁. Der nächste Satz ist strukturell ganz ähnlich; nur sind All- und Existenzquantor vertauscht.

⁵Ein wesentlicher Faktor, der bei der Betrachtung außen vor gelassen wird, ist noch die Intonation. Darüber hinaus ist festzuhalten, daß das Folgende eine Vereinfachung der Pafelschen Analyse darstellt.

⁶Ich abstrahiere von der von Pafel gemachten weitergehenden Unterscheidung zwischen Präzedenz aus dem Vorfeld und Präzedenz im Mittelfeld. Letztere wird bei ihm als noch stärker Skopus-determinierend klassifiziert. Es mag aber sein, daß hier noch weitere Faktoren eine Rolle spielen: Ein spezielles Intonationsmuster (“Hut-Kontur”, “I-Topikalisierung”), das oft Skopusinversion erzwingt, ist wesentlich leichter im Vorfeld zu realisieren als im Mittelfeld.

(11) Jeder von uns₁ hat eines der Bücher₂ gelesen

Jetzt sammelt QA₁ nicht nur bei (9-a) und (9-b) Punkte, sondern auch noch bei (9-c). Der Skopuswertvorsprung von QA₁ gegenüber QA₂ wird somit entscheidend, und der Satz wird nicht mehr als ambig eingestuft, sondern als skopuseindeutig: QA₁ hat weiten Skopus gegenüber QA₂. Ein letztes hier zu diskutierendes Beispiel ist (12):

(12) Eines der Bücher₁ hat jeder von uns₂ gelesen

In diesem Fall ist der Skopuswertabstand von QA₁ und QA₂ wiederum klein (QA₁ schneidet bei der syntaktischen Konstellation (9-a) besser ab, QA₂ dagegen bei der grammatischen Funktion (9-b) und bei der Distributivität (9-c)). Daher resultiert Ambiguität. Allerdings, und dies ist im gegenwärtigen Zusammenhang besonders interessant, ergibt sich ein leichter Vorteil für die Lesart, derzufolge QA₂ weiten Skopus über QA₁ hat, abweichend von der syntaktischen Oberflächenabfolge.

Auf den ersten Blick sieht es so aus, als könnten die Faktoren in (9) als Beschränkungen aufgefaßt werden, wie z.B. in (13):

- (13) a. SYN-KON (“Syntaktische Konstellation”):
Wenn QA₁ QA₂ auf der S-Struktur vorangeht, hat QA₁ Skopus über QA₂.
- b. GRAM-FUNK (“Grammatische Funktion”):
Wenn QA₁ QA₂ gemäß der Hierarchie der grammatischen Funktionen dominiert, hat QA₁ Skopus über QA₂.
- c. DIST (“Distributivität”):
Wenn QA₁ QA₂ gemäß einer unabhängig gegebenen Distributivitätsskala dominiert, hat QA₁ Skopus über QA₂.
- d. SPEZ (“Spezifizität”):
Wenn QA₁ spezifischer ist als QA₂, hat QA₁ Skopus über QA₂.

Diese Beschränkungen müssen im Prinzip verletzbar sein; denn anders wäre z.B. die Option von weitem Skopus von QA₂ gegenüber QA₁ in (12), die SYN-KON verletzt, nicht zu erklären. Analoges gilt für (10): Neben wiederum der nicht der Oberflächenreihenfolge entsprechenden reversen Lesart, die SYN-KON verletzt, kann QA₁ auch Skopus über QA₂ haben, obwohl dies DIST verletzt (*jede Frau* ist höher auf der Distributivitätsskala als *ein Mann*).

Der entscheidende Punkt ist nun, daß Pafel argumentiert, daß bestimmte Faktoren (bzw. reinterpretiert: bestimmte Beschränkungen) bei der Gesamtberechnung des relativen Skopus stärker gewichtet werden als andere. M.a.W.: Die Faktoren bzw. Beschränkungen sind geordnet. Vereinfacht ergibt sich die folgende Ordnung:⁷

⁷Konkret schlägt Pafel (1993) vor, den Skopuswert für Syntaktische Konstellation bei der Ermittlung des relativen Skopus immer zu verdoppeln, den für Spezifizität zu dritteln, und die anderen beiden gleich zu lassen.

- (14) *Ordnung der Beschränkungen für relativen Skopus* :
 SYN-KON \gg GRAM-FUNK, DIST \gg SPEZ

Dies bedeutet, daß die lineare Abfolge der wichtigste Faktor bei der Determination des relativen Skopus ist, und daß der Grad der Spezifität eines QA am unwichtigsten dafür ist. Diese Analyse mit verletzbaren und zueinander geordneten Beschränkungen kommt einem optimalitätstheoretischen System zwar schon nahe, sie ist aber noch keines. Der Unterschied ist, daß die Faktoren in (9) bei Pafel eben gerade *nicht* wie in (13) als direkte Beschränkungen der Grammatik aufgefaßt werden. Die relevante Beschränkung der Grammatik ist im Pafelschen System vielmehr (15):

- (15) SKOPUSBESCHRÄNKUNG:
- a. Wenn QA_1 einen wesentlich höheren Gesamtskopuswert hat als QA_2 , dann hat QA_1 Skopus über QA_2 .
 - b. Wenn QA_1 und QA_2 eng bei einander liegende Gesamtskopuswerte haben, dann kann QA_1 Skopus über QA_2 haben oder QA_2 Skopus über QA_1 ; präferiert ist dabei der relative Skopus, der einem absteigenden Gesamtskopuswert entspricht.

Die jeweiligen Faktoren (syntaktische Konstellation, grammatische Funktion, Distributivität, Spezifität) sind in diesem System also nur Hilfsbegriffe, die der Ermittlung des einzig relevanten Gesamtskopuswertes eines QA dienen. Dementsprechend ist eine direkte optimalitätstheoretische Übertragung dieser Analyse auch keineswegs unmittelbar auf der Hand liegend. Will man dies trotzdem versuchen und – wie hier angedeutet – die Faktoren als Beschränkungen reinterpretieren, so stellt sich heraus, daß die zentrale optimalitätstheoretische Annahme über den Wettbewerb aufgegeben werden muß, nämlich das Konzept der Grammatikalität selbst, das auf dem der Optimalität beruht; vgl. (4) und (5) aus Kapitel 1, beides hier in (16) wiederholt.

- (16) a. *Grammatikalität*:
 Ein optimaler Kandidat aus der Kandidatenmenge ist grammatisch, alle nicht-optimalen Kandidaten sind ungrammatisch.
- b. *Optimalität*:
 Ein Kandidat K_i ist optimal hinsichtlich einer Beschränkungsordnung $\langle B_1 \gg B_2 \gg \dots \gg B_n \rangle$ gdw. es keinen anderen Kandidaten K_j in derselben Kandidatenmenge gibt, der ein besseres Beschränkungsprofil hat.

Wäre (14) eine tatsächliche Ordnung in einer optimalitätstheoretischen Grammatik, dann würden wir erwarten, daß zur Erfüllung der höchstgeordneten Beschränkung SYN-KON tiefer geordnete Beschränkungen beliebig oft verletzt werden können. Dies ist aber bei Pafel gerade nicht intendiert: Im Gegenteil, die Annahme ist hier (in optimalitätstheoretische Terminologie übersetzt), daß die Kombination

von häufigen Verletzungen von tiefer geordneten Beschränkungen B, C und gleichzeitiger Erfüllung einer höher geordneten Beschränkung A bei einem Kandidaten die Kombination einer Verletzung von A und gleichzeitiger Erfüllung von B, C bei einem anderen Kandidaten aufwiegen kann. D.h., das in der Optimalitätstheorie an sich irrelevante Konzept der Kumulativität spielt bei Pafels Skopusanalyse eine Rolle. Dies läßt sich anhand des Beispiels in (12) zeigen.

In (12) ist wie gesehen die präferierte Lesart die mit weitem Skopus für QA₂; hier wird die Verletzung der höchstgeordneten Beschränkung SYN-KON mehr als aufgewogen durch die Erfüllung der tiefer geordneten Beschränkungen GRAM-FUNK und DIST. Im Gegensatz dazu erweist sich die an der Oberflächenabfolge orientierte Lesart mit weitem Skopus für QA₁, die zwar SYN-KON respektiert, aber die tiefer-rangigen Beschränkungen GRAM-FUNK und DIST verletzt, als stärker markiert, also gegenüber der ersten Lesart suboptimal.

Das Konzept der Kumulativität von Beschränkungsverletzungen ist somit ein wesentlicher Faktor, der Pafels Analyse von einer optimalitätstheoretischen Analyse unterscheidet. Hinzu kommt noch, daß Pafels System Grammatikalitätsgrade (kodiert als Lesartpräferenzen) vorsieht, was wie gesehen ebenfalls nicht ohne weiteres mit der klassischen Optimalitätstheorie kompatibel ist. Nichtsdestoweniger scheint klar, daß die empirische Evidenz im Bereich des relativen Quantorenskopus im Deutschen von einer Art ist, die im Prinzip eine optimalitätstheoretische Herangehensweise möglich und plausibel erscheinen läßt.⁸ Bereits Kapitel 2 hat gezeigt, daß der Pafelsche Faktor der syntaktischen Konstellation optimalitätstheoretisch eine Umsetzung als verletzbare und geordnete Beschränkung erfahren kann; vgl. Abschnitt 5.2. von Kapitel 2 und die dortige Diskussion der Beschränkungen SKOPUS aus Vikner (1997b) sowie PAR-VER aus Müller (1999a). Tatsächlich erweist es sich als möglich, eine umfassende Theorie des relativen Quantorenskopus im Deutschen, die wesentliche Erkenntnisse der Pafelschen Analyse integriert, in einem strikt optimalitätstheoretischen System zu entwickeln; vgl. Heck (1998; 1999) und Fischer (1999).⁹

⁸Das Datenprofil ist demjenigen bei der freien Wortstellung im Mittelfeld nicht unähnlich. Tatsächlich ist das in Dietrich (1994) entwickelte sog. Koalitionsmodell der deutschen Wortstellung im Prinzip genau wie Pafels Skopusmodell aufgebaut: Es werden Präzedenzwerte an Konstituenten vergeben, die durch Faktoren wie Definitheit, Agentivität und Fokus ermittelt werden. Eine "gute" Wortstellung ist dann eine, in der die Präzedenzwerte von links nach rechts abnehmen; dies entspricht der Pafelschen Skopusbeschränkung (15). Weitere Bemerkungen zu dieser augenscheinlichen Konvergenz zweier verschiedener Datenmengen finden sich bei Pafel (1998).

⁹Ein wesentlicher, in der Befolgung optimalitätstheoretischer Grundsätze begründeter Unterschied ist allerdings, daß bei Heck wie bei Fischer zwar korrekt das Vorhandensein von Skopusambiguität bzw. Skopuseindeutigkeit vorausgesagt wird, aber mögliche Lesartpräferenzen nicht erklärt werden. Was im übrigen die Existenz von Skopusambiguitäten betrifft, so impliziert das in einem optimalitätstheoretischen System das Vorhandensein zweier optimaler Kandidaten; dies wird bei Heck mit Hilfe der Konzepte der Pseudo-Optionalität (Präsenz bzw. Fehlen eines optionalen Skopusmarkers führt zu zwei verschiedenen Kandidatenmengen, die unterschiedliche Lesarten

4. Wettbewerb

Daß über die Wohlgeformtheit eines Satzes auch Eigenschaften anderer Sätze wesentlich mit entscheiden können und somit die Grammatikalität eines Satzes beeinflußt wird von seinem Verhalten im Wettbewerb mit anderen Sätzen, ist tatsächlich eine relativ alte Idee, die in der Syntaxforschung immer wieder einmal aufgetaucht ist.¹⁰ Im folgenden möchte ich drei besonders augenfällige Beispiele für Ansätze diskutieren, die ein Wettbewerbskonzept inkorporieren, nämlich (a) eine Version des Chomskyschen minimalistischen Programms, die transderivationelle Beschränkungen vorsieht, (b) die Theorie der Blockade, wie sie vor allem von Edwin Williams für die Syntax entwickelt worden ist, sowie (c) eine von Gisbert Fanselow ausgearbeitete Variante der Theorie der Blockade, die auf dem Prinzip der Vollständigen Inklusion (Proper Inclusion Principle) beruht. All diese Wettbewerbssyntaxen haben im Prinzip denselben Aufbau wie eine optimalitätstheoretische Syntax, nämlich den in (17) (vgl. das Bild auf Seite 14):

(17) *Struktur von Wettbewerbssyntaxen:*

- a. Gen erzeugt die Kandidatenmenge $\{K_1, K_2, \dots\}$.
- b. H-Eval determiniert den (die) optimalen Kandidaten K_i aus $\{K_1, K_2, \dots\}$.

Der wesentliche Unterschied zwischen den verschiedenen Wettbewerbssyntaxen betrifft die Frage, wie Optimalität in (17-b) bestimmt wird, d.h., welches Evaluationsmaß die H-Eval-Komponente verwendet. In der optimalitätstheoretischen Syntax klassifiziert H-Eval den Kandidaten mit dem besten Beschränkungsprofil als opti-

als Gewinner aufweisen können) und der Kopplung (beide Lesarten haben bei Kopplung zweier Beschränkungen ein identisches Beschränkungsprofil) erfaßt, bei Fischer nur unter Bezug auf Kopplungen; vgl. (19-a) und (19-c) aus Kapitel 1, respektive. Zu bemerken ist noch, daß Kumulativität bei Heck keine Rolle spielt, daß dieser Faktor jedoch in der Analyse von Fischer, die als möglichst akkurate Rekonstruktion der Pafelschen Theorie angelegt ist, mit Hilfe des Konzepts der lokalen Konjunktion (vgl. Smolensky (1995)) implementiert wird; siehe Abschnitt 3 von Kapitel 3.

¹⁰Nicht zuletzt entspricht dies auch den Erfahrungen, die Linguisten häufig während der Datenerhebung bei linguistisch nicht vorgebildeten Informanten machen. Letztere neigen dazu, zunächst einfache Ja/Nein-Urteile zu verweigern und bei kritischen Daten stattdessen gleich mehr oder weniger geglückte Reparaturen vorzuschlagen, oft zusammen mit einer expliziten Feststellung dergestalt, daß die Existenz der Reparatur die Ursache für die Unmöglichkeit des erfragten Satzes ist. ("Das würde doch niemand so sagen; es muß vielmehr so und so heißen.") Im optimalitätstheoretischen Kontext ist das Auftauchen eines Reparaturversuchs bei Informantenbefragungen (bei leichter Ausdehnung des Konzepts) natürlich als äußeres Anzeichen eines Wettbewerbs zu interpretieren: Der Reparaturversuch blockiert den erfragten Satz als suboptimalen Kandidaten. Insofern kann man bei Syntaxtheorien, die ein Konzept des Wettbewerbs inkorporieren, davon sprechen, daß hier eine Konvergenz von theoretischer Vorgabe und dem gesunden Menschenverstand vorliegt – wobei diese Konvergenz zwar möglicherweise suggestiv, strenggenommen jedoch irrelevant ist.

mal; es gibt also ein einziges, dafür aber einigermaßen komplexes Evaluationsmaß. Demgegenüber wählt H-Eval in einer minimalistischen Syntax mit transderivationellen Beschränkungen denjenigen Kandidaten als optimal aus, der von allen transderivationellen Beschränkungen bevorzugt wird; hier liegen also mehrere Evaluationsmaße vor, die jedoch, wie sich zeigen wird, an sich alle recht einfach sind.¹¹ In Theorien mit dem Blockade-Prinzip oder dem Prinzip der Vollständigen Inklusion schließlich bestimmt H-Eval denjenigen Kandidaten als optimal, der am spezifischsten ist bzw. die kleinsten Anwendungsdomänen von Merkmalszuweisungsmechanismen hat; hier haben wir es wiederum nur mit einem einzigen Evaluationsmaß zu tun, das darüber hinaus auch noch recht einfach ist.

4.1. Transderivationelle Ökonomiebeschränkungen im Minimalismus

In den klassischen Varianten des minimalistischen Programms (vgl. Chomsky (1991; 1993; 1995)) sind transderivationelle Ökonomiebeschränkungen vorgesehen. Will man ermitteln, ob eine derartige Beschränkung von einem Satz (Kandidaten) K erfüllt ist oder nicht, reicht es nicht, inhärente Eigenschaften von K zu betrachten; man muß vielmehr K mit anderen Kandidaten in derselben Kandidatenmenge (“Referenzmenge” bei Chomsky) vergleichen und somit einen Wettbewerb einführen. Das heißt, eine transderivationelle Beschränkung ist nicht lokal anhand der Eigenschaften eines einzelnen Kandidaten überprüfbar.¹² Zwei der wichtigsten transde-

¹¹Man mag vielleicht die mehrdeutige Verwendung des Begriffes “Optimalität” bemängeln. Dieser Begriff findet jedoch nicht nur in der optimalitätstheoretischen Syntax Verwendung, sondern taucht auch in Minimalismusvarianten mit transderivationellen Beschränkungen auf – dies vollkommen ohne Bezug auf die Optimalitätstheorie und, wie mir scheint, systematisch auch schon zu einer Zeit, wo optimalitätstheoretische Syntaxpapiere noch kaum existierten (vgl. dazu Chomsky (1993, 4), sowie zu expliziten Definitionen Collins (1994, 46) und Kitahara (1997, 18)).

¹²Eine Bemerkung zur Terminologie. Da im minimalistischen Programm mit dem Wegfall der Repräsentationsebene S-Struktur zugunsten einer strikt derivationalen Organisation der Syntax keine Möglichkeit mehr besteht, echt syntaktische repräsentationelle Beschränkungen zu postulieren (es gibt lediglich ein Überbleibsel davon als “bare output conditions” auf den Schnittstellen-Ebenen LF und PF), müssen alle lokal überprüfbaren syntaktischen Beschränkungen von vornherein derivational sein. Der Begriff “transderivationell” ergibt sich dann in offensichtlicher Weise. Allerdings sollte man sich darüber im klaren sein, daß die Thematik keineswegs auf derivationalen Grammatiken beschränkt ist. Bei einer rein repräsentationell orientierten Grammatik wie z.B. der Minimalismus-Variante in Brody (1995) würde sich im Prinzip ebenso die Möglichkeit ergeben, analoge “transrepräsentationelle” Beschränkungen einzuführen. Um die Gefahr zu vermeiden, hier jemanden auf eine falsche Fährte zu führen, schiene es mir daher ratsam, auf den Begriff der “transderivationellen Beschränkung” zu verzichten zugunsten des allgemeineren Begriffs der “translokalen” Beschränkung (wobei mit “lokal” dann “ohne Einbeziehung anderer Kandidaten bzw. Kandidat-externer Informationen überprüfbar” gemeint ist, was zunächst einmal sowohl mit einer repräsentationellen wie mit einer derivationalen Beschränkung vereinbar ist). Um mit der einschlägigen Literatur kompatibel zu sein, verwende ich jedoch im folgenden weiterhin meistens den Begriff “transderivationell”.

rivationellen Beschränkungen im minimalistischen Programm sind das Prinzip der wenigsten Schritte und das Prinzip der kürzesten Pfade. Diese Prinzipien werden in den folgenden beiden Unterabschnitten besprochen.

4.1.1. Das Prinzip der wenigsten Schritte

Das Prinzip der wenigsten Schritte (“Fewest Steps”) geht im wesentlichen auf Chomsky (1991) zurück. Die zugrundeliegende Intuition ist die, daß von mehreren konkurrierenden Kandidaten immer derjenige zu wählen (d.h., einzig wohlgeformt) ist, der insofern am ökonomischsten ist, als er die wenigsten syntaktischen Operationen aufweist:¹³

(18) *Prinzip der wenigsten Schritte:*

Falls zwei Derivationen D_1 und D_2 in derselben Kandidatenmenge sind und D_1 weniger Operationen als D_2 involviert, dann ist D_1 D_2 vorzuziehen.

In diesem Sinne als eine kostspielige “Operation” zählt syntaktische Bewegung; darüber hinaus aber z.B. auch noch Tilgung (vgl. hierzu insbesondere Chomsky & Lasnik (1993), die ein LF-Tilgungsverbot von intermediären Adjunktspuren gegenüber der LF-Tilgbarkeit von intermediären Argumentspuren aus dem Prinzip der wenigsten Schritte herleiten), Einsetzung, sowie möglicherweise Rekonstruktion. Eine elegante Anwendung von (18) findet sich bei Epstein (1992) (andere wichtige Arbeiten, die (18) verwenden, sind Fox (1995) und Kitahara (1997)).

Epsteins Ausgangsbeobachtung ist, daß W-Phrasen im Englischen nie topikalisiert werden dürfen, anders als andere XPs. Dies zeigen die folgenden Beispiele:

- (19) a. Who₁ t₁ said [CP that [IP John likes who₂]] ?
 b. Who₁ t₁ said [CP that Mary₂ [IP John likes t₂]] ?
 c. *Who₁ t₁ said [CP that who₂ [IP John likes t₂]] ?

In allen drei Sätzen ist der Matrixsatz eine Ergänzungsfrage; dies wird signalisiert durch die W-Bewegung des Subjekts in die lokale SpecC-Position. In (19-a) findet sich im eingebetteten Deklarativsatz in situ noch ein weiteres W-Element, und dies kann als Objekt per LF-Anhebung Matrix-Skopus nehmen ((19-a) wäre dann z.B. beantwortbar durch so etwas wie *Bill said that John likes Mary, Peter said that John likes Bill, and Fritz said that John likes Paul*). In (19-b) dagegen ist das eingebettete Objekt kein W-Element, sondern die [-w]-Phrase *Mary*; diese NP kann innerhalb des eingebetteten Satzes ungeachtet des W-Elements im Matrixsatz topikalisiert werden (zumindest gilt dies für jene Sprecher des Englischen, die überhaupt eingebettete Topikalisierung erlauben). Der interessante Fall ist nun (19-c). Dieser Satz ist ungrammatisch, und das zeigt, daß, obwohl an sich ein W-Objekt

¹³ Anders als lokal überprüfbare Beschränkungen setze ich hier und im folgenden zur einfacheren Unterscheidung transderivationale Beschränkungen nicht in Kapitälchen.

in diesem Kontext im eingebetteten Satz auftreten kann, und obwohl auch an sich Topikalisierung eines Objekts in diesem Kontext im eingebetteten Satz möglich ist, die Kombination nicht erlaubt ist: Ein W-Element darf nicht topikalisiert werden. Wie (20) zeigt, gilt im Deutschen dieselbe Restriktion (zumindest für solche Sprecher, die eine W-Phrase in situ im eingebetteten Satz, die Matrixskopus nimmt, tolerieren).

- (20) a. Wer₁ sagte t₁ [_{CP} daß [_{IP} Fritz wen₂ gesehen habe]] ?
 b. Wer₁ sagte t₁ [_{CP} den Fritz₂ habe [_{IP} er t₂ gesehen]] ?
 c. *Wer₁ sagte t₁ [_{CP} wen₂ habe [_{IP} Fritz t₂ gesehen]] ?

Unter der Annahme von LF-Bewegung der auf der Oberfläche in situ stehenden W-Elemente können wir postulieren, daß die LF-Repräsentationen der ungrammatischen Sätze wie folgt aussehen:¹⁴

- (21) a. [_{CP} who₂ [_{CP} who₁ t₁ said [_{CP} that [_{IP} John likes t₂]]]] ?
 b. [_{CP} wen₂ [_{CP} wer₁ t₁ sagte [_{CP} [_{IP} Fritz t₂ gesehen habe]]]] ?

Die Idee von Epstein ist nun, daß eine Derivation D₁ über W-in-situ wie in (19-a) und (20-a) mit einer Derivation D₂ über W-Topikalisierung wie in (19-c) und (20-c) im Wettbewerb steht, und daher D₁ aufgrund des Prinzips der wenigsten Schritte D₂ blockiert. Konkret setzt Epstein voraus, daß Kandidatenmengen über LF-Identität festgelegt sind; somit ist die Derivation, die (19-a) erzeugt, mit der Derivation, die (19-c) erzeugt, im selben Wettbewerb. In (19-a) erfolgt aus Skopus- und Selektionsgründen (der eingebettete Satz ist [-w] markiert) LF-Anhebung des eingebetteten W-Objekts aus der in-situ-Position in die Matrixposition; von S-Struktur-Bewegung ist dieses W-Objekt jedoch nicht affiziert worden. Im Unterschied dazu erfolgt in (19-c) zwar ebenso LF-Anhebung des W-Objekts in seine Skopusposition; dazu kommt jedoch noch ein S-struktureller Bewegungsschritt, die Topikalisierung. Unter der von Epstein weiter gemachten Voraussetzung, daß LF-Bewegung von Argumenten nicht sukzessiv-zyklisch erfolgen muß, ergibt sich, daß eine Derivation, die die LF (21-a) über die S-Struktur (19-a) erzeugt, eine Bewegungsoperation weniger aufweist als eine konkurrierende Derivation, die dieselbe LF (21-a) über die S-Struktur (19-c) erzeugt. Somit ist erklärt, warum ein Satz wie (19-c) ungrammatisch ist.¹⁵

¹⁴Hier ist gemäß Abschnitt 1.2. aus Kapitel 2 vorausgesetzt, daß W-Bewegung auf LF bei gefüllter SpecC-Position Adjunktion an CP involviert; dies verletzt PAR-VER, aber da es keine andere Wahl gibt und PAR-VER eine tief geordnete Beschränkung ist, ist diese Verletzung unproblematisch. Epstein (1992) nimmt demgegenüber an, daß W-Bewegung auf LF bei gefüllter SpecC-Position als Adjunktion an SpecC zu analysieren ist. Diese Unterscheidung ist für den hier betrachteten Fall nicht wichtig – bei Annahme von SpecC-Rechts- statt CP-Links-Adjunktion würden die im folgenden diskutierten Konsequenzen dieselben sein.

¹⁵Daß LF-Bewegung des Objekts aus der in-situ-Position in (19-a) in einem Rutsch erfolgen kann, ist zunächst einmal eine wesentliche Voraussetzung, denn bei sukzessiv-zyklischer LF-

Die von dieser Analyse nachgezeichnete Intuition ist, daß Topikalisierung in (19-c) in gewissem Sinne unökonomisch und somit überflüssig ist, da diese overte Operation ja auf dem Weg hin zur LF wieder rückgängig bzw. unsichtbar gemacht wird. Würde die eingebettete Topikalisierung auf LF bestehen bleiben, dann würde sie nicht als unökonomisch ausgefiltert. Man könnte nun meinen, eine Unterlassung von LF-Anhebung des W-Objekts in (19-c) würde dazu führen, daß der Kandidat nicht mehr blockiert wird. Dies ist zwar im Prinzip richtig. Allerdings verletzt eine solche Derivation die Forderung, daß auf LF alle W-Phrasen im Bereich eines [+w] markierten C-Kopfes sind. In Chomskyscher Terminologie können wir sagen, daß eine derartige Derivation nicht *konvergiert*, und wenn nur konvergierende Derivationen überhaupt in einem Wettbewerb zugelassen sein können, dann ist auch dieser Ausweg für die eingebettete W-Topikalisierung geschlossen. Zur Verfügung steht eine derartige Option jedoch solchen Elementen, die nicht auf LF in eine höhere Position angehoben werden müssen, also z.B. dem Eigennamen *Mary* in (19-b): Die Derivation, die (19-b) generiert, ist aufgrund von *Mary*≠*who* in einem anderen Wettbewerb und kann so (19-a)/(19-c) weder blockieren, noch davon blockiert werden. Der Grund dafür, daß (19-b) nicht durch eine in-situ-Strategie blockiert ist, d.h., daß ein optionaler Prozeß wie Topikalisierung überhaupt möglich ist, ist der, daß die Derivation, die (19-b) erzeugt, mit einer Derivation ohne Topikalisierung von *Mary* nicht im selben Wettbewerb ist, weil eine andere LF-Struktur vorliegt – intuitiv also, weil overte Topikalisierung hier nicht wieder durch LF-Anhebung zerstört wird.

Es soll an dieser Stelle nicht so sehr interessieren, ob sich die Epsteinsche Theorie aufrechterhalten läßt (in Müller & Sternefeld (1996) bringen wir empirische Gegen-evidenz, die nahelegt, daß diese Analyse zu restriktiv ist, und daß nur manche Fälle von durch LF-Anhebung zunichte gemachten overten Bewegungen von W-

Bewegung über die eingebettete SpecC-Position wäre ja doch ein zusätzlicher Bewegungsschritt notwendig, und die (19-a) erzeugende Derivation stünde beim Prinzip der wenigsten Schritte nicht besser da als die (19-c) erzeugende Derivation. Tatsächlich ist häufig davon ausgegangen worden, daß LF-Bewegung von Argumenten nicht sukzessiv-zyklisch zu sein braucht (d.h., daß LF-Bewegung keine Subjazeneffekte zeigt); vgl. etwa Chomsky (1981), Huang (1982) und Lasnik & Saito (1992). Darüber hinaus ist aber zum einen nicht ganz klar, daß, selbst wenn (19-a) noch einen zusätzlichen LF-Bewegungsschritt zur eingebetteten SpecC-Position bräuchte, nicht auch (19-c) diesen Schritt benötigte (denn da das eingebettete Topik hier dem Komplementierer folgt, bedeutet dies, daß Topikalisierung an sich zumindest im Englischen noch keine Bewegung nach SpecC sein kann); dann aber würde das Prinzip der wenigsten Schritte wieder wie erwünscht zwischen den beiden Derivationen unterscheiden. Zum anderen könnte ein Rekurs auf Chomskys (1991; 1993) Konzept der Kettenbildung ("Form Chain") auch bei Annahme sukzessiv-zyklischer LF-Bewegung von Argumenten die (19-a) erzeugende Derivation als die mit den wenigsten Schritten in ihrer Kandidatenmenge erweisen. Während bei der Operation "Bewegung" jeder einzelne Schritt gezählt wird, reanalysiert die Operation "Kettenbildung" beliebig viele Schritte sukzessiv-zyklischer Bewegung *in ein und derselben syntaktischen Komponente* (overte Syntax oder LF-Syntax) als einen einzigen Schritt. Man müßte hier lediglich postulieren, daß mit "Operation" im Sinne von (18) immer Kettenbildung, nicht Bewegung an sich, gemeint ist.

Phrasen in Ungrammatikalität enden). Der wesentliche Punkt ist, daß hier gerade so wie in der Optimalitätstheorie ein Wettbewerb von Kandidaten postuliert wird, und daß nur ein Kandidat, der diesen Wettbewerb als in gewissem Sinne “optimal” beendet, grammatisch sein kann. Bevor ich genauer auf das Verhältnis dieser beiden Wettbewerbstheorien zu sprechen komme, sei noch eine zweite transderivationelle Ökonomiebeschränkung im minimalistischen Programm diskutiert.

4.1.2. Das Prinzip der kürzesten Pfade

Das Prinzip der kürzesten Pfade (“Shortest Path Condition”) läßt sich wie folgt formulieren (vgl. Chomsky (1993; 1995), Collins (1994), Kitahara (1993; 1997) und Nakamura (1998), neben anderen):

(22) *Prinzip der kürzesten Pfade:*

Falls zwei Derivationen D_1 und D_2 in derselben Kandidatenmenge sind und D_1 kürzere Pfade als D_2 involviert, dann ist D_1 D_2 vorzuziehen.

Pfade sind minimale syntaktische Teilbäume, also kleinste Mengen von verbundenen Knoten, die zwei gegebene Knoten enthalten. Ein Pfad ist kürzer als ein anderer, wenn er weniger Knoten aufweist. Für diese transderivationelle Beschränkung sind einige Anwendungen in der Literatur vorgeschlagen worden. Drei davon seien hier kurz referiert.

Superioritätseffekte im Englischen Vielleicht die bekannteste und offensichtlichste Anwendung dieser Beschränkung betrifft den Superioritätseffekt. Die entsprechende Analyse ist in Chomsky (1993) skizziert und in Kitahara (1993) ausgearbeitet worden. Wir haben Superioritätseffekte bereits im Zusammenhang mit der Beschränkung PAR-VER in Abschnitt 4.3 von Kapitel 2 behandelt. Die einschlägigen Daten sind in (23) wiederholt.

- (23) a. I wonder [_{CP} who₁ C [_{IP} t₁ bought what₂]]
 b. *I wonder [_{CP} what₂ C [_{IP} who₁ bought t₂]]

Grundvoraussetzung ist, daß die Derivation D_1 in (23-a) mit der Derivation D_2 in (23-b) konkurriert. Da D_1 einen kürzeren W-Pfad hat als D_2 (der W-Pfad geht im ersten Fall von SpecI nach SpecC, im zweiten Fall von der Komplementposition von V nach SpecC), blockiert D_1 D_2 als ungrammatisch, und der Superioritätseffekt ist somit erklärt. In Chomsky (1981) wurde der Superioritätseffekt in (23) durch das ECP (“Empty Category Principle”) abgeleitet, demzufolge Spuren strikt regiert sein müssen: Die Annahme war hier, daß alle S-Struktur-Spuren in (23-ab) strikt regiert sind, daß die auf LF entstehende Objektspur in (23-a) strikt regiert ist, und daß die auf LF entstehende Subjektspur in (23-b) nicht strikt regiert ist. Diese Erklärung war immer nicht ganz unproblematisch, da sie nichts über den bei zwei (per Annahme immer strikt regierten) Objekten auftretenden Superioritätseffekt in (24) sagen konnte.

- (24) a. Whom₁ did John persuade t₁ [_{CP} to visit whom₂] ?
 b. *Whom₂ did John persuade whom₁ [_{CP} t'₂ to visit t₂] ?

Auch dieser Effekt ist jedoch durch das Prinzip der kürzesten Pfade ableitbar, denn der W-Pfad von *whom*₁ in (24-a) ist kürzer als der W-Pfad von *whom*₂ in (24-b).

Vergleicht man diese Erklärung mit Epsteins Analyse unmöglicher W-Topikalierungen, die auf dem Prinzip der wenigsten Schritte beruht, so ergibt sich eine interessante Divergenz. Während Epsteins Analyse darauf beruht, daß S-strukturelle und auf LF stattfindende Bewegungen gleichermaßen für transderivationelle Ökonomie zählen, setzt die Erklärung des Superioritätseffekts durch das Prinzip der kürzesten Pfade voraus, daß entweder die durch LF-W-Bewegung in den Satzpaaren (23) und (24) entstehenden Pfade für (22) irrelevant sind, oder daß von vornherein gar keine LF-Bewegung von W-Phrasen erfolgt. Der Grund ist, daß bei jedem Paar der Kandidat mit kurzer W-Bewegung auf der S-Struktur lange W-Bewegung auf LF erfordert, und der Kandidat mit langer W-Bewegung auf der S-Struktur mit kurzer W-Bewegung auf LF auskommt, so daß insgesamt über die gesamte Derivation verteilt die W-Pfad-Längen wieder gleich sind. Es ist nicht ganz offensichtlich, wie die beiden Erklärungen miteinander konsistent gemacht werden können (stipulativ wäre die Annahme, daß die eine Beschränkung alle Bewegungen betrifft, die andere nur S-strukturelle Bewegungen).¹⁶ Da es hier jedoch vor allem um den Vergleich optimalitätstheoretischer Analysen mit transderivationellen Analysen in der minimalistischen Syntax geht, und nicht in erster Linie um die interne Vereinbarkeit spezieller in der Literatur vorgeschlagener minimalistischer Analysen, werde ich mich im folgenden um dieses Problem nicht weiter kümmern. Stattdessen wende ich mich einem zweiten Bereich zu, für den eine Erklärung mit Hilfe des Prinzips der kürzesten Pfade vorgeschlagen worden ist.

Jojo-Bewegung im Ewe Mit "Jojo-Bewegung" ist eine Kombination von innerhalb einer syntaktischen Komponente (S-Struktur oder LF) bei demselben Element applizierenden Senkungs- und Anhebungsoperationen gemeint. Die Thematik ist z.T. bereits bei Chomsky (1986a) relevant, und sie wird sehr wichtig bei Lasnik & Saito (1992). Collins (1994, 53-57) führt zunächst ein empirisches Argument gegen die Existenz von Jojo-Bewegung in der westafrikanischen Sprache Ewe an und leitet

¹⁶Aber vgl. Sternefeld (1997, 92-93), wo ein Maß für Pfadlängen vorgeschlagen wird, demzufolge LF-Bewegung von Position α zu Position β kürzer sein kann als S-strukturelle Bewegung von α nach β (weil nicht überquerte Knoten, sondern überquerte Ketten gezählt werden). Durch diese Änderung läßt sich die Analyse von Superioritätsdaten mit Hilfe des Prinzips des kürzesten Pfades auch bei Annahme von LF-W-Bewegung aufrechterhalten. – Nebenbei sei bemerkt, daß in neueren minimalistischen Arbeiten gezeigt worden ist, daß man den Kern dieser Erklärung von Superioritätseffekten in einer nicht transderivationellen Variante der Analyse unangetastet lassen kann; diese Analyse beruht auf der derivationalen Bedingung des kürzesten Bindeglieds ("Minimal Link Condition"), die anders als die Bedingung der kürzesten Pfade lokal überprüfbar ist. Vgl. insbesondere Chomsky (1995) und Kitahara (1997).

dann das generelle Verbot einer solchen Bewegung aus dem Prinzip der kürzesten Pfade her.

Im Ewe gibt es wie in vielen anderen Sprachen auch einen (optionalen) Reflex von sukzessiv-zyklischer Bewegung durch eine eingebettete SpecC-Position: Genau dann, wenn A'-Bewegung durch SpecC stattgefunden hat, ist es möglich, daß ein eingebettetes Personalpronomen der 3. Person Singular in der adjazenten Subjektposition anstatt in der regulären Form *é* in der markierten Form *wo* erscheint; dieser Wechsel ist aber optional. Das zeigt das folgende Beispiel, in dem ein eingebettetes Objekt in den Matrixsatz A'-bewegt worden ist (als Instanz von Fokusbewegung).

(25) Kofi₁ ε me gblë [CP t'₁ be é/wo fo t₁]
 Kofi FOK ich sagte daß er schlug

Warum dieser Wechsel nur in diesem Kontext möglich ist, und warum er dann optional ist, ist an dieser Stelle nicht so wichtig; wichtig ist die Generalisierung, daß der Wechsel die Benutzung von SpecC als Landestelle von A'-Bewegung signalisiert. Aus prätheoretischer Perspektive ist es nun auch nicht weiter erstaunlich, daß dieser Reflex von sukzessiv-zyklischer Bewegung in einem eingebetteten Satz nicht auftreten kann, wenn das A'-bewegte Element von vornherein schon im Matrixsatz basisgeneriert wird und auch dort seine Zielposition hat. Wie (26) zeigt, ist in diesem Fall im eingebetteten Satz nur *é* in der Subjektposition möglich; *wo* ist verboten.

(26) Kofi₁ ε me gblë na t₁ [CP be é/*wo fo Kosi]
 Kofi FOK ich sagte zu daß er schlug Kosi

Allerdings gilt es sicherzustellen, daß (26) nicht eine Derivation haben kann, in der Jojo-Bewegung von *Kofi* erfolgt: zuerst in die eingebettete SpecC-Position (von wo dann als Reflex dieser Senkung *wo* in der eingebetteten Subjektposition lizenziert wird), und darauf über die Basisposition hinweg in die SpecC-Zielposition des Matrixsatzes. Eine solche Derivation sähe wie in (27) aus:

(27) *Kofi₁ ε me gblë na t₁ [CP t'₁ be wo fo Kosi]

Es kommen wohl einige Möglichkeiten in Betracht, Jojo-Bewegung wie in (27) auszuschließen (erwähnt seien nur die Bedingung des Strikten Zyklus (Strict Cycle Condition) oder eine Beschränkung, derzufolge alle Spuren an jedem Schritt der Derivation gebunden sein müssen; eine derivationale Version von STR-BD aus Kapitel 2), aber Collins (1994) versucht zu zeigen, daß viele naheliegende Erklärungen bei genauerem Hinsehen nicht unproblematisch sind. Er schlägt vor, daß Jojo-Bewegung wie in (27) systematisch durch das Prinzip der kürzesten Pfade ausgeschlossen werden kann: Die Derivation in (26), die sich von der Derivation in (27) nur dadurch unterscheidet, daß sie den ersten Senkungsschritt unterläßt, hat natürlich kürzere Bewegungspfade und blockiert somit die Derivation in (27) als unökonomisch.

Zwei Voraussetzungen sind bei dieser Analyse noch zu klären. Erstens muß

gewährleistet sein, daß Kandidaten mit *é* mit Kandidaten mit *wo* konkurrieren können. Dies bedeutet, daß Kandidatenmengen nicht über Identität der Numeration (desselben lexikalischen Materials) definiert werden dürfen (sondern z.B. wieder über LF-Identität zu definieren sind). Zweitens stellt sich die Frage, ob nicht bereits auch das Prinzip der wenigsten Schritte für den vorliegenden Fall den erwünschten Erfolg brächte: (27) involviert zwei A'-Bewegungsschritte, (26) dagegen nur einen. Eine genauere Betrachtung von regulärer sukzessiv-zyklischer Bewegung wie z.B. in (25) zeigt jedoch, daß eine solche Analyse sehr problematisch wäre. Denn dann sollten wir ja ebenfalls erwarten, daß niemals die Form *wo* als optionaler Reflex von sukzessiv-zyklischer Bewegung auftreten sollte – eine Variante, die die Bewegung in einem Rutsch vornimmt, wäre ja immer ökonomischer. Angesichts dieser Situation nimmt Collins (1994) auch an, daß mit dem Begriff der "Operation" in der Formulierung des Prinzips der wenigsten Schritte immer "Kettenbildung" (s.o.) und nicht schlichte "Bewegung" gemeint ist. Somit weisen (25), (26) und (27) jeweils nur genau eine syntaktische A'-Bewegungsoperation im Sinne von (18) auf; sie sind daher vom Prinzip der wenigsten Schritte nicht zu unterscheiden.¹⁷

Abgeleitete Subjekte im Englischen Eine dritte Anwendung der Bedingung der kürzesten Pfade betrifft den Inselstatus abgeleiteter Subjekte im Englischen. Die zu erklärende Generalisierung ist, daß bestimmte Typen von NPs im Englischen für W-Extraktion transparent sind, wenn sie die Objektposition einnehmen, aber eine Insel sind, wenn sie sich in der Subjektposition befinden:

- (28) a. Who₁ did you take [_{NP₂} a picture of t₁] ?
 b. *Who₁ did [_{NP₂} a picture of t₁] please John ?

Eine Subjekt-NP in einer Passivkonstruktion, die ein D-strukturelles Objekt ist, verhält sich in dieser Beziehung gerade so wie eine echte Subjekt-NP – sie blockiert W-Extraktion:

- (29) *Who₁ was [_{NP₂} a picture of t₁] taken t₂ by Bill ?

NPs in Subjektposition sind Barrieren (da sie keine Schwestern von lexikalischen Köpfen und somit nicht L-markiert sind; vgl. Chomsky (1986a)), NPs in Objektposition sind das nicht. Die folgende Lokalitätsbeschränkung verbietet das Überqueren von Barrieren durch syntaktische Bewegung (vgl. Huang (1982), Chomsky (1986a), sowie BAR aus Kapitel 2 und Kapitel 3.).

¹⁷Die Darstellung von Collins' (1994) Analyse ist hier leicht vereinfacht. Collins nimmt mit Chomsky (1986a) immer noch zusätzliche intermediäre VP-Adjunktion bei A'-Bewegung an. Das ändert an der Argumentation an sich allerdings nichts.

- (30) BFE (“Beschränkung für Extraktionsbereiche”, “Condition on Extraction Domain”, CED):
- a. Bewegung darf keine Barriere überqueren.
 - b. Eine XP ist eine Barriere, wenn sie nicht die Schwester eines lexikalischen Kopfes ist.

In einer repräsentationellen Grammatik, die nur auf die S-Struktur-Verhältnisse sieht, reicht diese Annahme bereits aus, um die Ungrammatikalität von (29) zu erklären. In einer strikt derivationellen Grammatik wie dem minimalistischen Programm jedoch kann man sich bei der Überprüfung der Legitimität von Bewegungen nicht auf die S-Struktur-Verhältnisse beziehen; vielmehr gilt es zu zeigen, daß jede Kombination von Bewegungen in (29) entweder eine derivationelle Beschränkung der Grammatik verletzt oder durch eine transderivationelle Beschränkung aussortiert wird. Auf den ersten Blick ergeben sich in (29) zwei Möglichkeiten. Wenn Passiv-induzierte NP-Anhebung W-Extraktion vorangeht, verläßt W-Extraktion NP₂ zu einem Zeitpunkt, wo sie schon in der Subjektposition und also eine Barriere ist; diese Derivation verletzt somit die BFE. Welche Beschränkung ist nun aber bei der umgekehrten Abfolge von Bewegungsoperationen verletzt, gemäß der zunächst W-Extraktion aus der Objektposition erfolgt, und darauf hin erst NP₂-Anhebung in die Subjektposition? Verschiedene Möglichkeiten kommen hier in Betracht. Collins (1994) und Chomsky (1995) stellen fest, daß die Bedingung des strikten Zyklus in der ein oder anderen Formulierung das erwünschte Resultat liefern kann (die nachfolgende NP-Anhebung endet in einer tieferen Position als die vorausgegangene W-Extraktion). Chomsky (1995, 328) weist aber darauf hin, daß unabhängig davon das Prinzip der kürzesten Pfade ausreicht, eine “azyklische”, aber Lokalität respektierende Derivation von (29) zu blockieren. Er schreibt: “Passiv [= NP-Anhebung] ist bei beiden Kandidaten dasselbe; W-Bewegung ist beim unmöglichen Kandidaten in einem offensichtlichen Sinne ‘länger’, da die Objektposition von SpecC weiter weg ist als die Subjektposition, was die Zahl der überquerten XPs betrifft.” Mit anderen Worten: Eine Derivation D₁, die (29) in der Reihenfolge NP-Anhebung vor W-Bewegung ableitet, hat kürzere Pfade als eine konkurrierende Derivation D₂, die (29) in der Reihenfolge W-Bewegung vor NP-Anhebung ableitet. Daher blockiert D₁ D₂, auch wenn – und das ist im gegenwärtigen Zusammenhang interessant – der somit “optimale” Kandidat D₁ anders als der blockierte Kandidat D₂ eine lokale Beschränkung der Grammatik (die BFE) verletzt. Somit ergibt sich, daß unter bestimmten Umständen ein ungrammatischer Kandidat durch das Prinzip der kürzesten Pfade einen anderen Kandidaten blockieren kann.¹⁸

¹⁸Im Unterschied hierzu hatte sich bei der Diskussion des Prinzips der wenigsten Schritte gezeigt, daß ein Kandidat, der eine bestimmte grammatische Forderung verletzt, nämlich die, daß eine W-Phrase auf LF in SpecC stehen muß, vom Wettbewerb ausgeschlossen werden muß. Dies bedeutet, daß in diesem System zwischen zwei Typen von aufgrund der Verletzung lokal überprüfbarer Beschränkungen ungrammatischen Derivationen unterschieden werden muß: Kandidaten,

4.1.3. *Konklusion*

Die zwei eingeführten transderivationellen Beschränkungen sind vielleicht die am meisten diskutierten; sie sind aber keineswegs die einzigen, die im minimalistischen Programm vorgeschlagen wurden. Andere Beschränkungen dieses Typs, wie sie etwa Chomsky (1993) verwendet, sind das Präferenzprinzip für Rekonstruktion oder, wie bereits in Abschnitt 5.3 von Kapitel 2 gezeigt wurde, das Zauderprinzip (“Procrastinate”) unter einem bestimmten Verständnis (vgl. Marantz (1995)). Insgesamt ergibt sich, daß eine optimalitätstheoretische Syntax und eine minimalistische Syntax mit transderivationellen Beschränkungen im Prinzip dieselbe Grundstruktur haben, nämlich die in (17), die hier wiederholt ist.

- (31) *Struktur einer transderivationellen minimalistischen Syntax* :
- a. Gen erzeugt die Kandidatenmenge $\{K_1, K_2, \dots\}$.
 - b. H-Eval determiniert den (die) optimalen Kandidaten K_i aus $\{K_1, K_2, \dots\}$.

Üblicherweise sind im Minimalismus mit transderivationellen Beschränkungen die Kandidaten K in (31) Derivationen, und Gen ist ein Regelsystem, daß keine Oberflächenfilter, sondern nur derivationelle Beschränkungen und “bare output conditions” auf den Schnittstellenebenen PF und LF aufweist. Dies ist jedoch, wie in Kapitel 1 ausgeführt wurde, kein grundlegender Unterschied, da eine optimalitätstheoretische Syntax ebensogut derivationell wie repräsentationell ausgerichtet sein kann (und umgekehrt schließt a priori nichts eine repräsentationelle Minimalismusvariante mit translokalen Beschränkungen aus; siehe Fußnote 12).

Es gibt allerdings zwei wichtige, miteinander in Zusammenhang stehende Unterschiede. Der erste Unterschied betrifft die Arbeitsteilung von Gen und H-Eval: In der minimalistischen Syntax wird die Hauptarbeit von Gen gemacht; H-Eval spielt eine untergeordnete Rolle. In der optimalitätstheoretischen Syntax wird demgegenüber die Hauptarbeit von H-Eval gemacht; Gen spielt eine untergeordnete Rolle.¹⁹ Der zweite und zentrale Unterschied betrifft das Evaluationsmaß für Optimalität. In der optimalitätstheoretischen Syntax ist der Begriff der Optimalität komplex: Optimal ist der Kandidat mit dem besten Beschränkungsprofil gemäß einer größeren Zahl zueinander geordneter verletzbarer Beschränkungen. In der minimalistischen Syntax ist dagegen der Begriff der Optimalität einfacher: Optimal ist der Kandidat mit den wenigsten Schritten, oder der Kandidat mit den kürzesten Pfaden, etc., je nachdem, welches Maß für Optimalität angenommen wird.

die W-Bewegung unterlassen, konvergieren nicht und können daher nicht ansonsten wohlgeformte andere Kandidaten blockieren; Kandidaten, die eine Inselverletzung aufweisen, konvergieren und können daher ansonsten wohlgeformte andere Kandidaten blockieren.

¹⁹Eine Ausnahme ist hier die Theorie von Pesetsky (1997; 1998), die eine echte Arbeitsteilung zwischen Gen und H-Eval vorsieht. Siehe Abschnitt 2 von Kapitel 3.

Die größere Einfachheit des Evaluationsmaßes im minimalistischen Programm besteht allerdings nur so lang, wie es nur eine einzige transderivationelle Beschränkung in einer minimalistischen Syntax gibt. Es stellt sich die Frage, wie die Interaktion mehrerer transderivationeller Beschränkungen in einer minimalistischen Syntax zu regeln ist, denn diese Beschränkungen können im Prinzip jeweils unterschiedliche Kandidaten als optimal klassifizieren. Die grundsätzlichen Optionen sind in (32) aufgelistet.

- (32) Wenn zwei transderivationelle Beschränkungen im Konflikt sind, weil sie zwei unterschiedliche Derivationen auswählen, kann der Konflikt aufgelöst werden durch
- a. Ordnung: Optimal ist die Derivation, die die höher geordnete transderivationelle Beschränkung präferiert.
 - b. Toleranz: Beide Derivationen sind optimal.
 - c. Zusammenbruch: Keine Derivation ist optimal.

Die Problematik kann illustriert werden anhand der zwei hier vorgestellten transderivationellen Beschränkungen und ihren Konsequenzen für die Ableitung des soeben diskutierten Inselstatus abgeleiteter Subjekte im Englischen:

- (33) *Who₁ was [_{NP₂} a picture of t₁] taken t₂ by Bill ?

Bislang haben wir zwei Derivationen für (33) betrachtet: D₁ involviert NP-Anhebung vor W-Bewegung und verletzt die BFE. D₂ involviert W-Bewegung vor NP-Anhebung und hat somit einen längeren W-Pfad als D₁. D₁ wird daher vom transderivationellen Prinzip der kürzesten Pfade gegenüber D₂ präferiert. Es sind nun jedoch noch mindestens zwei weitere Derivationen D₃ und D₄ auszuschließen, um die Ungrammatikalität des Satzes in (33) zu garantieren. Derivation D₃ ist von Browning (1989; 1991) und Collins (1994) betrachtet worden; hier erfolgt ein "Kettendurchschuß" ("chain interleaving") in dem Sinne, daß zunächst per Adjunktion des W-Elements aus der Objekt-NP heraus an VP ein Teil der W-Kette gebildet wird, dann die NP in die SpecI-Position angehoben wird, und schließlich in einem dritten Schritt das W-Element aus der VP-Adjunktionsposition in die SpecC-Position geschoben wird. Das Ergebnis ist dann (34).

- (34) *Who₁ was [_{NP₂} a picture of t₁] [_{VP} t'₁ [_{VP} taken t₂ by Bill]] ?

Collins (1994) argumentiert, daß D₃ gegenüber D₁ durch das Prinzip der wenigsten Schritte ausgeschlossen wird: In D₃ gibt es insgesamt drei Bewegungen für NP₂ und NP₁ (die beiden Bewegungen des W-Elementes *who*₁ können aufgrund von Kettendurchschuß nicht als eine Instanz von Kettenbildung angesehen werden), in D₁ nur zwei. Tatsächlich hat aber D₃ auch dieselben Pfadlängen wie D₂ (wenn auch die identischen Pfadlängen durch mehr syntaktische Bewegungen zustande kommen), so daß das Prinzip der kürzesten Pfade ebenfalls D₁ gegenüber D₃ präferiert. Theoretisch interessant ist nun eine vierte Derivation D₄. Nehmen wir einmal an, daß Adjunktion an eine NP möglich ist, aber nur dann, wenn die NP in situ steht.

Dann erfolgt in D_4 zunächst Adjunktion des W-Elements who_1 an NP_2 ; im nächsten Schritt wird NP_2 in die Subjektposition bewegt; und schließlich wird who_1 aus der NP-Adjunktionsposition nach SpecC geschoben. Das Resultat ist dann (35):

(35) *Who₁ was [_{NP₂} t'₁ [_{NP₂} a picture of t₁]] taken t₂ by Bill ?

Wichtig ist, daß der letzte Schritt nicht die BFE verletzt, weil intermediäre Adjunktion an eine Barriere eine Überquerung im technischen Sinne vermeidet (vgl. Chomsky (1986a)). Wie wird nun D_4 ausgeschlossen? Es wäre nicht unmöglich, eine Verletzung einer lokalen Beschränkung (z.B. gegen NP-Adjunktion) für das Verbot von D_4 verantwortlich zu machen. Davon abgesehen wird aber D_4 gegenüber D_1 durch das Prinzip der wenigsten Schritte ausgeschlossen, denn D_4 hat drei Operationen bei NP_1 und NP_2 gegenüber zwei Operationen im Fall von D_1 . Es ergibt sich insgesamt das Ergebnis in Bild 2:

Bild 2: Transderivationelle Beschränkungen im Konflikt

	LOKALE BESCHRÄNKUNGEN	Kürzeste Pfade	Wenigste Schritte
D_1	verletzt	präferiert	präferiert
D_2	erfüllt	ausgeschlossen	präferiert
D_3	erfüllt	ausgeschlossen	ausgeschlossen
D_4	erfüllt	präferiert	ausgeschlossen

Auf dieser Basis können wir nun wieder die drei Optionen in (32) betrachten. Zunächst ist festzustellen, daß (32-a) kein genereller Ausweg aus dem Dilemma sein kann, daß zwei transderivationelle Beschränkungen unterschiedliche Derivationen auswählen. Wenn bei konfligierenden Vorhersagen des Prinzips der kürzesten Pfade und des Prinzips der wenigsten Schritte ersteres sich durchsetzt, wären D_1 und D_4 gleichermaßen optimal, und da D_4 (per Annahme) keine lokale Beschränkung verletzt, würde fälschlicherweise prognostiziert werden, daß diese Derivation legitim ist. Somit könnte der Inselstatus abgeleiteter Subjekte nicht mehr erfaßt werden. Ein analoges Problem ergäbe sich unter der umgekehrten Konfliktauflösung, bei der sich das Prinzip der wenigsten Schritte gegen das Prinzip der kürzesten Pfade durchsetzt, denn jetzt wären D_1 und D_2 beide optimal, und D_2 wäre eine legitime Derivation.²⁰ Die zweite Option in (32), Toleranz, hat noch weniger erwünschte

²⁰Die hier betrachtete Konfliktauflösung durch Ordnung transderivationeller Beschränkungen ist nicht mit dem optimalitätstheoretischen Konfliktauflösungsmechanismus zu verwechseln; nur im letzteren Fall spielt bei einem Konflikt die untergeordnete Beschränkung überhaupt eine Rolle. Würde man dieses Verfahren auf transderivationelle Beschränkungen anwenden, ergäben sich für die hier betrachteten Daten dieselben Konsequenzen wie bei der dritten Option (Zusammenbruch), die gleich besprochen wird. Mir ist keine Evidenz bekannt, die zwischen diesen beiden Optionen unterscheidet; dies spricht für die einfachere Zusammenbruchslösung.

Konsequenzen: Jetzt wären sowohl D_2 als auch D_4 optimale Derivationen, und die Erklärung des Inselstatus abgeleiteter Subjekte wäre zweifach unterminiert. (Lediglich Derivation D_3 könnte noch blockiert werden.) Was als einzige Möglichkeit übrig bleibt, ist die dritte Option, Zusammenbruch. Dies bedeutet, daß eine Derivation nur dann optimal sein kann, wenn sie von jeder transderivationellen Beschränkung präferiert wird. Im vorliegenden Falle ist das D_1 , und das ist unproblematisch, weil D_1 ja wie gesehen die lokale BFE verletzt. Damit ergibt sich im transderivationellen Minimalismus der Grammatikalitätsbegriff in (36) (etwas anders in der Formulierung, aber substantiell identisch findet sich dies auch bei Collins (1994)).

(36) *Grammatikalität:*

Eine Derivation D ist grammatisch gdw. (a) und (b) gelten.

- a. D verletzt keine lokale Beschränkung.
- b. Es gibt keine Derivation D' in derselben Kandidatenmenge, die von einer transderivationellen Beschränkung gegenüber D vorgezogen wird.

Damit ist das Evaluationsmaß für Optimalität im transderivationellen Minimalismus zwar nicht mehr ganz so einfach, aber immer noch einfacher als in der Optimalitätstheorie. Darüber hinaus gilt, daß eine Minimalismusvariante mit, sagen wir, fünf transderivationellen Beschränkungen insgesamt fünf Durchläufe durch die Kandidatenmenge voraussetzt: Zunächst wird ermittelt, welcher Kandidat die wenigsten Schritte hat, darauf wird festgestellt, welcher Kandidat die kürzesten Pfade hat, usw.; nur wenn ein Kandidat alle Durchläufe als Sieger beendet, kann er optimal sein. Im Vergleich hierzu bedarf es in der klassischen Optimalitätstheorie nur eines Durchlaufes durch die Kandidatenmenge (aber vgl. Kapitel 8).

Als Konklusion bleibt festzuhalten, daß sich die hier behandelte Minimalismusvariante und die optimalitätstheoretische Syntax im Grundaufbau erstaunlich ähnlich sind. Die Hauptunterschiede liegen in der Verteilung der Aufgaben für Gen und H-Eval und, vor allem, in der Festlegung des Evaluationsmaßes für Optimalität: Das in der optimalitätstheoretischen Syntax verwendete Maß ist komplexer, aber damit natürlich auch flexibler.²¹

²¹Eine wissenschaftsgeschichtliche Fußnote: Ein Faktum ist, daß die meisten jüngeren Minimalismusvarianten ohne transderivationelle Beschränkungen auszukommen versuchen; vgl. etwa Collins (1997), Frampton (1997) und Frampton & Gutman (1999). Dieses Bestreben findet sich bereits bei Chomsky (1995). So wird der oben auf Seite 165 zitierte Vorschlag, das Prinzip der kürzesten Pfade bei einer Ableitung des Inselstatus von derivierten Subjekten einzusetzen, ergänzt durch die Anmerkung, daß "die Sache nicht trivial ist, zum Teil, weil wir hier einen 'globalen' [= transderivationellen] Begriff von Ökonomie verwenden, der von der Art ist, wie wir ihn zu vermeiden versucht haben". Es ist zumindest erwähnenswert, daß der Aufstieg der Optimalitätstheorie und der Fall des transderivationellen Minimalismus in den neunziger Jahren Hand in Hand gegangen sind.

4.2. Blockade-Syntax

4.2.1. Die Grundidee

Auf der Basis der morphologischen Arbeiten von Aronoff (1976) entwickeln DiSciullo & Williams (1987, 10-14) und Williams (1997) eine Theorie der morphologischen und syntaktischen Blockade ("blocking"). Dieses Konzept setzt voraus, daß die Wohlgeformtheit einer sprachlichen Einheit nicht nur von internen Eigenschaften abhängt, sondern auch von externen Faktoren, genauer, von Eigenschaften anderer sprachlicher Einheiten. Angenommen wird dabei, daß Blockade bei Identität von Bedeutungen auftritt und der Synonymvermeidung dient. Dies bedeutet: In Blockade-Syntaxen sind Kandidatenmengen über identische LFs bzw. identische Bedeutung festgelegt. Der eigentliche Begriff der Blockade kann wie in (37) definiert werden.

(37) *Blockade:*

Wenn bei zwei miteinander im Wettbewerb befindlichen sprachlichen Formen F_1 , F_2 gilt, daß F_1 spezifischer ist als F_2 , dann blockiert F_1 F_2 als ungrammatisch.

Dieses Prinzip wählt unter ansonsten durch das grammatische System zugelassenen, aber bedeutungsgleichen Formen die spezifischste aus und blockiert alle anderen.²² Es stellt sich natürlich die Frage, wann eine Form als spezifischer als eine andere zu gelten hat. Die Antwort ist im Einzelfall (vor allem bei rein morphologischen Anwendungen) nicht immer offensichtlich. Unter anderem gilt wohl für die Morphologie: Unregelmäßige, nicht prädiktale und somit extra im Lexikon zu speichernde Formen sind spezifischer als regelmäßige (*went* blockiert z.B. **goed*). Für die Syntax (bzw. für die Morphologie-Syntax-Schnittstelle) bietet sich dagegen die folgende Festlegung an: Von zwei konkurrierenden (d.h., bedeutungsgleichen) Formen F_1 , F_2 ist F_1 spezifischer als F_2 genau dann, wenn es stärkere bzw. mehr Beschränkungen für F_1 als für F_2 gibt, die somit die Distribution von F_1 gegenüber F_2 restringieren.²³ Die Blockade-Theorie soll nun anhand von zwei von Williams diskutierten Phänomenen aus dem Englischen illustriert werden, der Komparativbildung und der pränominalen Possessiva.

4.2.2. Komparativbildung im Englischen

Im Englischen existieren zwei Mittel, um den Komparativ zu bilden. Die erste Möglichkeit ist die morphologische Suffigierung durch die Endung *er*; die zweite

²²Dem Blockade-Prinzip sehr ähnlich und in ihren Effekten durchaus vergleichbar ist die von Paul Kiparsky für die Phonologie entwickelte "Elsewhere Condition" (vgl. z.B. Kiparsky (1982)).

²³Als in diesem Sinne die Distribution restringierende Beschränkungen zählen natürlich nur lokale Beschränkungen, nicht dagegen Meta-Beschränkungen wie das Prinzip des Blockade selbst.

Möglichkeit ist die syntaktische Hinzufügung von *more*. Wie die Beispiele in (38) zeigen, sind die beiden Strategien in komplementärer Verteilung:

- (38) a. hot → hotter, *more hot
 b. happy → happier, *more happy
 c. colorful → *colorfuller, more colorful

Die morphologische Strategie wird durch Regel A erfaßt, die syntaktische Strategie durch Regel B:

- (39) a. *Regel A* (morphologisch):
 Das Suffix *er* kann nur angefügt werden an einsilbige Adjektive oder an zweisilbige Adjektive, die auf *y* enden.
 b. *Regel B* (syntaktisch):
 Der Komparativ von Adjektiven kann in der Syntax durch Hinzusetzen von *more* gebildet werden.

Es ist offensichtlich, daß somit syntaktisch gebildete Komparativformen weniger stark in ihrer Distribution eingeschränkt sind als morphologisch gebildete Komparativformen, denn letztere dürfen gemäß Regel A nur bei einer ganz bestimmten Teilklasse von Adjektiven gebildet werden, während Regel B zunächst einmal erlaubt, daß syntaktische Komparativformen bei allen Adjektiven gebildet werden. Konkret schließt Regel A Komparative wie **colorfuller* aus. Somit ist ein morphologischer Komparativ immer die spezifischere Form. Das Prinzip der Blockade gewährleistet dann, daß in allen Fällen, wo die Regeln A und B beide erfüllt sind und somit zunächst einmal ein morphologischer Komparativ F_1 ebenso wie ein syntaktischer Komparativ F_2 möglich sind, F_1 das weniger spezifische (also weniger stark restringierte) F_2 blockiert. Mit anderen Worten: Die Anwendbarkeit von Regel A verhindert die erfolgreiche Anwendung von Regel B. Daher sind Beispiele wie **more hot* und **more happy* ausgeschlossen.

Daß es sich hier um eine elegante Erklärung handelt, wird deutlich, wenn man versucht, Regel B so zu formulieren, daß man ohne ein Wettbewerbsmodell wie das der Blockade auskommt. Regel B muß dann nämlich wie in (40) aussehen:

- (40) *Regel B'* (zu verwerfende Variante):
 Der Komparativ von Adjektiven kann in der Syntax durch Hinzusetzen von *more* gebildet werden, es sei denn, es handelt sich um ein einsilbiges Adjektiv oder ein zweisilbiges Adjektiv, das auf *y* endet.

Damit ist der Kontext für erlaubte Anwendungen von Regel A identisch als Kontext für nicht-erlaubte Anwendungen von Regel B' wiederholt worden; eine solche Vorgehensweise ist aber unattraktiv, weil sie Redundanzen involviert und eine offensichtliche Generalisierung auslöst, nämlich die, daß die syntaktische Komparativbildung in genau den Kontexten möglich ist, wo morphologische Komparativbildung unmöglich ist – diese Generalisierung ist bei Annahme von Regel B' ein bloßer Zufall.

4.2.3. Possessivkonstruktionen im Englischen

Als zweites Beispiel seien englische Possessivkonstruktionen im pränominalen Bereich betrachtet. Die relevanten Daten sind in (41) versammelt:

- (41) a. [_{NP₁} the man I was talking to]'s hat
 b. *_{[NP₁} ??] hat (= es gibt kein Pronomen für NP₁ in (41-a))
 c. *_{[NP₁} them]'s hat
 d. [_{NP₁} their] hat

Pränominale Possessiv-NPs können entweder durch geeignete lexikalische Pronomina (wie *their*) oder syntaktisch durch Anfügung von *s* an gebildet werden. Lexikalische Possessivpronomina sind dabei die spezifischere Form; wo diese Form möglich ist, blockiert sie die syntaktische Strategie; vgl. (41-cd). Für nicht-pronominale (im vorliegenden Fall sogar recht komplexe) NPs wie *the man I was talking to* gibt es natürlich keine Possessivpronomina, die per se dasselbe bedeuten; und daher ist in diesem Fall die weniger spezifische Form der syntaktischen Possessivbildung per *s*-Anhängung erlaubt. Die Analyse ist also im Prinzip dieselbe wie vorher: Syntaktische pränominalen Possessivkonstruktionen sind teurer als lexikalische pränominalen Possessivkonstruktionen; erstere sind daher nur möglich als letzter Ausweg, wenn die andere Variante unabhängig blockiert ist.

Bei den Possessivkonstruktionen wie bei den Komparativkonstruktionen erscheint eine Optimalitätstheoretische Rekonstruktion leicht möglich. Dies soll kurz für die Komparativbildung gezeigt werden.

4.2.4. Optimalitätstheoretische Rekonstruktion

Zum einen bedarf es einer Beschränkung, die die durch Regel A gegebenen Restriktionen für morphologische Komparativbildung kodiert. Dies sei der SUFFIX-FILTER in (42-a).²⁴ Zum anderen ist eine Beschränkung notwendig, die insofern die Arbeit des Blockade-Prinzips übernimmt, als sie ganz allgemein morphologische Komparativbildung syntaktischer Komparativbildung vorzieht.²⁵ Dies sei die Beschränkung

²⁴ σ_2^* in (42-a) soll für eine beliebige Anzahl von Silben stehen (mindestens für eine). In der gegenwärtigen Formulierung sieht der Filter natürlich sprachspezifisch aus; um dem Postulat universeller Beschränkungen im H-Eval-System der Grammatik nachzukommen, wären einige Verallgemeinerungen notwendig. Hier geht es jedoch vor allem um die grundsätzliche Übersetzbarkeit von Blockade-Analysen in optimalitätstheoretische Analysen und nicht so sehr um die detaillierte Rechtfertigung der dabei verwendeten Beschränkungen.

²⁵Dies setzt voraus, daß morphologische und syntaktische Prinzipien miteinander konkurrieren können. Vgl. auch Perlmutter (1996), der ähnliche Wettbewerbe (zwischen Phonologie und Syntax) betrachtet, sowie Bresnan (1997; 1998).

KOMP-REAL (“Komparativ-Realisierung”) in (42-b).²⁶

- (42) a. SUFFIX-FILTER:
 $*[{}_X \sigma_1 \sigma_2^* \text{-er}]$, es sei denn, σ_2^* ist eine Silbe, die auf $-y$ endet.
 b. KOMP-REAL (“Komparativrealisierung”):
 Komparativmorpheme sind Affixe.

Wenn nun die Ordnung SUFFIX-FILTER \gg KOMP-REAL gilt, dann ergibt sich wie erwünscht, daß in den Fällen, wo beide Kandidaten den hoch geordneten SUFFIX-FILTER erfüllen (der syntaktische Kandidat leer, der morphologische Kandidat, weil er entweder einsilbig oder zweisilbig und auf y endend ist), die syntaktischer Komparativbildung inhärente KOMP-REAL-Verletzung fatal ist; vgl. Tabelle T_{4.1}.

T_{4.1}: Englische Komparativbildung 1

Kandidaten	SUFFIX-FILTER	KOMP-REAL
☞ K ₁ : happier		
K ₂ : more happy		*!

Falls jedoch der morphologische Kandidat den SUFFIX-FILTER verletzt, kann die immer präsenste KOMP-REAL-Verletzung durch den syntaktischen Kandidaten in Kauf genommen werden; vgl. Tabelle T_{4.2}.

T_{4.2}: Englische Komparativbildung 2

Kandidaten	SUFFIX-FILTER	KOMP-REAL
K ₁ : colorfuller	*!	
☞ K ₂ : more colorful		*

DiSciullo & Williams (1987, 13) und Williams (1997, 578-586) führen noch eine Reihe von weiteren Phänomenen an, die eine Analyse mit Hilfe des Prinzips der Blockade sinnvoll erscheinen lassen. Auch in diesen Fällen ist eine optimalitätstheoretische Rekonstruktion ohne große Probleme möglich. Man kann also schließen, daß dieses Ergebnis ganz allgemein gilt (vgl. auch Russell (1997)). Dabei ist ein möglicher Vorteil von optimalitätstheoretischen Analysen, daß sie flexibler sind. In

²⁶Man könnte hier möglicherweise statt KOMP-REAL auch eine noch allgemeinere Bedingung annehmen, wie etwa (i):

- (i) VERMEIDE PHRASENAUFBAU:
 Syntaktischer Phrasenaufbau ist verboten.

Gemäß (i) ist jeder Phrasenaufbau mit einer Verletzung verbunden. Diese Beschränkung mißt somit die Kosten von “Merge”-Operationen in Chomskys (1995) Sinn, und sie zieht *ceteris paribus* morphologische Operationen syntaktischen vor.

Blockade-Theorien ist es kaum möglich, mit dem Phänomen der Optionalität zurecht zu kommen, denn die Analyse beruht auf strikt komplementärer Verteilung. Diese Einschränkung gilt für die Optimalitätstheorie nicht (vgl. Kapitel 1).²⁷

4.2.5. Konklusion

Wie schon die minimalistische Syntax mit transderivationellen Beschränkungen, so erweist sich auch die Theorie der Blockade als grundsätzlich vom selben Typ wie die optimalitätstheoretische Syntax:

(43) *Struktur einer Blockade-Syntax:*

- a. Gen erzeugt die Kandidatenmenge $\{K_1, K_2, \dots\}$.
- b. H-Eval determiniert den (die) optimalen Kandidaten K_i aus $\{K_1, K_2, \dots\}$.

Wie beim Minimalismus liegt der Schwerpunkt auf Gen, und der H-Eval-Teil der Grammatik ist sehr einfach: Der optimale Kandidat in einer Kandidatenmenge ist gemäß dem Blockade-Prinzip (37) immer der spezifischste. Die Kandidatenmengen sind dabei in den in der Literatur existierenden Blockade-Analysen anders als in der optimalitätstheoretischen Syntax und im Minimalismus immer gleich definiert, nämlich über identische Bedeutung. Dies liegt an der Herkunft des Blockade-Prinzips – es ist zunächst einmal ein Mittel, Synonyme zu vermeiden.

Während Blockade-Theorien üblicherweise an Schnittstellen applizieren, ist eine von Fanselow (1991) entwickelte Variante derselben Grundidee eher Syntax-intern.

4.3. Das Prinzip der Vollständigen Inklusion

4.3.1. Die Analyse

Die Ausgangsbeobachtung ist, daß im Deutschen normalerweise anaphorische (Reflexiv- und Reziprok-) Pronomina und Personalpronomina komplementär verteilt sind. Bei A-Bindung innerhalb desselben Satzes sind anaphorische Pronomina obligatorisch; vgl. (44).

- (44) a. daß Maria₁ auf sich₁/*sich₂ stolz ist
 b. daß Maria₁ auf *sie₁/sie₂ stolz ist

Liegt entweder keine A-Bindung vor (vgl. (45)), oder erfolgt A-Bindung aus einer Position außerhalb des minimalen Satzes (vgl. (46)), so sind Personalpronomina obligatorisch zum Ausdruck von Koreferenz (bzw., bei quantifizierten Antezedentien,

²⁷Um eine Blockade-Analyse angesichts von Optionalitätsphänomenen zu retten, ist man gezwungen, subtile semantische Unterschiede zu finden, die verschiedene Kandidatenmengen erzeugen, d.h., den Weg der Pseudo-Optionalität zu gehen. Tatsächlich ist dies in Blockade-Analysen die gängige Praxis.

von Variablenbindung).

- (45) a. daß Marias₁ Bruder auf *sich₁ stolz ist
 b. daß Marias₁ Bruder auf sie₁ stolz ist
- (46) a. Maria₁ denkt [_{CP} daß man₃ sie₁/sie₂ reingelegt hat]
 b. Maria₁ denkt [_{CP} daß man₃ *sich₁/*sich₂/sich₃ reingelegt hat]

Dieser Sachverhalt wird in der klassischen GB-Theorie (vgl. Chomsky (1981)) durch die Bindungsprinzipien A und B erfaßt, wobei wir hier vereinfacht annehmen können, daß mit “lokaler Domäne” der minimale Satz gemeint ist.

- (47) a. PRINZIP A:
 Anaphern müssen in ihrer lokalen Domäne gebunden sein.
 b. PRINZIP B:
 Pronomina müssen in ihrer lokalen Domäne frei sein.

Hier entsteht aber genau dasselbe Dilemma wie oben bei der Diskussion der syntaktischen Komparativregel B': Dieselbe Information über lokale Domänen geht in zwei verschiedene Beschränkungen ein, einmal als Lizenzierungskontext, einmal als Verbotskontext. Diese Redundanz ist oft bemerkt worden, und es ist verschiedentlich versucht worden, die Daten dadurch zu erfassen, daß (etwas in der Art von) PRINZIP A angenommen, auf PRINZIP B jedoch verzichtet wird, und die Effekte dieser Beschränkung durch Präferenzprinzipien abgeleitet werden, die anaphorische Pronomina Personalpronomen vorziehen; vgl. Burzio (1991), Fanselow (1991) und Richards (1997). Die folgende Darstellung orientiert sich an Fanselow (1991).

Es sei also zunächst vorausgesetzt, daß PRINZIP A gilt, PRINZIP B dagegen nicht.²⁸ Personalpronomen treten nun genau dann zum Ausdruck von Koreferenz (bzw. Bindung) auf, wenn Anaphern aufgrund von Prinzip A nicht auftreten können. Dies bedeutet: Innerhalb einer lokalen Domäne muß ein Pronomen nicht deshalb frei sein, weil es bei Bindung eine spezielle Beschränkung verletzt, sondern deshalb, weil der Kandidat mit dem Pronomen blockiert wird durch den konkurrierenden Kandidaten mit der Anapher. Fanselow implementiert diesen Wettbewerb durch das Prinzip der Vollständigen Inklusion (“Proper Inclusion Principle”, PIP), das ein anaphorisches Pronomen im Zweifelsfall einem Personalpronomen vorzieht.²⁹

²⁸Dies ist eine Vereinfachung, denn Fanselow (1991) ersetzt PRINZIP A durch ein anderes Prinzip (das der “Vollständigen Spezifikation”), das zum Teil etwas andere und – wie Fanselow zeigt – bessere Vorhersagen macht als PRINZIP A. Diese Unterschiede sind jedoch für die gegenwärtigen Zwecke irrelevant.

²⁹(48) ist die vereinfachte Variante; vgl. Fanselow (1991, 275). Das Prinzip ist sehr allgemein formuliert. Dies liegt daran, daß es bei Fanselow noch eine Reihe weiterer Konsequenzen hat, z.B. bezüglich des Auftauchens von Spuren und resumptiven Pronomina, oder aber bezüglich der unten noch zu diskutierenden VERMEIDE PRONOMINA-Effekte (vgl. Chomsky (1981)).

(48) *Prinzip der Vollständigen Inklusion* (PIP):

Konkurrieren *ceteris paribus* in einer strukturellen Konfiguration miteinander zwei Merkmalszuweisungsmechanismen α , β , so kann α nicht angewendet werden, falls allgemein die Domäne der Anwendung von α die Domäne der Anwendung von β echt umfaßt.

Relevante Merkmalszuweisungsmechanismen sind hier die Zuweisung des Merkmals [+anaphorisch] an eine NP (β) und die Zuweisung des Merkmals [+pronominal] an eine NP (α); kurz: Reflexivierung und Pronominalisierung. Die Domänen von Reflexivierung und Pronominalisierung sind nicht identisch. Pronominalisierung ist an sich überall möglich; die Domäne für Reflexivierung ist durch PRINZIP A festgelegt. Das bedeutet, daß Anaphern in weniger Positionen auftauchen können als Pronomina; die Domäne für Reflexivierung ist echt in der Domäne für Pronominalisierung enthalten. Somit blockiert das Prinzip der Vollständigen Inklusion Pronominalisierung überall da, wo Reflexivierung möglich ist.

Die Rolle der Domänen für Merkmalszuweisungsmechanismen in der Definition des Prinzips der Vollständigen Inklusion entspricht genau dem, was bei der Diskussion von Blockade-Theorien über das Konzept von Spezifität von syntaktischen Kandidaten gesagt wurde. Somit ist das Prinzip der Vollständigen Inklusion eine Variante des Blockade-Prinzips, und daher sind PIP-Syntaxen letztlich Varianten von Blockade-Syntaxen. Nicht verwunderlich ist dementsprechend auch, daß sich hier ebenfalls eine optimalitätstheoretische Rekonstruktion anbietet.

4.3.2. *Optimalitätstheoretische Rekonstruktion*

Eine erste Voraussetzung ist, daß Kandidatenmengen so definiert sind, daß zwei Kandidaten im Wettbewerb stehen können, wenn sie sich darin unterscheiden, daß einmal in einer gegebenen Position ein anaphorisches Pronomen erscheint, und ein andermal ein Personalpronomen. Wiederum scheint also der Begriff der identischen Numeration nicht geeignet. Die Analyse beruht dann auf der Einführung einer Beschränkung wie (49), die PRINZIP A untergeordnet ist (PRINZIP B bleibt aufgegeben):

(49) *PERS-PRON (“Vermeide Personalpronomina”):

Personalpronomina sind verboten.

Die Argumentation läuft komplett analog zur optimalitätstheoretischen Rekonstruktion der Blockade-Analyse von englischer Komparativbildung. *PERS-PRON übernimmt dabei die Rolle von KOMP-REAL; diese Beschränkung setzt den Preis fest, der immer für das Auftreten eines Personalpronomens zu zahlen ist, und präferiert so zunächst einmal anaphorische Pronomina. Wenn also anaphorische Pronomina und Personalpronomina in einer gegebenen Position beide die höher geordnete Beschränkung PRINZIP A erfüllen können (bei lokaler A-Bindung im ersteren Falle, leer im letzteren), dann gewinnt der Kandidat mit dem anaphorischen Pronomen; vgl. Tabelle T_{4.3}.

T_{4.3}: Reflexivierung bei Bindung in lokalen Domänen

Kandidaten	PRINZIP A	*PERS-PRON
☞K ₁ : ... Maria ₁ auf sich ₁ stolz ist		
K ₂ : ... Maria ₁ auf sie ₁ stolz ist		*!

Wenn dagegen das anaphorische Pronomen PRINZIP A verletzt, kann die Verletzung der tiefer geordneten Beschränkung *PERS-PRON durch Personalpronomina in Kauf genommen werden; vgl. Tabelle T_{4.4}.

T_{4.4}: Pronominalisierung ansonsten

Kandidaten	PRINZIP A	*PERS-PRON
K ₁ : ... Marias ₁ Bruder auf sich ₁ stolz ist	*!	
☞K ₂ : ... Marias ₁ Bruder auf sie ₁ stolz ist		*

Diese Argumentation setzt voraus, daß es im Wettbewerb T_{4.4} nicht noch andere Kandidaten gibt, die sowohl PRINZIP A als auch *PERS-PRON erfüllen. Insbesondere gälte es, etwas zur scheinbar vorliegenden optionalen Alternation von K₂ mit (50) zu sagen:

(50) daß Marias₁ Bruder auf Maria₁ stolz ist

Der (50) entsprechende Kandidat darf entweder nicht im selben Wettbewerb sein oder aber kein besseres Beschränkungsprofil als K₂ in T_{4.4} haben.

Diese Analyse hat eine mögliche weitere Konsequenz (vgl. Fanselow (1991)). Nehmen wir einmal an, es gibt in einem Wettbewerb keinen einzigen Kandidaten, der PRINZIP A erfüllen kann, z.B., weil es im morphologischen Paradigma Lücken gibt und an bestimmten Stellen gar keine anaphorischen Pronomina zur Verfügung stehen.³⁰ Dann wird korrekt vorhergesagt, daß trotz Bindung in der lokalen Domäne ein Personalpronomen benutzt werden kann:

- (51) a. daß er₁ auf sich₁ stolz ist
 b. *daß er₁ auf ihn₁ stolz ist
 c. daß du₁ auf dich₁ stolz bist
 d. daß wir₁ auf uns₁ stolz sind

Der Grund ist hier, daß durch das Fehlen bestimmter Formen im Paradigma die Kandidatenmenge kleiner wird, und so ein Kandidat mit Pronominalisierung, der normalerweise (durch das Prinzip der Vollständigen Inklusion wie durch die Optimalitätstheoretische Rekonstruktion) zugunsten eines Kandidaten mit Reflexivierung

³⁰Dies entspricht allerdings nicht unbedingt dem Slogan "Das funktionale Lexikon ist Sklave der Syntax" von Seite 79.

blockiert wird, möglich wird, weil es eben keine bessere Form gibt, die verfügbar wäre: Wenn z.B. der zu K_1 aus $T_{4.3}$ analoge Kandidat bei Pronomina der 2. Person gar nicht im Wettbewerb auftaucht, dann wird der zu K_2 analoge Kandidat erlaubt.

Damit ist die Betrachtung anderer Wettbewerbssyntaxen abgeschlossen. Im folgenden möchte ich zeigen, daß Konzepte wie Wettbewerb, Beschränkungsverletzbarkeit und Beschränkungsordnung nicht nur explizit in der Syntaxforschung eine Rolle gespielt haben, sondern daß es auch das Phänomen versteckter Optimalität gibt, also syntaktische Analysen, die mehr oder weniger implizit optimalitätstheoretische Argumentationsschemata aufweisen. Dazu habe ich vier Beispiele aus der GB- bzw. Prinzipien- und Parameter-Literatur ausgewählt, stellvertretend für viele weitere. In gewisser Weise kann man sagen, daß diese Analysen nur auf die Ankunft der Optimalitätstheorie gewartet haben und eigentlich mit den sie umgebenden theoretischen Annahmen kaum vereinbar waren.

5. Versteckte Optimalität

5.1. Vermeide Pronomina bei Chomsky (1981)

5.1.1. Die Analyse

Die klassische GB-Theorie in der Version von Chomsky (1981) enthält mindestens einen Bereich, in dem die (Un-)Grammatikalität eines Satzes nicht nur auf der Basis seiner internen Eigenschaften ermittelt werden kann. Normalerweise sind gemäß dieser Theorie PRO und lexikalische Pronomina komplementär verteilt. Das liegt daran, daß PRO aufgrund des PRO-Theorems auf rektionsfreie Positionen beschränkt ist (z.B. die Subjektposition von Kontrollinfinitiven), während lexikalische Pronomina Kasus erhalten müssen und, da die Zuweisung von Kasus normalerweise unter Rektion erfolgt, nur in regierten Positionen auftauchen. Es gibt jedoch eine Position, in der die GB-Theorie zunächst einmal sowohl das leere Element PRO zuläßt (weil sie unregiert ist), als auch ein lexikalisches Pronomen (weil es dort Kasus ohne Rektion zugewiesen bekommen kann, und zwar über eine spezielle Kasuszuweisungsregel). Diese Position ist die Subjektposition von Gerundivkonstruktionen im Englischen. Betrachten wir zunächst die Möglichkeit von PRO in dieser Position.

- (52) a. John₁ would much prefer [PRO₁ going to the movie]
 b. *John₁ would much prefer [PRO_{2/arb} going to the movie]

Wie (53-a) gegenüber (53-b) zeigt, muß PRO hier mit dem Matrixsubjekt koindiziert sein. Dieser Kontrast folgt aus der Annahme, daß ein in einem satzwertigen Objekt enthaltenes PRO immer im nächsthöheren Satz kontrolliert werden muß. Dies läßt sich als Kontrollregel festhalten (vgl. Manzini (1983)):

- (53) KONTROLLREGEL:
 Wenn PRO unmittelbar von einem deklarativen (d.h., [-w] markierten) satz-

wertigen Objekt α dominiert wird, dann muß es innerhalb der minimalen CP, die α dominiert, von einem Antezedens kontrolliert werden.

Wie (54-b) zeigt, ist auch ein lexikalisches Pronomen in der Subjektposition von englischen Gerundivkonstruktionen möglich. (54-a) ist nun der interessante Fall; er zeigt, daß das lexikalische Pronomen nicht mit dem Matrixsubjekt koindiziert sein kann.

- (54) a. *John₁ would much prefer [his₁ going to the movie]
 b. John₁ would much prefer [his₂ going to the movie]
 c. John₁ would much prefer [his₁ book]

Die Ungrammatikalität von (54-a) folgt erst einmal aus nichts. Insbesondere zeigt (54-c), daß *his*₂ in (54-a) Bindungsprinzip B erfüllen kann, und (54-b), daß *his*₂ in (54-a) Kasus erhalten kann. Chomsky (1981, 65) schlägt nun vor, daß die Ungrammatikalität aus einem Wettbewerb von (54-a) mit der PRO-Variante (52-a) folgt: Gerade weil letzteres möglich ist, ist ersteres unmöglich. Die Beschränkung, die diese Erklärung leistet, ist das Prinzip *Vermeide Pronomina* ("Avoid Pronoun"):

- (55) *Vermeide Pronomina* (Chomsky (1981)):
 Lexikalische Pronomina werden durch leere Pronomina blockiert, falls dies möglich ist.

Der Zusatz "falls dies möglich ist" macht klar, daß es sich hier nicht um eine übliche, lokale Beschränkung handelt. *Vermeide Pronomina* ist vielmehr translokal (im Einklang damit ist die Beschränkung auch ohne Kapitälchen gesetzt): Um entscheiden zu können, ob ein gegebener Satz mit lexikalischem Pronomen diese Beschränkung erfüllt, muß man herausfinden, ob es einen anderen wohlgeformten, konkurrierenden Satz gibt, in dem stattdessen ein PRO erscheint. Konkret nimmt Chomsky (1981) an, daß bei Koindizierung von Matrix- und Gerundivsubjekt, wo PRO die KONTROLLREGEL erfüllen kann (vgl. (52-a)), ein lexikalisches Pronomen durch *Vermeide Pronomina* blockiert wird (vgl. (54-a)), und daß bei Kontraindizierung von Matrix- und Gerundivsubjekt, wo PRO die KONTROLLREGEL nicht erfüllen kann (vgl. (52-b)), ein lexikalisches Pronomen nicht mehr durch *Vermeide Pronomina* blockiert werden kann und daher möglich wird (vgl. (54-b)).

5.1.2. Optimalitätstheoretische Rekonstruktion

Diese Analyse ist leicht optimalitätstheoretisch rekonstruierbar. *Vermeide Pronomina* kann wie folgt als lokale, verletzbare Beschränkung reformuliert werden:

- (56) *PRON ("Vermeide Pronomina"):
 Lexikalische Pronomina sind verboten.

Die relevante Ordnung lautet: KONTROLLREGEL \gg *PRON. Den Wettbewerb bei Koindizierung des Gerundivsubjekts mit dem Matrixsubjekt zeigt Tabelle T_{4.5}.

T_{4.5}: PRO vs. Pronomen in Gerundivkonstruktionen 1: Koindizierung

Kandidaten	KONTROLLREGEL	*PRON
K ₁ : John ₁ ... [his ₁ going to the movie]		*!
☞ K ₂ : John ₁ ... [PRO ₁ going to the movie]		

Da hier K₁ gar kein PRO enthält und K₂ die KONTROLLREGEL respektiert, ergibt sich, daß der Wettbewerb durch die Erfüllung bzw. Verletzung einer einzigen Beschränkung (*PRON) entschieden wird. Der Fall liegt anders, wenn keine Koindizierung von Gerundiv- und Matrixsubjekt besteht. Wie Tabelle T_{4.6} zeigt, ist hier die Ordnung KONTROLLREGEL \gg *PRON wesentlich. K₁ verletzt *PRON, ist aber dennoch optimal, weil K₂ gegen die höher geordnete KONTROLLREGEL verstößt.

T_{4.6}: PRO vs. Pronomen in Gerundivkonstruktionen 2: Fehlende Koindizierung

Kandidaten	KONTROLLREGEL	*PRON
☞ K ₁ : John ₁ ... [his ₂ going to the movie]		*
K ₂ : John ₁ ... [PRO ₂ going to the movie]	*!	

In einem optimalitätstheoretischen Rahmen läßt sich die Beschränkung *PRON somit analog zu etwa Grimshaw's (1997) ÖKON oder den oben eingeführten Beschränkungen KOMP-REAL und *PERS-PRON auffassen: *PRON ist tief geordnet und bleibt häufig konsequenzlos, weil die Effekte meistens von anderen, höher geordneten Beschränkungen überlagert werden. Wenn höher geordnete Beschränkungen aber nicht zwischen zwei Kandidaten unterscheiden, prognostiziert *PRON korrekt den unmarkierten Fall.³¹

Wie oben bei der optimalitätstheoretischen Rekonstruktion von PIP-Effekten, so muß auch im vorliegenden Fall gewährleistet sein, daß die Konkurrenten eines Beispiels wie (57-a) in (57-bcd) alle gegen noch höher als *PRON geordnete Beschränkungen verstoßen. Für (57-b) kommt hier das PRO-THEOREM infrage (PRO in (57-b) ist regiert).³² Für (57-c) könnten wir ganz traditionell annehmen, daß eine (hoch geordnete) Beschränkung PRO-LIZENSIERUNG fatal verletzt ist (die *pro* nur bei hinreichend reichhaltiger Verbmorphologie erlaubt). Alternativ dazu könnten wir aber auch mit Grimshaw & Samek-Lodovici (1995; 1998) und Samek-Lodovici (1996) annehmen, daß Pro-Drop keine leere Kategorie *pro* involviert, sondern die Nicht-Realisierung von Input-Elementen (vgl. Abschnitt 7.4.2 aus Kapitel 2). Die

³¹Man beachte, daß an sich nichts dagegen spricht, die allgemeine Beschränkung *PRON und die speziellere Beschränkung *PERS-PRON gleichzeitig anzunehmen. Diese beiden Beschränkungen stehen als Elemente einer Subhierarchie in einer Stringenzrelation.

³²Natürlich kann ein Theorem keine primitive Beschränkung sein. PRO-THEOREM steht hier für ein System von Beschränkungen, das PRO auf unregierte Positionen festlegt.

Beschränkung, die die Realisierung lexikalischer Pronomina verlangt (REAL) wäre dann nicht nur wie gesehen TOP-TIL übergeordnet, sondern auch *PRON. Was schließlich (57-d) betrifft, so ist ein offensichtlicher Kandidat für eine fatal verletzte höher geordnete Beschränkung STR-BD (“Strikte Bindung”), wodurch für jede Spur ein c-kommandierendes, koindiziertes Antezedens erfordert wird (vgl. Kapitel 2 und Vikner (1999)).

- (57) a. Heute hat sie gearbeitet
 b. *Heute hat PRO gearbeitet
 c. *Heute hat *pro* gearbeitet
 d. *Heute hat t gearbeitet

Eine interessante Konsequenz ergibt sich für die Definition von Kandidatenmengen. Die Wettbewerbe in $T_{4.5}$ und $T_{4.6}$ setzen voraus, daß zwei Kandidaten miteinander konkurrieren können, die sich dadurch unterscheiden, daß ein Pronomen das eine Mal lexikalisch und das andere Mal nicht-lexikalisch, als PRO, realisiert wird. Eine Definition des Konzepts der Kandidatenmenge, die auf dem Konzept der identischen Numeration (also einem identischen lexikalischen Material) beruht, scheint hier also nicht ausreichend (zumindest gilt dies, solange wir PRO als echte Kategorie betrachten, und nicht als das Resultat einer Nicht-Realisierung eines lexikalischen Pronomens aus dem Input). Darüber hinaus muß sichergestellt werden, daß K_1 aus $T_{4.6}$ nicht mit K_2 aus $T_{4.5}$ konkurriert, d.h., daß (52-a) nicht (54-b) als suboptimal blockieren kann. Dies kann gewährleistet werden durch die Annahme, daß unterschiedliche Indizes auf einer Kategorie bei zwei Kandidaten immer dazu führen, daß die beiden Kandidaten zu verschiedenen Kandidatenmengen gehören. Mit anderen Worten: PRO_1 und $Pronomen_1$ können in zwei verschiedenen Kandidaten in derselben Kandidatenmenge auftauchen, PRO_1 und $Pronomen_2$ oder PRO_1 und PRO_2 dagegen nicht.

Erwähnt sei abschließend noch, daß Haegeman (1994, 217) Chomskys (1981) Erklärung des Kontrastes von (52-a) und (54-a) in englischen Gerundivkonstruktionen auf Pro-Drop-Phänomene übertragen hat. Die so entstehende Analyse ist bei allen Unterschieden im Detail im Grundaufbau durchaus vergleichbar dem, was Grimshaw & Samek-Lodovici (1995; 1998) und Samek-Lodovici (1996) in einem optimalitätstheoretischen Rahmen vorgeschlagen haben. Dabei haben *Vermeide Pronomina* bei Haegeman und TOP-TIL bei Grimshaw und Samek-Lodovici eine ähnliche Funktion (insofern, als sie in normalen Pro-Drop-Kontexten in einer Sprache wie dem Italienischen lexikalische Pronomina verbieten bzw. deren Nicht-Realisierung im Output erzwingen).

5.2. Kontrolle bei Stechow & Sternefeld (1988)

5.2.1. Die Analyse

Betrachten wir nochmals die KONTROLLREGEL, wie sie in der GB-Theorie postuliert worden ist:

(58) KONTROLLREGEL:

Wenn PRO unmittelbar von einem deklarativen (d.h., [-w] markierten) satzwertigen Objekt α dominiert wird, dann muß es innerhalb der minimalen CP, die α dominiert, von einem Antezedens kontrolliert werden.

Einige klassische Daten, die durch diese Beschränkung erklärt werden, sind in (59) versammelt. (59-a) weist einen deklarativen Objektinfininitiv und daher auch obligatorische Kontrolle von PRO auf. In (59-bc) dagegen erscheint PRO in einem Subjektinfininitiv. Daher muß keine Kontrolle erfolgen, und PRO kann entweder frei auf im Diskurs präsenste Personen referieren oder aber arbiträr interpretiert werden. Dasselbe gilt für den interrogativen Objektinfininitiv im englischen Beispiel (59-d).³³

- (59) a. Karl₁ versuchte [PRO₁ sich zu benehmen]
 b. [PRO sich in der Öffentlichkeit besser zu benehmen] würde Karls Ruf gut tun
 c. Maria weiß [daß [PRO freundlich zu sein] dem Karl helfen würde]
 d. John asked Bill [how PRO to prove the theorem without using the axiom of choice]

Diese Daten stehen also im Einklang mit (58). Es gibt aber einige Ausnahmen. Eine ist (60-a), wo eine arbiträre Interpretation möglich ist, obwohl PRO in einem deklarativen Objektinfininitiv steckt. Tatsächlich zeigt (60-b), daß hier die gemäß (58) erwartbare Koindizierung mit dem Matrixsubjekt verboten ist.

- (60) a. Fritz₁ ordnete an [PRO_{arb} den Saal zu verlassen]
 b. *Fritz₁ ordnete an [PRO₁ den Saal zu verlassen]

Stechow & Sternfeld (1988, 315) bemerken, daß man angesichts eines Datums wie (60-a) entweder ein leeres, arbiträr interpretiertes Objektpronomen (*pro*) im Matrixsatz annehmen muß (was sie als theoretisch unattraktiv ablehnen), oder aber sagen muß, daß "lexikalische Bedingungen stärker sind als rein strukturelle; d.h., lexikalische Eigenschaften von Kontrollverben können die strukturelle Bedingung [(58)] gelegentlich außer Kraft setzen."

5.2.2. Optimalitätstheoretische Rekonstruktion

Letzteres ist nichts anderes als die Umschreibung einer optimalitätstheoretischen Analyse: Es liegt eine implizite Beschränkungsordnung vor, die man wie folgt explizit machen kann. Die Beschränkung, die Treue zu den lexikalisch vermerkten Kontrolleigenschaften verlangt (TREUE(LEX-KONTROLLE)), ist höher geordnet als die KONTROLLREGEL. Letztere wird somit nur erfüllt, soweit das möglich ist, ohne lexikalisch festgelegte Kontrolleigenschaften zu verletzen.

³³Das Deutsche besitzt keine W-Infinitive, daher ist der Bezug auf ein englisches Beispiel zur Illustration des Phänomens notwendig.

(61) TREUE(LEX-KONTROLLE):

Lexikalische Kontrolleigenschaften werden in der Syntax respektiert.

Tabelle T_{4.7} zeigt, daß bei Irrelevanz von TREUE(LEX-KONTROLLE) die KONTROLLREGEL den unmarkierten Fall bei deklarativen Objektinfininitiven erfaßt.

T_{4.7}: Strukturelle Kontrolle im Deutschen

Kandidaten	TREUE(LEX-KONTROLLE)	KONTROLLREGEL
☞K ₁ : Karl ₁ versuchte [PRO ₁ ...]		
K ₂ : Karl ₁ versuchte [PRO _{arb} ...]		*!

Den durch TREUE(LEX-KONTROLLE) bestimmten Fall zeigt Tabelle T_{4.8}.

T_{4.8}: Lexikalische Kontrolle im Deutschen

Kandidaten	TREUE(LEX-KONTROLLE)	KONTROLLREGEL
K ₁ : Fritz ₁ ordnete an [PRO ₁ ...]	*!	
☞K ₂ : Fritz ₁ ordnete an [PRO _{arb} ...]		*

Man beachte, daß diese Analyse anders als die vorigen voraussetzt, daß zwei Kandidaten miteinander konkurrieren können, die verschiedene Indizes auf ein und derselben Kategorie aufweisen (hier: PRO_{arb} und PRO₁), d.h., die eine unterschiedliche Interpretation haben. Dies bedeutet, daß eine Kompatibilität von z.B. Chomskys Analyse von gegenüber einem lexikalischen Pronomen präferiertem PRO in englischen Gerundivkonstruktionen mit der von Stechow und Sternefeld anvisierten Analyse von strukturell vs. lexikalisch determinierter Interpretation von PRO zunächst einmal nicht gegeben ist. Daraus können wir schließen, daß entweder eine der beiden Analysen aufzugeben ist oder noch weitere Annahmen (z.B. bzgl. Neutralisierung; vgl. Kapitel 3, Abschnitt 3) notwendig sind. Ich werde in Kapitel 8 auf dieses Thema zurückkommen.

Allgemein läßt sich aber festhalten, daß die soeben vorgestellte Analyse im Grundaufbau auf viele andere Phänomene übertragbar ist, bei denen es so aussieht, als überlagerten lexikalische Idiosynkrasien strukturelle Regularitäten; vgl. insbesondere auch die Ausführungen zur Kasustheorie (inhärenter vs. struktureller Kasus) in Kapitel 7.

5.3. W-Bewegung und Komplementierer-Finalität bei Kayne (1994)

5.3.1. Die Analyse

Viele Sprachen weisen keine S-strukturelle Bewegung von W-Phrasen nach SpecC auf. Dazu gehören u.a. das Japanische und das Koreanische. Kayne (1994, 54) schlägt eine neue Erklärung hierfür vor. Diese Erklärung beruht auf der Annahme des Axioms der linearen Korrespondenz ("Linear Correspondence Axiom", LCA),

das erzwingt, daß ein Kopf links von seinem Komplement steht. Dies bedeutet, daß ein satzfinaler Komplementierer, wie ihn etwa das Japanische besitzt, nicht rechts von seinem IP-Komplement basisgeneriert worden sein kann. Vielmehr ist hier, so Kayne, der Komplementierer wie auch im Englischen links von IP basisgeneriert worden; auf der S-Struktur wird jedoch IP im Japanischen wie in (62) gezeigt obligatorisch nach SpecC angehoben.

$$(62) \quad [_{CP} - [_{C'} C IP_1]] \rightarrow [_{CP} IP_1 [_{C'} C t_1]]$$

Die Idee ist nun, diesen Vorschlag in Beziehung zu setzen mit dem Verbot von S-struktureller W-Bewegung in dieser Sprache: Wenn IP obligatorisch in SpecC steht, dann wird damit die Landestelle für W-Bewegung blockiert, und, so die Annahme, das Auftreten von W-in-situ ist erklärt, ohne Rekurs auf spezielle Parametrisierung des W-Kriteriums oder andere Stipulationen; vgl. (63).

$$(63) \quad [_{CP} [_{IP_1} \dots NP_{[+w]} \dots] [_{C'} C t_1]]$$

Wenn nun die Tatsache, daß IP in SpecC steht, dafür verantwortlich gemacht werden soll, daß das Japanische keine W-Bewegung aufweist, dann ist klar, daß vorausgesetzt ist, daß es im Prinzip auch im Japanischen einen Auslöser für die Bewegung von W-Phrasen nach SpecC gibt – wäre dem nicht so, dann könnte es ja keine Beziehung zwischen IP-Anhebung und fehlender W-Bewegung geben, und es wäre schlicht ein Zufall, daß gerade die Sprache, die per Annahme keine Landestelle für W-Bewegung hat, auch keine braucht. Es scheint daher im Sinne der Kayneschen Argumentation zu sein, wenn wir annehmen, daß es im Japanischen wie im Englischen einen Auslöser für S-strukturelle W-Bewegung nach SpecC gibt. Zumindest für die Zwecke des Arguments sei dies das W-Kriterium.³⁴ In (63) ist eine solche Bewegung dann aufgrund der Besetzung von SpecC durch IP nicht möglich.

So weit, so gut. Wenn nicht mehr gesagt wird, bedeutet das in einer Standardtheorie mit unverletzbaren Beschränkungen nun aber nicht, daß die W-Phrase dann eben in situ verharrt; vielmehr wird die falsche Vorhersage gemacht, daß (63) ungrammatisch ist. Dies ist natürlich ein fatales Resultat: Es sollte nicht möglich sein, im Japanischen Ergänzungsfragen zu bilden. Was man also mindestens braucht, ist ein um eine Ausnahmeklausel erweitertes W-Kriterium: Eine W-Phrase muß nach SpecC bewegt werden, falls diese Position für Bewegung zu Verfügung steht und nicht anderweitig besetzt ist. Spätestens hier wird aber offensichtlich, daß dieses Argument implizit eine Beschränkungsordnung voraussetzt.

³⁴Derselbe Punkt ließe sich auch machen, wenn ein anderer Auslöser für W-Bewegung identifiziert würde, z.B. der Zwang, ein starkes W-Merkmal zu überprüfen, vgl. Chomsky (1993; 1995).

5.3.2. Optimalitätstheoretische Rekonstruktion

Die Beschränkung, die IP nach SpecC zwingt (hier IP-KRIT genannt), muß im Japanischen höher geordnet sein als W-KRIT, was S-strukturelle W-Bewegung auslöst (vgl. (25) aus Kapitel 1). Im Englischen gilt entsprechend die umgekehrte Ordnung (wobei IP-KRIT noch unter ÖKON stehen muß, um IP-Bewegung auch in Deklarativsätzen zu verhindern.).

- (64) a. IP-KRIT (“IP-Kriterium”):
IP muß auf der S-Struktur in SpecC stehen.
- b. W-KRIT:
Eine W-Phrase muß auf der S-Struktur in SpecC stehen.

Nummehr folgt, daß im Japanischen der Kandidat, in dem IP in situ bleibt und die W-Phrase bewegt werden kann, blockiert wird zugunsten des Kandidaten, in dem IP nach SpecC bewegt wird und W-Bewegung unterbleibt; vgl. Tabelle T_{4.9}.

T_{4.9}: Das Fehlen von W-Bewegung im Japanischen

Kandidaten	IP-KRIT	W-KRIT
☞ K ₁ : [CP [IP ₁ ... NP _[+w] ...] [C' C t ₁]]		*
K ₂ : [CP NP _[+w] [C' C [IP ₁ ... t _w ...]]]	*!	

Bei Kayne (1994, 54) hat man es also mit einer unvollständigen Argumentation zu tun, die erst dann vollständig ist, wenn die darin vorausgesetzte Verletzbarkeit und Geordnetheit von Beschränkungen explizit gemacht wird. Es versteht sich nun von selbst, daß aus der Korrektheit der Argumentation nicht folgt, daß auch die Analyse korrekt ist; darum geht es hier jedoch nicht primär. Wichtig ist, in diesem Argumentationstyp die versteckte Optimalität zu erkennen. Und gerade die unvollständigen Argumentationen vom Typ (65-a) sind in der GB-orientierten syntaktischen Literatur Legion, obwohl sie dort strenggenommen wie in (65-b) aussehen müßten.

- (65) a. In einem Satz müssen zwei Elemente α , β in die Position γ verschoben werden; sie können dort aber nicht beide stehen. Also bleibt β in situ.
- b. In einem Satz müssen zwei Elemente α , β in die Position γ verschoben werden; sie können dort aber nicht beide stehen. Also ist der Satz ungrammatisch.

5.4. Phonologische Realisierung in Kopf-Ketten bei Roberts (1997)

5.4.1. Die Analyse

Die Ausgangsbeobachtung ist hier, daß in italienischen kohärenten Konstruktionen (Restrukturierungskonstruktionen) zum einen Evidenz dafür besteht, daß die Satzgrenze des Infinitivs durch Kopf-Bewegung durchlässig gemacht wird (so daß

zum Beispiel lange Bewegung klitischer Pronomina möglich ist; vgl. Baker (1988)), und daß aber andererseits das eingebettete Infinitivverb und das Matrixverb keinen komplexen Kopf bilden zu können scheinen (weil manche Adverbien, W-Phrasen und herumtreibende Quantoren zwischen beiden Verben intervenieren können); vgl. den herumtreibenden Quantor *tutti* in (66):

- (66) Gianni li_1 vuole tutti leggere t_1
Gianni sie will alle lesen

Roberts' (1997) Idee ist, daß in (66) tatsächlich S-strukturelle Kopf-Bewegung erfolgt ist. Dies ist für die Transparenz des Infinitivs verantwortlich, so daß sich das klitische Pronomen *li* ohne Lokalitätsverletzung in den Matrixsatz bewegen kann. Das offensichtliche Problem ist dann, daß man in (66) keinen komplexen Kopf *vuole-leggere* sehen kann.

Um dieses Problem zu lösen, nimmt Roberts an, daß Kopf-Bewegung die Zurücklassung einer vollständigen Kopie als Spur involviert; dies setzt die Kopiertheorie der Bewegung voraus (vgl. Chomsky (1993; 1995)). An welcher Stelle die phonetische Realisierung in einer Kette erfolgt, ist zunächst einmal von der Bewegung unabhängig – im Prinzip kann jede der Kopien ausgesprochen werden. Die PF-Realisierung wird durch die folgenden beiden Beschränkungen geregelt.³⁵

- (67) a. *KOMP-KOPF (“Vermeide komplexe Köpfe”):
* $[_X W_1 W_2]$, wobei W_n ein morphologisches Wort ist.
b. KOPF-REAL (“Kopf-Realisierung”):
Ein Kopf wird in der höchsten Position seiner Kette phonologisch realisiert, die mit *KOMP-KOPF kompatibel ist.

Als Konsequenz kann nun in (66) tatsächlich Inkorporation des Infinitivs in das Matrixverb vorliegen. In der Zielposition ist aber wg. *KOMP-KOPF kein Platz mehr, also wird der Infinitiv in der nächsttieferen Kopfposition (bei Roberts AGR_S) realisiert. Eine zusätzliche Annahme, die Roberts (1997, Fußnote 1) noch macht, ist die, daß (67-a) parametrisiert, also in manchen Sprachen abschaltbar ist. Dies ist notwendig, um Phänomene wie Nominalinkorporation in polysynthetischen Sprachen zu erlauben, wo kein Weg an der Konklusion vorbei führt, daß ein morphologisch komplexer Kopf gebildet wird.

5.4.2. Optimalitätstheoretische Rekonstruktion

Es ist klar, wie die beiden Beschränkungen in (67) optimalitätstheoretisch zu rekonstruieren sind. Auf die Ausnahmeklausel in (67-b) kann verzichtet werden: KOPF-REAL läßt sich nunmehr wie in (68) fassen.

³⁵Vgl. Roberts (1997, 426). Die Namen dieser Beschränkungen sind von mir hinzugefügt.

(68) KOPF-REAL (revidiert):

Ein phonologisch realisierter Kopf darf in seiner Kette nicht von einer Kopie c-kommandiert werden.

Darüber hinaus muß *KOMP-KOPF in polysynthetischen Sprachen nicht “abgeschaltet” werden; der erwünschte Effekt entsteht schlicht durch Beschränkungsumordnung. T_{4.10} macht dann deutlich, warum die Realisierung des Infinitivs in der höchstmöglichen Position, die nicht *KOMP-KOPF verletzt, optimal ist, also K₁. Bei Umordnung der beiden Beschränkungen wäre demgegenüber immer eine phonologische Realisierung des Infinitivs in einem Wort zusammen mit dem Matrixverb optimal, also K₂.³⁶

T_{4.10}: *Restrukturierung im Italienischen*

Kandidaten	*KOMP-KOPF	KOPF-REAL
☞K ₁ : Gianni li ₁ vuole t ₂ tutti [AGR _S P leggere ₂ t ₂ t ₁]		*
K ₂ : Gianni li ₁ vuole leggere ₂ tutti [AGR _S P t ₂ t ₂ t ₁]	*!	
K ₃ : Gianni li ₁ vuole t ₂ tutti [AGR _S P t ₂ leggere ₂ t ₁]		**!

5.4.3. Mehrfachfragen bei Grewendorf (1997)

Eine ganz ähnliche optimalitätstheoretische Rekonstruktion bietet sich an bei Grewendorfs (1997) Vorschlag für W-Bewegung in Mehrfachfragen im Deutschen. Grewendorf argumentiert, daß in Mehrfachfragen im Deutschen gar nicht eine W-Phrase nach SpecC bewegt wird und die restlichen W-Phrasen in situ verharren, wie es traditionell angenommen wird (vgl. (69-a)), sondern daß tatsächlich alle W-Phrasen auf der S-Struktur bewegt werden (vgl. (69-b)); somit wäre Deutsch in diesem Bereich dem Bulgarischen ähnlich (vgl. Tabelle T_{2.4} aus Kapitel 2).

- (69) a. Wer₁ hat t₁ wem₂ was₃ gesagt ?
 b. Wer₁ wem₂ was₃ hat t₁ t₂ t₃ gesagt ?

Nehmen wir, wiederum für die Zwecke des Arguments, einmal an, daß sich diese Analyse tatsächlich unabhängig motivieren läßt.³⁷ Dann fällt zunächst auf, daß (69-b) selbstverständlich so im Deutschen niemals ausgesprochen werden kann. Grewendorfs Vorschlag ist, daß durch sukzessive Adjunktion in (69-b) ein komplexes W-Cluster entstanden ist (so wie es von Rudin (1988) und Epstein (1992) postuliert,

³⁶Die phonologisch leeren Kopien, die aus der Infinitivbewegung resultieren, sind hier der Einfachheit halber auch als Spuren t₂ notiert. Das höchste Element der Kette mit Index 2 soll dabei immer an das Matrixverb adjungiert sein, das mittlere in AGR_S stehen, und das dritte in situ. Letzteres ist eine Vereinfachung, die aber unproblematisch ist, weil sich hier die Kandidaten nicht unterscheiden.

³⁷Eine Diskussion dieses Vorschlags findet sich in Sabel (1998).

in Kapitel 2 allerdings verworfen wurde), und daß nur eine W-Phrase, also z.B. *wer*₁, im vorangestellten komplexen W-Cluster phonologisch realisiert werden kann – die restlichen W-Phrasen müssen trotz S-strukturell erfolgter Bewegung phonologisch in situ realisiert werden.

Dies läßt sich erklären, wenn es (i) eine Beschränkung gibt, die *KOMP-KOPF analog ist und mehrfache phonologische Realisierung in einem W-Cluster verbietet (*KOMP-W, “Vermeide komplexe W-Cluster”), und (ii) eine widerstreitende tiefer geordnete zweite Beschränkung existiert, die KOPF-REAL entspricht und die Aussprache von W-Phrasen in tiefen Kettenpositionen verbietet (so daß in Ergänzungsfragen eine W-Phrase in der höchsten mit der ersten Beschränkung vereinbaren Position realisiert wird, also idealerweise in der Zielposition SpecC) (W-REAL (“W-Realisierung”)).³⁸ Den optimalitätstheoretischen Wettbewerb für die PF-Realisierung von W-Phrasen in (69-b) veranschaulicht dann Tabelle T_{4.11}.³⁹

T_{4.11}: Mehrfache W-Bewegung im Deutschen

Kandidaten	*KOMP-W	W-REAL
☞K ₁ : Wer ₁ t ₂ t ₃ hat t ₁ wem ₂ was ₃ gesagt ?		**
K ₂ : Wer ₁ wem ₂ was ₃ hat t ₁ t ₂ t ₃ gesagt ?	*!*	
K ₃ : Wer ₁ wem ₂ t ₃ hat t ₁ t ₂ was ₃ gesagt ?	*!	*
K ₄ : t ₁ t ₂ t ₃ Hat wer ₁ wem ₂ was ₃ gesagt ?		***!

*KOMP-W und W-REAL müßten in einer Sprache wie dem Bulgarischen dann umgekehrt angeordnet sein, so daß die phonologische Realisierung von W-Phrasen in der jeweils höchsten Kettenposition im Zweifelsfall wichtiger ist als die Beschränkung auf eine W-Phrase im W-Cluster.

³⁸Dies entspricht allerdings nicht Grewendorfs Erklärung, die auf die unterschiedliche Stärke von Merkmalen von C und W-Phrasen im Deutschen Bezug nimmt. – Es wäre im übrigen wohl nicht aussichtslos, wollte man versuchen, *KOMP-KOPF und *KOMP-W einerseits, und KOPF-REAL und W-REAL andererseits, zu jeweils einer einzigen, allgemeineren Beschränkung zu vereinheitlichen. KOPF-REAL und W-REAL lassen sich als Teilinstanzen von Pesetskys (1998) Beschränkung ST-SP (“Stille Spur”, “Silent t”) auffassen, die in Abschnitt 2 von Kapitel 3 diskutiert worden ist. Tatsächlich schlägt Pesetsky (1998, 361-362) unabhängig eine Analyse vor, die im Grundprinzip genau der hier skizzierten Rekonstruktion der Grewendorfschen W-Bewegungstheorie für Sprachen wie Deutsch und Englisch entspricht. Da es im gegenwärtigen Kontext vor allem um das zugrundeliegende Argumentationsschema geht und nicht um eine wirkliche empirische Rechtfertigung der Idee, werde ich weitere Vereinheitlichungsversuche hier nicht unternehmen.

³⁹Vorausgesetzt ist hier natürlich, daß es andere Beschränkungen gibt, die überhaupt erst W-Bewegung aller W-Phrasen im Deutschen erzwingen. Die in Kapitel 2 in T_{2.2} bis T_{2.4} skizzierte Analyse der W-Bewegung würde dies leisten und wäre auch mit Grewendorfs Idee kompatibel, wenn das Deutsche die in (9-c) aus Kapitel 2 postulierte bulgarische Beschränkungsordnung W-KRIT » PUR-EP » ÖKON zugewiesen erhielte, und wenn *KOMP-W phonologische Realisierung von CP-Adjunkten bestrafen würde. – Zumindest für die hier betrachteten Fälle kann man annehmen, daß die Beschränkungen *KOMP-W und W-REAL unterhalb dieser vier anderen Beschränkungen angeordnet sind.

Kapitel 5

Optionalität

1. Das Phänomen und seine Konsequenzen für die Syntax

Syntaktische Optionalität ist bereits an einigen Stellen in diesem Buch aufgetaucht. In Abschnitt 10 von Kapitel 1 sind drei wesentliche Konzepte, dieses Phänomens Herr zu werden, eingeführt worden (Pseudo-Optionalität, echte Optionalität, Kopp- lung); in Abschnitt 8.1 aus Kapitel 2 ist vermerkt (aber noch nicht gerechtfertigt) worden, daß Optionalität ein potentielles Problem für die optimalitätstheoretische Syntax darstellt; und in den Abschnitten 1 und 2 aus Kapitel 3 sind die Analysen von Grimshaw und Pesetsky zur optionalen Komplementierertilgung in englischen Deklarativsätzen besprochen worden. In diesem Kapitel sollen nun die optima- litätstheoretischen Herangehensweisen an syntaktische Optionalität systematisch miteinander verglichen werden.

Beginnen wir mit einer prätheoretischen, möglichst allgemeinen Festlegung des- sen, was syntaktische Optionalität überhaupt ist: Syntaktische Optionalität liegt immer dann vor, wenn es zwei oder mehr Möglichkeiten zu geben scheint, ein und dieselbe Sache zu sagen. Ob eine solche Situation dann Probleme bereitet, hängt von der Art der gewählten syntaktischen Theorie ab. Die Transformationsgramma- tik der sechziger Jahre hat syntaktische Optionalität dadurch in den Griff bekom- men, daß sie zwischen obligatorischen und optionalen Transformationen unter- schieden hat. Syntaktische Optionalität ist auf optionale Transformationen zurückzu- führen. Einige Instanzen syntaktischer Optionalität sind im folgenden aufgeführt, zusammen mit der Benennung der entsprechenden optionalen Transformation in der klassischen generativen Grammatik. Zunächst illustriert (1) nochmals das mitt- lerweile wohl hinlänglich bekannte Phänomen des optionalen Komplementierers im Englischen. Eine klassische Annahme war schlicht, daß die Transformation Kom- plementierertilgung im Englischen optional angewendet wird.

- (1) a. I think – John will leave
b. I think that John will leave

Zweitens veranschaulicht (2) die Tatsache, daß Dativ-Anhebung im Englischen op- tional appliziert (die klassische Transformation der Dativ-Anhebung leitet (2-b) aus (2-a) durch Tilgung der Präposition *to* und Umstellung von direktem und in- direktem Objekt ab).

- (2) a. John gave [_{NP₁} a book] [_{PP₂} to Mary]
b. John gave [_{NP₂} Mary] [_{NP₁} a book] t₂

Drittens ist in Fragesätzen im (Standard-) Deutschen, die lange Abhängigkeiten involvieren, die Einsetzung eines Skopusmarkers *was* (mit partieller W-Bewegung) eine Alternative zur sukzessiv-zyklischen W-Bewegung.

- (3) a. Wen₁ glaubst du [_{CP} t'₁ daß man t₁ einladen sollte] ?
 b. Was₁ glaubst du [_{CP} wen₁ (daß) man t₁ einladen sollte] ?

Viertens appliziert W-Bewegung im Französischen optional (zumindest gilt dies im Hauptsatz und bei Argumenten; s.u.).

- (4) a. Qui₁ as-tu vu t₁ ?
 b. – Tu as vu qui₁ ?

Fünftens ist im Deutschen und in vielen anderen Sprachen Relativsatz-Extraposition aus einer NP optional:

- (5) a. daß [_{NP} eine Frau [_{CP₁} die ich mag]] zur Tür reingekommen ist
 b. daß [_{NP} eine Frau t₁] zur Tür reingekommen ist [_{CP₁} die ich mag]

Schließlich ist Scrambling optional in Sprachen mit freier Wortstellung wie dem Deutschen:

- (6) a. daß [_{NP₁} keiner] [_{NP₂} den Fritz] gesehen hat
 b. daß [_{NP₂} den Fritz] [_{NP₁} keiner] t₂ gesehen hat

Die GB- bzw. Prinzipien-und-Parameter-Theorie hat im wesentlichen nur eine einzige Transformation angenommen. Diese Transformation, *Affiziere* α (“Affect α ”) appliziert durchweg optional (vgl. etwa Chomsky (1981), Lasnik & Saito (1992)).¹ Die Anwendung von *Affiziere* α ist reguliert nur durch die Forderung, daß das resultierende syntaktische Objekt keine Beschränkungen der Grammatik verletzt; genau diese Forderung ist es natürlich, die dann doch oft den Effekt von obligatorischer Regelanwendung hervorruft. In einem derartigen System ist syntaktische Optionalität im Prinzip unproblematisch: Optionalität entsteht immer dann, wenn alle Beschränkungen der Grammatik sowohl dadurch erfüllt werden können, daß *Affiziere* α auf ein gegebenes syntaktisches Element α angewendet wird, als auch dadurch, daß α von *Affiziere* α nicht betroffen ist.²

Das minimalistische Programm (vgl. Chomsky (1995)) wählt eine entgegengesetzte Strategie: Hier sind alle syntaktischen Operationen obligatorisch. Bewegung

¹*Affiziere* α ist eine Verallgemeinerung von *Bewege* α , die neben Bewegung auch noch andere Transformationstypen wie Tilgung, Einsetzung etc. abdeckt.

²Im Einzelfall können im Rahmen der Prinzipien-und-Parameter-Theorie gemachte Vorschläge aber auch auf andere Ansätze zur Erfassung von Optionalität ausweichen. Tatsächlich behandeln Lasnik & Saito (1992, 4 & 187) die optionale W-Bewegung im französischen Hauptsatz nicht wie eben beschrieben, sondern auf eine Weise, die Ansätzen in Minimalismus und Optimalitätstheorie verwandt ist (s.u.): Im einen Fall fehlt das entscheidende [+w]-Merkmal, im anderen Fall ist es da; optional ist lediglich die Merkmalsausstattung eines C-Kopfes, nicht die Bewegung an sich.

erfolgt dann und nur dann, wenn sie durch Merkmale mit geeigneten Eigenschaften ausgelöst wird. Daher sieht es zunächst einmal so aus, als sei Optionalität für diesen Ansatz ein großes Problem. Es gibt jedoch eine Lösung: All jene Fälle, für die in der klassischen Analyse eine optionale Bewegungstransformation postuliert wurde, können im Minimalismus als obligatorische Bewegungen reanalysiert werden, die durch optionale Merkmale ausgelöst sind (vgl. die letzte Fußnote). Komplizierter wird die Situation, wenn wir Minimalismusvarianten mit transderivationalen Beschränkungen betrachten, wie sie in Kapitel 4 bereits diskutiert worden sind. Nehmen wir einmal eine minimalistische Syntax an, die das Prinzip der wenigsten Schritte beinhaltet. In einer solchen Syntax ist ein Kandidat grammatisch, wenn er keine lokale grammatische Beschränkung verletzt und darüber hinaus auch noch die wenigsten Instanzen von syntaktischen Operationen in seiner Kandidatenmenge aufweist. Ein Kandidat wie (6-b) mit Scrambling hat nun unter der Annahme, daß Scrambling eine syntaktische Bewegung ist, mehr Schritte als ein konkurrierender Kandidat wie (6-a), der ohne Scrambling auskommt. Dieses Problem ist ganz allgemein; eine ähnliche Situation liegt auch in vielen anderen hier angeführten Beispielen vor.

Was kann man tun, um angesichts syntaktischer Optionalität transderivationalle Beschränkungen aufrechtzuerhalten? Der naheliegende Ausweg ist der, sicherzustellen, daß zwei Derivationen, die Sätze generieren, die den Eindruck der Optionalität erwecken, nicht in derselben Kandidatenmenge sind. Das heißt, es müssen hinreichend viele Bedingungen für die Mitgliedschaft in ein und derselben Kandidatenmenge gestellt werden, so daß die Kandidatenmengen kleiner werden und syntaktische Optionalität sich als nur scheinbar erweist, weil jede "optionale" Derivation letztlich in ihrer eigenen Kandidatenmenge der Gewinner ist. Nimmt man etwa an, daß Kandidatenmengen nicht über Identität der Bedeutung (bzw. der LF), sondern über Identität der Numeration definiert werden, kann man argumentieren, daß jeder der Sätze in (1), (2) und (3) in seiner Kandidatenmenge die ökonomischste Derivation ist (vorausgesetzt, im Laufe der Derivation werden weder in der Numeration vorhandene lexikalische Elemente getilgt, noch Numerations-fremde lexikalische Elemente eingesetzt). Nimmt man weiter an, daß unterschiedliche Merkmalspezifikationen auf lexikalischen Elementen zu unterschiedlichen lexikalischen Elementen selbst führen, können die Fälle in (4), (5) und (6) auf dieselbe Weise erfaßt werden: Derivationen, die sich nur hinsichtlich von optionalen Merkmalen unterscheiden (die Bewegung auslösen), sind nicht mehr in derselben Kandidatenmenge. Wie aber Sternefeld (1997) feststellt, führt ein solcher Schritt direkt in ein Dilemma hinein: Wenn man in minimalistischen Syntaxen mit transderivationalen Beschränkungen Kandidatenmengen signifikant verkleinert, löst das zwar das Problem der syntaktischen Optionalität. Als unerwünschter Nebeneffekt ergibt sich jedoch, daß der Begriff der transderivationalen Beschränkung selbst leer wird: Derivationen, die bislang durch transderivationalle Beschränkungen wie etwa das Prinzip der wenigsten Schritte ausgeschlossen werden konnten, können nunmehr überleben, weil die ökonomischere Derivation durch die Kandidatenmengenverklei-

nerung jetzt nicht mehr Teil derselben Kandidatenmenge ist. Hier einen Mittelweg zu finden (also Kandidatenmengen, die hinreichend klein sind, um Optionalität zu erfassen, und hinreichend groß, um transderivationelle Beschränkungen noch Arbeit verrichten zu lassen), mag nicht unmöglich sein; aber es ist alles andere als einfach (vgl. Müller & Sternefeld (1996)).

Das Optionalitätsproblem entsteht nicht nur im transderivationellen Minimalismus. Wie bereits in Kapitel 4 festgehalten wurde, ist es der Blockade-Syntax inhärent (der dort üblicherweise gewählte Ausweg besteht in dem Versuch des Nachweises, daß scheinbar optional ko-existierende Formen bei genauerem Hinsehen nicht bedeutungsgleich sind und somit zu zwei verschiedenen Kandidatenmengen gehören). Ganz allgemein ergibt sich das Problem für jede Wettbewerbssyntax, die zwischen einem Generator und einem Harmonie-Evaluationsbereich unterscheidet. Und natürlich ist Optionalität zunächst einmal auch in der optimalitätstheoretischen Syntax unerwartet: Eigentlich sollte immer nur ein Kandidat der beste sein und so alle konkurrierenden Kandidaten als suboptimal blockieren. Aufgrund ihrer komplexeren internen Struktur erweist sich jedoch die Optimalitätstheorie als flexibler bei der Behandlung dieses Phänomens. Die folgende Aufstellung von Vorschlägen zur Behandlung von Optionalität in der optimalitätstheoretischen Syntax ist eine beim Kopplungsbegriff verfeinerte und um das Konzept der Neutralisierung ergänzte Variante von (19) aus Kapitel 1.

(7) *Optionalität zweier Kandidaten K_i , K_j :*

a. *Pseudo-Optionalität:*

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. *Echte Optionalität:*

K_i , K_j haben ein identisches Beschränkungsprofil.

c. *Kopplung ('tie'):*

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

(i) Global hierarchische Kopplung

(ii) Lokal hierarchische Kopplung

(iii) Lokal konjunktive Kopplung

(iv) Lokal disjunktive Kopplung

(v) Global disjunktive Kopplung

d. *Neutralisierung:*

K_i , K_j gehören zu verschiedenen Kandidatenmengen, interagieren aber.

Die vier Konzepte werden in den folgenden vier Abschnitten der Reihe nach abgehandelt.³

³Ein terminologischer Hinweis: Man darf nicht erwarten, Begriffe wie "Pseudo-Optionalität" oder die diversen Kopplungsbegriffe (z.B. "global hierarchische Kopplung") so in der Literatur zu

2. Pseudo-Optionalität

Diese Methode, Optionalität zu erfassen, ist in der optimalitätstheoretischen Syntax u.a. von Grimshaw & Samek-Lodovici (1995; 1998), Choi (1996; 1999), Samek-Lodovici (1996), Buring (1997; 1999), Costa (1998; 1999), Legendre (1998) und Heck (1998; 1999; 2000) verwendet worden. Die Grundstrategie ist ähnlich derjenigen, die in minimalistischen Syntaxen mit transderivationellen Beschränkungen und in Blockadesyntaxen verfolgt wird: Die beobachtete Optionalität ist nur scheinbar vorhanden. Tatsächlich sind die verschiedenen Kandidaten jeweils einzig optimale Gewinner in verschiedenen Wettbewerben. Folgerichtig muß in Pseudo-Optionalitätsansätzen die Kandidatenmenge entsprechend verkleinert werden. Um dies zu sehen, nehme man beispielsweise Grimshaws (1997) relativ weit gefaßten Kandidatenmengenbegriff (1) aus Kapitel 3 an, der hier wiederholt ist.

(8) *Kandidatenmenge* (Grimshaw (1997, 375-376)):

- Zwei Kandidaten K_i , K_j sind in derselben Kandidatenmenge gdw. gilt:
- a. K_i und K_j sind Realisierungen von identischen Prädikat-/Argument-Strukturen.
 - b. K_i und K_j haben identische LFs.

Nach dieser Definition sieht es so aus, als seien in jedem der bisher angeführten Beispiele in (1) bis (6) die verschiedenen Varianten miteinander im Wettbewerb. Um diese Konsequenz zu vermeiden, kann zunächst einmal ein engeres Konzept der Kandidatenmenge gewählt werden. Ersetzen wir "identische Prädikat-/Argument-Strukturen" durch "dieselbe Numeration" wie bei Chomsky (1995) (vgl. (9)), ist ein erster Schritt getan (wie oben sei für die Zwecke des Arguments vorausgesetzt, daß der Generator nicht Numerations-interne Elemente tilgen oder Numerations-fremde Elemente einfügen kann).

(9) *Kandidatenmenge*:

- Zwei Kandidaten K_i , K_j sind in derselben Kandidatenmenge gdw. gilt:
- a. K_i und K_j haben dieselbe Numeration (= identisches lexikalisches Material).
 - b. K_i und K_j haben identische LFs.

Diese Annahme erzeugt dann zwei Kandidatenmengen in den Beispielpaaren in (1) (Vorhandensein vs. Fehlen des Komplementierers *that*), (2) (Vorhandensein vs. Fehlen der Kasus-zuweisenden Präposition *to*) und (3) (Vorhandensein vs. Fehlen des W-Skopusmarkers *was*). Daher können in jedem der drei Fälle beide Kandidaten optimal sein. Dies reicht jedoch für die Beispielpaare in (4), (5) und (6) noch nicht aus, denn hier sind die lexikalischen Elemente ja identisch. Im Prinzip könnte man hier wie im Minimalismus auf bewegungsauslösende Merkmale

finden. Ich führe diese Begriffe hier zur leichteren Systematisierung ein.

rekurrieren, die zu unterschiedlichen lexikalischen Elementen und somit zu unterschiedlichen Kandidatenmengen führen; aber in der Praxis haben Vertreter von Pseudo-Optionalitätsanalysen für Fälle wie die vorliegenden oft zu argumentieren versucht, daß subtile Bedeutungsunterschiede vorliegen, die gemäß (9-b) unterschiedliche Kandidatenmengen erzeugen.

Derartige subtile Bedeutungsunterschiede sind oft für die Wahrheitsbedingungen irrelevant; sie beziehen sich typischerweise auf pragmatische oder funktionale Begriffe wie Topik, Fokus usw. Für Scrambling im Deutschen und anderen Sprachen beispielsweise ist oft vorgeschlagen worden, daß die Variante mit Scrambling und die ohne nicht vollständig bedeutungsgleich sind; für den vorliegenden Fall (6) scheint es naheliegend, unterschiedliche Topik-/Fokus-Gliederungen anzunehmen. Ebenso findet sich in der Literatur hin und wieder die Beobachtung, daß Extraposition eines Relativsatzes aus einer NP in geringem Maße semantische Effekte hat; vgl. (5). Ebenso (allerdings bei Annahme von (9-a) im gegenwärtigen Zusammenhang redundanterweise) ist häufig beobachtet worden, daß Dativ-Anhebung im Englischen wie in (2) eine gewisse "Affiziertheit" des indirekten Objekts bewirkt; vgl. Speas (1990) und die dort angegebene Literatur.⁴ Im Fall der optionalen W-Bewegung im Hauptsatz im Französischen (vgl. (4)) scheint allerdings die Identifikation eines semantisch relevanten Unterschiedes schwieriger. Es ist z.B. nicht so (wie man auf der Basis des Deutschen vielleicht zunächst einmal vermuten könnte), daß nur (4-a) ein echter Fragesatz ist und (4-b) eine Echofrage; auch (4-b) kann als reguläre Frage verstanden werden. Nichtsdestoweniger ist es wohl nicht aussichtslos, für (9-b) einschlägige funktionale Unterschiede zwischen den beiden Kandidaten zu finden.

Pseudo-Optionalitätsanalysen haben, das kann man aus dem soeben Gesagten schließen, kein Problem damit, Optionalität zu erfassen: Irgendein Unterschied zwischen zwei Kandidaten wird sich immer finden lassen, und die einzige Aufgabe besteht dann darin, diesen Unterschied als für die Etablierung zweier Kandidatenmengen relevant zu erweisen. Das Problem für Pseudo-Optionalitätsanalysen liegt in einem anderen Bereich und ist dem vergleichbar, was oben zum transderivationalen Minimalismus gesagt wurde: Dadurch, daß weniger Formen miteinander im Wettbewerb stehen, wird dem Gesamtsystem einiges von seiner Kraft genommen.

⁴Derselbe Effekt ist auch im Deutschen zu beobachten, am deutlichsten wohl bei der Akkusativ-Anhebung (Applikativbildung) bei *be*-Verben:

- (i) a. Sie lud Heu auf den Wagen
 b. Sie belud den Wagen mit Heu

Vgl. hierzu Wunderlich (1987; 1997) und Stechow (1992). Wunderlich konstatiert einen Bedeutungsunterschied (der Wagen ist in (i-b) notwendigerweise affiziert, in (i-a) dagegen nicht), der auf zwei verschiedene Lexikoneinträge zurückgeht und bei Verfolgen einer Pseudo-Optionalitätsanalyse ausreicht, um zwei Kandidatenmengen zu erzeugen. Stechow nimmt dagegen an, daß (i-ab) bedeutungsgleich sind (die Sätze erhalten dieselbe D-Struktur), womit sich im gegenwärtigen Kontext das Optionalitätsproblem stellt.

Am klarsten wird dies bei dem Phänomen, das ich – in loser Anlehnung an den Gebrauch in der Literatur – *syntaktische Alternation* nennen möchte und das man sich als so etwas wie die “Schattenseite” der Optionalität vorstellen kann:

(10) *Alternation:*

Zwei (oder mehr) Konstruktionen können miteinander koexistieren (= Optionalität), aber unter gewissen Bedingungen kann die eine die andere blockieren.

Alternation impliziert also den Zusammenbruch von Optionalität in bestimmten Kontexten. Es zeigt sich, daß sämtliche in (1) bis (6) aufgelisteten Fälle von Optionalität im Sinne von (10) Alternationen sind. Bei der englischen Komplementierung (vgl. (1)) haben wir das bereits gesehen: Bei eingebetteter Topikalisierung oder in Subjektsätzen bricht die Optionalität zusammen, und nur die Variante mit lexikalisch realisiertem Komplementierer ist möglich:

- (11) a. *I think – [PP to John]₁ Mary gave a book t₁
 b. I think that [PP to John]₁ Mary gave a book t₁
 c. *[He left so early] shows that he was tired
 d. [That he left so early] shows that he was tired

Ebenso ist Dativ-Anhebung (vgl. (2)) nicht immer optional. In manchen Kontexten ist die Operation obligatorisch (z.B. in (12-ab)), in anderen Kontexten ist sie blockiert (z.B. in (12-cd) und in (12-ef)); vgl. Perlmutter & Soames (1979), Baker (1988), Larson (1988) sowie die dort angegebene Literatur.

- (12) a. *The orange socks cost [NP two dollars] [PP to/for Linda]
 b. The orange socks cost [NP Linda] [NP two dollars]
 c. I donated [NP money] [PP to charity]
 d. *I donated [NP charity] [NP money]
 e. The boss gave [NP it] [PP to John]
 f. *The boss gave [NP John] [NP it]

Ebenso ist Skopusmarkierung plus partielle W-Bewegung im Deutschen nicht immer wie in (3) optional. In Kapitel 2 wurde bereits erwähnt, daß die Optionalität zusammenbricht in Subjekt-Insel-Kontexten (vgl. McDaniel (1989)); das relevante Beispielpaar ist hier (in der Reihenfolge umgetauscht) wiederholt.

- (13) a. ?[PP Mit wem]₁ ist es schade [CP t'₁ daß Hans t₁ gesprochen hat] ?
 b. *Was₁ ist es schade [CP [PP mit wem]₁ Hans t₁ gesprochen hat] ?

Ebenso geht, wie Höhle (1990) und Rizzi (1992) beobachtet haben, die Optionalität verloren bei Präsenz einer Negation im Matrixsatz:

- (14) a. ?Wen₁ glaubst du nicht [CP t'₁ daß man t₁ einladen sollte] ?
 b. *Was₁ glaubst du nicht [CP wen₁ (daß) man t₁ einladen sollte] ?

Dasselbe Bild ergibt sich bei der W-Bewegung im Französischen (vgl. (4)). Optional ist W-Bewegung nur in Hauptsätzen, nicht im eingebetteten Fall (vgl. (15-ab)), und auch nur bei solchen W-Phrasen, die Argumente sind, nicht bei Adjunkten (vgl. (15-cd)). In allen anderen Fällen ist W-Bewegung wie im Englischen obligatorisch.

- (15) a. Je me demande [_{CP} qui₁ C tu as vu t₁]
 ich frage mich wen du hast gesehen
 b. *Je me demande [_{CP} – (que) tu as vu qui₁]
 ich frage mich daß du hast gesehen wen
 c. Pourquoi₁ es-tu venu t₁ ?
 warum bist-du gekommen
 d. *– Tu es venu pourquoi₁ ?
 du bist bekommen warum

Auf dieselbe Art und Weise sind auch Extraposition (vgl. (5)) und Scrambling im Deutschen (vgl. (6)) als Alternationen erweisbar: Unter bestimmten Umständen geht die Optionalität verloren.

An und für sich kann nun syntaktische Alternation in der optimalitätstheoretischen Syntax leicht erfaßt werden: Es gibt einen Wettbewerb, aus dem in bestimmten Kontexten nur einer der Kandidaten als optimal hervorgeht. Dies setzt natürlich voraus, daß die zwei Kandidaten in den Beispielpaaren in (11)–(15) jeweils einer Kandidatenmenge angehören – aber genau diese Annahme ist ja im Pseudo-Optionalitätsansatz aufgegeben. Damit stecken wir in einem Dilemma: Es ist nicht klar, durch welchen Kandidaten ein ungrammatischer Satz wie, z.B., (11-a) (mit Komplementierertilgung bei eingebetteter Topikalisierung) blockiert wird, wenn (11-b) (mit Kookkurrenz von Komplementierer und eingebetteter Topikalisierung) nicht Teil derselben Kandidatenmenge ist; denn wir wissen ja unabhängig, daß Komplementierertilgung an sich im Englischen möglich ist (vgl. (1)).

Es läßt sich somit festhalten, daß unter einer Pseudo-Optionalitätsanalyse von syntaktischer Optionalität das Problem entsteht, zu erklären, wie Optionalität unter gewissen Umständen zu Obligatorizität werden kann. Das Phänomen der Alternation zeigt, daß Kandidatenmengen so definiert sein sollten, daß die Sätze in den Beispielpaaren in (1) bis (6) jeweils miteinander im Wettbewerb stehen. Dann muß die Optionalität anders erfaßt werden. Die einfachste Möglichkeit wäre, daß Optionalität einfach “zufällig” und ohne Zusatzannahmen entsteht. Dieser Möglichkeit ist der nächste Abschnitt gewidmet. (Tatsächlich wird sich zeigen, daß auch eine Variante der Pseudo-Optionalitätsanalyse, der Neutralisierungsansatz, zum Ziel führt. Aber dazu später.)

3. Echte Optionalität

Die Definition von Optimalität in (5) aus Kapitel 1 gewährleistet, daß es im Prinzip mehr als einen optimalen Kandidaten in einer Kandidatenmenge geben kann, und zwar dann, wenn zwei (oder mehr) Kandidaten ein identisches Beschränkungsprofil

aufweisen, das besser ist als das aller Mitbewerber. Alternation folgt dann schlicht aus der Tatsache, daß aus Optionalität in bestimmten Kontexten Obligatorizität wird: Sobald im neuen Kontext eine (beliebig geordnete) Beschränkung relevant wird, die die beiden Kandidaten unterscheidet (und die “normalerweise” irrelevant ist), blockiert der eine Kandidat den anderen Kandidaten als suboptimal. Derartige Ansätze sind von Grimshaw (1997) und Vikner (1997a; 1999) verfolgt worden.⁵

Die Komplementierertilgungstheorie von Grimshaw (1997) ist bereits in Kapitel 3 im Detail besprochen worden. Einschlägige Daten sind hier nochmals mit der von Grimshaw angenommenen Struktur wiederholt (vgl. (27) und (29) aus Kapitel 3).

- (16) a. I think [_{IP} John will leave]
 b. I think [_{CP} that [_{IP} John will leave]]
 c. *I think [_{IP} most of the time [_{IP} John could accept this solution]]
 d. I think [_{CP} that [_{IP} most of the time [_{IP} John could accept this solution]]]

In von Brückenverben selegierten Objektsätzen ist IP-Einbettung ebenso optimal wie CP-Einbettung, weil die einschlägigen Beschränkungen in (17) gleich gut (und besser als bei allen anderen Mitbewerbern) erfüllt werden; vgl. Tabelle T_{5.1} (dies ist eine vereinfachte Version von Tabelle T_{3.15} aus Kapitel 3).

- (17) a. PUR-EP:
 Adjunktion an die höchste XP einer eingebetteten erweiterten Projektion und Bewegung in deren Kopf sind verboten.
 b. OB-KOPF:
 Eine Projektion hat einen (nicht-leeren) Kopf.
 c. ÖKON:
 Bewegung ist verboten (Spur ist nicht erlaubt).

T_{5.1}: Komplementierertilgung im Englischen

Kandidaten	PUR-EP	OB-KOPF	Ö-KON
☞K ₁ : think [_{CP} that [_{IP} John ₁ will [_{VP} t ₁ leave]]]			*
☞K ₂ : think [_{IP} John ₁ will [_{VP} t ₁ leave]]			*
K ₃ : think [_{CP} e [_{IP} John ₁ will [_{VP} t ₁ leave]]]		*!	*

Bei einem eingebetteten IP-Adjunkt wird jedoch der Kandidat mit IP-Einbettung (also “Komplementierertilgung”) durch den Kandidat mit CP-Einbettung blockiert, weil ersterer fatal PUR-EP verletzt; vgl. Tabelle T_{5.2} (die auf T_{3.16} aus Kapitel 3 basiert).

Soweit Grimshaws (1997) Analyse der Komplementierertilgungsalternation mit

⁵Aber vgl. auch die bei der Darstellung von Pesetskys (1997; 1998) Theorie der PF-Realisierung in Kapitel 3 zu Tabelle T_{3.26} gemachten Bemerkungen.

T_{5.2}: Komplementierertilgung und Satzadjunkte

Kandidaten	PUR-EP	OB-KOPF	Ö-KON
☞ K ₁ : V [_{CP} that [_{IP} Adjunkt [_{IP} NP ₁ I t ₁ V NP]]]			*
K ₂ : V [_{IP} Adjunkt [_{IP} NP ₁ I t ₁ V NP]]	*!		*
K ₃ : V [_{CP} e [_{IP} Adjunkt [_{IP} NP ₁ I V t ₁ NP]]]		*!	*

Hilfe von echter Optionalität. Diese Analyse ist von Vikner (1999) so erweitert worden, daß eine interessante sprachvergleichende Korrelation erfaßt wird: Es sieht so aus, als könnten nur solche Sprachen Komplementierertilgung aufweisen, denen S-strukturelle V-nach-I-Bewegung fehlt (z.B. Englisch, Dänisch), nicht dagegen solche Sprachen, die S-strukturelle V-nach-I-Bewegung aufweisen (wie z.B. Französisch oder Isländisch). Wie Vikner (1999) den V-nach-I-Bewegungsparameter optimalitätstheoretisch erfaßt, haben wir bereits in Abschnitt 7.4.2 von Kapitel 2 gesehen, nämlich durch die relative Ordnung der Beschränkung LEX-ÖKON von Grimshaw (1997) (die Bewegung von Vollverben verbietet) und der Beschränkung STR-BD (die allgemein ungebundene Spuren, speziell die Absenkung der Flexionsmorphologie von I nach V verbietet).

- (18) a. LEX-ÖKON:
Bewegung von lexikalischen Köpfen ist verboten.
b. STR-BD:
Spuren müssen auf der S-Struktur gebunden sein.

Für Grimshaws (1997) Analyse von Komplementierertilgung im Englischen hat dies keinerlei Konsequenzen: IP-Einbettung und CP-Einbettung können in Objektsätzen ohne IP-Adjunktion nach wie vor beide aufgrund eines identischen Beschränkungsprofils optimal sein. Dies zeigt Tabelle T_{5.3}.⁶

T_{5.3}: Optionaler Komplementierer ohne V-nach-I im Englischen

Kandidaten	PUR-EP	OB-KOPF	LEX-ÖKON	STR-BD
K ₁ : ... think [_{CP} that [_{IP} John V ₁ ... t ₁ ...			*!	
K ₂ : ... think [_{CP} - [_{IP} John V ₁ ... t ₁ ...		*!	*	
☞ K ₃ : ... think [_{CP} that [_{IP} John ... V ₁ ...				*
K ₄ : ... think [_{CP} - [_{IP} John ... V ₁ ...		*!		*
K ₅ : ... think [_{IP} John V ₁ ... t ₁ ...	*!		*	
☞ K ₆ : ... think [_{IP} John ... V ₁ ...				*

Hier sind gegenüber T_{5.1} der Übersichtlichkeit halber die für den vorliegenden

⁶Man beachte, daß Vikner (1999) anders als Grimshaw (1997) annimmt, daß alle finiten Sätze mindestens IP-Status haben müssen.

Wettbewerb nicht entscheidenden ÖKON-Verletzungen entfernt worden. Darüber hinaus ist für jeden der drei relevanten Kandidaten in $T_{5.1}$ (CP-Einbettung mit Komplementierer, IP-Einbettung, CP-Einbettung mit leerem C) noch sein Gegenstück mit V-nach-I-Bewegung aufgelistet, das wie bei Grimshaw (1997) LEX-ÖKON verletzt. Die V-in-situ-Kandidaten verletzen demgegenüber bei Vikner (anders als bei Grimshaw) STR-BD. Wie K_1 vs. K_3 zeigt, ist eine Verletzung von LEX-ÖKON dabei aufgrund der Ordnung LEX-ÖKON \gg STR-BD im Englischen im Zweifelsfall (d.h., wenn nicht ohnehin schon höherrangige Beschränkungen involviert sind) fatal. Man sieht aber bereits bei K_5 , daß bei IP-Einbettung eine Verletzung von LEX-ÖKON durch V-nach-I-Bewegung *nicht* fatal ist, weil hier gleichzeitig noch die hoch geordnete Beschränkung PUR-EP verletzt ist. Allgemein gilt, daß bei V-nach-I-Bewegung PUR-EP nur dann respektiert werden kann, wenn IP nicht die höchste erweiterte Projektion des eingebetteten Satzes ist. Dies macht aber eine klare Prognose. Sprachen, die sich vom Englischen hinsichtlich der hier betrachteten Beschränkungen nur durch die Ordnung von LEX-ÖKON und STR-BD unterscheiden und so V-nach-I-Bewegung aufweisen, sollten keine IP-Einbettung – und daher auch keine Komplementierertilgung – aufweisen. Die folgenden Daten aus dem Französischen bestätigen diese Vorhersage:

- (19) a. Je crois que l'acteur voit vraiment le film
 b. *Je crois – l'acteur voit vraiment le film

Der Wettbewerb ist in Tabelle $T_{5.4}$ dargestellt; man vergleiche K_1 mit K_5 . Die Rolle, die die Kontextveränderung bei innersprachlichen Alternationsphänomenen spielt, wird hier von der Beschränkungsreihenfolge übernommen.

$T_{5.4}$: Obligatorischer Komplementierer bei V-nach-I im Französischen

Kandidaten	PUR-EP	OB-KOPF	STR-BD	LEX-ÖKON
☞ K_1 : ... crois $[_{CP}$ que $[_{IP}$ l'acteur V_1 ... t_1 ...				*
K_2 : ... crois $[_{CP}$ – $[_{IP}$ l'acteur V_1 ... t_1 ...		*!		*
K_3 : ... crois $[_{CP}$ que $[_{IP}$ l'acteur ... V_1 ...			*!	
K_4 : ... crois $[_{CP}$ – $[_{IP}$ l'acteur ... V_1 ...		*!	*	
K_5 : ... crois $[_{IP}$ l'acteur V_1 ... t_1 ...	*!			*
K_6 : ... crois $[_{IP}$ l'acteur ... V_1 ...			*!	

Soweit sieht es so aus, als sei die echte Optionalität der Pseudo-Optionalität auf empirischem Gebiet überlegen, weil sie ohne Probleme Alternation erfassen kann. Darüber hinaus ist dieser Ansatz auch konzeptuell attraktiv, weil er keinerlei Zusatzannahmen benötigt, ist einmal Optimalität so definiert, daß mehr als ein Gewinner pro Kandidatenmenge erlaubt ist (dies ist, wie gesagt, bei Grimshaw (1997) nicht der Fall). Nichtdestoweniger gibt es ein ernsthaftes Problem mit diesem Ansatz, das in Baković (1997) und Baković & Keer (1999) herausgearbeitet, in Grimshaw (1997, 411) bereits zur Kenntnis genommen und in Abschnitt 1.7.3.

von Kapitel 3 auch schon erwähnt worden ist: Es ist im Einzelfall gar nicht so einfach zu gewährleisten, daß zwei optimale Kandidaten *exakt* dieselben Beschränkungen verletzen bzw. erfüllen. Das Vorhandensein eines Komplementierers etwa muß vollständig kostenfrei sein (*that* darf z.B. VOLL-INT nicht verletzen); ebenso darf die gegenüber IP-Einbettung zusätzliche funktionale Projektion bei CP-Einbettung nicht per se Kosten verursachen (mit einer Beschränkung, die Ökonomie der Projektion verlangt, wie Grimshaws (1993) MIN-PROJ (“Minimal Projection”), ist diese Analyse daher inkompatibel – aber MIN-PROJ wird in Grimshaw (1997) auch nicht mehr angenommen). Wie niedrig auch immer eine Beschränkung geordnet ist, die zwischen den beiden wohlgeformten, optionalen Kandidaten diskriminiert, eine Verletzung dieser Beschränkung wird immer fatal sein und den Zusammenbruch der Optionalität bewirken. Dieses Problem ist nicht nur abstrakt vorhanden: In Grimshaws (1997) Analyse wird wie in Kapitel 3 gesehen aus unabhängigen Gründen ein Beschränkungspaar KOPF-LINKS/KOPF-RECHTS (“Der Kopf steht in seiner Projektion links/rechts”) vorgeschlagen. KOPF-RECHTS ist im Englischen tief geordnet (Köpfe sind de facto nicht rechts-peripher), aber die bloße Präsenz reicht aus, um die Optionalität in $T_{5.1}$ unerwünschterweise zu unterminieren: Die CP-Variante hat immer eine zusätzliche links-köpfige XP und verletzt damit einmal mehr KOPF-RECHTS als die konkurrierende IP/VP-Variante. Das falsche Ergebnis, daß Komplementierertilgung bei gewöhnlichen Objektsatzeinbettungen im Englischen obligatorisch sein sollte, illustriert Tabelle $T_{5.5}$.

$T_{5.5}$: Komplementierertilgung im Englischen mit KOPF-RECHTS

Kandidaten	PUR-EP	OB-KOPF	Ö-KON	KOPF-RECHTS
K_1 : think [$_{CP}$ that [$_{IP}$ John ₁ will [$_{VP}$ t ₁ leave]]]			*	*!
$\rightarrow K_2$: think [$_{IP}$ John ₁ will [$_{VP}$ t ₁ leave]]			*	
K_3 : think [$_{CP}$ e [$_{IP}$ John ₁ will [$_{VP}$ t ₁ leave]]]		*!	*	

Tatsächlich geht Grimshaw (1999) aus Gründen wie diesem von dieser Analyse von Optionalität bei Komplementierertilgung wieder ab.

4. Kopplungen

Eine dritte Möglichkeit, optimalitätstheoretisch syntaktische Optionalität zu erfassen, beruht auf dem Konzept der Kopplung von Beschränkungen (‘constraint tie’). Die Grundidee ist, daß zwei (oder mehr) Beschränkungen gleich gewichtet sind, also gekoppelt: $A \circ B$. Wenn sich zwei (oder mehr) Kandidaten nur bei der Kopplung unterscheiden, dann können sich die Vor- und Nachteile der Kandidaten aufheben, und beide (bzw. alle) sind optimal, auch wenn sie nicht ein vollständig identisches Beschränkungsprofil aufweisen. Alternation wird erfaßt durch entweder die Intervention einer anderen Beschränkung, bei der sich die Kandidaten dann doch

unterscheiden, oder, im Falle von Parametrisierung, durch Auflösung der Kopplung in die eine oder andere Dominanzrichtung.

Es lassen sich mindestens fünf verschiedene Kopplungskonzepte unterscheiden, die in der Literatur vorgeschlagen worden sind. Allen Kopplungsbegriffen gemeinsam ist, daß in einem abstrakten Wettbewerb wie in $T_{5.6}$ bei Vorliegen einer Hierarchie $A \gg B \circ C \gg D$ K_1 und K_2 als optimal prognostiziert werden, und K_3 und K_4 als suboptimal.

$T_{5.6}$: *Kopplung von Beschränkungen*

Kandidaten	A	B	C	D
$\Leftarrow K_1$			*	
$\Leftarrow K_2$		*		
K_3		*(!)	*(!)	
K_4	*!			

Ansonsten aber haben die diversen Kopplungsbegriffe unterschiedliche Konsequenzen. Eine Unterscheidung kann man zunächst einmal vornehmen zwischen “globalen” und “lokalen” Kopplungen. Globale Kopplungen lassen sich als Abkürzungen für die simultane Präsenz verschiedener Beschränkungsordnungen in einer Sprache auffassen; lokale Kopplungen sind als spezielle Beschränkungstypen zu betrachten. Der erste zu diskutierende Kopplungstyp ist die global hierarchische Kopplung.

4.1. Global Hierarchische Kopplung

Global hierarchische Kopplungen sind verwendet worden von Sells, Rickford & Wasow (1996), Ackema & Neeleman (1998), Schmid (1999) und anderen. Sie beruhen auf einem Kopplungskonzept, das von Prince & Smolensky (1993) vorgeschlagen worden ist. Die Grundidee ist, daß eine Beschränkungsordnung mit einer Kopplung $B_1 \circ B_2$ unterspezifiziert ist; es handelt sich im Grunde um eine abkürzende Schreibweise für zwei verschiedene Hierarchien, in denen einmal $B_1 \gg B_2$ gilt, und einmal $B_2 \gg B_1$. Als Konsequenz daraus können K_1 und K_2 beide optimal sein, wenn sie unter der Auflösung einer Kopplung optimal sind, auch wenn z.B. K_1 sich bei tieferrangigen Beschränkungen anders als K_2 verhält.

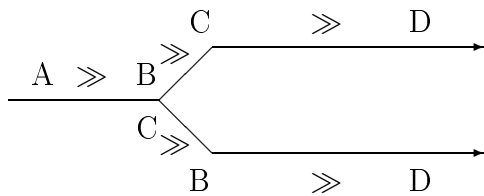
Wenn in der Definition von Optimalität in (5) aus Kapitel 1 von einer “Beschränkungsordnung“ die Rede ist, hinsichtlich derer kein Kandidat sich besser verhalten darf als ein optimaler Kandidat, so können wir annehmen, daß dies eine *vollständige* Ordnung der Beschränkungen voraussetzt. Dann läßt sich eine Ordnung mit global hierarchischen Kopplungen als unterspezifizierte, *partielle* Ordnung der Beschränkungen verstehen:⁷

⁷Eine vollständige Beschränkungsordnung \gg ist im technischen Sinne *strikt* (transitiv, irreflexiv und asymmetrisch) und *konnex* (d.h., für alle Paare von Beschränkungen B_i, B_j gilt $B_i \gg$

(20) *Global hierarchische Kopplung:*

Es sei $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine partielle Beschränkungsordnung der Sprache L und B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$. Dann gilt: Für jede mögliche Subordnung O der Beschränkungen in B_i ist auch Γ_O eine Beschränkungsordnung der Sprache L , wobei Γ_O sich von Γ nur dadurch unterscheidet, daß B_i durch O ersetzt ist.

Da Optimalität nur hinsichtlich vollständiger Beschränkungsordnungen definiert ist und nicht hinsichtlich partieller Beschränkungsordnungen, impliziert (20), daß ein Kandidat in einer Sprache grammatisch ist genau dann, wenn er optimal ist unter wenigstens einer vollständigen Ordnung, die durch Auflösung aller Kopplungen in der Sprache generierbar ist. Das Diagramm in (21) zeigt, wie global hierarchische Kopplungen funktionieren:

(21) *Diagramm der global hierarchischen Kopplung $B \circ C$* 

Angenommen also, es kommen zwei Kopplungen von jeweils zwei Beschränkungen in einer partiellen Ursprungsordnung vor. Dann wird zuerst die eine Kopplung aufgelöst, und zwei immer noch partielle Beschränkungsordnungen Γ_{O_1} , Γ_{O_2} resultieren. In beiden wird dann noch die andere Kopplung aufgelöst, so daß insgesamt vier vollständige Beschränkungsordnungen entstehen. Ein Kandidat ist dann grammatisch, wenn er gemäß einer dieser vier vollständigen Ordnungen optimal ist.

Als Beispiel sei die Analyse von optionaler W -Bewegung im Hauptsatz des Französischen betrachtet, die Ackema & Neeleman (1998) als Teil ihrer umfassenden Studie zum W -Bewegungsparameter entwickeln. Das Gesamtsystem fußt wesentlich auf den drei (S -strukturellen) Beschränkungen in (22). Es soll vor einer Darstellung der Kopplungsanalyse im folgenden kurz eingeführt werden.

(22) a. Q-MARK ("Q-Merkmal-Markierung"):

In einem Fragesatz muß ein $[+Q]$ -Merkmal der Konstituente zugewiesen werden, die der Proposition entspricht.

b. Q-SKOP ("Q-Skopus"):

$[+Q]$ -Elemente müssen die Konstituente, die der Proposition entspricht, c -kommandieren.

c. BKP ("Bedingung der kürzesten Pfade"):

Jeder bei Bewegung echt überschrittene Knoten erzeugt einen Stern.

B_j oder $B_j \gg B_i$). Eine partielle Beschränkungsordnung ist nicht konnex. Vgl. z.B. Partee, ter Meulen & Wall (1993, 51).

Was Q-MARK betrifft, so nehmen Ackema und Neeleman an, daß die geforderte Q-Markierung nur durch einen lexikalisch gefüllten Kopf unter Schwesternschaft erfolgen kann; die Fähigkeit, eine Konstituente für ein Q-Merkmal zu markieren, erwirbt ein Kopf per Spezifikator/Kopf-Kongruenz von der W-Phrase in seinem Spezifikator, die dieses Q-Merkmal trägt. Nehmen wir vereinfachend an, daß die Kategorie IP immer der Proposition entspricht, dann erzwingt diese Beschränkung somit simultan die Bewegung eines W-Elements nach SpecC und Kopf-Bewegung nach C (bei Grimshaw (1997) leistet dies wie gesehen die Interaktion von OP-SPEZ und OB-KOPF). Die zweite Beschränkung Q-SKOP ist komplementär zur ersten: Sie verlangt, daß W-Elemente die IP c-kommandieren, egal ob aus einer Adjunkt- oder Spezifikatorposition (mindestens aber aus einer IP-Adjunktposition). Die dritte Beschränkung BKP schließlich vermerkt jede bei einem Bewegungsschritt erfolgende Überquerung eines Knotens als Verletzung. Eine Bewegung über z.B. VP, I', IP und C' bewirkt somit vier BKP-Verletzungen.⁸

Je nach Ordnung dieser Beschränkungen resultieren verschiedene Festlegungen des W-Bewegungsparameters in verschiedenen Sprachen.

(23) *Beschränkungsordnungen:*

- | | | |
|----|-------------------------------|---------------|
| a. | Q-MARK \gg BKP \gg Q-SKOP | (Englisch) |
| b. | Q-SKOP \gg BKP \gg Q-MARK | (Tschechisch) |
| c. | Q-MARK, Q-SKOP \gg BKP | (Bulgarisch) |
| d. | BKP \gg Q-MARK, Q-SKOP | (Chinesisch) |

Unter der englischen Ordnung gewinnt in einer mehrfachen Hauptsatzfrage mit zwei W-Elementen der Kandidat, der ein W-Element nach SpecC schiebt und das Auxiliar nach C bewegt (aufgrund der Ordnung Q-MARK \gg BKP – das Auxiliar deriviert das W-Merkmal von dem W-Element in seinem Spezifikator und Q-markiert so die IP), und das andere W-Element in situ läßt (aufgrund der Ordnung BKP \gg Q-SKOP – anders als Q-SKOP verlangt ja Q-MARK nicht per se, daß W-Phrasen bewegt werden; wichtig ist nur, daß IP Q-markiert wird, und dafür reicht die Bewegung einer einzigen W-Phrase schon aus); vgl. (24-a).

- (24) a. Which books₁ will₂ they t₂ give t₁ to whom₃ ?
 b. *Which books₁ will₂ to whom₃ they t₂ give t₁ t₃ ?

⁸BKP heißt bei Ackema und Neeleman anders, nämlich "Stay" (ÖKON); die Beschränkung ist auch genau so formuliert wie Grimshaws (1997) ÖKON. Allerdings liegt hier ein Fall unglücklicher terminologischer Ambiguität vor, denn Ackema & Neeleman (1998, 448) meinen etwas ganz anderes als Grimshaw (1997) oder Legendre et al. (1995; 1998) mit dieser Beschränkung: ÖKON soll eben nicht den Bewegungsschritt an sich verbieten ("Spur ist nicht erlaubt"), sondern Bewegungspfade in ihrer Länge minimieren. Es handelt sich also um eine lokale Variante des in Kapitel 4 bereits besprochenen transderivationalen *Prinzips der kürzesten Pfade*, wie es im minimalistischen Programm von Chomsky (1993; 1995), Collins (1994), Kitahara (1993; 1997) und Nakamura (1998) postuliert worden ist (vgl. (22) aus Kapitel 4). Daher die Umbenennung.

- c. *Which books₁ to whom₃ they will₂ give t₁ t₃ ?
- d. *Which books₁ they will₂ give t₁ to whom₃ ?
- e. *Which books₁ to whom₃ will₂ they give t₁ t₃ ?
- f. *They will₃ give which books₁ to whom₂ ?

Den Wettbewerb illustriert Tabelle T_{5.7}.⁹

T_{5.7}: Mehrfachfragen im Englischen

Kandidaten	Q-MARK	BKP	Q-SKOP
☞K ₁ : [<i>CP</i> which b. ₁ Aux ₂ [<i>IP</i> NP t ₂ V t ₁ to whom ₃]]		*****	*
K ₂ : [<i>CP</i> which b. ₁ Aux ₂ to whom ₃ [<i>IP</i> NP t ₂ V t ₁ t ₃]]		*****!***	
K ₃ : [<i>IP</i> which b. ₁ to whom ₃ [<i>IP</i> NP Aux ₂ V t ₁ t ₃]]	*!	***	
K ₄ : [<i>IP</i> which b. ₁ [<i>IP</i> NP Aux ₂ V t ₁ to whom ₃]]	*!	**	*
K ₅ : [<i>CP</i> which b. ₁ to whom ₃ Aux ₂ [<i>IP</i> NP t ₂ V t ₁ t ₃]]		*****!	
K ₆ : [<i>IP</i> they Aux ₂ ... which b. ₁ to whom ₃]]	*!		**

K₁ ist der Kandidat, der sowohl Q-MARK erfüllt, als auch die damit unumgänglichen BKP-Verletzungen minimiert, auf Kosten einer Q-SKOP-Verletzung. Die sechs eingetragenen BKP-Verletzungen resultieren aus einer Kombination von W-Bewegung (über VP, I', IP, C') und Subjekt-Auxiliar-Inversion nach C (über I', IP). K₂ hat zusätzliche, fatale BKP-Verletzungen, weil noch die zweite W-Phrase über VP und I' hinweg an IP adjungiert wird (wodurch Q-SKOP erfüllt wird).¹⁰ K₅ hat ebenfalls eine zusätzliche, fatale BKP-Verletzung. Hier ist per Annahme die untere W-Phrase zuerst an die obere adjungiert worden (wodurch nur V' überquert wird), und dann läuft alles ab wie bei K₁. Die erste, extrem kurze Bewegung ist jedoch bereits fatal. K₃ und K₄ adjungieren zwei bzw. eine W-Phrase an IP, lassen SpecC und C leer und verletzen so fatal Q-MARK. K₆ schließlich bewegt gar kein Element und verletzt so ebenfalls fatal Q-MARK.

Unter einer einer Vertauschung von Q-MARK und Q-SKOP, also unter der tschechischen Ordnung, wird K₃ (= (24-c), mit mehrfacher W-Bewegung, aber ohne

⁹Die Darstellung hier ist und im folgenden vereinfacht. Zum Beispiel ist die genaue Zahl der BKP-Verletzungen approximativ angegeben; die exakte Ermittlung im System von Ackema und Neeleman ist nur möglich unter ihren speziellen Annahmen über funktionale Projektionen, deren Darlegung hier zu weit führen würde; unter einer Klärung der Frage, ob bei Adjunktion an XP die XP als überquert gilt oder nicht; sowie unter einer Entscheidung bzgl. der Basisposition von Subjekten und anderen Argumenten. Darüber hinaus ist hier von dem Kandidaten abstrahiert, der die tiefere W-Phrase nach SpecC bewegt und die höhere in situ läßt, allgemeiner noch von allen Kandidaten, die die Ordnung der W-Phrasen vertauschen. Durch BKP kann in diesem Ansatz im Prinzip der Superioritätseffekt abgeleitet werden (was angesichts der Herkunft aus dem Prinzip der kürzesten Pfade nicht verwunderlich ist).

¹⁰Anders als OP-SPEZ bei Grimshaw (1997) (aber so wie W-KRIT bei Müller (1997)) kann Q-SKOP also mehrfache W-Bewegung auslösen.

V-nach-C-Bewegung) optimal (vgl. auch die Bemerkungen zum Russischen in Kapitel 3): Hier wird Q-SKOP unter Minimierung von BKP-Verletzungen erfüllt, auf Kosten einer Verletzung von Q-MARK. Dominieren Q-MARK und Q-SKOP (in beliebiger Ordnung) beide BKP, wie im Bulgarischen, dann gewinnt K_5 (= (24-e), mit mehrfacher W-Bewegung und V-nach-C-Bewegung): Wie bei K_2 werden beide hoch geordneten Beschränkungen erfüllt, aber weil die zweite W-Phrase in K_5 anders als in K_2 den Großteil des Weges “auf dem Rücken” der ersten W-Phrase zurücklegt, werden BKP-Verletzungen bei K_5 minimiert. Dominiert schließlich BKP sowohl Q-MARK als auch Q-SKOP (letztere wiederum in beliebiger Ordnung), so ist K_6 (= (24-f)) optimal, weil hier gar nicht bewegt wird, auf Kosten von Verletzungen von Q-MARK und Q-SKOP. Somit resultiert für jede Ordnung eine tatsächlich dokumentierte Sprache, und dem von Ackema & Neeleman (1998) akzeptierten Postulat der faktoriellen Typologie ist entsprochen. Man beachte, daß K_2 und K_4 unter keiner Ordnung optimal sind und die Sätze in (24-b) und (24-d) daher als allgemein unmöglich vorausgesagt werden. K_2 ist wie gesehen durch K_5 harmonisch begrenzt; K_4 dagegen ist ein “komplexer Verlierer” im Sinne von Samek-Lodovici & Prince (1999) (vgl. Kapitel 2). Die K_4 harmonisch begrenzende Menge von Beschränkungen ist $\{K_3, K_6\}$: In jeder Ordnung, in der Q-SKOP \gg BKP gilt, gewinnt K_3 mit mehrfacher IP-Adjunktion gegen K_4 ; und in jeder Ordnung, in der umgekehrt BKP \gg Q-SKOP gilt, gewinnt K_6 mit generellem W-in-situ gegen K_4 .

Soweit die Theorie des W-Bewegungsparameters bei Ackema & Neeleman (1998). Um die Verhältnisse in eingebetteten Sätzen korrekt zu erfassen, müssen noch weitere Annahmen gemacht werden. Dies soll uns gegenwärtig aber nicht interessieren. Vielmehr soll es darum gehen, wie Ackema und Neeleman die Optionalität von W-Bewegung von Argumenten im französischen Hauptsatz in dieses System integrieren.¹¹ Die Daten sind hier wiederholt (vgl. (4)).

- (25) a. Qui₁ as-tu vu t₁ ?
 b. – Tu as vu qui₁ ?

Für das Französische wird folgende Beschränkungsordnung vorgeschlagen.

- (26) *Beschränkungsordnung im Französischen:*
 BKP \circ Q-MARK \gg Q-SKOP

Gemäß dem Konzept der global hierarchischen Kopplung in (20) bedeutet dies nun nichts anderes, als daß im Französischen ein Kandidat grammatisch ist, wenn er entweder unter der englischen Ordnung Q-MARK \gg BKP \gg Q-SKOP optimal ist,

¹¹Es sei allerdings darauf hingewiesen, daß die nun vorzustellende Analyse des Französischen letzten Endes von Ackema & Neeleman verworfen wird zugunsten der Annahme, daß die scheinbare Alternation unterschiedliche Grammatiken involviert. Dies ist jedoch gar nicht so weit von ihrem Kopplungskonzept entfernt, denn unterschiedliche Grammatiken sind in der Optimalitätstheorie durch unterschiedliche Beschränkungsordnungen definiert, und global hierarchische Kopplungen tun genau dies: unterschiedliche Beschränkungsordnungen einzuführen.

oder aber unter der chinesischen Ordnung $BKP \gg Q\text{-MARK} \gg Q\text{-SKOP}$. Genau dies ist das erwünschte Ergebnis, wie Tabelle T_{5.8} zeigt.¹²

T_{5.8}: Optionale W-Bewegung im Französischen

Kandidaten	BKP	Q-MARK	Q-SKOP
K_1 : qui ₁ Aux ₂ NP t ₂ V t ₁	*(!)*****		
K_2 : -- NP Aux ₂ V qui ₁		*(!)	*
K_3 : qui ₁ - NP Aux ₂ V t ₁	*(!)*	*(!)	
K_4 : - Aux ₂ NP t ₂ V qui ₁	*(!)*	*(!)	*

Ein wesentlicher Unterschied zur echten Optionalität ist, daß ein unterschiedliches Verhalten zweier Kandidaten bei tiefer geordneten Beschränkungen nicht von vornherein ausschließt, daß beide optimal sein können. Im vorliegenden Fall verletzt K_2 anders als K_1 die tief geordnete Beschränkung Q-SKOP, aber beide Kandidaten sind grammatisch, weil jeder unter einer bestimmten Ordnung optimal ist.

Trotz diesem Vorteil gibt es ein Problem mit global hierarchischen Kopplungen: die Komplexität. Eine Kopplung zweier Beschränkungen resultiert in einem unterspezifizierten Objekt, das zwei Beschränkungsordnungen generiert, und somit zwei Grammatiken. Diese sind simultan bei einem Sprecher vorhanden. Zwei Kopplungen generieren vier Grammatiken, drei Kopplungen acht Grammatiken, usw. Darüber hinaus gibt es nichts, das ausschließt, daß auch drei oder noch mehr Beschränkungen gekoppelt sind; tatsächlich ist eine Kopplung von drei Beschränkungen von Sells, Rickford, & Wasow (1996) in ihrer Studie der negativen Inversion im African American Vernacular English, also einer unter der schwarzen Bevölkerung verbreiteten Varietät des Englischen, vorgeschlagen worden. Bei einer Dreierkopplung werden sechs Grammatiken erzeugt, gibt es noch einmal eine Kopplung, liegen sechsunddreißig Grammatiken vor, und das alles simultan in einer einzigen Sprache. Man sieht, daß so bereits einige wenige Optionalitätsphänomene in einer Sprache eine ungeheure Vervielfältigung von gleichzeitig präsenten Grammatiken zur Folge haben. Es ist nicht ganz klar, wie schwerwiegend dieses Resultat aus empirischer Perspektive ist; aber es scheint unzweifelhaft, daß global hierarchische Kopplungen damit konzeptuell nicht unproblematisch sind – problematisch genug jedenfalls, um nach restriktiveren Kopplungskonzepten Ausschau zu halten.

4.2. Lokal Hierarchische Kopplung

Pesetsky (1997; 1998) und Broihier (1995) schlagen ein lokales Konzept hierarchischer Kopplungen vor, das das Problem der Vervielfältigung von Grammatiken vermeidet. Wie vorher wird eine Kopplung $A \circ B$ in die beiden möglichen Subordnungen $A \gg B$ und $B \gg A$ aufgelöst. Jetzt fungiert die Kopplung aber nicht als

¹²T_{5.8} kodiert zwei Tabellen: In T_{5.8a} gilt $BKP \gg Q\text{-MARK}$; in T_{5.8b} gilt $Q\text{-MARK} \gg BKP$.

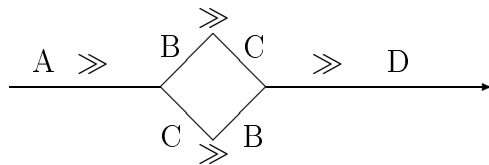
Abkürzung für das gleichzeitige Vorhandensein von zwei (oder mehr) Beschränkungsordnungen in einer Sprache; vielmehr ist die Kopplung ein spezieller Typ einer verletzbaren und geordneten Beschränkung (eine “Metabeschränkung” in der Terminologie von Broihier (1995)). Pesetskys (1998, 348) Vorschlag liest sich wie folgt: “Der Output einer Menge von gekoppelten Beschränkungen ist die Vereinigungsmenge der Outputs einer jeden möglichen Ordnung dieser Beschränkungen.” Dies ist etwas vage formuliert, aber es wird durch Beispiele noch klar werden, was damit gemeint ist. In erster Annäherung können wir lokal hierarchische Kopplungen im Sinne von Pesetsky und Broihier wie folgt definieren:

(27) *Lokal hierarchische Kopplung:*

Es seien $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine vollständige Beschränkungsordnung der Sprache L, B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$, und O_1, \dots, O_n die möglichen Subordnungen in B_i . Dann gilt: B_i ist von einem Kandidaten K verletzt gdw. es keine Subordnung O_i gibt, gemäß der K optimal ist.

Dies bedeutet, daß eine Kopplung wie jede andere Beschränkung im Harmonie-Evaluationsteil der Grammatik auch von einem Kandidaten verletzt oder erfüllt werden kann; insofern ist der Kopplungsbegriff hier strikt lokal. Optionalität kann entstehen, weil es mehr als eine Möglichkeit gibt, eine Kopplung zu erfüllen, je nachdem, wie die Kopplung in Subordnungen aufgelöst wird. Schematisch können lokal hierarchische Kopplungen wie in (28) repräsentiert werden.

(28) *Diagramm der lokal hierarchischen Kopplung $B \circ C$*



In Abschnitt 2.3 von Kapitel 3 haben bereits gesehen, wie in Pesetskys (1997; 1998) System optionale Komplementierung in Objektsätzen im Englischen mit Hilfe einer Kopplung abgeleitet wird. Die relevanten Beispiele sind mit Pesetskys Strukturzuweisungen hier wiederholt (vgl. (48) aus Kapitel 3; man erinnere sich, daß Pesetsky anders als Grimshaw (1997) bloße IP-Einbettung nicht zulassen darf).

- (29) a. I think [_{CP} [C -] [_{IP} John will leave]]
 b. I think [_{CP} [C that] [_{IP} John will leave]]

Was in Kapitel 3 jedoch noch nicht behandelt wurde, ist Pesetskys Analyse des Zusammenbruchs der Optionalität in Subjektsätzen (vgl. (11)):

- (30) a. *[He left so early] shows that he was tired
 b. [That he left so early] shows that he was tired

Zu den bereits bekannten Beschränkungen WIED, LR(CP) und TEL kommt zur Erfassung der Alternation die Beschränkung TCP in (31-d) hinzu.

- (31) a. WIED:
Ein bedeutungstragendes syntaktisches Element muß PF-realisiert werden, es sei denn, es hat ein hinreichend lokales Antezedens.
- b. LR(CP):
Das erste PF-realisierte Wort in einer CP ist der Komplementierer.
- c. TEL:
Funktionswörter dürfen nicht PF-realisiert werden.
- d. TCP (“Tilgung in CP”, “Deletion in CP”, “DCP”):
Der Kopf oder Spezifikator einer CP kann nur getilgt werden, wenn die CP ein Komplement ist.

Die für das Englische angenommene Beschränkungsordnung ist (32):

- (32) *Beschränkungsordnung im Englischen:*
WIED \gg LR(CP) \circ TEL \gg TCP

Für Komplementierertilgung in Objektsätzen ändert sich dadurch nichts. Tabelle T_{5.9} unterscheidet sich von Tabelle T_{3.25} aus Kapitel 3 im wesentlichen nur dadurch, daß TCP hinzugekommen ist. Diese Beschränkung verbietet C-Tilgung aber nur in Nicht-Komplementsätzen; C-Tilgung in einem Objektsatz ist unproblematisch. K₁ und K₂ können hier beide optimal sein, weil sie gemäß (27) beide die lokal hierarchische Kopplung LR(CP) \circ TEL erfüllen. Unter der Auflösung in die Teilordnung O₁: LR(CP) \gg TEL ist K₁ innerhalb von O₁ optimal, und deswegen erfüllt K₁ die Kopplung LR(CP) \circ TEL. Unter der Auflösung in die umgekehrte Teilordnung O₂: TEL \gg LR(CP) ist dagegen K₂ innerhalb von O₂ optimal, und deswegen erfüllt auch K₂ die Kopplung LR(CP) \circ TEL.

T_{5.9}: *Komplementierer in englischen Objektsätzen*

Kandidaten	WIED	LR(CP) TEL	TCP
☞ K ₁ : ... think [CP that John ...]		*	
☞ K ₂ : ... think [CP – John ...]		*	

Bei Subjektsätzen wird TCP relevant und verbietet Komplementierertilgung. Wie Tabelle T_{5.10} zeigt, ist die Verletzung von TCP durch K₂ fatal, auch wenn beide Kandidaten die Kopplung LR(CP) \circ TEL erfüllen. Der Grund ist, daß TCP hier die höchst-geordnete Beschränkung ist, bei der sich K₁ und K₂ unterscheiden (denn bei der Kopplung unterscheiden sie sich nicht).

T_{5.10}: *Komplementierer in englischen Subjektsätzen*

Kandidaten	WIED	LR(CP) TEL	TCP
☞ K ₁ : ... [CP that he left ...] shows ...		*	
K ₂ : ... [CP – he left ...] shows ...		*	*!

Ein zweites Beispiel zeigt, daß (27) noch nicht ganz ausreicht, um Pesetskys Konzept der lokal hierarchischen Kopplung zu erfassen. Das Beispiel involviert Relativsätze im Polnischen; es ist im Detail diskutiert bei Broihier (1995, 23ff.) und Pesetsky (1998, 373-377). Wichtig wird eine weitere, in Kapitel 3 im Zusammenhang mit resumptiven Pronomina in Inselkontexten bereits diskutierte Beschränkung, nämlich ST-SP (“Stille Spur”), die – bei Annahme der Kopiertheorie der Bewegung – komplette PF-Tilgung von Spuren einer Realisierung als resumptives Pronomen vorzieht, und letzteres wiederum einer Realisierung als vollständige Kopie.

(33) ST-SP:

Spuren dürfen nicht PF-realisiert werden.

Die Generalisierung im Polnischen ist nun folgende. Wenn ein Relativpronomen einen obliquen Kasus (wie Dativ) trägt, gibt es zwei Möglichkeiten der PF-Realisierung. Zum einen können das Relativpronomen ausgesprochen und der Komplementierer und die Spur des Relativpronomens getilgt werden; vgl. (34-a). Zum anderen kann das Relativpronomen getilgt, der Komplementierer ausgesprochen und die Spur des Relativpronomens als resumptives Pronomen realisiert werden; vgl. (34-b). Alle anderen Kombinationen, von denen einige in (34-c-f) aufgelistet sind, sind ungrammatisch.

- (34) a. On spotkał studenta [_{CP} któremu₁ – on dał t₁ piątke]
 er traf Student dem er gab gute Note
- b. On spotkał studenta [_{CP} –₁ co mu₁ on dał piątke]
 er traf Student daß ihm er gab gute Note
- c. *On spotkał studenta [_{CP} któremu₁ – on mu₁ dał piątke]
 er traf Student dem er ihm gab gute Note
- d. *On spotkał studenta [_{CP} –₁ co on dał t₁ piątke]
 er traf Student daß er gab t₁ gute Note
- e. *On spotkał studenta [_{CP} któremu₁ co mu₁ on dał piątke]
 er traf Student dem daß ihm er gab gute Note
- f. *On spotkał studenta [_{CP} któremu₁ co on dał t₁ piątke]
 er traf Student dem daß er gab gute Note

Broihier und Pesetsky schlagen vor, daß die hier zu beobachtende Optionalität auf eine Dreierkopplung ST-SP ◦ LR(CP) ◦ TEL zurückgeführt wird; die Beschränkungen WIED und TCP können im Polnischen genau wie im Englischen geordnet sein. Somit resultiert die folgende Ordnung (im Englischen wäre ST-SP höher einzuordnen, da es keine Strategie mit resumptiven Pronomina bei satzgebundener Relativierung gibt):

(35) *Ordnung im Polnischen:*

WIED ≫ ST-SP ◦ LR(CP) ◦ TEL ≫ TCP

Der (34) zugrundeliegende Wettbewerb ist in Tabelle T 5.11 skizziert.

$T_{5.11}$: Relativierung im Polnischen

Kandidaten	WIED	ST-SP	LR(CP)	TEL	TCP
\Rightarrow K ₁ : wh ₁ – ... t ₁			*		*
\Rightarrow K ₂ : \neg_1 co ... mu ₁		*		*	*
K ₃ : wh ₁ – ... mu ₁		*	*		*
K ₄ : \neg_1 co ... t ₁	*!			*	*
K ₅ : wh ₁ co ... mu ₁		*	*	*	
K ₆ : wh ₁ co ... t ₁			*	*	

Die Idee ist hier, daß K₁ und K₂ optimal sind, weil sie (a) mit vielen anderen Konkurrenten die höchst-geordnete Beschränkung WIED erfüllen (die per Annahme nur bei vollständiger Tilgung eines obliquen Arguments *und* seiner Spur wie in K₄ verletzt wird); weil (b) nur K₁ und K₂ die zweithöchste Beschränkung (also die Kopplung) erfüllen; und weil (c) K₁ und K₂ sich bei TCP (sowie allen tiefer geordneten Beschränkungen) gleich verhalten. Das Problem ist, daß Behauptung (b), und somit dieses Ergebnis, nicht aus der Definition der lokal hierarchischen Kopplung in (27) folgt. Zu zeigen ist, daß K₁ und K₂ in $T_{5.11}$ die Dreierkopplung erfüllen, K₃, K₅ und K₆ jedoch nicht (K₄ ist ja durch WIED ausfilterbar). Wenn die Kopplung in eine Subordnung aufgelöst wird, in der TEL am höchsten steht, ist immer K₁ innerhalb der Subordnung optimal. K₁ erfüllt also wie erwünscht die Kopplung. Steht bei der Auflösung ST-SP am höchsten, wird die Ordnung der anderen Beschränkungen der Kopplung wichtig: Bei einer Auflösung in ST-SP \gg TEL \gg LR(CP) erfüllt wieder K₁ allein die Kopplung; und bei einer Auflösung in ST-SP \gg LR(CP) \gg TEL erfüllt K₄ die Kopplung. Letzteres ist an sich unproblematisch, weil K₄ ja fatal WIED verletzt. Ist schließlich bei einer Auflösung der Kopplung in Subordnungen LR(CP) am höchsten plazierte, so hoffen wir, daß jetzt K₂ die Kopplung erfüllen kann. Diese Hoffnung erfüllt sich leider nicht: Unter jeder Subordnung von ST-SP und TEL steht K₄ besser da als K₂. Dies ist nicht deshalb problematisch, weil Kandidat K₄ dadurch grammatisch würde (das wird er wie gesehen aufgrund der fatalen WIED-Verletzung nicht), sondern deshalb, weil damit K₂ nicht grammatisch werden kann und K₁ als der einzig optimale Kandidat in diesem Wettbewerb vorhergesagt wird. Dieses unerwünschte Ergebnis der sechs möglichen Auflösungen der Kopplung ist nochmals in (36) zusammengefaßt:

(36) Innerhalb der Subordnung

- | | |
|---------------------------------|-----------------------------|
| a. TEL \gg ST-SP \gg LR(CP) | ist K ₁ optimal. |
| b. TEL \gg LR(CP) \gg ST-SP | ist K ₁ optimal. |
| c. ST-SP \gg TEL \gg LR(CP) | ist K ₁ optimal. |
| d. ST-SP \gg LR(CP) \gg TEL | ist K ₄ optimal. |
| e. LR(CP) \gg ST-SP \gg TEL | ist K ₄ optimal. |
| f. LR(CP) \gg TEL \gg ST-SP | ist K ₄ optimal. |

Das Problem ist, daß K_4 innerhalb der Kopplung optimal sein kann und somit eine Verletzung der Kopplung durch K_2 bewirkt, obwohl K_4 bereits eine höher geordnete Beschränkung verletzt. Um dieses Problem zu umgehen, müssen wir annehmen, daß ein Kandidat wie K_4 , der eine höher geordnete Beschränkung verletzt, nicht mehr an tieferen Kopplungs-internen Optimierungsprozeduren teilnehmen darf. Den Begriff der lokal hierarchischen Kopplung kann man somit wie in (37) modifizieren.

(37) *Lokal hierarchische Kopplung* (revidiert):

Es seien $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine vollständige Beschränkungsordnung der Sprache L , B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$, und O_1, \dots, O_n die möglichen Subordnungen in B_i . Dann gilt: B_i ist von einem Kandidaten K verletzt gdw. es keine Subordnung O_i gibt, so daß K gemäß $\langle B_1 \gg \dots \gg B_{i-1} \gg O_i \rangle$ optimal ist.

Diese Modifikation macht nun die korrekten Vorhersagen. Allerdings ist dafür ein Preis gezahlt worden: Um herauszufinden, ob ein Kandidat eine Kopplung erfüllt oder verletzt, müssen alle Beschränkungen, die die Kopplung dominieren, ebenfalls betrachtet werden. Damit sind lokal hierarchische Kopplungen nicht mehr ganz so lokal wie bei der ursprünglichen Definition in (27).¹³ Ungeachtet dieser Komplikation läßt sich jedoch festhalten, daß lokal hierarchische Kopplungen das durch global hierarchische Kopplungen hervorgerufene Komplexitätsproblem vermindern. Allerdings entsteht wiederum das Problem, das wir von der Betrachtung der echten Optionalität bereits kennen: Um Optionalität abzuleiten, muß gewährleistet werden, daß es keine tief geordneten Beschränkungen gibt, die die Koexistenz zweier optimaler Kandidaten in einer Kandidatenmenge unterminieren. Das Problem ist aber nicht ganz so drängend wie bei echter Optionalität. Der Grund ist, daß es in einem Ansatz, der mit lokal hierarchischen Kopplungen arbeitet, sehr wohl möglich ist, Beschränkungen zu postulieren, die sich auf die unterschiedlichen Eigenschaften alternierender Kandidaten beziehen, solange eben diese Beschränkungen gekoppelt sind. Beispielsweise kann im vorliegenden Fall eine Beschränkung Komplementierrealisierung verbieten (TEL), eine andere Beschränkung Komplementierertilgung (LR(CP)); so etwas ist bei echter Optionalität nicht möglich.

Abschließend sei festgehalten, daß global hierarchische Kopplungen und lokal hierarchische Kopplungen nicht nur konzeptuell verschieden sind, sondern auch empirisch andere Vorhersagen machen. So wird unter der global hierarchischen Kopplung (20) K_2 in $T_{5.10}$ fälschlicherweise als optimal vorausgesagt (weil die Verletzung der tief geordneten Beschränkung TCP erst erfolgt, als sich die Beschränkungshierarchien bereits gespalten haben). Umgekehrt wird unter der lokal hierarchischen

¹³Interessanterweise wirft auch gerade der in $T_{5.11}$ dargestellte Wettbewerb Lernbarkeitsprobleme für Kopplungen auf; vgl. Brohier (1995, 57). Möglicherweise deutet dies darauf hin, daß der von Brohier und Pesetsky intendierte Kopplungsbegriff in (37) unter konzeptuellen Gesichtspunkten nicht ganz unproblematisch ist.

Kopplung (27) K_2 in $T_{5,8}$ fälschlicherweise als suboptimal vorausgesagt (weil die Verletzung der tief geordneten Beschränkung Q-SKOP fatal sein sollte).

4.3. Lokal Konjunktive Kopplung

Neben global hierarchischen Kopplungen fassen Prince & Smolensky (1993, 51) noch ein zweites, lokales Konzept ins Auge, demgemäß Kopplungen eine wesentliche Ungeordnetheit ('crucial nonranking') der beteiligten Beschränkungen involvieren. Tesar (1998, 428-429) formuliert dasselbe Konzept; er redet von der Mitgliedschaft zweier (oder mehrerer) Beschränkungen im selben "Stratum". Ich nenne diesen Kopplungsbegriff "lokal konjunktive Kopplung". In der Syntax sind lokal konjunktive Kopplungen von Legendre et al. (1995), Legendre, Smolensky & Wilson (1998), Heck (1999) und Müller (1997) verwendet worden. In den ersten beiden Arbeiten geht es u.a. darum, die Alternation von resumptiven Pronomina und Spuren bei Argumenttopikalisierung im Chinesischen aus einer Kopplung der Beschränkungen ÖKON (was Spuren verbietet) und FILL (was die Einsetzung von resumptiven Pronomina verbietet – vgl. auch Grimshaws (1997) VOLL-INT und Pesetskys (1997; 1998) ST-SP) abzuleiten. Ich diskutiere diesen Ansatz hier nicht weiter im Detail, weil er auch mit anderen (globalen wie lokalen) Kopplungsbegriffen vereinbar ist.¹⁴

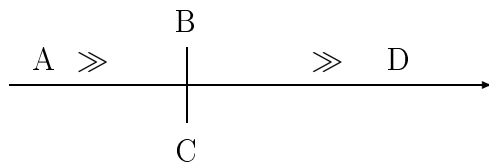
Die Idee ist die folgende. Anders als bei den bisherigen beiden Kopplungskonzepten wird die Kopplung $A \circ B$ *nicht* in die beiden Subordnungen aufgelöst. Vielmehr verschmelzen die beiden Beschränkungen zu einer konjunktiven Beschränkung. Ein Kandidat verletzt die Kopplung, wenn er eine Beschränkung der Kopplung verletzt; mehrfache Verletzungen in diesem Bereich addieren sich auf.

(38) *Lokal konjunktive Kopplung:*

Es sei $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine vollständige Beschränkungsordnung der Sprache L und B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$. Dann gilt: B_i ist von einem Kandidaten K verletzt gdw. es eine Beschränkung B_{i_j} in der Kopplung B_i gibt, die K verletzt.

(39) illustriert die Arbeitsweise lokal konjunktiver Kopplungen.

(39) *Diagramm der lokal konjunktiven Kopplung $B \circ C$*



¹⁴Tatsächlich sprechen Legendre et al. (1995; 1998) zwar von strikter Gleichordnung ("equal ranking") der gekoppelten Beschränkungen VOLL-INT und ÖKON im Chinesischen; eindeutig identifiziert wird das von ihnen zugrundegelegte Konzept allerdings erst durch den expliziten Verweis auf Tesar (1998).

Als Beispiel sei die W-Skopusmarker-Alternation bei Fragesätzen im Deutschen betrachtet, die in Müller (1997) mit Hilfe von lokal konjunktiven Kopplungen analysiert wird. Die einschlägigen Daten sind hier wiederholt. (40-ab) (= (3)) veranschaulichen die Optionalität von sukzessiv-zyklischer W-Bewegung und Skopusmarkereinsatzung mit partieller W-Bewegung in transparenten Kontexten; (40-cd) (= (13)) zeigen, daß in Inselkontexten (hier: bei Subjekt-Inseln) Skopusmarkereinsatzung durch sukzessiv-zyklische Bewegung blockiert wird.

- (40) a. Wen₁ glaubst du [_{CP} t'₁ daß man t₁ einladen sollte] ?
 b. Was₁ glaubst du [_{CP} wen₁ (daß) man t₁ einladen sollte] ?
 c. ?[_{PP} Mit wem]₁ ist es schade [_{CP} t'₁ daß sie t₁ gesprochen hat] ?
 d. *Was₁ ist es schade [_{CP} [_{PP} mit wem]₁ sie t₁ gesprochen hat] ?

Die Grundzüge der Analyse sind bereits in Kapitel 2 (Abschnitt 3.2) dargelegt worden. Die Beschränkungen, die dabei eine Rolle spielen, sind nochmals in (41) aufgelistet; die für das Deutsche angenommene Ordnung mit der Kopplung von ÖKON und VOLL-INT ist in (42) angegeben.

- (41) a. W-KRIT:
 (i) Eine W-Phrase muß auf der S-Struktur in der Domäne von C stehen.
 (ii) C_[+w] verlangt ein geeignetes lexikalisches Element in seiner Domäne.
 b. BAR:
 Kettenbildung darf keine Barriere überschreiten.
 c. ÖKON:
 Bewegung ist verboten (Spur ist nicht erlaubt).
 d. VOLL-INT:
 Expletiveinsatzung ist verboten.

- (42) *Beschränkungsordnung im Deutschen:*
 W-KRIT \gg BAR \gg ÖKON \circ VOLL-INT

Tabelle T_{5.12} (die im wesentlichen Tabelle T_{2.14} aus Kapitel 2 entspricht) zeigt, daß K₁ und K₂ beide optimal sind, weil sie ein besseres Beschränkungsprofil haben als alle anderen Kandidaten des Wettbewerbs und sich darüber hinaus auch hinsichtlich der Kopplung ÖKON \circ VOLL-INT gleich verhalten: K₂ verletzt durch zweifache W-Bewegung ÖKON zweimal; K₁ verletzt durch partielle W-Bewegung ÖKON einmal, und durch W-Skopusmarkereinsatzung VOLL-INT ebenfalls einmal. Zusammengenommen ergibt das ein identisches Verhalten bei der Kopplung.¹⁵

Unter der Annahme, daß partieller W-Bewegung auf der S-Struktur W-

¹⁵Kandidat K₆ (der in Kapitel 2 noch nicht betrachtet worden ist) weist zweifache Skopusmarkereinsatzung, aber keine W-Bewegung auf und hat somit durch zwei VOLL-INT-Verletzungen bei der Kopplung ein zu K₁ und K₂ identisches Verhalten. Er verletzt jedoch fatal die hoch geordnete Beschränkung W-KRIT, da die eigentliche W-Phrase in situ verharret.

T_{5.12}: Partielle und lange W-Bewegung im Deutschen

Kandidaten	W-KRIT	BAR	ÖKON VOLL-INT
☞ K ₁ : [CP was ... [CP wann ₁ ... t ₁			* *
☞ K ₂ : [CP wann ₁ ... [CP t' ₁ ... t ₁			**
K ₃ : [CP was ... [CP - ... wann ₁	*!		*
K ₄ : [CP - ... [CP - ... wann ₁	*!*		
K ₅ : [CP - ... [CP wann ₁ ... t ₁	*!		*
K ₆ : [CP was ... [CP was ... wann ₁	*!		**

Bewegung zum Skopusmarker auf LF folgt, ergibt sich der Zusammenbruch der Optionalität in Subjekt-Inselkontexten: In T_{5.13} verletzt K₁ zweimal BAR (auf der S-Struktur und auf LF); K₂ verletzt BAR dagegen nur einmal (auf der S-Struktur).

T_{5.13}: Verbot partieller W-Bewegung in Subjektsätzen im Deutschen

Kandidaten	W-KRIT	BAR	ÖKON VOLL-INT
K ₁ : [CP was ... [CP mit wem ₁ ... t ₁		**!	* *
☞ K ₂ : [CP mit wem ₁ ... [CP t' ₁ ... t ₁		*	**
K ₃ : [CP was ... [CP - ... mit wem ₁	*!	**	*
K ₄ : [CP - ... [CP - ... mit wem ₁	*!*	*	
K ₅ : [CP - ... [CP mit wem ₁ ... t ₁	*!	*	*

Bis hier könnte die Skopusmarkeralternation im Deutschen ebenso gut mit Hilfe von global hierarchischen oder lokal hierarchischen Kopplungen erfaßt werden. Unterschiedliche Vorhersagen ergeben sich jedoch im Falle längerer W-Abhängigkeiten, die zwei CPs überschreiten. Es zeigt sich, daß hier nur lokal konjunktive Kopplungen die Daten korrekt erfassen. Man betrachte die folgenden Sätze:

- (43) a. Wann₁ meinst du [CP t''₁ daß sie gesagt hat [CP t'₁ daß sie t₁ kommen würde]] ?
 b. Was₁ meinst du [CP wann₁ sie gesagt hat [CP t'₁ daß sie t₁ kommen würde]] ?
 c. Was₁ meinst du [CP was₁ sie gesagt hat [CP wann₁ sie t₁ kommen würde]] ?

Alle drei Sätze sind gleichermaßen grammatisch. (43-a) weist sukzessiv-zyklische W-Bewegung bis in die Zielposition hinein auf. Dadurch entstehen drei Verletzungen von ÖKON, aber es resultiert keine Verletzung von VOLL-INT. Demgegenüber erfolgt in (43-b) zweifache W-Bewegung bis in die SpecC-Position des höchsten eingebetteten Satzes, sowie Skopusmarkereinsatz in der Zielposition. Das ergibt zwei Verletzungen von ÖKON und eine Verletzung von VOLL-INT. (43-c) schließlich hat kurze, satzgebundene W-Bewegung mit zweifacher Skopusmarkereinsatz. Dadurch resultieren eine Verletzung von ÖKON und zwei Verletzungen von

VOLL-INT. Insgesamt ergeben sich in allen drei Fällen drei Verletzungen der Kopplung $\ddot{O}KON \circ VOLL-INT$. Daher werden in Tabelle T_{5.14} bei Annahme eines lokal konjunktiven Kopplungsbegriffs wie in (38) K₁, K₂ und K₃ alle korrekt als optimal vorhergesagt (suboptimale Kandidaten, die höher geordnete Beschränkungen verletzen, sind hier weggelassen).

T_{5.14}: *Skopusmarkeralternation bei längeren Abhängigkeiten im Deutschen*

Kandidaten	W- KRIT	BAR-	Ö- VOLL- KON INT
☞K ₁ : [CP wann ₁ ... [CP t' ₁ daß ... [CP t' ₁ daß ... t ₁]]			***
☞K ₂ : [CP was ₁ ... [CP wann ₁ ... [CP t' ₁ daß ... t ₁]]			** *
☞K ₃ : [CP was ₁ ... [CP was ₁ ... [CP wann ₁ ... t ₁]]			* **

Dieses Ergebnis läßt sich nicht unter hierarchischen Kopplungskonzepten erzielen. Global hierarchische Kopplungen machen für T_{5.14} die falsche Vorhersage, daß K₂ suboptimal sein sollte: K₁ wäre optimal unter der Auflösung $VOLL-INT \gg \ddot{O}KON$, K₃ wäre optimal unter der Auflösung $\ddot{O}KON \gg VOLL-INT$, aber es gibt keine Subordnung, unter der K₂ optimal würde. Aus den gleichen Gründen machen auch lokal hierarchische Kopplungen für T_{5.14} eine inkorrekte Prognose: Es gibt keine Kopplungs-interne Subordnung, unter der K₂ optimal wäre, also würde K₂ die Kopplung verletzen.

Andererseits ist es auch nicht so, daß lokal konjunktive Kopplungen mit der für global hierarchische oder lokal hierarchische Kopplungen angeführten Evidenz vollständig vereinbar wären. Wie lokal hierarchische Kopplungen machen lokal konjunktive Kopplungen für den durch global hierarchische Kopplungen korrekt erfaßten Wettbewerb in Tabelle T_{5.8} (optionale W-Bewegung in französischen Hauptsätzen) eine falsche Vorhersage. Erstaunlicherweise ist es bei lokal konjunktiven Kopplungen aber nicht Kandidat K₂, der suboptimal sein sollte (wie bei lokal hierarchischen Kopplungen), sondern (aufgrund einer Anhäufung von Verletzungen von BKP) K₁. Darüber hinaus machen lokal konjunktive Kopplungen zwar dieselben Prognosen wie lokal hierarchische Kopplungen bei den Wettbewerben in T_{5.9} und T_{5.10}; für T_{5.11} (Relativierung im Polnischen) gilt dies jedoch nicht: K₂ sollte hier aufgrund einer zusätzlichen Verletzung der verschmolzenen Beschränkungen gegenüber K₁ suboptimal sein.

Allgemeiner läßt sich festhalten, daß lokal konjunktive Kopplungen dieselbe schwächere Version des inhärenten Problems für echte Optionalität hervorrufen wie lokal hierarchische Kopplungen: Man muß sicherstellen, daß tiefer geordnete Beschränkungen nicht doch ungewollt freie Alternation blockieren. Darüber hinaus ist nicht ganz klar, wie zwei völlig unterschiedliche Beschränkungen in manchen Sprachen zu einer verschmelzen können, in anderen dagegen nicht, und welche Konsequenzen diese Verschmelzung für Spracherwerb und Sprachwandel hat. Wie kann ein Kind herausfinden, ob zwei universelle, angeborene Beschränkungen buchstäblich als eine gelten, wie kann eine Sprache jemals eine Kopplung wieder loswerden,

wenn die Kopplung keine von außen wahrnehmbare interne Struktur hat? Antworten auf diese Art von Fragen finden sich zwar in der Literatur (vgl. Tesar & Smolensky (1993; 1998; 2000), Tesar (1998) und, aus anderem Blickwinkel, Müller (1997)), aber sie sind doch z.T. spekulativ und jedenfalls nicht auf der Hand liegend.

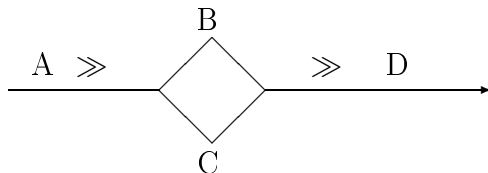
4.4. Disjunktive Kopplungen

Die drei bisher diskutierten Kopplungskonzepte sind die in der Praxis der Optimalitätstheoretischen Syntax vorherrschenden. Die Menge der logisch möglichen und linguistisch plausiblen Kopplungskonzepte ist damit jedoch noch nicht erschöpft. Zum Schluß dieses Abschnitts möchte ich noch das Konzept der disjunktiven Kopplung vorstellen, von dem es in der Literatur eine lokale Version (vgl. Brohier (1995)) und eine globale Version (vgl. Müller (1999c)) gibt. Der Kern beider Ansätze ist, daß ein Kandidat unter einer Kopplung optimal sein kann nur dann, wenn er wenigstens eine der bei der Kopplung beteiligten Beschränkungen erfüllt. Lokal disjunktive Kopplungen können wie in (44) definiert und wie in (45) illustriert werden (vgl. die Definition lokal konjunktiver Kopplungen in (38)).

(44) *Lokal disjunktive Kopplung:*

Es sei $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine vollständige Beschränkungsordnung der Sprache L und B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$. Dann gilt: B_i ist von einem Kandidaten K verletzt gdw. es keine Beschränkung B_{i_j} in der Kopplung B_i gibt, die K erfüllt.

(45) *Diagramm der lokal disjunktiven Kopplung $B \circ C$*



Brohier (1995) erwähnt zwar dieses Kopplungskonzept, verwendet es dann aber in seiner Analyse nicht. Tatsächlich sind mir keine Ansätze bekannt, in denen lokal disjunktive Kopplungen eine Rolle spielen.¹⁶ Etwas anders liegt der Fall bei den global disjunktiven Kopplungen.

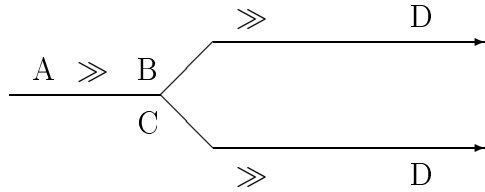
Die Definition einer global disjunktiven Kopplung sieht eher aus wie bei der global hierarchischen Kopplung; dasselbe gilt auch für die schematische Darstellung in (46) (mit einer Verzweigung der Hierarchien, die immer Globalität signalisiert).

¹⁶Es gibt allerdings eine Verwandtschaft zur disjunktiv interpretierten lokalen Konjunktion; vgl. Kapitel 3.

(46) *Global disjunktive Kopplung:*

Es sei $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine partielle Beschränkungsordnung der Sprache L und B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$. Dann gilt: Für jede Beschränkung B_{i_j} in B_i ist auch Γ_{i_j} eine Beschränkungsordnung der Sprache L , wobei Γ_{i_j} sich von Γ nur dadurch unterscheidet, daß B_i durch B_{i_j} ersetzt ist.

(47) *Diagramm der global disjunktiven Kopplung $B \circ C$*



Global disjunktive Kopplungen werden in Müller (1999c) vorgeschlagen, um die Optionalität von Scrambling im Deutschen zu erfassen; vgl. etwa (6), hier wiederholt als (48):¹⁷

- (48) a. daß $[_{NP_1} \text{keiner}] [_{NP_2} \text{den Fritz}]$ gesehen hat
- b. daß $[_{NP_2} \text{den Fritz}] [_{NP_1} \text{keiner}] t_2$ gesehen hat

Angenommen sei, daß Scrambling nur dann unter Verletzung von ÖKON möglich ist, wenn dies zu einer Erfüllung wenigstens einer aus einer Vielzahl von gekoppelten Linearisierungsbeschränkungen führt; dies ist durchaus der Theorie von Uszkoreit (1984; 1986) vergleichbar, die in Abschnitt 2 von Kapitel 4 referiert wurde. Zu diesen gekoppelten Beschränkungen gehören die folgenden (vgl. (53) aus Kapitel 2).

- (49) a. NOM (“Nominativbedingung”):
 Eine Nominativ-NP geht einer NP mit Objektkasus voran.
- b. DEF (“Definitheitsbedingung”):
 Eine definite NP geht einer indefiniten NP voran.

Die Optionalität von Scrambling in (48) folgt nun auf die in Tabelle T_{5.15} gezeigte Weise:

T_{5.15}: *Optionalität bei Scrambling im Deutschen*

Kandidaten	NOM DEF	ÖKON
$\Leftrightarrow K_1: \dots \text{keiner}_1 \text{den Fritz}_2 \dots$	*	
$\Leftrightarrow K_2: \dots \text{den Fritz}_2 \text{keiner}_1 t_2 \dots$	*	*

¹⁷Die folgende Darstellung ist extrem vereinfacht; sie dient einzig und allein der Darstellung des involvierten Kopplungskonzepts. Zu weiteren Bemerkungen zur Wortstellung vgl. Kapitel 6.

K_1 und K_2 unterscheiden sich bei der unter der Kopplung geordneten Beschränkung ÖKON, da K_2 Scrambling involviert, K_1 nicht. Dies ist unproblematisch unter einer global disjunktiven Kopplung; ein Problem würde allerdings unter einer lokal disjunktiven Kopplung (sowie unter lokal hierarchischen und lokal konjunktiven Kopplungen) entstehen. Es sieht jedoch so aus, als ergäbe sich das Ergebnis in $T_{5.15}$ auch unter dem besser etablierten Konzept der global hierarchischen Kopplung. Dies ist zwar für den hier betrachteten Fall wirklich so, aber das Resultat läßt sich nicht auf alle Instanzen von Optionalität ausweiten, die in Müller (1999c) behandelt werden. Man nehme z.B. eine weitere gekoppelte Beschränkung FOK (“Fokusbedingung”) an, derzufolge ein nicht-fokussiertes Element einem fokussierten Element vorangeht (d.h., Fokusplatzierung ist ein weiterer Auslöser für Scrambling). FOK wird nun z.B. in einer ersten Variante von K_1 (K_{1a}) respektiert, in der das Objekt NP_2 Fokusexponent ist; aber FOK wird in der zweiten Variante von K_1 (K_{1b}) verletzt, in der das Subjekt NP_1 Fokusexponent ist. Angenommen nun, es gibt keine anderen Beschränkungen, bei denen sich K_{1a} und K_{1b} unterschiedlich verhalten. Da K_{1b} jetzt alle Beschränkungen verletzt, die K_{1a} verletzt, aber nicht umgekehrt, ergibt sich das Resultat, daß K_{1a} K_{1b} harmonisch begrenzt. Damit wird aber unter einer global hierarchischen Kopplung (de facto unter jeder nicht-disjunktiven Kopplung) K_{1b} durch K_{1a} als suboptimal blockiert. Dieses Problem entsteht nicht unter einer global disjunktiven Kopplung, denn in einer vollständigen Beschränkungsordnung, in der die Kopplung durch NOM allein ersetzt ist, sind K_{1a} und K_{1b} mit identischem Beschränkungsprofil beide optimal. Vgl. Tabelle $T_{5.16}$.¹⁸

$T_{5.16}$: Fokus und Scrambling im Deutschen

Kandidaten	NOM	DEF	FOK	ÖKON
☞ K_{1a} : ... keiner ₁ DEN FRITZ ₂ ...		*		
☞ K_{1b} : ... KEINER ₁ den Fritz ₂ ...		*	*	

Somit ergibt sich, daß global disjunktive Kopplungen Aufgaben übernehmen können, an denen andere Kopplungskonzepte scheitern müssen. Umgekehrt fällt aber auch der Nachweis wieder leicht, daß beide Typen von disjunktiven Kopplungen nicht für die Aufgaben geeignet sind, für die die anderen Kopplungskonzepte verwendet werden. Beispielsweise sagt im Wettbewerb $T_{5.11}$ (Relativierung im Polnischen) die lokal hierarchische Kopplung von Pesetsky (1997; 1998) und Broihier (1995) korrekt vorher, daß genau K_1 und K_2 optimal sind. Gemäß der lokal disjunktiven Kopplung dagegen sollte merkwürdigerweise nur der ungrammatische Kandidat K_6 optimal sein, denn außer Kandidat K_5 (der alle Beschränkungen der Kopplungen verletzt) erfüllen nach (44) alle anderen Wettbewerber problemlos die Kopplung (weil sie wenigstens eine Beschränkung der Kopplung erfüllen), und dann

¹⁸Fokussierung ist wie üblich durch Kapitälchen angezeigt.

entscheidet zwischen K_1 , K_2 , K_3 und K_6 erst die tief geordnete Beschränkung TCP, die nur K_6 erfüllt. Ähnlich falsch ist auch das Ergebnis unter einer global disjunktiven Kopplung: Jetzt sollten in $T_{5.11}$ zwar immerhin die beiden grammatischen Kandidaten K_1 und K_2 optimal sein, leider aber auch noch die ungrammatischen Kandidaten K_3 und K_6 ; denn für alle diese Kandidaten gibt es gemäß (46) bei Ersetzung der Kopplung durch eine beliebige Teilbeschränkung der Kopplung eine vollständige Ordnung, unter der sie optimal sind (die ersetzende Beschränkung ist bei K_1 und K_3 TEL, bei K_2 LR(CP) und bei K_6 ST-SP).

Generell werfen disjunktive Kopplungen all die Probleme auf, wie wir bei den anderen Kopplungskonzepten schon gesehen haben. Lokal disjunktive Kopplungen machen es wie andere lokale Kopplungen schwer, sicherzustellen, daß tief geordnete Beschränkungen nicht unerwünschterweise Optionalität zerstören können. Global disjunktive Kopplungen lassen wie andere globale Kopplungen die Komplexität der Grammatik ansteigen. Zusätzlich haben beide Typen von disjunktiven Kopplungen noch eine Eigenschaft, die sie zumindest als einziges Mittel, Optionalität zu erfassen, unbrauchbar macht: Für einen Kandidaten, der eine Beschränkung einer disjunktiven Kopplung erfüllt, werden alle anderen Beschränkungen in der Kopplung vollkommen irrelevant. Wenn z.B. eine so allgemeine Beschränkung wie ÖKON und eine beliebige Beschränkung X disjunktiv gekoppelt sind, dann bedeutet das, daß ein Kandidat, der X erfüllt, ÖKON beliebig oft und in ganz unabhängigen Bereichen verletzen kann, weil ÖKON für diesen Kandidaten gar nicht mehr sichtbar ist. Dies ist kaum ein akzeptables Ergebnis.

Insgesamt ergibt sich als Konklusion aus diesem Abschnitt, daß die hier betrachteten Kopplungsbegriffe jeweils ganz gut für den Bereich funktionieren, für den sie entwickelt bzw. ursprünglich verwendet wurden. Allerdings hat es sich gezeigt, daß sie sämtlich Probleme aufwerfen – Komplexitätsprobleme bei globalen Kopplungen und durch tief geordnete Beschränkungen ausgelöste Interferenzprobleme bei lokalen Kopplungen. Weiterhin hat sich herausgestellt, daß die verschiedenen Kopplungskonzepte empirische Vorhersagen machen, die radikal unterschiedlich und inkompatibel sind. Es gibt kein Kopplungskonzept, das sämtliche in der Literatur vorgeschlagenen Kopplungsanalysen für syntaktische Optionalität erfassen könnte. Das ist Grund genug, sich nochmals nach einer anderen Behandlung von Optionalität und Alternation umzusehen. Eine Möglichkeit ist, den Pseudo-Optionalitätsansatz aus einer anderen Perspektive zu betrachten, der der Neutralisierung.

5. Neutralisierung

Man erinnere sich, daß das Hauptproblem des Pseudo-Optionalitätsansatzes war, daß das Phänomen der Alternation, also der Zusammenbruch von Optionalität in bestimmten syntaktischen Kontexten, nicht erfassbar war: Um Optionalität zweier Konstruktionen α , β abzuleiten, dürfen α und β nicht in derselben Kandidatenmenge sein; aber damit α β unter bestimmten Umständen als suboptimal blockieren kann, müssen α und β doch wieder miteinander im Wettbewerb stehen. Dieses

Dilemma sieht so formuliert zunächst einmal unlösbar aus. Wir haben aber dasselbe Problem schon bei der Behandlung absoluter Ungrammatikalität kennengelernt: In bestimmten Kontexten soll eine Konstruktion α eine Konstruktion β als absolut ungrammatisch blockieren; aber in anderen Kontexten sollen beide Konstruktionen optimal sein, obwohl sie unmöglich ein identisches Beschränkungsprofil haben können. Eine für absolute Ungrammatikalität in Kapitel 2 (Abschnitt 8.3.5) und Kapitel 3 (Abschnitt 3) diskutierte Lösung ist die Neutralisierung: α und β sind einerseits Input-treue Kandidaten in “ihren” Kandidatenmengen, andererseits Input-untreue Kandidaten in der Kandidatenmenge “des jeweils anderen”. Absolute Ungrammatikalität heißt nichts anderes, als daß in bestimmten Kontexten β nicht nur als Input-treuer Kandidat in seiner eigenen Kandidatenmenge gewinnt, sondern auch noch als Input-untreuer Kandidat in der Kandidatenmenge von α ; mit anderen Worten: Eine im Input gemachte Unterscheidung wird im Output neutralisiert. Man erinnere sich an die Beispiele in (80) aus Kapitel 2, die hier wiederholt sind:

- (50) a. *Was₁ ist Fritz eingeschlafen [CP nachdem er t₁ gelesen hat] ?
 b. Fritz ist eingeschlafen [CP nachdem er was₁ gelesen hat] (= etwas)

In einer Neutralisierungsanalyse von absoluter Ungrammatikalität gewinnt (50-b), mit einer Realisierung von *was* als Indefinitpronomen, sowohl den Wettbewerb, in dem im Input eine [-w]-Spezifikation auf *was* vorliegt, als auch (aufgrund einer hoch geordneten Lokalitätsbeschränkung) den Wettbewerb, in dem im Input eine [+w]-Spezifikation auf *was* vorhanden ist (sowie auf dem Hauptsatz-C-Kopf). In bewegungstransparenten Kontexten wie in (51) setzt sich dagegen immer der Input-treue Kandidat durch; hier hat z.B. (51-b) keine Chance, (51-a) in “dessen” Kandidatenmenge (also bei [+w]-Spezifikation im Input) zu blockieren.

- (51) a. Was₁ mache ich t₁ ?
 b. Ich mache was₁ (= etwas)

In einer Neutralisierungsanalyse der Optionalität z.B. bei Komplementierertilgung im Englischen verhält sich nun das Optionalität zeigende Beispielpaar in (52-cd) wie (51), und das Beispielpaar in (52-ab), in dem die Optionalität zusammenbricht, wie (50).

- (52) a. *I think [IP most of the time [IP John could accept this solution]]
 b. I think [CP that [IP most of the time [IP John could accept this solution]]]
 c. I think [IP John will leave]
 d. I think [CP that [IP John will leave]]

Gibt es dann überhaupt noch einen theoretisch signifikanten Unterschied zwischen Neutralisierung bei absoluter Ungrammatikalität und Neutralisierung bei (zusammenbrechender) Optionalität? Eigentlich nicht. Der einzige Unterschied betrifft die konventionalisierte Redeweise: Da üblicherweise von Optionalität nur gesprochen wird, wenn es “zwei oder mehr Möglichkeiten zu geben scheint, ein und dieselbe

Sache zu sagen" (vgl. die Festlegung am Anfang des Kapitels), fällt es uns schwer, bei Satzpaaren wie dem in (51) von Optionalität zu reden.

Nach diesen Klärungen können wir uns einer tatsächlichen Neutralisierungsanalyse von Alternationen wie der in (52) zuwenden. Derartige Analysen sind für Komplementierertilgung entwickelt worden von Legendre, Wilson, Smolensky, Homer & Raymond (1995), Baković (1997), Keer & Baković (1997) und Baković & Keer (1999).¹⁹ Die folgende Darstellung orientiert sich an Baković (1997).

Zunächst werden zwei Beschränkungen gebraucht: Neben Grimshaws (1997) Markiertheitsbeschränkung PUR-EP, die hier aus (17-a) wiederholt ist, erscheint die Treuebeschränkung SUB-TREUE, die die Aufrechterhaltung eines im Input spezifizierten Selektionsmerkmals [\pm Sub] (für Subordination) im Output verlangt; die Ordnung im Englischen ist PUR-EP \gg SUB-TREUE.²⁰

(53) a. PUR-EP:

Adjunktion an die höchste XP einer eingebetteten erweiterten Projektion und Bewegung in deren Kopf sind verboten.

b. SUB-TREUE ("Subordinationstreue", "Faith[Sub]"):

Ein Kandidat muß die [\pm Sub]-Spezifikation des Inputs realisieren.

([+Sub]/input \rightarrow [+Sub]/output; [-Sub]/input \rightarrow [-Sub]/output)

Das Merkmal [\pm Sub] spezifiziert die Selektionsanforderung eines Satz-einbettenden Prädikats.²¹ Per Annahme verlangt [+Sub] eine CP (die aufgrund einer hoch geordneten Beschränkung OB-KOPF wie bei Grimshaw (1997) dann ihrerseits Füllung von C verlangt), während [-Sub] eine IP fordert. Brückenverben wie *think* können alternativ mit einem [+Sub]-Merkmal oder einem [-Sub]-Merkmal ausgestattet sein, und dieser Unterschied resultiert in zwei verschiedenen Inputs, also auch zwei verschiedenen Kandidatenmengen.²² Wie T_{5.17} und T_{5.18} zeigen, sind die Kandidaten (52-c) und (52-d) unter diesen Annahmen beide optimal, jeder in "seiner" Kandidatenmenge, also ohne Änderung der [\pm Sub]-Spezifikation des Inputs.

¹⁹Während in Schmid (1999) Optionalität bei bestimmten Ersatzinfinitivkonstruktionen mit Hilfe globaler Kopplungen abgeleitet wird, enthält Schmid (1998) eine Neutralisierungsanalyse dieses Phänomens.

²⁰In Baković & Keer (1999) ist [\pm Sub] durch das analoge Merkmal [\pm Comp] ersetzt.

²¹Tatsächlich nehmen Baković & Keer (1999, 7) an, daß dieses Merkmal im Input nicht auf dem Verb, sondern auf dem eingebetteten Satz selbst auftaucht ("an embedded clause may be freely specified in the input as [+Comp] or as [-Comp]"). Dies macht aber einerseits dieselben empirischen Vorhersagen; andererseits erscheint mir eine solche Annahme konzeptuell problematisch (es ist unklar, wo genau das Merkmal [\pm Sub]/[\pm Comp] auf dem "clause" im Input lokalisiert ist, wenn es nicht auf einer X⁰-Kategorie auftritt).

²²Legendre et al. (1995, 618) vermeiden den Umweg über ein Merkmal [\pm Sub] und nehmen an, daß bei Brückenverben wie *think* von vornherein zwei Varianten *think*_{IP} und *think*_{CP} existieren.

T_{5.17}: Optionaler Komplementierer im Englischen: [+Sub] im Input

Kandidaten	PUR-EP	SUB-TREUE
☞ K ₁ : ... think _[+Sub] [CP that [IP ...		
K ₂ : ... think _[+Sub] [IP ...		*!

T_{5.18}: Optionaler Komplementierer im Englischen: [-Sub] im Input

Kandidaten	PUR-EP	SUB-TREUE
K ₁ : ... think _[-Sub] [CP that [IP ...		*!
☞ K ₂ : ... think _[-Sub] [IP ...		

Soweit ist alles wie in einer reinen Pseudo-Optionalitätsanalyse. Das Problem für eine solche Analyse war, daß *that* obligatorisch wird, wenn IP-Adjunktion vorliegt. Dies folgt nun im Neutralisierungsansatz: Bei IP-Adjunktion wird die höher geordnete Markiertheitsbeschränkung PUR-EP relevant. Damit gewinnt (52-b) nun nicht nur "seinen" Wettbewerb, wo dieser Kandidat Input-treu ist, sondern auch den Wettbewerb, in dem (52-a) Input-treu ist, (52-b) aber nicht. Mit anderen Worten: Durch die hohe Ordnung von PUR-EP findet in Kontexten, wo Input-Treue gemäß SUB-TREUE die Verletzung ersterer Beschränkung impliziert, eine Neutralisierung der [±Sub]-Spezifikation statt. Dies zeigen die Tabellen T_{5.19} und T_{5.20}.

T_{5.19}: Obligatorischer Komplementierer bei IP-Adjunktion: [+Sub] im Input

Kandidaten	PUR-EP	SUB-TREUE
☞ K ₁ : ... think _[+Sub] [CP that [IP XP [IP ...		
K ₂ : ... think _[+Sub] [IP XP [IP ...	*!	*

Soweit der Kern der Analyse von Baković (1997), der auch in den zitierten Arbeiten von Baković und Keer wieder auftaucht.²³ Es scheint nun nicht ausgeschlossen, daß ein Großteil der Ansätze, die auf Kopplungen beruhen, auf diese Weise rekonstruiert werden könnte. Zudem spricht für Neutralisierungsanalysen von Optionalität, daß mit demselben Mittel das Problem der absoluten Ungrammatikalität in der optimalitätstheoretischen Syntax erfolgreich angegangen werden kann. Andererseits bleiben die bereits in Abschnitt 8.3.5 von Kapitel 2 und Abschnitt 3.2 von Kapitel 3 für Neutralisierungsanalysen von Ungrammatikalität vermerkten Probleme natürlich auch bei Anwendungen auf Optionalität voll erhalten. Das Hauptproblem ist nach wie vor der Komplexitätszuwachs auf mehreren Ebenen.

²³Der Ansatz von Legendre et al. (1995) ist im Prinzip ganz ähnlich und, wie in der letzten Fußnote vermerkt, in gewisser Weise konzeptuell einfacher; dort wird allerdings der Zusammenbruch von Optionalität anhand von komplexeren Extraktionsdaten illustriert, nicht anhand des einfacheren Phänomens der IP-Adjunktion.

$T_{5.20}$: *Obligatorischer Komplementierer bei IP-Adjunktion: [-Sub] im Input*

Kandidaten	PUR-EP	SUB-TREUE
\Leftarrow K ₁ : ... think _[-Sub] [CP that [IP XP [IP ...		*
K ₂ : ... think _[-Sub] [IP XP [IP ...	*!	

Erstens ist zur Vermeidung derivationaler Ambiguitäten das Meta-Verfahren der Input-Optimierung notwendig. Zweitens muß der Generator die Fähigkeit besitzen, massiv unter Verletzung von Treuebeschränkungen das Inputmaterial zu manipulieren. Dies betrifft nicht nur wie bei Baković und Keer abstrakte Merkmale, sondern, wie wir in den Kapiteln 2 und 3 gesehen haben, auch bedeutungstragende Elemente. Letzten Endes wird man zur Konsequenz gezwungen, daß jeder Satz mit jedem beliebigen anderen Satz konkurriert (auch wenn die hinreichend untreuen Kandidaten niemals eine Chance haben, optimal zu werden). Damit werden Kandidatenmengen extrem mächtig. Drittens schließlich werden die Kandidatenmengen in Neutralisierungsanalysen nicht nur komplexer, sondern auch mehr: Da Inputs, die sich z.B. nur hinsichtlich eines abstraken Selektionsmerkmals unterscheiden, bereits zu verschiedenen Wettbewerben führen, müssen zunächst einmal (d.h., vor Input-Optimierung) mehr Kandidatenmengen für ein und denselben grammatischen Satz betrachtet werden.

6. Konklusion

Das Vorliegen von Optionalität ist per se ein Problem für die Optimalitätstheorie. Dieses Problem ist durch Pseudo-Optionalitätsanalysen gewaltsam lösbar: Hier resultiert Optionalität einfach daraus, daß verschiedene Kandidaten verschiedene Wettbewerbe gewinnen, daß also gar kein Wettbewerb besteht. Diese Lösung wird durch die Tatsache infrage gestellt, daß viele Fälle von Optionalität als Alternationen zu reanalysieren sind, in dem Sinne, daß die Optionalität in bestimmten Kontexten zusammenbricht. Dieses Phänomen suggeriert stark das Vorliegen eines Wettbewerbs. Die auf echter Optionalität, globalen oder lokalen Kopplungen sowie Neutralisierung beruhenden Ansätze gehen alle von einem Wettbewerb bei syntaktischer Optionalität aus, und sie können alle im Prinzip mit Optionalität und ihrem Zusammenbruch zurecht kommen. Aber auch diese Ansätze sind nicht problemlos: Bei echter Optionalität und, in geringerem Maße, bei lokalen Kopplungen muß gewährleistet werden, daß niedrig geordnete Beschränkungen nicht interferieren und ein identisches Beschränkungsprofil zweier Kandidaten unerwünscht unterminieren. Das ist nicht immer eine leichte Aufgabe. Globale Kopplungen werfen demgegenüber das Problem der Komplexitätszunahme auf: Eine Sprache mit nur wenigen Fällen syntaktischer Alternation muß bereits über viele simultan vorhandene verschiedene Grammatiken verfügen. Schließlich läßt auch Neutralisierung die Komplexität des System ansteigen, weil damit mehr und größere Kandidatenmengen vorausgesetzt werden und derivationalle Ambiguitäten entstehen, die Meta-

Optimierungsverfahren verlangen.

Es scheint mir daher die Konklusion vertretbar, daß das Phänomen der Optionalität nach wie vor ein Problem für die optimalitätstheoretische Syntax darstellt: Idealerweise würde man so etwas in einer Theorie, in der nur der beste Kandidat gut genug ist, nicht erwarten. Die Situation ist jedoch keineswegs hoffnungslos. Die Interferenzprobleme bei echter Optionalität und lokalen Kopplungen könnten sich möglicherweise als nicht so ernst wie vermutet herausstellen, wenn man einmal größere optimalitätstheoretische Fragmente entwickelt. Die Komplexitätsprobleme bei globalen Kopplungen und bei Neutralisierung könnten sich letztlich als irrelevant oder zumindest weniger schwerwiegend erweisen, wenn die Forschung in diesem Bereich vorangeschritten ist und wir mehr über diese Thematik wissen. Ebenfalls könnte es sein, daß der Pseudo-Optionalitätsansatz dort, wo Optionalität nicht eine Instanz einer Alternation ist, das probate Mittel ist. Zum gegenwärtigen Zeitpunkt ist es für den praktizierenden optimalitätstheoretischen Syntaktiker daher wohl das Beste, sich nicht auf ein einziges Mittel zur Behandlung von Optionalität festzulegen, sondern jeweils das für das untersuchte Phänomen am geeignetsten erscheinende Mittel zu wählen. Dies kann dann zur Kombination von verschiedenen Methoden in einer einzigen Analyse führen; und genau dies passiert auch in einem Teil der in diesem Kapitel zitierten Literatur.

Kapitel 6

Wortstellung

1. Einleitung

Ich habe bereits mehrfach erwähnt, daß sich viele Wortstellungsphänomene zunächst einmal für eine optimalitätstheoretische Herangehensweise anzubieten scheinen, weil in diesem Bereich Konzepte wie Verletzbarkeit und Gewichtetheit von Beschränkungen schon traditionell als relevant betrachtet worden sind. Tatsächlich existieren bereits einige optimalitätstheoretische Arbeiten zur Wortstellungsvariation. Samek-Lodovici (1996) ist eine ausführliche Studie über die variable Position von Subjekten in verschiedenen Sprachen (darunter Italienisch, Englisch, Arabisch und die afro-asiatische Sprache Kanakuru), in Abhängigkeit von der relativen Ordnung von Beschränkungen, die sich auf strukturelle (Subjektposition, Kasus) sowie diskursbasierte (Fokus, Topik) Begriffe stützen und die ein Subjekt sozusagen in verschiedene Richtungen zerren wollen (die Beschränkungen heißen u.a. SUBJEKT (“Subject”), KASUS-ADJAZENZ (“Case Adjacency”), TOPIK-TILGUNG (“Drop Topic”) und FOKUS-AUSRICHTUNG_{links/rechts} (“Align Focus_{left/right}”). In eine ähnliche Richtung gehen die Arbeiten von Costa (1998; 1999): Auf der Basis romanischer Sprachen wie Portugiesisch, Spanisch und Italienisch, sowie u.a. des Arabischen, Griechischen, Malagasy und Irischen wird ein System entwickelt, in dem die Argumente des Verbs wiederum gemäß der Ordnung von Beschränkungen plaziert werden, die auf strukturellen wie diskursbasierten Begriffen beruhen, und die selbsterklärende Namen tragen wie SUBJEKT-KASUS (“Subject-Case”), OBJEKT-KASUS (“Object-Case”), TOPIK-ZUERST (“Topic-First”) und FOKUS-AUSRICHTUNG (“Align-Focus”). Vergleichbare Studien sind die von Legendre (1998) über stilistische Inversion im Französischen (die auf der Interaktion von strukturellen Beschränkungen wie EPP (Erweitertes Projektionsprinzip, “Extended Projection Principle”), AKKUSATIVREALISIERUNG (“Accusative”), NOMINATIVREALISIERUNG (“Nominative”), OP-SPEZ (“Op-Spec”), und der diskursbasierten Beschränkung FOKUS-AUSRICHTUNG (“AlignFocus”) beruht); sowie die von Gutiérrez-Bravo (2000) über Subjektinversion im Spanischen (hier spielen neben strukturellen Beschränkungen wie SUBJEKT (“Subject”) und diskursbasierten Beschränkungen wie FOKUS (“Focus”) und GEGEBENHEIT (“GIVENness”) auch noch intonationsbezogene Beschränkungen wie die klassische “Nuclear Stress Rule” eine Rolle).

Eine Gemeinsamkeit all dieser Arbeiten ist, daß Wortstellungsfreiheit innerhalb einer Sprache als scheinbar betrachtet wird (d.h., daß hier Pseudo-Optionalität vorliegt): Abhängig von den Erfordernissen diskursbasierter Beschränkungen ist (im

Idealfall) eine gegebene Abfolge die einzig optimale. Dies sei kurz anhand von Costas (1998; 1999) Analyse der variablen Subjektstellung im (europäischen) Portugiesischen illustriert. Die Sätze in (1-ab) sind zunächst einmal beide wohlgeformt; das Subjekt kann entweder an erster Stelle stehen (SVO) oder dem Verb folgen (VSO).

- (1) a. $[_{IP} O \text{ Paulo}_1 \text{ comeu}_2 [_{VP} t_1 t_2 a \text{ sopa }]]$
 Paulo aß die Suppe
- b. $[_{IP} - \text{ Comeu}_2 [_{VP} o \text{ Paulo}_1 t_2 a \text{ sopa }]]$
 aß Paulo die Suppe

Costa nimmt an, daß die SVO-Stellung Bewegung des Subjekts nach SpecI signalisiert, die VSO-Stellung ein Verharren des Subjekts in der VP. Vorausgesetzt ist weiterhin, daß hochgeordnete, unabhängige Beschränkungen im Portugiesischen immer die Bewegung des finiten Verbs nach I verlangen. Wichtig ist nun die Ordnung der Beschränkungen SUBJEKT-KASUS und FOKUSAUSRICHTUNG. Die erste Beschränkung verlangt, daß Subjekte in SpecI stehen, weil sie nur so Nominativ erhalten können; die zweite Beschränkung ist so formuliert, daß sie ein fokussiertes Subjekt zwar innerhalb der VP, aber nicht in SpecI erfüllen kann. Somit ist bei einer Ordnung FOKUSAUSRICHTUNG \gg SUBJEKT-KASUS im Portugiesischen allein (1-a) optimal, wenn das Subjekt nicht im Input als Fokus markiert ist, während allein (1-b) optimal ist, wenn das Subjekt im Input als Fokus markiert ist.

Im folgenden möchte ich mich auf die Darstellung von einigen Arbeiten beschränken, die die Wortstellung von nicht-pronominalen NPs im deutschen Mittelfeld behandeln; manche dieser Arbeiten sind den soeben angeführten im Prinzip ähnlich, andere nicht. Das Thema selbst ist in den vorangehenden Kapiteln bereits einige Male aufgetaucht. In Kapitel 2 habe ich darauf verwiesen, daß die Tatsache, daß die häufig zur Behandlung der freien Wortstellung im Deutschen vorgeschlagenen Beschränkungen miteinander unverträgliche Anforderungen stellen, zunächst einmal als Argument für eine optimalitätstheoretische Analyse angesehen werden kann. In Kapitel 4 wurde als Vorläufer einer optimalitätstheoretischen Analyse der Ansatz von Uszkoreit (1984; 1986) zur Wortstellung im Deutschen vorgestellt; dieser Ansatz beruht wesentlich auf der Idee verletzbarer Beschränkungen. In Kapitel 5 schließlich wurde bei der Einführung global disjunktiver Kopplungen darauf hingewiesen, daß sie möglicherweise im Bereich der deutschen Wortstellung Verwendung finden können; ich werde darauf unten noch zurückkommen.

In diesem Kapitel soll nun eine systematischere Darstellung der theoretischen Optionen im Rahmen der Optimalitätstheorie erfolgen. Die zu besprechenden Ansätze sind Choi (1996; 1999), Buring (1997; 1999), Müller (1999c) und Heck (2000). Der Hintergrund all dieser Arbeiten ist, daß in prä-optimalitätstheoretischen Analysen häufig beobachtet worden ist, daß sich die variable Wortstellung im deutschen Mittelfeld aus dem Zusammenspiel verschiedener Faktoren ergibt; zu diesen Faktoren gehören Definitheit, Belebtheit, Fokus, Kasus, usw. Weiterhin ist oft vorgeschlagen worden, daß diese Faktoren für sich allein genommen "schwach" sind,

also nicht per se über Grammatikalität oder Ungrammatikalität eines Satzes entscheiden, und daß sie unterschiedliche Wichtigkeit aufweisen. Eine (unvollständige) Liste der einschlägigen Literatur umfaßt Lenerz (1977, 27 & 62f), Hoberg (1981, 62), Lötscher (1981, 58-59), Uszkoreit (1984, 174ff; 1986, 896-899), Lerot (1985, 141), Abraham (1986, 27ff; 1995, 565-649), Reis (1986, 27ff), Jacobs (1988, 17ff), Stechow & Sternefeld (1988, 455), Siewierska (1993, 830ff), Dietrich (1994, 41), und Primus (1994, 40-48). Auf den ersten Blick sieht es so aus, als seien "Schwäche" und "unterschiedliche Wichtigkeit" von Faktoren direkt als Verletzbarkeit und Geordnetheit von Beschränkungen interpretierbar und eine optimalitätstheoretische Rekonstruktion der in den soeben zitierten Arbeiten gewonnenen empirischen Erkenntnisse somit eine recht einfache Angelegenheit. Wir werden jedoch sehen, daß dem nicht so ist.

2. Ein Pseudo-Optionalitäts-Ansatz: Choi (1996; 1999)

2.1. Die Analyse

Chois (1996; 1999) Analyse der Wortstellung im deutschen Mittelfeld gehört in eine Klasse mit den oben erwähnten Arbeiten zu anderen Sprachen, die von einer Interaktion von diskursbasierten und grammatisch-strukturellen Beschränkungen ausgehen und im Idealfall für jede im Input vorgegebene Diskursspezifikation nur einen optimalen Kandidaten ableiten. Die Analyse fußt auf den folgenden Beschränkungen.

- (2) a. KN₁ ("Kanonische Abfolge₁", CN₁):
SUBJ ist strukturell prominenter als (c-kommandiert) Nicht-SUBJ-Funktionen.
- b. KN₂ ("Kanonische Abfolge₂", CN₂):
Nicht-SUBJ-Funktionen richten sich auf der Oberfläche umgekehrt zur folgenden funktionalen Hierarchie aus:
SUBJ > D.OBJ > I.OBJ > OBL > ADJUNKT
- c. NEU ("New"):
[-Neu] geht [+Neu] voran.
- d. PROM ("Prominence"):
[+Prom] geht [-Prom] voran.

Chois Analyse ist im Rahmen der Lexikalisch-Funktionalen Grammatik (LFG) abgefaßt, derzufolge grammatische Funktionen wie SUBJ, D(irektes).OBJ, I(indirektes).OBJ nicht Positionen im Baum sind, sondern Primitive der Theorie, die – in einer um die Optimalitätstheorie erweiterten Variante der LFG – als Informationen im Input kodiert sind. KN₁ besagt dann, daß ein Subjekt auf der syntaktischen Oberfläche prominenter als alle anderen grammatischen Funktionen sein sollte; für Sprachen wie das Deutsche gilt, daß Prominenz immer asymmetrisches C-Kommando impliziert, und dies wiederum aufgrund der Struktur des

deutschen Satzes lineare Präzedenz. Die für uns hier allein wichtige Konsequenz aus KN_2 ist, daß ein indirektes Objekt einem direkten Objekt vorangehen sollte. Diese zwei grammatisch-strukturellen Beschränkungen sagen zunächst einmal vorher, daß die Wortstellung im prototypischen deutschen Satz Subjekt vor indirektem Objekt vor direktem Objekt sein sollte, wie in (3-a). Steht dagegen z.B. das direkte Objekt zwischen Subjekt und indirektem Objekt, wird KN_2 verletzt, steht das indirekte Objekt vor dem Subjekt, wird KN_1 verletzt; vgl. (3-bc). Die folgenden Sätze akkumulieren noch mehr Verletzungen dieser beiden Beschränkungen.¹

- (3) a. daß der Fritz der Maria das Buch gegeben hat
 b. daß der Fritz das Buch der Maria gegeben hat
 c. daß der Maria der Fritz das Buch gegeben hat
 d. daß der Maria das Buch der Fritz gegeben hat
 e. daß das Buch der Fritz der Maria gegeben hat
 f. daß das Buch der Maria der Fritz gegeben hat

Die Abfolgen in (3-b-f) sind nun nicht per se unmöglich. Choi argumentiert jedoch, daß sie möglich (also optimal) werden nur dann, wenn diskursbasierte Beschränkungen wie NEU und PROM ins Spiel kommen, die die von KN_1 und KN_2 abweichenden Abfolgen erzwingen. Zunächst bedürfen die Merkmale $[\pm Neu]$ und $[\pm Prom]$ einer Erläuterung. Mit Hilfe dieser Merkmale können durch Kreuzklassifikation die Begriffe Topik ($[+Prom], [-Neu]$), komplettierender Fokus ($[-Prom], [+Neu]$), kontrastiver Fokus ($[+Prom], [+Neu]$) und Ende ("tail") ($[-Prom], [-Neu]$) definiert werden. Ein komplettierender (quasi "regulärer") Fokus wird also betrachtet als etwas, das gleichzeitig noch nicht im Diskurs prominent und neue Information ist; ein Topik ist etwas, das prominent und nicht neue Information ist; ein kontrastiver Fokus ist etwas, das sowohl prominent, als auch neue Information ist; und ein Ende ist nichts von alledem.² Aus NEU folgt somit, daß ein Argument, das ein komplettierender Fokus ist, möglichst weit rechts steht; aus PROM folgt umgekehrt, daß ein Argument, das ein Topik ist, möglichst den Nicht-Topiks vorangeht. Interessant ist der Fall des kontrastiven Fokus: Als $[+Prom]$ -markiertes Argument wird er von PROM nach links gezwungen, als $[+Neu]$ -markiertes Argument wird er von NEU nach rechts gedrängt. Hier erweist sich die relative Ordnung der beiden diskursbasierten Beschränkungen als wesentlich. Die von Choi (1999) diskutierten zwei möglichen Ordnungen der vier Beschränkungen im Deutschen sind in (4) aufgelistet; der einzige Unterschied betrifft die Frage, ob NEU und KN_2 gekoppelt sind,

¹Für die meisten Fälle kann man Chois Beschränkungen über die Anordnung grammatischer Funktionen unmittelbar verstehen als Beschränkungen über die Anordnung von Kasus: Nominativ vor Dativ vor Akkusativ. Die Unterschiede zwischen LFG und anderen syntaktischen Theorien, die grammatische Funktionen nicht als gegeben, sondern als ableitbar ansehen, sind im vorliegenden Zusammenhang nicht der Rede wert.

²Diese Einteilung geht auf Vallduví (1992) zurück.

oder ob die Ordnung in ein Dominanzverhältnis aufgelöst wird.

- (4) a. PROM \gg KN₁ \gg NEU \circ KN₂
 b. PROM \gg KN₁ \gg NEU \gg KN₂

Betrachten wir dieses System bei der Arbeit. Angenommen sei zuerst, daß keines der drei Argumente in (3) als Topik oder (komplettierender bzw. kontrastiver) Fokus zu interpretieren ist. Dann sind bei allen Abfolgen die diskursbasierten Beschränkungen PROM und NEU leer erfüllt, und die Entscheidung über den optimalen Kandidaten fällt allein den Beschränkungen KN₁ und KN₂ zu, die wie gesehen die “normale”, “unmarkierte” Abfolge Subjekt vor indirektem Objekt vor direktem Objekt wählen; alle anderen Abfolgen sind als suboptimal blockiert.³ Dies zeigt Tabelle T_{6.1}.⁴

T_{6.1}: Normalabfolge: Irrelevanz von diskursbasierten Beschränkungen bei Choi

Kandidaten	PROM	KN ₁	NEU KN ₂
☞ K ₁ : Subj IO DO			
K ₂ : Subj DO IO			*!
K ₃ : IO Subj DO		*!	
K ₄ : IO DO Subj		*!*	
K ₅ : DO Subj IO		*!	*
K ₆ : DO IO Subj		*!*	*

Hier entsprechen die Kandidaten K₁ bis K₆ der Reihe nach den Sätzen in (3). Da im Prinzip aber alle Abfolgen im Deutschen möglich sind, besteht die Aufgabe nun darin, zu zeigen, daß jede Abfolge unter bestimmten Zuordnungen von Merkmalen wie [\pm Prom],[\pm Neu] zu Argumenten des Inputs in Optimalität resultiert. Choi nimmt z.B. an, daß Kandidat K₅ einzig optimal wird, wenn das direkte Objekt im Input als Topik ausgezeichnet ist, also als [+Prom],[–Neu]. Denn dann zwingt die höchst-geordnete Beschränkung PROM das direkte Objekt vor alle anderen, nicht-prominenten Argumente; die dadurch hervorgerufenen Verletzungen von KN₁ und KN₂ werden nicht-fatal. Dies zeigt Tabelle T_{6.2}. Die Vorhersage ist somit, daß ein Satz wie (3-e) dann möglich ist, wenn das direkte Objekt als Topik interpretiert wird.

Eine weitere Frage ist, wie die Variation in der Abfolge von direktem und indirektem Objekt in (3-a) und (3-b) zu erklären ist. Choi rekuriert hier auf die

³Der Begriff der “Unmarkiertheit” von Sätzen, wie er hier und im Rest des Kapitels verwendet wird, hat nichts mit dem Konzept des “unmarkierten Falls” zu tun, wie es in den vorangegangenen Kapiteln benutzt wurde. Bei letzterem geht es um Eigenschaften von Kandidaten, die durch niedrig geordnete Beschränkungen vorhergesagt werden, die nur greifen, wenn alle höheren Beschränkungen irrelevant sind; bei ersterem geht es dagegen um die empirische Bewertung der Güte eines Kandidaten.

⁴Ich habe hier die Beziehung von NEU und KN₂ als Kopplung notiert; aber siehe unten.

T_{6.2}: Andere Abfolgen: Relevanz von diskursbasierten Beschränkungen bei Choi

Kandidaten	PROM	KN ₁	NEU KN ₂
K ₁ : Subj IO DO	*!*		
K ₂ : Subj DO IO	*!		*
K ₃ : IO Subj DO	*!*	*	
K ₄ : IO DO Subj	*!	**	
☞K ₅ : DO Subj IO		*	*
K ₆ : DO IO Subj		**!	*

unterschiedlichen Erfordernisse von NEU und KN₂ und rekonstruiert damit im wesentlichen die Analyse von Lenerz (1977). Wie gesehen ist indirektes Objekt vor direktem Objekt die Abfolge, die optimal wird, wenn keine diskursbasierten Beschränkungen für die Objekte eine Rolle spielen. Angenommen sei nun aber, daß eines der beiden Objekte ein komplettierender Fokus ist und somit NEU unterliegt. Dann gilt bei Lenerz (und, wie wir in Kapitel 4 gesehen haben, auch bei Uszkoreit): Ist das indirekte Objekt der Fokus, kann es entweder (gemäß KN₂) dem direkten Objekt vorausgehen oder ihm (gemäß NEU) folgen; vgl. (5-ab). Ist dagegen das direkte Objekt der Fokus, kann es dem indirekten Objekt nicht gut vorausgehen; vgl. (5-cd).⁵

- (5) a. daß der Fritz der MARIA das Buch gegeben hat
 b. daß der Fritz das Buch der MARIA gegeben hat
 c. daß der Fritz der Maria das BUCH gegeben hat
 d.?*daß der Fritz das BUCH der Maria gegeben hat

Die Unmöglichkeit von (5-d) folgt wie in Tabelle T_{6.3} gezeigt daraus, daß der entsprechende Kandidat K₂ von K₁ (das (5-c) entspricht) harmonisch begrenzt wird – K₂ tut nichts, was K₁ nicht besser täte.⁶

T_{6.3}: Abfolge bei komplettierendem Fokus auf DO bei Choi

Kandidaten	PROM	KN ₁	NEU KN ₂
☞K ₁ : Subj IO DO			
K ₂ : Subj DO IO			*! *

Interessant ist nun die Ableitung der Optionalität in (5-ab). Sind NEU und KN₂ gekoppelt, erscheinen K₁ und K₂ in Tabelle T_{6.4} beide als optimal (unter allen Standardkonzepten von Kopplungen): Wo K₁ bei KN₂ besser abschneidet, gewinnt K₂ bei NEU durch Verlagerung des fokussierten Arguments nach rechts.

⁵Komplettierender Fokus ist wie vorher durch Kapitälchen signalisiert.

⁶Weitere suboptimale Kandidaten, die ein Objekt oder beide Objekte vor das Subjekt stellen, sind hier der Einfachheit halber weggelassen.

T_{6.4}: Abfolge bei komplettierendem Fokus auf IO bei Choi

Kandidaten	PROM	KN ₁	NEU KN ₂
☞ K ₁ : Subj IO DO			*
☞ K ₂ : Subj DO IO			*

Dies ist die Erklärung von Choi (1996), die auch in Choi (1999) noch erwogen wird. Letztendlich entscheidet sich jedoch Choi (1999) gegen diese Kopplungsanalyse der Optionalität und versucht, auch bei (5-ab) noch eine unterschiedliche Diskurspezifikation ausfindig zu machen; diese sei “subtil”, aber ausreichend dafür, daß unter der strikten Ordnung in (4-b) (anstatt von (4-a)) die beiden Sätze in (5-ab) jeweils Gewinner separater Wettbewerbe werden.

2.2. Diskussion

Dies mag als Illustration des Modells von Choi (1996; 1999) hier ausreichen. Der zuletzt angesprochene Punkt führt bereits zu einer interessanten Eigenschaft des Systems, die unter Umständen problematisch werden kann: Man ist verpflichtet, für jede auch noch so geringfügige Wortstellungsvariation einen speziellen diskursbezogenen Grund beizuschaffen. Dies ist nicht immer einfach. Ist man bereit, Kopplungen anzunehmen, läßt sich diese Schwierigkeit etwas eingrenzen. Aber man kann wohl argumentieren, daß eine möglichst weitreichende und restriktive Theorie der Wortstellungsvariation, in der diese Variation auf diskursbezogene Faktoren zurückgeführt wird und nicht auf grammatik-interne Zufälligkeiten, idealerweise auf Kopplungen in diesem Bereich ganz verzichten sollte. Tatsächlich scheint mir der verzeichnete Wandel von Choi (1996) zu Choi (1999) genau diesem Bestreben zu entsprechen.

Als weiteres Beispiel für die Schwierigkeit, für jede Abfolge einen diskursbezogenen Grund auszumachen, sei nochmals der Kontrast in (5-cd) betrachtet. (5-d), hier wiederholt als (6-a), ist sicherlich schlechter oder “weniger normal” als (5-c), aber es ist oft festgestellt worden, daß der Satz eigentlich nicht ungrammatisch im klassischen Sinne ist. Uszkoreit (1986) schreibt, daß ein Satz dieses Typs “immer noch akzeptabler erscheint als Sätze, in denen Beschränkungen verletzt sind, die strikt eine Abfolge erzwingen” (vgl. auch Abschnitt 2 von Kapitel 4); und Buring (1999) stellt fest, daß ein Satz wie “[6-a)] merkwürdig ist, aber immer noch viel besser als [(6-b)], und daß [(6-a)] daher mit Fragezeichen versehen, aber nicht gesternt werden sollte.”

- (6) a. ?*daß der Fritz das BUCH der Maria gegeben hat
 b. *daß der Fritz der Maria gegeben das BUCH hat

Choi argumentiert nun, daß (6-a) tatsächlich ein wohlgeformter Satz ist, aber nicht mit komplettierendem Fokus auf dem direkten Objekt, sondern mit kontrastivem Fokus, wie in (7).

(7) daß der Fritz das BUCH der Maria gegeben hat (und nicht die ZEITUNG)

Der einzige Unterschied zwischen komplettierendem und kontrastivem Fokus ist in Chois Analyse die Auszeichnung eines Elementes als [-Prom] bzw. [+Prom]. Kontrastiver Fokus unterliegt neben NEU auch noch der höher geordneten Beschränkung PROM, die ihn vor alle [-Prom]-markierten Elemente drängt. Wenn wir nun noch bereit sind, anzunehmen, daß das Subjekt in (7) selbst im Input als Topik markiert ist (so daß keine Voranstellung des kontrastiven Fokus vor das Subjekt erzwungen wird, anders als beim oben diskutierten Beispiel (3-e) mit dem direkten Objekt als Topik), dann erscheint (7) tatsächlich als Gewinner eines Wettbewerbs; vgl. Tabelle T_{6.5}.

T_{6.5}: Abfolge bei kontrastivem Fokus auf DO bei Choi

Kandidaten	PROM	KN ₁	NEU KN ₂
K ₁ : Subj IO DO	*!		
☞K ₂ : Subj DO IO			* *

Wie ein Vergleich mit Tabelle T_{6.3} zeigt, behält K₂ seine Verletzungen von NEU und KN₂; weil jedoch nunmehr PROM relevant wird, spielen diese Verletzungen keine Rolle mehr, und K₁ wird als suboptimal blockiert.

Chois Lösung des durch den Status von (6-a) erzeugten Dilemmas ist nicht unproblematisch. Zum einen scheint klar, daß (6-a) zwar mit kontrastivem Fokus noch besser wird, aber eben auch mit "regulärem", komplettierendem Fokus nicht unmöglich ist. Dies bleibt unerklärt. Zum anderen kann kein Zweifel daran bestehen, daß auch bei kontrastivem Fokus auf dem direkten Objekt die gegenüber (7) umgetauschte Stellung in (8) möglich (und für die meisten Sprecher wohl sogar präferiert) ist; auch dies ist aber unerwartet, weil (8) K₁ entspricht und somit ein suboptimales Beschränkungsprofil aufweist.

(8) daß der Fritz der Maria das BUCH gegeben hat (und nicht die ZEITUNG)

Choi (1996; 1999) nimmt dieses Problem partiell zur Kenntnis. Sie versucht jedoch nicht mehr, eine grammatik-interne Lösung zu entwickeln. Verwiesen wird vielmehr auf Unterschiede zwischen Sprechern bezüglich der Kontextperzeption. Warum (7) und (8) im selben Kontext als Alternativen koexistieren können, bleibt damit ein Geheimnis. Festzuhalten ist somit, daß die Aufgabe, für jeden Satz eine Diskursspezifikation zu finden, unter der er einzig optimal ist, keineswegs trivial ist; möglicherweise gibt es eben doch Sätze, die zwar keine perfekten Realisierungen für irgendeine Diskursspezifikation sind, die aber nichtsdestoweniger im Prinzip grammatisch (wenn auch stärker markiert) sind.

Dies führt zum letzten Punkt. Betrachtet man die Sätze in (3), so ist klar, daß zwar alle Abfolgen im Prinzip möglich (d.h., grammatisch) sind, daß sie aber nicht denselben Status haben; manche Abfolgen sind "markierter" oder "weniger normal" als andere. So ist (3-a) die "Normalabfolge"; (3-b) ist demgegenüber markierter;

und (3-f) ist intuitiv noch stärker markiert. Es stellt sich die Frage, ob die Syntax zusätzlich zu Entscheidungen über die Grammatikalität von Sätzen auch derartige relative Markiertheitsgrade voraussagen können soll bzw. muß. Angesichts von Tabelle T_{6.1} sieht es zunächst einmal fast so aus, als könne Chois System dies leisten: Die relative Markiertheit der Sätze in (3) scheint mit der jeweiligen Güte des Beschränkungsprofils zu korrelieren. Das Problem ist jedoch, daß K₂ bis K₆ in T_{6.1} als suboptimal und somit als vollkommen ungrammatisch klassifiziert werden, und daß daher ein offensichtliches Grammatik-internes Maß für Markiertheitsgrade nicht zur Verfügung steht. Warum also Sprecher üblicherweise Intuitionen darüber haben, welche von zwei Abfolgen im Mittelfeld (mit jeweilig fixem Intonationsmuster) die weniger markierte, "bessere" ist, wird bei Choi von der Grammatik nicht erklärt: Alles, was überhaupt möglich ist, ist auch eine perfekte Realisierung vis-à-vis einer gegebenen diskursbezogenen Input-Spezifikation.

3. Ein anderer Pseudo-Optionalitäts-Ansatz: Büring (1996; 1997; 1999)

3.1. Analyse

Unabhängig hat Büring (1997; 1999, z.T. auch 1996) eine Analyse der freien Wortstellung im Deutschen entwickelt, die der von Choi in mancherlei Hinsicht vergleichbar ist, und die ebenfalls ganz wesentlich auf der Arbeit von Lenerz (1977) beruht. Die in Büring (1997) wichtigen Beschränkungen sind in (9) aufgelistet.

- (9) a. IND ("Indefinita", "Indefinites"):
Existentiell interpretierte Indefinita müssen vollständig in der VP enthalten sein.
- b. ENDFOK ("EndFokus", "Final Focus", FF):
Der Fokus ist satzfinal.
- c. ÖKON ("Ökonomie", "Stay"):
Bewegung ist verboten.

Bürings Grundannahme ist, daß es eine fixe Basisabfolge Subjekt \succ indirektes Objekt \succ direktes Objekt in der VP gibt, die nur durch Bewegung (Scrambling) z.B. in VP-Adjunktionspositionen geändert werden kann. Derartiges Scrambling verletzt immer die Ökonomiebeschränkung ÖKON. Da an VP adjungierte Elemente nicht mehr vollständig in der VP enthalten sind, blockiert auch IND Scrambling von gewissen Elementen, nämlich von solchen Indefinita, die existenziell (also nicht generisch) interpretiert werden. Ein potentieller Auslöser für Scrambling ist dagegen ENDFOK – ein direktes Objekt kann über ein als fokussiert markiertes indirektes Objekt gescrambelt werden, um letzteres näher an das Satzende zu bringen.⁷ Die

⁷Eine Unterteilung in komplettierenden und kontrastiven Fokus nimmt Büring nicht vor; tatsächlich nimmt er an (vgl. Büring (1999)), daß es keinen empirisch manifesten Unterschied zwischen den beiden Fokustypen gibt. Von dieser Distinktion wird daher im folgenden abgesehen.

für das Deutsche postulierte Ordnung ist (10).

(10) $IND \gg \ddot{O}KON \circ ENDFOK$

Hier ist IND undominiert, also von keinem wohlgeformten Satz verletzbar. $\ddot{O}KON$ und $ENDFOK$ sind gekoppelt; der präsupponierte Kopplungsbegriff ist der einer global hierarchischen Kopplung (für die im folgenden betrachteten Fälle würden allerdings auch lokale Kopplungskonzepte ausreichen). Es nimmt nicht wunder, daß die Kopplung von $\ddot{O}KON$ und $ENDFOK$ bei Büring (1997) im wesentlichen dieselbe Funktion hat wie die Kopplung von KN_2 und NEU bei Choi (1996) (vgl. (4-a)): $ENDFOK$ entspricht NEU unmittelbar, und $\ddot{O}KON$ entspricht KN_2 indirekt, insofern, als beide Beschränkungen auf der syntaktischen Oberfläche die Abfolge indirektes Objekt vor direktem Objekt verlangen. Tatsächlich funktioniert Bürings (1997) Analyse der Daten in (5) (hier in (11) wiederholt) ganz ähnlich wie Chois Modell.

- (11) a. daß der Fritz der MARIA das Buch gegeben hat
 b. daß der Fritz das Buch der MARIA gegeben hat
 c. daß der Fritz der Maria das BUCH gegeben hat
 d.*daß der Fritz das BUCH der Maria gegeben hat

Betrachten wir zunächst wiederum den (11-cd) zugrundeliegenden Wettbewerb. In Tabelle $T_{6.6}$ verletzt K_2 (= (11-d)) sowohl $ENDFOK$ (weil das fokusmarkierte direkte Objekt links vom indirekten Objekt steht), als auch $\ddot{O}KON$ (weil das direkte Objekt vor das indirekte Objekt gescrambelt worden ist). K_1 (= (11-c)) verletzt keine dieser Beschränkungen. Da IND nicht aktiv ist (alle Argumente sind definit), ist der Fall klar.

$T_{6.6}$: Abfolge bei Fokus auf DO bei Büring

Kandidaten	IND_1	$\ddot{O}KON$ $ENDFOK$
$\Leftarrow K_1$: Subj IO_1 DO_2		
K_2 : Subj DO_2 IO_1 t_2		*! *

Die in (11-ab) vorliegende Optionalität wird durch die Kopplung abgeleitet. K_1 in Tabelle $T_{6.7}$ (= (11-a)) verletzt $ENDFOK$, erfüllt aber dafür $\ddot{O}KON$, und K_2 (= (11-b)) verletzt zwar $\ddot{O}KON$, schneidet jedoch bei $ENDFOK$ besser ab.

$T_{6.7}$: Abfolge bei Fokus auf IO bei Büring

Kandidaten	IND_1	$\ddot{O}KON$ $ENDFOK$
$\Leftarrow K_1$: Subj IO_1 DO_2		*
$\Leftarrow K_2$: Subj DO_2 IO_1 t_2		*

Die Relevanz von IND zeigt sich bei der Einbeziehung von Daten mit indefiniten Argumenten. Wie Lenerz (1977) festgestellt hat, stellt sich der soeben behandel-

te Optionalitätseffekt nicht ein, wenn zur Erfüllung von ENDFOK ein existentiell interpretiertes Indefinitum gescrambelt werden muß. Man betrachte die folgenden Beispiele.⁸

- (12) a. daß der Fritz der MARIA ein Buch geschenkt hat
 b.?*daß der Fritz ein Buch der MARIA geschenkt hat
 c. daß der Fritz einer Frau das BUCH geschenkt hat
 d.?*daß der Fritz das BUCH einer Frau geschenkt hat

Hier wird (12-d) durch (12-c) blockiert; die Blockade ist vollständig analog zum Wettbewerb in Tabelle T_{6.6} (IND ist von beiden Kandidaten erfüllt). Anders liegt der Fall bei (12-ab). (12-a) verletzt ENDFOK und erfüllt ÖKON, (12-b) verletzt ÖKON und erfüllt ENDFOK. Diese END-FOK-Erfüllung ist aber hier zu teuer erkaufte, denn zusätzlich zu ÖKON verletzt (12-b) durch Scrambling des existentiell interpretierten Indefinitums auch noch fatal IND.⁹ Dies zeigt Tabelle T_{6.8}.

T_{6.8}: Abfolge bei Fokus auf IO und existentiell DO-Indefinitum bei Büring

Kandidaten	IND ₁	ÖKON ENDFOK
☞K ₁ : Subj IO ₁ DO ₂		*
K ₂ : Subj DO ₂ IO ₁ t ₂	*!	*

Wird ein Indefinitum nicht als existentiell, sondern als generisch interpretiert, unterliegt es nicht mehr IND, und die Vorhersage ist, daß Optionalität im selben Kontext wieder möglich ist. In (12-ab) ist nun eine generische Interpretation des Indefinitums vielleicht nicht so naheliegend. Etwas anders liegt der Fall beim bloßen Plural in (13). Hier sind bei fokussiertem indirektem Objekt und generisch interpretiertem direktem Objekt beide Abfolgen möglich; d.h., beide Sätze können so etwas bedeuten wie: “Wenn etwas ein Buch ist, dann schenkt der Fritz es üblicherweise der MARIA”.¹⁰

- (13) a. daß der Fritz der MARIA Bücher schenkt
 b. daß der Fritz Bücher der MARIA schenkt

Eine eigene Tabelle, die diesen Wettbewerb illustriert, ist an dieser Stelle nicht notwendig, denn die Argumentation läuft wie in Tabelle T_{6.7} für definite direkte Objekte angegeben.

⁸Choi (1996; 1999) diskutiert ebenfalls derartige Daten.

⁹IND steht zu ÖKON als speziell zu generell in einer Stringenzrelation; vgl. (65) aus Kapitel 2.

¹⁰Diese Dateneinschätzung ist nicht vollständig vereinbar mit dem, was in Diesing (1992) über die Verteilung von generischen und existentiellen Lesarten von Indefinita festgehalten ist. Ausführlich wird das Thema aus empirischer Perspektive behandelt in Büring (1996). Vgl. auch Lenerz (1999) zu weiterer Diskussion.

In Büring (1999) wird diese Analyse zwar in verschiedener Hinsicht weiterentwickelt, im Kern jedoch beibehalten. Erstens tritt an die Stelle von ÖKON die Beschränkung DATIV (“Dativ-NPs gehen Akkusativ-NPs voran”). DATIV hat für die hier betrachteten Daten dieselben Konsequenzen (insofern, als diese Beschränkung Stellungen mit direktem Objekt vor indirektem Objekt bestraft). Diese Beschränkung setzt aber nicht mehr voraus, daß es eine uniforme Basisabfolge mit indirektem Objekt vor direktem Objekt gibt (siehe dazu unten), und sie beruht auch nicht mehr notwendigerweise auf der Existenz einer Scrambling-Operation als syntaktischer Bewegung. Zweitens, und das ist eine weit weniger triviale Änderung, dekonstruiert Büring (1999) die Beschränkung ENDFOK in ein geordnetes System von unabhängig in der prosodischen Phonologie motivierten Beschränkungen, die die Haupteffekte von ENDFOK als Epiphänomene abzuleiten gestatten. Unter dieser Sichtweise ist Fokus nach wie vor im Input markiert; aber die phonologische Realisierung einer Abfolge mit einer gegebenen Fokusmarkierung ergibt sich aus der Interaktion prosodischer Beschränkungen (die natürlich partiell auf syntaktische Fokusmarkierung Bezug nehmen). Erhalten bleibt in diesem Modell die Notwendigkeit der Kopplung: DATIV ist wie vorher ÖKON gekoppelt, allerdings nicht mehr mit einer einfachen Beschränkung, sondern mit sämtlichen relevanten prosodischen Beschränkungen mit intern fixer Ordnung.¹¹ Die Kopplung wird unter global hierarchischer Interpretation aufgelöst in zwei Einzelordnungen, in der einmal DATIV höher geordnet ist als die damit gekoppelten prosodischen Beschränkungen, und einmal alle prosodischen Beschränkungen höher geordnet sind als DATIV.

3.2. Diskussion

Bei allen Unterschieden im Detail teilt Bürings Analyse wesentliche Merkmale mit Chois Analyse, und es erstaunt daher auch nicht, daß sie tendentiell dieselben Eigenschaften aufweist, die in der Diskussion von Chois Modell als potentiell problematisch aufgezeigt worden sind. Erstens ist das System so aufgebaut, daß idealerweise einer gegebenen diskursbezogenen Inputspezifikation (bei Büring ist dies in erster Linie die Fokusmarkierung von Elementen) nur eine einzige optimale Realisierung entsprechen sollte: Optionalität in der Wortstellung im Mittelfeld wird als scheinbar erwiesen. Interessanterweise muß jedoch wie bei Choi (1996) auch hier für einen Teilbereich der Wortstellungsvariation neben der (erwünschten) Pseudo-Optionalität auch noch das (eigentlich unerwünschte) Mittel der Kopplung verwendet werden.¹²

¹¹“Sämtliche” ist allerdings nicht so schlimm, wie es sich auf den ersten Blick darstellen mag. Es handelt sich hier um zwei prosodische Beschränkungen.

¹²Daß auch Büring (1999) Kopplungen in der Endversion seines Systems als Zusatzmechanismen versteht, ergibt sich aus seiner Diskussion der Variation in den romanischen Sprachen, in denen eine Alternation von Beispielen vom Typ (11-a)/(11-b) obligatorisch zugunsten des Kandidaten mit rechts-peripherem Fokus (11-b) zusammenbricht (vgl. Gutiérrez-Bravo (2000)). Dies

Zweitens muß Bürings Analyse ebenso wie Chois Analyse die Gefahr der Unter-
generierung gewärtigen: Es ist nicht immer offensichtlich, wie für jeden grammati-
schen Satz mit Wortstellungsvariation eine diskursbezogene Input-Spezifikation als
Auslöser gefunden werden kann. Zudem scheint es angemessener, Sätze wie (11-d)
(Scrambling eines fokussierten direkten Objekts) oder (12-b) (Scrambling eines existi-
entuell interpretierten indefiniten direkten Objekts) anstatt als ungrammatisch als
grammatisch, aber stark markiert anzusehen; die beiden Sätze sind hier wiederholt.

- (14) a.?*daß der Fritz das BUCH der Maria gegeben hat
b.?*daß der Fritz ein Buch der MARIA geschenkt hat

Tatsächlich ist dies, wie bei der Diskussion von Chois Ansatz gesehen, ja auch
Bürings (1999) Meinung. Der von ihm ins Auge gefaßte Ausweg ist folgender: Per
Annahme ist ein Kandidat K_1 , der gegenüber einem optimalen Kandidaten K_2
suboptimal ist nur deshalb, weil er fatal eine Beschränkung verletzt, die zu einer
bestimmten Klasse Σ gehört, nur markiert, aber nicht ungrammatisch. Σ umfaßt
im vorliegenden Fall natürlich gerade die hier diskutierten Beschränkungen, die die
Wortstellung im deutschen Mittelfeld (genauer, in der VP) steuern, aber nicht z.B.
die Beschränkung, die V/2-Bewegung auslöst und deren fatale Verletzung im Deut-
schen zu uneingeschränkter Ungrammatikalität führen muß. Es versteht sich von
selbst, daß ein solcher Schritt eine nicht-triviale Modifikation der klassischen Op-
timalitätstheorie impliziert. Ich werde auf diesen Punkt bei der Diskussion meiner
eigenen Analyse (Müller (1999c)) noch zurückkommen.

Drittens und letztens gilt bei Büring ebenso wie bei Choi, daß (abgesehen vom
gerade behandelten Fall) Markiertheits- bzw. Akzeptabilitätsunterschiede zwischen
in ihren jeweiligen Kandidatenmengen optimalen Kandidaten mit unterschiedlichen
Wortstellungen bzw. Fokus- (Akzent-)verteilungen nicht leicht ausgedrückt werden
können. Büring lehnt dies auch ganz explizit ab: In (11) soll zwar der suboptimale
Kandidat (11-d) markierter sein als (11-a), (11-b) und (11-c); diese drei optima-
len Kandidaten sollen jedoch alle denselben Markiertheitsstatus haben; sie sind
sämtlich optimale Realisierungen einer Input-Spezifikation. Was könnte man dann
auf der Basis von Bürings (oder Chois) Analyse noch sagen, wenn es sich erwie-
se, daß Sprecher Intuitionen haben darüber, daß etwa (11-c) "weniger markiert"
bzw. "normaler" erscheint als (11-a)? Systemintern nichts, denke ich. Allerdings
gibt es eine systemexterne Möglichkeit, auf die Hans Kamp anlässlich der Prä-
sentation von Büring (1997) auf einer Tagung hingewiesen hat: Diskursbezogene
Input-Spezifikationen kodieren die möglichen Kontexte, in denen ein Satz Verwen-
dung finden kann. Wenn nun Sprecher der Meinung sind, sie beurteilten die relative
Markiertheit von Sätzen, dann, so die Idee, reflektieren sie tatsächlich nur die re-
lative Markiertheit der Kontexte, in denen die betreffenden Sätze optimal sind.

wird zunächst einmal als der zu erwartende Normalfall eingeführt, bevor für das Deutsche die
Komplikation mit Kopplungen entwickelt wird.

4. Ein erweiterter Kopplungs-Ansatz: Müller (1999c)

4.1. Die Analyse

4.1.1. Voraussetzungen

Der in Müller (1999c) entwickelte Ansatz weicht von den beiden bisher dargestellten wesentlich dadurch ab, daß nicht davon ausgegangen wird, daß die große Wortstellungsfreiheit im deutschen Mittelfeld nur scheinbar ist. Vielmehr wird eine Variante des Kopplungsbegriffs eingeführt, derzufolge die Existenz mehrerer grammatischer Kandidaten mit unterschiedlicher Wortstellung und Fokuszuweisung in einer Kandidatenmenge eher der Normalfall als die Ausnahme ist; ein Untergenerierungsproblem kann daher nicht entstehen. Zusätzlich sollen Markiertheitsunterschiede zwischen verschiedenen optimalen Kandidaten ein und desselben Wettbewerbs dadurch erfaßt werden, daß die miteinander gekoppelten Wortstellungsbeschränkungen in gewisser Weise doch eine Ordnung zugewiesen bekommen. Dies geschieht durch eine Aufspaltung des Optimalitätskonzepts: Optimalität als Grammatikalität gegenüber Optimalität als Unmarkiertheit. Wie wir sehen werden, erfordert diese Zweiteilung einige nicht-triviale Modifikationen der Optimalitätstheorie.

Eine erste Grundannahme des Modells ist, daß die D-strukturelle Abfolge im Deutschen fest ist; alle Veränderungen der Abfolge sind das Resultat einer S-strukturellen Scrambling-Operation. Anders als bei Lenerz (1977), Büring (1997) und den meisten anderen Arbeiten, die von einer festen Basisabfolge ausgehen, wird jedoch postuliert, daß D-strukturell das direkte Objekt das indirekte Objekt asymmetrisch c-kommandiert: Subj \succ direktes Objekt \succ indirektes Objekt. Evidenz für diese Annahme kommt u.a. aus dem Bereich der anaphorischen Bindung. Wie von Grewendorf (1988), Webelhuth (1992), und Müller & Sternefeld (1994) beobachtet wurde, kann ein direktes Objekt ein nachfolgendes indirektes Objekt A-bindern und so als anaphorisches Element lizensieren; ein indirektes Objekt kann demgegenüber ein nachfolgendes direktes Objekt nicht A-bindern. Dies ist anhand von Rezipropronomen für einige Verben in (15) gezeigt.¹³

- (15) a. daß man die Gäste₁ einander vorstellte
 b. *daß man den Gästen₁ einander t₁ vorstellte
 c. daß ich diese Leute₁ einander₁ von Herzen gönne
 d. *daß ich diesen Leuten₁ einander₁ t₁ von Herzen gönne
 e. daß die UNO die Feinde₁ einander₁ nicht ausliefern darf
 f. *daß die UNO den Feinden₁ einander t₁ nicht ausliefern darf

Nehmen wir nun an, daß die Prinzipien der Bindungstheorie rein strukturell sind und Bindung aus einer A-Position erfordern (vgl. Chomsky (1981)), und daß durch

¹³Reflexivpronomen verhalten sich zumindest in der Tendenz ebenso, aber es gibt hier geringfügige Variation unter Sprechern.

Scrambling – zumindest im Deutschen – keine neue Bindungsoptionen geschaffen werden können.¹⁴ Dann ist diese Beobachtung ein Argument für eine D-strukturelle Abfolge direktes Objekt \succ indirektes Objekt – wenn ein indirektes Objekt einem direkten Objekt im Mittelfeld vorangeht, muß es gescrambelt worden sein und kann somit das direkte Objekt nicht mehr A-bindern.

Die zweite Grundannahme betrifft den Begriff der Markiertheit. Vorausgesetzt wird, daß die Grammatik im Bereich der freien Wortstellung systematisch mit konkurrierenden, an sich grammatischen Kandidaten relative Grade der Markiertheit korrelieren sollte. Wie können diese Markiertheitsgrade empirisch überprüft werden? Der Vorschlag ist, daß dies zunächst einmal ebenso erfolgen kann wie üblicherweise bei der Beurteilung von Grammatikalität: Sprecher haben oft relativ einheitliche Intuitionen darüber, welcher von zwei grammatischen konkurrierenden Kandidaten weniger normal, d.h., markierter, klingt. Darüber hinaus wird Bezug genommen auf eine Idee von Höhle (1982), der seinerseits Überlegungen von Lenerz (1977) weiterentwickelt. Die Idee ist in (16) wiedergegeben.

- (16) Von zwei Sätzen α , β , die sich nur hinsichtlich von Wortstellung und/oder Akzentzuweisung unterscheiden, ist α weniger markiert als β , falls α mehr Foki als β haben kann (d.h., in mehr Kontexttypen auftreten kann).

Tatsächlich scheint es so zu sein, daß zumindest tendentiell ein Satz, der in mehr Kontexten natürlich klingt als ein konkurrierender Satz, auch insgesamt von Sprechern als weniger markiert eingestuft wird. Betrachten wir ein Standardbeispiel mit Varianten von Sätzen, die oben schon besprochen wurden: Auf die Frage “Wem hat der Fritz das Buch gegeben?”, die in der Antwort einen Fokus auf dem indirekten Objekt erfordert, kann man (17-a) gut sagen, (17-b) weniger gut. Auf die Frage “Was hat der Fritz der Maria gegeben?” kann man umgekehrt (17-b) sagen, aber kaum (17-a). Wenn nun jedoch der Fragekontext so ist, daß eine größere Konstituente als eine NP im Fokus steht (bzw., daß der Fokus auf eine größere Konstituente projiziert wird), erweist sich, daß (17-b) durchweg adäquater ist als (17-a). Dies gilt z.B. für die Fragen “Was hat der Fritz gemacht?” (eine Konstituente, die indirektes und direktes Objekt sowie das Verb enthält, ist der Fokus) und “Was ist passiert?” (sogenannte maximale Fokusprojektion, d.h., der ganze Satz ist der Fokus).

¹⁴Dies ist zwar manchmal bestritten worden, aber eigentlich immer nur auf der Basis von Daten wie den wohlgeformten Sätzen in (15), und nur deshalb, weil man deren Wohlgeformtheit mit der Annahme versöhnen mußte, daß die Basisabfolge indirektes Objekt vor direktem Objekt ist. Daß man unter dieser Annahme nie befriedigend erklären konnte, wieso dann ein indirektes Objekt in seiner angeblichen Basisposition kein direktes Objekt binden kann, spricht aber ebenso gegen diese Idee wie das Faktum, daß alle unkontroversen Beispiele von Scrambling im Deutschen niemals neue A-Bindungsoptionen erzeugen. Dies gilt z.B. für Scrambling eines direkten Objekts über ein Subjekt hinweg; hierdurch werden keine nominativischen Anaphern möglich; vgl.:

(i) *daß den Fritz₁ sich₁ t₁ mag

- (17) a. Der Fritz hat der MARIA das Buch gegeben
 b. Der Fritz hat der Maria das BUCH gegeben

Oft sieht es also so aus, als konvergierten die Ergebnisse gemäß (16) und der naiven Sprecherintuition; denn fraglos wird (17-a) normalerweise von Sprechern auch ohne ein Raisonement gemäß (16) als markierter beurteilt als (17-b). Die Annahme ist somit, daß (relative) Markiertheit von Sätzen ebenso wie Grammatikalität empirisch determinierbar ist. Dann stellt sich die Frage, ob und wie die empirisch ermittelten Markiertheitsgrade von der syntaktischen Komponente systemintern vorhergesagt werden können. Hierfür gibt es in der Literatur ein klassisches Konzept, das u.a. von Lenerz (1977), Frey & Tappe (1991), Haider (1992) und Fortmann & Frey (1997) verfolgt worden ist: Die D-strukturelle Argumentabfolge ist unmarkiert; Scrambling erzeugt Markiertheit. Diese Annahme ist ein Abkömmling der derivationalen Theorie der Komplexität (DTK) der sechziger Jahre, derzufolge Transformationen für die Sprachverarbeitung Kosten aufwerfen, so daß Strukturen, die das Resultat der Anwendung einer kleineren Zahl von Transformationen sind, gegenüber Strukturen, deren Erzeugung einer größeren Menge von Transformationen bedarf, einfacher zu verstehen und zu generieren sind. Man kann daher den zur Debatte stehenden Ansatz als DTK-basierte Theorie der Markiertheit bezeichnen. Wie Haider (1992) bemerkt, legen bei Annahme der DTK-basierten Theorie der Markiertheit die folgenden Daten den Schluß nahe, daß es für unterschiedliche Verbtypen verschiedene D-Struktur-Abfolgen gibt: Die Verben in (18) haben demnach eine Basisabfolge indirektes Objekt \succ direktes Objekt; die Verben in (19) dagegen weisen eine Basisabfolge direktes Objekt \succ indirektes Objekt auf.

- (18) a. daß er {der Maria den Erfolg gönnt/dem Fritz die Zeitung überließ }
 b. ?daß er {den Erfolg der Maria gönnt/die Zeitung dem Fritz überließ }
 (19) a. ?daß man diesem Einfluß die Kinder {ausgeliefert/entzogen } hat
 b. daß man die Kinder diesem Einfluß {ausgeliefert/entzogen } hat

Hiergegen wird allerdings in Müller (1999c) argumentiert. Ein erstes Gegenargument ist, daß die Beispiele in (15) zeigen, daß Verben wie *gönnen* und *ausliefern* ein identisches Bindungsverhalten der beiden Objekte aufweisen, das uniform für eine Basisabfolge direktes Objekt \succ indirektes Objekt spricht, obwohl sie gemäß der DTK-basierten Theorie der Markiertheit aufgrund der Evidenz in (18) und (19) verschiedene Basisabfolgen haben sollten. Ein zweites Gegenargument ist, daß sich Belebtheit als ein intervenierender Faktor bei der unmarkierten Abfolge erweist: Bei Verben wie *entziehen*, die scheinbar als Grundabfolge direktes Objekt \succ indirektes Objekt haben, tauscht sich die unmarkierte Argumentabfolge um, falls das indirekte Objekt nicht wie in (19) unbelebt, sondern belebt ist. Dies zeigen die Beispiele in (20) (vgl. Fanselow (1995) und Vogel & Steinbach (1998)).

- (20) a. daß man der Maria {die Kinder/den Führerschein/das Buch } entzogen hat

- b. ?daß man {die Kinder/den Führerschein/das Buch} der Maria entzogen hat

Diese Überlegungen zeigen, daß es nicht möglich ist, lexikalisch eine von Verb zu Verb unterschiedliche, aber jeweils feste Grundabfolge der Argumente zu spezifizieren, deren Beibehaltung in der Syntax dann eine Voraussetzung für Unmarkiertheit ist. Wenn nun einerseits die D-strukturelle Abfolge bei allen Verben im Deutschen gleich ist (ein direktes Objekt c-kommandiert und präzediert ein indirektes Objekt), andererseits auf der S-Struktur manchmal diese, manchmal die umgekehrte Abfolge unmarkiert ist, dann kann man daraus schließen, daß Scrambling nicht per se Markiertheit erzeugt, sondern manchmal geradezu hin zur Unmarkiertheit führt. Welche der beiden Möglichkeiten vorliegt, das hängt, so die Annahme, allein ab von der Güte des resultierenden Beschränkungsprofils: In einem Verständnis sind zwar fast alle Abfolgen optimal (und somit grammatisch), in einem anderen, restriktiveren Verständnis ist jedoch nur noch eine Abfolge optimal (und somit unmarkiert). Dies führt uns zur eigentlichen Analyse.

4.1.2. Das Scrambling-Kriterium und die Aufspaltung des Optimalitätsbegriffs

Für die gegenwärtigen Zwecke können wir annehmen, daß die im Wettbewerb befindlichen Kandidaten im vorliegenden Ansatz S-Struktur-Repräsentationen sind, die auf dieselbe D-Struktur-Repräsentation als Input zurückgehen. Ob D-Struktur-Repräsentationen direkt von einer Standardgrammatik ohne verletzbare und geordnete Beschränkungen erzeugt werden oder ihrerseits Resultat einer vorgelagerten Optimierungsprozedur sind, spielt für die Analyse der freien Wortstellung keine Rolle, solange gewährleistet ist, daß die D-Struktur eine uniforme Argumentordnung Subjekt \succ direktes Objekt \succ indirektes Objekt aufweist.¹⁵

Die Analyse beruht dann auf lediglich zwei Beschränkungen. Auf der einen Seite verbietet die Beschränkung ÖKON syntaktische Bewegung. Auf der anderen Seite löst eine höher geordnete Beschränkung SCR-KRIT ("Scrambling-Kriterium") Mittelfeld-internes Scrambling (innerhalb der VP) aus. SCR-KRIT ist jedoch keine einfache Beschränkung, wie wir sie bisher kennengelernt haben. Vielmehr handelt es sich um ein komplexes Prinzip, das aus verschiedenen verletzbaren, zueinander geordneten Teilbeschränkungen besteht. Diese Einzelbeschränkungen kodieren ihrerseits die bei der freien Wortstellung relevanten Faktoren Definitheit, Kasus, Belebtheit, usw.:¹⁶

¹⁵Dies ist eine geringfügige Vereinfachung. In Müller (1999c) wird tatsächlich davon ausgegangen, daß der Input eine Numeration und die im Wettbewerb befindlichen Outputs \langle D-Struktur, S-Struktur \rangle -Paare sind. Unter dieser Perspektive muß die D-Struktur ebenfalls Resultat einer Optimierung sein, und ein \langle D-Struktur, S-Struktur \rangle -Paar ist optimal, wenn es sowohl hinsichtlich der D-Struktur, als auch hinsichtlich der S-Struktur optimal ist. Siehe Kapitel 8.

¹⁶Hier sind nur einige der Teilbeschränkungen von SCR-KRIT aus Müller (1999c) aufgelistet.

- (21) SCR-KRIT (“Scrambling-Kriterium”): Im VP-Bereich gilt:
- a. NOM (“Nominativbedingung”): [+nom] vor [-nom] >
 - b. DEF (“Definitheitsbedingung”): [+def] vor [-def] >
 - c. BEL (“Belebtheitsbedingung”): [+belebt] vor [-belebt] >
 - d. FOK (“Fokusbedingung”): [-fok] vor [+fok] >
 - e. DAT (“Dativbedingung”): [+dat] vor [+akk] >

Diese komplexe Beschränkung wird in Müller (1999c) als Subhierarchie bezeichnet. Dieser Begriff ist oben bereits mehrfach aufgetaucht, so z.B. in Abschnitt 2 von Kapitel 2 und in Abschnitt 3 von Kapitel 3; vgl. Legendre et al. (1995; 1998), Baković (1995; 1998), Grimshaw (1998), Schmid (1998; 1999) und Müller (2000a). In all diesen Arbeiten bedeutet die Eingruppierung von Beschränkungen in eine Subhierarchie eigentlich immer nur, daß die Beschränkungen in einem festen Ordnungsverhältnis zueinander stehen und daß sie in gewisser Weise miteinander verwandt (bzw. auseinander abgeleitet) sind. Unter dieser Perspektive wäre das Ordnungssymbol $>$ in (21) dem normalen Ordnungssymbol \gg in der Optimalitätstheorie gleichzusetzen. Dies ist jedoch in Müller (1999c) nicht intendiert: Abgeleitet werden soll z.B., daß (22-b) zwar markierter ist als (22-a), aber nicht, daß (22-b) ungrammatisch ist, weil im Konflikt zwischen NOM und DEF die Verletzung von NOM fatal wird.

- (22) a. daß eine Frau den Fritz geküßt hat
 b. ?daß den Fritz₂ eine Frau₁ t₂ geküßt hat

Daß unter der Gleichsetzung von $>$ und \gg (22-b) ungrammatisch werden sollte, zeigt Tabelle T_{6.9}.

T_{6.9}: NOM \gg DEF, *unerwünschtes Ergebnis*:

Kandidaten	NOM	DEF	BEL	FOK	DAT	ÖKON
☞ K ₁ : Subj[-def] ₁ DO[+def] ₂		*				
K ₂ : DO[+def] ₂ Subj[-def] ₁ t ₁	*!					*

Wenn man will, daß die NOM-Verletzung von K₂ zwar zu Markiertheit, aber nicht zu Ungrammatikalität führt, könnte eine naheliegende Lösung darin bestehen, daß man die Annahme aufgibt, daß nur ein optimaler Kandidat grammatisch sein kann. Man könnte postulieren, daß zwar nur ein optimaler Kandidat sowohl grammatisch als auch vollkommen unmarkiert ist, daß aber der nächstbeste Wettbewerber zwar stärker markiert, aber immer noch grammatisch ist, und so weiter, d.h., daß ein Kandidat umso markierter in seiner Kandidatenmenge ist, je schlechter sein Beschränkungsprofil ist. Es gäbe dann statt absoluter Ungrammatikalität nur noch extreme Markiertheit. Dies entspricht der Theorie der Suboptimalität, die in Keller (1996) zur Behandlung von gradueller Akzeptabilität bei der Extraktion aus NPs entwickelt worden ist. Suboptimalität ist wie in (23) definiert.¹⁷

¹⁷Keller redet statt von “stärker markierten” Strukturen von “weniger grammatischen” Struk-

(23) *Suboptimalität* (Keller (1996, 50)):

Eine Struktur S_i ist gegenüber einer Struktur S_j suboptimal, wenn es Teilmengen R_i und R_j in der Kandidatenmenge gibt, wobei S_i in R_i und S_j in R_j optimal ist, und wobei gilt: $R_i \subset R_j$. Eine Struktur S_i ist stärker markiert als eine Struktur S_j , wenn S_i gegenüber S_j suboptimal ist.

(23) macht für den Wettbewerb in Tabelle T_{6.9} die erwünschte Vorhersage: K_2 ist gegenüber K_1 suboptimal (die Teilmenge der Kandidatenmenge, in der K_2 optimal ist, darf nicht K_1 enthalten, während K_1 in der gesamten Kandidatenmenge optimal ist) und somit stärker markiert. Eine derartige Lösung ist aber mit praktisch sämtlichen existierenden optimalitätstheoretischen Analysen unvereinbar: Denn im Normalfall ist es gerade nicht so, daß der zweitbeste Kandidat akzeptabler oder weniger markiert erscheint als der drittbeste, usw. Dies mag ein willkürlich herausgegriffenes Beispiel illustrieren. Wie wir in Kapitel 3 gesehen haben, erklärt die Interaktion der drei geordneten Beschränkungen OP-SPEZ, OB-KOPF und ÖKON bei Grimshaw (1997), wieso in einem einfachen Matrixfragesatz im Englischen sowohl W-Bewegung, als auch V-Inversion stattfindet:

- (24) a. What will John read ?
 b. *What John will read ?
 c. *John will read what ?
 d. *Will John read what ?

Tabelle T_{6.10} zeigt, daß (24-d) gegenüber (24-c) suboptimal ist, daß (24-c) gegenüber (24-b) suboptimal ist, und daß (24-b) gegenüber (24-a) suboptimal ist.¹⁸ Die unterschiedliche Qualität der Beschränkungsprofile suboptimaler Kandidaten findet hier jedoch keinen empirischen Niederschlag: (24-b) etwa ist als Matrixfragesatz im Englischen genauso unmöglich wie (24-c).

T_{6.10}: *Matrix-Fragesätze*

Kandidaten	OP-SPEZ	OB-KOPF	ÖKON
$\Leftarrow K_1$: [CP what ₁ will ₂ [IP NP t ₂ [VP read t ₁]]]			**
K_2 : [CP what ₁ e [IP NP will [VP read t ₁]]]		*!	*
K_3 : [IP NP will [VP read what ₁]]	*!		
K_4 : [CP will ₂ [IP NP t ₂ [VP read what ₁]]]	*!		*

Angesichts dieses Dilemmas ist der Schluß in Müller (1999c), daß die Beschränkungen, die Scrambling auslösen, nicht auf einer Ebene mit den üblichen optimalitätstheoretischen Beschränkungen stehen (bei denen eine fatale Verletzung zu

turen; aus Gründen der terminologischen Einheitlichkeit habe ich diesen Begriff hier vermieden.

¹⁸T_{6.10} ist gegenüber Tabelle T_{3.1} aus Kapitel 3 leicht vereinfacht und gemäß der Abfolge in (24) umgruppiert.

Ungrammatikalität führt), sondern zu einer anderen Ebene gehören (bei denen eine fatale Verletzung nur zu Markiertheit führt). Um also die deutsche Wortstellung optimalitätstheoretisch zu erfassen, ohne dabei Grundannahmen der Optimalitätstheorie aufzugeben, wird das Gesamtsystem insofern komplizierter gemacht, als es eine Unterscheidung zweier Konzepte von Optimalität integriert – Optimalität als Grammatikalität einerseits, und Optimalität als Unmarkiertheit andererseits.

Der entscheidende Schritt besteht darin, dem Ordnungssymbol $>$ in Subhierarchien wie SCR-KRIT eine andere Interpretation zu geben als dem Ordnungssymbol \gg der üblichen Hierarchie von Beschränkungen, die nunmehr Matrixhierarchie genannt wird.¹⁹ Konkret kann $>$ sowohl als logisches \vee verstanden werden (dann geht es um Grammatikalität), als auch als klassisches \gg (dann geht es um (Un-)Markiertheit). Optimalität als Grammatikalität ist wie in (25) definiert.²⁰

(25) *Grammatikalität* (“Optimalität₁”):

Ein Kandidat K_i ist optimal₁ (grammatisch) gdw. es keinen anderen Kandidaten K_j in derselben Kandidatenmenge gibt, der hinsichtlich der Matrixhierarchie $\langle B_1 \gg B_2 \dots \gg \dots B_n \rangle$ ein besseres Beschränkungsprofil hat, wobei B_i durch ein C_i aus $\langle C_1 > C_2 > \dots > C_n \rangle$ ersetzt ist, falls B_i eine Subhierarchie $\langle C_1, C_2, \dots C_n \rangle$ ist.

Dies bedeutet, zur Ermittlung der Grammatikalität eines Kandidaten in einer Matrixhierarchie, die eine Subhierarchie wie SCR-KRIT enthält, muß herausgefunden werden, ob der Kandidat unter Ersetzung wenigstens einer Teilbeschränkung für die Subhierarchie insgesamt ein optimales Beschränkungsprofil aufweist. Dies ist äquivalent zu dem, was in (46) in Kapitel 5 als global disjunktive Kopplung eingeführt wurde. Die empirische Konsequenz ist, daß der Auslöser für Scrambling im Deutschen an sich relativ schwach ist: Scrambling ist möglich, wenn es durch mindestens eine der Beschränkungen der SCR-KRIT-Subhierarchie legitimiert ist. Dies ist im Effekt durchaus vergleichbar dem, was die Theorie von Uszkoreit (1984; 1986) vorhersagt. Das zweite Konzept der Optimalität ist das der Unmarkiertheit.

(26) *Unmarkiertheit* (“Optimalität₂”):

Ein Kandidat K_i ist optimal₂ (unmarkiert) gdw. es keinen anderen Kandidaten K_j in derselben Kandidatenmenge gibt, der hinsichtlich der Matrixhierarchie

¹⁹Dies soll natürlich nicht für das gelten, was in Legendre et al. (1995; 1998) und den anderen oben erwähnten Arbeiten Subhierarchie genannt wird. Die Terminologie ist damit an diesem Punkt nicht so eindeutig, wie das wünschenswert wäre; aber wenn in diesem Kapitel im folgenden von Subhierarchien die Rede ist, ist immer nur das Konzept in Müller (1999c) gemeint.

²⁰Die Definitionen aus Müller (1999c), die auf Grimshaws (1997) Optimalitätsbegriff in (i) aus Fußnote 3 aus Kapitel 1 basieren, sind hier und im folgenden an den Optimalitätsbegriff in (5) aus Kapitel 1 adaptiert, um mehrere Gewinner als Resultat eines identischen Beschränkungsprofils zuzulassen. Darüber hinaus ist hier die Ausführung dessen, was es heißt, ein besseres Beschränkungsprofil zu haben, weggelassen.

archie $\langle B_1 \gg B_2 \dots \gg \dots B_n \rangle$ ein besseres Beschränkungsprofil hat, wobei B_i durch $\langle C_1 \rangle C_2 \rangle \dots \rangle C_n \rangle$ ersetzt ist, falls B_i eine Subhierarchie $\langle C_1, C_2, \dots C_n \rangle$ ist.

Zur Ermittlung der Unmarkiertheit wird also de facto die Unterscheidung von Matrix- und Subhierarchie wieder ignoriert; unmarkiert ist der Kandidat, der unter Ersetzung aller Teilbeschränkungen einer Subhierarchie (inkl. ihrer Ordnung) für die Subhierarchie selbst das beste Beschränkungsprofil hat. Was nun noch fehlt, ist ein Mittel, um relative Grade der Markiertheit abzuleiten. Hier wird auf Kellers (1996) Theorie der Suboptimalität zurückgegriffen: Je schlechter gemäß (26) das Beschränkungsprofil eines nach (25) grammatischen Kandidaten im Vergleich zu den Konkurrenten ist, desto markierter ist er. Betrachten wir einige Beispiele.

4.1.3. Empirie

Die Sätze in (22) zeigen, daß Scrambling eines definiten Objekts über ein indefinites Subjekt zwar möglich ist, aber in Markiertheit resultiert; sie sind in (27) wiederholt.

- (27) a. daß eine Frau den Fritz geküßt hat
 b. ?daß den Fritz₂ eine Frau₁ t₂ geküßt hat

Aus Tabelle T_{6.11} geht zunächst einmal hervor, daß beide Kandidaten grammatisch sind: K_1 ist grammatisch (optimal₁) unter einer Ordnung, in der SCR-KRIT durch z.B. NOM ersetzt ist; K_2 ist dagegen grammatisch (optimal₁) unter einer Ordnung, in der SCR-KRIT durch DEF ersetzt ist. Letztere Beschränkung läßt sich somit als Auslöser für Scrambling auffassen.²¹ Zusätzlich zur Grammatikalität kann man nun aus Tabelle T_{6.11} die relative Markiertheit ableiten. Hierzu wird SCR-KRIT durch die Gesamtordnung der Teilbeschränkungen ersetzt. In diesem Fall gewinnt K_1 gegen K_2 ; die NOM-Verletzung von K_2 erweist sich nun als fatal, und K_2 wird korrekt als suboptimal₂ klassifiziert, d.h., als markierter als K_1 . In Tabellen wird Grammatikalität (Optimalität₁) eines Kandidaten weiterhin durch ☞ signalisiert (und eine fatale Verletzung durch "!"), Unmarkiertheit (Optimalität₂) dagegen durch einen Pfeil → (und eine "fatale" Verletzung eines grammatischen Kandidaten, die ihn suboptimal₂ macht, durch "?").

Als ein weiteres Beispiel betrachte man das folgende Satzpaar, das über die SCR-KRIT-interne Ordnung von DEF und BEL Aufschluß gibt,

- (28) a. daß der Verkäufer den Wein₁ einem Kunden₂ empfahl
 b. ?daß der Verkäufer einem Kunden₂ den Wein₁ t₂ empfahl

²¹Tatsächlich stellt sich heraus, daß K_1 gegenüber K_2 einen inhärenten Vorteil mitbringt: Weil K_1 die D-strukturelle Abfolge aufweist und somit ohne eine durch Scrambling hervorgerufene ÖKON-Verletzung auskommt, gewinnt K_1 gegen K_2 bei der Ersetzung von SCR-KRIT durch jede Teilbeschränkung, wo sich die beiden Kandidaten nicht unterscheiden, also auch bei BEL, FOK und DAT.

$T_{6.11}$: NOM > DEF bei Müller

Kandidaten	SCR-KRIT					ÖKON
	NOM	DEF	BEL	FOK	DAT	
☞ $\rightarrow K_1$: Subj[-def] ₁ DO[+def] ₂		*				
☞ K_2 : DO[+def] ₂ Subj[-def] ₁ t ₂	*?					*

Beide Sätze sind möglich. Wie jedoch Stechow & Sternefeld (1988, 453) bemerken, wird im Konfliktfall vom unmarkierten Kandidaten DEF respektiert und BEL verletzt. Grammatikalität und Unmarkiertheit ergeben sich aus Tabelle T_{6.12}.

$T_{6.12}$: DEF > BEL bei Müller

Kandidaten	SCR-KRIT					ÖKON
	NOM	DEF	BEL	FOK	DAT	
☞ $\rightarrow K_1$: DO[+def,-bel] ₁ IO[-def,+bel] ₂			*		*	
☞ K_2 : IO[-def,+bel] ₂ DO[+def,-bel] ₁ t ₂		*?				*

Man beachte, daß hier zum einen hin zur Unmarkiertheit Scrambling erfolgt (per Annahme c-kommandiert und präzediert das direkte Objekt ja uniform das indirekte Objekt auf der D-Struktur), und daß zum anderen der unmarkierte Kandidat K_1 aufgrund des zusätzlichen Verstoßes gegen die tief geordnete Beschränkung DAT insgesamt mehr Verletzungen von SCR-KRIT-Teilbeschränkungen aufweist als der markiertere Kandidat K_2 . Dies zeigt, daß Kumulativität von Beschränkungsverletzungen im System von Müller (1999c) keine Rolle spielt.

Die unterschiedliche unmarkierte Abfolge bei den Beispielen in (18), (19) und (20) wird auf die relative Ordnung der Teilbeschränkungen BEL und DAT in der SCR-KRIT-Subhierarchie zurückgeführt. DAT löst Scrambling des indirektes Objekts über das direkte Objekt aus; BEL verlangt Belebtes vor Unbelebtem. Ist nun das indirekte Objekt belebt und das direkte Objekt unbelebt wie in (29) (vgl. (18) und z.T. (20)), so verlangen BEL und DAT dasselbe. Wie Tabelle T_{6.13} zeigt, führt Scrambling von K_1 hier somit zur Unmarkiertheit; K_2 ist zwar markiert, aber nichtsdestoweniger grammatisch, da dieser Kandidat ohne ÖKON-Verletzung unter Ersetzung durch jede SCR-KRIT-Teilbeschränkung optimal₁ ist, wo die beiden Kandidaten sich nicht unterscheiden.

- (29) a. daß er der Maria den Erfolg gönnt
 b. ?daß er den Erfolg der Maria gönnt

$T_{6.13}$: BEL > DAT: IO[+bel], DO[-bel] bei Müller

Kandidaten	SCR-KRIT					ÖKON
	NOM	DEF	BEL	FOK	DAT	
☞ $\rightarrow K_1$: IO[+bel] ₂ DO[-bel] ₁ t ₂						*
☞ K_2 : DO[-bel] ₁ IO[+bel] ₂			*?		*	

Entsteht dagegen ein Konflikt von BEL und DAT, weil das indirekte Objekt unbelebt und das direkte Objekt belebt ist, setzt sich, wie (30) (vgl. (19)) zeigt, hinsichtlich der Unmarkiertheit BEL durch; vgl. Tabelle T_{6.14}.

- (30) a. ?daß man diesem Einfluß die Kinder entzogen hat
 b. daß man die Kinder diesem Einfluß entzogen hat

T_{6.14}: BEL > DAT: IO[-bel], DO[+bel] bei Müller

Kandidaten	SCR-KRIT					ÖKON
	NOM	DEF	BEL	FOK	DAT	
☞ K ₁ : IO[-bel] ₂ DO[+bel] ₁ t ₂			*?			*
☞ →K ₂ : DO[+bel] ₁ IO[-bel] ₂					*	

Sind beide Objekte belebt, differenziert BEL die Kandidaten nicht mehr, und die Entscheidung über Unmarkiertheit fällt (bei ebenfalls vorausgesetzter Irrelevanz der anderen höher geordneten Beschränkungen in SCR-KRIT) der tiefer geordneten Beschränkung DAT zu. Dies ergibt sich aus der Betrachtung von (31) (vgl. (20)) und der dazugehörigen Tabelle T_{6.15}.

- (31) a. daß man der Maria₂ die Kinder₁ t₂ entzogen hat
 b. ?daß man die Kinder₁ der Maria₂ entzogen hat

T_{6.15}: DAT: IO[+bel], DO[+bel] bei Müller

Kandidaten	SCR-KRIT					ÖKON
	NOM	DEF	BEL	FOK	DAT	
☞ →K ₁ : IO[+bel] ₂ DO[+bel] ₁ t ₂						*
☞ K ₂ : DO[+bel] ₁ IO[+bel] ₂					*?	

Natürlich sollte sich genau dasselbe Ergebnis einstellen, wenn beide Objekte unbelebt sind. Daten wie die in (32) bestätigen dies.²²

- (32) a. daß man dem Film₂ den Preis₁ zuerkannt hat
 b. ?daß man den Preis₁ dem Film₂ zuerkannt hat

²²Allerdings scheinen zunächst Beispiele wie (i-ab) in die entgegengesetzte Richtung zu weisen.

- (i) a. daß man das Buch der Prüfung unterzogen hat
 b. *daß man der Prüfung das Buch unterzogen hat

Will man das bisher dargestellte System angesichts von (i) aufrecht erhalten, muß man nachweisen, daß es hier einen interferierenden Faktor gibt. In Müller (1999c) wird angenommen, daß Funktionsverbgefüge wie *der Prüfung unterziehen* einer höher als SCR-KRIT geordneten Beschränkung unterliegen, die die syntaktische Trennung von Phraseologismen im Mittelfeld verbietet, und deren Verletzung dann in (i-b) nicht nur zu Suboptimalität₂ (Markiertheit), sondern auch zu Suboptimalität₁ (Ungrammatikalität) führt.

Unter dieser Perspektive gibt es also keine unterschiedlichen Verbklassen mit unterschiedlicher D-struktureller Abfolge im Deutschen. Eine Abfolge mit indirektem Objekt vor direktem Objekt führt im Deutschen in der Mehrzahl der Fälle zu Unmarkiertheit. Dies sagt aber nichts über die D-strukturelle Abfolge aus. Vielmehr folgt diese Verteilung daraus, daß zum einen DAT direkt diese Abfolge favorisiert, und daß zum anderen die Verben des Gebens und Nehmens im weiteren Sinne die größte Verbklasse mit indirektem und direktem Objekt im Deutschen bilden (vgl. etwa Eisenberg (1986)); gerade diese Verben realisieren jedoch üblicherweise das indirekte Objekt als belebt und das direkte Objekt als unbelebt. Somit tendiert auch BEL dazu, die Abfolge indirektes Objekt \succ direktes Objekt zu favorisieren.

Abschließend sei noch auf die Effekte von FOK eingegangen. Diese Teilbeschränkung von SCR-KRIT entspricht NEU bei Choi (1996; 1999) und ENDFOK (bzw. der Dekonstruktion von ENDFOK durch prosodische Beschränkungen) bei Büring (1997; 1999). Die Grundannahme ist bei Müller (1999c) anders als bei Choi und Büring, daß eine syntaktische Fokusmarkierung nicht im Input präsent ist; sie wird vielmehr auf der S-Struktur hinzugefügt. Dies bedeutet, daß Kandidaten mit unterschiedlichen Fokusmarkierungen miteinander im Wettbewerb sind. Was die Einordnung von FOK in die SCR-KRIT-Subhierarchie betrifft, so wird angenommen, daß FOK unterhalb von BEL und oberhalb von DAT steht. Die Ordnung $BEL > FOK$ hat zur Folge, daß die vier Sätze in (11), die hier nochmals in (33) wiederholt sind, im Unterschied zu dem, was wir bei Choi und z.T. bei Büring gesehen haben, zwar alle (auch (33-d)) als grammatisch vorhergesagt werden; daß aber die Markiertheit von oben nach unten zunimmt.²³

- (33) a. daß der Fritz der Maria das BUCH gegeben hat
 b. ?daß der Fritz der MARIA das Buch gegeben hat
 c. ??daß der Fritz das Buch der MARIA gegeben hat
 d. ?*daß der Fritz das BUCH der Maria gegeben hat

Dies zeigt Tabelle T_{6.16}. Was die Grammatikalität aller Kandidaten betrifft, so beachte man, daß K_4 die D-strukturelle Abfolge aufweist und so $optimal_1$ ist unter Ersetzung von SCR-KRIT durch NOM und DEF. Hinsichtlich der Markiertheit gilt: K_1 respektiert BEL, FOK und DAT und ist unmarkiert, also $optimal_2$. K_2 verletzt nur FOK und ist markierter als K_1 . K_3 verletzt BEL (sowie DAT) und ist seinerseits markierter als K_2 . K_4 schließlich verletzt BEL, FOK und DAT und ist am stärksten markiert.

Aus der Ordnung $FOK > DAT$ ergibt sich demgegenüber die Markiertheitsverteilung in (34), wo gegenüber (33) der Belebtheitsfaktor ausgeschaltet ist:

- (34) a. daß der Fritz der Maria den KARL vorgestellt hat

²³Dies ist eine empirische Frage. Erwiese sich z.B. (33-c) als weniger markiert als (33-b), wären innerhalb von SCR-KRIT BEL und FOK zu vertauschen.

$T_{6.16}$: BEL > FOK bei Müller

Kandidaten	SCR-KRIT					ÖKON
	NOM	DEF	BEL	FOK	DAT	
☞ $\rightarrow K_1$: IO[+bel,-fok] ₂ DO[-bel,+fok] ₁ t ₂						*
☞ K ₂ : IO[+bel,+fok] ₂ DO[-bel,-fok] ₁ t ₂				*?		*
☞ K ₃ : DO[-bel,-fok] ₁ IO[+bel,+fok] ₂			*?		*	
☞ K ₄ : DO[-bel,+fok] ₁ IO[+bel,-fok] ₂			*?	*?	*	

b. ?daß der Fritz den Karl der MARIA vorgestellt hat

c.??daß der Fritz der MARIA den Karl vorgestellt hat

d.*daß der Fritz den KARL der Maria vorgestellt hat

Den Wettbewerb in (34) illustriert Tabelle T 6.17.

$T_{6.17}$: FOK > DAT bei Müller

Kandidaten	SCR-KRIT					ÖKON
	NOM	DEF	BEL	FOK	DAT	
☞ $\rightarrow K_1$: IO[+bel,-fok] ₂ DO[+bel,+fok] ₁ t ₂						*
☞ K ₂ : DO[+bel,-fok] ₁ IO[+bel,+fok] ₂					*?	
☞ K ₃ : IO[+bel,+fok] ₂ DO[+bel,-fok] ₁ t ₂				*?		*
☞ K ₄ : DO[+bel,+fok] ₁ IO[+bel,-fok] ₂				*?	*?	

Es ist in diesem System gar nicht so einfach für einen Kandidaten, allein aufgrund einer Mittelfeld-internen Abfolge (bzw. Fokusverteilung) ungrammatisch zu werden. Ganz unmöglich ist es jedoch auch nicht: Vollständig leeres, ketteninvariantes Scrambling wird ebenso als suboptimal₁ ausgeschlossen (aufgrund einer fatalen ÖKON-Verletzung, die nicht zu einer Verbesserung bei irgendeiner der höher geordneten SCR-KRIT-Beschränkungen führt; vgl. Abschnitt 4.1 von Kapitel 2) wie eine Abfolge, die eine höher geordnete, SCR-KRIT-externe Beschränkung (wie z.B. ein Bindungsprinzip) verletzt.

4.2. Diskussion

Der Ansatz in Müller (1999c) vermeidet zwei der Schwierigkeiten für Pseudo-Optionalitäts-Analysen: Zum einen kann es kaum zu Untergenerierung kommen, weil der Auslöser für Wortstellungsvariation die disjunktive Interpretation der Einzelbeschränkungen von SCR-KRIT und somit relativ schwach ist. Zum anderen kann neben der Grammatikalität bzw. Ungrammatikalität eines gegebenen Kandidaten auch seine relative Markiertheit im Vergleich zu anderen Kandidaten optimalitätstheoretisch vorhergesagt werden.

Allerdings ist auch diese Analyse nicht problemlos. Die Annäherung an die in den zu Beginn dieses Abschnitts erwähnten Arbeiten manifeste empirische Adäquatheit ist teuer erkaufte durch eine nicht-triviale und weitreichende Modifikation der Optimalitätstheorie: Solange Beschränkungsordnungen keine Subhierarchien im Sinne von SCR-KRIT enthalten, ändert sich in diesem System zwar nichts gegenüber

Prince & Smolensky (1993) (und insbesondere ist das Modell im Prinzip kompatibel mit anderen Arbeiten zur optimalitätstheoretischen Syntax); wenn aber Subhierarchien wie SCR-KRIT ins Spiel kommen, hat man es sofort mit zwei Begriffen der Optimalität und, damit verbunden, mit einer Aufspaltung der Beschränkungsordnung in zwei disparate Ebenen zu tun. Die wesentliche Neuerung – eine disjunktive Interpretation von Beschränkungen in der Subhierarchie zwecks Determination von Grammatikalität – läßt sich nun zwar, wie in Kapitel 5 ausgeführt, als eine spezielle Form von globaler Kopplung auffassen. Wie jedoch schon im letzten Kapitel gezeigt worden ist, ist dieser Kopplungsbegriff als generelles Mittel, Optionalität abzuleiten, ungeeignet, weil dadurch eine gekoppelte Beschränkung für einen Kandidaten unerwünschterweise vollkommen irrelevant werden kann. Es überrascht daher nicht, daß auch in Müller (1999c) zur Behandlung eines SCR-KRIT-unabhängigen Wortstellungsphänomens (der optionalen Anhebung von Subjekt-NPs in eine externe Subjektposition) noch auf einen anderen Kopplungstyp (global hierarchische Kopplungen) zurückgegriffen werden muß.²⁴

Wenn nun auch unstrittig zu sein scheint, daß Subhierarchien in dem hier spezifizierten Sinne nicht als generelles Kopplungskonzept taugen, so stellt sich doch die Frage, ob von diesem Mittel noch irgendwo sonst in der Grammatik Gebrauch gemacht wird, oder ob es sich hier um eine nur für den Fall der freien Wortstellung im deutschen Mittelfeld notwendige konstruktionsspezifische Lösung handelt. Wenn letzteres der Fall wäre, dann könnte man auf der Basis der deutschen Wortstellungsdaten kein besonders gutes Argument für eine optimalitätstheoretische Analyse gewinnen. Tatsächlich scheinen mir Anwendungen von Subhierarchien in anderen empirischen Bereichen (wie z.B. der Extraktionstheorie, in der oft Grade der Wohlgeformtheit eine Rolle spielen) nicht ausgeschlossen; sie sind jedoch auch nicht völlig offensichtlich. In eine ähnliche Richtung gehen weitere Fragen: Wie kann ein Kind beim Spracherwerb wissen, ob eine Beschränkung zur Matrix- oder zu einer Subhierarchie gehört? Ist die Menge der in eine gegebene Subhierarchie integrierten Beschränkungen fest, oder kann von Sprache zu Sprache variieren, was in einer Subhierarchie auftaucht, und was in der Matrixhierarchie? Gibt es Sprachvariation, die auf Beschränkungsreihenfolge innerhalb einer Subhierarchie zurückgeht? (Ist z.B. eine vom Deutschen minimal unterschiedliche Sprache möglich, in der DEF innerhalb von SCR-KRIT am höchsten oder am tiefsten geordnet ist, mit entsprechenden Konsequenzen für den Markiertheitsstatus gescrembelter definiter NPs?) Diese Fragen sind alle offen.

²⁴Die relevante Kopplung betrifft ÖKON und das Erweiterte Projektionsprinzip EPP. Wäre diese Kopplung global disjunktiv, so wie die Interpretation von > für die Ermittlung von Grammatikalität gemäß SCR-KRIT, ergäbe sich hier die fatale Konsequenz, daß für einen Kandidat, der EPP erfüllt, ÖKON an keiner Stelle mehr relevant ist.

5. Ein D-Struktur-Ansatz: Heck (2000)

5.1. Die Analyse

Abschließend sei noch der Ansatz von Heck (2000) eingeführt. Anders als bei Choi, Büring und Müller wird bei Heck nicht angenommen, daß Oberflächenstrukturen Gegenstand der Wortstellungsoptimierung sind. Vielmehr erfolgt diese Optimierung auf der D-Struktur. Die (auf Vogel & Steinbach (1998) zurückgehende) Grundannahme ist, daß zwar Subjekt und direktes Objekt in festen VP-internen Positionen erzeugt werden, daß aber die Position des indirekten Objekts variabel ist – das indirekte Objekt kann im Prinzip über dem Subjekt, zwischen Subjekt und direktem Objekt oder unterhalb des direkten Objekts basisgeneriert werden. Welche der somit potentiell erlaubten Abfolgen die einzig D-strukturell optimale ist, wird durch die folgenden vier Beschränkungen festgelegt, die hier in leicht vereinfachter Form aufgelistet sind.

- (35) a. DEF (“Definitheitsbedingung”):
 NP_[+def] geht NP_[-def] voran.
 b. ANIM (“Belebtheitsbedingung”):
 NP_[+bel] geht NP_[-bel] voran.
 c. AGENS (“Agentivitätsbedingung”):
 NP_[+Agens] geht NP_[-Agens] voran.
 d. ADJA (“Adjazenzbeschränkung”):
 Eine NP mit strukturellem Kasus ist dem Kasuszuweiser adjazent.

Die für das Deutsche angenommene Ordnung entspricht der Reihenfolge in (35). DEF und ANIM (= BEL) haben wir schon kennengelernt. AGENS ist in erster Linie für die relative Abfolge von Subjekt und indirektem Objekt auf der D-Struktur relevant: Subjekte sind üblicherweise eher agentiv und werden so vor indirekte Objekte gezwungen. Insofern hat AGENS also einen Effekt, der dem der Beschränkung NOM in der Analyse von Müller (1999c) vergleichbar ist. Man beachte aber, daß AGENS anders als NOM nichts sagt über Fälle, in denen Subjekt und indirektes Objekt beide kein Agens sind. In diesem Fall fordert die tiefstgeordnete Beschränkung ADJA eine Abfolge indirektes Objekt \succ Subjekt, da per Annahme der Nominativkasus des Subjekts strukturell ist (und dieses daher so nah wie möglich an der Flexionsmorphologie als Kasuszuweiser stehen muß), während der Dativkasus des indirekten Objekts als nicht-strukturell eingeschätzt wird. ADJA ist darüber hinaus in diesem System dafür verantwortlich, daß bei Irrelevanz von ANIM das indirekte Objekt dem direkten Objekt vorangeht (der Dativkasus des indirekten Objekts ist per Annahme nicht strukturell, der Akkusativkasus des direkten Objekts ist dies).

Vor der Betrachtung einiger Fälle im Rahmen dieses Modells muß noch ein offensichtliches Problem ausgeräumt werden, nämlich das der scheinbaren Untergenerierung: Hecks (2000) Wortstellungssystem kennt (anders als das LF-Skopussystem in Heck (1999)) keine Kopplung und gibt tatsächlich für jeden Input (der mit Grim-

shaw (1997) als Prädikat-/Argument-Struktur betrachtet wird) nur einen einzigen optimalen D-Struktur-Output aus. Die suboptimalen Abfolgen sollen aber nicht per se ausgeschlossen, sondern als grammatisch, aber markiert klassifiziert werden. Dieses Dilemma wird gelöst durch eine Rückkehr zur DTK-basierten Theorie der Markiertheit, derzufolge nur die D-strukturelle Argumentabfolge unmarkiert ist und Scrambling immer Markiertheit erzeugt. Die optimalen D-strukturellen Abfolgen können durch S-strukturelle Scrambling-Operationen wieder vertauscht werden. Diese Scrambling-Operationen sollen (v.a.) durch Skopus- und Fokus-bezogene Beschränkungen auf der S-Struktur erzwungen werden. Die so entstehenden optimalen S-Strukturen sind zwar grammatisch; trotz Optimalität gelten sie jedoch per DTK-Stipulation als markiert.²⁵ Nun zu einigen Beispielen.

Die Beschränkung DEF und ihre Ordnung über ANIM und ADJA läßt sich illustrieren an Beispielen wie denen in (28), hier wiederholt in (36).

- (36) a. daß der Verkäufer den Wein einem Kunden empfahl
 b. ?daß der Verkäufer einem Kunden den Wein empfahl

Die Analyse ist im Prinzip dem ähnlich, was wir bei der Besprechung von Müller (1999c) in Tabelle T_{6.12} gesehen haben; vgl. Tabelle T_{6.18}.

T_{6.18}: DEF ≫ ANIM bei Heck

Kandidaten	DEF	ANIM	AGENS	ADJA
☞ K ₁ : DO[+def,-bel] IO[-def,+bel]		*		*
K ₂ : IO[-def,+bel] DO[+def,-bel]	*!			

In T_{6.12} und T_{6.18} entsprechen sich DEF und DEF, BEL und ANIM sowie DAT und ADJA. (ÖKON fehlt natürlich in T_{6.18}, da es hier um D-Struktur-Optimierung geht und es auf der D-Struktur noch keine Bewegung gibt.) Der wesentliche Unterschied ist, daß nunmehr K₂ ungrammatisch ist, aber eben nur als D-Struktur. Vorausgesetzt ist, daß die S-Struktur in (36-b) durch Scrambling erzeugt werden kann; somit ist diese Abfolge möglich, aber markiert. Eine zweite Illustration für die Wirkung der Beschränkung DEF, diesmal in Interaktion mit AGENS, ergibt sich bei der Betrachtung von (37).

- (37) a. daß dem Patienten eine Schwester geholfen hat
 b. ?daß eine Schwester dem Patienten geholfen hat

Die S-Struktur (37-a) ist unmarkiert und muß daher die optimale D-Struktur reflektieren; die S-Struktur (37-b) ist demgegenüber markiert und muß somit durch Scrambling aus der optimalen D-Struktur gewonnen worden sein. Daß auf der D-Struktur hier das indirekte Objekt dem Subjekt vorangeht, folgt aus der angenommenen Beschränkungsordnung: Eine Abfolge mit dem indirekten Objekt vor

²⁵Diese S-Struktur-Bewegungen spielen in Heck (2000) allerdings praktisch keine Rolle.

dem Subjekt fordert nicht nur die tief geordnete Beschränkung ADJA, sondern im vorliegenden Fall auch die am höchsten geordnete Beschränkung DEF; daher wird die dazwischen geordnete Beschränkung AGENS, die das Subjekt vor das indirekte Objekt gliedern möchte, verletzt. Dies zeigt Tabelle T_{6.19}.²⁶

T_{6.19}: DEF \gg AGENS bei Heck

Kandidaten	DEF	ANIM	AGENS	ADJA
☞ K ₁ : IO[+def,+bel] Subj[-def,+bel]			*	
K ₂ : Subj[-def,+bel] IO[+def,+bel]	*!			*

Die Evidenz für die Beschränkung ANIM entspricht zum Teil dem, was in Müller (1999c) als Evidenz für BEL beigebracht wird (vgl. T_{6.13}, T_{6.14}). Zusätzlich macht die Analyse korrekte Vorhersagen für die Beispiele in (38).

- (38) a. daß einem Patienten ein Medikament geholfen hat
 b. ?daß ein Medikament einem Patienten geholfen hat

Hier ist DEF irrelevant. ANIM verlangt (wie ADJA) auf der D-Struktur das indirekte Objekt vor dem Subjekt. Lediglich die tiefer geordnete Beschränkung AGENS verlangt hier auf der D-Struktur das Subjekt vor dem indirekten Objekt. Dies endet aber, wie Tabelle T_{6.20} zeigt, in D-struktureller Ungrammatikalität (die Abfolge ist nur markiert auf der S-Struktur durch Scrambling zu erzeugen).

T_{6.20}: ANIM \gg AGENS bei Heck

Kandidaten	DEF	ANIM	AGENS	ADJA
☞ K ₁ : IO[-def,+bel] Subj[-def,-bel]			*	
K ₂ : Subj[-def,-bel] IO[-def,+bel]		*!		*

Ist gegenüber (38) der Belebtheitsfaktor (ebenso wie der Definitheitsfaktor) ausgeschaltet, wird AGENS relevant und sagt korrekt die optimale D-Struktur (und somit unmarkierte S-Struktur) Subjekt \succ indirektes Objekt vorher, die ADJA verletzen muß (weil das Subjekt von seinem Kasuszuweiser durch das indirekte Objekt getrennt ist); vgl. (39) und Tabelle T_{6.21}.

- (39) a. daß ein Arzt einem Patienten geholfen hat
 b. ?daß einem Patienten ein Arzt geholfen hat

²⁶Diese Analyse ist nicht unverträglich mit der in Müller (1999c) für die Daten in (27) in Tabelle T_{6.11} vorgeschlagenen. Im vorliegenden Fall geht es um Subjekt und indirektes Objekt; in T_{6.11} geht es dagegen um Subjekt und direktes Objekt. Für letzteren Fall nehmen beide Analysen an, daß Subjekt vor direktem Objekt ungeachtet von Definitheitseffekten die unmarkierte Abfolge sein muß – im einen Fall aufgrund der Ordnung NOM $>$ DEF, in anderen Fall, weil die Positionen von Subjekt und direktem Objekt nicht als D-strukturell variabel betrachtet werden.

$T_{6.21}$: AGENS \gg ADJA bei Heck

Kandidaten	DEF	ANIM	AGENS	ADJA
☞ K ₁ : Subj[-def,+bel] IO[-def,+bel]				*
K ₂ : IO[-def,+bel] Subj[-def,+bel]			*!	

Sind schließlich alle drei höher geordneten Beschränkungen nicht in der Lage, zwischen zwei Abfolgen eine Entscheidung herbeizuführen, wird ADJA relevant. Dies gilt für Beispiele wie die in (40).

- (40) a. daß man der Maria die Kinder entzogen hat
 b. ?daß man die Kinder der Maria entzogen hat

Tabelle $T_{6.22}$ zeigt, daß hier ADJ dieselbe Funktion wahrnimmt wie DAT in der Tabelle $T_{6.15}$ im Ansatz von Müller (1999c). Der Unterschied ist lediglich, daß so im einen Fall die Abfolge direktes Objekt \succ indirektes Objekt D-strukturell ausgeschlossen und nur S-strukturell als durch Scrambling ableitbar markiert möglich ist, während im anderen Fall diese Abfolge S-strukturell als hinsichtlich der Grammatikalität optimal, hinsichtlich der Unmarkiertheit jedoch suboptimal klassifiziert wird.

$T_{6.22}$: ADJA bei Heck

Kandidaten	DEF	ANIM	AGENS	ADJA
☞ K ₁ : IO[+def,+bel] DO[+def,+bel]				
K ₂ : DO[+def,+bel] IO[+def,+bel]				*!

5.2. Diskussion

Die Hauptprobleme bei den Ansätzen von Choi und Büring waren Untergenerierung und die fehlende Möglichkeit, Markiertheitsgrade syntaktisch abzuleiten; das Hauptproblem bei dem Ansatz von Müller war, daß er nicht-triviale Modifikationen und Zusatzannahmen gegenüber der klassischen Optimalitätstheorie erfordert. Man kann sagen, daß der Ansatz von Heck hier irgendwo in der Mitte liegt. Für das Untergenerierungsproblem ist eine Lösung vorgesehen, nämlich Scrambling auf der S-Struktur. (Da dieses Scrambling dann jeweils obligatorisch durch eine Beschränkung erzwungen wird, hat man es hier dann wiederum mit einem Pseudo-Optionalitätsansatz zu tun.) Allerdings wird hierzu wenig gesagt, und es stellt sich die Frage, ob die Skopus- und Fokus-bezogenen S-Struktur-Beschränkungen tatsächlich ausreichen, für jeden Satz ein optimales Profil abzuleiten, oder ob es nicht doch für manche grammatischen (aber markierten) Sätze schwerfällt, einen S-strukturellen Auslöser zu finden. Die Vermeidung des Untergenerierungsproblems beruht also partiell darauf, daß von vornherein nur ein Teilbereich der relevanten Daten genauer betrachtet wird.

Demgegenüber werden Markiertheitsunterschiede zwar explizit abgeleitet, aber nur durch die Stipulation, daß Scrambling Markiertheit erzeugt (also durch das DTK-basierte Konzept der Markiertheit). Wie bereits Höhle (1982) festgestellt hat, folgt diese Annahme zunächst einmal aus nichts. Dieser Umstand ist bereits für Standard-Syntaxen ein Problem; dort hat man aber kein offensichtlich einfacheres und weniger stipulatives Maß für Unmarkiertheit. Anders liegt der Fall jedoch in der Optimalitätstheorie. Man hat mit dem Konzept der Optimalität *im Prinzip* ein einfaches Maß für Unmarkiertheit; dieses Maß, so scheint mir, sollte man dann auch zu verwenden versuchen. In jedem Fall läßt sich festhalten, daß auch der Ansatz von Heck auf nicht-triviale Zusatzannahmen (wie die DTK-basierte Theorie der Markiertheit) angewiesen ist, die über ein klassisches optimalitätstheoretisches System hinausgehen und ihm zunächst einmal fremd sind.

Zu diesen allgemeinen Erwägungen kommen noch potentielle empirische Probleme. Bei der Annahme von je nach Definitheit, Belebtheit, Agentivität und Kasus unterschiedlichen D-strukturellen Argumentabfolgen bei ein und demselben Verb können nicht nur Schwierigkeiten bei der adäquaten Formulierung von Prinzipien entstehen, die die Abbildung von lexikalisch vermerkten Argumentordnungen auf die Syntax betreffen. Es gibt auch Probleme mit einer strukturellen Behandlung der Bindungsdaten (vgl. (15)), die ja für eine einheitliche D-strukturelle Abfolge direktes Objekt \succ indirektes Objekt sprechen. Heck (2000) nimmt dies zur Kenntnis und schlägt vor, die Bindungstheorie unter Bezug auf nicht-strukturelle Faktoren zu formulieren, wie dies z.B. auch bei Pollard & Sag (1994) gemacht wird.

Ein weiteres empirisches Problem ergibt sich bei der Betrachtung der Daten in (33), die an dieser Stelle zum endgültig letzten Mal in diesem Kapitel wiederholt sind.

- (41) a. daß der Fritz der Maria das BUCH gegeben hat
 b. ?daß der Fritz der MARIA das Buch gegeben hat
 c. ??daß der Fritz das Buch der MARIA gegeben hat
 d. ?*daß der Fritz das BUCH der Maria gegeben hat

Durch die Beschränkung ANIM wird die Abfolge indirektes Objekt \succ direktes Objekt auf der D-Struktur als einzig optimale klassifiziert. Damit werden (41-ab) als weniger markiert klassifiziert als (41-cd). Offengelassen ist aber, wieso die Fokusverschiebung von (41-a) zu (41-b) zu größerer Markiertheit führt. (Die DTK-basierte Theorie der Markiertheit wäre mindestens dadurch zu ergänzen, daß neben Scrambling auch Fokusverschiebung automatisch zu Markiertheit führt.) Und schließlich ist, wie Heck (2000) einräumt, nicht ganz klar, wie der S-strukturelle Auslöser für Scrambling in (41-d) aussieht (das Untergenerierungsproblem). Daß im übrigen mehr als ein Markiertheitsunterschied von (41-ab) gegenüber (41-cd) nicht abgeleitet werden kann, liegt daran, daß Fokus-motivierte Wortstellungsvariation als S-strukturelle Bewegung betrachtet wird und Fokus-bezogene Beschränkungen somit aus prinzipiellen Gründen nicht mit den D-strukturellen Linearisierungsbeschränkungen interagieren können.

6. Konklusion

Hiermit bin ich am Ende des Überblicks über die Versuche angelangt, die deutsche Wortstellung mit den Mitteln der Optimalitätstheorie anzugehen. Dabei sind längst nicht alle einschlägigen Arbeiten abgehandelt worden. Erwähnen möchte ich hier nur noch Lenerz (1999), der sich mit dem Versuch einer optimalitätstheoretischen Erfassung von Wortstellungsregularitäten kritisch auseinandersetzt; Keller (2000), der optimalitätstheoretische und andere Wettbewerbsmodelle der freien Wortstellung im Deutschen in einer experimentellen Untersuchung überprüft; sowie Haftka (1999). In letzterer Arbeit wird das Konzept der Kumulativität vorausgesetzt, zusammen mit der Annahme aus Müller (1999c), daß die die Wortstellung determinierenden Beschränkungen in einer Gruppe zusammengefaßt sind, deren Ordnung anders interpretiert wird als die Ordnung sonstiger Beschränkungen in der Optimalitätstheorie. Konkret schlägt Haftka vor, daß ein Kandidat optimal bzgl. der in einer Gruppe (bzw. "Subhierarchie") zusammengefaßten Wortstellungsbeschränkungen ist, wenn die Gesamtzahl der Verletzungen dieser Beschränkungen im Vergleich aller Kandidaten minimal ist. Haben mehrere Kandidaten gleichermaßen die minimale Verletzungszahl in diesem Bereich, können sie alle grammatisch sein. Unter den so gewonnenen grammatischen Kandidaten wird dann gemäß Kellers Konzept der Suboptimalität relative Markiertheit vorhergesagt: Der grammatische Kandidat mit dem besten Beschränkungsprofil ist unmarkiert, der nächstbeste grammatische Kandidat etwas markierter, usw. Mit der unterschiedlichen Bewertung des Faktors der Kumulativität liegt zwar ein fundamentaler Unterschied gegenüber dem Modell in Müller (1999c) vor;²⁷ und die untersuchte empirische Domäne ist auch eine etwas andere, da Haftka nur mit dem vorderen Bereich des Mittelfeldes beschäftigt ist. Die oben für den Ansatz in Müller (1999c) aufgeworfenen konzeptuellen Schwierigkeiten findet man jedoch im Ansatz von Haftka (1999) alle wieder.

Es ergibt sich somit das Gesamtergebnis, daß eine einfache Anwendung der Standard-Optimalitätstheorie auf die Fakten der deutschen Wortstellung zu Pseudo-Optionalitäts-Analysen führt, die untergenerieren und Markiertheitsgrade nicht grammatisch erfassen können; Analysen, die diese Schwierigkeiten umgehen, scheinen demgegenüber nicht-triviale Modifikationen der Optimalitätstheorie notwendig zu machen, die entweder Unvereinbarkeit mit der existierenden Forschung zur optimalitätstheoretischen Syntax implizieren oder aber das Gesamtsystem einigermaßen komplex werden lassen. Hinzu kommt, daß in den hier besprochenen optimalitätstheoretischen Analysen bestimmte Eigenschaften und Faktoren systematisch ausgeblendet werden, die zumindest manchmal in den zu Beginn dieses Kapitels erwähnten empirisch orientierten Arbeiten auftauchen und als relevant betrachtet werden. Neben der Eigenschaft der Kumulativität (die nur bei Haftka

²⁷Wie oben gezeigt, ist die Idee, daß die Zahl von Beschränkungsverletzungen im SCR-KRIT-Bereich relevant ist, mit der Analyse in Müller (1999c) unvereinbar.

(1999) von Bedeutung und mit der Standard-Optimalitätstheorie inkompatibel ist) ist hier insbesondere zu erwähnen der Faktor der Konstituentenlänge, also Behagels Gesetz der wachsenden Glieder. Dies spielt bei Lenerz (1977) eine untergeordnete Rolle, ist aber bei Primus (1994) einigermaßen prominent. Ungeachtet der Frage, ob dieses Gesetz in gleichem Maße wie die in diesem Kapitel besprochenen Faktoren die Wortstellung beeinflusst, läßt sich jedenfalls festhalten, daß eine direkte optimalitätstheoretische Implementierung dieses Faktors nicht unproblematisch ist (denn die Beschränkung sieht zunächst einmal translokal aus; vgl. Kapitel 4 und Kapitel 8).

Angesichts dieser Sachlage möchte ich den Schluß ziehen, daß eine Behandlung der freien Wortstellung im deutschen Mittelfeld in der Optimalitätstheorie zwar keineswegs unmöglich ist, daß sich eine solche Analyse jedoch nicht in so offensichtlicher Weise anbietet, wie man das vielleicht auf den ersten Blick vermuten mag.

Kapitel 7

Kasus

1. Einleitung

In den vorangegangenen Kapiteln ist nicht viel zur Kasustheorie gesagt worden. Im wesentlichen wurde durchweg präsupponiert, daß eine gegebene NP einen bestimmten Kasus trägt; wie der Kasus auf die NP kommt, ist jedoch mit einer Ausnahme im dunkeln geblieben. Die Ausnahme betrifft Abschnitt 4 von Kapitel 2. Dort wurde gezeigt, daß die Generalisierung, daß im Deutschen im unmarkierten Fall (also wenn nichts Gegenteiliges gefordert wird) eine NP Nominativ erhält, gut optimalitätstheoretisch abzuleiten ist, und zwar durch eine tiefgeordnete Beschränkung DEFAULT-NOM, die auf jeder NP Nominativ verlangt (und die im Normalfall durch höher geordnete Beschränkungen überlagert wird, die anderes verlangen); vgl. (45) aus Kapitel 2. Es ist nun zu beobachten, daß nach der Pionierarbeit von Legendre, Raymond & Smolensky (1993) gerade in jüngster Zeit einige optimalitätstheoretische Arbeiten zur Kasustheorie entwickelt worden sind. Deren Hauptmerkmale vorzustellen ist das Ziel dieses Kapitels.

2. Inhärente vs. strukturelle Kasus: Woolford (1999)

In der klassischen GB-Kasustheorie von Chomsky (1981) wird zwischen inhärentem (lexikalischem) und strukturellem Kasus unterschieden. Die strukturellen Kasus sind für die Argumente des Verbs zunächst einmal der Nominativ und der Akkusativ. Strukturelle Kasus sind nicht an spezielle Argumente gekoppelt; sie werden bestimmten syntaktischen Positionen zugewiesen (vgl. den Nominativ für z.B. ein Thema-Argument im Passiv und den Akkusativ für z.B. ein Agens-Argument in der A.c.I.-Konstruktion). Inhärente Kasus sind demgegenüber an bestimmte Argumente gebunden und im Lexikoneintrag des Verbs vermerkt. Üblicherweise werden im Deutschen der Genitiv und der Dativ bei einfach transitiven Verben als inhärente Kasus betrachtet. Ein offensichtliches Problem entsteht in diesem System dadurch, daß strukturelle und lexikalische Kasuszuweisung oft Unvereinbares verlangen. In (1) etwa erhält das Objekt gemäß struktureller Kasuszuweisung Akkusativ, aber lexikalisch ist ein inhärenter Genitiv für die Thetarolle Thema beim Verb *gedenken* spezifiziert.

(1) daß sie des Mannes/*den Mann gedachte

Wie hier, so setzt sich offenbar systematisch der inhärente gegen den strukturellen Kasus durch. Der vorliegende Beschränkungskonflikt wird bei Chomsky (1981) wie

folgt aufgelöst: Inhärenter Kasus wird auf der D-Struktur vergeben, struktureller Kasus auf der S-Struktur. Ein einmal zugewiesener Kasus kann dann nicht mehr geändert werden. Dies bedeutet jedoch, daß entweder die Regeln für die strukturelle Kasuszuweisung verletzbar sein müssen, oder daß sie optional applizieren. Tatsächlich ist die Formulierung in Chomsky (1981) so, daß diese Regeln nicht optional applizieren. Somit bietet sich in diesem Bereich eine optimalitätstheoretische Rekonstruktion an. Der Ansatz in Woolford (1999) soll dies leisten, und er soll darüber hinaus eine systematische Erfassung von für klassische Kasustheorien notorisch schwierigen Fällen (wie inhärentem Kasus auf Subjekten) aus einer Reihe typologisch verschiedener Sprachen ermöglichen. Ich werde mich hier auf Woolfords Analyse des inhärenten Dativs konzentrieren, die anhand des Isländischen, Japanischen und Englischen entwickelt wird. Eine Grundannahme ist, daß Chomskys (1981) Kasusfilter als unverletzbar Beschränkung gilt: Jede NP muß Kasus tragen. Gegenstand der Optimierung ist nur, welchen Kasus eine NP trägt. Die Beschränkungen, die im folgenden eine Rolle spielen, sind in (2) zusammengefaßt (die Formulierung ist z.T. leicht vereinfacht).

- (2) a. *DAT (“*Dative”):
Vermeide Dativ.
b. *AKK (“*Accusative”):
Vermeide Akkusativ.
c. *NOM (“*Nominative”):
Vermeide Nominativ.
d. LEX-TREUE (“Treue zum lexikalischen Kasus”, “Faith-lex”):
Ein lexikalisch spezifiziertes inhärentes Kasusmerkmal muß realisiert werden.
e. LEX-TREUE_{tr}:
Ein lexikalisch auf einem transitiven Verb spezifiziertes inhärentes Kasusmerkmal muß realisiert werden.

Die Beschränkungen *DAT, *AKK und *NOM sind Markiertheitsbeschränkungen, deren Ordnung als universell fest postuliert wird: *DAT \gg *AKK \gg *NOM. Angenommen wird, daß der Dativ inhärent ist, Akkusativ und Nominativ strukturell. Die Beschränkungen, die strukturellen Kasus verbieten, sind also leichter verletzbar als die Beschränkungen, die inhärenten Kasus verbieten. Da es für andere (wie der Dativ als inhärent klassifizierte) Kasus wie Ergativ und Genitiv ebenfalls hoch geordnete Beschränkungen gibt, die sie verbieten, kann keine NP existieren, ohne eine dieser Kasusbeschränkungen zu verletzen. Da das Nominativverbot die tiefstgeordnete Beschränkung ist, ergibt sich, daß der Nominativ der unmarkierte Kasus ist, den eine NP trägt, wenn nichts anderes erzwungen wird.¹

¹Paradoxerweise hat so das durch die Beschränkung *NOM ausgedrückte Nominativverbot ganz ähnliche Effekte wie das durch die in Kapitel 2 eingeführte Beschränkung DEFAULT-NOM

Diesen Markiertheitsbeschränkungen steht eine Treuebeschränkung gegenüber. LEX-TREUE verlangt, im Lexikon auf einem Verb spezifizierte inhärente Kasusmerkmale in der Syntax zu realisieren. Unterschieden wird zwischen einer allgemeinen Form der Beschränkung und einer universell höher geordneten, damit in Stringenzbeziehung (vgl. (65) aus Kapitel 2) stehenden speziellen Variante LEX-TREUE_{tr}. Aus dieser Ordnung folgt, daß es bei transitiven Verben wichtiger ist als bei intransitiven, lexikalische Kasusmerkmale zu realisieren; d.h., im transitiven Fall überleben inhärente Kassuspezifikationen eher als im intransitiven.

Für das Isländische wird nun die Ordnung in (3-a) postuliert, für das Japanische die Ordnung in (3-b), und für das Englische die Ordnung in (3-c).

- (3) a. *Ordnung im Isländischen:*
 LEX-TREUE_{tr} >> LEX-TREUE >> *DAT >> *AKK >> *NOM
 b. *Ordnung im Japanischen:*
 LEX-TREUE_{tr} >> *DAT >> LEX-TREUE >> *AKK >> *NOM
 c. *Ordnung im Englischen:*
 *DAT >> LEX-TREUE_{tr} >> LEX-TREUE >> *AKK >> *NOM

Wenn auf einem Verb im Lexikon ein inhärentes Dativmerkmal spezifiziert ist, gibt es Konflikte zwischen den Treuebeschränkungen, die Realisierung des Dativs verlangen, und der Markiertheitsbeschränkung *DAT, die diesen Kasus im Output verbietet. Die folgenden Beispiele aus dem Isländischen illustrieren, daß ein lexikalisch spezifizierter Dativ auf dem Subjekt von intransitiven wie transitiven Verben auftreten kann; im transitiven Fall erhält dann das Objekt den Nominativ.

- (4) a. Bátnum hvolfdi
 Boot_{dat} kenterte
 b. Barninu batnaði veikin
 Kind_{dat} genas von Krankheit_{nom}

Betrachten wir zunächst den Fall (4-a). Hier ist im Lexikon vermerkt, daß das einzige Argument des Verbs ein Dativmerkmal trägt. LEX-TREUE verlangt Dativrealisierung, *DAT verbietet Dativrealisierung. Aus der Ordnung der beiden Beschränkungen ergibt sich, daß hier der inhärente Dativ im optimalen Kandidaten erscheint; eine Nominativ- oder Akkusativrealisierung verletzt zwar tiefer geordnete Markiertheitsbeschränkungen als *DAT, aber auch die höher geordnete Treuebeschränkung LEX-TREUE. Dies zeigt Tabelle T_{7.1}.

In den Fällen, wo nicht lexikalisch ein inhärentes Kasusmerkmal spezifiziert ist, wird demgegenüber korrekt vorhergesagt, daß der Kandidat mit dem strukturellen Nominativ auf dem einzigen Argument gewinnt. Im Prinzip dieselbe Argumentation wie für den intransitiven Fall gilt für den transitiven Fall (4-b). Per Annahme ist

ausgedrückte Nominativgebot. Der Grund hierfür ist, daß erstere Beschränkung durch einen unverletzbaren Kasusfilter und höher geordnete Verbote gegen alle anderen Kasus ergänzt ist.

T_{7.1}: Intransitives V im Isländischen mit inhärentem Dativ

Kandidaten	L-TREUE _{tr}	L-TREUE	*DAT	*AKK	*NOM
☞ K ₁ : NP _{dat} V _[+dat]			*		
K ₂ : NP _{nom} V _[+dat]		*!			*
K ₃ : NP _{akk} V _[+dat]		*!		*	

hier im Lexikon vermerkt, daß V zwei Argumente hat (die wir für die gegenwärtigen Zwecke <Experienter, Thema> nennen können), und daß dem ersten Argument (das in der Syntax als Subjekt erscheint) ein Dativmerkmal eignet. Der optimale Kandidat realisiert dieses Dativmerkmal auf dem Experienter und respektiert dadurch LEX-TREUE und LEX-TREUE_{tr} (was nunmehr nicht mehr leer von allen Kandidaten erfüllbar ist), unter Verletzung von *DAT. Für das Thema ist kein inhärenter Kasus vorgezogen. Daher entscheiden allein Markiertheitsbeschränkungen, und aus denen ergibt sich für das zweite Argument der am wenigsten markierte Kasus: Nominativ. Der Wettbewerb wird illustriert durch Tabelle T_{7.2}.

T_{7.2}: Transitives V im Isländischen mit inhärentem Dativ auf dem Subjekt

Kandidaten	L-TREUE _{tr}	L-TREUE	*DAT	*AKK	*NOM
☞ K ₁ : NP _{dat} V _[+dat] NP _{nom}			*		*
K ₂ : NP _{dat} V _[+dat] NP _{acc}			*	*!	
K ₃ : NP _{nom} V _[+dat] NP _{acc}	*!	*		*	*

Das Japanische unterscheidet sich vom Isländischen dadurch, daß ein inhärenter Dativ zwar bei transitiven Verben auftreten kann, aber nicht bei intransitiven:

- (5) a. Akatyan-ga/*-ni moo arukeru
 Baby_{nom/dat} schon gehen kann
 b. Taroo-ni eigo-ga hanaseru
 Taro_{dat} Englisch_{nom} sprechen kann

Das Verbot von inhärentem Kasus auf Subjekten intransitiver Verben im Japanischen ergibt sich daraus, daß gegenüber dem Isländischen *DAT und LEX-TREUE umgeordnet sind (und die höher geordnete Beschränkung LEX-TREUE_{tr} hier irrelevant ist). Dies zeigt Tabelle T_{7.3}.

T_{7.3}: Intransitives V im Japanischen mit inhärentem Dativ

Kandidaten	L-TREUE _{tr}	*DAT	L-TREUE	*AKK	*NOM
K ₁ : NP _{dat} V _[+dat]		*!			
☞ K ₂ : NP _{nom} V _[+dat]			*		*
K ₃ : NP _{akk} V _[+dat]			*	*!	

Der Gewinner dieses Wettbewerbs ist also ein Neutralisierungskandidat: Derselbe optimale Output ergibt sich natürlich auch ohne Treueverletzung, wenn von

vornherein kein inhärentes Kasusmerkmal spezifiziert ist.²

Die Möglichkeit von inhärentem Kasus auf dem Subjekt transitiver Verben im Japanischen ergibt sich nun daraus, daß die Beschränkung *DAT zwar LEX-TREUE dominiert, aber ihrerseits von der spezifischeren Beschränkung LEX-TREUE_{tr} dominiert wird (tatsächlich motiviert das unterschiedliche Verhalten von transitiven und intransitiven Verben im Japanischen bei Woolford (1999) erst die Aufteilung in eine allgemeine und eine speziellere Form von LEX-TREUE). Der Wettbewerb ist in diesem Fall also analog zu dem in Tabelle T_{7.2}; vgl. Tabelle T_{7.4}.

T_{7.4}: *Transitives V im Japanischen mit inhärentem Dativ auf dem Subjekt*

Kandidaten	L-TREUE _{tr}	*DAT	L-TREUE	*AKK	*NOM
☞K ₁ : NP _{dat} V _[+dat] NP _{nom}		*			*
K ₂ : NP _{dat} V _[+dat] NP _{akk}		*		*!	
K ₃ : NP _{nom} V _[+dat] NP _{akk}	*!		*	*	*

Im Englischen ist nun *DAT höher geordnet als LEX-TREUE und LEX-TREUE_{tr}. Damit kann in dieser Sprache bei intransitiven wie bei transitiven Verben niemals ein inhärenter Dativ erscheinen, auch wenn er auf einem Verb im Input spezifiziert ist. Daß dies im Englischen so ist, muß also nicht auf eine lexikalische Stipulation (wie “Verben können im Englischen keinen inhärenten Kasus zuweisen”) zurückgeführt werden; es folgt aus der postulierten Beschränkungsordnung.

Da unter den Markiertheitsbeschränkungen *DAT, *AKK und *NOM ebenso wie unter den Treuebeschränkungen LEX-TREUE und LEX-TREUE_{tr} feste Ordnungsbeziehungen herrschen, ist durch die drei beschriebenen Sprachen die faktorielle Typologie in diesem Bereich erschöpft: Eine Unterordnung von LEX-TREUE oder LEX-TREUE_{tr} unter *AKK oder *NOM macht gegenüber dem Englischen keine neuen Vorhersagen mehr, da die Treuebeschränkungen nur für inhärenten Kasus relevant sind und Akkusativ und Nominativ per Annahme strukturelle Kasus sind.

Über die systematische Unterdrückung inhärenter Kasus hinaus macht Woolfords System noch eine interessante Vorhersage für das Englische und andere Sprachen: Bei unakkusativischen intransitiven Verben, deren einziges Argument Objekteigenschaften aufweist, kann nie der typische Objektkasus, Akkusativ, realisiert werden (vgl. “Burzios Generalisierung”); vielmehr tritt wie bei unergativen intransitiven Verben der Nominativ auf. Dies liegt an der fixen Ordnung *AKK ≫ *NOM und ist in Tabelle T_{7.5} illustriert; die Gesamtordnung ist hier aus Illustrationsgründen die des Englischen.

²Dementsprechend wird hier wieder das Konzept der Input-Optimierung relevant: Eine [+dat]-Spezifikation auf intransitivem V im Input kann im Japanischen nie im Output optimal werden, und so kann der Sprachlerner durch das Meta-Verfahren der Input-Optimierung rückschließen, daß eine solche Inputspezifikation in der Sprache vorn vornherein nicht betrachtet werden muß.

*T*_{7.5}: *Intransitives unakkusativisches V*

Kandidaten	*DAT	L-TREUE _{tr}	L-TREUE	*AKK	*NOM
☞ K ₁ : NP _{nom} V					*
K ₂ : NP _{akk} V				*!	

Wenn ein Verb zwei Argumente, aber keine inhärente Kasusspezifikation hat, tritt normalerweise ein Nominativ-Akkusativ-Muster auf.³ Es stellt sich die Frage, warum das so ist, denn zunächst einmal sollte eine zweifache Verwendung des Nominativs ein besseres Beschränkungsprofil haben. Woolfords Antwort ist hier, daß von vornherein nur Kandidaten im Wettbewerb sind, die pro Satz nicht mehr als einen Nominativ aufweisen.⁴ Wenn dem so ist, und wenn es keine vergleichbare Restriktion für Akkusative gibt, dann ist die Vorhersage, daß in einer Doppel-Objekt-Konstruktion mit einem ditransitiven Verb ohne lexikalische Kasusspezifikation der optimale Kandidat ein Nominativ-Akkusativ-Akkusativ-Muster aufweist. Woolford nimmt an, daß das für das Englische korrekt ist. Für Sprachen wie das Japanische (oder Deutsche), die in diesem Kontext typischerweise ein Nominativ-Dativ-Akkusativ-Muster realisieren, könnte man dann einen lexikalisch auf dem dritten Argument (Ziel, Quelle, etc.) vermerkten Dativ annehmen. Das erklärt dann aber noch nicht, warum es im Japanischen überhaupt keine Doppel-Akkusativ-Konstruktionen gibt (und im Deutschen nur wenige, wie in *jemanden eine Sprache lehren*, wo ein Akkusativ offensichtlich inhärent ist). Um dies generell auszuschließen, postuliert Woolford (1999) eine Beschränkung OCP_{Kas} (“Obligatory Countour Principle_{Case}”), derzufolge zwei Instanzen desselben Kasus auf Ko-Argumenten zu vermeiden sind. Diese an das in der Phonologie geltende “Obligatory Contour Principle” angelehnte Beschränkung hat eine spezielle Variante nur für den Akkusativ, und letztere ist im Japanischen (anders als im Englischen) sehr hoch geordnet, so daß in Doppel-Objekt-Konstruktionen nur ein Nominativ-Dativ-Akkusativ-Muster entstehen kann. Wie Woolford bemerkt, legt dies den Schluß nahe, daß auch der Dativ ein struktureller Kasus sein kann (denn ist er bei ditransitiven Verben immer inhärent am Verb markiert, braucht man keine weitere Erklärung mehr für das Fehlen des Nominativ-Akkusativ-Akkusativ-Musters).

Ein letztes allgemeines Problem ist noch zu klären: Wie ist in diesem System eigentlich gewährleistet, daß bei einfach transitiven Verben ohne inhärente Kasusspezifikation ein struktureller Nominativ an das höhere und ein struktureller Akkusativ an das tiefere Argument geht? Eine Standardlösung etwa gemäß Chomsky

³Woolford behandelt den Ergativ in Ergativsprachen als inhärenten Kasus, und den Unterschied zwischen “echten” Ergativsprachen und “gespaltene[n]” Ergativsprachen (bzw. “aktiven” Sprachen) in Analogie zum Unterschied bzgl. des Dativs zwischen Japanisch und Isländisch. Der Absolutiv wird dem Nominativ gleichgesetzt.

⁴Mehrfache Nominative z.B. im Japanischen sind dann zunächst einmal ein Problem.

(1981) ist die, daß bei einem Verb mit den Thetarollen <Agens,Thema> das Agens in der Subjektposition realisiert wird, das Thema in der Objektposition; im Normalfall erhält die Subjektposition dann Nominativ zugewiesen, die Objektposition Akkusativ. Ganz so einfach kann der Fall im vorliegenden System nicht sein, denn wir haben gesehen, daß im Isländischen und Japanischen struktureller Nominativ in einer tieferen, VP-internen Position realisiert werden kann, bei gleichzeitiger Realisierung eines inhärenten Kasus in der Subjektposition. Die korrekte Verteilung der Kasus an die Argumente muß hier vielmehr so vonstatten gehen, daß immer das höchste nicht inhärent markierte Argument den Nominativ erhält, das tiefste den Akkusativ, und ein eventuelles drittes Argument, das syntaktisch in eine mittlere Position abgebildet wird, den Dativ (in Sprachen, wo OCP_{Kas} dies erzwingt). Die Konklusion hieraus ist, daß Woolfords Ansatz zwar qua Optimierung erklären kann, welche Kasus in einem Satz vorkommen (und wie inhärente und strukturelle Kasus interagieren), daß er jedoch nicht erklären will, wie bei Optimalität mehrerer struktureller Kasus diese ihre Argumente finden.

In Woolfords (1999) Ansatz lassen sich einige der typischen Merkmale identifizieren, die eine optimalitätstheoretische Analyse stützen (vgl. Kapitel 2): Es treten zwischen den Beschränkungen für inhärenten Kasus und den Beschränkungen für strukturellen Kasus fast unvermeidbar Konflikte auf. Diese werden manchmal zugunsten des inhärenten Kasus entschieden (dann sehen wir inhärenten Kasus an der syntaktischen Oberfläche), manchmal zugunsten des strukturellen Kasus (so bei intransitiven Verben im Japanischen, und durchweg im Englischen). Durch eine feste Ordnung unter den Markiertheitsbeschränkungen für die unterschiedlichen Kasus wird gewährleistet, daß bei Fehlen lexikalischer Spezifikationen immer strukturelle Kasus den inhärenten Kasus vorgezogen werden, und daß unter den strukturellen Kasus im unmarkierten Fall immer der Nominativ erscheint. Schließlich hat sich gezeigt, daß im Bereich der Kasussyntax neben Beschränkungskonflikten und dem Konzept des unmarkierten Falles auch noch die parametrische Variation eine optimalitätstheoretische Analyse stützt: Durch die relative Ordnung von zwei Treuebeschränkungen für die Realisierung lexikalischer Kasusmerkmale und einer Beschränkung, die Dativ verbietet, läßt sich das unterschiedliche Auftreten des inhärenten Dativs im Isländischen, Japanischen und Englischen auf einfache Weise erfassen.

3. Nochmals inhärente vs. strukturelle Kasus: Fanselow (1999)

Auf der Basis u.a. deutscher Daten verfolgt Fanselow (1999) ganz ähnliche Ziele wie Woolford, und seine Grundausrichtung ist auch vergleichbar. Im Einzelfall sind die Vorschläge jedoch unterschiedlich. Zunächst einmal postuliert Fanselow wie Woolford, daß die Anforderung an eine NP, Kasus zu tragen, unverletzbar ist, und daß von Gen alle möglichen Kasuskombinationen in einem Satz erzeugt werden, von denen dann bei Fehlen inhärenter Kasuspezifikationen die am wenigsten markierte Variante als optimal gewählt wird. Anders als Woolford versucht Fanselow

low jedoch, auf eine extern stipulierte Ordnung der Art $*DAT \gg *AKK \gg *NOM$ zu verzichten. Vielmehr schlägt er vor, die Kasus nicht als Primitive zu sehen, sondern als Kombinationen von Merkmalen. So ist der Nominativ der “Nullkasus”, der nur das abstrakte Merkmal [Kasus] aufweist. Der Akkusativ hat demgegenüber ein weiteres Merkmal f ([Kasus,f]), und der Dativ noch ein zusätzliches Merkmal g ([Kasus,f,g]).⁵ Somit folgt, daß in gewisser Weise der Akkusativ komplexer ist als der Nominativ, und der Dativ komplexer als der Akkusativ. Um die Effekte der Woolfordschen Beschränkungshierarchie abzuleiten, benötigt man nun noch eine einzige Markiertheitsbeschränkung $*KASUS$, die das Auftreten von Kasusmerkmalen (also [Kasus], [f] und [g]) bestraft. Daß somit bei einem transitiven Verb ohne inhärente Kasusspezifikation ein Nominativ-Akkusativ-Muster optimal wird, zeigt Tabelle T_{7.6}.⁶

T_{7.6}: *Transitives V*

Kandidaten	$*KASUS$
\textcircled{K}_1 : NP _{nom} NP _{akk} V	* **
K ₂ : NP _{nom} NP _{dat} V	* ***!
K ₃ : NP _{akk} NP _{dat} V	** **!*

Ebenfalls wie Woolford nimmt Fanselow an, daß es bei Präsenz inhärenter Kasusspezifikationen auf einem Verb unvermeidlich zu Konflikten mit dem strukturellen Kasussystem kommt, und daß der resultierende Wettbewerb durch ein Optimierungsverfahren aufgelöst wird. Wie gesehen schlägt Woolford vor, daß die Realisierung eines gegebenen inhärenten Kasus, z.B. des Dativs, durch LEX-TREUE gefordert wird, und daß die Ordnung dieser Treuebeschränkung über der Markiertheitsbeschränkung, die den Dativ verbietet, für das Auftreten dieses Kasus sorgt. Fanselow lehnt eine derartige Interaktion ab, weil er davon ausgeht, daß es nur Evidenz für die Ordnung LEX-TREUE $\gg *KASUS$ gibt, und nicht für die umgekehrte Ordnung $*KASUS \gg$ LEX-TREUE (gemäß der in einer Sprache alle inhärenten Kasus blockiert werden) oder für eine Kopplung der beiden Beschränkungen (gemäß der die Wahl zwischen inhärentem und strukturellem Kasus frei sein und Optimalität entstehen sollte). Seine Konklusion ist, daß sich inhärente Kasus immer gegen strukturelle Kasus durchsetzen, und daß dieser Wettbewerb nicht durch eine Standardoptimierung mit LEX-TREUE als geordneter, verletzbarer Beschränkung

⁵Die Natur der Merkmale f und g wird offen gelassen, aber vgl. die Bemerkungen zu Wunderlich (2000) und Stiebels (2000) unten.

⁶Wie vorher stellt sich die Frage, warum nicht ein Kandidat mit mehrfachem Nominativ gewinnen kann. Die Antwort ist, daß bei Fanselow jeder strukturelle Kasus durch einen funktionalen (VP-externen) Kasusuweiser legitimiert werden muß, so daß Kandidaten mit mehreren Instanzen eines strukturellen Kasus im selben minimalen Satz in den betrachteten Sprachen nicht generiert werden.

entschieden wird, sondern durch das Optimierungsprinzip der Blockade, das bereits in Kapitel 4 besprochen wurde und hier wiederholt ist (vgl. (37) aus Kapitel 4):

(6) *Blockade:*

Wenn bei zwei miteinander im Wettbewerb befindlichen sprachlichen Formen F_1 , F_2 gilt, daß F_1 spezifischer ist als F_2 , dann blockiert F_1 F_2 als ungrammatisch.

Die Annahme ist, daß eine Form mit lexikalisch determinierter Kasusmarkierung immer spezifischer ist als eine Form mit rein struktureller Kasusmarkierung. Daher blockiert der inhärente Kasus immer den strukturellen Kasus. Wie Blockadeprinzip und Optimalitätstheorie interagieren, wird bei Fanselow (1999) nicht ausgeführt. Es scheint aber klar, daß zwischen zwei Optimierungsverfahren unterschieden werden muß: Der Standardoptimierung muß ein Blockadeverfahren entweder vorangehen oder folgen. Da bei letzterer Variante zunächst sichergestellt werden müßte, daß die Form mit strukturellem Kasus und die Form mit inhärentem Kasus beide optimal sind, und da dies Schwierigkeiten für die Definition der Kandidatenmenge aufwirft, ist erstere Variante vermutlich die erwünschte: Gen erzeugt die Kandidaten, und vor der eigentlichen Optimierung entfernt das Blockadeprinzip bereits all jene Kandidaten, die den inhärenten Kasus zugunsten eines strukturellen Kasus aufgegeben haben. Die nachfolgende Optimierung wählt dann bei z.B. einem inhärenten Dativ bei einem transitiven Verb nur noch aus zwischen einem Kandidaten mit Nominativ-Dativ-Muster und einem anderen Kandidaten mit Akkusativ-Dativ-Muster.

Eine solche Modifikation des Grammatikmodells mit zwei unterschiedlichen Optimierungsverfahren ist ein radikaler Schritt, und man kann sich angesichts der Tatsache, daß Blockade-Analysen typischerweise optimalitätstheoretisch rekonstruierbar sind (vgl. Kapitel 4), fragen, ob er durch die empirische Evidenz gerechtfertigt ist. Tatsächlich hat ja Woolford (1999) dafür argumentiert, daß LEX-TREUE im Prinzip z.B. unter *DAT geordnet sein kann, so daß in manchen Sprachen (Englisch) und in bestimmten Konstruktionen anderer Sprachen (Japanisch) eigentlich erwartbare inhärente Kasus niemals auftauchen. Anstatt diese Frage weiter zu verfolgen, möchte ich mich nun aber drei Generalisierungen in Fanselow (1999) zuwenden, die vor allem die Eigenschaften von inhärentem Kasus im Deutschen betreffen.

Fanselow nimmt an, daß im Deutschen die Muster in (7-a-c) die regulären für intransitive, transitive und ditransitive Verben sind, d.h., daß hier jeder Kasus strukturell ist. Die Muster in (7-d-i) sind demgegenüber das Resultat inhärenter Kassusspezifikation, die jeweils den strukturellen Kasus blockiert.

- (7) a. daß der Fritz_{nom} weint
 b. daß die Maria_{nom} das Buch_{akk} liest
 c. daß die Maria_{nom} dem Fritz_{dat} ein Buch_{akk} gibt
 d. daß mich_{akk} friert
 e. daß mir_{dat} davor graut

- f. daß die Maria_{nom} dem Fritz_{dat} hilft
- g. daß die Maria_{nom} des Mannes_{gen} gedachte
- h. daß die Maria_{nom} den Mann_{akk} des Verbrechens_{gen} beschuldigt
- i. daß die Maria_{nom} den Fritz_{akk} diese Sprache_{akk} lehrt

Bestimmte andere Muster scheinen jedoch im Deutschen systematisch zu fehlen. Erstens kann inhärenter Kasus zwar bei intransitiven Verben auftreten (vgl. (7-de)), aber nicht, wenn das Verb die Thetarolle Agens vergibt. Zweitens sieht es so aus, als könne immer nur höchstens ein Argument des Verbs inhärenten Kasus erhalten; die anderen Argumente bekommen strukturellen Kasus zugewiesen. Drittens schließlich, und dies ist eine Verschärfung der zweiten Generalisierung, muß offenbar ein inhärenter Kasus immer dem niedrigsten Argument eines Verbs zugeordnet werden. Fanselow macht nun eine ungewöhnliche Annahme: Anders als bei Chomsky (1981), Woolford (1999) und der großen Mehrheit von Arbeiten zu dieser Thematik wird bei ihm nicht davon ausgegangen, daß ein im Lexikon vermerkter inhärenter Kasus explizit einer Thetarolle zugeordnet ist. Vielmehr steht ein inhärentes Kasusmerkmal im Lexikon einfach so am Verb; daß der inhärente Kasus dann korrekt einem Argument zugewiesen wird, ist auf der Basis der unabhängig lexikalisch vermerkten Thetahierarchie bei jedem Prädikat durch die Syntax zu gewährleisten.

Die Ableitung der ersten Generalisierung ergibt sich nun in einfacher Weise auf der Basis der Annahme, daß ein Agens nicht von V, sondern von einem VP-externen Kopf *v* eingeführt wird: Wenn ein inhärentes Kasusmerkmal [K] im Lexikon auf $V_{[K]}$ vermerkt ist, ein agentive NP jedoch im Spezifikator von *v* erzeugt wird, ergibt sich keine lokale Beziehung von $V_{[K]}$ und NP_{Agens} , und NP_{Agens} kann nicht den inhärenten Kasus K zugewiesen bekommen.

Die Ableitung der zweiten Generalisierung macht sich die konzeptuelle Vereinfachung zunutze, daß es keine a priori stipulierte Beziehung von Thetarolle und inhärentem Kasus im Lexikoneintrag mehr gibt. Angenommen, es gäbe z.B. bei einem ditransitiven Verb zwei inhärente Kassuspezifikationen. Da per Voraussetzung nicht im Lexikon festgelegt werden kann, welcher inhärente Kasus welcher Thetarolle zugewiesen wird, ist in diesem Fall die eindeutige Zuordnung von Kasus und Thetarolle unterminiert, und es entsteht Ambiguität. Betrachten wir ein abstraktes Beispiel. Angenommen, es gäbe im Lexikon ein ditransitives Verb BESCHULDIGEN, das sich von dem tatsächlich existierenden Verb *beschuldigen* dadurch unterscheidet, daß es zwar ein Thetaraster wie <Patiens,Thema> besitzt (das Agens wird ja in diesem System erst durch *v* eingeführt), aber nicht ein inhärentes Kasusmerkmal [gen], sondern zwei inhärente Kasusmerkmale [dat], [gen]. Dann wäre die Zuordnung von Kasus und Thetarolle uneindeutig, und wir würden sowohl (8-a) wie (8-b) erwarten.⁷

⁷Dativ und Genitiv müssen hier beide inhärent sein, denn ansonsten würde *KASUS einen Akkusativ durchsetzen.

- (8) a. *daß die Maria_{nom} dem Mann_{dat} des Verbrechens_{gen} BESCHULDIGT
 b. *daß die Maria_{nom} des Mannes_{gen} dem Verbrechen_{dat} BESCHULDIGT

Fanselow schlägt vor, daß derartige Ambiguitäten von natürlichen Sprachen nicht toleriert werden. Dies können wir explizit als Beschränkung notieren:

- (9) KAS- θ -EIN:

Die Zuordnung von Kasus und Thetarolle muß bei jedem Verb eindeutig sein.

KAS- θ -EIN hat einen ähnlichen Status wie WIED bei Pesetsky (1997; 1998); insbesondere ist diese Beschränkung offenbar undominiert. Angenommen sei nun, daß anstatt auf Fanselows Blockade-Erklärung für die Präferenz von inhärentem Kasus gegenüber strukturellem Kasus dieser Effekt im Sinne von Woolford (1999) auf eine Ordnung LEX-TREUE \gg *KASUS zurückführbar ist. Dann können wir Fanselows Erklärung für die Nicht-Existenz von Verben wie BESCHULDIGEN wie in Tabelle T_{7.7} gezeigt explizit machen: LEX-TREUE wird zugunsten einer Erfüllung von KAS- θ -EIN verletzt (vgl. K₃ vs. K₁, K₂), allerdings nur minimal, so daß hin zum Muster des Verbs *beschuldigen* neutralisiert wird, aber nicht ein vollkommen reguläres Muster mit Dativ und Akkusativ erscheinen kann (vgl. K₃ vs. K₄).⁸

T_{7.7}: Das Verbot mehrerer inhärenter Kasus

Kandidaten	KAS- θ -EIN	LEX-TREUE	*KASUS
K ₁ : NP _{1_{nom}} NP _{2_{dat}} NP _{3_{gen}} V _[+dat,+gen]	*!		* *** **
K ₂ : NP _{1_{nom}} NP _{2_{gen}} NP _{3_{dat}} V _[+dat,+gen]	*!		* *** **
☞ K ₃ : NP _{1_{nom}} NP _{2_{akk}} NP _{3_{gen}} V _[+dat,+gen]		*	* ** **
K ₄ : NP _{1_{nom}} NP _{2_{dat}} NP _{3_{akk}} V _[+dat,+gen]		**!	* *** **

Dies erklärt, warum es nur einen inhärenten Kasus pro Verb geben kann; der optimale Kandidat bei einem Verb mit zwei inhärenten Kasus ist immer identisch mit einem optimalen Kandidaten bei einem Verb mit nur einem inhärenten Kasus. Was damit noch nicht abgeleitet ist, ist die Verschärfung dieser Beobachtung zur dritten der oben angeführten Generalisierungen, derzufolge ein inhärenter Kasus im Deutschen immer an das tiefste Argument gehen muß; vgl. (10-a) vs. (10-b).⁹

- (10) a. daß die Maria_{nom} den Mann_{akk} des Verbrechens_{gen} beschuldigt
 b. *daß die Maria_{nom} das Verbrechen_{akk} des Mannes_{gen} beschuldigt

⁸Hier stehen die NP-Indizes 1, 2, 3 für die Thetarollen Agens, Patiens, Thema (respektive). Angenommen ist weiterhin, daß der Genitiv ebenso viele *KASUS-Verletzungen hat wie der Dativ. Daß der Genitiv im V-System nicht strukturell sein kann, muß dann daran liegen, daß im Unterschied zum Dativ im modernen Deutsch ein entsprechender funktionaler Kopf fehlt.

⁹Wenn man sich vorstellen kann, daß ein Verbrechen beschuldigt wird, gibt es natürlich eine hier irrelevante Lesart von (10-b), bei der der Genitiv *des Mannes* vom Nomen *Verbrechen* zugewiesen wird und in der der Satz wohlgeformt ist.

Die naheliegende Lösung, (10-b) auszuschließen, besteht darin, lexikalisch der Thetarolle, die syntaktisch durch *des Verbrechens* realisiert wird, fest den Genitiv zuzuordnen. Eine derartige Stipulation soll bei Fanselow (1999) ja aber gerade vermieden werden, und sie würde auch nicht nur wie gesehen die Ableitung des Verbots mehrerer exzeptioneller Kasus unterminieren, sondern darüber hinaus auch nicht erklären können, wieso der inhärente Kasus *generell* an das tiefstmögliche Argument eines Verbs geht. Fanselows Lösung beruht auf der Annahme, daß struktureller Kasus immer durch VP-externe funktionale Köpfe zugewiesen wird, während inhärenter Kasus als einziger durch V kontrolliert wird. Wie üblich wird postuliert, daß Thetarollen in der Syntax auf Argumente gemäß ihrer Ordnung im Lexikoneintrag des Verbs abgebildet werden; für (10) ergibt sich somit die D-Struktur (11).¹⁰ Bei lexikalischer Spezifikation von *beschuldigen* für ein Genitivmerkmal muß der optimale Kandidat die Kasus Nominativ, Akkusativ und Genitiv aufweisen, und das Agens kann nicht inhärenten Kasus durch V erhalten. Die Frage ist jedoch, warum das Patiens hier Akkusativ zugewiesen bekommen muß, und das Thema den inhärenten Genitiv.

(11) $[_{VP} \text{ Agens } [_{VP} \text{ Patiens } [_{V'} \text{ Thema } V_{[+gen]}]] v]$

Die VP-externen Kasuspositionen für strukturellen Nominativ und Akkusativ können wir für die gegenwärtigen Zwecke mit Chomsky (1995) als Spezifikator von I und zweiten Spezifikator (bzw. Adjunktposition) von v identifizieren. Was nun benötigt wird, ist ein Prinzip, demzufolge die postulierten Kasusbewegungen die ursprüngliche Abfolge der Argumente nicht durcheinanderbringen dürfen, so daß das am höchsten stehende Agens in die höchste Kasusposition (SpecI) bewegt wird, wo es Nominativ erhält, das in der Mitte stehende Patiens in die zweite externe Kasusposition (Specv bzw. Adj-vP) bewegt wird, wo es Akkusativ erhält, und das zunächst am tiefsten stehende Thema in der VP verharret, wo es dann inhärenten Genitiv erhält. Als Prinzip, das dieses leistet, identifiziert Fanselow (1999) die Beschränkung PAR-VER ("Parallelverschiebung") aus Müller (1999a), die in Kapitel 2 im Zusammenhang mit W-Bewegung im Bulgarischen und Englischen schon eingeführt worden ist und hier wiederholt wird:

(12) PAR-VER:

Wenn $\alpha \beta$ auf der Ebene L_n c-kommandiert, dann c-kommandiert $\alpha \beta$ auch auf der Ebene L_{n+1} (wobei α, β Argumente sind).

Wir haben in Kapitel 2 gesehen, daß PAR-VER verletzbar und relativ tief geordnet sein muß; bereits in einfachen Fragesätzen mit einem W-Objekt, das sich über ein Nicht-W-Subjekt bewegt, ist die Beschränkung verletzt. Wenn sich aber zwei Kandidaten hinsichtlich aller höher geordneten Beschränkungen gleich verhalten, wird

¹⁰Die lineare Einordnung von v und V spielt für das folgende im Grunde keine Rolle; die Abfolge ist nur eine von verschiedenen Möglichkeiten, das Deutsche zu analysieren.

PAR-VER relevant und sagt die korrekte Abfolge voraus. Genau dieser Fall liegt vor in (10), wo die beiden Kandidaten hinsichtlich höher geordneter Beschränkungen wie LEX-TREUE (Genitiv wird realisiert) und *KASUS (Nominativ und Akkusativ werden verwendet) ein gleiches Profil haben und PAR-VER auch bei Tiefstordnung den Unterschied erfaßt. Dies zeigt Tabelle T_{7.8}. Hier entsprechen K₁ und K₂ den beiden Sätzen in (10); K₂ verletzt fatal PAR-VER, weil das D-strukturell tiefere Thema-Argument auf der S-Struktur in die Akkusativposition bewegt wird und das D-strukturell höhere Patiens-Argument in der VP von V Genitiv erhält. K₃ verletzt durch Befolgen des regulären Nominativ-Dativ-Akkusativ-Musters fatal LEX-TREUE. K₄ schließlich verletzt durch Wahl eines Nominativ-Dativ-Genitiv-Musters (mit strukturellem Dativ) fatal *KASUS.

T_{7.8}: Das Verbot von nicht-tiefstem inhärentem Kasus

Kandidaten	LEX-TREUE	*KASUS	PAR-VER
☞ K ₁ : NP _{1_{nom}} NP _{2_{akk}} t ₁ t ₂ NP _{3_{gen}} V _[+gen]		* ** ***	
K ₂ : NP _{1_{nom}} NP _{3_{akk}} t ₁ NP _{2_{gen}} t ₃ V _[+gen]		* ** ***	*!
K ₃ : NP _{1_{nom}} NP _{2_{dat}} NP _{3_{akk}} t ₁ t ₂ t ₃ V _[+gen]	*	* *** **	
K ₄ : NP _{1_{nom}} NP _{2_{dat}} t ₁ t ₂ NP _{3_{gen}} V _[+gen]		* *** ***!	

Da PAR-VER eine Beschränkung ist, die etwas über das Verhältnis von D-Struktur und S-Struktur sagt, setzt diese Analyse zunächst einmal voraus, daß im Deutschen offene Kasusanhebung existiert; tatsächlich hat das Deutsche ja aber eine relativ freie Wortstellung, die durch eine Reihe von Faktoren determiniert zu werden scheint, von denen Kasus nur einer ist (vgl. Kapitel 6). Fanselow (1999) diskutiert einige Möglichkeiten, dieses Problem zu umgehen. Dazu zählen die Kopiertheorie der Bewegung zusammen mit einer Theorie der optimalen PF-Realisierung im Sinne von Pesetsky (1997; 1998) (derzufolge K₁ in T_{7.8} zwar einzig optimal ist, aber die Aussprache für jede bewegte NP auch an der Stelle der Spur erfolgen kann), sowie gemäß Chomsky (1995) die Annahme bloßer Merkmalsbewegung.¹¹

Auf diese (und weitere von Fanselow ins Auge gefaßte) Komplikationen möchte ich hier jedoch nicht mehr eingehen. Abschließend sei lediglich darauf hingewiesen, daß die drei nunmehr abgeleiteten Generalisierungen zwar für das Deutsche adäquat scheinen, zum Teil für andere Sprachen jedoch Schwierigkeiten bereiten können. Was die erste Generalisierung betrifft, so muß man sicherstellen, daß Dativ-

¹¹Eine weitere Möglichkeit wäre, daß die Output-Kandidaten in T_{7.8} nicht S-Struktur-Repräsentationen sind, sondern NP-Struktur-Repräsentationen im Sinne von Riemsdijk & Williams (1981). Die NP-Struktur ist eine Repräsentationsebene zwischen D-Struktur und S-Struktur, auf der zwar Kasusbewegung erfolgt ist, aber noch keine anderen Bewegungen; letztere (wie W-Bewegung oder Scrambling) applizieren unter dieser Perspektive auf dem Weg zur S-Struktur und können unter Verletzung von PAR-VER die auf der NP-Struktur festgelegte Argumentordnung wieder umstoßen.

Subjekte im Japanischen nie agentiv sind (tatsächlich ist ja das Dativ-Argument in (5-b) auch kein Agens); ebenso kann man gemäß dieser Generalisierung Woolfords (1999) Einschätzung nicht folgen, daß der Ergativ ein inhärenter Kasus ist (denn der Ergativ kookkurriert oft mit Agentivität). Zumindest auf den ersten Blick problematischer ist die dritte Generalisierung angesichts von Beispielen mit einem inhärenten Dativ auf dem Subjekt und einem strukturellen Nominativ auf dem Objekt im Isländischen; vgl. (4-b). Hier weist Fanselow jedoch zurecht darauf hin, daß die Generalisierungen nicht per se gelten, sondern nur insoweit, wie sie durch die Theorie ableitbar sind. Fälle wie (4-b) sind unter dieser Perspektive das Ergebnis intervenierender Faktoren; konkret scheinen erstens bei bestimmten Verben lexikalisch markierte Restriktionen für strukturellen Nominativ relevant, die als hoch geordnete Beschränkungen kodierbar sind und den Nominativ in eine Verb-nahe Position zwingen, und zweitens eine Beschränkung, die Füllung der Subjektposition durch eine (an sich beliebige, hier Dativ-) NP verlangt (SUBJEKT bzw. EPP).

4. Kasus in freien Relativsätzen: Vogel (1999)

Anders als Woolford und Fanselow beschäftigt sich Vogel (1999) nicht vorrangig mit Kasuskonflikten im einfachen Satz; er nimmt an, daß Kasuszuweisung mehr oder weniger so wie in der klassischen GB-Theorie erfolgt: Eine NP in der lokalen Domäne eines Kopfes α (I, V) erhält von α den Kasus, für den α spezifiziert ist. Viele der Fragen, die Woolford und Fanselow beschäftigen (z.B., warum inhärenter Kasus strukturellem Kasus vorgezogen wird, warum im Zweifelsfall der Nominativ dem Akkusativ vorgezogen zu werden scheint – vgl. u.a. Burzios Generalisierung –, warum inhärenter Kasus in manchen Sprachen häufiger auftritt als in anderen und in manchen Sprachen ganz fehlt, oder warum inhärenter Kasus pro Verb immer auf ein Argument beschränkt zu sein scheint), sollen in Vogels Ansatz dementsprechend nicht optimalitätstheoretisch beantwortet werden. Worum es stattdessen geht, sind Kasuskonflikte in Konstruktionen mit freien Relativsätzen und ihre optimalitätstheoretische Auflösung. Ein einfacher freier Relativsatz im Deutschen ist (13-a):

- (13) a. Ich lese [_{CP} was₁ ich t₁ will]
 b. Ich lese das [_{CP} was₁ ich t₁ will]

Vogel nimmt an, daß in (13-a) direkt eine CP vom Matrixverb eingebettet wird. Im vorliegenden Fall verlangen das Matrixverb und das eingebettete Verb beide den Akkusativ; ein Kasuskonflikt tritt somit für *was* nicht auf. Allerdings ergibt sich in jedem freien Relativsatz die Situation, daß ein W-Pronomen Diener zweier Herren ist, anders als in der Standardkonstruktion (13-b). Jeder freie Relativsatz verletzt daher die Beschränkung (14); dies ist der Preis, der für die Konstruktion zu bezahlen ist.

- (14) K-EIN (“Kasuseindeutigkeit”, “UniCase”):
 Jeder Kasuszuweiser weist genau einer NP lokal Kasus zu, und jede NP erhält Kasus von genau einem Kasuszuweiser.

In Sprachen, in denen eine Verletzung dieser Beschränkung immer fatal für einen Kandidaten ist, gibt es keine freien Relativsätze des Typs (13-a); im Deutschen sind K-EIN-Verletzungen nicht fatal. Wenn der vom Matrixverb und vom eingebetteten Verb geforderte Kasus nicht derselbe ist, kommt es zu Kasuskonflikten. Vogel stellt fest, daß ein solcher Konflikt im Deutschen (i) nur zugunsten des vom eingebetteten Verb kontrollierten Kasus auflösbar ist, und (ii) nur dann nicht zu Ungrammatikalität führt, wenn das Matrixverb strukturellen Kasus verlangt. Eigenschaft (i) wird durch die Kontraste in (15-ac) und (15-bd) dokumentiert.¹²

- (15) a. Ich lade ein_[+akk] [CP wem_{dat} ich t vertraue_[+dat]]
 b. Ich lade ein_[+akk] [CP wer_{nom} mir t sympathisch ist_[+nom]]
 c. *Ich lade ein_[+akk] [CP wen_{akk} ich t vertraue_[+dat]]
 d. *Ich lade ein_[+akk] [CP wen_{akk} mir t sympathisch ist_[+nom]]

Angenommen nun, ein Verb weist seinen Kasus bei A'-Abhängigkeiten obligatorisch an den Fuß einer Kette zu (von wo er an den Kopf der Kette vererbbar und so PF-realisiert ist). Dann haben in (15) die Spuren t immer den Kasus des eingebetteten Verbs. Trägt das W-Pronomen als Kopf der Kette nicht diesen Kasus (sondern den des Matrixverbs), wie in (15-c) und (15-d), ist die Beschränkung in (16) verletzt.

- (16) 1K1K (“Ein Kette, ein Kasus”, “1C1C”):
 Die Glieder einer Kette haben nicht verschiedene Kasus.

Verletzungen von 1K1K sind im Deutschen immer fatal. Eigenschaft (ii) illustrieren die Sätze in (17).

- (17) a. Ich folge_[+dat] [CP wem_{dat} ich t vertraue_[+dat]]
 b. *Ich folge_[+dat] [CP wen_{akk} ich t mag_[+akk]]

(17-a) zeigt, daß ein Matrixverb, das Dativ zuweist, einen freien Relativsatz zu sich nehmen kann, wenn das eingebettete Verb ebenfalls Dativ zuweist. Weist das eingebettete Verb wie in (17-b) jedoch einen anderen Kasus zu, so kommt es anders als bei Matrixverben, die strukturellen Kasus fordern (vgl. (15-a), (15-b)), immer zu Ungrammatikalität. Was (15-a) bzw. (15-b) und (17-b) gemeinsam haben, ist, daß der vom Matrixverb geforderte Kasus nicht realisiert wird. Dies wird durch die folgenden Beschränkungen bestraft, die zueinander in einer invarianten Ordnung stehen sollen, die der Reihenfolge ihrer Präsentation in (18) entspricht.

¹²Es gibt noch eine restriktivere Variante des Deutschen, in der (15-b) nicht wohlgeformt ist. Diese Variante wird bei Vogel (1999) auch erfaßt, soll im folgenden aber ausgeblendet bleiben.

- (18) a. REALOBL (“Realize Oblique”):
Ein Zuweiser von obliquem Kasus (wie Dativ) steht in einer lokalen Beziehung zu einer NP, die diesen Kasus realisiert.
- b. REALAKK (“Realize Accusative”):
Ein Akkusativzuweiser steht in einer lokalen Beziehung zu einer NP, die Akkusativ realisiert.
- c. REALNOM: (“Realize Nominative”):
Ein Nominativzuweiser steht in einer lokalen Beziehung zu einer NP, die Nominativ realisiert.

Die Daten zeigen dann, daß im Deutschen Verletzungen von REALOBL bei optimalen Kandidaten nicht möglich sind, anders als Verletzungen von REALAKK. (Letzteres gilt auch für REALNOM.)

Somit ergibt sich für das Deutsche, daß in freien Relativsätzen Verletzungen von 1K1K und REALOBL nicht in Kauf genommen werden können, Verletzungen von K-EIN, REALAKK und REALNOM dagegen schon. Was nun eigentlich nur noch zur Ableitung der Daten fehlt, ist eine Klärung der Natur der optimalen Kandidaten, die fatale Verletzungen von 1K1K und REALOBL im Deutschen vermeiden und so ungrammatische Sätze wie (15-c), (15-d) und (17-b) blockieren können. Vogel nimmt an, daß diese Kandidaten Neutralisierungskandidaten des Typs in (13-b) sind: Freien Relativsätzen wie (13-a) fehlt im Input zunächst einmal die funktionale D-Projektion, die Standardrelativsätze wie (13-b) aufweisen. Daß sich in vielen Fällen beide Strategien durchsetzen können, liegt eben an dieser unterschiedlichen Inputspezifikation. Standardrelativsatzvarianten mit realisiertem D-Pronomen sind aber in den Wettbewerben der freien Relativsätze als Input-untreue Kandidaten ebenfalls vorhanden. Diese Kandidaten fügen vom Input zum Output eine D-Schale hinzu und verletzen so die Treuebeschränkung in (19).

- (19) F-TREUE (“DepFunc”):
Jeder funktionale Kopf im Output muß auch im Input vorhanden sein.

Oft ist diese Verletzung bei Input-untreuen Standardrelativsätzen fatal; aber weil im Deutschen per Annahme F-TREUE unterhalb von 1K1K und REALOBL geordnet ist, gewinnt der normale Relativsatz gegen den freien Relativsatz, wenn letzterer eine dieser höher geordneten Beschränkungen verletzen müßte (in Standardrelativsätzen können beide Kasus realisiert werden). Mit anderen Worten: Hier wird neutralisiert.¹³ Damit ergibt sich als (eine mögliche) Ordnung für das Deutsche:

¹³Anstatt des Neutralisierungskandidaten, der F-TREUE nicht-fatal verletzt, kämen als alternative Mittel, hier absolute Ungrammatikalität abzuleiten, auch infrage der optimale leere Kandidat, der die in Kapitel 2 angeführte Beschränkung * \emptyset (= (74) aus Kapitel 2) nicht-fatal verletzt, oder ein optimaler schlechter Gewinner (im Sinne von Kapitel 2). Dies würde am Kern der Analyse nichts ändern. Vgl. auch die Diskussion von Wunderlich (2000) in Abschnitt 5.3 unten.

(20) *Ordnung der Beschränkungen im Deutschen:*

1K1K \gg REALOBL \gg F-TREUE \gg REALAKK \gg REALNOM \gg K-EIN

Anhand von Tabelle T_{7.9} wird zunächst deutlich, daß in solchen Fällen, wo der Matrix- und der eingebettete Kasus identisch sind (wo sog. "Matching" vorliegt), wie z.B. in (17-a), außer der tief geordneten Beschränkung K-EIN nichts verletzt wird. Der Neutralisierungskandidat (mit untreuem D-Pronomen) ist somit blockiert.

T_{7.9}: *Optimale freie Relativsätze ohne Kasuskonflikt im Deutschen*

Kandidaten	1K 1K	REAL OBL	F- TR	REAL AKK	REAL NOM	K- EIN
☞ K ₁ : V _[+dat] W _{dat} t _{dat} V _[+dat]						*
K ₂ : V _[+dat] D _{dat} W _{dat} t _{dat} V _[+dat]			*!			

Weiterhin zeigt Tabelle T_{7.10}, wie sich beim Kasuskonflikt in (15), z.B. bei (15-b) vs. (15-d) (Matrixverb verlangt Akkusativ, eingebettetes Verb verlangt Nominativ), Kandidat K₁ (= (15-b)) mit Realisierung des Nominativs gegenüber Kandidat K₂ (= (15-d)) mit Realisierung des Akkusativs und gegenüber dem Neutralisierungskandidaten K₃ durchsetzt.

T_{7.10}: *Optimale freie Relativsätze bei Kasuskonflikt im Deutschen*

Kandidaten	1K 1K	REAL OBL	F- TR	REAL AKK	REAL NOM	K- EIN
☞ K ₁ : V _[+akk] W _{nom} t _{nom} V _[+nom]				*		*
K ₂ : V _[+akk] W _{akk} t _{nom} V _[+nom]	*!				*	*
K ₃ : V _[+akk] D _{akk} W _{nom} t _{nom} V _[+nom]			*!			

Liegt ein Kasuskonflikt zwischen einem obliquen (inhärenten) Kasus (wie Dativ) und einem strukturellen Kasus (wie Akkusativ oder Nominativ) vor und ist der strukturelle Kasus der vom Matrixverb geforderte, so kann ein freier Relativsatz optimal werden; vgl. (15-a). In Fällen wie (17-b) jedoch, wo der strukturelle Kasus vom eingebetteten Verb verlangt wird und der oblique Kasus vom Matrixverb, ergibt sich Ungrammatikalität. Nichtrealisierung des obliquen Matrix-Kasus in (17-b) verletzt fatal REALOBL (vgl. K₁), Realisierung des obliquen Matrix-Kasus in K₂ verletzt fatal 1K1K, und so bleibt als Gewinner nur der Neutralisierungskandidat K₃ übrig. Diese Argumentation wird in Tabelle T_{7.11} nachgezeichnet.

Das so entstehende System macht eine Reihe von interessanten typologischen Vorhersagen. Solange der Neutralisierungskandidat höhere Beschränkungen nicht stärker verletzt als der freie Relativsatz, markiert F-TREUE eine Grenze; links dieser Grenze sind Verletzungen fatal. Angenommen nun, K-EIN ist in einer Sprache höher als F-TREUE geordnet. Dann ergibt sich die Prognose, daß es in dieser Sprache freie Relativsätze des hier betrachteten Typs nicht gibt. Angenommen weiter,

T_{7.11}: Blockierte freie Relativsätze bei Kasuskonflikt im Deutschen

Kandidaten	1K 1K	REAL OBL	F- TR	REAL AKK	REAL NOM	K- EIN
K ₁ : V _[+dat] W _{akk} t _{akk} V _[+akk]		*!				*
K ₂ : V _[+dat] W _{dat} t _{akk} V _[+akk]	*!			*		*
☞ K ₃ : V _[+dat] D _{dat} W _{akk} t _{akk} V _[+akk]			*			

eine Sprache weicht vom Deutschen (in diesem Bereich) nur insofern ab, als alle drei REAL-Beschränkungen über F-TREUE geordnet sind. Dann sind freie Relativsätze nur bei Kasusübereinstimmung (Matching) erlaubt; dies ist möglicherweise die korrekte Ordnung für das Englische. Weicht die Ordnung vom Deutschen nur durch die lokale Umordnung von REALOBL und F-TREUE ab, so erwarten wir, daß auch ein Matrixdativ bei Kasuskonflikt im freien Relativsatz unterdrückt werden kann. Aus der Annahme der festen internen Ordnung der REAL-Beschränkungen ergibt sich weiterhin als typologische Implikation, daß jede Sprache, die Kasuskonflikte zwischen Dativ und Akkusativ im freien Relativsatz zuläßt, auch solche zwischen Akkusativ und Nominativ zuläßt. Bemerkenswerte Effekte resultieren schließlich auch noch bei einer Absenkung von 1K1K unter F-TREUE: Wenn das Gebot der Kasushomogenität in der eingebetteten W-Kette nicht mehr stark ist, kann sich der Matrixkasus gegenüber dem eingebetteten Kasus durchsetzen. Dies erfaßt das Phänomen der sog. Kasusattraktion, wie es z.B. das Gotische kennt.

Vogel (1999) entwickelt diesen Ansatz noch in einer Reihe von Richtungen weiter.¹⁴ An dieser Stelle möchte ich jedoch, anstatt die Feinheiten der Analyse weiter zu verfolgen, kurz auf ein Faktum hinweisen, das zumindest auf den ersten Blick erstaunlich ist. Wie gesehen schlägt Woolford (1999) eine fixe Hierarchie von Kasusverboten vor: *DAT ≫ *AKK ≫ *NOM; bei Fanselow (1999) folgt Vergleichbares aus der Beschränkung *KASUS. Vogel (1999) nimmt nun demgegenüber eine fixe Hierarchie von Kasusgeböten an: REALOBL ≫ REALAKK ≫ REALNOM. Das Merkwürdige ist hierbei, daß die Verbote und Geböte bzgl. der Kasus nicht eine entgegengesetzte Ordnung aufweisen, sondern dieselbe. Der Grund hierfür liegt darin, daß Woolford und Fanselow auf der einen und Vogel auf der anderen Seite

¹⁴Eine wichtige zusätzliche Beschränkung ist (i).

- (i) 1K1R ("Eine Kette, eine Realisierung", "1C1W"):
Nur ein Kettenglied darf PF-realisiert werden.

Ist diese Beschränkung hoch geordnet, ändert sich gegenüber dem bisher Gesagten nichts. Ist sie jedoch zusammen mit 1K1K tief geordnet, so kann ein optimaler Kandidat durch eine PF-Realisierung der Spur als resumptives Pronomen mit einem vom Kopf der Kette unterschiedlichen Kasus 1K1R und 1K1K verletzen und dafür REAL-Beschränkungen und K-EIN erfüllen; das resumptive Pronomen ist hier eine Reparaturform, die in anderen Kontexten durch 1K1R ausgefiltert wird. Vogel argumentiert, daß diese Situation im Griechischen auftritt.

unterschiedliche Ziele verfolgen. Bei ersteren gilt der Kasusfilter als unverletzbar Beschränkung; jede NP muß eine Kasusverbotsbeschränkung verletzen, und verletzt wird dann die tiefstmögliche, im Idealfall also der Nominativ. Bei Vogel geht es dagegen darum, die Nichtrealisierung eines gegebenen Kasus zu bestrafen; und Nichtrealisierung z.B. des Dativs ist weniger leicht in Kauf zu nehmen als Nichtrealisierung des Akkusativs oder Nominativs. Tatsächlich entspricht Vogels Hierarchie der Kasusgebote in Woolfords Analyse von der Funktion her nicht die Hierarchie der Kasusverbote, sondern eher die Beschränkung LEX-TREUE, zusammen mit der Annahme, daß es eine entsprechende Beschränkung für strukturelle Kasus nicht gibt. Führt man sich dann noch vor Augen, daß Vogel am Ende des Papiers vorschlägt, REALAKK und REALNOM zu REALKASUS zusammenzufassen, dann scheint eine Vereinheitlichung der Ansätze gar nicht mehr so unmöglich, wie das zunächst einmal scheinen mag.

Soviel zu optimalitätstheoretischen Ansätzen zur Kasustheorie, die hauptsächlich auf der GB-Theorie (und ihren Weiterentwicklungen) basieren. Die letzten in diesem Kapitel vorzustellenden Ansätze sind beide im Rahmen der lexikalischen Dekompositionsgrammatik ausgearbeitet. Im folgenden Abschnitt wird daher diese Theorie kurz im Hinblick auf das, was im gegenwärtigen Kontext relevant ist, eingeführt.

5. Lexikalische Dekompositionsgrammatik

5.1. Der Ansatz

Die lexikalische Dekompositionsgrammatik (LDG) ist entwickelt worden vor allem von Manfred Bierwisch, Paul Kiparsky und Dieter Wunderlich. Standardreferenzen sind u.a. Bierwisch (1988) und Wunderlich (1997); in diesem Unterabschnitt folge ich letzterer Arbeit. Eine wesentliche Aufgabe des Modells ist die Zuordnung von Kasus zu im Lexikoneintrag von Prädikaten vermerkten Thetarollen. Diese Lexikoneinträge beinhalten an Ideen der Generativen Semantik angelehnte semantische Formen, die z.B. ein Verb wie *geben* in semantische Primitivprädikate dekomponieren und einen offenen Satz der Art “x macht, daß es so wird, daß y z besitzt” erzeugen (über den dann λ -abstrahiert wird). Die genaue Dekomposition ist für uns im gegenwärtigen Zusammenhang jedoch nicht zentral; wichtig ist nur, daß mit Hilfe der Dekomposition die Ordnung der Thetarollen für jedes Prädikat ableitbar ist: Im vorliegenden Fall ergibt sich, daß die Thetarolle von x der Thetarolle von y übergeordnet ist, und diese wiederum der Thetarolle von z. Diese Information soll uns hier ausreichen; wir können dann festhalten, daß sich aus der Dekomposition ergibt, daß *geben* ein Thetaraster $(\theta_1, \theta_2, \theta_3)$ hat.¹⁵ Die Thetarollen eines Rasters

¹⁵Tatsächlich wird in der LDG-Terminologie zwischen Thetarollen und Argumenten unterschieden; x, y und z sind in der LDG Argumente, und die Thetarollen entstehen durch λ -Abstraktion in umgetauschter Reihenfolge. Durch die Funktionsweise der λ -Konversion ist damit gewährlei-

können mit Hilfe der kontextuellen Merkmale [hr] (“es gibt eine höhere Thetarolle”) und [tr] (“es gibt eine tiefere Thetarolle”, im engl. Original [lr] für “lower role”) charakterisiert werden. Für intransitive, transitive und ditransitive Verben wie *weinen*, *lesen* und *geben* ergeben sich somit die Einträge in (21) (die Notation ist gegenüber der Standard-LDG wie erwähnt vereinfacht).¹⁶

- (21) a. *weinen*: $\langle \theta_1 \rangle$
 [-hr,-tr]
 b. *lesen*: $\langle \theta_1, \theta_2 \rangle$
 [-hr,+tr] [+hr,-tr]
 c. *geben*: $\langle \theta_1, \theta_2, \theta_3 \rangle$
 [-hr,+tr] [+hr,+tr] [+hr,-tr]

Begriffe wie Agens, Thema usw. haben in diesem System keinen relevanten Status. Das “Agens” z.B. ist in (21-bc) die Thetarolle mit der Charakterisierung [-hr,+tr], in (21-a) die Thetarolle mit der Charakterisierung [-hr,-tr]. Die wesentliche Annahme ist nun, daß die strukturellen Kasus ebenso wie die Thetarollen mit den Merkmalen [hr] und [tr] charakterisierbar sind. (Letztlich sind also strukturelle Kasus in diesem Modell auch lexikalisch.) Konkret schlägt Wunderlich (1997) folgende Identifikation vor: Der Dativ entspricht [+hr,+tr], der Akkusativ entspricht [+hr], der Ergativ entspricht [+tr], und der Nominativ entspricht der leeren Auszeichnung []. Der Dativ und die mittlere Thetarolle θ_2 bei ditransitiven Verben sind somit gleichgesetzt. Der Akkusativ, der Ergativ und insbesondere der Nominativ sind jedoch unterspezifiziert. Auf dieser Basis soll abgeleitet werden, welchen Kasus ein Argument mit einer gegebenen Thetarolle erhält. Die Kasuszuweisung unterliegt verschiedenen Abbildungsregeln, insbesondere dem Spezifitätsprinzip:

(22) *Spezifitätsprinzip*:

Jeder Thetarolle wird der spezifischste damit kompatible Kasus zugeordnet.

Ein Kasus ist spezifischer als ein anderer, wenn er durch mehr Merkmale charakterisiert ist (Akkusativ ist spezifischer als Nominativ, Dativ ist spezifischer als Akkusativ). Ein Kasus ist kompatibel mit einer Thetarolle, wenn es keine unterschiedlichen Werte für gemeinsame Merkmale gibt (Kompatibilität bzgl. eines Merkmals ist also automatisch gegeben, wenn ein Merkmal einer Thetarolle kein Merkmal des Kasus ist). Betrachten wir einige Beispiele (und sehen wir fürs erste vom Ergativ ab). Bei einem intransitiven Verb ist die einzige Thetarolle [-hr,-tr] markiert. Da Akkusativ

stet, daß die Abfolge der Argumente im Lexikon der Abfolge in der Syntax entspricht; dies muß man in der Vereinfachung, die ich hier vornehme, extra stipulieren (so wie man die umgetauschte Reihenfolge bei der λ -Abstraktion in der LDG stipulieren muß).

¹⁶Es sei vermerkt, daß die in (21-c) postulierte syntaktische Abfolge der Argumente nicht mit dem vereinbar ist, wofür in Müller (1999c) anhand von Bindungsdaten argumentiert worden ist; vgl. Kapitel 6.

und Dativ das Merkmal [+hr] haben, sind sie hiermit inkompatibel, und es bleibt nur der Nominativ für ein Argument mit dieser Thetarolle übrig; vgl. (23-a). Bei einem einfach transitiven Verb hat die erste Thetarolle die Markierung [-hr,+tr]. Ein Argument mit dieser Thetarolle erhält wiederum den Nominativ, denn Akkusativ und Dativ sind mit der [-hr]-Spezifikation unvereinbar.¹⁷ Die zweite Thetarolle hat die Spezifikation [+hr,-tr]. Dies ist inkompatibel mit dem Dativ und kompatibel mit Nominativ und Akkusativ. Die Entscheidung zwischen diesen beiden Kasus wird durch das Spezifitätsprinzip getroffen, denn der Akkusativ ist spezifischer als der Nominativ; vgl. (23-b). Bei einem ditransitiven Verb schließlich erhält das Argument mit der höchsten Thetarolle wiederum wegen Inkompatibilität der anderen Kasus den Nominativ. Die mittlere Thetarolle ist mit Dativ, Akkusativ und Nominativ kompatibel; folglich trägt ein Argument mit dieser Thetarolle den spezifischsten der drei Kasus, den Dativ. Die tiefste Thetarolle ist inkompatibel mit dem Dativ, und wie beim einfach transitiven Verb entscheidet das Spezifitätsprinzip dann für den Akkusativ, nicht für den merkmallösen Nominativ; vgl. (23-c).

- (23) a. daß der Fritz_{nom} weint
 b. daß die Maria_{nom} das Buch_{akk} liest
 c. daß die Maria_{nom} dem Fritz_{dat} ein Buch_{akk} gibt

Es stellt sich die Frage, wie sich inhärenter Kasus in diesen Ansatz einfügt. Die Thetarollen-Charakterisierungen in (21) werden als Defaults behandelt, die durch besondere lexikalische Spezifikationen überschrieben werden können. So wird z.B. für die Ableitung des inhärenten Dativs bei *helfen* angenommen, daß die tiefere Thetarolle exzeptionell anstatt als [-tr] als [+tr] charakterisiert wird; d.h., das Besondere am inhärenten Kasus ist, daß kontextuell definierte Merkmale dort auftreten, wo sie vom Kontext her eigentlich nicht sein können.¹⁸ Um das konzeptuelle Problem zu vermeiden, nunmehr zwei [+tr] markierte Thetarollen zu haben, obwohl es doch eigentlich für eine der beiden Thetarollen keine tiefere geben kann, wird die [tr]-Markierung auf der eigentlich höheren Thetarolle entsprechend umgestellt:¹⁹

- (24) a. daß die Maria_{nom} dem Fritz_{dat} geholfen hat

¹⁷Man beachte, daß diese beiden Fälle bereits erklären, warum es nicht möglich ist, den Nominativ wie z.B. den Dativ vollständig bzgl. [tr] zu spezifizieren. Daß man den Nominativ auch nicht bzgl. [hr] spezifizieren möchte (“[-hr]”), liegt, was das Deutsche betrifft, an seiner Verwendbarkeit als Default-Kasus. Siehe Abschnitt 5.3 unten.

¹⁸Die exzeptionelle Auszeichnung mache ich hier und im folgenden durch Unterstreichung kenntlich; bei Wunderlich (2000) und Stiebels (2000) erfolgt sie per Fettdruck und expliziter Annotation: “lexikalisch”.

¹⁹Der Preis ist jedoch ein neues konzeptuelles Problem: Für die erste, höhere Thetarolle wird nun gesagt, daß es keine tiefere gibt, und für die zweite, tiefste Thetarolle wird gesagt, daß es noch eine tiefere gibt. Ich werde in Abschnitt 5.3 hierauf zurückkommen.

- b. *helfen*: $\langle \theta_1, \theta_2 \rangle$
 $[-hr, -tr]$ $[+hr, +tr]$

Jetzt erhält ein Argument von *helfen* mit der Thetarolle θ_2 den spezifischsten mit dieser Thetarolle vereinbaren Kasus, und das ist nicht mehr der Akkusativ, sondern der Dativ. Das Argument mit der Thetarolle θ_1 wird demgegenüber mit dem Nominativ markiert, da dies der einzig kompatible Kasus ist.

Als weitere Illustration betrachte man ein psychologisches Prädikat wie *gefallen*. Wunderlich (1997) nimmt an, daß hier der Dativ dem Argument mit der höheren Thetarolle und der Nominativ dem Argument mit der tieferen Thetarolle zugewiesen wird. Dazu wird die höhere Thetarolle exzeptionell mit dem Merkmal $[+hr]$ ausgezeichnet, die tiefere Thetarolle entsprechend auf $[-hr]$ umgestellt.²⁰

- (25) a. daß der Maria_{dat} der Plan_{nom} gefallen hat
 b. *gefallen*: $\langle \theta_1, \theta_2 \rangle$
 $[+hr, +tr]$ $[-hr, -tr]$

Somit erhält ein Argument mit der Thetarolle θ_1 den spezifischsten damit kompatiblen Kasus, also den Dativ; ein Argument mit der Thetarolle θ_2 erscheint demgegenüber im unspezifischsten Nominativ, weil Akkusativ und Dativ mit der $[-hr]$ -Auszeichnung inkompatibel sind. Soviel zum inhärenten Kasus in der Standard-LDG.²¹

Dieser Ansatz hat einige Eigenschaften, die eine optimalitätstheoretische Reanalyse möglich und sinnvoll erscheinen lassen. Zum einen läßt sich die Ermittlung des für eine Thetarolle geeignetsten Kasus gemäß dem Spezifizitätsprinzip kaum anders denn als Optimierung verstehen. Zum anderen ist bei der Behandlung in-

²⁰Nach Wunderlich (1997, 51) zeigen "Wortstellung und Bindung ... daß das Dativargument die höhere Rolle ist". Was die Wortstellung betrifft, so beruht die Evidenz auf der Annahme der in Kapitel 6 diskutierten DTK-basierten Theorie der Markiertheit. Tatsächlich ist die gegenüber (25-a) vertauschte Argumentabfolge markierter. Ist jedoch das Nominativargument belebt, dreht sich die Normalabfolge um. Besonders deutlich wird dies, wenn das Dativargument trotz Experienter-Status selbst nicht einfach als belebt interpretiert werden kann; vgl. (i-a). Aus der Normalabfolge auf die Argumentstruktur zu schließen, erweist sich somit wieder einmal als problematisch. Ebenso sprechen die einfachsten Bindungsdaten nicht für die hier postulierte Argumentordnung (und alle komplexeren Bindungsdaten, z.B. solche mit Überkreuzungseffekten, sagen per se nichts über die Basisabfolge, sondern im Höchstfall etwas über die formale Natur von Scrambling). Ein vorangehendes Dativargument kann keine Nominativanapher möglich machen, während umgekehrt die Bindung unproblematisch ist; vgl. (i-b).

- (i) a. daß der neue Intendant_{nom} dem ganzen Ensemble_{dat} gefallen hat
 b. daß $\{*\text{dem Intendanten}_{dat} \text{ sich}_{nom}\} / \{\text{der Intendant}_{nom} \text{ sich}_{dat}\}$ gut gefällt

Im folgenden werde ich von diesen empirischen Problemen mit Wunderlichs Klassifikation des Thetarasters bei Verben wie *gefallen* jedoch absehen.

²¹Der inhärente Genitiv scheint mit den beiden Merkmalen $[hr]$ und $[tr]$ nicht beschreibbar. Tatsächlich sind für seine Erfassung Zusatzannahmen notwendig; vgl. Wunderlich (2000).

härenter Kasus vorausgesetzt, daß dadurch der per Default erwartbare strukturelle Kasus blockiert wird. Die im folgenden vorgestellten optimalitätstheoretischen Weiterentwicklungen der LDG sind Stiebels (2000) und Wunderlich (2000).²²

5.2. Struktureller Kasus und Spezifität bei Stiebels (2000)

Eine optimalitätstheoretische Ableitung der Effekte des Spezifitätsprinzips wird in Stiebels (2000) vorgenommen. Dabei spielen Treuebeschränkungen eine zentrale Rolle. Wie bereits in Kapitel 1 erwähnt wurde, ersetzen McCarthy & Prince (1995) die in Prince & Smolensky (1993) angenommenen Treuebeschränkungen PARSE (keine Weglassung von Inputmaterial im Output) und FILL (keine Einfügung von Material im Output, das nicht im Input war) durch die etwas allgemeiner formulierten Treuebeschränkungen MAX bzw. DEP. Hinzu kommen als dritter Typ IDENT-Beschränkungen, die Merkmalsveränderung von Input zu Output verbieten. Wichtig sind bei Stiebels nun IDENT-Beschränkungen und MAX-Beschränkungen für die Merkmale [hr] und [tr]:

- (26) a. IDENT([hr]):
Der [hr]-Wert einer Thetarolle im Input darf nicht mit einem [hr]-Wert des Arguments mit dieser Thetarolle im Output konfigrieren.
- b. IDENT([tr]):
Der [tr]-Wert einer Thetarolle im Input darf nicht mit einem [tr]-Wert des Arguments mit dieser Thetarolle im Output konfigrieren.
- c. MAX([+hr]):
Eine [+hr]-Spezifikation einer Thetarolle im Input muß auf dem Argument mit dieser Thetarolle im Output erscheinen.
- d. MAX([+tr]):
Eine [+tr]-Spezifikation einer Thetarolle im Input muß auf dem Argument mit dieser Thetarolle im Output erscheinen.

Wir können hier annehmen, daß der Input eine Numeration ist, also eine Ansammlung von lexikalischen Kategorien (inkl. ihrer Thetaraster), und daß der Output ein Satz ist (inkl. Kasusspezifikation aller NP-Argumente).²³ Die IDENT-Beschränkungen kodieren das Kompatibilitätsgebot aus (22): Ein Dativ ([+hr,+tr]) ist z.B. inkompatibel mit einer Inputspezifikation [-hr,+tr]. Wird der Dativ im Deutschen einem ranghöchsten Argument zugewiesen, resultiert daher normaler-

²²Vgl. auch Kiparsky (1999), eine in mancherlei Hinsicht mit Stiebels (2000) vergleichbare Arbeit, auf die ich hier nicht weiter eingehe – zum einen aus Platzgründen, zum anderen, weil sie mir nur als Handout vorliegt.

²³Dies ist eine Vereinfachung gegenüber Stiebels (2000); entsprechend dieser Vereinfachung sind die Treuebeschränkungen hier auch leicht reformuliert.

weise eine (fatale) Verletzung von IDENT([hr]).²⁴ Die MAX-Beschränkungen erfassen demgegenüber die Spezifitätseffekte an sich: Bei einer Inputspezifikation wie [+hr,+tr] verletzt ein Nominativ ([]) im Output MAX([+hr]) und MAX([+tr]) (aber keine IDENT-Beschränkung, da nicht der Wert eines Merkmals verändert wurde), ein Akkusativ ([+hr]) im Output immerhin noch MAX([+tr]). Klar ist, daß in der Beschränkungsordnung die IDENT-Beschränkungen die MAX-Beschränkungen dominieren müssen. Wie die Ordnung unter den IDENT- und MAX-Beschränkungen jeweils aussieht, spielt für die gegenwärtigen Zwecke keine Rolle (ich habe mich hier für eine der möglichen Ordnungen entschieden). In Tabelle T 7.12 ist illustriert, wie bei einem einfach transitiven Verb wie *lesen* der Kandidat mit Nominativ-Akkusativ-Muster im Output optimal wird.²⁵

T_{7.12}: Ableitung von Spezifitätseffekten in LDG. Im Input:

lesen: $\langle \theta_1, \theta_2 \rangle$
[-hr,+tr],[+hr,-tr]

Kandidaten:	IDENT([hr])	IDENT([tr])	MAX([+hr])	MAX([+tr])
☞ K ₁ : NP _{nom} NP _{akk} V				*
K ₂ : NP _{nom} NP _{nom} V			*!	*
K ₃ : NP _{akk} NP _{nom} V	*!		*	*
K ₄ : NP _{akk} NP _{akk} V	*!			*
K ₅ : NP _{nom} NP _{dat} V		*!		*
K ₆ : NP _{dat} NP _{akk} V	*!			
K ₇ : NP _{dat} NP _{dat} V	*!	*		

Man beachte, daß in diesem System anders als z.B. bei Fanselow (1999) im Prinzip jeder Kasus beliebig oft vorkommen kann. Da jedoch Kasus und Thetarollen mit demselben Merkmalsvokabular beschrieben werden und die Thetarollen normalerweise aufgrund kontextueller Klassifikation unterschiedlich spezifiziert sind, ergibt sich, daß eine uniforme Verwendung eines Kasus grundsätzlich einer Kasusmischung unterlegen ist, bei der sich Thetarolle und Kasus möglichst gut entsprechen (vgl. K₂, K₄, K₇, und s.u.). Ebenfalls bemerkenswert ist, daß ein Nominativ auf einem Argument mit der ranghöchsten Thetarolle bei transitiven und ditransitiven Verben zwar der Normalfall ist, aber immer eine Verletzung von MAX([+tr]) hervorruft, weil die [+tr]-Spezifikation im Input durch den unterspezifizierten ([]) Nominativ nicht realisiert werden kann. Es gibt nur zwei grundsätzliche Möglichkeiten, diese Verletzung zu umgehen, durch den Dativ ([+hr,+tr]) oder den Ergativ ([+tr]). Der Dativ verletzt in diesen Fällen aber fatalerweise die höher geordnete Beschränkung

²⁴Normalerweise, weil ja bei Wunderlich (1997) wie gesehen eine inhärente Dativzuweisung für das höchste Argument bei psychologischen Prädikaten wie *gefallen* angenommen wird, die auf exzeptionelle Markierung der höchsten Thetarolle als [+hr,+tr] zurückgeführt wird. Siehe unten.

²⁵Die Reihenfolge der Argumente entspricht hier der Ordnung im Thetaraster.

IDENT([hr]) (vgl. K₆); und der Ergativ steht im Deutschen prinzipiell nicht zur Verfügung (s.u.).

Tabelle T_{7.13} zeigt, wieso bei einem regulären ditransitiven Verb wie *geben* ein Nominativ-Dativ-Akkusativ-Muster auftaucht.

T_{7.13}: *Ableitung von Spezifitätseffekten in LDG. Im Input:*

geben: $\langle \theta_1, \theta_2, \theta_3 \rangle$
[-hr,+tr],[+hr,+tr],[+hr,-tr]

Kandidaten:	IDENT([hr])	IDENT([tr])	MAX([+hr])	MAX([+tr])
☞ K ₁ : NP _{nom} NP _{dat} NP _{akk} V				*
K ₂ : NP _{nom} NP _{akk} NP _{akk} V				**!
K ₃ : NP _{nom} NP _{dat} NP _{nom} V			*!	*
K ₄ : NP _{nom} NP _{nom} NP _{akk} V			*!	**
K ₅ : NP _{nom} NP _{dat} NP _{dat} V		*!		*
K ₆ : NP _{dat} NP _{dat} NP _{akk} V	*!			
K ₇ : NP _{akk} NP _{dat} NP _{akk} V	*!			*

Bis hierhin wird also die gesamte Arbeit bei der Kasuswahl von Treuebeschränkungen verrichtet.²⁶ Daß jedoch auch interpolierte Markiertheitsbeschränkungen wichtig sind, zeigt sich, wenn man den als [+tr] klassifizierten Ergativ betrachtet. Wie gesehen ruft ein Nominativ auf dem höchsten Argument eines regulären transitiven Verbs (das die Markierung [-hr,+tr] trägt) immer eine MAX([+tr])-Verletzung hervor. Durch die Wahl des Ergativs anstatt des Nominativs könnte diese Verletzung nun vermieden werden, und wir würden kontrafaktisch erwarten, daß im Deutschen bei transitiven Verben immer ein Ergativ-Akkusativ-Muster auftritt. Damit diese Konsequenz für Nominativ-Akkusativ-Sprachen (und eine entsprechend unerwünschte Konsequenz bzgl. des Akkusativs für Ergativ-Nominativ(Absolutiv)-Sprachen) vermieden wird, führt Stiebels Markiertheitsbeschränkungen ein.

- (27) a. *[+hr]:
Die Spezifikation [+hr] darf nicht im Output erscheinen.
b. *[+tr]:
Die Spezifikation [+tr] darf nicht im Output erscheinen.
c. EIN ("Eindeutigkeit", "Uniqueness"):
Jeder Kasus kann pro Kasusdomäne nur einmal auftreten.

*[+hr] interagiert mit MAX([+hr]), *[+tr] mit MAX([+tr]). Vorausgesetzt, all diese Beschränkungen sind unter den IDENT-Beschränkungen geordnet, entscheidet die

²⁶Man kann sich fragen, was passieren würde, wenn es neben MAX-Beschränkungen für positiv spezifizierte Merkmale auch solche für negativ spezifizierte Merkmale gäbe. Tatsächlich hätte dies keine Konsequenzen: Da die Kasus alle nur durch positiv spezifizierte Merkmale ausgezeichnet sind, würden alle Kandidaten unabhängig von dem gewählten Kasusmuster diese zusätzlichen MAX-Beschränkungen gleichermaßen verletzen.

relative Ordnung dieser vier Beschränkungen (sowie von EIN) über das Kasusinventar einer Sprache, und somit über grundsätzliche Sprachtypen. Gelten $\text{MAX}([+hr]) \gg * [+hr]$ und $* [+tr] \gg \text{MAX}([+tr])$, resultiert eine Nominativ-Akkusativ-Sprache des deutschen Typs: $[+hr]$ im Input muß im Output sichtbar gemacht werden, und hierfür bietet sich bei einfach transitiven Verben ohne IDENT-Verletzung nur der Akkusativ an; $[+tr]$ im Input darf dagegen nicht sichtbar gemacht werden, weil der infrage kommende Kasus hier der Ergativ ist und das Verbot gegen diesen Kasus höher geordnet ist. Im Wettbewerb in Tabelle T_{7.12} z.B. verliert ein Ergativ-Kandidat gegen den Nominativ-Kandidat K_1 . Dies zeigt Tabelle T_{7.14}, wo bereits in Tabelle T_{7.12} als suboptimal erwiesene Kandidaten (bis auf einen) ebenso wie die IDENT-Beschränkungen weggelassen sind und eine von verschiedenen möglichen Ordnungen angenommen wurde, die den erwünschten Effekt erzielen.

T_{7.14}: *Nominativ-Akkusativ-Muster. Im Input:*

$V: < \theta_1, \theta_2 >$
 $[-hr, +tr], [+hr, -tr]$

Kandidaten:	EIN	MAX([+hr])	*[+hr]	*[+tr]	MAX([+tr])
$\leftarrow K_1: NP_{nom} NP_{akk} V$			*		*
$K_2: NP_{erg} NP_{nom} V$		*!		*	
$K_3: NP_{erg} NP_{akk} V$			*	*!	
$K_4: NP_{nom} NP_{nom} V$	*!	*			*

Hier ist K_1 optimal. Unter einer Beschränkungsordnung, in der umgekehrt $\text{MAX}([+tr]) \gg * [+tr]$ und $* [+hr] \gg \text{MAX}([+hr])$ gelten, ergibt sich eine klassische Ergativsprache mit einem Ergativ-Nominativ-Muster bei transitiven Verben; K_2 wird optimal. Dominieren beide MAX-Beschränkungen beide Markiertheitsbeschränkungen, wird K_3 optimal, und es entsteht eine Sprache mit Ergativ-Akkusativ-Muster. Sind schließlich $* [+tr]$ und $* [+hr]$ beide über den entsprechenden Treuebeschränkungen geordnet, kann bei Tiefordnung von EIN K_4 und somit ein Nominativ-Nominativ-Muster optimal werden (ansonsten entscheidet die relative Ordnung von $* [+hr]$ und $* [+tr]$ zwischen K_1 und K_2).

Bis zu diesem Punkt gibt es noch keine positive Evidenz für EIN. Tatsächlich sind ja in den bisherigen Wettbewerben unmögliche Kandidaten mit mehreren Instanzen eines Kasus allein aufgrund ihres Beschränkungsprofils ausgeschlossen worden (T_{7.12}, T_{7.13}). Darüber hinaus gefährdet EIN bei Hochordnung sogar die beobachtete Wohlgeformtheit solcher Muster in manchen Sprachen (T_{7.14}). Direkte Evidenz für EIN liefert jedoch der Dativ in ditransitiven Konstruktionen im Deutschen bei Annahme der Markiertheitsbeschränkung $* [+tr]$. Das Problem ist, daß die Ordnung $* [+tr] \gg \text{MAX}([+tr])$ nicht nur erwünschterweise den Ergativ im Deutschen blockiert, sondern zunächst einmal unerwünschterweise auch den Dativ (der ja $[+hr, +tr]$ spezifiziert und somit hier dem Ergativ zu ähnlich ist): Um in T_{7.13} eine Verletzung der höher geordneten Beschränkung $* [+tr]$ durch K_1 mit einem Nominativ-Dativ-Akkusativ-Muster zu vermeiden, kann nun die Verletzung

der tiefer geordneten Beschränkung $\text{MAX}([+tr])$ durch K_2 mit einem Nominativ-Akkusativ-Akkusativ-Muster in Kauf genommen werden. Dieser Effekt (der für das Quechua erwünscht ist) wird im Deutschen vermieden, wenn K_2 noch zusätzlich die hoch (zumindest über $*[+tr]$) geordnete Beschränkung EIN verletzt. Dies zeigt Tabelle $T_{7.15}$, eine Erweiterung von Tabelle $T_{7.13}$, in der IDENT -Beschränkungen ebenso fehlen wie die anderen, hier zu vernachlässigenden Kandidaten, und die dafür durch die drei neuen Markiertheitsbeschränkungen angereichert ist.

$T_{7.15}$: *Nominativ-Dativ-Akkusativ-Muster. Im Input:*

$V: \langle \theta_1, \theta_2, \theta_3 \rangle$
 $[-hr, +tr], [+hr, +tr], [+hr, -tr]$

Kandidaten:	EIN	$\text{MAX}([+hr])$	$*[+hr]$	$*[+tr]$	$\text{MAX}([+tr])$
$\Leftarrow K_1: \text{NP}_{nom} \text{NP}_{dat} \text{NP}_{akk} V$				*	*
$K_2: \text{NP}_{nom} \text{NP}_{akk} \text{NP}_{akk} V$	*!				**

Wie der Dativ mit dem Ergativ das Merkmal $[+tr]$ teilt, so hat er mit dem Akkusativ das Merkmal $[+hr]$ gemein. Für Ergativ-Nominativ-Sprachen gelten daher analoge Überlegungen: Aufgrund der Ordnung $*[+hr] \gg \text{MAX}([+hr])$ wird zunächst einmal nicht nur ein Akkusativ im einfach transitiven Fall, sondern auch ein Dativ beim mittleren Argument bei ditransitiven Verben blockiert, und zwar zugunsten des Ergativs (der zwar $\text{MAX}([+hr])$ verletzt, dafür aber $*[+hr]$ erfüllt). Dieses Ergebnis ist wiederum für manche Ergativ-Sprachen korrekt; für Sprachen wie das Baskische mit Ergativ-Dativ-Nominativ-Muster braucht man jedoch wieder eine hohe Einordnung von EIN , um den Ergativ-Ergativ-Nominativ-Kandidaten zu blockieren.

Stiebels zeigt, daß ihre Beschränkungen unter faktorieller Typologie transitive und ditransitive Konstruktionen in einer ganzen Reihe von Sprachen korrekt erfassen können. Darüber hinaus entwickelt sie das System so weiter, daß Personen- und Belebtheits-konditionierte Unterschiede bei der Kasuszuweisung in manchen Sprachen beschrieben werden können (vgl. auch Aissen (1999) und Kapitel 8). Das hier dargestellte Material mag jedoch ausreichen, um einen Eindruck davon zu erhalten, wie man per Optimierung gemäß verletzbarer und geordneter Beschränkungen in der LDG sowohl das Kasusinventar einer gegebenen Sprache (z.B. Ergativ vs. Akkusativ), als auch die Wahl zwischen den a priori möglichen Kasus in jedem Konstruktionstyp dieser Sprache (z.B. Dativ vs. Akkusativ) ableiten kann.

5.3. Inhärenter Kasus und Sichtbarkeit bei Wunderlich (2000)

Geht es bei Stiebels (2000) in erster Linie um strukturellen Kasus, so beschäftigt sich Wunderlich (2000) vor allem mit dem inhärenten Kasus. Zunächst einmal ist festzuhalten, daß Wunderlich (2000) (wie auch Stiebels (2000)) eine Änderung gegenüber der Thetarasterspezifikation bei inhärentem Kasus in Wunderlich (1997) vornimmt. Nach wie vor wird die lexikalische Idiosynkrasie bei Verben

wie *helfen* und *gefallen* darin gesehen, daß eine (Dativ-erzwingende) [+tr]- bzw. [+hr]-Spezifikation auf der tieferen bzw. höheren Thetarolle die Defaultspezifikation überschreibt. Eine Umstellung des entsprechenden Merkmals auf der anderen Thetarolle von + auf – erfolgt jedoch nicht mehr; rein formal ergibt sich also die Situation, daß die zwei Argumente eines transitiven Verbs beide eine [+tr]- bzw. [+hr]-Spezifikation tragen können. Die Einträge für *helfen* und *gefallen* sehen somit wie in (28) aus (vgl. (24-b), (25-b)).

- (28) a. *helfen*: $\langle \theta_1, \theta_2 \rangle$
 [-hr,+tr] [+hr,+tr]
 b. *gefallen*: $\langle \theta_1, \theta_2 \rangle$
 [+hr,+tr] [+hr,-tr]

In der LDG unterscheiden sich inhärenter und struktureller Kasus wie gesehen v.a. bzgl. der Frage, ob die Thetarollenklassifikation im Lexikoneintrag per Default erfolgt oder exzeptionell markiert wird (beide Kasus sind strenggenommen lexikalisch). Wenn also einmal ein Thetaraster für ein Verb ermittelt worden ist, sollten die Treue- und Markiertheitsbeschränkungen für Kasusrealisierung idealerweise ohne weitere Annahmen die korrekten Ergebnisse liefern. Dies ist der Fall bei Verben wie *helfen*. Das Thetaraster in (28-a) sagt unter den Treuebeschränkungen in T_{7.12} (sowie unter den Markiertheitsbeschränkungen aus T_{7.14}) für ein Argument mit der höheren, [-hr,+tr] markierten Thetarolle den Nominativ vorher (Dativ und Akkusativ rufen hier aufgrund ihrer [+hr]-Markierung fatale IDENT-Verletzungen hervor), und für ein Argument mit der tieferen, [+hr,+tr] markierten Rolle den Dativ (Nominativ und Akkusativ lösen fatale MAX-Verletzungen aus).

Bei *gefallen* gibt es unter der neuen Perspektive jedoch zunächst einmal ein interessantes Problem. Ein Thetaraster wie (28-b) führt zwar wie erwünscht zu einer Kasusrealisierung des Arguments mit der höheren, [+hr,+tr] markierten Thetarolle als Dativ. Die tiefere Thetarolle ist nun jedoch nicht mehr [-hr,-tr] markiert, sondern [+hr,-tr]. Statt dem unspezifischen Nominativ sollte aufgrund von MAX([+hr]) beim zweiten Argument von *gefallen* also der spezifischere Akkusativ auftreten. Dies scheint zwar für manche Sprachen die korrekte Vorhersage zu sein;²⁷ für das Deutsche jedoch nicht. Im Spezifizitätsansatz von Wunderlich (1997) wäre unklar, wie man das Problem lösen könnte; aber in der optimalitätstheoretischen Analyse von Wunderlich (2000) (sowie von Stiebels (2000)) kann die Blockade des Akkusativs durch den Nominativ bei *gefallen* auf eine intervenierende Beschränkung zurückgeführt werden, die im Deutschen höher als MAX([+hr]) geordnet ist.

²⁷Dieter Wunderlich (p.M.) verweist auf Malayalam. Neben der konzeptuellen Vereinfachung ist die Existenz solcher Dativ-Akkusativ-Muster im übrigen einer von mehreren empirischen Gründen für die Änderung bei den Thetarastern mit exzeptioneller Markierung in der LDG.

Wunderlich (2000) nennt diese Beschränkung NOM!.²⁸

(29) NOM!:

Jedes Verb hat einen Nominativkasus.

Tabelle T_{7.16} illustriert den Wettbewerb bei Verben wie *gefallen* und zeigt, wie der inhärente Dativ auf dem höheren Argument und der strukturelle Nominativ auf dem tieferen Argument optimal werden.

T_{7.16}: *Inhärenter Kasus in LDG. Im Input:*

gefallen: $\langle \theta_1, \theta_2 \rangle$
 $[\underline{+hr}, +tr], [+hr, -tr]$

Kandidaten:	IDENT([hr])	IDENT([tr])	NOM!	MAX([+hr])	MAX([+tr])
☞ K ₁ : NP _{dat} NP _{nom} V				*	
K ₂ : NP _{dat} NP _{akk} V			*!		
K ₃ : NP _{nom} NP _{akk} V				*	*!
K ₄ : NP _{nom} NP _{dat} V		*!		*	*
K ₅ : NP _{akk} NP _{nom} V				*	*!

Bei den bisher betrachteten Konstruktionen werden inhärenter und struktureller Kasus in der optimalitätstheoretischen LDG gleich behandelt: Anders als bei Woolford oder Fanselow fällt die wesentliche Entscheidung über inhärenten vs. strukturellen Kasus *vor* der eigentlichen syntaktischen Optimierung, nämlich bei der Konstruktion des Lexikoneintrags. Bei der syntaktischen Optimierung sorgen dann Treuebeschränkungen dafür, daß den Thetarollenspezifikationen möglichst nah verwandte (oder identische) Kasusspezifikationen entsprechen. Markiertheitsbeschränkungen (wie das Ergativverbot *[tr], die Einzigkeitsbedingung EIN oder die gerade erwähnte Beschränkung NOM!) können dieses Bild stören. Die bisher betrachteten Markiertheitsbeschränkungen unterscheiden jedoch nicht zwischen inhärenten und strukturellen Kasus.²⁹

²⁸NOM! findet bei Wunderlich auch an anderer Stelle Anwendung. Beim Passiv wird dadurch der von MAX([+hr]) geforderte Akkusativ in (i-a) zugunsten des Nominativs in (i-b) blockiert.

- (i) a. *daß der Maria_{dat} den Mann_{akk} vorgestellt wird
 b. daß der Maria_{dat} der Mann_{nom} vorgestellt wird

Allgemein dient NOM! auch der Ableitung von Burzios Generalisierung; der Effekt von NOM! bei Wunderlich ist dabei (bei Fehlen einer entsprechenden Beschränkung AKK!) der fixen Ordnung *AKK \gg *NOM bei Woolford vergleichbar. NOM! ist im übrigen nicht dasselbe wie die in Kapitel 2 eingeführte Beschränkung DEFAULT-NOM ("Eine NP trägt Nominativ." = "Wenn α eine NP ist, dann trägt α Nominativ."). DEFAULT-NOM fordert für *jede* NP den Nominativ und muß daher tief geordnet sein, um nicht falsche Vorhersagen zu machen; NOM! fordert demgegenüber nur, daß *eine* NP eines gegebenen Verbs den Nominativ trägt, und kann daher relativ hoch geordnet sein.

²⁹Dies bedeutet wie gesehen nicht, daß es in LDG keinen Wettbewerb von inhärentem und strukturellem Kasus gibt. Dieser Wettbewerb erfolgt aber eben nicht in der Syntax, sondern im

Grundsätzlich spricht nun nichts dagegen, bei der syntaktischen Optimierung in der LDG Beschränkungen zu haben, die zwischen regulären, per Default zugewiesenen Spezifikationen und exzeptionellen, durch spezielle Auszeichnung entstandenen Spezifikationen unterscheiden; denn wenn auch z.B. eine reguläre [+hr,+tr]-Spezifikation auf dem Thetaraster eines Verb sich hinsichtlich der geforderten Dativrealisierung genauso verhält wie eine exzeptionelle [+hr,+tr]- bzw. [+hr,+tr]-Spezifikation, so ist doch aufgrund der Unterstreichung immer noch ein Unterschied im Input vorhanden, auf den sich Beschränkungen beziehen können. Eine solche Beschränkung für inhärenten Kasus ist (30).

- (30) SICHTBARKEIT (“Visibility”):
Inhärenter Kasus ist in allen Derivationen sichtbar.

Diese Beschränkung entspricht LEX-TREUE bei Woolford (1999). Die zugrundeliegende Annahme ist, daß inhärenter Kasus weniger leicht durch grammatische Operationen verloren gehen kann als struktureller (vgl. etwa das einfache Passiv im Deutschen, das den strukturellen Akkusativ absorbiert, oder das Rezipientenpassiv mit *kriegen*, das den strukturellen Dativ verschluckt). Die Relevanz von SICHTBARKEIT illustriert Wunderlich anhand von Kontroll- und A.c.I.-Konstruktionen im Deutschen und Isländischen; ich werde mich hier auf erstere beschränken.

Die betrachteten Daten weisen alle obligatorische Subjektkontrolle auf. Eine Grundannahme ist, daß die Subjektkontrolleigenschaft eines Verbs nicht verletzbar ist: Es sind nur Kandidaten in einem Wettbewerb, die eine identische Koindizierung des Matrixsubjekts und eines Infinitivarguments aufweisen.³⁰ Zwischen dem Deutschen und dem Isländischen gibt es dann bei Kontrolle von Infinitivargumenten mit inhärentem Kasus einen interessanten Unterschied. Dieser Unterschied läßt sich anhand von passivierten Nominativ-Dativ-Konstruktionen zeigen, in denen nur noch das Dativargument mit inhärentem Kasus verfügbar ist.³¹

- (31) a. *Ich hoffe [PRO_{dat} geholfen zu werden]

Lexikon: Hier wird angenommen, daß die exzeptionelle Markierung immer die Defaultmarkierung außer Kraft setzt. Will man dies in der LDG optimalitätstheoretisch erfassen, so ist der offensichtliche Weg der, zusätzlich zur syntaktischen Optimierung eine prä-syntaktische Optimierung der Thetaraster im Lexikon zu postulieren (vgl. Kapitel 8), in der die Entscheidung über inhärenten vs. strukturellen Kasus vorgegeben wird. Eine solche Vorgehensweise wäre mit der allgemeinen Architektur von LDG vereinbar, wird aber in Wunderlich (2000) und Stiebels (2000) nicht gewählt.

³⁰Vgl. die Rekonstruktion von Stechow & Sternefeld (1988) in Kapitel 4, wo dies gerade nicht so war.

³¹Das fehlende Argument des Kontrollinfinitivs notiere ich anders als Wunderlich im folgenden als PRO. Dies geschieht nur der besseren Übersichtlichkeit halber; Wunderlich nimmt nicht an, daß das fehlende Argument in Kontrollinfinitiven als leere Kategorie syntaktisch präsent ist.

- b. Ég vonast [til að PRO_{dat} verða hjálpað]
 ich hoffe für zu werden geholfen

Wunderlichs Vorschlag ist, daß der Kontrast mit der unterschiedlichen Ordnung von SICHTBARKEIT im Deutschen und Isländischen zu tun hat. In (31-ab) ist der inhärente Dativ beidesmal nicht sichtbar; dieses Faktum ist fatal für (31-a), kann aber in Kauf genommen werden in (31-b), weil SICHTBARKEIT im Isländischen nicht sehr hoch geordnet ist. An Beschränkungen erweisen sich neben SICHTBARKEIT die folgenden als relevant:

- (32) a. *INF-PROP (“NoProp”):
 Infinitiven fehlt ein Argument (sie sind keine Propositionen).
 b. *NOM-INF (“NoNom”):
 Infinitive erlauben keinen Nominativ.
 c. *∅ (“Vermeide leere Kandidaten”):
 Der Input darf nicht vollständig unrealisiert sein.
 d. *HOCH-INF: (“NoHigh”):
 Infinitiven fehlt das höchste Argument.

*INF-PROP beschreibt die allgemeine Eigenschaft von Kontrollinfinitiven, ein Argument obligatorisch unrealisiert zu lassen; diese Beschränkung ist in beiden Sprachen undominiert.

*∅ entspricht der Beschränkung (74) aus Kapitel 2, wie sie von Ackema & Neeleman (1998) vorgeschlagen worden ist.³² Diese Beschränkung bestraft den leeren Kandidaten ∅, der per Annahme Teil eines jeden Wettbewerbs ist. Es handelt sich um eine Möglichkeit, absolute Ungrammatikalität in einer Kandidatenmenge abzuleiten (andere Möglichkeiten wie schlechte Gewinner oder Neutralisierung kämen im Prinzip aber genauso infrage, ohne daß sich an Wunderlichs Analyse substantiell etwas ändern müßte).

*NOM-INF ist im Deutschen nicht verletzbar (in der GB-Theorie folgte das daraus, daß Nominativ nur von finitem I zuweisbar war), aber wir werden sehen, daß dem im Isländischen nicht so ist. *NOM-INF steht zu NOM! als speziell zu generell in einer Pāninischen Relation (vgl. (63) aus Kapitel 2): Jede Erfüllung von *NOM-INF führt zu einer Verletzung von NOM!, aber nicht notwendigerweise umgekehrt. Aufgrund von Pāninis Theorem (vgl. (64) aus Kapitel 2) muß dann *NOM-INF über NOM! geordnet sein, um Arbeit zu verrichten.

*HOCH-INF schließlich verbietet das Auftauchen des höchsten Arguments eines Verbs im Infinitiv. In Fällen, wo das höchste Argument Nominativ trägt, verlangt diese Beschränkung dasselbe wie *NOM-INF; aber bei inhärentem Kasus auf dem höchsten Argument gibt es Unterschiede. Darüber hinaus steht *INF-PROP

³²Wie Ackema und Neeleman nennt Wunderlich diese Beschränkung PARSE. Vgl. die terminologischen Bemerkungen hierzu in Kapitel 2.

zu *HOCH-INF als speziell zu generell in einer Stringenzbeziehung (vgl. (65) aus Kapitel 2): Jede Verletzung von *INF-PROP ist auch eine Verletzung von *HOCH-INF, aber nicht notwendigerweise umgekehrt (und dies, obschon die Formulierung von *INF-PROP auf den ersten Blick allgemeiner scheint als die Formulierung von *HOCH-INF).³³ Zwei für das Deutsche und Isländische mögliche Ordnungen sind (33-a) und (33-b).

- (33) a. *Ordnung im Deutschen:*
 *INF-PROP \gg SICHTBARKEIT \gg *NOM-INF \gg * \emptyset \gg *HOCH-INF
 b. *Ordnung im Isländischen:*
 *INF-PROP \gg *HOCH-INF \gg * \emptyset \gg SICHTBARKEIT \gg *NOM-INF

Tabelle T_{7.17} zeigt dann, wie im Deutschen K₁ (= (31-a)) mit fataler Verletzung von SICHTBARKEIT ebenso wie K₂ mit fataler Verletzung von *INF-PROP durch den leeren Kandidaten K₃ blockiert wird, der nur * \emptyset verletzt. * \emptyset markiert wie F-TREUE bei Vogel (1999) eine Grenze: Alle Verletzungen links davon sind bei einem Kandidaten automatisch fatal (vorausgesetzt, der leere Kandidat sammelt keine zusätzlichen Verletzungen in diesem Bereich an).

T_{7.17}: *Inhärenter Dativ im passivierten Infinitiv im Deutschen :*

Kandidaten:	*INF-PROP	SICHTBARKEIT	*NOM-INF	* \emptyset	*HOCH-INF
K ₁ : ... [PRO _{dat} geholfen zu werden]		*!			
K ₂ : ... [mir _{dat} geholfen zu werden]	*!				*
☞ K ₃ : \emptyset				*	

Aufgrund der tieferen Ordnung von SICHTBARKEIT im Isländischen erscheint in Tabelle T_{7.18} K₁ (= (31-b)) als optimal; K₂ mit fataler *INF-PROP-Verletzung und der leere Kandidat K₃ werden als suboptimal blockiert.

T_{7.18}: *Inhärenter Dativ im passivierten Infinitiv im Isländischen :*

Kandidaten:	*INF-PROP	*HOCH-INF	* \emptyset	SICHTBARKEIT	*NOM-INF
☞ K ₁ : ... [til að PRO _{dat} verða hjálpað]				*	
K ₂ : ... [til að mér _{dat} verða hjálpað]	*!	*			
K ₃ : \emptyset			*!		

Die Beschränkungen *NOM-INF und *HOCH-INF haben bislang noch keine entscheidende Rolle gespielt – erstere nicht, weil im Input gar kein Argument vorkam,

³³Wunderlich nimmt an, daß *HOCH-INF ableitbar ist aus einer allgemeineren Beschränkung SEMHIER (“Semantische Hierarchie”), die von derselben Art ist wie das bereits mehrfach angesprochene PAR-VER. SEMHIER verlangt, daß sich die lexikalisch determinierte Argumentordnung in der Syntax widerspiegelt. Zu Details der Ableitung vgl. Wunderlich (2000, 3.1).

das Realisierung als Nominativ verlangte; letztere nicht, weil bei einstelligen Verben *HOCH-INF und *INF-PROP dasselbe fordern und jede *HOCH-INF-Verletzung sofort eine Verletzung der (an sich spezielleren) undominierten Beschränkung *INF-PROP mit sich bringt. Dies ändert sich bei der Betrachtung von Verben wie *gefallen*, wo per Annahme im Deutschen wie im Isländischen das höhere Argument inhärenten Dativ, das tiefere strukturellen Nominativ erhält.

Die zu erklärende Beobachtung ist, daß im Deutschen bei *gefallen* das tiefere (Nominativ-) Argument kontrolliert werden kann, das höhere (Dativ-) Argument jedoch nicht, und daß die Situation im Isländischen genau umgekehrt ist:

- (34) a. Ich hoffe [ihr_{dat} PRO_{nom} zu gefallen]
 b. *Ich hoffe [PRO_{dat} das Buch_{nom} zu gefallen]
 "Ich hoffe, daß mir das Buch gefällt."
 c. *Ég vonast [til að henni_{dat} líka PRO_{nom}]
 ich hoffe für zu ihr mögen
 "Ich hoffe, ihr zu gefallen."
 d. ?Ég vonast [til að PRO_{dat} líka þessi bók_{nom}]
 ich hoffe für zu mögen dieses Buch
 "Ich hoffe, daß mir das Buch gefällt."

Eine einfache Beschreibung des Sachverhalts ist, daß nur Subjekte kontrolliert werden, und daß das Subjekt im Deutschen immer das Nominativargument ist, im Isländischen dagegen das höchste Argument.³⁴ Dies wird bei Wunderlich nun abgeleitet. Die relevante Ordnung im Deutschen ist SICHTBARKEIT \gg * \emptyset \gg *HOCH-INF, so daß zwar nicht das höchste Argument kontrolliert werden muß, aber auf keinen Fall bei der Kontrolle ein inhärenter Kasus verloren gehen darf. Zwei Wettbewerbe sind zu betrachten: einer, in dem das Matrixsubjekt mit dem eingebetteten Nominativargument koindiziert ist und *HOCH-INF verletzt werden muß (hier gewinnt (34-a) gegen den leeren Kandidaten), und einer, in dem das Matrixsubjekt mit dem eingebetteten Dativargument koindiziert ist und SICHTBARKEIT verletzt werden müßte (hier gewinnt der leere Kandidat gegen (34-b)). Die Wettbewerbe werden illustriert durch die Tabellen T_{7.19} und T_{7.20}.

Im Isländischen gilt gegenüber dem Deutschen jedoch die Ordnung *HOCH-INF \gg * \emptyset \gg SICHTBARKEIT, so daß zwar unbedingt das höchste Argument kontrolliert werden muß, ein inhärenter Kasus aber im Zweifelsfall verloren gehen darf. Wiederrum sind zwei Wettbewerbe zu betrachten. Im ersten Wettbewerb ist das Matrixsubjekt mit dem eingebetteten Nominativargument koindiziert; hier gewinnt, weil *HOCH-INF verletzt werden müßte, anders als im Deutschen der leere Kandidat

³⁴Nähme man an, daß bei *gefallen* im Deutschen tatsächlich das Nominativargument über dem Dativargument steht (vgl. Fußnote 20), würde sich das Problem von vornherein nicht stellen: In beiden Sprachen wäre es das höchste Argument, das kontrolliert werden kann. (Die Evidenz für den Subjektstatus von Dativargumenten wie den hier vorliegenden ist im Isländischen weit weniger kontrovers als im Deutschen.)

T_{7.19}: Dativ-Nominativ-Infinitive im Deutschen, Kontrolle des Nominativs :

Kandidaten:	*INF-PROP	SICHT-BARKEIT	*NOM-INF	*Ø	*HOCH-INF
☞ K ₁ : ... [ihr _{dat} PRO _{nom} zu gefallen]					*
K ₂ : ... [ihr _{dat} ich _{nom} zu gefallen]	*!		*		*
K ₃ : Ø				*!	

T_{7.20}: Dativ-Nominativ-Infinitive im Deutschen, Kontrolle des Dativs :

Kandidaten:	*INF-PROP	SICHT-BARKEIT	*NOM-INF	*Ø	*HOCH-INF
K ₁ : ... [PRO _{dat} d. Buch _{nom} zu gefallen]		*!	*		
K ₂ : ... [mir _{dat} d. Buch _{nom} zu gefallen]	*!		*		*
☞ K ₃ : Ø				*	

gegen (34-c). Im zweiten Wettbewerb ist das Matrixsubjekt mit dem eingebetteten Dativsubjekt koindiziert; hier gewinnt (34-d) trotz Verletzung von SICHTBARKEIT gegen den leeren Kandidaten.³⁵ Die Wettbewerbe sind in den Tabellen T_{7.21} und T_{7.22} dargestellt (auf den fatal *INF-PROP verletzenden Kandidaten ist jeweils verzichtet worden).

T_{7.21}: Dativ-Nominativ-Infinitive im Isländischen, Kontrolle des Nominativs :

Kandidaten:	*INF-PROP	*HOCH-INF	*Ø	SICHT-BARKEIT	*NOM-INF
K ₁ : ... [til að henni _{dat} líka PRO _{nom}]		*!			
☞ K ₂ : Ø			*		

Man beachte, daß es für das Überleben von K₁ in T_{7.22} wichtig ist, daß *NOM-INF unter *Ø geordnet ist, d.h., daß neben der Vertauschung von SICHTBARKEIT und *HOCH-INF vom Deutschen zum Isländischen auch noch *NOM-INF abgesenkt ist. Andernfalls würde auch in T_{7.22} der leere Kandidat gewinnen, und wir hätten eine Sprache, die bei einem Konflikt zwischen Kontrolle des höchsten Arguments und Kontrolle des Nominativs gar keine Kontrolle mehr zuläßt.³⁶

³⁵(34-d) ist nicht ganz perfekt, und vergleichbare Konstruktionen mit belebtem Nominativargument sind noch schlechter. Wunderlich (2000, 3.2.) spekuliert, daß dies darauf zurückzuführen sein könnte, daß der optimale Kandidat sowohl SICHTBARKEIT, als auch *NOM-INF verletzt, d.h., daß bei der Ermittlung von Markiertheitsgraden bei eigentlich optimalen Kandidaten kumulative Effekte eine Rolle spielen; vgl. Kapitel 6.

³⁶Die hohe Einordnung von *NOM-INF im Deutschen ist übrigens durch die in diesem Abschnitt betrachteten Daten nicht motiviert; T_{7.17}, T_{7.19} und T_{7.20} würden dasselbe Ergebnis liefern, wenn *NOM-INF wie im Isländischen eingeordnet wäre oder sogar vollkommen fehlen würde. Evidenz

$T_{7.22}$: *Dativ-Nominativ-Infinitive im Isländischen, Kontrolle des Dativs* :

Kandidaten:	*INF-PROP	*HOCH-INF	*Ø	SICHTBARKEIT	*NOM-INF
☞ K ₁ : ... [til að PRO _{dat} líka þessi bók _{nom}]				*	*
K ₂ : Ø			*!		

Abschließend sei nochmals kurz auf den Status von SICHTBARKEIT eingegangen. Grundsätzlich ist SICHTBARKEIT nichts anderes als eine MAX-Beschränkung für inhärente Kasusmerkmale; für den Fall des inhärenten Dativmerkmals heißt das: MAX([+tr]). Da nur Daten mit inhärentem Dativ betrachtet wurden, stellt sich die Frage, ob anstelle von SICHTBARKEIT nicht einfach die allgemeinere Beschränkung MAX([+hr,+tr]) angenommen werden könnte. Ist eine [+hr,+tr]-Spezifikation im Thetaraster eines Verbs im Input nicht auf einem Argument im Output präsent, wie aufgrund von undominiertem *INF-PROP im Falle von Infinitiven, die das Dativargument weglassen, ist MAX([+hr,+tr]) immer verletzt. Tatsächlich macht eine Ersetzung von SICHTBARKEIT durch MAX([+hr,+tr]) in den Tabellen dieses Abschnitts dieselben Vorhersagen. Die Weglassung eines strukturellen Dativs in Infinitiven würde nun genauso MAX([+hr,+tr]) verletzen (aber nicht SICHTBARKEIT). Dies wäre jedoch unproblematisch: Im Deutschen könnte (35-b) allein wegen der hohen Ordnung von MAX([+hr,+tr]) nicht gegenüber (35-a) optimal werden.

- (35) a. Ich versuche [PRO_{nom} dem Fritz_{dat} das Buch_{acc} zu schenken]
 b. *Ich versuche [der Fritz_{nom} PRO_{dat} das Buch_{acc} zu schenken]

Und auch im Isländischen hätte das Analog zu (35-b) keine Chance, jemals optimal zu werden, da die Beschränkungen, die Argumentweglassung im Infinitiv fordern (*INF-PROP, *NOM-INF, *HOCH-INF), ohnehin die Weglassung des höheren Nominativarguments erzwingen, nicht des tieferen Dativarguments. (Im übrigen gäbe es ja ansonsten im Isländischen auch schon Probleme mit Nominativ-Dativ-Mustern, in denen der Dativ inhärent ist.)

Was bei einer Ersetzung von SICHTBARKEIT durch MAX([+hr,+tr]) verloren ginge, wäre jedoch, wie Barbara Stiebels (p.M.) bemerkt, eine Erklärung dafür, daß ein Satz wie (36) mit einem inhärenten Genitiv ebenso unmöglich ist wie der analoge Fall mit einem inhärenten Dativ, der oben besprochen wurde (vgl. (31-a)).

- (36) *Ich hoffe [PRO_{gen} gedacht zu werden]

Dies folgt aus SICHTBARKEIT, aber nicht aus MAX([+hr,+tr]), denn welche Spezifikation ein inhärenter Genitiv auch letztendlich hat (vgl. Fußnote 21), es ist nicht die des Dativs. Darüber hinaus verlöre man die zugrundeliegende Intuition, daß

für *NOM-INF und für seine hohe Einordnung im Deutschen kommt aus anderen Konstruktionen (Kontrollinfinitive mit zwei strukturellen Kasus und A.c.I.-Konstruktionen).

struktureller Kasus leichter “verschluckt” werden kann als inhärenter. Das Rezipientenpassiv im Deutschen etwa, in dem bei einem Passivauxiliar *kriegen* nicht das Akkusativobjekt, sondern das Dativobjekt zum Nominativargument wird, ist nach allgemeinem Dafürhalten sensitiv für die Unterscheidung von strukturellem und inhärentem Kasus:

- (37) a. Fritz_{nom} kriegt ein Buch_{akk} gekauft
 b. Fritz_{nom} kriegt ?geholfen/??gedankt/*gefolgt

(37-ab) verletzen beide MAX([+hr,+tr]), aber nur (37-b) verletzt SICHTBARKEIT. Bei einer Ersetzung von SICHTBARKEIT durch MAX([+hr,+tr]) muß also eine andere Beschränkung als MAX([+hr,+tr]) für die Ungrammatikalität von (37-b) verantwortlich sein. Somit können wir festhalten, daß eine Reduktion von SICHTBARKEIT auf unabhängig etablierte Beschränkungen, die nicht für inhärenten Kasus formuliert sind, zwar konzeptuell wünschenswert, aber nicht unproblematisch ist.

6. Konklusion

Im Detail unterscheiden sich die hier vorgestellten Ansätze zur Kasuszuweisung sowohl in ihrer empirischen wie in ihrer theoretischen Ausrichtung. Empirisch geht es bei Woolford (1999), Fanselow (1999) und Stiebels (2000) vor allem um die Kasusverteilung in einfachen Strukturen (ein Verb und seine Argumente); Vogel (1999) und Wunderlich (2000) dagegen unterwerfen Kasusmuster speziellen Kontexten, die die Unterdrückung eines Kasus verlangen (freie Relativsätze bzw. Infinitive). Als ein Beispiel für unterschiedliche theoretische Orientierung sei nur das Problem der Aufteilung der Arbeit zwischen Lexikon und Syntax erwähnt: Bei Stiebels (2000) und Wunderlich (2000) wird ein Großteil der Arbeit bei der Kasus-theorie bereits im Lexikon gemacht: Hier finden sich durch die ausführliche Spezifikation von Thetarollen bereits wesentliche Informationen über inhärenten wie strukturellen Kasus, und Treuebeschränkungen sorgen dafür, daß die lexikalische Spezifikation oft in der Syntax umgesetzt wird. Allerdings spielt auch in diesem Ansatz die Syntax bei der Kasuszuweisung eine Rolle: Markiertheitsbeschränkungen sorgen dafür, daß bestimmte, aufgrund lexikalischer Spezifikation erwartbare Kasus nicht auftreten (so wird z.B. im Deutschen der Ergativ generell, der zweite Akkusativ bei ditransitiven Verben und der Nominativ in Kontrollinfinitiven durch Markiertheitsbeschränkungen unterdrückt). Woolford (1999) (partiell auch Vogel (1999)) postuliert demgegenüber eine Arbeitsteilung von Lexikon und Syntax, wie sie in der GB-Theorie angenommen wurde: Struktureller Kasus wird von syntaktischen Markiertheitsbeschränkungen erfaßt, inhärenter Kasus dagegen durch die explizite Kassuspezifikation einer Thetarolle im Lexikon; hoch geordnete Treuebeschränkungen sorgen im letzteren Fall dafür, daß sich der inhärente Kasus in der Syntax durchsetzt. Fanselow (1999) schließlich versucht die Rolle des Lexikons zu minimieren: Inhärenter Kasus wird durch eine entsprechende Markierung auf dem Verb im Lexikon notiert (jedoch nicht mit einer Thetarolle korreliert) und setzt

sich per Blockadepinzip (bzw. invariant hoch geordneter Treuebeschränkung) immer gegen strukturellen Kasus in der Syntax durch. Die restliche Arbeit (inkl. der Frage, welche Thetarolle dem inhärenten Kasus assoziiert wird) macht allein die Syntax.

Ungeachtet solcher Unterschiede ist jedoch festzuhalten, daß die dargestellten Ansätze in einigen wichtigen Punkten konvergieren. Erstens sind alle Arbeiten explizit sprachvergleichend ausgerichtet, und sie zeigen, daß parametrische Variation zwischen Sprachen hinsichtlich des Kasussystems auf einfache Weise durch Beschränkungsumordnung erfaßt werden kann. Zweitens wird die Generalisierung, daß inhärenter, im Lexikon exceptionell markierter Kasus in der Syntax eher realisiert werden muß als struktureller Kasus (was beinhaltet, daß sich zum einen inhärenter Kasus im Konfliktfall gegen strukturellen Kasus durchsetzt, und daß sich zum anderen inhärenter Kasus weniger leicht in besonderen syntaktischen Kontexten – wie freien Relativsätzen und Kontrollinfinitiven – unterdrücken läßt als struktureller Kasus), uniform durch entsprechende hoch geordnete Treuebeschränkungen abgeleitet. Drittens schließlich stimmen die Analysen im wesentlichen darin überein, daß nicht nur strukturelle Kasus weniger markiert sind als inhärente, sondern daß unter den strukturellen Kasus der Nominativ auch weniger markiert ist als der Akkusativ (worunter auch die Effekte von Burzios Generalisierung fallen), und der Akkusativ weniger markiert als der Dativ. Dies folgt bei Woolford (1999) aus einer festen Hierarchie von Kasusverboten; bei Fanselow (1999) aus dem allgemeinen Kasusverbot *KASUS in Interaktion mit einer vom Nominativ zum Dativ ansteigenden Merkmalskomplexität der Kasus; und bei Stiebels (2000) und Wunderlich (2000) aus einer ebensolchen unterschiedlichen Merkmalskomplexität (kodiert mit Hilfe der kontextuellen Merkmale [\pm hr, \pm tr]) in Interaktion mit IDENT-Treuebeschränkungen und dem Kasusgebot NOM!. Insofern, als diese drei Punkte – parametrische Variation, Primat des inhärenten Kasus bei besonderer lexikalischer Auszeichnung, Primat des Nominativs über den Akkusativ und des Akkusativs über den Dativ ansonsten – mit Hilfe geordneter und verletzbarer Beschränkungen sowie einem Konzept des Wettbewerbs gut (und besser als in Standard-Syntaxen) zu erfassen sind, können sie als Argumente für eine optimalitätstheoretische Herangehensweise betrachtet werden.³⁷

³⁷Explizite Argumente gegen einen optimalitätstheoretischen Ansatz im Bereich der deutschen Kasussyntax finden sich bei Bayer, Bader & Meng (1999; 2000). Allerdings wird in diesen Arbeiten auch davon ausgegangen, daß ein indirektes Objekt anstatt durch den Dativ mit Hilfe einer Präposition realisiert werden kann, wenn der Dativ unabhängig blockiert ist; dieser Effekt wird durch die Elsewhere Condition abgeleitet. Diese Bedingung ist jedoch ihrerseits einer optimalitätstheoretischen Rekonstruktion zugänglich; vgl. die Bemerkungen zum Blockade-Prinzip in Kapitel 4. Darüber hinaus verwenden Bayer, Bader & Meng (2000) ein Prinzip der minimalen Anbindung (“Minimal Attachment”), das translokal formuliert ist und dementsprechend nicht nur einen Wettbewerb voraussetzt, sondern ebenfalls optimalitätstheoretisch rekonstruierbar scheint; vgl. Kapitel 4 und Fanselow, Schlewsky & Kliegl (1999).

Kapitel 8

Neuere Entwicklungen

Zum Abschluß dieses Buches sollen noch einige neuere Entwicklungen in der optimalitätstheoretischen Syntax skizziert werden.

1. Beschränkungen

Wie in Kapitel 2 bemerkt wurde, ist ein oft geäußelter Einwand gegen optimalitätstheoretische Analysen eine fehlende Restriktivität hinsichtlich der Natur der verwendeten Beschränkungen. Es ist unklar, so die Kritik, was eine mögliche optimalitätstheoretische syntaktische Beschränkung ist und was nicht. Meine eigene Konklusion in Kapitel 2 war, daß dieses Problem vielleicht nicht größer ist als in anderen syntaktischen Theorien auch, und daß die Kriterien für eine "gute" Beschränkung, die sich anderswo durchgesetzt haben (wie Einfachheit und Allgemeinheit), auch in der Optimalitätstheorie Anwendung finden sollten.

Nichtsdestoweniger ist festzustellen, daß verschiedentlich, und verstärkt in den letzten Jahren, Versuche unternommen worden sind, Restriktionen für Beschränkungen zu formulieren. Insbesondere sind hier die Arbeiten von Grimshaw (1998; 1999) zu nennen. Dort wird zum einen eine detailliertere Untergliederung von möglichen Klassen von Beschränkungen vorgenommen, die über eine einfache Einteilung in Treue- und Markiertheitsbeschränkungen hinausgeht. Grimshaw schlägt vor, zu unterscheiden zwischen Treuebeschränkungen, Markiertheitsbeschränkungen, Ökonomiebeschränkungen, Strukturbeschränkungen und Abbildungsbeschränkungen (die das Verhältnis von syntaktischer Struktur zu anderen Repräsentationsebenen betreffen). Zu dieser Taxonomie gesellt sich eine Forderung nach Einfachheit: Eine Beschränkung in der optimalitätstheoretischen Syntax soll soweit wie möglich ohne Boolesche Operationen in der Formulierung auskommen und minimale logische Komplexität aufweisen. Auf dieser Basis kritisiert Grimshaw (1998) ihre eigene Beschränkung PUR-EP aus Grimshaw (1997), die hier nochmals wiederholt ist:

(1) PUR-EP:

Adjunktion an die höchste XP einer eingebetteten erweiterten Projektion und Bewegung in deren Kopf sind verboten.

Diese Beschränkung ist komplex: P ist verboten, und Q ist verboten. Dazu redet die Beschränkung von der "höchsten" XP einer erweiterten Projektion (bzw. deren Kopf), und sie unterscheidet zwischen "eingebetteten" erweiterten Projektionen und anderen. Grimshaw (1998) schlägt daher vor, auf PUR-EP in dieser Form zu verzichten und die Effekte von PUR-EP durch die Interaktion einer Vielzahl einfacher

Beschränkungen abzuleiten. Die Grundidee ist, daß die bereits in Grimshaw (1997) auftauchenden Beschränkungen OB-KOPF und KOPF-LINKS anstatt als Einzelbeschränkungen als Familien von Beschränkungen angesehen werden, deren Mitglieder zueinander in Stringenzbeziehung stehen. Die OB-KOPF-Familie umfaßt neben der allgemeinen Beschränkung OB-KOPF auf bestimmte Domänen relativierte Versionen, nämlich OB-KOPF^{mtx} und OB-KOPF^{exp}. Hier steht “mtx” für “Matrix” (gemeint ist der Wurzelknoten) und “exp” für die höchste Schicht einer erweiterten Projektion. In derselben Weise ist die KOPF-LINKS-Familie unterteilt.

- (2) a. (i) OB-KOPF:
Eine Projektion hat einen (nicht-leeren) Kopf.
(ii) OB-KOPF^{exp}:
Die höchste Schicht einer erweiterten Projektion hat einen (nicht-leeren) Kopf.
(iii) OB-KOPF^{mtx}:
Die Wurzel-Projektion hat einen (nicht-leeren) Kopf.
- b. (i) KOPF-LINKS:
Ein Kopf steht am linken Rand seiner Projektion.
(ii) KOPF-LINKS^{exp}:
Ein Kopf steht am linken Rand der höchsten Schicht seiner erweiterten Projektion.
(iii) KOPF-LINKS^{mtx}:
Ein Kopf steht am linken Rand der Wurzel-Projektion.

Eine für das Englische mögliche Ordnung ist (3).

- (3) *Ordnung im Englischen:*
OB-KOPF^{mtx} \gg KOPF-LINKS^{mtx} \gg KOPF-LINKS^{exp} \gg OB-KOPF^{exp} \gg OB-KOPF \gg KOPF-LINKS

Angenommen nun, es gibt im System von Grimshaw (1997) eine durch höher geordnete Beschränkungen (wie OP-SPEZ oder KASUS, SUBJEKT) erzwungene Besetzung eines Spezifikators SpecX. Dann erfordern jeweils die für die vorliegende Domäne einschlägigen OB-KOPF-Beschränkungen Füllung der Kopfposition X, während umgekehrt die jeweiligen KOPF-LINKS-Beschränkungen diese Füllung von X durch einen lexikalischen Kopf verbieten (ein Kopf, der einem Spezifikator folgt, ist nicht links-peripher in seiner Projektion). Durch die relative Ordnung der allgemeinsten Beschränkungen OB-KOPF und KOPF-LINKS wird so normalerweise eine Kopfposition unter Verletzung von KOPF-LINKS gefüllt; vgl. etwa die *do*-Einsetzung in Negationskontexten:

- (4) a. *I think [_{CP} that [_{IP} John₁ [_I -] [_{NegP} not [_{VP} t₁ left]]]]
b. I think [_{CP} that [_{IP} John₁ [_I did] [_{NegP} not [_{VP} t₁ leave]]]]

In Bereichen, wo zusätzlich noch KOPF-LINKS^{exp} und OB-KOPF^{exp} relevant sind, also in der höchsten Schicht einer erweiterten Projektion, wird durch die umgekehr-

te Ordnung dieser Beschränkungen auch das Ergebnis umgedreht: Hier verletzt der optimale Kandidat OB-KOPF^{exp} (und OB-KOPF). Das ist ein klassischer PUR-EP-Effekt:

- (5) a. I wonder [CP what₁ – [IP she₂ will [VP t₂ say t₁]]]
 b. *I wonder [CP what₁ will₃ [IP she₂ t₃ [VP t₂ say t₁]]]

Handelt es sich bei der fraglichen Kopfposition jedoch um die Wurzel domäne, dann gewinnt aufgrund der Ordnung OB-KOPF^{mtx} ≫ KOPF-LINKS^{mtx} wieder der Kandidat mit Füllung der Kopfposition, unter Verletzung von KOPF-LINKS^{mtx} (sowie KOPF-LINKS^{exp} und KOPF-LINKS):

- (6) a. *[CP What₁ – [IP she will₃ [VP t say t₁]]] ?
 b. [CP What₁ will₃ [IP she t₃ [VP t say t₁]]] ?

Auf diese Weise ist ein Teil von PUR-EP abgeleitet (mehr wäre zu sagen zu den Fällen, wo PUR-EP Adjunktion einer XP blockiert). Die involvierten Beschränkungen sind alle einfacher; der zu bezahlende Preis ist, daß es auch mehr sind.¹ Darüber hinaus stellt sich die Frage, ob die für das Englische postulierte Ordnung ein reiner Zufall ist. Unter faktorieller Typologie erwarten wir z.B., daß es genauso gut eine Sprache geben kann, in der die Füllung von Kopfpositionen zwar in der Peripherie (Wurzel und höchste Schicht der erweiterten Projektion) obligatorisch, ansonsten aber verboten ist.

Interessant und für eine generelle Tendenz in der optimalitätstheoretischen Syntax charakteristisch ist, daß in der soeben besprochenen Ableitung mit KOPF-LINKS eine Beschränkung wesentlich involviert war, die zur Klasse der sog. Ausrichtungsbeschränkungen (“alignment constraints”) gehört. Diese Beschränkungen spielen in der Phonologie eine große Rolle. Sie sind sämtlich von der Art, daß für ein bestimmtes Element eine lineare Ausrichtung mit dem rechten oder linken Rand einer übergeordneten Konstituente gefordert wird – im vorliegenden Fall für einen Kopf die lineare Ausrichtung mit dem linken Rand seiner Projektion. In der Syntax spielen derartige Beschränkungen außer bei Grimshaw (1997; 1998; 1999) u.a. eine Rolle bei Costa (1998; 1999), Pesetsky (1997; 1998), Grimshaw & Samek-Lodovici (1995; 1998), Samek-Lodovici (1996) und Legendre (1998); vgl. auch Kapitel 6. Wenn es gelänge, diesen Beschränkungstyp in der Syntax als unverzichtbar zu erweisen, dann hätte das natürlich eine wichtige allgemeine Konsequenz: Optimalitätstheoretische Phonologie und Syntax wären einander nicht nur im Grundaufbau ähnlich, sondern darüber hinaus auch hinsichtlich der Form der Beschränkungen, und somit auch der Form der Kandidaten, über die die Beschränkungen reden.

Tatsächlich stärkt Grimshaw (1999) die Rolle von Ausrichtungsbeschränkungen in der Syntax. Sie zeigt, daß nicht nur (ein Teil von) PUR-EP mit Hilfe dieser Be-

¹Allerdings erfolgt bei den Exp-Beschränkungen immer noch der Bezug auf die *höchste* Schicht einer erweiterten Projektion.

schränkungen ableitbar ist, sondern auch ÖKON. Für den Fall der Kopfbewegung gilt z.B., daß ein Kandidat mit der abstrakten Struktur (7-a) und X-in-situ immer besser hinsichtlich der Ausrichtungsbeschränkungen KOPF-LINKS und KOPF-RECHTS abschneidet als der konkurrierende Kandidat (7-b) mit X-Bewegung; denn (7-b) fügt zu den Verletzungen von (7-a), die nun durch die Spur *t* ausgelöst werden, immer noch die neuen Verletzungen von X in der Zielposition hinzu. Kopfbewegung muß daher immer durch höher geordnete Beschränkungen erzwungen werden.²

- (7) a. $[_{YP} \text{Spez } Y [_{XP} \text{Spez } X_1 [\dots]]]$
 b. $[_{YP} \text{Spez } X_1+Y [_{XP} \text{Spez } t_1 [\dots]]]$

Diese Argumentation zeigt aber auch bereits, daß die Analogie von Ausrichtungsbeschränkungen in der Phonologie und in der Syntax nicht komplett ist. In der Phonologie sind diese Beschränkungen strikt Oberflächen-orientiert: Die Elemente, die es auszurichten gilt, haben konkrete phonologische Realisierungen. Für Spuren gilt dies normalerweise nicht. Daß Spuren als phonologisch leere Elemente in der Syntax phonologisch in einer höheren Domäne an deren Rand ausgerichtet werden müssen, ist daher zunächst einmal unerwartet und eine Abweichung vom Idealfall: Ganz gleich verhalten sich dann Phonologie und Syntax bei Ausrichtungsbeschränkungen eben doch nicht. Tatsächlich ist in manchen Arbeiten zur Syntax, die eine zur Phonologie identische Anwendung von Ausrichtungsbeschränkungen vorsehen, auch davon ausgegangen worden, daß Spuren für diese Beschränkungen unsichtbar sind (vgl. u.a. Legendre (1998), Costa (1999)). Nehmen wir für die Zwecke des Arguments einmal an, daß dem so ist, und daß die Beschränkungen, die Subjektbewegung nach SpecI und W-Bewegung auslösen (z.B. KASUS und OP-SPEZ bei Grimshaw (1997)) als Ausrichtungsbeschränkungen zu reformulieren sind: (i) Eine Nominativ-NP steht am linken Rand der Projektion des Finitums. (ii) Eine W-Phrase steht am linken Rand der höchsten Schicht der erweiterten Projektion, über die sie Skopus nehmen soll. Betrachten wir nun das Beispiel in (8).

- (8) $Who_1 \text{ do you think } [_{CP} t_1'' C t_1' \text{ was kissed } t_1] ?$

Gemäß der klassischen GB-Analyse ist t_1' notwendig, damit die Kette Kasus erhält; t_1'' wird aus Lokalitätsgründen gebraucht; und who_1 erfüllt in seiner Position das W-Kriterium. Sehen wir einmal von der Lokalitätsfrage ab. Dann ist immer noch klar, daß unter der Annahme, daß Spuren für Ausrichtungsbeschränkungen unsichtbar sind, (8) einen Beschränkungskonflikt illustriert, den Beschränkung (ii) zu ihren Gunsten entscheidet. Das aber bedeutet: Im Konfliktfall verletzt eine W-Phrase das Gebot, Kasus zu erhalten. Dies hat unerwünschte Konsequenzen. In einem Satz wie (9) sollte nämlich, wenn nicht mehr gesagt wird, keine Bindung des Reflexivs mehr möglich sein, weil es keine Spur von who_1 in einer A-Position geben kann, die

²Damit ÖKON als in toto abgeleitet gelten kann, muß natürlich noch Analoges für XP-Bewegung gezeigt werden.

die Anapher *c*-kommandiert.

(9) Who₁ do you think [_{CP} C (*t*'₁) seems to himself₁ [_{IP} *t*₁ to be intelligent]] ?

Generell ergibt sich unter dieser Perspektive, daß keine Bewegungsoperation mehr auf einer anderen Bewegungsoperation aufbauen kann; sukzessiv-zyklische Bewegung, also "Feeding" unter Bewegungsoperationen (desselben oder unterschiedlichen Typs), ist unmöglich. Insofern als sukzessiv-zyklische Bewegung in der Syntax als gut motiviert gelten kann, spricht dies potentiell gegen eine direkte Angleichung von Ausrichtungsbeschränkungen in Phonologie und Syntax – Ausrichtungsbeschränkungen in der Syntax müssen anders als ihre phonologischen Gegenstücke mindestens über abstrakte Elemente (Spuren) reden können.

2. Harmonische Ausrichtung

2.1. Hintergrund

Bislang haben wir eine Methode kennengelernt, in der optimalitätstheoretischen Syntax Beschränkungen zu generieren, nämlich die lokale Konjunktion (vgl. Abschnitt 3 von Kapitel 3). Eine zweite Methode, systematisch neue Beschränkungen zu erzeugen, ist die harmonische Ausrichtung ("harmonic alignment"). Bevor ich mich diesem Konzept näher zuwende, ist eine terminologische Klärung am Platze: Harmonische Ausrichtung darf nicht verwechselt werden mit den Ausrichtungsbeschränkungen ("alignment constraints"), die gerade im letzten Abschnitt behandelt wurden. Ausrichtungsbeschränkungen verlangen, daß bestimmte Elemente mit dem linken bzw. rechten Rand einer höheren Domäne in Deckung gebracht werden. Harmonische Ausrichtung dagegen ist ein Meta-Prinzip, das ausgehend von als universell postulierten Skalen in systematischer Weise Beschränkungen mit fester Ordnung erzeugt; was hier ausgerichtet wird, ist also nicht ein objektsprachliches Element mit einer Domäne, sondern eine Skala mit einer anderen.

Das Konzept der harmonischen Ausrichtung ist von Prince & Smolensky (1993, 136)) eingeführt worden und wie folgt definiert.

(10) *Harmonische Ausrichtung*:

Es sei D_1 eine binäre Dimension mit einer Skala $X > Y$ auf ihren Elementen $\{X, Y\}$ und D_2 eine andere Dimension mit einer Skala $a > b > \dots > z$ auf ihren Elementen $\{a, b, \dots, z\}$. Die *harmonische Ausrichtung* von D_1 und D_2 ist dann das Paar von Harmonieskalen H_X, H_Y .

a. $H_X: X/a \succ X/b \succ \dots \succ X/z$

b. $H_Y: Y/z \succ \dots \succ Y/b \succ Y/a$

Die *Beschränkungsausrichtung* ist das Paar von Beschränkungssubhierarchien C_X, C_Y .

a. $C_X: *X/z \gg \dots \gg *X/b \gg *X/a$

b. $C_Y: *Y/a \gg *Y/b \gg \dots \gg *Y/z$

Gemäß (10) werden zunächst aus zwei unabhängig gegebenen Skalen (Hierarchien), von denen eine nur zwei Elemente hat, sog. Harmonieskalen erzeugt, eine für das erste Element der zwei-elementigen Skala, eine für das zweite Element dieser Skala. Das Ordnungssymbol \succ in (10) ist zu lesen als “ist weniger markiert als” oder “ist harmonischer als”. Am harmonischsten bzw. wenigsten markiert sind in (10) also der Fall, wo ein X ein a ist (also die oberen Enden der beiden Ausgangsskalen miteinander ausgerichtet sind), und der Fall, wo ein Y ein z ist (also die unteren Enden der beiden Ausgangsskalen miteinander ausgerichtet sind). Am stärksten markiert ist es für ein X, ein z zu sein, und für ein Y, ein a zu sein: Hier sind jeweils das höchste Element der einen Skala und das tiefste Element der anderen Skala miteinander verbunden worden. Was im letzten Schritt, der Beschränkungsausrichtung, erfolgt, ist eine Präfigierung der Elemente der Harmonieskalen durch *, eine Umtauschung der Reihenfolge und eine Ersetzung von \succ durch \gg . Somit ergeben sich fixe Hierarchien von Beschränkungen wie $*\alpha/\beta$ (“Vermeide ein α als β ”). Am höchsten geordnet (und somit am wenigsten leicht zu verletzen) ist dabei die Beschränkung, die die markierteste Kombination verbietet (in (10) z.B. $*X/z$); am tiefsten geordnet (und somit am leichtesten zu verletzen) ist die Beschränkung, die die am wenigsten markierte Kombination verbietet (in (10) z.B. $*X/a$).

Ein Beispiel mag dies verdeutlichen. Der Kontext, in dem Prince & Smolensky (1993) harmonische Ausrichtung einführen, ist die Diskussion ihrer phonologischen Beschränkung H-NUK (“Harmonie des Nukleus”). Diese Beschränkung wählt den Kandidaten mit der größten Sonorität im Silbennukleus aus; alle anderen Kandidaten mit geringerer Sonorität im Nukleus verletzen diese Beschränkung in unterschiedlich starkem Maße, je nach Sonoritätsgrad. Wie ein Kandidat K_i H-NUK erfüllt, ist somit nicht lokal, unter Einbeziehung lediglich von Eigenschaften von K_i selbst, überprüfbar; stattdessen entscheidet sich das Verhalten von K_i bzgl. H-NUK auf der Basis der Eigenschaften anderer Kandidaten. Dies bedeutet jedoch, daß H-NUK eine translokale Beschränkung ist (vgl. Kapitel 4). Dies ist aus Komplexitätserwägungen unattraktiv, weil es sich bei der Ermittlung des Verhaltens der Kandidaten hinsichtlich von H-NUK um ein zusätzliches (wenn auch einfaches) Optimierungsverfahren handelt. Als Alternative schlagen Prince & Smolensky (1993) die harmonische Ausrichtung auf der Basis zweier Hierarchien für Position und Sonorität in der Silbe in (11) vor:

- (11) a. *Positionsprominenz:*
 $P > M$ (Silbengipfel (“Peak”) vor Silbenrand (“Margin”))
 b. *Sonoritätshierarchie:*
 $a > i > \dots > t$

Harmonische Ausrichtung dieser beiden Skalen (von denen die erste zwei-elementig ist) erzeugt dann die beiden Harmonieskalen in (12-a). Die erste Harmonieskala H_P besagt, daß es am wenigsten markiert für einen Silbengipfel ist, ein “a” zu sein, und am markiertesten, ein “t” zu sein (und daß jedenfalls ein Vokal als Silbengipfel immer besser als ein Konsonant ist). Die zweite Harmonieskala H_M besagt für den

Silbenrand das Gegenteil (hier ist ein Konsonant immer weniger markiert als ein Vokal). Per Beschränkungsausrichtung werden auf dieser Basis Beschränkungshierarchien mit fixer Ordnung erzeugt; vgl. (12-b).

- (12) a. *Harmonische Ausrichtung:*
 (i) $H_P: P/a \succ P/i \succ \dots \succ P/t$
 (ii) $H_M: M/t \succ \dots \succ M/i \succ M/a$
 b. *Beschränkungsausrichtung:*
 (i) $C_P: *P/t \gg \dots \gg *P/i \gg *P/a$
 (ii) $C_M: *M/a \gg *M/i \gg \dots \gg *M/t$

Die Beschränkungshierarchie C_P bewirkt, daß das Verbot von stimmlosen Plosiven wie “t” im Silbengipfel immer stärker ist als das Verbot von stimmhaften Plosiven, daß dies wiederum stärker ist als das Verbot eines jeden Vokals, und daß schließlich das Verbot von “a” im Silbengipfel am schwächsten ist. Wenn also ein Segment im Silbengipfel plaziert werden muß, ergibt sich aus C_P , daß ein “a” hier die Verletzung der am tiefsten geordneten Beschränkung auslöst, ein “t” dagegen die Verletzung der höchstgeordneten Beschränkung. Auf diese Weise leistet die Silbengipfelsubhierarchie C_P alles, was vorher H-NUK geleistet hat (C_M ist für diesen Zweck nicht notwendig). Zudem gibt es einen konzeptuellen und einen empirischen Vorteil: Einerseits ist die Verwendung einer translokalen Beschränkung vermieden; andererseits ist das resultierende System flexibler, weil andere Beschränkungen an beliebigen Stellen der Subhierarchie interpoliert werden können.

2.2. GF-Realisierung bei Aissen (1999)

Diese zunächst einmal für die Phonologie entwickelte Methode der Beschränkungs-erzeugung hat Aissen (1999) auf die Syntax angewendet.³ Die Grundfrage, der Aissen nachgeht, betrifft die Zuordnung von Argumenten des Verbs zu grammatischen Funktionen (GFs) wie Subjekt, direktes Objekt usw. Wenn wir das Deutsche betrachten, dann richtet sich diese Zuordnung im wesentlichen nach den Thetarollen der Argumente. Bei einem zweistelligen Verb wird im Aktivsatz die höhere Thetarolle (z.B. Agens) dem Argument in der Subjektposition zugeordnet, die tiefere Thetarolle (z.B. Thema) dem Argument in der Objektposition. Im Passivsatz wird dagegen beim selben Verb das Thema zum Subjekt, und das Agens kann als oblique PP (*von*-Phrase) realisiert werden. Diese Thetarollen-basierte Zuordnung von Argumenten zu GFs erweist sich jedoch bei Einbeziehung von Evidenz aus anderen, typologisch fernen Sprachen als zu einfach. Insbesondere scheint in manchen Sprachen die Zuordnung von Argumenten zu GFs nicht nur von der Thetarolle, sondern auch von der durch das Argument realisierten Person-Markierung abzuhängen.

Aissens Idee ist nun, daß es drei Skalen gibt, die jeweils durch harmonische Aus-

³Vgl. auch Artstein (1998), eine Arbeit, die auf (einer Vorversion von) Aissen (1999) beruht.

richtung mit einer vierten Skala, der GF-Hierarchie, Beschränkungsausrichtungen mit fester interner Ordnung erzeugen. Durch die feste Ordnung können implikative Universalien abgeleitet werden. Weil die Subhierarchien aber frei miteinander interpoliert werden können, ist das System gleichzeitig sehr flexibel. Die Basisskalen sind in (13) aufgelistet.⁴

(13) *Skalen:*

- a. *GF-Hierarchie(n):*
Subjekt > Nicht-Subjekt (Subjekt > Objekt, Subjekt > Obliques Objekt)
- b. *Theta-Hierarchie:*
Agens > Patiens
- c. *Prominenz-Hierarchie:*
X > x (diskursprominentes Argument > nicht prominentes Argument)
- d. *Person-Hierarchie:*
Lokale Pers. (1, 2) > 3. Pers. (1 > 2, 1 > 3)

Hierzu zwei Bemerkungen. Erstens sind die Elemente der Hierarchien in (13-bcd) auf einem Argument im Input vermerkt; die GF entsteht dagegen erst im Output. Zweitens muß bei harmonischer Ausrichtung immer eine Skala binär sein. Daher ist es manchmal notwendig, größere Hierarchien zu zerlegen (wie z.B. für die GF-Hierarchie in (13-a) gezeigt). Im folgenden sind jeweils unter (b) die durch (in (a) dokumentierte) harmonische Ausrichtung der GF-Hierarchie mit einer der drei anderen Skalen entstehenden Beschränkungshierarchien dargestellt.

(14) *Harmonische Ausrichtung von GF-Hierarchie und Theta-Hierarchie :*

- a. (i) $H_{Su}: Su/Agens \succ Su/Patiens$
(ii) $H_{Nicht-Su}: Nicht-Su/Patiens \succ Nicht-Su/Agens$
- b. (i) $C_{Su}: *SU/PATIENS \gg *SU/AGENS$
(ii) $C_{Nicht-Su}: *NICHT-SU/AGENS \gg *NICHT-SU/PATIENS$

(15) *Harmonische Ausrichtung von GF-Hierarchie und Prominenz-Hierarchie :*

- a. (i) $H_{Su}: Su/X \succ Su/x$
(ii) $H_{Nicht-Su}: Nicht-Su/x \succ Nicht-Su/X$
- b. (i) $C_{Su}: *SU/x \gg *SU/X$
(ii) $C_{Nicht-Su}: *NICHT-SU/X \gg *NICHT-SU/x$

(16) *Harmonische Ausrichtung von GF-Hierarchie und Person-Hierarchie :*

- a. (i) $H_{Su}: Su/Lokal \succ Su/3$
(ii) $H_{Nicht-Su}: Nicht-Su/3 \succ Nicht-Su/Lokal$
- b. (i) $C_{Su}: *SU/3 \gg *SU/LOKAL$
(ii) $C_{Nicht-Su}: *NICHT-SU/LOKAL \gg *NICHT-SU/3$

⁴Diese Skalen sind z.T. unvollständig. Zum Beispiel fehlt bei der GF-Hierarchie das indirekte Objekt; bei der Thetahierarchie fehlen andere Rollen wie Experiencer, Thema, etc.

Aus (14) ergibt sich, daß das Verbot, ein Patiens im Input zum Subjekt im Output zu machen, universell höher geordnet ist als das Verbot, ein Agens im Input zum Subjekt im Output zu machen; gerade umgekehrt liegt der Fall bei Nicht-Subjekten. Wenn weiter nichts gesagt wird, heißt das zunächst einmal, daß ein Agens zum Subjekt und ein Patiens zum Nicht-Subjekt (Objekt) wird. Dies ist in Sprachen wie dem Deutschen für Aktivsätze ja auch erwünscht. In Passivsätzen wird jedoch ein Patiens ein Subjekt. Folgerichtig muß es eine Beschränkung geben, die *SU/PATIENS dominiert. Aissen nimmt an, daß diese Situation entstehen kann, wenn die Beschränkung *SU/X aus der Beschränkungshierarchie in (15) über *SU/PATIENS steht. Ist ein Patiens in einem gegebenen Input als prominenter als ein Agens klassifiziert (wobei Prominenz ein diskurspragmatisches Merkmal ist, das mit "X" kodiert wird, fehlende Prominenz mit "x"), sorgt die Ordnung *SU/X \gg *SU/PATIENS dafür, daß das Patiens zum Subjekt wird, und es entsteht Passiv (das Agens wird dann zum obliquen Argument).⁵ Die Beschränkungen in (16) schließlich können das so entstehende Bild wieder verwischen. Aus der festen Ordnung dieser Beschränkungen folgt, daß das Verbot von Subjekten in der 3. Person immer höher steht als das Verbot von Subjekten in einer lokalen (1. oder 2.) Person, und daß das Verbot von Nicht-Subjekten in einer lokalen Person immer höher steht als das Verbot von Nicht-Subjekten in der 3. Person. Mit anderen Worten: Im Idealfall sind Subjekte 1. oder 2. Person und Objekte 3. Person. Ist nun z.B. die Beschränkung *OBL/LOKAL ("Vermeide oblique Argumente, die lokale Personen sind"; zurückführbar auf die Skala Subjekt > Obliques Objekt, die eine Instanz der Skala Subjekt > Nicht-Subjekt ist) über *SU/X \gg *SU/PATIENS geordnet, dann kann ein prominentes Patiens (bei nicht-prominentem Agens) nur dann zum Subjekt werden, wenn das nunmehr als obliques Argument zu realisierende Agens nicht 1. oder 2. Person ist. Diese und andere Konsequenzen seien nun kurz anhand der Ordnungen einiger Sprachen dargestellt.⁶

Fox In der nordamerikanischen Indianersprache Fox gibt es kein Passiv. Natürlich kann aber in dieser Sprache die Situation eintreten, daß ein Patiens prominenter ist als ein Agens. Passiv, also die Zuordnung eines Patiens-Argumentes zur Subjektposition, wird jedoch unterdrückt bei einer Ordnung, in der die *GF/Prom-Beschränkungen (insbesondere *SU/X) sowie die *GF/Pers-Beschränkungen unterhalb der *GF/Theta-Beschränkungen (insbesondere *SU/PATIENS) stehen.

⁵Die Grundintuition ist hier also, daß die Operation Passiv dadurch legitimiert ist, daß das prominenteste Argument zum Subjekt wird.

⁶Aissen (1999) verzichtet auf die Anführung natürlich-sprachlicher Beispiele und begnügt sich jeweils mit einem Verweis auf die einschlägige Literatur. Zum leichteren Verständnis gebe ich hier die deutschen Entsprechungen wieder.

(17) *Ordnung im Fox:*

*Su/Patiens \gg ... \gg { *GF/Pers., *GF/Prom }

Den Wettbewerb zeigt Tabelle T_{8.1}. Selbst wenn das Agens nicht-prominent (x) und das Patiens prominent (X) ist (was “normalerweise” Passiv auslöst), gewinnt der Aktivkandidat K₁. Natürlich gewinnt K₁ auch, wenn das Agens prominent ist oder wenn sowohl Agens, als auch Patiens nicht-prominent sind. Im vorliegenden Fall erfolgt also so etwas wie Neutralisierung bzgl. der X/x-Spezifikation, und zwar aufgrund der hohen Ordnung von *SU/PAT.⁷ Die *GF/PERS-Beschränkungen spielen hier aufgrund tiefer Ordnung noch keine Rolle.

T_{8.1}: *Unterdrücktes Passiv im Fox*

Input: V (Agens/1/x, Patiens/3/X)	*SU/PAT	*SU/X	*GF/PERS
☞ K ₁ : Ich liebe sie		*	**
K ₂ : Sie wird von mir geliebt	*!		**

Englisch/Deutsch Angenommen, Englisch und Deutsch unterscheiden sich von Fox hinsichtlich der hier relevanten Beschränkungen nur dadurch, daß *SU/X höher geordnet ist als *SU/PATIENS:

(18) *Ordnung im Englischen/Deutschen:*

*SU/X \gg *SU/PATIENS \gg ... \gg *GF/PERS.

Dann resultiert eine klassische Aktiv-Passiv-Diathese. Ist das Agens (allgemeiner, die höchste Thetarolle) im Diskurs genauso (viel oder wenig) prominent wie das Patiens (bzw. die zweithöchste Thetarolle), oder auch prominenter, kann *SU/X nicht zwischen den Kandidaten diskriminieren, und die Entscheidung wird der Beschränkung *SU/PATIENS übergeben, die zugunsten des Aktivkandidaten (Realisierung des Agens als Subjekt) entscheidet. Vgl. Tabelle T_{8.2}.

T_{8.2}: *Aktiv im Englischen/Deutschen*

Input: V (Agens/1/x, Patiens/3/x)	*SU/X	*SU/PAT	*GF/PERS
☞ K ₁ : Ich liebe sie	*		**
K ₂ : Sie wird von mir geliebt	*	*!	**

Liegt dagegen im Input ein Patiens (die zweithöchste Thetarolle) vor, das im Diskurs prominenter ist als das Agens (die höchste Thetarolle), dann fällt die Ent-

⁷Strenggenommen erfolgt *keine* Neutralisierung der X/x-Merkmale, denn diese Merkmale werden an sich vom Input zum Output unangetastet gelassen. Der Neutralisierungseffekt liegt hier daran, daß die typische, durch *SU/X vermittelte Wirkung von Patiens/X- und Agens/x-Spezifikation im Input (nämlich Passiv) nicht eintritt.

scheidung *SU/X zu, und das Patiens wird als Subjekt realisiert: Das Verbot von Patiens-Subjekten ist nicht so stark wie das Verbot von nicht-prominenten Subjekten. Dies veranschaulicht Tabelle T_{8.3}.

T_{8.3}: *Passiv im Englischen/Deutschen*

Input: V (Agens/1/x, Patiens/3/X)	*SU/X	*SU/PAT	*GF/PERS
K ₁ : Ich liebe sie	*!		**
☞K ₂ : Sie wird von mir geliebt		*	**
K ₃ : Sie liebt von mir		*	**

Lushootseed Das so entstehende Bild kann wie erwähnt verwischt werden durch den Einfluß hoch geordneter *GF/PERSON-Beschränkungen. Die bereits beschriebene Situation (hoch geordnetes *OBL/LOKAL) wird durch das Lushootseed instantiiert, eine Indianersprache im amerikanischen Westen, die zur Salish-Gruppe gehört. Lushootseed verhält sich hinsichtlich der Aktiv-Passiv-Diathese fast wie Englisch oder Deutsch. Der zentrale Unterschied ist jedoch, daß lokale Personen nicht Agens im Passiv sein können, d.h., nicht als oblique Objekte realisiert werden können. Die von Aissen (1999) angenommene Ordnung im Lushootseed ist die in (19).

(19) *Ordnung im Lushootseed*:

*OBL/LOKAL ≫ *SU/X ≫ *SU/PATIENS ≫ *GF/PERS.

Für das Aktiv ändert sich somit gegenüber dem Englischen/Deutschen nichts (die hoch geordnete Beschränkung *OBL/LOKAL ist hier irrelevant). Beim Passiv gibt es jedoch dann einen Unterschied, wenn das Agens eine lokale (1. oder 2.) Person ist. Wenn das Agens nicht-prominent und das Patiens prominent ist, erweist sich anstelle des von *SU/X favorisierten Passiv-Kandidaten K₂ der von *OBL/LOKAL (und, irrelevanterweise von *SU/PATIENS) geforderte Aktiv-Kandidat K₁ als optimal. Hier findet also wieder Neutralisierung statt; vgl. Tabelle T_{8.4}.

T_{8.4}: *Unterdrücktes Passiv bei lokalem Agens im Lushootseed*

Input: V (Agens/1/x, Patiens/3/X)	*OBL/LOKAL	*SU/X	*SU/PAT	*GF/PERS
☞K ₁ : Ich liebe sie		*		**
K ₂ : Sie wird von mir geliebt	*!		*	*

Lummi Noch einen Schritt weiter als Lushootseed geht das Lummi, das ebenfalls zur Familie der Salish-Sprachen gehört. Lummi ist fast wie Lushootseed: Wenn ein Agens eine lokale Person ist, kann es im Passiv nicht als obliques Argument auftreten; im Zweifelsfall (bei entsprechender Relevanz von Prominenz) wird wie gesehen zum Aktiv neutralisiert. Zusätzlich gilt aber noch, daß *OBJ/LOKAL hoch

geordnet ist und somit ein lokales (1. oder 2. Person) Patiens nicht Objekt im Aktiv sein kann (hier wird zum Passiv neutralisiert). Nicht-komplementäre Verteilung von Aktiv und Passiv entsteht somit nur, wenn beidesmal die 3. Person vorliegt. Die für das Lummi angenommene Ordnung ist (20).

(20) *Ordnung im Lummi:*

*OBL/LOKAL \gg *OBJ/LOKAL \gg *SU/X \gg *SU/PATIENS \gg *GF/PERS.

Die gegenüber dem Lushootseed zusätzliche Konsequenz im Lummi – nicht durch Diskursprominenz getriebenes Passiv bei lokalem Patiens in Input – illustriert Tabelle T_{8.5}.

T_{8.5}: *Unterdrücktes Aktiv bei lokalem Patiens im Lummi*

Input: V (Agens/3/X, Patiens/1/x)	*OBL/ LOK	OBJ/ LOK	*SU/ X	*SU/ PAT	*GF/ PERSON
K ₁ : Sie liebt mich		*!			*
☞K ₂ : Ich werde von ihr geliebt			*	*	**

Es stellt sich die Frage, was passiert, wenn Agens und Patiens beide lokale Personen sind. An dieser Stelle wird die Ordnung *OBL/LOKAL \gg *OBJ/LOKAL wichtig, denn es findet immer Neutralisierung hin zum Aktiv statt. Dies zeigt Tabelle T_{8.6}, wo die Prominenzmarkierung der Argumente im Input zunächst einmal einen Passivsatz favorisieren würde (ebenso wie *OBJ/LOKAL), wo sich aber die höchstgeordnete Beschränkung *OBL/LOKAL zugunsten des Aktivsatzes durchsetzt.

T_{8.6}: *Unterdrücktes Passiv bei lokalem Agens und Patiens im Lummi*

Input: V (Agens/1/x, Patiens/2/X)	*OBL/ LOK	OBJ/ LOK	*SU/ X	*SU/ PAT	*GF/ PERSON
☞K ₁ : Ich liebe dich		*	*		*
K ₂ : Du wirst von mir geliebt	*!			*	*

Es ergibt sich somit bzgl. der betrachteten Sprachen eine ansteigende Sensitivität der (grundsätzlich durch den Faktor der Diskursprominenz gesteuerten) Aktiv-Passiv-Diathese für Person: Von Fox über Englisch/Deutsch, Lushootseed und Lummi zur hier nicht mehr besprochenen Salish-Sprache Squamish (die noch stärkere Personensensitivität aufweist).

So weit, so gut. Was damit noch nicht klar ist, ist, warum Aktiv und Passiv nicht nur einfach die Plazierung der Argumente betreffen, sondern eben auch noch einen sichtbaren Reflex haben: Aktiv- und Passivsätze haben typischerweise unterschiedliche morphosyntaktische Realisierungen; vgl. K₂ vs. K₃ in T_{8.3}, die zunächst einmal dasselbe Beschränkungsprofil aufweisen. Dieses Problem wird so bei Aissen (1999) nicht behandelt, dafür aber ganz ähnliche Probleme mit dem sichtbaren Reflex markierter (aber an sich erfolgreicher) Strategien (Kasusmarkierung, inverse

Verben) in Sprachen wie Dyirbal und Nocte. Die Grundidee ist, daß die Verletzung hoch geordneter Markiertheitsbeschränkungen in einem optimalen Kandidaten eher einen sichtbaren morphosyntaktischen Reflex induziert als die Verletzung tief geordneter Markiertheitsbeschränkungen in einem optimalen Kandidaten. Um diese Idee umzusetzen, wird auf das Konzept der lokalen Konjunktion Bezug genommen. Postuliert wird insbesondere, daß die Beschränkungen $*\emptyset$ und $*\text{STRUK}$ mit einer durch harmonische Ausrichtung entstandenen Hierarchie von Beschränkungen lokal konjugiert werden können, wobei das Ordnungsgesetz in (22) gilt.⁸

- (21) a. $*\emptyset$:
Morphologische Kategorien müssen ausgedrückt werden.
b. $*\text{STRUK}$:
Sprachliche Struktur ist verboten.
- (22) Lokale Konjunktion von C_1 mit einer Beschränkungshierarchie [$C_2 \gg C_3 \gg \dots \gg C_n$] ergibt die Beschränkungshierarchie [$C_1 \ \& \ C_2 \gg C_1 \ \& \ C_3 \gg \dots \gg C_1 \ \& \ C_n$].

Angenommen nun, zwei Markiertheitsbeschränkungen $*A/C$, $*A/B$ sind durch harmonische Ausrichtung entstanden, mit $*A/C$ als inhärent höher geordneter Beschränkung. Wir wollen ableiten, daß nicht-fatale Verletzungen von $*A/C$ einen morphologischen Reflex M (wie z.B. eben im Passiv) haben, von $*A/B$ nicht (wie z.B. im Aktiv). Wir konjugieren $*\emptyset_M$ mit $*A/C \gg *A/B$ und erhalten wg. (22) $*\emptyset_M \ \& \ *A/C \gg *A/B$. Wegen der allgemeinen Regel für lokale Konjunktion (vgl. Kapitel 3) gilt: $*\emptyset_M \ \& \ *A/C \gg \dots \gg *A/B$, $*A/C \gg \dots \gg *A/B$. Jetzt folgt zunächst einmal, daß der optimale Kandidat, der $*A/C$ verletzt, und der optimale Kandidat, der $*A/B$ verletzt, beide einen morphologischen Reflex haben müssen. Es muß also nun noch über $*\emptyset_M \ \& \ *A/B$ die Beschränkung $*\text{STRUK}_M$ interpoliert werden: $*\emptyset_M \ \& \ *A/C \gg \dots \gg *A/B$, $*\text{STRUK}_M \gg \dots \gg *A/B$, $*A/C \gg \dots \gg *A/B$. Jetzt hat nur noch der optimale Kandidat, der $*A/C$ verletzt, einen morphologischen Reflex, wie erwünscht.

Zusammengenommen sieht es so aus, als sei harmonische Ausrichtung ein gutes Mittel, um flexibel und doch mit typologischen Vorhersagen die Frage anzugehen, mit welcher GF ein Input-Argument abhängig von Thetarolle, Person und Prominenz im Output realisiert wird. Ableitbar ist, daß Sprachen in manchen Fällen markierte Strategien wählen müssen, die hoch geordnete, durch harmonische Ausrichtung gewonnene Beschränkungen verletzen. In solchen Fällen entsteht oft ein morpho-syntaktischer Reflex. Dieser wird mit Hilfe des Konzepts der lokalen Konjunktion abgeleitet; die resultierenden Analysen werden jedoch sehr schnell sehr komplex und unübersichtlich.

⁸ $*\emptyset$ hier ist nicht zu verwechseln mit der oben mehrfach erwähnten Beschränkung $*\emptyset$, die leere Kandidaten verbietet.

3. Mehrfache Optimierung

Prince & Smolensky (1993) unterscheiden zwischen zwei grundsätzlichen Möglichkeiten der Optimierung, harmonischem Parallelismus und harmonischem Serialismus. Gemäß harmonischem Parallelismus werden Kandidaten beliebiger interner Komplexität – z.B. in der Syntax: $\langle D\text{-Struktur}, S\text{-Struktur}, LF \rangle$ -Tupel wie in der GB-Theorie – in einem einzigen Durchgang dem Wettbewerb ausgesetzt, und ein Kandidat, der diesen Wettbewerb als bester beendet, ist grammatisch. Prince & Smolensky (1993) entscheiden sich für dieses Optimierungsverfahren, und es ist seither in der Phonologie wie in der Syntax Standard. Im Einklang damit ist auch in diesem Buch bisher (bis auf wenige Andeutungen) immer vorausgesetzt worden, daß syntaktische Optimierung eine Instanz von harmonischem Parallelismus ist. In jüngster Zeit hat aber auch der harmonische Serialismus einige Beachtung gefunden. Diesem Verfahren zufolge wird nicht in einem Rutsch optimiert, sondern mehrfach hintereinander. Dadurch läßt sich in der Phonologie das klassische Konzept der Derivation optimalitätstheoretisch umsetzen; vgl. McCarthy (1999), Rubach (2000) und die Beiträge in Hermans & van Oostendorp (2000). Die auf harmonischem Serialismus basierende Alternative zur syntaktischen Parallel-Optimierung von z.B. $\langle D\text{-Struktur}, S\text{-Struktur}, LF \rangle$ -Tupeln wäre entsprechend eine Optimierung zunächst von D-Strukturen; dann von S-Strukturen; und schließlich von LF-Repräsentationen.⁹ Im folgenden sollen kurz einige Vorschläge für harmonischen Serialismus in der Syntax vorgestellt werden.

3.1. Bindungstheorie bei Wilson (1998)

In Kapitel 4 haben wir bereits gesehen, daß es im Bereich der Kontrolle widerstrebende Evidenz bezüglich der Kandidatenmengen gibt. Für die optimalitätstheoretische Rekonstruktion der Annahme von Stechow & Sternefeld (1988), daß lexikalische Kontrolleigenschaften die strukturellen Kontrollbedingungen außer Kraft setzen können, ist es notwendig, daß Kandidaten mit unterschiedlichen Indizes (und also auch unterschiedlicher Interpretation) konkurrieren können. Für die optimalitätstheoretische Rekonstruktion von Chomskys (1981) Prinzip “Vermeide Pronomina” ist es dagegen notwendig, daß zwar Kandidaten mit unterschiedlicher Rea-

⁹Auf den ersten Blick mag es so scheinen, als sei der Unterschied zwischen harmonischem Parallelismus (Einfachoptimierung) und harmonischem Serialismus (Mehrfachoptimierung) gleichzusetzen dem Unterschied zwischen repräsentationeller und derivationeller Grammatikorganisation. Das ist jedoch nicht ganz so. Zum einen kann man sich serielle Optimierungsverfahren vorstellen, die nicht inhärent gerichtet sind (s.u.); das resultierende System ist dann repräsentationell. Zum anderen ist es, wie bereits in Kapitel 1 vermerkt wurde, möglich, daß Gen komplette Derivationen erzeugt, die dann einer Einfachoptimierung ausgesetzt werden, die die optimale Derivation für sich entscheidet (vgl. Müller (1997)). Was man somit nur gleichsetzen kann, sind harmonischer Parallelismus und “repräsentationelle Harmonieevaluation”, sowie gerichteter harmonischer Serialismus und “derivationelle Harmonieevaluation”.

lisierung von Argumenten (Pronomen vs. PRO) konkurrieren können, aber nicht Kandidaten mit unterschiedlichen Indizes. Wilson (1998) konstatiert auf analoge Weise Evidenz für zwei unterschiedliche Kandidatenmengendefinitionen im Bereich der Bindungstheorie. Seine Konklusion ist, daß in der Syntax tatsächlich zweimal optimiert wird: zuerst bzgl. der Interpretation, dann bzgl. des formalen Ausdrucks. Etwas vereinfacht läßt sich das serielle Optimierungssystem von Wilson wie folgt darstellen.

Am Anfang steht ein komplexer Input I_i . Diesen Input können wir uns für die gegenwärtigen Zwecke so vorstellen wie bei Legendre, Smolensky & Wilson (1998) beschrieben: als Ansammlung von Inhaltswörtern, die sowohl syntaktisch strukturiert als auch mit semantischer Interpretation versehen ist. Insbesondere liegen im Input bereits (i) eine Unterscheidung von Anaphern (Reflexiv- und Rezipropronomen) und (Personal-) Pronomen und (ii) eine vollständige Indizierung von NPs vor. Auf der Basis dieses Inputs wird von Gen die Kandidatenmenge erzeugt (die "universale Basis"). Die Elemente dieser Kandidatenmenge sind $\langle \text{Form}, \text{Interpretation} \rangle$ -Paare. Im ersten Schritt erfolgt *interpretative Optimierung*: Es werden nur solche Kandidaten betrachtet, die die syntaktische Form bewahren (also z.B. eine Input-Anapher nicht durch ein Output-Pronomen ersetzen), die semantische Interpretation (und hier vor allem die Indizes von NPs) aber abändern können. Je nach Ordnung von Markiertheits- und (semantischen) Treuebeschränkungen in einer Sprache behält der optimale Output O_i der ersten, interpretativen Optimierung dann entweder die Ausgangsbedeutung (in den von Wilson betrachteten Fällen heißt das: die Ausgangsindizierung) bei, oder er ändert sie ab.

Der optimale Kandidat der interpretativen Optimierung, O_i , ist der Input I_f für die *expressive Optimierung*. Es wird wieder eine Kandidatenmenge konstruiert. Die Kandidaten dieser neuen Kandidatenmenge sind erneut $\langle \text{Form}, \text{Interpretation} \rangle$ -Paare, aber nur solche, die erstens die Interpretation mit dem optimalen Kandidaten der interpretativen Optimierung identisch haben (also insbesondere die Indizes von NPs gleich lassen), und die zweitens – und diese nicht-triviale Zusatzannahme wird sich als wesentlich erweisen – selbst Gewinner *wenigstens einer* Interpretationsoptimierung sind (wohlgemerkt, nicht notwendigerweise derjenigen mit I_i als Input). Diese zweite Kandidatenmenge, die eine Teilmenge der universalen Basis ist, nennt Wilson das "Zwischeninventar". Nun wird wiederum gemäß der Ordnung von Markiertheits- und (syntaktischen) Treuebeschränkungen der syntaktische Ausdruck optimiert. Grammatisch ist ein Satz (ein $\langle \text{Form}, \text{Interpretation} \rangle$ -Output O_f), wenn er die expressive Optimierung als Gewinner beendet. Die Funktionsweise dieses Modells soll nun an (abstrakt gehaltenen) Beispielen gezeigt werden. Zwei für die interpretative Optimierung entscheidende Beschränkungen, die Wilson annimmt, sind (23-ab).

- (23) a. LOK-ANT ("Lokales Antezedens", "Local Antecedent"):
 Wenn eine Bindungsdomäne δ eine Anapher enthält, muß die Anapher in δ ein Antezedens haben.

b. BIND-TREUE (“Bindungstreue”; “Binding Faithfulness”):

Wenn zwischen zwei Elementen eine Bindungsbeziehung im Input vorliegt, muß sie auch im Output vorliegen.

LOK-ANT ist eine Markiertheitsbeschränkung, BIND-TREUE eine Treuebeschränkung. Unter der Ordnung $\text{LOK-ANT} \gg \text{BIND-TREUE}$ wird, wie Tabelle T_{8.7} zeigt, bei interpretativer Optimierung aus einem Input, in dem eine Anapher ein nicht-lokales Antezedens im Matrixsatz hat, der untreue optimale Output K_2 , in dem die Anapher an das nähere Antezedens in derselben Bindungsdomäne gebunden ist; der treue Output K_1 verletzt fatal LOK-ANT. Derselbe Gewinner ergibt sich für einen minimal unterschiedlichen Input, in dem die Anapher von vornherein innerhalb ihrer Bindungsdomäne gebunden ist; vgl. Tabelle T_{8.8}. Mit anderen Worten: Der Input-Unterschied bzgl. Koindizierung der Anapher wird neutralisiert.

T_{8.7}: Interpretative Optimierung bei nicht-lokaler Anaphernbindung

Input: [$\alpha_1 \dots$ $\delta \beta_2 \dots$ Anapher ₁ ...]	LOK-ANT	BIND-TREUE
K_1 : [$\alpha_1 \dots$ $\delta \beta_2 \dots$ Anapher ₁ ...]	*!	
$\hookrightarrow K_2$: [$\alpha_1 \dots$ $\delta \beta_2 \dots$ Anapher ₂ ...]		*

T_{8.8}: Interpretative Optimierung bei lokaler Anaphernbindung

Input: [$\alpha_1 \dots$ $\delta \beta_2 \dots$ Anapher ₂ ...]	LOK-ANT	BIND-TREUE
K_1 : [$\alpha_1 \dots$ $\delta \beta_2 \dots$ Anapher ₁ ...]	*!	*
$\hookrightarrow K_2$: [$\alpha_1 \dots$ $\delta \beta_2 \dots$ Anapher ₂ ...]		

Auf diese Weise werden viele auf Prinzip A der Bindungstheorie zurückführbare Effekte erfaßt. Man beachte aber, daß Wilsons Ansatz noch eine weitere Vorhersage macht: Angenommen, eine Anapher verletzt aufgrund eines nicht-lokalen Antezedens LOK-ANT, aber es gibt kein näheres Antezedens in derselben Bindungsdomäne. Dann kann auch keine Neutralisierung hin zur Bindung an ein anderes Antezedens erfolgen, und der optimale Kandidat kann LOK-ANT verletzen. Auf diese Weise werden trotz der Formulierung von LOK-ANT als strikte Minimalitätsbedingung relativierte Minimalitätseffekte für anaphorische Bindung z.B. im Marathi (einer indoeuropäischen, in Indien gesprochenen Sprache) erfaßt: Eine Anapher wird an das nächstmögliche Antezedens gebunden; wenn ein Satz kein potentielles Antezedens aufweist, kann eine Anapher von einem Argument im Matrixsatz gebunden werden.¹⁰

Auch für die expressive Optimierung sind eine Markiertheits- und eine Treuebeschränkung relevant:

¹⁰Zu Sprachen wie dem Deutschen ist dann noch mehr zu sagen: Hier sind Subjektanaphern oder Subjekt-interne Anaphern in finiten Sätzen nicht von einem Matrixargument bindbar, auch wenn kein potentielles Antezedens interveniert; d.h., es besteht im Bereich der anaphorischen Bindung rigide Minimalität, nicht relativierte Minimalität. Siehe unten.

- (24) a. REF-ÖKON (“Referentielle Ökonomie”, “Referential Economy”):
Ein Argument darf keine lexikalische Φ -Merkmalspezifikation besitzen.
- b. M-TREUE (“Merkmalstreue”, “Feature Faithfulness”):
Ein Argument mit der Merkmalspezifikation [F] im Input muß auch im Output für [F] spezifiziert sein.

Die für REF-ÖKON wesentliche Annahme ist, daß Anaphern typischerweise keine Φ -Merkmale (Person, Numerus, Genus) aufweisen, im Unterschied zu Pronomina. REF-ÖKON bestraft also grundsätzlich die Verwendung von Pronomina. Insofern hat diese Beschränkung eine ganz ähnliche Funktion wie *PERS-PRON in (49) von Kapitel 4: In solchen Positionen, wo an sich eine Anapher und ein Pronomen beide vorkommen können, wird die Anapher dem Pronomen vorgezogen (vgl. Burzio (1991), Fanselow (1991) und Richards (1997)). REF-ÖKON wird ausbalanciert durch die entgegengesetzte Beschränkung M-TREUE, die die Realisierung eines Input-Pronomens als Anapher verbietet. Unter der Ordnung REF-ÖKON \gg M-TREUE kann jedoch der Effekt der Verdrängung von Pronomina durch Anaphern entstehen. Die Tabellen T_{8.9} und T_{8.10} zeigen, daß bei dieser Ordnung ein Input-Pronomen ebenso wie eine Input-Anapher als Output-Anapher erscheint, d.h., daß hier wiederum Neutralisierung erfolgt.

T_{8.9}: *Expressive Optimierung bei Input-Pronomen*

Input: [α_1 ... Pronomen ₁ ...]	REF-ÖKON	M-TREUE
K ₁ : [α_1 ... Pronomen ₁ ...]	*!	
☞ K ₂ : [α_1 ... Anapher ₁ ...]		*

T_{8.10}: *Expressive Optimierung bei Input-Anapher*

Input: [α_1 ... Anapher ₁ ...]	REF-ÖKON	M-TREUE
K ₁ : [α_1 ... Pronomen ₁ ...]	*!	*
☞ K ₂ : [α_1 ... Anapher ₁ ...]		

Dieser Effekt ist in vielen Sprachen zu beobachten. Zum Wettbewerb in T_{8.9} ist aber noch mehr zu sagen. So, wie es bisher aussieht, ergibt sich die Konsequenz, daß bei expressiver Optimierung in allen Wettbewerben, wo ein Anaphernkandidat teilnimmt, dieser den Pronominakandidat blockiert, und zwar ungeachtet dessen, ob der Input für die expressive Optimierung ein Pronomen hat oder nicht. Die einzige Möglichkeit, daß eine Anapher ein Pronomen zum Ausdruck von Bindungsbeziehungen nicht blockiert, ist die, daß der Anaphernkandidat nicht der Gewinner irgendeiner interpretativen Optimierung ist, so daß er nicht mehr bei der expressiven Optimierung teilnehmen kann. Diese Konsequenz scheint nun für manche Sprachen korrekt (z.B. für das Chamorro, s.u.), für Sprachen wie das Deutsche aber nicht. An dieser Stelle wird eine weitere Annahme von Wilson wichtig: Alle Beschränkungen sind Teil einer einzigen Hierarchie, unabhängig davon, ob sie eher von der interpretativen oder der expressiven Optimierung benutzt werden. Nun

hat zwar REF-ÖKON für die interpretative Optimierung keine Konsequenzen (da per Annahme hier die syntaktische Form gleich bleiben muß); aber LOK-ANT kann sehr wohl die expressive Optimierung beeinflussen, da durch diese Beschränkung nicht nur interpretativ die Indizes bei Anaphern umgestellt, sondern auch expressiv Pronomina Anaphern vorgezogen werden können. Konkret gilt, daß bei einer Ordnung $\text{LOK-ANT} \gg \text{REF-ÖKON} \gg \text{M-TREUE} \gg \text{BIND-TREUE}$ nur dann wie in $T_{8.9}$ gezeigt im Einklang mit REF-ÖKON ein Input-Pronomen zu einer Output-Anapher neutralisiert wird, wenn die Anapher nicht LOK-ANT verletzt. Ebenso wird bei einer solchen Ordnung aus einer Input-Anapher, die LOK-ANT verletzt, aber trotz langer Bindung die interpretative Optimierung überlebt hat (weil es kein intervenierendes mögliches Antezedens gab), bei der expressiven Optimierung ein Output-Pronomen; genau diese Konsequenz ist für Sprachen wie das Deutsche (mit rigider Minimalität bei anaphorischer Bindung) erwünscht (vgl. Fußnote 10).

Zu den Sprachen, die interpretative wie expressive Neutralisierungseffekte aufweisen, weil jeweils die Markiertheitsbeschränkung die entsprechende Treuebeschränkung dominiert, gehört das austronesische Chamorro, das auf Guam gesprochen wird. Hier erfolgt also sowohl eine Neutralisierung langer anaphorischer Bindung zu kurzer anaphorischer Bindung an ein intervenierendes potentiell Antezedens, als auch eine Neutralisierung des Unterschiedes zwischen Pronomen und Anapher zur letzteren Variante bei kurzer Bindung. Da das Chamorro anders als z.B. das Deutsche relativierte Minimalität bei anaphorischer Bindung aufweist, kann hier LOK-ANT nicht hoch geordnet sein. Wilson schlägt für diese Sprache die Ordnung $\text{REF-ÖKON} \gg \text{M-TREUE} \gg \text{LOK-ANT} \gg \text{BIND-TREUE}$ vor. Es ist instruktiv, sich anzusehen, als welche optimalen Outputs unter dieser Ordnung verschiedene Inputs nach Mehrfachoptimierung erscheinen. Die Möglichkeiten sind in (25) zusammengestellt; Inputs, die im Output per Neutralisierung unmöglich sind, werden hier durch einen Stern markiert.

- (25) a. [$\alpha_1 \dots [\delta \dots \text{Anapher}_1 \dots]$]
 b. *[$\alpha_1 \dots [\delta \dots \text{Pronomen}_1 \dots]$]
 c. [$\alpha_1 \dots [\delta \beta_2 \dots \text{Anapher}_2 \dots]$]
 d. *[$\alpha_1 \dots [\delta \beta_2 \dots \text{Pronomen}_2 \dots]$]
 e. *[$\alpha_1 \dots [\delta \beta_2 \dots \text{Anapher}_1 \dots]$]
 f. [$\alpha_1 \dots [\delta \beta_2 \dots \text{Pronomen}_1 \dots]$]

(25-ab) überleben beide die interpretative Optimierung als optimal (sie konkurrieren da nicht, weil nur Form-identische Kandidaten betrachtet werden). Bei der darauffolgenden expressiven Optimierung blockiert jedoch (25-a) (25-b) wie gesehen als suboptimal (wg. REF-ÖKON). Dasselbe gilt für (25-cd). Der interessante Fall betrifft das Paar (25-ef). (25-f) ist unter interpretativer Optimierung optimal (denn LOK-ANT ist leer erfüllt), während (25-e) unter interpretativer Optimierung suboptimal ist (LOK-ANT ist fatal verletzt, und es wird zu (25-c) hin neutralisiert). Nun ist (25-f) Input für die expressive Optimierung; es wird eine Kandidatenmenge konstruiert, die die Indizierung gleich hält, aber die Form variieren lassen kann.

Wichtig ist nun, daß (25-e) *nicht* Teil dieser Kandidatenmenge sein darf; denn wenn das möglich wäre, dann sollte (25-f) im letzten Schritt noch von (25-e) blockiert werden (weil das Pronomen in (25-f) immer REF-ÖKON verletzt und diese Beschränkung daher zugunsten von (25-e) entscheidet). Daß (25-e) im letzten Schritt nicht mehr mit (25-f) konkurrieren kann, folgt nun aus Wilsons Stipulation bzgl. der Natur des “Zwischeninventars”: (25-e) kann deshalb nicht mit (25-f) bei expressiver Optimierung konkurrieren, weil (25-e) selbst nicht der Gewinner irgendeiner interpretativen Optimierung ist.

Diese Argumentation setzt zum einen die Gerichtetheit der beiden Optimierungen voraus, zum anderen die Annahme, daß nur interpretativ optimale Outputs an expressiver Optimierung noch teilnehmen können. Betrachten wir zunächst den ersten Punkt, und nehmen wir für die Zwecke des Arguments einmal an, daß die Kandidaten als $\langle \text{Form, Interpretation} \rangle$ -Paare nicht seriell, sondern parallel, also klassisch, evaluiert werden. Dann würde unter denselben Beschränkungen und derselben Beschränkungsordnung eine falsche Vorhersage gemacht werden, daß nämlich (25-f) kein optimaler Output werden kann. Der Grund ist folgender: Wie gesehen übersteht bei serieller Optimierung zwar (25-f) die interpretative Optimierung ohne Verletzung (der hier betrachteten Beschränkungen); bei expressiver Optimierung muß jedoch REF-ÖKON verletzt werden, was möglich ist, weil zu diesem Zeitpunkt kein Konkurrent mehr da ist, der diese Verletzung umgehen könnte. Wenn nun Form und Interpretation parallel evaluiert werden, dann ist der optimale Kandidat einer, der REF-ÖKON erfüllt (durch Wahl einer Anapher gegenüber einem Pronomen), und der darüber hinaus noch LOK-ANT respektiert (durch Umstellung des Bindungsindex auf der Anapher), also (25-c). Generell wäre die falsche Prognose, daß im vorliegenden Sprachtyp keine lange Bindung möglich sein sollte, und daß darüber hinaus Pronomina überhaupt nicht gebunden vorkommen könnten. Da es hierfür keine Evidenz gibt, spricht dies, wie Wilson betont, für serielle und gegen parallele Optimierung in der Bindungstheorie.¹¹

Weniger ausführlich geht Wilson auf seine zweite wesentliche Grundannahme ein, daß nur interpretativ optimale Outputs an expressiver Optimierung partizipieren. Wie gesehen würde sein System ohne diese Annahme *ceteris paribus* ebenfalls eine falsche Prognose machen: (25-f) sollte zwar nicht zu (25-c) neutralisiert werden, dafür jedoch zu (25-e), was nun selbst ganz unmöglich ist. Wilson benötigt hier somit eine zusätzliche Stipulation, die zwar durch die empirische Evidenz begründet scheint, aber konzeptuell keinesfalls unschuldig ist: Um herauszufinden, ob bei einem gegebenen Wettbewerb der expressiven Optimierung mit Input I_1 neben dem I_1 -treuen Output O_1 ein I_1 -untreuer Kandidat O_2 teilnimmt oder nicht,

¹¹An dieser Stelle sei vermerkt, daß Wilson selbst sein System “bidirektionale” Optimierung nennt. Dieser Begriff ist nicht ganz glücklich gewählt, da das System wie soeben gesehen nicht bidirektional, sondern inhärent unidirektional ist. Ähnliches gilt im übrigen für das im Grundaufbau partiell vergleichbare “bidirektionale” Mehrfachoptimierungssystem in Kuhn (2000).

muß man herausfinden, ob O_2 Gewinner irgendeines anderen Wettbewerbs der interpretativen Optimierung ist; und dieses Verfahren ist für *jeden* mit O_1 potentiell konkurrierenden Output durchzuführen. Das bedeutet, daß die Entscheidung über die Aufnahme eines Outputs in einen Wettbewerb der expressiven Optimierung nicht auf der Basis lokal zugänglicher Information (wie Rückführbarkeit auf denselben Input oder identische Interpretation) erfolgen kann, und daß somit ein Komplexitätsanstieg unvermeidbar scheint.

Abschließend sei eine Variante von Wilsons (1998) Modell betrachtet, die unter Umständen konzeptuell etwas einfacher ist, das eben angesprochene Problem aber noch deutlicher zutage treten läßt. Man nehme an, daß Gen auf der Basis des Ursprungs-Inputs zunächst Kandidaten erzeugt, die die syntaktische Form des Inputs intakt lassen und nur (potentiell) die Interpretation (die Indizes) manipulieren. Per interpretativer Optimierung wird dann wie vorher der optimale Output erzeugt. Dieser fungiert als Input, auf dessen Basis – und das ist die wesentliche Änderung – Gen nochmals appliziert, um die Kandidaten für die zweite, expressive Optimierung zu erzeugen. Die Kandidaten halten jetzt die Interpretation fest und ändern nur (potentiell) die syntaktische Form ab, wobei sie – dies integriert Wilsons Stipulation bzgl. des “Zwischeninventars” – nur solche Formen haben dürfen, die in irgendeiner anderen Interpretationsoptimierung als optimal erschienen sind; an dieser Stelle wird die Nicht-Lokalität des Verfahrens deutlich. Der optimale Output der zweiten Optimierung ist dann ein grammatischer Satz. Diese Variante von Wilsons Vorschlag kommt einer optimalitätstheoretischen Generativen Semantik sehr nahe: Das Resultat der ersten, interpretativen Optimierung kann als artikulierte D-Struktur angesehen werden (die Eigenschaften der LF in der GB-Theorie inkorporiert); das Resultat der zweiten, expressiven Optimierung läßt sich als S-Struktur betrachten.

3.2. Wortstellung und Quantorenskopus bei Heck (1999)

Wenn man die GB-Annahme macht, daß jeder Satz Repräsentationen auf der D-Struktur, auf der S-Struktur und auf LF hat, gibt es zwei naheliegende Möglichkeiten der optimalitätstheoretischen Umsetzung. Die erste Möglichkeit ist, daß die Kandidaten als $\langle \text{D-Struktur}, \text{S-Struktur}, \text{LF} \rangle$ -Tupel betrachtet werden. Die Optimierung erfolgt parallel, in einem Durchgang: Ein Kandidat ist optimal, wenn er die optimale D-Struktur, die optimale S-Struktur und die optimale LF-Repräsentation hat. Die zweite Möglichkeit ist dagegen serielle Optimierung; ein solches Verfahren ist von Heck (1999) anhand von deutschen Daten zur Wortstellung und zum relativen Quantorenskopus entwickelt worden. In diesem Ansatz erzeugt Gen auf der Basis einer Ansammlung lexikalischen Materials (der Numeration) zunächst eine Kandidatenmenge mit D-Struktur-Outputs $\{O_{d/1}, O_{d/2}, \dots, O_{d/n}\}$. Der optimale D-Struktur-Kandidat $O_{d/i}$ wird dann als Input $I_{s/i}$ benutzt für die Generierung von S-Strukturkandidaten $\{O_{s/1}, O_{s/2}, \dots, O_{s/n}\}$. Hierauf erfolgt S-Struktur-Optimierung. Der Gewinner dieser Optimierung $O_{s/j}$ entspricht dem, was ausgesprochen wird;

außerdem ist er der Input $I_{lf/j}$ für eine letzte Applikation von Gen. Es werden die Kandidaten der LF-Kandidatenmenge $\{O_{lf/1}, O_{lf/2}, \dots, O_{lf/n}\}$ erzeugt, und der optimale Output $O_{lf/k}$ legt die Interpretation fest. In diesem Modell ist also das Verhältnis von syntaktischer und semantischer Optimierung gegenüber dem Modell von Wilson (1998) vertauscht: Zunächst wird die Oberflächenform optimiert (S-Struktur), danach die Interpretation (LF), und nicht umgekehrt.

Wie Heck (1999) sich die D-Struktur-Optimierung vorstellt, haben wir bereits bei der Diskussion von Heck (2000) in Kapitel 6 gesehen; relevant waren hier die Beschränkungen DEF, ANIM, AGENS und ADJA. Die auf der Basis von optimalen D-Strukturen erfolgende S-Struktur-Optimierung arbeitet demgegenüber mit Beschränkungen, die per Annahme zwar Teil derselben Hierarchie, aber höher geordnet sind. Hierzu gehören neben dem Bewegungsverbot ÖKON Beschränkungen, die Bewegung auslösen, u.a. die Beschränkung TYP ("Type"), die Topikalisierung von (entsprechend markierten) XPs erzwingt, und die Beschränkung SP ("Skopus-Prinzip", "Scope Principle"), die fordert, daß ein Quantor an einen abstrakten Skopusmarker adjungiert sein muß, wenn der mit ihm koindiziert ist. Dies setzt voraus, daß abstrakte indizierte Q-Skopusmarker optional (allerdings unter Respektierung bestimmter Restriktionen, die semantisch leere Skopusmarkierung verbieten) eingesetzt werden können. Ist ein Skopusmarker mit einem Quantor koindiziert, wird durch die Ordnung $SP \gg \text{ÖKON}$ S-strukturelle Bewegung des Quantors erzwungen. Hinter dieser Analyse steht die Annahme, daß Scrambling oft semantisch getrieben zu sein scheint, und daß es im Deutschen manchmal so aussieht, als würden Quantoren zur Desambiguierung des Skopus bereits auf der S-Struktur bewegt. Ein Beispiel: Die Beschränkungen der D-Struktur (insbesondere ANIM und ADJA) legen fest, daß im Deutschen (26-a) eine optimale D-Struktur ist, (26-b) dagegen nicht. Per Annahme kann in der Numeration optional ein abstrakter Q-Skopusmarker auftauchen. Weil Präsenz vs. Fehlen des Q-Skopusmarkers zu zwei unterschiedlichen D-Struktur-Wettbewerben führt, und weil der Q-Skopusmarker nicht mit den für D-Struktur-Optimierung relevanten Beschränkungen interagiert, ist (26-c) als D-Struktur wie (26-a) optimal und (26-d) als D-Struktur wie (26-b) suboptimal.

- (26) a. daß Jakob jedem Kind ein Märchen erzählt hat
 b. *daß Jakob ein Märchen jedem Kind erzählt hat
 c. daß Jakob Q^1 jedem Kind ein Märchen¹ erzählt hat
 d. *daß Jakob Q^1 ein Märchen Q^1 jedem Kind erzählt hat

(26-a) als Input für S-Struktur-Optimierung führt bei der Ordnung $SP \gg \text{ÖKON}$ wegen Irrelevanz der höheren Beschränkung zur optimalen S-Struktur (27-a), in der das direkte Objekt in situ verharrt; (26-c) dagegen führt als Input für S-Struktur-Optimierung wegen Relevanz von SP zur optimalen S-Struktur (27-b), in der Scrambling erfolgt ist. Auf diese Weise ergibt sich in diesem System eine

Pseudo-Optionalitätsanalyse von Scramblingphänomenen im Deutschen.¹²

- (27) a. daß Jakob jedem Kind ein Märchen erzählt hat
 b. daß Jakob [ein Märchen¹ Q¹] jedem Kind t₁ erzählt hat

Die optimalen S-Strukturen in (27) sind dann Inputs für LF-Optimierungsprozeduren. Wichtig sind unter anderem die Beschränkung QUIB (“Quantifier Induced Barrier”), derzufolge Bewegung über einen Quantor verboten ist (eine Variante von relativierter Minimalität, vgl. Beck (1996)); wiederum die Beschränkung SP, die auf LF die Konsequenz hat, daß ein an einen Q-Skopusmarker adjungierter Quantor nicht mehr wegbewegt werden kann; die Beschränkung REK (“Rekonstruktion”, “Rec”), die Rekonstruktion jedweder Bewegung fordert; sowie eine QR (“Quantorenanhebung”, “Quantifier Raising”) genannte Beschränkung, die jedoch hier nur verlangt, daß bestimmte “starke” Quantoren (wie *jeder*), die auch in Pafels (1993; 1998) System als zu weitem Skopus neigend klassifiziert sind (vgl. Kapitel 4), oberhalb ihrer Basisposition adjungiert werden. QUIB und SP sind beide Beschränkungen, die LF-Bewegungen tendentiell verbieten; da sie per Annahme im Deutschen oberhalb von REK und QR geordnet sind, sorgen sie dafür, daß LF-Struktur (und damit Interpretation) und S-Struktur in dieser Sprache oft nicht allzu sehr voneinander abweichen. Syntaktische Topikalisierung, die durch TYP und nicht durch SP ausgelöst wurde, muß aber normalerweise unter Beachtung von QUIB und SP rekonstruiert werden, wodurch Skopusinversionseffekte zustande kommen können. Ein interessanter Fall der LF-Skopusinversion gegenüber einer S-Struktur ist (28-a). Dieser Satz kann die LF-Repräsentation in (28-b) haben.

- (28) a. [CP Ein Schüler₁ hat [IP t₁ jedes Kind₂ verprügelt]] (S-Struktur)
 b. [CP – [IP jedes Kind₂ [IP ein Schüler₁ t₂ verprügelt hat]]] (mögliche LF)

In (28-a) ist NP₁ aufgrund von TYP topikalisiert worden; ein Q-Skopusmarker ist hier nicht präsent (und SP daher im vorliegenden Fall irrelevant). Auf dem Weg zur LF in (28-b) ist zuerst NP₂ aufgrund von QR in eine IP-Adjunktionsposition geschoben worden (dies ist möglich, weil die Bewegung über t₁ erfolgt und somit nicht QUIB verletzt), und dann NP₁ in seine Basisposition unterhalb der nunmehr IP-adjungierten NP₂ zurückgeschoben worden, wie dies REK fordert. Somit ergibt sich das auf den ersten Blick überraschende, aber empirisch wohl korrekte Resultat, daß im Deutschen in genau diesem Kontext, wo die Skopusverhältnisse durch Topikalisierung “destabilisiert” worden sind und ein starker Quantor vorliegt, ein Quantor Q₂ über einen anderen Quantor Q₁ Skopus nehmen kann, obwohl auf der

¹²In (27-b) geht ein (existentiell interpretiertes) indefinites Objekt einem definiten voran. Diese Abfolge ist per se markiert; vgl. Kapitel 6. Bei Heck (1999; 2000) wird das mit Hilfe der Annahme erfaßt, daß Scrambling immer zu Markiertheit führen muß (vgl. die DTK-basierte Theorie der Markiertheit). Im gegenwärtigen Kontext geht es aber nicht so sehr um die Markiertheit von (27-b), sondern vielmehr um die Skopusverhältnisse in den beiden Sätzen.

S-Struktur Q_2 weder Q_1 noch eine Spur von Q_1 c-kommandiert.

Zu den Einzelheiten der Skopustheorie von Heck (1999) wäre noch weit mehr zu sagen. Ich möchte hier nur noch kurz zwei Argumente erwähnen, die Heck für einen seriellen Ansatz und gegen die zu Beginn dieses Abschnitts erwähnte Option, parallel in einem Durchgang $\langle D\text{-Struktur}, S\text{-Struktur}, LF \rangle$ -Tupel zu optimieren, vorbringt. Das erste Argument ist konzeptuell. Beim seriellen System muß bei der Optimierung einer S-Struktur nur eine einzige D-Struktur betrachtet werden; beim parallelen System müssen dagegen bei der Optimierung einer S-Struktur sämtliche a priori möglichen D-Struktur-Kandidaten betrachtet werden, also auch die vielen D-Struktur-Outputs, die als suboptimal erwiesen sind. Analoges gilt für LF-Optimierung. Ein einfaches Beispiel mag dies verdeutlichen. Angenommen, es gibt nur drei D-Struktur-Kandidaten $O_{d/1}$, $O_{d/2}$ und $O_{d/3}$; $O_{d/1}$ sei optimal. Dann müssen bei serieller Optimierung nur noch solche S-Struktur-Outputs betrachtet werden, die auf $O_{d/1}$ zurückgehen, nicht solche, die auf den suboptimalen D-Struktur-Outputs $O_{d/2}$ oder $O_{d/3}$ basieren. Die auf $O_{d/1}$ zurückgehenden S-Strukturen seien wiederum nur drei Kandidaten: $O_{s/1}$, $O_{s/2}$ und $O_{s/3}$; $O_{s/1}$ sei wieder optimal. Dann sind schließlich bei LF-Optimierung nur Abkömmlinge von $O_{s/1}$ zu berücksichtigen, nicht Abkömmlinge von $O_{s/2}$ oder $O_{s/3}$. Wäre das System nun parallel anstatt seriell, wie in der Standardoptimalitätstheorie, so müßten bei $\langle D\text{-Struktur}, S\text{-Struktur}, LF \rangle$ -Tupeln S-Struktur-Nachfolger von $O_{d/2}$ und $O_{d/3}$ und LF-Nachfolger von $O_{s/2}$ und $O_{s/3}$ ebenfalls betrachtet werden, obschon klar ist, daß solche Tupel nicht optimal werden können. Harmonischer Serialismus verkleinert hier also die Kandidatenmenge und reduziert auf diese Weise die Komplexität des Systems.

Das zweite Argument ist empirisch. Die Existenz mehrerer hintereinandergeschalteter Optimierungen eröffnet die Möglichkeit, daß von einer Harmonieevaluation zur nächsten die universellen und stets gültigen Beschränkungen ungeordnet werden können; dies ist bei Einfachoptimierung von Tupeln unmöglich. Im Einklang damit versucht Heck (1999) Evidenz für Beschränkungsumordnung beizubringen und auf diese Weise ein Argument für harmonischen Serialismus in der Syntax zu entwickeln: Es tritt der Fall ein, daß zwei Beschränkungen B_1 , B_2 auf S-Struktur wie auf LF gebraucht werden, und daß auf der S-Struktur $B_1 \gg B_2$ gelten muß, auf LF dagegen $B_2 \gg B_1$.

Zum Abschluß dieses Kapitels seien noch kurz weitere Ansätze zur Mehrfachoptimierung angeführt – nicht in der Ausführlichkeit, die nötig wäre, um sie zu richtig zu erfassen, aber doch hoffentlich so, daß man einen ersten Eindruck gewinnen kann.

3.3. Weitere Ansätze

Interpretative Optimierung bei Hendriks & de Hoop (1999) Wir haben gesehen, daß die Ansätze von Wilson (1998) und Heck (1999) zwar beide Mehrfachoptimierung involvieren, sich aber u.a. hinsichtlich des postulierten Syntax-/Semantik-Verhältnisses unterscheiden. Bei Wilson geht die interpretative Optimierung der

Form-Optimierung voran; bei Heck geht die Form-Optimierung der interpretativen Optimierung voran. Unter den letzteren Theorientyp subsumierbar ist auch der Ansatz von Hendriks & de Hoop (1999): Zunächst erfolgt (auf der Basis von Prädikat-/Argument-Strukturen) syntaktische Optimierung, und der optimale Output dieser ersten Optimierung ist der Input für die nachfolgende semantische Optimierung, die die Interpretation determiniert. Hendriks und de Hoop legen im Unterschied zu Heck allerdings nicht das klassische GB-Modell zugrunde, und auch ansonsten unterscheiden sich die Arbeiten in ihrer Ausrichtung wesentlich. Ein interessanter Aspekt ist dabei, daß Hendriks und de Hoop bei der interpretativen Optimierung Beschränkungen verwenden, die partiell pragmatisch sind, weil die Einbeziehung kontextueller Information notwendig ist, um zu entscheiden, ob sie erfüllt oder verletzt sind. Insofern als man davon ausgehen kann, daß Beschränkungen, die nur für die Form-Optimierung relevant sind, diese Eigenschaft nicht besitzen, scheint die Relevanz pragmatischer Faktoren bei der interpretativen Optimierung mit dem Ansatz von Wilson schlechter vereinbar.

Echt bidirektionale Optimierung bei Blutner (2000) und Jäger & Blutner (2000)
 Alle bisher behandelten Ansätze zur Mehrfachoptimierung involvieren einen gerichteten Serialismus: Von zwei Optimierungen muß die eine immer der anderen vorangehen. Diese Annahme wird aufgegeben in der Theorie der Mehrfachoptimierung, die Blutner (2000) und Jäger & Blutner (2000) entwickeln. Blutner und Jäger gehen davon aus, daß es zwei verschiedene Optimierungen gibt, die sie das Q-Prinzip und das I-Prinzip nennen. Ein Kandidat ist bei der Q-Optimierung der Gewinner, wenn er maximal explizit ist; ein Kandidat gewinnt bei der I-Optimierung, wenn er maximal ökonomisch ist. Die beiden Optimierungsverfahren sind nun aber nicht geordnet, sondern sie beeinflussen sich gegenseitig: Ein Kandidat erfüllt das Q-Prinzip, wenn keinen Konkurrenten mit derselben Interpretation, aber harmonischerer Form gibt, und ein Kandidat erfüllt das I-Prinzip, wenn es keinen Konkurrenten mit derselben Form, aber harmonischerer Interpretation gibt. Ein Kandidat zählt schließlich als "super-optimal", wenn er bei Q- und I-Optimierung gleichermaßen als Gewinner erscheint. Dieser Ansatz ist also anders als der von Wilson (1998) (vgl. Fußnote 11) echt bidirektional.

Parsing und serielle Optimierung bei Fanselow, Schleswesky & Kliegl (1999)
 Fanselow, Schleswesky & Kliegl (1999) beschäftigen sich mit den häufig konstatierten Unterschieden zwischen Grammatik und Parser. Sie schlagen vor, daß bei Annahme einer optimalitätstheoretischen Organisation der Syntax Asymmetrien zwischen Sprachkompetenz und Sprachperzeption nicht wie üblicherweise vorausgesetzt zu der Konklusion zwingen, daß die den beiden Systemen zugrundeliegenden Beschränkungen unterschiedlich sind. Vielmehr kann es sich um dieselben verletzbaren Beschränkungen handeln (wie z.B. eine Minimalitätsbedingung in der Art von Chomskys (1995) MLC ("Minimal Link Condition"), derzufolge Bewegung von α nach δ unmöglich ist, wenn ein α hinreichend ähnliches β interveniert, ÖKON, OB-KOPF

und VOLL-INT), und sie müssen in Grammatik und Parser nicht einmal unterschiedlich geordnet sein – Grammatik und Parser sind identisch. Die Asymmetrien, so die Idee, gehen letztlich darauf zurück, daß die aktuelle Sprachperzeption wiederholte Optimierungen involviert, die erst dann ihr Ende finden, wenn das letzte Wort eines Satzes in den Parser eingelesen ist. Die zentrale Einsicht ist dabei, daß sich bei der Perzeption eines gegebenen Satzes mit jeder eingelesenen syntaktischen Einheit die Kandidatenmenge ändert. Da mehr und mehr lexikalisches Material hinzu kommt, wird die Kandidatenmenge immer spezifischer und kleiner, und ein Kandidat K_1 , der am Beginn des Parsingprozesses noch von einem anderen Kandidaten K_2 als suboptimal blockiert wird, kann an einem späteren Punkt optimal werden, weil K_2 aus der Kandidatenmenge entfernt worden ist. Auf diese Weise entstehen an unterschiedlichen Punkten der Sprachperzeption unterschiedliche optimale Kandidaten. Erst im letzten Schritt erscheint der vollständige Satz. Vom Parser wird das Endergebnis notwendigerweise als optimal klassifiziert; die Grammatik konstruiert dagegen eine Kandidatenmenge gemäß einem der in Kapitel 1 vorgestellten Prinzipien (z.B., wie Fanselow, Schlesewsky & Kliegl (1999) annehmen, über identische Numerationen und identische LF-Repräsentation) und entscheidet dann über die Grammatikalität des Satzes.¹³

Ein Beispiel: Angenommen, der Parser hat bisher lediglich eine Satz-initiale W-Phrase *welche Frau* erhalten; diese W-Phrase ist mit Akkusativ- und Nominativmarkierung verträglich und somit lokal ambig. Der Parser konstruiert nun mögliche Fortsetzungen, die Kandidatenmenge. Darunter sind K_2 , wo eine Subjektspur postuliert wird, und K_1 , wo die Spur der W-Phrase in der Objektposition lokalisiert ist. Aufgrund der Minimalitätsbedingung ist K_2 optimal. Später erhält der Parser jedoch ein eindeutiges Subjekt wie *der Mann* oder ein Plural-kongruierendes Verb. Damit wird K_2 aus der Kandidatenmenge entfernt, und K_1 kann optimal werden. Die Revision eines optimalen Kandidaten beim Parsingvorgang läßt sich als Ursache für die beobachtbare Verarbeitungsverzögerung identifizieren, die nicht auftritt, wenn der anfänglich als optimal klassifizierte Kandidat K_2 optimal bleibt (weil der Parser z.B. statt eines Subjekts ein Objekt einliest).

Lokale Optimierung bei Heck & Müller (2000ab) und Müller (1999d) In den bisher in diesem Buch behandelten Optimierungsverfahren (seien sie parallel oder seriell) ist vorausgesetzt worden, daß die Kandidaten vollständige syntaktische Strukturen sind und die Optimierung somit global erfolgt. Eine Alternative dazu ist die lokale Optimierung von Satzteilen. Da am Ende natürlich über die Grammatikalität des Gesamtsatzes entschieden werden muß, präsupponiert dieses Verfahren, daß Optimierung mehrfach erfolgt, also seriell ist; und da sich die Satzstruktur von unten nach oben bis hin zur Wurzel verengt, wird weiterhin vorausgesetzt, daß

¹³Siehe zu diesem verbleibenden Unterschied zwischen Perzeption und Grammatik (bzw. Produktion) insbesondere Smolensky (1996).

die Optimierung zyklisch von unten nach oben durchgeführt wird (würde man mit dem Wurzelknoten beginnen, hätte man automatisch bereits den Gesamtsatz optimiert). Zyklische Optimierung in der Syntax wird als Möglichkeit angedeutet bei Archangeli & Langendoen (1997, 214) und in einer Fußnote bei Ackema & Neeleman (1998, 478). Ein allgemeines Modell der zyklischen lokalen Optimierung wird in Heck & Müller (2000ab) und Müller (1999d) entwickelt.

Eine Grundannahme ist, daß die Syntax derivationell ist, und daß sich Bewegung (“Move”) und Phrasenaufbau (“Merge”) wie bei Chomsky (1995) beschrieben abwechseln. Jede vollendete XP ist ein zyklischer Knoten, und syntaktische Optimierung betrifft auch nur genau diesen extrem lokalen Bereich: die Teilderivation von einer XP α bis zur unmittelbar dominierenden XP β . Optimal ist eine XP, wenn die Teilderivation, die sie erzeugt, das beste Beschränkungsprofil hat. Ein optimaler XP-Output β wird dann (zusammen mit einem aus dem Lexikon genommenen Kopf γ und potentiell weiteren optimalen XPs, je nachdem, ob γ keinen, einen oder mehrere Spezifikatoren hat) zum Input für die nächste Optimierung, usw., bis die optimale Wurzel-XP erreicht und ein wohlgeformter Gesamtsatz entstanden ist. Der Ansatz beruht auf dem (unverletzbaren) Prinzip des strikten Zyklus: Ist eine XP einmal optimiert, können spätere Operationen nicht mehr exklusiv auf diesen zyklischen Bereich zugreifen.

Konzeptuell ergeben sich zwischen lokaler (und somit inhärent serieller) und globaler (serieller oder, meistens, paralleler) Optimierung große Unterschiede. Auf der einen Seite ist bei der lokalen Optimierung die Komplexität des Verfahrens gegenüber der globalen (auch der seriell globalen) Optimierung reduziert, weil bei der Optimierung einer XP β ausschließlich solche Kandidaten im Wettbewerb sind, die sich im Bereich zwischen β und dem Kopf von β unterscheiden. Wenn z.B. der Kopf von β mit der XP α verknüpft wird, dann werden alte suboptimale Konkurrenten von α nicht mehr betrachtet. Dies oder etwas Analoges (je nachdem, ob die Grammatik derivationell oder repräsentationell ausgerichtet ist) muß man aber bei der globalen Optimierung (auch bei der seriellen Variante) tun. Auf der anderen Seite muß beim lokalen Verfahren natürlich wieder und wieder optimiert werden; dem steht bei den existierenden globalen Verfahren entweder nur eine einzige Optimierung gegenüber (parallel), oder es handelt sich um eine kleinere Zahl von Optimierungen (seriell – die vorgestellten Ansätze schwanken hier zwischen 2 und 3).

Auf der empirischen Seite liegt der Schwerpunkt in Heck & Müller (2000ab) und Müller (1999d) auf dem Phänomen der Reparaturbewegung, das in Abschnitt 2 von Kapitel 2 schon kurz angesprochen worden ist. Hiermit ist die folgende Situation gemeint: Eine gegebene Bewegungsoperation OP appliziert normalerweise in einer Sprache nicht auf eine Kategorie K. Unter bestimmten Umständen, in wohldefinierten syntaktischen Kontexten, kann OP aber doch auf K applizieren; tatsächlich *muß* K in diesen Fällen sogar von OP bewegt werden. Also z.B.: W-Phrasen können im Deutschen normalerweise nicht gescrambelt werden. W-Phrasen können und müssen aber gescrambelt werden, um Schmarotzerlücken zu lizensieren, um schwache

Überkreuzungseffekte zu vermeiden, und um Negationsinterventionseffekte zu umgehen. Die Analyse ist hier im Prinzip analog zu anderen Reparaturanalysen in der optimalitätstheoretischen Syntax: Es gibt eine Beschränkung, die das Scrambling von W-Phrasen verbietet, und es gibt höher geordnete Beschränkungen, die in bestimmten Kontexten verletzt würden, wenn eine W-Phrase in situ verharret. Daher wird beim optimalen Kandidaten das Verbot von W-Scrambling verletzt. Dies ist an und für sich auch gut durch globale Optimierung zu erfassen. Was jedoch den vorgestellten Fall (sowie andere Fälle Reparatur-getriebener Bewegung) auszeichnet, ist, daß die Reparatur unmittelbar und lokal erfolgt, nämlich an der Stelle, wo das "Problem" auftritt, und nicht früher oder später in der Derivation (bzw., repräsentationell gesehen, tiefer oder höher in der Struktur). Aus diesem Faktum lassen sich dann empirische Argumente gegen globale Optimierung ableiten: Unter globaler Optimierung sollte Reparatur-getriebene Bewegung weit öfter (an anderen Stellen im Baum) möglich sein, als dies tatsächlich der Fall ist.

4. Möglichkeiten und Grenzen

Die vorangegangenen Kapitel haben zwar viele, aber längst noch nicht alle wichtigen Aspekte der optimalitätstheoretischen Syntax behandelt. Ausgeblendet geblieben ist zum Beispiel der Bereich der diachronen optimalitätstheoretischen Syntax, wo es bislang noch recht wenig Literatur gibt (aber vgl. Kiparsky (1999) und Müller (2000a)), obwohl es so aussieht, als würde gerade hier die Optimalitätstheorie mit dem Konzept des Sprachwandels durch Beschränkungsumordnung interessante neue Einsichten liefern können. Ebenso vernachlässigt worden ist in diesem Buch das Thema des Erwerbs von optimalitätstheoretischen Grammatiken (vgl. hierzu Kager (1999) und insbesondere Tesar & Smolensky (2000)). Nichtsdestoweniger ist es Zeit für ein Fazit.

Zentrale theoretische Konzepte der optimalitätstheoretischen Syntax sind zum gegenwärtigen Zeitpunkt ungeklärt; dies betrifft unter anderem den Input, die Kandidatenmenge, die Natur der Kandidaten und die Natur möglicher Beschränkungen. Phänomene wie absolute Ungrammatikalität, Grammatikalitätsgrade und Optionalität können zwar optimalitätstheoretisch erfaßt werden, erfordern aber spezielle Mechanismen, die man als Abweichung vom Idealfall betrachten kann. Ob sich die Hypothese der faktoriellen Typologie aufrechterhalten läßt, ist unklar. Erwünscht scheinen universell feste Ordnungen unter einem Teil der Beschränkungen; aber die Mittel, solche festen Ordnungen abzuleiten (Subhierarchien, lokale Konjunktion, harmonische Ausrichtung), involvieren nicht-triviale Erweiterungen des Systems. Schließlich läßt sich allgemein die der optimalitätstheoretischen Syntax eigene Komplexität als Nachteil gegenüber Standardsyntaxen auffassen.

Dem steht gegenüber, daß in bestimmten Bereichen die Evidenz stark für eine optimalitätstheoretische Organisation der Syntax spricht: Wie wir gesehen haben, sind die Konzepte der Reparatur, des Wettbewerbs und der Blockade von Sätzen, des unmarkierten Falles und des Beschränkungskonflikts empirisch motiviert; die-

se Konzepte sind optimalitätstheoretisch systematisch erfaßbar, bereiten aber in Standardsyntaxen große Probleme. Darüber hinaus leitet die Theorie der Parametrisierung als Beschränkungsumordnung zwischensprachliche Variation auf einfache Weise ab.

Aus alledem mag jeder Leser seine eigene Konklusion hinsichtlich der Möglichkeiten und Grenzen einer optimalitätstheoretischen Syntax ziehen. Mein persönlicher Schluß ist, daß die Argumente pro die Argumente contra mehr als aufwiegen. Ich denke, die optimalitätstheoretische Syntax ist zwar aufgrund der ihr inhärenten Probleme unter konstantem Rechtfertigungsdruck und auch in Zukunft auf eine Stützung durch weitere empirische Evidenz angewiesen; aber sie hat bereits wesentlich zum Verständnis von manchen syntaktischen Phänomenen beigetragen und bleibt eines der faszinierendsten Forschungsfelder der Grammatiktheorie in den nächsten Jahren.

Literaturangaben

- Abraham, Werner. 1986. Word Order in the Middle Field of the German Sentence. In *Topic, Focus, and Configurationality*, eds. Werner Abraham & Sjaak de Meij, 15-38. Amsterdam: Benjamins.
- Abraham, Werner. 1995. *Deutsche Syntax im Sprachenvergleich*. Tübingen: Stauffenburg.
- Ackema, Peter, & Ad Neeleman. 1998. Optimal Questions. *Natural Language and Linguistic Theory* 16:443-490.
- Aissen, Judith (1999). Markedness and Subject Choice in Optimality Theory. *Natural Language and Linguistic Theory* 17:673-711.
- Altmann, Hans. 1981. *Formen der Herausstellung im Deutschen. Rechtsversetzung, Linksversetzung, Freies Thema*. Tübingen: Niemeyer.
- Archangeli, Diana, & D. Terence Langendoen. 1997. Afterword. In *Optimality Theory. An Overview*, eds. Diana Archangeli & D. Terence Langendoen, 200-215. Oxford: Blackwell.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.
- Artstein, Ron. 1998. Hierarchies. Ms., Rutgers University. In *Proceedings of Console VII*, ed. Tina Cambier-Langeveld, Anikó Lipták, Michael Redford & Erik Jan van der Torre, 1-15. Leiden: SOLE.
- Baker, Mark. 1988. *Incorporation*. Chicago: Chicago University Press.
- Baković, Eric. 1995. A Markedness Subhierarchy in Syntax. Ms., Rutgers University.
- Baković, Eric. 1997. Faithfulness & Optionality. Ms., Rutgers University.
- Baković, Eric. 1998. Optimality and Inversion in Spanish. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 35-58. Cambridge, Mass.: MIT Press & MITWPL.
- Baković, Eric, & Ed Keer. 1999. Optionality and Ineffability. Ms., Harvard University & UMass., Amherst. Erscheint in *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw & Sten Vikner. Cambridge, Mass.: MIT Press.
- Bayer, Josef. 1984. COMP in Bavarian Syntax. *The Linguistic Review* 3:209-275.
- Bayer, Josef, Markus Bader & Michael Meng. 1999. Conflicting Cases in German. Ms., Universität Jena.
- Bayer, Josef, Markus Bader & Michael Meng. 2000. Morphological Underspecification Meets Oblique Case: Syntactic and Processing Effects in German. Ms., Universität Jena.
- Beck, Sigrid. 1996. *Wh-Constructions and Transparent Logical Form*. Dissertation, Universität Tübingen.
- Bennis, Hans. 1986. *Gaps and Dummies*. Dordrecht: Foris.
- Bierwisch, Manfred. 1988. On the Grammar of Local Prepositions. In *Syntax, Semantik und Lexikon*, eds. Manfred Bierwisch, Wolfgang Motsch & Ilse Zimmermann, 1-65. Berlin: Akademie-Verlag.
- Billings, Loren, & Catherine Rudin. 1996. Optimality and Superiority: A New Approach to Multiple Wh-Ordering. In *Formal Approaches to Slavic Linguistics 4*, ed. Jindřich Toman, 35-60. Ann Arbor: Michigan Slavic Publications.
- Blutner, Reinhard. 2000. Some Aspects of Optimality in Natural Language Interpretation. Ms., Humboldt-Universität Berlin.

- Borer, Hagit. 1984. *Parametric Syntax*. Dordrecht: Foris.
- Borsley, Robert. 1997. *Syntax-Theorie. Ein zusammengefaßter Zugang*. Deutsche Bearbeitung von Peter Suchsland. Tübingen: Niemeyer.
- Bošković, Željko. 1999. What is Special about Multiple Wh-Fronting? Ms., University of Connecticut. Erscheint in *Proceedings of NELS 30*.
- Bresnan, Joan. 1982. (ed.) *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: MIT Press.
- Bresnan, Joan. 1997. Optimal Syntax. Ms., Stanford University. Erscheint in *Conceptual Studies in Optimality Theory*, eds. Joost Dekkers et al., Oxford: Oxford University Press.
- Bresnan, Joan. 1998. Morphology Competes With Syntax: Explaining Typological Variation in Weak Crossover Effects. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 59-92. Cambridge, Mass.: MIT Press & MITWPL.
- Brody, Michael. 1995. *Lexico-Logical Form. A Radically Minimalist Theory*. Cambridge, Mass.: MIT Press.
- Broekhuis, Hans, & Joost Dekkers. 1997. The Minimalist Program and Optimality Theory: Derivations and Filters. Ms., HIL, Amsterdam. Erscheint in *Conceptual Studies in Optimality Theory*, eds. Joost Dekkers et al., Oxford: Oxford University Press.
- Broihier, Kevin. 1995. Optimality Theoretic Rankings with Tied Constraints: Slavic Relatives, Resumptive Pronouns and Learnability. Ms., MIT, Cambridge, Mass.
- Browning, M.A. 1989. ECP \neq CED. *Linguistic Inquiry* 20:481-491.
- Browning, M.A. 1991. Bounding Conditions on Representations. *Linguistic Inquiry* 22:541-562.
- Büring, Daniel. 1996. Interpretation and Movement: Towards an Economy-Theoretic Treatment of German 'Mittelfeld' Word Order. Ms., Universität Frankfurt.
- Büring, Daniel. 1997. Towards an OT Account of German Mittelfeld Word Order. Ms., Universität Köln.
- Büring, Daniel. 1999. Let's Phrase It! Ms., University of California, Santa Cruz. Erscheint in *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld. Berlin: Mouton de Gruyter.
- Burzio, Luigi. 1991. The Morphological Basis of Anaphora. *Journal of Linguistics* 27:81-105.
- Cardinaletti, Anna. 1990. *Es, pro, and Sentential Arguments in German*. *Linguistische Berichte* 126:135-164.
- Choi, Hye-Won. 1996. *Optimizing Structure in Context: Scrambling and Information Structure*. Dissertation, Stanford University.
- Choi, Hye-Won. 1999. *Optimizing Structure in Context. Scrambling and Information Structure*. Stanford: CSLI Publications.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1973. Conditions on Transformations. In *A Festschrift for Morris Halle*, eds. Stephen Anderson & Paul Kiparsky, 232-286. New York: Academic Press.
- Chomsky, Noam. 1977. On Wh-Movement. In *Formal Syntax*, eds. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 232-286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: MIT Press.

- Chomsky, Noam. 1986a. *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1986b. *Knowledge of Language*. New York: Praeger.
- Chomsky, Noam. 1991. Some Notes on Economy of Derivation and Representation. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 417-454. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20*, eds. Kenneth Hale & Samuel Jay Keyser, 1-52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1995. Categories and Transformations. Chapter 4 of *The Minimalist Program*, 219-394. Cambridge, Mass.: MIT Press.
- Chomsky, Noam, & Howard Lasnik. 1977. Filters and Control. *Linguistic Inquiry* 8:425-504.
- Chomsky, Noam, & Howard Lasnik. 1993. Principles and Parameters Theory. In *Syntax, vol. I*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 506-569. Berlin: de Gruyter.
- Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, Mass.: MIT Press.
- Collins, Chris. 1994. Economy of Derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25:45-61.
- Collins, Chris. 1997. *Local Economy*. Cambridge, Mass.: MIT Press.
- Costa, João. 1998. *Word Order Variation*. Dissertation. Universiteit Leiden. (LOT dissertations, 14)
- Costa, João. 1999. The Emergence of Unmarked Word Order. Ms., Universidade de Lisboa & University of Leiden. Erscheint in *Optimality Theoretic Syntax*, eds. Jane Grimshaw, Géraldine Legendre & Sten Vikner. Cambridge, Mass.: MIT Press.
- d'Avis, Franz, & Jürgen Pafel. 1992. Ein Skopusmodell und -Algorithmus. Unter besonderer Berücksichtigung des Skopus von w-Phrasen. In *W-Phrasen, W-Merkmale, Skopusberechnung*, eds. Marga Reis et al., 35-67. Stuttgart & Tübingen: Arbeitspapiere des SFB 340, 7.
- Dekkers, Joost. 1999. *Derivations and Evaluations*. Dissertation. Universiteit Amsterdam. (LOT dissertations, 21)
- Déprez, Viviane. 1991. Economy and the *that-t* Effect. In *Proceedings of the Western Conference on Linguistics* 4:74-87.
- Dickey, Michael. 1995. Inversion in Child English and Acquisition in Optimality Theory. In *Papers in Optimality Theory*, eds. Jill Beckman, Laura Walsh-Dickie & Suzanne Urbanczyk, 575-587. Amherst, Massachusetts: UMass Occasional Papers in Linguistics 18.
- Dietrich, Rainer. 1994. Wettbewerb – aber wie? Skizze einer Theorie der freien Wortstellung. In *Was determiniert die Wortstellungsvariation?*, ed. Brigitta Haftka, 33-47. Opladen: Westdeutscher Verlag.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.
- Diesing, Molly. 1997. Yiddish VP Order and the Typology of Object Movement in Germanic. *Natural Language and Linguistic Theory* 15:369-427.
- DiSciullo, Anna-Maria, & Edwin Williams. 1987. *On the Definition of Word*. Cambridge, Mass.: MIT Press.
- Dietrich, Rainer. 1994. Wettbewerb – aber wie? Skizze einer Theorie der freien Wortstellung. In *Was determiniert die Wortstellungsvariation?*, ed. Brigitta Haftka, 33-47. Opladen: Westdeutscher Verlag.

- Déprez, Viviane. 1991. Economy and the 'that'-t Effect. *Proceedings of the Western Conference on Linguistics* 4:74-87. California State University, Fresno.
- Dowty, David, Robert Wall & Stanley Peters. 1981. *Introduction to Montague Semantics*. Dordrecht:Reidel.
- Dresher, Elan. 1996. The Rise of Optimality Theory in First Century Palestine. *GLOT International* 2-1/2:8.
- Eisenberg, Peter. 1986. *Grundzüge der deutschen Grammatik*. Stuttgart: Metzler.
- Emonds, Joseph. 1978. The Verbal Complex V'-V in French. *Linguistic Inquiry* 9:151-175.
- Engdahl, Elisabet. 1996. A Comparison of Movement Theory in GB and HPSG. Handout eines Vortrags bei der SFB-Konferenz "Interfaces of Grammar", Tübingen.
- Epstein, Samuel David. 1992. Derivational Constraints on \bar{A} -Chain Formation. *Linguistic Inquiry* 23:235-259.
- Epstein, Samuel David, Ruriko Kawashima, Hisatsugu Kitahara & Erich Groat. 1998. *A Derivational Approach to Syntactic Relations*. New York & Oxford: Oxford University Press.
- Fanselow, Gisbert. 1983. *Zu einigen Problemen von Kasus, Rektion und Bindung in der deutschen Syntax*. Magisterarbeit, Universität Konstanz.
- Fanselow, Gisbert. 1988. German Word Order and Universal Grammar. In *Natural Language Parsing and Linguistic Theories*, eds. Uwe Reyle & Christian Rohrer, 317-355. Dordrecht: Kluwer.
- Fanselow, Gisbert. 1990. Scrambling as NP-movement. In *Scrambling and Barriers*, ed. Günther Grewendorf & Wolfgang Sternefeld, 113-140. Amsterdam: Benjamins.
- Fanselow, Gisbert. 1991. *Minimale Syntax*. Habilitation thesis, Universität Passau.
- Fanselow, Gisbert. 1993. The Return of the Base Generators. Ms., Universität Passau. (Also in *Groninger Arbeiten zur Germanistischen Linguistik* 36:1-74.)
- Fanselow, Gisbert. 1995. A Minimalist Approach to Free Constituent Order. Ms., Universität Potsdam.
- Fanselow, Gisbert. 1997. The Proper Interpretation of the Minimal Link Condition. Ms., Universität Potsdam.
- Fanselow, Gisbert. 1999. Optimal Exceptions. Ms., Universität Potsdam.
- Fanselow, Gisbert, Matthias Schlesewsky & Reinhold Kliegl. 1999. Optimal Parsing. Ms., Universität Potsdam.
- Fanselow, Gisbert, & Sascha Felix. 1987. *Sprachtheorie, Bd.2: Die Rektions- und Bindungstheorie*. Tübingen: Francke (UTB).
- Fanselow, Gisbert, & Damir Ćavar. 2000. Remarks on the Economy of Pronunciation. Ms., Universität Potsdam. Erscheint in *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld. Berlin: Mouton de Gruyter.
- Felix, Sascha. 1983. Parasitic Gaps in German. *Groninger Arbeiten zur Germanistischen Linguistik* 22:1-46.
- Fischer, Silke. 1999. On the Integration of Cumulative Effects into Optimality Theory. Ms., Universität Tübingen. Erscheint in *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld. Berlin: Mouton/de Gruyter.
- Fortmann, Christian, & Werner Frey. 1997. Konzeptuelle Struktur und Grundabfolge der Argumente im Deutschen. In *Zur Satzstruktur im Deutschen*, eds. Franz-Josef d'Avis & Uli Lutz, 143-170. Tübingen: Arbeitspapiere des SFB 340.
- Fox, Danny. 1995. Economy and Scope. *Natural Language Semantics* 3:283-341.
- Frampton, John. 1997. Expletive Insertion. In *Economy in Linguistic Theory*, eds. Chris

- Wilder, Hans-Martin Gärtner & Manfred Bierwisch, 36-57. Berlin: Akademie Verlag.
- Frampton, John, & Sam Gutman. 1999. Cyclic Computation. *Syntax* 2:1-27.
- Frank, Robert, & Giorgio Satta. 1998. Optimality Theory and the Generative Complexity of Constraint Violability. *Computational Linguistics* 24:277-299.
- Frey, Werner, & Thilo Tappe. 1991. Zur Interpretation der X-bar-Theorie und zur Syntax des Mittelfeldes. Grundlagen eines GB-Fragmentes. Ms., Universität Stuttgart.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum & Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
- Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax*. Tübingen: Stauffenburg.
- Grewendorf, Günther. 1989. *Ergativity in German*. Dordrecht: Foris.
- Grewendorf, Günther. 1997. Multiple Wh-Fronting. Ms., Universität Frankfurt/Main.
- Grewendorf, Günther, & Joachim Sabel. 1994. Long Scrambling and Incorporation. *Linguistic Inquiry* 25:263-308.
- Grimshaw, Jane. 1986. Subjacency and the S/S' Parameter. *Linguistic Inquiry* 17:364-369.
- Grimshaw, Jane. 1991. Extended Projection. Ms., Brandeis University, Waltham, Mass.
- Grimshaw, Jane. 1993. Minimal Projection, Heads, and Optimality. Technical Reports of the Rutgers Center for Cognitive Science 4. Rutgers University, New Brunswick, New Jersey.
- Grimshaw, Jane. 1994. Heads and Optimality. Handout, Universität Stuttgart.
- Grimshaw, Jane. 1997. Projection, Heads, and Optimality. *Linguistic Inquiry* 28:373-422.
- Grimshaw, Jane. 1998. Constraints on Constraints in Optimality Theoretic Syntax. Ms., Rutgers University. New Brunswick, New Jersey.
- Grimshaw, Jane. 1999. Heads and Clauses. Ms., Rutgers University. New Brunswick, New Jersey.
- Grimshaw, Jane, & Vieri Samek-Lodovici. 1995. Optimal Subjects. In *Papers in Optimality Theory*, eds. Jill Beckmann et al., 589-605. University of Massachusetts, Amherst: UMass Occasional Papers in Linguistics 18.
- Grimshaw, Jane, & Vieri Samek-Lodovici. 1998. Optimal Subjects and Subject Universals. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 193-219. Cambridge, Mass.: MIT Press & MITWPL.
- Gutiérrez-Bravo, Rodrigo. 2000. An OT Account of the Interaction between Intonation and Focus-Related Word Order Variation. Ms., University of California, Santa Cruz.
- Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory*. Oxford: Blackwell.
- Haegeman, Liliane. 1995. *The Syntax of Negation*. Cambridge: Cambridge University Press.
- Haftka, Brigitta. 1981. Reihenfolgebeziehungen im Satz. In *Grundzüge einer deutschen Grammatik*, eds. Heidolph, Karl-Erich et al., 702-764. Berlin: Akademie Verlag.
- Haftka, Brigitta. 1999. Topik-Constraints im Deutschen. In *Linguistische Arbeitsberichte* 74. Universität Leipzig.
- Haider, Hubert. 1983. Connectedness Effects in German. *Groninger Arbeiten zur Germanistischen Linguistik* 23:82-119.
- Haider, Hubert. 1992. Branching and Discharge. Ms., Universität Stuttgart.
- Haider, Hubert. 1993. *Deutsche Syntax, generativ*. Tübingen: Narr.
- Haspelmath, Martin. 2000. Optimality and Diachronic Adaptation. *Zeitschrift für Sprachwissenschaft* 18 (1999!):180-205.
- Heck, Fabian. 1997. Komplementierer & ihre Spezifikatoren. Seminararbeit, Universität Tübingen.

- Heck, Fabian. 1998. *Relativer Quantorenskopos im Deutschen – Optimalitätstheorie und die Syntax der Logischen Form*. Magisterarbeit, Universität Tübingen.
- Heck, Fabian. 1999. Quantifier Scopepe in German and Cyclic Optimization. Ms., Universität Stuttgart. Erscheint in *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld. Berlin: Mouton/de Gruyter.
- Heck, Fabian. 2000. Tiefenoptimierung. Ms., Universität Stuttgart. Erscheint in *Linguistische Berichte*.
- Heck, Fabian, & Gereon Müller. 2000a. Repair-driven Movement and the Local Optimization of Derivations. Ms., Universität Stuttgart & Universität Tübingen. Kurzform in: *Glow Newsletter* 44:26-27.
- Heck, Fabian, & Gereon Müller. 2000b. Successive Cyclicity, Long-Distance Superiority, and Local Optimization. Ms., Universität Stuttgart & Universität Tübingen. Erscheint in *Proceedings of WCCFL 19*. Somerville, Mass.: Cascadilla Press.
- Hendriks, Petra, & Helen de Hoop. 1999. Optimality Theoretic Semantics. Ms., University of Groningen. (Cognitive Science and Engineering Prepublications 98-3.)
- Hermans, Ben & Mark van Oostendorp (eds). 2000. *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: Benjamins.
- Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. München: Hueber.
- Hoekstra, Teun. 1984. *Transitivity*. Dordrecht: Foris.
- Höhle, Tilman. 1982. Explikation für "normale Betonung" und "normale Wortstellung". In *Satzglieder im Deutschen*, ed. Werner Abraham, 75-153. Tübingen: Stauffenburg.
- Höhle, Tilman. 1990. Die w ... w-Konstruktion im Deutschen. Handout für einen Vortrag bei der Jahrestagung der DGfS in Saarbrücken, Februar 1990. Revidierte Fassung in Lutz, Müller & von Stechow (2000).
- Huang, Cheng-Teh James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Dissertation. MIT, Cambridge, Mass.
- Huck, Geoffrey, & John A. Goldsmith. 1995. *Ideology and Linguistic Theory*. London: Routledge.
- Jacobs, Joachim. 1988. Probleme der freien Wortstellung im Deutschen. *Sprache und Pragmatik* 5, 8-37.
- Jäger, Gerhard & Reinhard Blutner. 2000. Against Lexical Decomposition in Syntax. Ms., ZAS & Humboldt-Universität Berlin.
- Jaeggli, Osvaldo, & Ken Safir. 1989. The Null Subject Parameter and Parametric Theory. In *The Null Subject Parameter*, eds. Osvaldo Jaeggli & Ken Safir, 1-44. Dordrecht: Kluwer.
- Janiszewski, Horst. 1998. *Straßenverkehrsrecht*. München: Deutscher Taschenbuch Verlag & C.H. Beck.
- Johnson, Kyle. 1991. Object Positions. *Natural Language and Linguistic Theory* 9:577-636.
- Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Keer, Ed, & Eric Baković. 1997. Have FAITH in Syntax. *Proceedings of WCCFL 16*.
- Keller, Frank. 1996. *Extraction from Complex Noun Phrases. A Case Study in Graded Grammaticality*. Diplomarbeit, Universität Stuttgart.
- Keller, Frank. 2000. Evaluating Competition-Based Models of Word Order. Ms., University of Edinburgh.

- Kiparsky, Paul. 1982. From Cyclic Phonology to Lexical Phonology. In *The Structure of Phonological Representations, vol 1*, eds. Harry van der Hulst & Neil Smith, 131-175. Dordrecht: Foris.
- Kiparsky, Paul. 1999. Analogy and OT: Morphological Change as Emergence of the Unmarked. Vortrag auf der 21. Jahrestagung der DGfS, Konstanz.
- Kiss, Tibor. 1994. Bemerkungen zum Vorkommen des 2. Status. *Linguistische Berichte* 154, 461-484.
- Kitahara, Hisatsugu. 1993. Deducing 'Superiority' Effects from the Shortest Chain Requirement. *Harvard Working Papers in Linguistics* 3:109-119.
- Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, Mass.: MIT Press.
- Koster, Jan. 1986. The Relation between pro-drop, Scrambling, and Verb Movements. Ms., Rijksuniversiteit Groningen.
- Koster, Jan. 1987. *Domains and Dynasties*. Dordrecht: Foris.
- Kroch, Anthony. 1974. *The Semantics of Scope in English*. Dissertation, MIT, Cambridge, Mass.
- Kuhn, Jonas. 2000. On Faithfulness Violations and the 'Direction' of Optimization in a Declarative Approach to OT Syntax. Ms., Universität Stuttgart.
- Lakoff, George. 1971. On Generative Semantics. In *Semantics*, eds. Danny Steinberg & Leon Jakobovits, 232-296. Cambridge: Cambridge University Press.
- Lasnik, Howard, & Juan Uriagereka. 1988. *A Course in GB Syntax*. Cambridge, Mass.: MIT Press.
- Lasnik, Howard, & Mamoru Saito. 1992. *Move α* . Cambridge, Mass.: MIT Press.
- Legendre, Géraldine. 1998. Why French Stylistic Inversion is Optimal. Ms., Johns Hopkins University.
- Legendre, Géraldine, William Raymond & Paul Smolensky. 1993. An Optimality-Theoretic Typology of Case and Grammatical Voice Systems. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 19:464-478.
- Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer & William Raymond. 1995. Optimality and Wh-Extraction. In *Papers in Optimality Theory*, eds. Jill Beckman, Laura Walsh-Dickie & Suzanne Urbanczyk, 607-635. Amherst, Massachusetts: UMass Occasional Papers in Linguistics 18.
- Legendre, Géraldine, Paul Smolensky & Colin Wilson. 1998. When is Less More? Faithfulness and Minimal Links in Wh-Chains. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 249-289. Cambridge, Mass.: MIT Press & MITWPL.
- Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Stauffenburg.
- Lenerz, Jürgen. 1992. Zur Syntax der Pronomina im Deutschen. *Sprache und Pragmatik* 29, University of Lund.
- Lenerz, Jürgen. 1999. Word Order Variation: Competition or Cooperation? Ms., Universität Köln. Erscheint in *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld. Berlin: Mouton/de Gruyter.
- Lerot, Jacques. 1985. Zur Wortstellungsnorm im Deutschen. *Deutsche Sprache*, 137-142.
- Lötscher, Andreas. 1981. Abfolgeregeln für Ergänzungen im Mittelfeld. *Deutsche Sprache*, 44-60.
- Lutz, Uli, Gereon Müller & Arnim von Stechow (eds.). 2000. *Wh-Scope Marking*. Amsterdam: Benjamins.

- McCarthy, John, & Alan Prince. 1994. The Emergence of the Unmarked: Optimality in Prosodic Morphology. *Proceedings of NELS* 24:333-379.
- McCarthy, John, & Alan Prince. 1995. Faithfulness and Reduplicative Identity. In *Papers in Optimality Theory*, eds. Jill Beckman, Laura Walsh-Dickie & Suzanne Urbanczyk, 249-384. Amherst, Massachusetts: UMass Occasional Papers in Linguistics 18.
- McCarthy, John. 1999. Harmonic Serialism and Parallelism. Ms., University of Massachusetts, Amherst.
- McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics*. Dordrecht: Reidel.
- McCloskey, James. 1988. Syntactic Theory. In *Linguistics: The Cambridge Survey, vol. 1*, ed. Frederick J. Newmeyer, 18-59. Cambridge: Cambridge University Press.
- McDaniel, Dana. 1989. Partial and Multiple *Wh*-movement. *Natural Language and Linguistic Theory* 7:565-604.
- Manzini, Rita. 1983. On Control and Control Theory. *Linguistic Inquiry* 14:421-446.
- Manzini, Rita, & Kenneth Wexler. 1987. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18:413-444.
- Marantz, Alec. 1995. The Minimalist Program. In *Government and Binding Theory and the Minimalist Program*, ed. Gert Webelhuth, 351-382. Oxford: Blackwell.
- May, Robert. 1985. *Logical Form*. Cambridge, Mass.: MIT Press.
- Meinunger, André. 1995. Prominence Hierarchy and Phrase Ordering. *FAS Working Papers* 2:95-121. ZAS: Berlin.
- Müller, Gereon. 1995. *A-bar Syntax*. Berlin: Mouton/de Gruyter.
- Müller, Gereon. 1997. Partial *Wh*-Movement and Optimality Theory. *The Linguistic Review* 14:249-306.
- Müller, Gereon. 1999a. Order Preservation, Parallel Movement, and the Emergence of the Unmarked. Ms., Universität Stuttgart. Erscheint in *Optimality Theoretic Syntax*, eds. Jane Grimshaw, Géraldine Legendre & Sten Vikner. Cambridge, Mass.: MIT Press.
- Müller, Gereon. 1999b. Optionality in Optimality-Theoretic Syntax. *GLOT International* 4.5:3-8.
- Müller, Gereon. 1999c. Optimality, Markedness, and Word Order in German. *Linguistics* 37:777-818.
- Müller, Gereon. 1999d. Shape Conservation and Remnant Movement. Ms., Universität Tübingen. Erscheint in *Proceedings of NELS 30*. Amherst, Mass.: GLSA.
- Müller, Gereon. 2000a. Das Pronominaladverb als Reparaturphänomen. *Linguistische Berichte* 182:139-178.
- Müller, Gereon. 2000b. On Common-Sense Justifications of Optimality-Theoretic Constraints. *Zeitschrift für Sprachwissenschaft* 18 (1999!):230-234.
- Müller, Gereon, & Wolfgang Sternefeld. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24:461-507.
- Müller, Gereon, & Wolfgang Sternefeld. 1994. Scrambling as A-bar Movement. In *Studies on Scrambling*, eds. Norbert Corver & Henk van Riemsdijk, 331-385. Berlin: Mouton de Gruyter.
- Müller, Gereon, & Wolfgang Sternefeld. 1996. A-bar Chain Formation and Economy of Derivation. *Linguistic Inquiry* 27:480-511.
- Nakamura, Masanori. 1998. Reference Set, Minimal Link Condition, and Parametrization. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 291-313. Cambridge, Mass.: MIT Press & MITWPL.

- Neeleman, Ad. 1994. Scrambling as a D-structure Phenomenon. In *Studies on Scrambling*, eds. Norbert Corver & Henk van Riemsdijk, 387-429. Berlin: Mouton de Gruyter.
- Pafel, Jürgen. 1993. Scope and Word Order. In *Syntax, vol I*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 867-879. Berlin: DeGruyter.
- Pafel, Jürgen. 1998. *Skopus und logische Struktur – Studien zum Quantorenskopos im Deutschen*. Habilitationsschrift, Universität Tübingen.
- Partee, Barbara, Alice ter Meulen & Robert Wall. 1993. *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.
- Perlmutter, David, & Scott Soames. 1979. *Syntactic Argumentation and the Structure of English*. Berkeley: The University of California Press.
- Perlmutter, David. 1996. How Simple Universal Constraints Make Morphology More Complex. Ms., UC San Diego.
- Pesetsky, David. 1997. Optimality Theory and Syntax: Movement and Pronunciation. In *Optimality Theory. An Overview*, eds. Diana Archangeli & D. Terence Langendoen, 134-170. Oxford: Blackwell.
- Pesetsky, David. 1998. Some Optimality Principles of Sentence Pronunciation. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 337-383. Cambridge, Mass.: MIT Press & MITWPL.
- Pollard, Carl, & Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford: CSLI Publications.
- Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 30:365-424.
- Primus, Beatrice. 1994. Grammatik und Performanz: Faktoren der Wortstellungsvariation im Mittelfeld. *Sprache und Pragmatik* 32, 39-86.
- Prince, Alan & Paul Smolensky. 1993. *Optimality Theory. Constraint Interaction in Generative Grammar*. Ms., Rutgers University. Erscheint: Cambridge, Mass.: MIT Press.
- Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. London: Croom Helm.
- Reis, Marga. 1986. Die Stellung der Verbargumente im Deutschen. Stilübungen zum Grammatik:Pragmatik-Verhältnis. Ms., Universität Tübingen. Erschienen 1987 in *Sprache und Pragmatik*, ed. Inger Rosengren, 139-178. Stockholm: Almqvist & Wiksell International.
- Reis, Marga. 1995. Über infinite Nominativkonstruktionen im Deutschen. In *Festvorträge anlässlich des 60. Geburtstags von Inger Rosengren*, ed. Olaf Önnersfors, 114-156. Lund: Sprache & Pragmatik.
- Reuland, Eric, & Wim Kosmeijer. 1993. Projecting Inflected Verbs. In *The Parametrization of Universal Grammar*, ed. Gisbert Fanselow, 37-71. Amsterdam: Benjamins.
- Richards, Norvin. 1997. Competition and Disjoint Reference. *Linguistic Inquiry* 28:178-187.
- Riemsdijk, Henk van. 1982. Correspondence Effects and the Empty Category Principle. *Tilburg Papers in Language and Literature* 12. University of Tilburg.
- Riemsdijk, Henk van, & Edwin Williams. 1981. NP-Structure. *The Linguistic Review* 1:171-217.
- Rizzi, Luigi. 1982. Violations of the Wh-Island Constraint and the Subjacency Condition. In ders., *Issues in Italian Syntax*, 49-76. Dordrecht: Foris.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- Rizzi, Luigi. 1992. Argument/Adjunct (A)symmetries. *Proceedings of NELS* 22:365-381.
- Roberts, Ian. 1993. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer.

- Roberts, Ian. 1997. Restructuring, Head Movement, and Locality. *Linguistic Inquiry* 28:423-460.
- Rochemont, Michael. 1989. Topic Islands and the Subjacency Parameter. *Canadian Journal of Linguistics* 34:145-170.
- Rohrbacher, Bernhard. 1994. *The Germanic Languages and the Full Paradigm: A Theory of V to I Raising*. Dissertation, University of Massachusetts, Amherst.
- Rohrbacher, Bernhard. 1999. *Morphology-Driven Syntax. A Theory of V to I Raising and Pro Drop*. Amsterdam: Benjamins.
- Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Dissertation, MIT, Cambridge, Mass. Erschienen 1986 als *Infinite Syntax*. Norwood, New Jersey: Ablex Publishing Corporation.
- Rubach, Jerzy. 2000. Glide and Glottal Stop Insertion in Slavic Languages: A DOT Analysis. *Linguistic Inquiry* 31:271-317.
- Rudin, Catherine. 1985. *Aspects of Bulgarian Syntax: Complementizers and Wh-Constructions*. Columbus, Ohio: Slavica.
- Rudin, Catherine. 1988. On Multiple Questions and Multiple WH Fronting. *Natural Language and Linguistic Theory* 6:445-501.
- Russell, Kevin. 1997. Optimality Theory and Morphology. In *Optimality Theory. An Overview*, eds. Diana Archangeli & D. Terence Langendoen, 102-133. Oxford: Blackwell.
- Sabel, Joachim. 1998. *Principles and Parameters of Wh-Movement*. Habilitationsschrift, Universität Frankfurt/Main.
- Samek-Lodovici, Vieri. 1996. *Constraints on Subjects. An Optimality-Theoretic Analysis*. Dissertation, Rutgers University: New Brunswick, New Jersey.
- Samek-Lodovici, Vieri, & Alan Prince. 1999. Optima. Technical Reports of the Rutgers Center for Cognitive Science 57, Rutgers University.
- Sauerland, Uli. 1995. Early Features. *MIT Working Papers in Linguistics* 27:223-242.
- Savigny, Eike von. 1983. *Zum Begriff der Sprache. Konvention, Bedeutung, Zeichen*. Stuttgart: Reclam.
- Schmid, Tanja. 1998. West Germanic "Infinitivus Pro Participio" (IPP) Constructions in Optimality Theory. In *Proceedings of Console VII*, ed. Tina Cambier-Langeveld, Anikó Lipták, Michael Redford & Erik Jan van der Torre, 229-244. Leiden: SOLE.
- Schmid, Tanja. 1999. Die Ersatzinfinitivkonstruktion im Deutschen. Ms., Universität Stuttgart. Erscheint in *Linguistische Berichte*.
- Sells, Peter, John Rickford & Thomas Wasow. 1996. An Optimality Theoretic Approach to Variation in Negative Inversion in AAVE. *Natural Language and Linguistic Theory* 14:591-627.
- Shlonsky, Ur. 1992. Resumptive Pronouns as a Last Resort. *Linguistic Inquiry* 23:443-468.
- Siewierska, Anna. 1993. On the Interplay of Factors in the Determination of Word Order. In *Syntax, vol I*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 826-846. Berlin: DeGruyter.
- Smolensky, Paul. 1995. On the Internal Structure of Con, the Constraint Component of UG. Ms., Johns Hopkins University.
- Smolensky, Paul. 1996. On the Comprehension/Production Dilemma in Child Language. *Linguistic Inquiry* 27:720-731.
- Speas, Margaret. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- Speas, Margaret. 1995. Generalized Control and Null Objects in Optimality Theory. In *Papers in Optimality Theory*, eds. Jill Beckman, Laura Walsh-Dickie & Suzanne Ur-

- banczyk, 637-653. Amherst, Massachusetts: UMass Occasional Papers in Linguistics 18.
- Stechow, Arnim von. 1980. Wie interessant ist die Syntaxforschung heute? *Studium Linguistik* 8/9:32-59.
- Stechow, Arnim von. 1992. Kompositionsprinzipien und grammatische Struktur. In *Biologische und soziale Grundlagen der Sprache*, ed. Peter Suchsland, 175-248. Tübingen: Niemeyer.
- Stechow, Arnim von. 1993. Die Aufgaben der Syntax. In *Syntax, vol. I*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann, 1-88. Berlin: de Gruyter.
- Stechow, Arnim von, & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.
- Sternefeld, Wolfgang. 1990. Seminarskript zum Kurs "Bausteine syntaktischen Wissens". Ms., Universität Konstanz.
- Sternefeld, Wolfgang. 1991. *Syntaktische Grenzen*. Opladen: Westdeutscher Verlag.
- Sternefeld, Wolfgang. 1997. Comparing Reference Sets. In *Economy in Linguistic Theory*, eds. Chris Wilder, Hans-Martin Gärtner & Manfred Bierwisch, 81-114. Berlin: Akademieverlag.
- Stiebels, Barbara. 2000. Linker Inventories, Linking Splits, and Lexical Economy. In *Lexicon in Focus*, ed. Barbara Stiebels & Dieter Wunderlich, 213-247. Berlin: Akademie Verlag. (= *Studia Grammatica* 45)
- Stowell, Tim. 1981. *Origins of Phrase Structure*. Dissertation, MIT, Cambridge, Mass.
- Tesar, Bruce. 1998. Error-Driven Learning in Optimality Theory via the Efficient Computation of Optimal Forms. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 421-435. Cambridge, Mass.: MIT Press & MITWPL.
- Tesar, Bruce, & Paul Smolensky. 1993. The Learnability of Optimality Theory.: An Algorithm and Some Basic Complexity Results. Technical Report, Computer Science Department, University of Colorado, Boulder.
- Tesar, Bruce, & Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29:229-268.
- Tesar, Bruce, & Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, Mass.: MIT Press.
- Thiersch, Craig. 1978. *Topics in German Syntax*. Dissertation, MIT, Cambridge, Mass.
- Uszkoreit, Hans Jürgen. 1984. *Word Order and Constituent Structure in German*. Dissertation, University of Austin, Texas.
- Uszkoreit, Hans Jürgen. 1986. Constraints on Order. *Linguistics* 24:883-906.
- Vallduví, Enric. 1992. *The Informational Component*. New York: Garland.
- Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York & Oxford: Oxford University Press.
- Vikner, Sten. 1997a. V-to-I Movement, 'Do'-Insertion, and Negation in Optimality Theory. Ms., Universität Stuttgart.
- Vikner, Sten. 1997b. The Interpretation of Object Shift, Optimality Theory, and Minimalism. *Working Papers in Scandinavian Syntax* 60:1-24.
- Vikner, Sten. 1999. V-to-I Movement and 'Do'-Insertion in Optimality Theory. Ms., Universität Stuttgart. Erscheint in *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw & Sten Vikner. Cambridge, Mass.: MIT Press.
- Vikner, Sten. 2000. Checking Strong Verbal Inflection in Optimality Theory. Ms., Uni-

- versität Stuttgart.
- Vogel, Ralf. 1999. The Typology of Case Conflicts in Free Relative Constructions. Ms., Universität Stuttgart. Erscheint in *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld. Berlin: Mouton/de Gruyter.
- Vogel, Ralf, & Markus Steinbach. 1998. The Dative – an Oblique Case. *Linguistische Berichte* 173:65-90.
- Wartena, Christian. 2000. A Note on the Complexity of Optimality Systems. Ms., Universität Potsdam.
- Watanabe, Akira. 1992. Wh-in-Situ, Subjacency, and Chain Formation. Ms., MIT, Cambridge, Mass. (*MIT Occasional Papers in Linguistics* 2.)
- Webelhuth, Gert. 1987. Eine universale Scrambling-Theorie. Paper presented at the annual meeting of the DGfS, Augsburg.
- Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation*. Oxford: Oxford University Press.
- Williams, Edwin. 1997. Blocking and Anaphora. *Linguistic Inquiry* 28:577-628.
- Williams, Edwin. 1999. Economy as Shape Conservation. Ms., Princeton University.
- Wilson, Colin. 1998. Bidirectional Optimization and the Theory of Anaphora. Ms., Johns Hopkins University. Erscheint in *Optimality Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw & Sten Vikner. Cambridge, Mass.: MIT Press.
- Woolford, Ellen. 1999. Case Patterns. Ms., University of Massachusetts, Amherst. Erscheint in *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw & Sten Vikner. Cambridge, Mass.: MIT Press.
- Wunderlich, Dieter. 1987. An Investigation of Lexical Composition: The Case of German *be*-Verbs. *Linguistics* 25:283-331.
- Wunderlich, Dieter. 1997. Cause and the Structure of Verbs. *Linguistic Inquiry* 27:27-68.
- Wunderlich, Dieter. 2000. The Force of Lexical Case: German and Icelandic Compared. In *Papers on Argument Linking*, ed. Dieter Wunderlich, 81-102. Düsseldorf: Arbeiten des SFB 282 'Theorie des Lexikons', Nr.112.

Index

OT-Beschränkungen

- *[+hr], 283
- *[+tr], 283
- *Ø (morphologisch), 309
- *Ø (Vermeide leere Kandidaten), 84, 289
- *AKK, 260
- *DAT, 260
- *GE-MV (*ge-Modalverben), 41, 42
- *HOCH-INF, 289
- *INF-PROP, 289
- *KASUS, 266
- *KOMP-KOPF (Vermeide komplexe Köpfe), 186
- *KOMP-W (Vermeide komplexe W-Cluster), 188
- *NICHT-SU/3, 304
- *NICHT-SU/AGENS, 304
- *NICHT-SU/LOKAL, 304
- *NICHT-SU/PATIENS, 304
- *NOM-INF, 289
- *NOM, 260
- *PERS-PRON (Vermeide Personalpronomina), 176
- *PRON (Vermeide Pronomina), 179
- *STRUK, 309
- *SU/3, 304
- *SU/AGENS, 304
- *SU/LOKAL, 304
- *SU/PATIENS, 304
- 1K1K (Ein Kette, ein Kasus), 273
- 1K1R (Eine Kette, eine Realisierung), 276
- ADJA (Adjazenzbeschränkung), 251
- ADJUNKT-INSEL, 83
- AGENS (Agentivitätsbedingung), 251
- ANIM (Belebtheitsbedingung), 251
- BAR^{2[-ref]} (Zwei-Barrieren-Bedingung für Adjunktketten), 132, 136
- BAR (Barrierenbedingung), 55, 136, 213
- BARⁿ&_lREF, 139
- BAR&_lBAR, 138
- BEL (Belebtheitsbedingung), 66, 242
- BFE (Beschränkung für Extraktionsbereiche), 165
- BIND-TREUE (Bindungstreue), 312
- BKP (Bedingung der kürzesten Pfade), 202
- BL-EIN (Blaulicht-Einsatzhorn), 6
- DAT (Dativbedingung), 242
- DEFAULT-NOM (Default-Nominativ), 60
- DEF (Definitheitsbedingung), 66, 217, 242, 251
- DEP, 29
- DIST (Distributivität), 153
- ECP (Empty Category Principle), 19
- EIN (Eindeutigkeit), 283
- ENDFOK (EndFokus), 233
- EPP (Erweitertes Projektionsprinzip), 250
- F-TREUE, 274
- FILL, 29
- FOKUSAUSRICHTUNG, 226
- FOK (Fokusbedingung), 66, 242
- GRAM-FUNK (Grammatische Funktion), 153
- H-NUK (Harmonie des Nukleus), 302
- IP-KRIT (IP-Kriterium), 185
- IDENT-F(EATURE), 31
- IDENT([hr]), 281
- IDENT([tr]), 281
- IND (Indefinita), 233
- INSEL, 127
- K-EIN (Kasuseindeutigkeit), 272
- KAS-THETA-EIN, 269
- KASUS, 22, 100
- KN₁ (Kanonische Abfolge₁), 227
- KN₂ (Kanonische Abfolge₂), 227
- KOMP-REAL (Komparativrealisierung), 173
- KOND (Konditionalbedingung), 109
- KONTROLLREGEL, 178, 182
- KONVERGENZ, 70
- KOPF-LINKS, 114, 298
- KOPF-LINKS^{exp}, 298
- KOPF-LINKS^{mtx}, 298
- KOPF-REAL (Kopf-Realisierung), 186
- KOPF-RECHTS, 114
- L-ZEI (Lichtzeichen), 6
- LR(CP) (Linker Rand(CP)), 120, 208
- LETZT-AUS (Letzter Ausweg), 19
- LEX-TREUE, 260
- LEX-TREUE_{tr}, 260
- LEX-ÖKON (Bewegungsökonomie für lexikalische Köpfe), 22, 78, 100, 198
- LOK-ANT (Lokales Antezedens), 311
- M-TREUE (Merkmalstreue), 313
- MAX, 29
- MAX([+hr]), 281
- MAX([+tr]), 281
- NEU, 227
- NOM!, 287
- NOM (Nominativbedingung), 217, 242

- O-KASUS (Objektkasusbedingung), 66
 OCP_{Kas}, 264
 OB-KOPF (Obligatorische Köpfe), 22, 52, 95, 197, 298
 OB-KOPF^{exp}, 298
 OB-KOPF^{mtx}, 298
 OBJEKT-LIZ (Objekt-Lizensierung), 68
 ÖKON (Ökonomie), 19, 22, 30, 52, 55, 70, 95, 197, 213, 233
 OP-SPEZ (Operator im Spezifikator), 95, 108
 PF-KONVERGENZ, 70
 PP-INSEL, 43
 PAR-VER (Parallelverschiebung), 60, 270
 PARALLELISMUS, 51
 PARSE, 29
 PRINZIP A, 175
 PRINZIP B, 175
 PROJ-P (Projektionsprinzip), 41, 43
 PROM (Prominenz), 227
 PRON-KRIT (Pronomen-Kriterium), 43
 PRON-KRIT^{schwach}, 47
 PRON-KRIT^{unbetont}, 47
 PUR-EP (Reinheit der erweiterten Projektion), 37, 52, 104, 197, 221, 297
 Q-MARK (Q-Merkmal-Markierung), 202
 Q-SKOP (Q-Skopus), 202
 Q-ÖKON (Quantorenökonomie), 51
 QR, 318
 QUIB, 318
 REALAKK (Realisiere Akkusativ), 274
 REALNOM (Realisiere Nominativ), 274
 REALOBL (Realisiere obliquen Kasus), 274
 REAL (Realisierung, Parse), 81
 REF-ÖKON (Referentielle Ökonomie), 313
 REF (Referentialität), 138
 REK, 318
 RVL (Rechts vor Links), 6
 S-LEX-REK (Lexikalische Spur-Rektion), 111
 S-REK (Spur-Rektion), 111
 SP (Skopus-Prinzip), 317
 SCR-KRIT (Scrambling-Kriterium), 242
 SEL (Selektion), 132
 SEMHIER (Semantische Hierarchie), 290
 SICHTBARKEIT, 288
 SKOP-TREUE (Skopus-Treue), 132
 SKOPUS, 68
 SPECV/2, 49
 SPEZ (Spezifität), 153
 SPRICH-KOPF, 126
 ST-SP (Stille Spur), 127, 209
 STR-BD (Strikte Bindung), 78, 198
 SUB-TREUE (Subordinationstreue), 221
 SUBJEKT-KASUS, 226
 SUBJEKT, 22
 SUFFIX-FILTER, 173
 SVF (Straße vor Feldweg), 6
 SYN-KON (Syntaktische Konstellation), 153
 TCP (Tilgung in CP), 208
 TEL (Telegraph), 120, 208
 TOP-TIL (Topik-Tilgung), 80
 TOP-TREUE (Topik-Treue), 136
 TREUE(LEX-KONTROLLE), 183
 TYP, 317
 V-POL (Verkehrspolizist), 6
 V-ZEI(A) (Verkehrszeichen(a)), 6
 V-ZEI(B) (Verkehrszeichen(b)), 6
 VERMEIDE PHRASENAUFBAU, 173
 VOLL-INT (Vollständige Interpretation), 22, 29, 30, 55, 98, 213
 W-KRIT (W-Kriterium), 31, 36, 55, 185, 213
 W-REAL (W-Realisierung), 188
 W-TREUE, 87, 135
 WIED (Wiederauffindbarkeit), 30, 120, 208
 ZAUDERPRINZIP (Procrastinate), 70

Sachindex

- A-Bewegung, 61
A.c.I.-Konstruktionen, 60, 146, 259, 288, 293
absolute Ungrammatikalität, 44, 82, 136, 220
 im Generator, 83
 und leere Kandidaten, 83, 289
 und Neutralisierung, 88, 134, 274
 und Reparatur, 86
 und schlechte Gewinner, 85
Adjazenz, 251, 253, 254
Adjunkt-Insel, 82, 142, 151
Affix Hopping, 19, 79
Affiziere Alpha (Affect Alpha), 190
Affiziertheit, 194
Agentivität, 251–253
Alternation, 195
Ambiguitäten
 derivationelle, 90, 134
 lokale, 321
Applikativ, 194
Attraktion (Attract), 38, 71
bare output conditions, 157
Barrieren
 Definition von, 133, 147
 Zahl der überschrittenen, 143
Basisabfolge der Argumente, 58, 236, 238, 240, 251, 278
Belebtheit, 66, 226, 240, 241, 245, 247, 248, 251–253, 285
Beschränkungen
 Ökonomie-, 297
 Abbildungs-, 32, 297
 aktive, 73
 als globale Regeln, 60
 als Oberflächenfilter, 60
 Ausrichtungs- (alignment), 32, 299
 derivationelle, 61
 diskursbasierte, 225, 227
 funktionale Motivation für, 77
 Geordnetheit von, 9
 lokale, 157
 Markiertheits- (markedness), 31
 Restriktionen für, 76, 297–301
 spezielle vs. generelle, 72
 Sprachspezifität von, 9
 Struktur-, 32, 297
 transderivationelle, 70, 157, 167, 191
 transderivationelle – im Konflikt, 167
 translokale, 157, 179, 295, 302
 transrepräsentationelle, 157
 Treue- (faithfulness), 29
 Ungeordnetheit von, 9
 Universalität von, 8
 Verletzbarkeit von, 9
Beschränkungsausrichtung
 Definition, 301
Beschränkungsordnung
 vollständige vs. partielle, 201
Beschränkungsprofil, 10
Bindungstheorie, 146, 175, 238, 255, 278, 280, 310–316
Blockade, 170, 267
Blockade-Syntax
 Struktur einer, 174
Burzios Generalisierung, 263, 272, 287, 295
crucial nonranking, 212
D-Schale, 274
D-Struktur
 Optimierung der, 251, 317
Datenerhebung, 156
Dativ-Anhebung, 189, 194, 195
Default-Kasus, 59, 261
 bei Kasuskongruenz mit PRO, 59
 bei Linksversetzung, 59
 bei prädikativen und-Konstruktionen, 59
Defaultmarkierung in LDG, 279, 281, 286, 288
Definitheit, 66, 155, 217, 226, 241, 245, 251, 252, 318
derivationelle Theorie der Komplexität (DTK), 240, 252, 280, 318
Diskursprominenz, 304
Distributivität, 152
Do-Einsetzung
 und Emphase, 98
 und Fragesätze, 98, 103
 und Negation, 18, 40, 97, 103
Doppel-Comp-Filter, 114, 115, 125, 126
Echofrage, 194
ECP (Empty Category Principle), 82, 143, 151, 161
Elsewhere Condition, 170, 295
Ergativ, 260, 264, 272, 278, 282–285, 287, 294
Ersatzinfinitiv, 40
erweiterte Projektion, 23, 94, 106, 109, 199, 298
erweiterte V-Kette, 42
Exceptional Case Marking (ECM), 60

- Expletiva
 und Allgemeinheit der Bedeutung, 98
 Expletiversetzung (expletive replacement),
 72
 expletives es, 48
 Extraposition, 190, 194, 196
- faktorielle Typologie, 39, 77, 116, 125, 129,
 205, 263, 285, 299, 323
 Feeding, 301
 Fokus, 149, 155, 163, 194, 218, 226, 233, 239,
 241, 248, 255
 komplettierender vs. kontrastiver, 228
 Fokusexponent, 149, 218
 Fokusprojektion
 maximale, 239
 Funktionsverbgefüge, 247
- GB-Theorie, 1–3, 9, 11, 82, 129, 145, 146,
 148, 151, 175, 178, 181, 185, 190,
 259, 272, 277, 289, 294, 300, 310,
 316, 320
 Generalisierte Phrasenstrukturgrammatik
 (GPSG), 1, 3, 148
 Generative Semantik, 15, 131, 277, 316
 Generator (Gen), 13
 Genitiv, 259, 260, 268–271, 280, 293
 Gerundivkonstruktionen, 178
 Gesetz der wachsenden Glieder, 257
 GF-Realisierung, 303
 Gier (Greed), 71
 Grammatikalität, 9, 154
 im minimalistischen Programm, 169
 Grammatikalitätsgrade, 57, 67, 82, 143, 151,
 239, 242, 250, 255
 Grenzknoten, 147
- Harmonie-Evaluation (H-Eval), 13
 harmonische Ausrichtung, 301–309
 Definition, 301
 harmonische Begrenzung, 59, 62, 63, 97, 125,
 205, 218, 230
 Definition, 59
 durch mehrere Kandidaten, 63, 205
 Head-Driven Phrase Structure Grammar
 (HPSG), 1, 3, 255
 Hypothese der prädikats-internen Subjekte,
 23
- I-Prinzip, 320
 I-Topikalisierung, 152
 Indefinita
 existenziell vs. generisch interpretierte,
 233, 235
 Indefinitheit, 233, 318
 Index, 131
 Input, 13, 131, 241, 274, 281, 304, 311, 316,
 317, 320
 Input-Optimierung, 91, 134, 223
 Definition von, 134
 Intonation, 152
- Jojo-Bewegung, 162
- Kandidaten
 Definition von, 11, 93, 118, 130, 241,
 271, 310, 311, 316
 Grammatikalität von, 9
 Wettbewerb von, 9
 Kandidatenmengen
 Definition von, 12, 83, 93, 118, 130, 164,
 170, 193, 267, 311, 321
 Finitheit von, 119
 Größe von, 76, 115, 119, 223, 319, 321
 Zahl von, 223
 Kasus, 59–60, 259–295
 -attraktion, 276
 Default-, *siehe* Default-Kasus
 inhärenter, 259–261, 266–269, 279, 280,
 285–294
 inhärenter vs. struktureller, 259
 inhärenter vs. struktureller in LDG, 286
 lexikalischer, *siehe* inhärenter Kasus
 Null-, 266
 struktureller, 259, 260, 266, 269, 281–
 285
 Kasusanhebung, 271
 Kasusfilter, 260, 261, 277
 Kettenbildung (Form Chain), 160, 164
 Kettendurchschuß (chain interleaving), 167
 Koalitionsmodell, 155
 kohärente Konstruktion, 185
 Kohärenzfeld, 42
 Komparativ, 170
 Komparativregel A, 171
 Komparativregel B, 171
 Komparativregel B', 171
 Komplementierer
 satzfinaler, 184
 Komplementierer-Spur-Effekt, 111
 Anti-, 113
 Komplementiererertilgung, 27, 51, 110, 189,
 197, 207, 220
 als IP-Einbettung, 110

- als PF-Tilgung, 119
- und Brückenprädikate, 53, 110
- und eingebettete Adjunkte, 54, 110, 197, 222
- und negative Operatoren, 110
- und Subjekt-Relativierung, 111, 113
- und Subjektsätze, 53, 110, 195, 207
- und Topikalisierung, 53, 110, 195
- komplexer Verlierer, 63, 205
- Komplexität, 74, 144, 206, 222, 319
- Konditionalsätze, 108
- Kontexttypen, 239
- Kontrollinfinitive, 146, 182, 288, 289, 291–295, 310
- Konvergenz, 70, 160, 166
- Kopiertheorie der Bewegung, 30, 127, 271
- Kopplung, 26, 52, 123, 124, 156, 192, 200, 228, 234
 - global disjunktive, 217, 250
 - global hierarchische, 202, 250
 - globale vs. lokale, 201
 - lokal disjunktive, 216
 - lokal hierarchische, 207, 211
 - lokal konjunktive, 212
 - und Spracherwerb, 215, 250
 - von mehr als zwei Beschränkungen, 206, 209
- Kumulativität, 17, 137, 149, 155, 156, 246, 256, 292
- L-Markierung, 133, 147
- Lambda-Abstraktion, 277
- leerer Operator, 131
- Lexikalisch-Funktionale Grammatik (LFG), 1, 2, 227, 228
- lexikalisch-konzeptuelle Struktur, 98
- Lexikalische Dekompositionsgrammatik (LDG), 277–294
- lexikalische Idiosynkrasien, 183
- Lexikon
 - als Sklave der Syntax, 79, 102, 177
- Lexikon-Optimierung, 91, 134
- LF-Tilgung, 158
- Linear Correspondence Axiom (LCA), 183
- lokale Konjunktion, 138, 139, 143, 144, 156, 301, 309, 323
 - als Disjunktion, 138
 - Definition, 137
 - mit Beschränkungshierarchien, 309
 - reflexive, 137
- Lokalitätstheorie, 129, 130, 136, 140, 143, 220, 250, 300
- Matching, 275, 276
- Merge, 173, 322
- Merkmalsbewegung, 271
- Minimal Attachment, 295
- Minimal Link Condition, 65, 162, 320
- Minimale Projektion, 200
- minimalistisches Programm, 2, 12, 38, 66, 70–72, 76, 129, 156–169, 174, 190, 191, 193, 203
- Modalverb, 40
- Morphologie, 170
- Negation, 18, 195
 - negative Inversion, 206
 - negative Operatoren, 106
- Neutralisierung, 88, 90–92, 131, 133–136, 141–144, 192, 219, 262, 274, 275, 289, 306–308, 312–314
 - bei absoluter Ungrammatikalität vs. bei Optionalität, 220
- NP-Schale, 142
- NP-Struktur, 271
- Numeration, 12, 14, 15, 21, 28, 31, 164, 176, 181, 191, 193, 241, 281, 316, 317, 321
- Object Shift, 68
- Objektbewegung, 68
- Obligatory Contour Principle, 264
- Optimalität, 10, 154
 - als Grammatikalität, 244
 - als Unmarkiertheit, 244
 - im minimalistischen Programm, 157, 167
- Optimierung
 - bidirektionale, 315, 320
 - expressive, 311
 - globale, 321
 - interpretative, 311, 319
 - lokale, 321
 - mehrfache, 310
 - parallele, 310, 315, 316, 319
 - serielle, 310, 315, 316, 319, 320
- Optionalität, 26, 82, 90, 152, 174, 177, 189
 - als obligatorische Bewegung mit optionalen Merkmalen, 191
 - echte, 26, 52, 124, 192, 196
 - Pseudo-, 26, 155, 174, 192, 193, 225, 227, 233, 254, 318
 - Zusammenbruch von, 195
- Output, 13

- Pāninis Theorem zur Beschränkungsordnung, 73
- Pāninische Relation, 73, 74, 289
Definition, 72
- Parametrisierung
als Beschränkungsordnung, 17, 33
in der GB-Theorie, 146
lexikalische, 147
und faktorielle Typologie, 77
und morphologische Fundierung, 77
- Parsing, 295, 320
- partielle W-Bewegung, 54, 190, 195, 213
- Passiv, 164, 259, 303, 305–309
Rezipienten-, 288, 294
- PF-Realisierung, 118, 271
in Kopf-Ketten, 185
in W-Clustern, 187
von Spuren, 127
- Pfad, 161
- Phi-Merkmale, 313
- Phraseologismus, 247
- Pointing Finger, 16
- polysynthetische Sprachen, 187
- Positionsprominenz, 302
- Possessivkonstruktion, 172
- Postpositionsstranden, 44
- Prädikat-/Argument-Struktur, 12, 15, 21, 28, 29, 31, 93, 94, 130, 193
- Präferenzprinzip für Rekonstruktion, 166
- Präpositionsstranden, 44, 122
- Prinzip der kürzesten Pfade (Shortest Path Condition), 161, 203
- Prinzip der Vollständigen Inklusion (Proper Inclusion Principle), 175
- Prinzip der Vollständigen Spezifikation, 175
- Prinzip der wenigsten Schritte (Fewest Steps Condition), 158, 191
- Prinzipien-und-Parameter-Theorie, *siehe* GB-Theorie
- Pro-Drop-Parameter, 80, 81, 181
- PRO-Theorem, 11, 178
- Projektionsprinzip
als globale Regel, 61
- Pronomen
schwaches, 43, 44
starkes, 44
- Pronominaladverb, 42
- psychologisches Prädikat, 280
- Q-Prinzip, 320
- Quantor
herumtreibender, 186
- Quantorenanhebung
als Reparatur, 50
- Quantorenskopus, 50, 61, 68, 152, 316, 317
in der Optimalitätstheorie, 155
- R-Pronomen, 42
- Rattenfang (pied piping), 45, 121, 122, 124
- Referenzmenge, 157
- Reflexivpronomen, 146, 174, 238, 311
morphologische Form, 147
- Regel R, 79
- Relativsatz, 111, 113, 119–124, 126, 130, 148, 209, 215, 218, 274
als Adjunkt, 113
extraponierter, 190, 194
freier, 272–277, 294, 295
mit obliquem Pronomen, 209
- Reparatur, 18, 39, 116, 128, 156, 322, 323
- Restrukturierung, 185
- resumptives Pronomen, 86, 127, 209
- Rezipropronomen, 238, 311
- Scrambling, 190, 194, 196, 217, 238, 241, 252, 271, 280, 318
ketteninvariantes, 58, 249
- Senkung (lowering), 163
- Skala, 302
GF-Hierarchie, 304
Person-Hierarchie, 304
Prominenz-Hierarchie, 304
- Skopusbeschränkung, 154
- Skopusinversion, 152
- Skopusmarker, 55, 214, 317
abstrakte, 131, 134
- Skopuswert, 152
- Sonoritätshierarchie, 302
- spezifische Form, 170
- Spezifizität, 152
- Spezifizitätsprinzip, 278
- Sprachen
African American Vernacular English, 206
Arabisch, 225
Baskisch, 285
Bulgarisch, 35, 37–39, 61–65, 96, 97, 187, 188, 203, 205, 270
Chamorro, 314
Chinesisch, 33, 35, 81, 131–133, 139, 203, 212
Dänisch, 78, 198
Dyirbal, 309
Ewe, 162–164

- Fox, 305
 Französisch, 21, 78–80, 118–125, 129, 190, 194, 196, 198, 199, 202, 205, 225
 Gotisch, 276
 Griechisch, 225, 276
 Haitianisch, 125
 Hindi, 98
 Irisch, 225
 Isländisch, 68–69, 78, 146, 147, 198, 260–262, 264, 265, 272, 288–293
 Italienisch, 80, 81, 146–148, 181, 185, 187, 225
 Japanisch, 33, 35, 61, 81, 116, 183–185, 260–265, 267, 272
 Jiddisch, 78
 Kanakuru, 225
 Koreanisch, 33–35, 37, 39, 73, 81, 116, 183
 Lummi, 307
 Lushootseed, 307
 Malagasy, 225
 Malayalam, 286
 Marathi, 312
 Modernes Hebräisch, 81, 128
 Niveau, 5
 Nocte, 309
 Normalem Ulem, 219
 Polnisch, 63, 128, 209, 210, 215, 218
 Portugiesisch, 225, 226
 Quechua, 285
 Russisch, 63, 102, 103, 116, 128, 205
 Schwedisch, 78
 Spanisch, 118, 225
 Squamish, 308
 Tschechisch, 63, 203, 204
 Ungarisch, 98
 Stratum, 212
 Straßenverkehrsordnung, 5
 Strikter Zyklus (Strict Cycle Condition), 163, 165, 322
 Stringenzrelation, 111, 138, 180, 235, 261, 290, 298
 Definition der, 74
 Subhierarchie, 41, 47, 77, 137, 138, 140, 180, 242, 244, 245, 249, 250, 256, 303, 304, 323
 als Kopplung, 250
 alternative Definition, 244
 der minimalen Kettenglieder, 137, 138
 der minimalen nicht-referentiellen Kettenglieder, 139
 von Linearisierungsbeschränkungen, 244
 Subjazenbedingung, 82, 143, 147, 151, 160
 Subjekt-Auxiliar-Inversion, 94, 106, 204
 Subjekt-Insel, 56, 142, 195, 213, 214
 abgeleitete, 164
 Suboptimalität, 243
 sukzessiv-zyklische W-Bewegung, 56, 141, 159, 163, 190, 213, 214, 301
 Super-Optimalität, 320
 Superiorität, 61, 64, 140, 161, 204

 Tableau, 16
 Theorem der harmonischen Begrenzung, 59
 Thetaraster, 277
 Thetarollen
 kontextuelle Definition von, 278
 Topikalisierung
 von W-Phrasen, 158
 vs. W-Bewegung, 148
 Transformationen
 obligatorische vs. optionale, 189

 Unaussprechbarkeit (ineffability), 83
 universale Basis, 311
 unmarkierter Fall, 57
 Unmarkiertheit von Sätzen, 229

 v, 268
 V-I-Bewegungsparameter, 21, 78
 und Komplementierertilgung, 198
 V/2-Nebensätze, 117
 Verletzung
 fatale, 16
 Zahl vs. Qualität, 17
 Vermeide Pronomina, 179, 310
 versteckte Optimalität, 178
 Voranstellung negativer Operatoren, 106
 Vorfeld-es, 48
 VP-Ellipse, 50
 VSO-Abfolge, 226

 W-Bewegung, 31, 33, 61, 82, 95, 111, 127, 131, 136, 139, 148, 164, 187, 190, 194, 196, 202, 205, 271, 321
 auf LF, 56, 159, 162, 213
 von Subjekten, 101
 W-Bewegungsparameter, 33, 104, 132, 205
 W-Expletiv, 55, 98
 W-Infinitive, 182
 W-Insel, 141, 147
 W-W-Konstruktion, 54, 190, 195, 213
 Wackernagelposition, 43, 44, 65, 150

- Wettbewerb, 9, 156, 192
- Wettbewerbssyntax
 - Struktur einer, 156
 - und Optionalität, 192
- Wortstellung, 66, 148, 155, 217, 225–257,
271, 280, 316

- Zwischeninventar, 311, 315, 316