

Zur Verarbeitung koordinierter Strukturen

D i s s e r t a t i o n

zur

Erlangung des akademischen Grades

Doktor der Philosophie

in der Philosophischen Fakultät

der Eberhard Karls Universität Tübingen

vorgelegt von

Ilona Steiner

aus

Reutlingen

2022

Gedruckt mit Genehmigung der Philosophischen Fakultät
der Eberhard Karls Universität Tübingen.

Dekan: Prof. Dr. Jürgen Leonhardt

Hauptberichterstatterin: Prof. Dr. Veronika Ehrich

Mitberichterstatter: Prof. Dr. Hanspeter A. Mallot
Prof. Dr. Ingo Reich
PD Dr. Sam Featherston

Tag der mündlichen Prüfung: 25.06.2010

Universitätsbibliothek Tübingen: TOBIAS-lib

Danksagung

Allen voran möchte ich mich bei Prof. Dr. Veronika Ehrich für ihre Unterstützung und Betreuung meiner Doktorarbeit bedanken. Insbesondere danke ich ihr für den wissenschaftlichen Freiraum, den sie mir gelassen hat.

Weiterhin danke ich meinem Zweitbetreuer Prof. Dr. Hanspeter Mallot für seine Unterstützung während meines Promotionsstipendiums im Graduiertenkolleg *Kognitive Neurobiologie*. Als Sprecher des interdisziplinären Graduiertenkollegs schuf er ein Umfeld, das mir einen fruchtbaren Austausch mit anderen Wissenschaftsdisziplinen ermöglichte.

Herzlichen Dank an PD Dr. Oliver Bott, der mir bei der statistischen Auswertung meiner Daten hilfreich zur Seite stand. Desweiteren danke ich Dr. habil. Stephan Kepser, der während meiner Doktorarbeit ein wichtiger Gesprächspartner für mich war. Vor allem bei technischen Problemen konnte ich mich stets auf seine Hilfe verlassen.

Schließlich danke ich allen, die zum Gelingen dieser Arbeit beigetragen haben!

Bei der vorliegenden Arbeit handelt es sich um die überarbeitete Fassung meiner Dissertation. Sie basiert auf dem Forschungsstand von 2009.

Inhaltsverzeichnis

1	Einleitung	1
2	Koordinierte Strukturen	7
2.1	Koordinationstypen	7
2.1.1	Symmetrische Koordination	7
2.1.2	Asymmetrische Koordination	10
2.2	Symmetrische und asymmetrische Eigenschaften	12
2.2.1	Kategoriale (A)Symmetrie	12
2.2.2	Semantische (A)Symmetrie	14
2.2.3	(A)Symmetrische Substitution	14
2.2.4	(A)Symmetrische Extraktion	18
2.2.5	(A)Symmetrische Kongruenz	20
2.3	Syntaktische Analyse der Koordination	22
2.3.1	Der Reduktionsansatz	22
2.3.2	Die phrasale Analyse	26
2.3.3	Mehrdimensionale Modelle	31
2.4	Zusammenfassung und Motivation	37
3	Die Verarbeitung koordinierter Strukturen	41
3.1	Grundlegendes zur syntaktischen Verarbeitung	42
3.1.1	Lokale vs. globale Ambiguitäten	43
3.1.2	Der Parsingprozess	44
3.1.3	Präferenzen bei Ambiguitäten	45

3.2	Parallelitätseffekte bei Koordination	47
3.2.1	Parallelitätseffekte bei der Interpretation	47
3.2.2	Parallel-Struktur-Effekte auf sententialer Ebene	50
3.2.3	Parallel-Struktur-Effekte auf phrasaler Ebene	53
3.3	Präferenzen bei Ambiguitäten	59
3.3.1	Lokale Ambiguitäten	59
3.3.2	Globale Ambiguitäten	63
3.4	Das Copy- α -Modell	66
3.4.1	Komplexitätseffekt bei Gapping	67
3.4.2	Fehlende Komplexitätseffekte bei ‘ <i>either or</i> ’	68
3.4.3	Probleme des Copy- α -Modells	70
3.5	Zusammenfassung und Ziele der eigenen Untersuchungen	71
4	Das Iterationsmodell	73
4.1	Das Modell	73
4.2	Modellierung der Verarbeitungseffekte	78
4.2.1	Parallel-Struktur-Effekt auf phrasaler Ebene	78
4.2.2	Parallel-Struktur-Effekt auf sententialer Ebene	85
4.2.3	Präferenzen bei Ambiguitäten	90
4.2.4	Fehlende Komplexitätseffekte bei ‘ <i>either or</i> ’	95
4.2.5	Parallelitätseffekte bei der Interpretation	97
4.3	Sprachverständnis bei agrammatischer Aphasie	100
4.3.1	Die agrammatische Sprache	100
4.3.2	Agrammatische Verarbeitung im Iterationsmodell	104
4.4	Zusammenfassung und Diskussion	107
5	Das Verhältnis von Parsingpräferenzen und Korpusfrequenzen	113
5.1	Einleitung	114
5.2	Erstellung des Koordinations-Datensatzes	118
5.3	Cross-kategoriale Koordination	120

5.3.1	Korpusanalyse I: Cross-kategoriale vs. gleiche syntaktische Kategorien . . .	121
5.4	Parallel-Struktur-Effekt im Phrasenaufbau	122
5.4.1	Korpusanalyse II: Strukturelle Ähnlichkeit der Konjunkte	122
5.4.2	Verteilung der parallelen Vorkommen	127
5.5	Präferenzen bei lokalen Ambiguitäten	134
5.5.1	Korpusanalyse III: NP- vs. S-Koordination	135
5.6	Methodische Überlegungen	136
5.7	Zusammenfassung und Diskussion	137
6	Simulationen mit künstlichen neuronalen Netzen	141
6.1	Künstliche neuronale Netze	142
6.2	Lernen in neuronalen Netzen	144
6.2.1	Die Hebbsche Lernregel	145
6.2.2	Überwachtes vs. unbeaufsichtigtes Lernen	148
6.3	Simulationen zur syntaktischen Verarbeitung	154
6.3.1	Elman-Netzwerk	155
6.3.2	Erste Simulationen (Elman, 1990, 1991)	160
6.3.3	Vergleich mit probabilistischen Verfahren (Towsey et al., 1998)	168
6.3.4	Vergleich mit menschlicher Performanz (Christiansen und Chater, 1999, 2001)	173
6.4	Eigene Simulationen zur Verarbeitung koordinierter Strukturen . .	182
6.4.1	Trainingsdaten	183
6.4.2	Lernaufgabe und Kodierung der Trainingsmuster	188
6.4.3	Simulationen	194
6.4.4	Evaluation des Netzes	197
6.4.5	Performanz des Netzes	207
6.5	Zusammenfassung und Diskussion	225

7	Asymmetrische Kongruenz: Ein Verarbeitungseffekt?	229
7.1	Einleitung	230
7.2	Korpusuntersuchungen	234
7.3	Experimentelle Studien	239
7.3.1	Testmaterial für Experiment 1 und 2	240
7.3.2	Experiment 1: Inkrementelle Grammatikalitätsurteile . . .	244
7.3.3	Experiment 2: Lesezeitexperiment	248
7.4	Zusammenfassung und Diskussion	255
8	Allgemeine Diskussion	259
	Literaturverzeichnis	269

Kapitel 1

Einleitung

In dieser Arbeit soll die syntaktische Verarbeitung koordinierter Strukturen untersucht werden. Koordinierte Strukturen, wie in (1), sind Konstruktionen, in denen zwei oder mehrere Ausdrücke miteinander verknüpft werden. Ich betrachte dabei ausschließlich Strukturen, in denen die Verknüpfung durch eine koordinierende Konjunktion, wie *und*, *oder*, etc., erfolgt.

(1) a. Kardinal Ratzinger wird heute 78 und nächste Woche vielleicht Papst.

(Quelle: Radiomoderator SWR3, 16.04.2005)

b. Zu Risiken und Nebenwirkungen lesen Sie die Packungsbeilage und fragen Sie Ihren Arzt oder Apotheker.

(Quelle: Werbung der Pharmaindustrie)

Der Prozess der Koordination ist nach Chomsky (1957) einer der produktivsten Mechanismen, um neue Sätze einer Sprache zu bilden (vgl. hierzu das folgende Zitat). Es handelt sich dabei um einen rekursiven Mechanismus, d.h. die Konjunkte können, wie in Beispiel (1b), rekursiv ineinander eingebettet sein.

”One of the most productive processes for forming new sentences is the process of conjunction. If we have two sentences $Z + X + W$ and $Z + Y + W$, and if X and Y are actually constituents of these sentences, then we can generally form a new sentence $Z - X + Y - W \dots$ ”

(Chomsky, 1957)

Da koordinierte Strukturen ein fundamentaler und hochfrequenter Bestandteil unserer Sprache sind, liegt die Vermutung nahe, dass diese Strukturen während des Sprachverstehens leicht zu verarbeiten sind. In diesem Zusammenhang stellt sich die Frage, wie Sprachverstehensprozesse empirisch untersucht werden können. Dabei ist das menschliche Sprachsystem mit dem Problem konfrontiert, dass eine kohärente phonologische, syntaktische und semantische Repräsentation für einen Satz aufgebaut werden muss, obwohl die einzelnen Elemente des Satzes nicht zeitgleich präsentiert werden. Dieses Problem wurde von Hagoort (2006) als das *temporale Binding-Problem der Sprache* bezeichnet. Der Begriff 'Binding' bezieht sich in diesem Kontext nicht auf ein Konzept aus einer linguistischen Theorie, sondern auf ein kognitives Problem, das ursprünglich aus dem Bereich der visuellen Wahrnehmung stammt. Dieses ursprünglich räumliche Problem wurde von Hagoort auf die Sprache und somit auf die zeitliche Dimension übertragen.

“How is information that is incrementally retrieved from the mental lexicon unified into a coherent overall interpretation of a multi-word utterance? Most likely, unification needs to take place at the conceptual, syntactic and phonological levels, as well as between these levels.”

(Hagoort, 2006)

Wie aus obigem Zitat deutlich wird, erstreckt sich das Binding-Problem der Sprache über zwei zeitliche Dimensionen. Auf der einen Seite müssen die einzelnen Elemente eines Satzes strukturiert werden, obwohl sie nicht zeitgleich präsentiert werden. Das betrifft sowohl die phonologische als auch die syntaktische und semantische bzw. konzeptuelle Ebene. Auf der anderen Seite müssen die unterschiedlichen Informationen (phonologisch, syntaktisch, semantisch) für jedes Element eines Satzes miteinander verknüpft werden. Letzteres lässt sich beispielsweise gut mittels EEG-Studien ('Elektroenzephalographie') untersuchen, da es für jede Informations-Ebene definierte messbare ERP-Komponenten gibt (ERP = ereigniskorrelierte Potentiale). Entsprechende Effekte ergeben sich bei der Präsentation von Sätzen, die phonologische, syntaktische oder semantische Regelverletzungen enthalten oder bei komplexen Sätzen, die Reanalyseprozesse erfordern (z.B.

Friederici et al., 1993; Friederici, 2002; Hagoort, 2003; Kaan et al., 2000). Auf diese Weise kann der zeitliche Verlauf und das Zusammenspiel von phonologischer, syntaktischer und semantischer Information untersucht werden.

Zur Untersuchung des inkrementellen Strukturierungsprozesses innerhalb einer, z.B. der syntaktischen, Ebene haben sich insbesondere Lesezeit-Studien etabliert. Lesezeiten können entweder über die Methode des selbstgesteuerten Lesens ('Self-paced Reading') erhoben werden oder durch die Messung der Augenbewegungen beim Lesen ('Eye Tracking'). Die inkrementell erhobenen Lesezeiten für die einzelnen Segmente eines Satzes ergeben Aufschluss über die Komplexität der zugrundeliegenden Prozesse und der zugrundeliegenden Strukturen (siehe Kapitel 3.1). Insbesondere wenn Satzkonstruktionen im Mittelpunkt stehen, die grammatisch völlig unauffällig sind, d.h. weder Verletzungen enthalten noch zwingend Reanalyseprozesse erfordern, bietet sich methodisch die Erhebung von Lesezeiten an.

In dieser Arbeit soll der inkrementelle Strukturierungsprozess innerhalb der syntaktischen Ebene untersucht werden. Dabei muss das menschliche Sprachsystem für einen Satz eine syntaktische Struktur nach den Regeln der Grammatik aufbauen (Parsingprozess). Die Struktur eines Satzes bildet die Grundlage, um die Bedeutung eines Satzes ermitteln zu können. Im Rahmen von syntaktischen Verarbeitungsmodellen wird beschrieben, mit welchen Mechanismen und nach welchen Prinzipien eine solche Struktur inkrementell aufgebaut wird. Diese Modelle machen im Allgemeinen Vorhersagen für Lesezeiten.

Die Untersuchung koordinierter Strukturen stellt dabei in verschiedener Hinsicht eine Herausforderung dar. Die erste Schwierigkeit besteht bereits darin, dass keinerlei Konsens darüber besteht, wie die syntaktische Struktur von Koordination beschaffen ist. Es existiert eine Vielzahl unterschiedlicher Vorschläge zu ihrer syntaktischen Analyse (siehe Kapitel 2). Aus diesem Grunde wurden koordinierte Strukturen auch lange Zeit bei der Untersuchung syntaktischer Verarbeitungsprozesse ausgeklammert. Bestehende Verarbeitungsmodelle sind nur unzureichend in der Lage, die beobachteten Effekte und Präferenzen bei der Verarbeitung koordinierter Strukturen zu beschreiben (siehe Kapitel 3).

Diese Problematik zeigt sich auch bei der automatischen Sprachverarbeitung. Koordinierte Strukturen stellen für maschinelle Parsingsysteme eine der größten Fehlerquellen dar (siehe z.B. Argamon et al., 1998; Collins, 1999). Die Ursache hierfür könnte darauf zurückzuführen sein, dass entweder ungeeignete syntaktische Strukturen zugrundegelegt oder ungeeignete Parsingmechanismen verwendet wurden. Dabei ist nicht geklärt, inwiefern allgemeine Parsingmechanismen überhaupt auf koordinierte Strukturen angewandt werden können oder ob koordinationspezifische Mechanismen verwendet werden müssen. Koordinierte Strukturen stellen somit eine große Herausforderung sowohl für syntaktische Theorien als auch für syntaktische Verarbeitungsmodelle dar.

Das Ziel der vorliegenden Arbeit ist es, die syntaktischen Prozesse bei der Verarbeitung koordinierter Strukturen unter verschiedenen Gesichtspunkten zu untersuchen und auf diese Weise zusätzliche Evidenz für deren syntaktische Analyse zu erhalten. Die Arbeit ist wie folgt strukturiert: In *Kapitel 2* wird der Phänomenbereich der Koordination vorgestellt und die unterschiedlichen Ansätze zur syntaktischen Analyse dieser Konstruktionen beschrieben. In *Kapitel 3* werden Verarbeitungseffekte und Präferenzen aus der Literatur präsentiert, die beim Sprachverstehen koordinierter Strukturen im Englischen beobachtet wurden. Es wird aufgezeigt, dass sich diese Effekte nicht oder nur eingeschränkt aus bestehenden Parsingmechanismen und -modellen ableiten lassen. In *Kapitel 4* stelle ich ein eigenes Modell zur syntaktischen Verarbeitung koordinierter Strukturen vor, mit dem Ziel, die in Kapitel 3 beschriebenen Effekte zu erklären. Es wird außerdem geprüft, ob sich das Modell auch auf die Verarbeitung koordinierter Strukturen bei agrammatischer Aphasie übertragen lässt. Die Implikationen für syntaktische Theorien werden jeweils aufgezeigt. In *Kapitel 5* wird untersucht, ob die in Kapitel 3 beschriebenen Effekte auch in der Sprachproduktion in Form von Korpushäufigkeiten auftreten. Hierfür werden englische spontansprachliche Korpusdaten statistisch ausgewertet, die außerdem als Trainingsdaten für die Simulationen in Kapitel 6 verwendet werden. Die Ergebnisse erlauben Rückschlüsse auf das Verhältnis von Sprachverstehens- und Sprachproduktions-Mechanismen. In *Kapitel 6* wird die menschliche Satzverarbeitung im Rahmen eines kognitiven Systems simuliert. Hierfür werden künstliche neuronale Netze trainiert, selbständig grammatische

Regularitäten und Parsingmechanismen aus spontansprachlichen Korpusdaten zu lernen. Nach abgeschlossener Trainingsphase wird das Verhalten des Netzes bei der Verarbeitung koordinierter Strukturen überprüft und mit der menschlichen Leseverarbeitung verglichen. In *Kapitel 7* wird das Phänomen der asymmetrischen Kongruenz bei koordinierten Subjekten untersucht. Dabei soll die Frage beleuchtet werden, ob die asymmetrische Kongruenz einen Verarbeitungseffekt darstellt und möglicherweise außerhalb der Syntax erklärt werden kann. Hierzu werden eigene Experimente und Korpusuntersuchungen zum Deutschen durchgeführt. Schließlich werden in *Kapitel 8* die Ergebnisse dieser Arbeit diskutiert.

Kapitel 2

Koordinierte Strukturen

In diesem Kapitel werden koordinierte Strukturen und Ansätze zu ihrer syntaktischen Beschreibung vorgestellt. Das Kapitel ist wie folgt strukturiert: In Abschnitt 2.1 werden die verschiedenen Koordinationstypen beschrieben und in Abschnitt 2.2 ihre symmetrischen und asymmetrischen Eigenschaften. Die unterschiedlichen Ansätze zur syntaktischen Beschreibung dieser Konstruktionen stelle ich in Abschnitt 2.3 vor. Schließlich wird in Abschnitt 2.4 motiviert, warum koordinierte Strukturen in der vorliegenden Arbeit aus der Perspektive der Verarbeitung beleuchtet werden.

2.1 Koordinationstypen

Koordinierte Strukturen werden klassischerweise in zwei Bereiche unterteilt, in “Symmetrische Koordination” und “Asymmetrische Koordination”. Die Begriffe Symmetrie und Asymmetrie können unter vielerlei Gesichtspunkten betrachtet werden. Dies wird im Verlauf dieses Kapitels noch präzisiert werden.

2.1.1 Symmetrische Koordination

Koordinierte Strukturen im Bereich der “Symmetrischen Koordination” umfassen den Phänomenbereich der Konstituenten-Koordination (Constituent Coordination) sowie der Nicht-Konstituenten-Koordination (Non-Constituent Coordination) (einen Überblick hierzu bietet Wesche, 1995). Die

Konstituenten-Koordination ist dadurch charakterisiert, dass es sich bei den Konjunkten um vollständige Konstituenten handelt. Dies wird an den folgenden Beispielen aus dem Deutschen (1a–e) und Englischen (2a–f) illustriert.¹ Die deutschen Beispielsätze in (1a,c,d) sind Korpusbelege aus Cosmas (*Mannheimer Morgen*), die Sätze in (1b,e) sind aus *Spiegel Online*. Die englischen Beispiele in (2) stammen aus Wilder (1994).

- (1) a. Wir wünschen Ihnen allen [_{AP} frohe] und [_{AP} erholsame] Urlaubstage.
- b. Der Täter hinterließ [_{NP} DNA-Spuren] und [_{NP} die Tatwaffe].
- c. Viele Zuschauer hatten schon [_{PP} an den Fenstern] und [_{PP} auf den Bürgersteigen] gewartet.
- d. Das große Schulfest der Haupt- und Realschule [_{C'} beginnt um 11 Uhr] und [_{C'} endet gegen 17 Uhr].
- e. [_{CP} Die Videospielemesse E3 ist in vollem Gange], und [_{CP} die Hersteller übertrumpfen sich mit Ankündigungen].

In (1b) werden beispielsweise zwei vollständige Nominalphrasen² koordiniert, in (1c) sind es zwei vollständige Präpositionalphrasen, usw. Die Koordination von Sätzen ist darunter als Spezialfall anzusehen (siehe (1e) und (2f)). Die Konstituenten-Koordination wird auch häufig als *phrasale Koordination* bezeichnet, obwohl es sich bei den Konjunkten nicht immer um vollständige Phrasen handeln muss. (1d) kann beispielsweise als Koordination der Kategorie *C'* aufgefasst werden und (2e) als Koordination von *I'*.

¹ Da in der vorliegenden Arbeit Untersuchungen zum Englischen und Deutschen durchgeführt werden, präsentiere ich Beispiele aus beiden Sprachen.

² Mir ist bewusst, dass mittlerweile meist nicht mehr von Nominalphrasen (NPs) sondern eher von Determinatorenphrasen (DPs) ausgegangen wird (siehe Abney, 1987). Dennoch verwende ich die traditionelle NP-Analyse, um eine einheitliche Notation im Verlaufe dieser Arbeit zu gewährleisten.

- (2) a. [_{NP} John] and [_{NP} Mary] are coming
- b. [_{PP} In London] and [_{PP} in Berlin], it is still cold
- c. the [_{AP} red] and [_{AP} blue] flag
- d. Mary has [_{VP} left] and [_{VP} gone to England]
- e. John [_{I'} has left] and [_{I'} will return]
- f. I don't know if [_{IP} Mary left] and [_{IP} Peter returned]

Die **Nicht-Konstituenten-Koordination** ist dadurch charakterisiert, dass die Konjunkte keine Konstituenten bilden, sondern in einer sogenannten reduzierten Form vorkommen. Man spricht hier auch von Koordinationsellipsen (zum Reduktionsansatz siehe Abschnitt 2.3.1). Je nach Art und Position des ausgelassenen Materials wird unterschieden zwischen *Gapping* (Ross, 1967, 1970) und *Right Node Raising* (RNR, vgl. Postal, 1974). Wird das finite Verb (und evtl. weitere Konstituenten) im nicht-initialen Konjunkt unter Identität mit entsprechenden Ausdrücken in den anderen Konjunkten ausgelassen, spricht man von einer Gapping-Konstruktion (siehe (3)). Bei dieser Ellipseform bilden die Rest-Phrasen ('remnants') Konstituenten.

- (3) a. [Hans trinkt Bier] und [Maria ~~trinkt~~ Wein].
- b. [John gave a book to Sarah] and [Peter ~~gave a book~~ to Sue].

Bei RNR-Konstruktionen, wie in (4a–c), sind ein oder mehrere Elemente am rechten Rand des ersten Konjunks ausgelassen. Während die Auslassung im Englischen auf eine Hauptkonstituente beschränkt ist (4a), können im Deutschen mehrere Elemente ausgelassen werden, die nicht unbedingt eine Konstituente bilden müssen ((4b), siehe z.B. Klein, 1993). RNR-Konstruktionen werden auch als *Linkstilgung* bezeichnet, weil hier offenbar Elemente unmittelbar links der Konjunktion ausgelassen werden.

- (4) a. [John bought ~~the book~~] and [Mary read the book].
- b. [Karl soll vier ~~Brote kaufen~~] und [Maria will drei Brote kaufen].
(Klein, 1993)
- c. [Die Billardspieler lieben ~~sie~~] und [die Autofahrer fürchten sie].
(Quelle: Kreuzworträtsel *TV Spielfilm*, Lösungswort: Karambolage)

Es ist auch möglich, Elemente unmittelbar rechts der Konjunktion auszulassen (*Rechtstilgung*), wie in (5). Diese Konstruktionen werden jedoch häufig nicht als Ellipseform betrachtet, sondern als einen Fall von phrasaler Koordination (wie u.a. in Reich (2009) argumentiert wird, vgl. auch Höhle (1991) und Hartmann (2002)).

- (5) [Petra telefonierte mit ihrer Freundin] und [~~Petra~~ vergaß die Zeit].

Nach diesem kurzen Überblick über den Phänomenbereich der “Symmetrischen Koordination” skizziere ich im Folgenden den Phänomenkomplex der “Asymmetrischen Koordination”.

2.1.2 Asymmetrische Koordination

Der Phänomenkomplex der “Asymmetrischen Koordination”, der bereits in den grundlegenden Arbeiten von Höhle (1983, 1990) und Wunderlich (1988) zum Deutschen beschrieben wurde, umfasst satzwertige Koordinationen. Die Asymmetrie kann in erster Annäherung an der Verbstellung der beteiligten Konjunkte festgemacht werden. Beispielsweise wird in (6a) offenbar ein Verb-End-Satz (VE) mit einem Verb-Zweit-Satz (V2) verknüpft. Die symmetrische Variante (VE + VE) ist in (6b) dargestellt.

- (6) a. wenn [[jemand nach Hause kommt] und [da steht der Gerichtsvollzieher vor der Tür]], ...
- b. wenn [[jemand nach Hause kommt] und [da der Gerichtsvollzieher vor der Tür steht]], ...

Die asymmetrische Koordination in (7a) weist neben einer Verbstellungs-Asymmetrie (VE + V1) eine weitere Besonderheit auf, nämlich eine Subjektücke. Das zweite Konjunkt weist kein overt Subject auf, das Subject des ersten Satzes wird als inhaltliches Subject des zweiten Konjunks mitverstanden. Die entsprechende symmetrische Variante (VE + VE) ist in (7b) dargestellt.

- (7) a. wenn jemand [[nach Hause kommt] und [sieht den Gerichtsvollzieher]], ...
- b. wenn jemand [[nach Hause kommt] und [den Gerichtsvollzieher sieht]], ...

Asymmetrische Koordination kann auch in selbständigen Strukturen beobachtet werden. In (8a) wird beispielsweise ein V2-Satz mit einem V1-Satz koordiniert, wobei das zweite Konjunkt eine Subjektücke aufweist. Hierbei handelt es sich in erster Linie um eine Wortstellungs-Asymmetrie. Eine modifizierte Wortstellung des ersten Konjunks (mit dem Subject im Vorfeld anstatt im Mittelfeld), wie in (8b), führt zu einer symmetrischen Koordination.

- (8) a. [In den Wald ging der Jäger] und [fing einen Hasen].
- b. Der Jäger [ging in den Wald] und [fing einen Hasen].

Konstruktionen mit Subjektücke, wie in (7a) und (8a), werden als "Asymmetrische Koordination mit Subjektücke in F-strukturierten, nicht-initialen Konjunkten" (SLFK) bezeichnet (vgl. Höhle, 1990; Reich, 2009). Wie bereits in Höhle (1990) gezeigt wird, können in asymmetrischen Koordinationen ausschließlich Subjekte ausgelassen werden (siehe (9a) und (10a)). Objektücken, wie in (9b) und (10b), sind hingegen nicht möglich.

- (9) a. wenn ich den Hund suche und finde ihn nicht, ...
- b. *wenn ich den Hund suche und finde ich nicht, ...

- (10) a. Vielleicht sucht sie den Hund und findet ihn nicht.
 b. *Vielleicht sucht sie den Hund und findet sie nicht.

Reich (2009) zeigt, dass die Subjektücke in Asymmetrischen Koordinationen nicht auf einen Ellipseprozess zurückgeführt werden kann. Desweiteren sind Ellipseprozesse, wie Linkstilgung und Gapping, in Asymmetrischen Koordinationen generell ausgeschlossen, was diesen Phänomenbereich deutlich von der Symmetrischen Koordination abgrenzt.

2.2 Symmetrische und asymmetrische Eigenschaften

Im letzten Abschnitt wurde der Phänomenbereich der Koordination dargestellt und bereits eine erste grobe Einteilung in Symmetrische und Asymmetrische Koordination vorgenommen. Die verschiedenen Koordinationstypen weisen jedoch darüber hinaus eine Vielzahl an symmetrischen und asymmetrischen Eigenschaften auf, die im Folgenden beschrieben werden.

2.2.1 Kategoriale (A)Symmetrie

Es ist seit langem bekannt, dass koordinierte Strukturen grundlegenden Symmetriebedingungen unterliegen. Nach dem ‘*Coordination of Likes Constraint*’ (Chomsky, 1957; Williams, 1978) dürfen beispielsweise nur gleiche syntaktische Kategorien miteinander verknüpft werden. Daraus folgt, dass z.B. Nominalphrasen nur mit Nominalphrasen koordiniert werden können, Präpositionalphrasen nur mit Präpositionalphrasen, usw. (siehe (11a–c)).

- (11) a. Wir wünschen Ihnen allen [_{AP} frohe] und [_{AP} erholsame] Urlaubstage.
 b. Der Täter hinterließ [_{NP} DNA-Spuren] und [_{NP} die Tatwaffe].
 c. Viele Zuschauer hatten schon [_{PP} an den Fenstern] und [_{PP} auf den Bürgersteigen] gewartet.

Der im letzten Abschnitt skizzierte Phänomenbereich der Symmetrischen Koordination unterliegt im Allgemeinen dieser Bedingung (vgl. die Beispiele zu phrasaler Koordination in (1) und (2)). Bei Koordinationsellipsen ist dies nur der Fall, wenn man tatsächlich von Reduktion ausgeht und die Beispiele in (3) und (4) auf die Koordination von Sätzen zurückführt. Die Verknüpfung von ungleichen Kategorien sollte dementsprechend ungrammatisch sein (vgl. die Kontraste in (12) und (13)).³

- (12) a. The scene [_{PP} of the movie] and [_{PP} of the play] was in Chicago.
 b. *The scene [_{PP} of the movie] and [_{CP} that I wrote] was in Chicago.
 (Chomsky, 1957)

- (13) a. John ate [_{AdvP} quickly] and [_{AdvP} greedily].
 b. *John ate [_{AdvP} quickly] and [_{NP} a grilled cheese sandwich].
 (Schachter, 1977)

Auch wenn intuitiv die Koordination von gleichen syntaktischen Kategorien der Normalfall zu sein scheint, gibt es doch eine Reihe von wohlbekannten Gegenbeispielen, d.h. *cross-kategoriale Koordinationen*, wie in (14a–d) (siehe u.a. Gazdar et al., 1985; Sag et al., 1985; Munn, 2000). Auch der Phänomenbereich der Asymmetrischen Koordination verletzt den ‘*Coordination of Likes Constraint*’.

- (14) a. Pat is either [_{AP} stupid] or [_{NP} a liar]. (Sag et al., 1985)
 b. She walked [_{AdvP} slowly] and [_{PP} with great care].
 (Gazdar et al., 1985)
 c. Bewirb Dich [_{AdvP} allein] oder [_{PP} mit Deinem besten Freund]!
 (Quelle: Werbung für *Marlboro Summer Jobbing*)
 d. Da hat er sich die A8 [_{PP} in Ruhe] und [_{AP} intensiver] angucken können.
 (Quelle: Radiomoderator im Kontext einer Staumeldung, 5.10.2008)

³ Schachter (1977) schlägt eine stärkere Restriktion vor, wonach neben der kategorialen auch die semantische bzw. funktionale Gleichartigkeit gefordert wird (‘*Coordinate Constituent Constraint*’).

2.2.2 Semantische (A)Symmetrie

Alternativ besteht die Möglichkeit, den Symmetriebegriff in die Semantik zu verlagern. So geht Munn (1993, 1999) beispielsweise davon aus, dass es keine Bedingung zur Verknüpfung von gleichen syntaktischen Kategorien gibt. Die Konjunkte müssen lediglich vom gleichen semantischen Typ sein (*'semantic type identity'*). Dabei unterscheidet er Typen wie *'manner'*, *'location'*, *'time'*, etc. Gestützt wird diese These durch Daten wie in (15). In beiden Fällen werden ungleiche syntaktische Kategorien koordiniert. Während jedoch in (15a) die Konjunkte vom selben semantischen Typ sind (*manner + manner*), wird in (15b) eine *manner*- mit einer *goal*-Spezifizierung verknüpft.

- (15) a. John walked [*AdvP* slowly] and [*PP* with great care]. (manner + manner)
 b. #John walked [*AdvP* slowly] and [*PP* to the park]. (manner + goal)

Symmetrie auf semantischer Ebene scheint eine grundlegende Eigenschaft koordinierter Strukturen zu sein (siehe auch Lang (1977) zur gemeinsamen Einordnungsinstanz (*'common integrator'*) und das Symmetrieaxiom in Reich (2009, S. 10)). Dennoch ist es umstritten, ob – wie Munn es vorschlägt – auf einen Symmetriebegriff innerhalb der Syntax vollständig verzichtet werden kann. Im Allgemeinen werden Symmetriebedingungen auf syntaktischer und auf semantischer Ebene angenommen. In der vorliegenden Arbeit konzentriere ich mich auf die syntaktischen Aspekte der Koordination.

2.2.3 (A)Symmetrische Substitution

Auf syntaktischer Ebene ist eine symmetrische – und meines Erachtens sehr grundlegende – Restriktion in Höhle (1990) formuliert, wonach die Koordination durch jedes einzelne Konjunkt ohne Grammatikalitätsverlust ersetzt werden kann (*'External Homogeneity Condition'*, im Folgenden EHC genannt). Mit diesem Substitutions-Kriterium steht nicht mehr die Symmetrie zwischen den Konjunkten im Mittelpunkt, sondern die Symmetrie der Konjunkte zu ihrer syntaktischen Umgebung. Das heißt, nach der EHC müssen alle Konjunkte die Anforderungen der syntaktischen Umgebung erfüllen, und alle Konjunkte haben

somit denselben Stellenwert. Die Symmetrie zwischen den Konjunkten folgt auf indirekte Weise. So sollte nach Höhle (1990) aus der EHC folgen, dass die einzelnen Konjunkte typischerweise derselben syntaktischen Kategorie angehören und in derselben grammatischen Relation zur ihrer syntaktischen Umgebung stehen. Cross-kategoriale Koordination wird jedoch nicht ausgeschlossen.

Entsprechend lässt sich dieser Substitutions-Test in allen (symmetrischen) Beispielen zu phrasaler Koordination in (1) und (2) sowie in allen Beispielen zu cross-kategorialer Koordination in (14) erfolgreich anwenden. Dies ist in (16) illustriert, wobei (14d) in (16a) wiederholt wird.

- (16) a. Da hat er sich die A8 [[*PP* in Ruhe] und [*AP* intensiver]] angucken können.
- b. Da hat er sich die A8 [*PP* in Ruhe] angucken können.
- c. Da hat er sich die A8 [*AP* intensiver] angucken können.

Höhle verwendet das Substitutions-Kriterium zur Abgrenzung von Symmetrischer und Asymmetrischer Koordination, wobei Koordinationsellipsen, wie Gapping oder Right Node Raising, von der Betrachtung explizit ausgenommen werden. Verstellungsasymmetrien und Subjektlücken-Konstruktionen verletzen die EHC und sind daher Asymmetrische Koordinationen (siehe (17)). Die Koordination kann zwar durch das erste Konjunkt (17b), nicht jedoch durch das zweite Konjunkt ersetzt werden (17c). Die Variante in (18a) ist hingegen eine symmetrische Koordination. Hier kann die Koordination durch jedes einzelne Konjunkt ersetzt werden (18b,c).

- (17) a. wenn jemand [[nach Hause kommt] und [sieht den Gerichtsvollzieher]] ...
- b. wenn jemand [nach Hause kommt] ...
- c. *wenn jemand [sieht den Gerichtsvollzieher] ...

- (18) a. wenn jemand [[nach Hause kommt] und [den Gerichtsvollzieher sieht]] ...
 b. wenn jemand [nach Hause kommt] ...
 c. wenn jemand [den Gerichtsvollzieher sieht] ...

Wie die folgenden beiden Beispiele zeigen, lassen sich auch auf phrasaler Ebene asymmetrische Konstruktionen beobachten, die die EHC verletzen (siehe Gazdar et al., 1985; Sag et al., 1985). In den Beispielen (19a–c) erfüllt jeweils nur das erste Konjunkt die Subkategorisierungsbedingungen der syntaktischen Umgebung (*asymmetrische Subkategorisierung*). Entsprechend kann die Koordination nur durch das erste Konjunkt, nicht aber durch das zweite Konjunkt ersetzt werden (siehe (20a–c)).

- (19) a. You can depend on [[*NP* my assistant] and [*CP* that he will be on time]].
 b. Pat was annoyed by [[*NP* the children's noise] and [*CP* that their parents did nothing to stop it]].
 c. We talked about [[*NP* Mr. Colson] and [*CP* that he had worked at the White House]].
- (20) a. *You can depend on [*CP* that my assistant will be on time].
 b. *Pat was annoyed by [*CP* that their parents did nothing to stop the noise].
 c. *We talked about [*CP* that Mr. Colson had worked at the White House].

Die bisher aufgeführten Beispiele zeigen, dass im symmetrischen (Normal-)Fall jedes einzelne Konjunkt die Gesamtkoordination ersetzen kann und somit jedes einzelne Konjunkt die Anforderungen der syntaktischen Umgebung erfüllt. Im asymmetrischen Fall kann die Koordination nur durch das erste Konjunkt ersetzt werden, d.h. nur das erste Konjunkt erfüllt die Anforderungen der syntaktischen Umgebung, das zweite Konjunkt kann manchmal davon abweichen.

Es gibt jedoch auch (grammatische) Beispiele, in denen die Koordination durch keines der Konjunkte ersetzt werden kann, wie z.B. bei koordinierten Subjekten oder bei bloßen Singularen. Die Substitution bei koordinierten Subjekten, wie in (21), ist problematisch hinsichtlich der Subjekt-Verb-Kongruenz. Die Substitution bei bloßen Singularen, wie in (22), ist nicht möglich, da der Determinator nur in der koordinierten Struktur wegfallen kann (zum bloßen Singular in koordinativen Verknüpfungen siehe Heycock und Zamparelli, 2003; Ehrich, 2007).

- (21) a. [[Ein Vater] und [eine Mutter]] stehen mal wieder vor dem Lehrerzimmer.
- b. *[Ein Vater] stehen mal wieder vor dem Lehrerzimmer.
- c. *[Eine Mutter] stehen mal wieder vor dem Lehrerzimmer.
- (22) a. Er schleppte [[Koffer] und [Rucksack]] zum Bahnhof.
- b. *Er schleppte [Koffer] zum Bahnhof.
- c. *Er schleppte [Rucksack] zum Bahnhof.

Obwohl die EHC in den Beispielen (21) und (22) nicht erfüllt ist, würde man diese Beispiele als symmetrisch und nicht als asymmetrisch klassifizieren wollen. In gewisser Hinsicht lässt sich die Symmetrie daran erkennen, dass die Substitution für keines der Konjunkte funktioniert. Ich gehe davon aus, dass *symmetrische Substitution* dann vorliegt, wenn die Gesamtkoordination entweder durch jedes einzelne Konjunkt oder eben durch keines der Konjunkte ersetzt werden kann. *Asymmetrische Substitution* liegt vor, wenn die Koordination nur durch eines der Konjunkte ersetzt werden kann.

Auch wenn in diesem Abschnitt deutlich wurde, dass die Möglichkeit der Substitution kein ausreichendes Kriterium für Grammatikalität darstellt, so ist sie doch ein gutes Werkzeug, um grammatische Strukturen in symmetrische und asymmetrische Koordinationen (hinsichtlich der Symmetrie zur syntaktischen Umgebung) einzuteilen.

2.2.4 (A)Symmetrische Extraktion

Eine zentrale Eigenschaft koordinierter Strukturen wurde in Ross (1967) beobachtet, die er als ‘*Coordinate Structure Constraint*’ (CSC) formuliert hat. Grundlage dieser Beschränkung sind Daten wie in (23). Das Beispiel in (23a) zeigt, dass in einer koordinierten Struktur die Extraktion eines Konjunks ungrammatisch ist, und (23b) zeigt, dass die Extraktion aus einem Konjunkt ungrammatisch ist.

Coordinate Structure Constraint (CSC): “In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.” (Ross, 1967)

- (23) a. *Which surgeon_i did Kim date t_i and a lawyer?
 b. *Which surgeon_i did Kim date friends of t_i and a lawyer?

Ross (1967) beobachtet jedoch weiter, dass zum CSC eine systematische Ausnahme existiert: Extraktion ist dann erlaubt, wenn sie aus allen Konjunkten gleichzeitig erfolgt (*‘Across-the-Board (ATB) Movement’*). Dies ist eine grundlegend symmetrische Eigenschaft koordinierter Strukturen, welche in (24) und (25) illustriert wird.

- (24) a. Which surgeon_i did Kim date friends of t_i and enemies of t_i ?
 (25) a. Welchen Kuchen_i hat Ilse t_i gebacken und Georg t_i verspeist?
 b. Diesen Vorschlag_i hat Franz t_i vorgelegt und Angela t_i abgeschmettert.

Auf der anderen Seite wird beispielsweise in Lakoff (1986) argumentiert, dass es sich beim CSC (/ATB) um eine falsche Generalisierung handelt. Basis dieser Argumentation sind asymmetrische Extraktionen, wie z.B. in (26) (vgl. Höhle, 1983; Goldsmith, 1985; Heycock und Kroch, 1994).

- (26) a. How much_{*i*} can you drink *t_i* and still stay sober?
 b. Das Gepäck_{*i*} ließ er *t_i* fallen und rannte zum Hinterausgang.
 c. Diesem Vorschlag_{*i*} will die Kommission *t_i* folgen und eine Unterkommission einsetzen.

Grundsätzlich ist daher davon auszugehen, dass sowohl symmetrische als auch asymmetrische Extraktionen existieren. Entsprechend wird in Reich (2009) die Möglichkeit von ATB-Bewegung als Unterscheidungskriterium für Symmetrische vs. Asymmetrische Koordination verwendet. Dieses Kriterium ist eingeschränkt auf satzwertige koordinierende Strukturen.

Symmetrische (sententiale) Koordination: “Eine satzwertige koordinierende Struktur heißt dann ‘symmetrisch’, wenn sie (prinzipiell) ATB-Bewegung erlaubt.” (Reich, 2009, S. 5)

Demnach ist Asymmetrische Koordination dadurch charakterisiert, dass ATB-Bewegung grundsätzlich ausgeschlossen ist. Das dies tatsächlich der Fall ist, wird in Reich (2009) gezeigt (vgl. hierzu die Asymmetrische Koordination mit Subjektücke (SL) in (27a)). In Beispiel (27a) ist das Objekt [*die Briefmarken*] ATB-bewegt worden, was zu einem ungrammatischen Resultat führt (siehe auch Höhle, 1983). Weiterhin zeigt Reich (2009), dass asymmetrische Extraktion aus dem ersten Konjunkt hingegen systematisch zulässig ist (siehe Beispiel (27b)).

- (27) a. *Die Briefmarken_{*i*} zeigt Karl dem Onkel *t_i* und bietet *SL t_i* ihm zum Verkauf an.
 b. Die Briefmarken_{*i*} zeigt Karl dem Onkel *t_i* und bietet *SL sie_i* ihm zum Verkauf an.

Reich (2009) schlägt vor, dass die Einteilung in Symmetrische Koordination (nur ATB-Extraktion) und Asymmetrische Koordination (nur asymmetrische Extraktion) – zumindest im Deutschen – um eine dritte Klasse erweitert werden muss. Diese Klasse erlaubt prinzipiell sowohl symmetrische (28a) als auch asymmetrische Extraktion (28b).

- (28) a. Diesen Vorschlag_i will die Kommission [[*t_i* prüfen] und [möglichst bald dem Bundestag *t_i* vorlegen]].
- b. Diesen Vorschlag_i will die Kommission [[*t_i* prüfen] und [möglichst bald dem Bundestag einen Gesetzentwurf vorlegen]].

Nach obigem Symmetrie-Kriterium handelt es sich bei den Konstruktionen in (28) um Symmetrische Koordinationen, da sie prinzipiell ATB-Bewegung zulassen. Daraus ergibt sich nach Reich (2009) eine Trichotomie mit i. Symmetrischer Koordination (Typ I), die ausschließlich ATB-Bewegung zulässt, ii. Symmetrischer Koordination (Typ II), die sowohl ATB-Bewegung als auch asymmetrische Extraktion zulässt und iii. Asymmetrischer Koordination, die ausschließlich asymmetrische Extraktion erlaubt.

2.2.5 (A)Symmetrische Kongruenz

Besteht das Subjekt aus mehreren Konjunkten, werden für die Subjekt-Verb-Kongruenz im Allgemeinen alle Konjunkte, d.h. die Gesamtkoordination, herangezogen (symmetrische Kongruenz wie in (29) und (30)). Das ist unabhängig davon, ob das Subjekt vor oder nach dem finiten Verb steht.

- (29) a. [A man] and [a woman] *are* at the front door.
- b. There *are* [a man] and [a woman] at the front door.
- (30) a. [Ein Mann] und [eine Frau] *stehen* dort.
- b. Dort *stehen* [ein Mann] und [eine Frau].

Es gibt jedoch auch Beispiele, in denen nur ein Konjunkt für die Kongruenz relevant ist (asymmetrische Kongruenz wie in (31)). Die asymmetrische Kongruenz ist allerdings nur möglich, wenn das finite Verb dem Subjekt vorangeht. Befindet sich das Subjekt in präverbaler Position, wie in (32), ist nur symmetrische Kongruenz möglich. Wie die Beispiele zeigen, scheint

die asymmetrische Kongruenz außerdem optional zu sein. Befindet sich das Subjekt in postverbaler Position, ist sowohl symmetrische als auch asymmetrische Kongruenz möglich. Dieses Muster wurde auch für andere Sprachen dokumentiert, u.a. für das marokkanische und libanesische Arabisch (Aoun et al., 1994), für das Russische (Babyonyshev, 1997), für Swahili (Marten, 2005) und für das Polnische (Citko, 2004).

(31) a. There *is* [a man] and [a woman] at the front door.

b. Dort *steht* [ein Mann] und [eine Frau].

(32) a. *[A man] and [a woman] *is* at the front door.

b. *[Ein Mann] und [eine Frau] *steht* dort.

Es wird im Allgemeinen davon ausgegangen, dass in den asymmetrischen Fällen das erste (und nicht etwa das zweite) Konjunkt für die Herstellung von Kongruenz relevant ist. Diese Auffassung wird durch Daten wie in (33) gestützt.

(33) a. There *is/are* [a man] and [three children] at the front door.

b. There **is/are* [three children] and [a man] at the front door.

In diesem Abschnitt wurde gezeigt, dass koordinierte Strukturen grundlegenden Symmetriebedingungen unterschiedlicher Art unterliegen. Es wurde aufgezeigt, dass auf der anderen Seite jedoch auch jeweils die entsprechenden Asymmetrien zu beobachten sind. Auch wenn intuitiv die jeweilige symmetrische Variante der Normalfall zu sein scheint, d.h. vermutlich häufiger verwendet wird, so steht doch eine quantitative Untersuchung bislang aus. Eine große Auswahl an Vorschlägen zur syntaktischen Analyse von Koordination reflektiert dieses Spannungsfeld zwischen Symmetrie und Asymmetrie. Abhängig davon, ob bei der Modellbildung den symmetrischen oder eher den asymmetrischen Eigenschaften den Vorzug gegeben wurde, führte das in der Vergangenheit zu symmetrischen bzw. zu asymmetrischen Modellen. Im nächsten Abschnitt wird ein Überblick über die verschiedenen syntaktischen Ansätze gegeben.

2.3 Syntaktische Analyse der Koordination

Bei der syntaktischen Analyse von Koordination können drei grundlegende Beschreibungsansätze unterschieden werden: der Reduktionsansatz, die phrasale Analyse sowie mehrdimensionale Modelle. Diese Ansätze werden im Folgenden kurz beschrieben.

2.3.1 Der Reduktionsansatz

Der Reduktionsansatz basiert auf der Beobachtung, dass offenbar Elemente unmittelbar links und unmittelbar rechts der Konjunktion ausgelassen werden können, bei Identität mit den entsprechenden Ausdrücken in den anderen Konjunkten (siehe u.a. Chomsky, 1957; Ross, 1967; Klein, 1993; van Oirsouw, 1987). In den Beispielen (34a,b) werden Elemente links der Konjunktion ausgelassen (*Linkstilgung*), in (35a,b) sind es Elemente rechts der Konjunktion (*Rechtstilgung*). Weiterhin kann im zweiten Konjunkt das finite Verb (optional mit weiteren Ausdrücken) ausgelassen werden (*Gapping*, wie in (35c), siehe auch Abschnitt 2.1.1 zu Koordinationsellipsen).

- (34) a. [Die Täter versteckten ~~die Tatwaffe~~] und [die Polizisten suchten die Tatwaffe].
 b. [Die Täter ~~suchten die Tatwaffe~~] und [die Komplizen suchten die Tatwaffe].
- (35) a. [Der Täter versteckte die Tatwaffe] und [~~der Täter~~ hinterließ DNA-Spuren].
 b. [Der Täter hinterließ DNA-Spuren] und [~~der Täter hinterließ~~ die Tatwaffe].
 c. [Der Täter hinterließ die Tatwaffe] und [der Komplize ~~hinterließ~~ DNA-Spuren].

In reduktionistischen Ansätzen wird nun versucht, alle koordinierten Strukturen auf die Koordination von vollständigen Sätzen zurückzuführen. Die Koordination von nicht-sententialen Konjunkten wird über koordinationspezifische Tilgungsprozesse ('conjunction reduction') abgeleitet. Der Reduktionssansatz umfasst traditionell die Tilgung von syntaktischen Strukturen. In neueren Arbeiten versteht man unter Reduktion die Auslassung von Elementen auf der Ebene der phonologischen Form (Wilder, 1994; Aoun et al., 1994; Schwarz, 1999).

Für den Reduktionsansatz spricht insbesondere die Nicht-Konstituenten-Koordination, d.h. Konstruktionen wie Right Node Raising (RNR) (siehe (34a)) und Gapping (35c), aber auch Strukturen, wie in (36), in denen die Konjunkte aus mehreren VP-Töchtern bestehen. Diese Konstruktionen können nicht über Phrasenstrukturregeln abgeleitet werden, da die Konjunkte keine Konstituenten bilden (vgl. u.a. Klein, 1993; Wilder, 1994). Es scheint daher unumgänglich, koordinationspezifische Tilgungsprozesse anzunehmen. Unter Ökonomie-Gesichtspunkten wäre es wünschenswert, sowohl Konstituenten- als auch Nicht-Konstituenten-Koordination auf einheitliche Weise behandeln zu können.

(36) John drinks [beer at lunchtime] and [wine in the evening].

Für den Reduktionsansatz sprechen außerdem Asymmetrien wie cross-kategoriale Koordination und asymmetrische Kongruenz bei koordinierten Subjekten. Wird aus unabhängigen Gründen angenommen, dass nur gleiche syntaktische Kategorien miteinander verknüpft werden können, lässt sich cross-kategoriale Koordination, wie in (37), dennoch problemlos aus dem Reduktionsansatz ableiten, da die ungleichen Kategorien in zwei strukturell voneinander unabhängigen Sätzen eingebettet sind (vgl. Wilder, 1994). Die asymmetrische Kongruenz bei koordinierten Subjekten, wie in (38), ergibt sich ebenfalls aus einem Reduktionsansatz und wird als Argument für eine sententiale Analyse der Koordination betrachtet (siehe Aoun et al., 1994).⁴

⁴ Asymmetrische Kongruenz bei koordinierten Subjekten lässt sich nur dann aus einem Reduktionsansatz ableiten, wenn zwei singularische NPs verknüpft werden. Wird eine Plural-NP mit einer Singular-NP koordiniert (*zwei Väter und eine Mutter*), kann dies nicht auf eine sententiale Analyse zurückgeführt werden (vgl. Munn, 2000).

- (37) a. [John walked slowly] and [~~John walked~~ with great care].
 b. [Pat is a Republican] and [~~Pat is~~ proud of it].

- (38) a. Vor dem Lehrerzimmer steht ein Vater und eine Mutter.
 b. [Vor dem Lehrerzimmer steht ein Vater] und
 [~~vor dem Lehrerzimmer steht~~ eine Mutter].

Allerdings ergibt sich nun das umgekehrte Problem mit der symmetrischen Kongruenz. Im symmetrischen Normalfall führt ein koordiniertes Subjekt, wie in (39a), zu einer Pluralmarkierung am Verb. Diese Form lässt sich aus einem Reduktionsansatz jedoch nicht ableiten (siehe (39b)).

- (39) a. Ein Vater und eine Mutter stehen vor dem Lehrerzimmer.
 b. *[Ein Vater ~~stehen vor dem Lehrerzimmer~~] und [eine Mutter stehen vor dem Lehrerzimmer].

Auch bloße Singulare in koordinierten Strukturen, wie in (40), und andere Konstruktionen, die das Substitutions-Kriterium nicht erfüllen (siehe Abschnitt 2.2), sind für den Reduktionsansatz problematisch und sprechen für eine phrasale Analyse (zur Analyse des bloßen Singulars siehe Heycock und Zamparelli, 2003; Ehrich, 2007).

- (40) a. Er schleppte Koffer und Rucksack zum Bahnhof
 b. *[Er schleppte Koffer ~~zum Bahnhof~~] und [~~er schleppte~~ Rucksack zum Bahnhof].

Weitere Argumente gegen den Reduktionsansatz sind semantischer Natur. So macht der Reduktionsansatz beispielsweise problematische Vorhersagen in Bezug auf kollektive und reziproke Prädikate (bzw. kollektive Adverbiale), wie in (41). Diese Prädikate erfordern ein pluralisches Subjekt und führen zu einer kollektiven Lesart. Selbst wenn man von der Subjekt-Verb-Kongruenz absieht, lassen sich diese Konstruktionen offensichtlich nicht auf die Verknüpfung von Sätzen zurückführen (siehe z.B. Klein, 1993; van Oirsouw, 1993).

- (41) a. John und Mary sind ein Ehepaar.
- b. John und Mary treffen sich in der Stadt.
- c. John und Bill tragen gemeinsam eine schwere Kiste.

Aufgrund der oben aufgeführten Probleme werden oft abgeschwächte reduktionistische Formen angenommen, in denen neben sententialen auch NP-Koordinationen zugelassen sind. Dies führt allerdings dazu, dass sich die asymmetrische Kongruenz nicht mehr (ohne Weiteres) über Reduktion erklären lässt.

Ein weiteres Problem für den Reduktionsansatz sind Konstruktionen mit quantifizierenden Subjekten, wie in (42a). Ließe sich dieses Beispiel auf sententiale Koordination, wie in (42b), zurückführen, dürften sich keine Unterschiede in der semantischen Interpretation der beiden Konstruktionen ergeben. Dies ist jedoch nicht der Fall. In (42a) werden die Subjekte der beiden Konjunkte als koreferent interpretiert, d.h. dieselbe Person kam und ging. In der nicht-elliptischen Variante (42b) wird dagegen die nicht-koreferente Interpretation bevorzugt (vgl. Klein, 1993). Dieser Kontrast spricht dafür, dass die sogenannte Rechtstilgung keinen Ellipseprozess darstellt, sondern als phrasale Koordination analysiert werden sollte (vgl. hierzu die Diskussion in Reich (2009, S. 31–33); siehe auch Höhle (1991) und Hartmann (2002)).

- (42) a. Jemand kam um vier Uhr und ging um fünf Uhr.
- b. Jemand kam um vier Uhr und jemand ging um fünf Uhr.

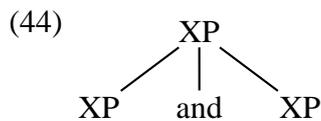
Zusammenfassend lässt sich sagen, dass es auf der einen Seite gute Gründe für den Reduktionsansatz gibt, insbesondere bei der Analyse der Nicht-Konstituenten-Koordination kann scheinbar nicht darauf verzichtet werden. Allerdings lassen sich auf der anderen Seite etliche Konstruktionen beobachten, die nicht auf sententiale Koordinationen zurückgeführt werden können. Man wird daher davon ausgehen müssen, dass neben einem (abgeschwächten) Reduktionsansatz auch phrasale Koordination benötigt wird.

2.3.2 Die phrasale Analyse

Wie im letzten Abschnitt diskutiert wurde, lassen sich Konstruktionen beobachten, die nicht auf sententiale Koordination plus Reduktion zurückgeführt werden können. Daraus kann auf die Existenz phrasaler Koordination geschlossen werden. Im Rahmen der phrasalen Analyse werden koordinierte Phrasen unterschiedlicher Kategorie direkt über Phrasenstrukturregeln basisgeneriert. Diese Ansätze können in zwei Gruppen unterteilt werden. In der ersten Gruppe werden die Konjunkte als Köpfe der Gesamtkoordination betrachtet, woraus sich mehrköpfige Konstruktionen ergeben. In der zweiten Gruppe von Ansätzen wird die Konjunktion als Kopf einer Konjunktionsphrase betrachtet. Die beiden Gruppen werden im Folgenden exemplarisch vorgestellt (siehe Progovac, 1998a,b).

Gruppe 1: Konjunkte als Kopf In frühen Ansätzen wurde eine mehrköpfige flache Analyse vorgeschlagen, die sich aus einer Phrasenstrukturregel wie in (43) ergibt, wobei X eine syntaktische Kategorie ist (Jackendoff, 1977; Chomsky, 1981). Daraus folgt, dass beispielsweise eine Nominalphrase (NP) nur mit einer Nominalphrase verknüpft werden kann, und die Konjunktion der beiden Phrasen ebenfalls eine Nominalphrase ist, etc. Die dazugehörige Struktur ist in (44) dargestellt.

(43) $XP \rightarrow XP \text{ Conj } XP$



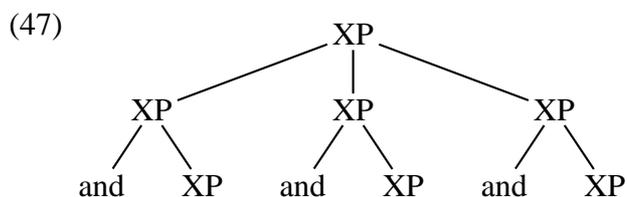
Es handelt sich hierbei um eine symmetrische Struktur, da alle Konjunkte denselben syntaktischen Status haben. Alle Konjunkte c-kommandieren sich gegenseitig und alle Konjunkte müssen derselben syntaktischen Kategorie angehören. Diese Struktur ist jedoch in verschiedener Hinsicht problematisch, insbesondere in Bezug auf die Beobachtung, dass die Konjunktion mit dem zweiten Konjunkt eine Konstituente bildet (Ross, 1967). Grundlage dieser Beobachtung bilden Daten wie in (45).

- (45) a. John left and he didn't even say good-bye.
 b. John left. And he didn't even say good-bye.
 c. *John left and. He didn't even say good-bye.

Auch auf nicht-sententialer Ebene wurde anhand von Extrapositionsdaten gezeigt, dass die Konjunktion mit dem zweiten Konjunkt eine Konstituente bildet (siehe z.B. Munn, 1993). Wie in (46) illustriert wird, kann die Konjunktion zusammen mit dem zweiten Konjunkt extraponiert werden (46b), nicht jedoch das erste Konjunkt mit der Konjunktion (46c).

- (46) a. John hat das Buch und die heutige Zeitung gelesen.
 b. John hat das Buch gelesen und die heutige Zeitung.
 c. *John hat die heutige Zeitung gelesen das Buch und.

Eine alternative mehrköpfige Analyse, die dieser Asymmetrie Rechnung trägt, wurde von Lakoff und Peters (1969) vorgeschlagen (siehe (47)). Diese etwas hierarchischer aufgebaute Struktur enthält zwei strukturelle Ebenen. Die obere Ebene ist nach wie vor mehrköpfig und symmetrisch. Alle Konjunkte haben den gleichen syntaktischen Status. Im Gegensatz zu (44) c-kommandieren die Konjunkte sich aber nicht gegenseitig. In der unteren asymmetrisch aufgebauten Struktur-Ebene ist jedem Konjunkt eine Kopie der Konjunktion vorangestellt. Dabei bildet die Konjunktion mit dem nachfolgenden Konjunkt jeweils eine Konstituente (für eine vergleichbare Struktur auf sententialer Ebene siehe Ross, 1967).



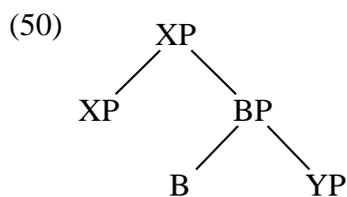
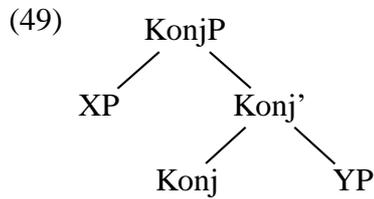
Durch die Symmetrie der oberen Struktur-Ebene sind symmetrische Phänomene, wie z.B. kategoriale Identität der Konjunkte, gleicher syntaktischer Status und symmetrische Kongruenz, gesichert. Durch die asymmetrische Strukturierung der unteren Ebene ist dennoch gewährleistet, dass die Konjunktion mit dem zweiten Konjunkt eine Konstituente bildet. Außerdem kann aus diesem Ansatz das sogenannte ‘*Conjunction Doubling*’ abgeleitet werden, dabei kann die Konjunktion vor jedem Konjunkt wiederholt werden. Dieses Phänomen wurde in einigen Sprachen beobachtet, wie z.B. im Französischen, Italienischen und Japanischen (siehe z.B. Kayne, 1994). In (48) ist ein Beispiel aus dem Französischen illustriert. Im Deutschen und Englischen ist dies jedoch nicht möglich.

(48) *Jean connait et Paul et Michel*
 Jean kennt und Paul und Michel

Entsprechend problematisch für mehrköpfige Konstruktionen sind cross-kategoriale Koordinationen sowie alle Asymmetrien, bei denen das erste Konjunkt einen besonderen Status hat, wie z.B. die asymmetrische Kongruenz, die asymmetrische Subkategorisierung, etc.

Gruppe 2: Konjunktion als Kopf Der Nachteil mehrköpfiger Konstruktionen ist, dass diese Strukturen nicht mit Standardannahmen, wie binärer Verzweigung und dem Kopfprinzip, im Einklang stehen. Neuere Ansätze behandeln daher Konjunktionen als funktionale Elemente, die eine eigene funktionale Projektion (*KonjP* oder *BP*) aufspannen (z.B. Munn, 1993, 2000; Kayne, 1994; Zoerner, 1995; Johannessen, 1998). Im Folgenden werden exemplarisch die Ansätze von Johannessen (1998) und Munn (1993, 2000) skizziert.

In Johannessen (1998) werden die Konjunkte als Spezifikator und Komplement einer Konjunktionsphrase aufgefasst. Die entsprechende Struktur ist in (49) dargestellt. In Munn (1993, 2000) wird die Konjunktionsphrase (*BP* = Boolean Phrase), die das zweite Konjunkt enthält, an das erste Konjunkt adjungiert (siehe die Struktur in (50)).



Diese Ansätze sind asymmetrisch, da das erste Konjunkt die nicht-ersten Konjunkte c-kommandiert. Das heißt, die beteiligten Konjunkte haben nicht mehr den gleichen syntaktischen Status. Es ist vielmehr das erste Konjunkt, das die Kategorie und die syntaktische Distribution der koordinierten Struktur bestimmt. Dies ergibt sich in Johannessen (1998) über die Spezifikator-Kopf-Kongruenz und in Munn (1993, 2000) durch die Annahme, dass das erste Konjunkt der Kopf der koordinierten Struktur ist.

Für diese Ansätze sprechen verschiedene Asymmetrien. Beispielsweise lassen sich daraus die Extrapositionsdaten in (46) problemlos ableiten, da die Konjunktion mit dem zweiten Konjunkt eine Konstituente bildet. Ein weiteres Argument ist die cross-kategoriale Koordination, da die Konjunkte in ihrer syntaktischen Kategorie nicht übereinstimmen müssen. Auch die asymmetrische Subkategorisierung, wie in (51), folgt aus obigen Ansätzen, da nur das erste Konjunkt die Anforderungen der syntaktischen Umgebung erfüllen muss.

- (51) a. You can depend on $[[_{NP}$ my assistant] and $[_{CP}$ that he will be on time]].
 b. *You can depend on $[_{CP}$ that he will be on time].

Auf der anderen Seite ergibt sich nun in Bezug auf die Subkategorisierung das umgekehrte Problem der Übergeneralisierung. In den meisten Fällen müssen beide Konjunkte die Anforderungen der syntaktischen Umgebung erfüllen, was mit den Analysen von Munn (1993, 2000) und Johannessen (1998) nicht erklärt werden kann. Entsprechend werden ungrammatische Strukturen, wie in (52), nicht mehr ausgeschlossen. Ebenfalls problematisch für die Analyse in (49) und (50) ist das ‘*Conjunction Doubling*’, d.h. das Wiederholen der Konjunktion vor jedem Konjunkt.

(52) *Die Täter versteckten [[*NP* die Tatwaffe] und [*PP* am Montag]].

Johannessen (1998) macht die die korrekte Vorhersage für symmetrische Kongruenz bei koordinierten Subjekten. In der Adjunktionsanalyse von Munn (1993, 2000) dagegen werden die Eigenschaften der koordinierten Struktur nur durch das erste Konjunkt bestimmt. Daraus ergibt sich eine direkte Erklärung für die asymmetrische Kongruenz, allerdings auf Kosten des regulären Symmetrie-Musters. Das symmetrische Phänomen der ATB-Bewegung lässt sich ebenfalls nicht direkt aus asymmetrischen Ansätzen wie in Munn (1993, 2000) und Johannessen (1998) ableiten.⁵

Insgesamt zeigen die verschiedenen phrasalen Ansätze das Spannungsfeld zwischen Symmetrie und Asymmetrie auf. Aus mehrköpfigen Strukturen lassen sich im Wesentlichen die symmetrischen Phänomene ableiten. Diese Ansätze sind allerdings mit dem Problem konfrontiert, dass sie nicht mit den Standardannahmen für Phrasenstrukturen vereinbar sind. Die asymmetrischen Ansätze, die die Konjunktion als Kopf einer Konjunktionsphrase betrachten, machen im Wesentlichen die richtige Vorhersage für asymmetrische Beobachtungen. Der symmetrische Normalfall kann jedoch nicht mehr (ohne Weiteres) erklärt werden, der asymmetrische Fall wird in gewisser Weise zum Normalfall erklärt.

⁵ In Reich (2009) wird jedoch gezeigt, wie ATB-Bewegung in einer Johannessenschen Struktur auf bekannte Bewegungs- und Tilgungsprozesse zurückgeführt werden kann. Dabei handelt es sich um einzelne lokale Bewegungen, die zusammengenommen den Eindruck eines symmetrischen Prozesses entstehen lassen.

2.3.3 Mehrdimensionale Modelle

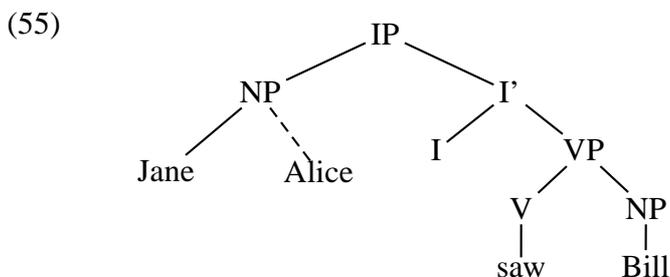
In mehrdimensionalen Modellen werden koordinierte syntaktische Strukturen als drei-dimensionale Bäume repräsentiert, wobei die Konjunkte strukturell hintereinander (anstatt nebeneinander) angeordnet sind (Goodall, 1987; Muadz, 1991; Moltmann, 1992; Wilder, 1999; De Vries, 2003). In Goodall (1987) entstehen drei-dimensionale Strukturen, indem reguläre zwei-dimensionale Phrasenstrukturen übereinandergelegt und die redundanten Knoten miteinander verschmolzen werden. Werden beispielsweise die zwei-dimensionalen Satzstrukturen für (53a) und (53b) übereinandergelegt und identische Knoten fusioniert, resultiert daraus eine drei-dimensionale Struktur für den Satz ‘*Jane and Alice saw Bill*’.

(53) a. Jane saw Bill.

b. Alice saw Bill.

Schematisch ergibt sich eine Anordnung wie in (54), wobei die Konjunkte auf parallelen Ebenen existieren. Die dazugehörige Struktur ist in (55) dargestellt, wobei durch die gestrichelten Linien angedeutet wird, dass sich ‘*Jane*’ und ‘*Alice*’ auf verschiedenen Ebenen befinden. Die Konjunkte werden auf der Ebene der Phonologischen Form linearisiert. In Goodalls Ansatz hat die Konjunktion keinen syntaktischen Status, sie wird während des Linearisierungsprozesses zwischen die Konjunkte eingefügt.

(54) Jane
 saw Bill.
 Alice



In einem drei-dimensionalen Baum, wie in (55), stehen die Konjunkte zu allen anderen Elementen im Satz in einer Präzedenz- und Dominanz-Relation. Zwischen den Konjunkten besteht jedoch keine Präzedenz- oder Dominanz-Relation. Entsprechend c-kommandieren sich die Konjunkte auch nicht gegenseitig. Nach De Vries (2003) besteht zwischen den Konjunkten eine (lokale) ‘Behindance’-Relation. Das zweite Konjunkt (‘*Alice*’) befindet sich *hinter* dem ersten Konjunkt (‘*Jane*’). Somit ist ein drei-dimensionaler Baum ein gerichteter Graph, in dem es neben einer Präzedenz-Relation (links–rechts) und einer Dominanz-Relation (oben–unten) noch zusätzlich eine ‘Behindance’-Relation (vorne–hinten) gibt. In drei-dimensionalen Bäumen werden Knoten, die Elemente auf verschiedenen Ebenen dominieren, als ‘*Splitting Nodes*’ bezeichnet. Der NP-Knoten, der ‘*Jane*’ und ‘*Alice*’ dominiert, ist ein solcher ‘*Splitting Node*’. Knoten, die zu beiden Ebenen gehören, werden als ‘*Shared Nodes*’ bezeichnet. Das sind die redundanten Knoten der einzelnen Ausgangssätze, die fusioniert und somit von beiden Konjunkten geteilt werden. Das betrifft in (55) alle Knoten mit Ausnahme der Terminalknoten ‘*Jane*’ und ‘*Alice*’.

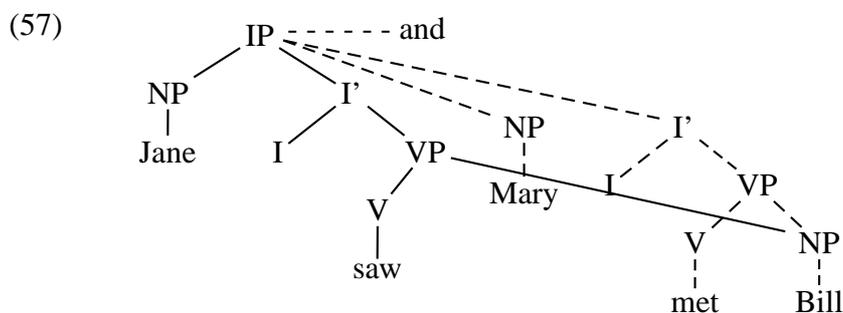
Grammatische Prinzipien, wie z.B. die Bedingungen der X-Bar-Theorie, müssen nach Goodall (1987) in den einzelnen Ausgangssätzen (53a,b) vor der Fusionierung der Strukturen erfüllt sein. Auf diese Weise wird das Problem der Mehrköpfigkeit bzw. das Problem mit mehrfach-verzweigenden Strukturen umgangen. Die Syntax der Koordination wird letztlich auf die Syntax von nicht-koordinierten Sätzen reduziert.

In Muadz (1991) wird ein ähnlicher Ansatz verfolgt, allerdings mit dem Unterschied, dass die drei-dimensionalen Phrasenstrukturen basisgeneriert werden. Auch Konjunktionen werden als Teil einer drei-dimensionalen Struktur basisgeneriert, sie werden allerdings keiner der parallelen Ebenen, d.h. keinem der Konjunkte, explizit zugeordnet und haben daher einen besonderen syntaktischen Status (vgl. die Struktur in (57)). Grammatische Prinzipien werden in Bezug auf zwei-dimensionale Schnitte evaluiert. Diese zwei-dimensionale Schnitte entsprechen den Ausgangssätzen bei Goodall (1987). In gewisser Hinsicht handelt es sich hierbei um das umgekehrte Verfahren. Bei Goodall (1987) werden zuerst die zwei-dimensionalen Ausgangssätze in Bezug auf die Grammatik evaluiert und anschließend die entsprechenden Strukturen fusioniert. Bei

Muadz (1991) werden die drei-dimensionalen Strukturen basisgeneriert und anschließend zwei-dimensionale Schnitte evaluiert. In beiden Ansätzen bilden die zwei-dimensionalen Strukturen die Basis für syntaktische Bedingungen und für die semantische Interpretation. Somit wird die Grundidee der Koordinationsreduktion beibehalten, nämlich dass koordinierte Strukturen letztlich auf der Verknüpfung von Sätzen basieren. Im Gegensatz zum Reduktionsansatz ist hierfür jedoch keine Reduktion erforderlich und es müssen keine leeren Strukturen angenommen werden.

Besonders interessant sind mehrdimensionale Modelle daher auch für die Analyse von Koordinationsellipsen, wie z.B. RNR-Konstruktionen in (56a). Die entsprechende drei-dimensionale Struktur nach Muadz (1991) ist in (57) dargestellt. In (57) sind zwei Sätze (IPs) koordiniert, wobei sich das zweite Konjunkt auf einer parallelen Ebene befindet (angedeutet durch die gestrichelten Linien). Die Objekt-NP 'Bill' wird von beiden Konjunkten geteilt. Der entsprechende NP-Knoten wird auch als 'Joining Node' bezeichnet, weil er von mehr als einem Knoten dominiert wird. Der Vorteil dieser Analyse ist, dass keine Tilgungsprozesse angenommen werden müssen. Grammatische Bedingungen werden in Bezug auf die zwei-dimensionalen Längsschnitte (56b,c) überprüft.

- (56) a. Jane saw and Mary met Bill.
 b. Jane saw Bill.
 c. Mary met Bill.



Weiterhin lässt sich das symmetrische Phänomen der ATB-Bewegung (symmetrische Extraktion) direkt aus einem mehrdimensionalen Ansatz ableiten. In (58) ist die drei-dimensionale Anordnung einer Konstruktion mit ATB-Bewegung schematisch dargestellt. Diese Konstruktion ist wohlgeformt, da die zugrundeliegenden zwei-dimensionalen Schnitte in (59a,b) wohlgeformt sind. Die *W*-Phrase ‘*welchen Kuchen*’ wird von beiden Ebenen geteilt.

(58) Welchen Kuchen_{*i*} hat $\left\langle \begin{array}{l} \text{Ilse } t_i \text{ gebacken} \\ \text{und} \\ \text{Georg } t_i \text{ verspeist} \end{array} \right\rangle$

- (59) a. Welchen Kuchen_{*i*} hat Ilse *t_i* gebacken?
 b. Welchen Kuchen_{*i*} hat Georg *t_i* verspeist?

Asymmetrische Extraktion, wie in (60), ist nicht wohlgeformt, da einer der zugrundeliegenden zwei-dimensionalen Schnitte (61b) nicht wohlgeformt ist.

(60) *Welchen Kuchen_{*i*} hat $\left\langle \begin{array}{l} \text{Ilse } t_i \text{ gebacken} \\ \text{und} \\ \text{Georg ihn verspeist} \end{array} \right\rangle$

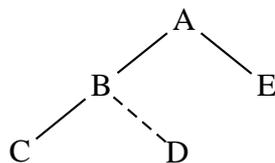
- (61) a. Welchen Kuchen_{*i*} hat Ilse *t_i* gebacken?
 b. *Welchen Kuchen_{*i*} hat Georg ihn verspeist?

Wenn jedoch zwei-dimensionale Schnitte die Basis für syntaktische Bedingungen und die semantische Interpretation darstellen, so ergeben sich auf der anderen Seite auch ähnliche Probleme wie beim Reduktionsansatz. So kann beispielsweise die ebenenweise Auswertung der einzelnen zwei-dimensionalen Schnitte nicht zu kollektiven Lesarten in (62a,b) führen.

- (62) a. John und Mary treffen sich in der Stadt.
 b. Peter und Bill tragen eine Leiter.

Eine Lösung hierfür bietet der mehrdimensionale Ansatz von Moltmann (1992), der auf Muadz (1991) aufbaut. Ein wesentlicher Unterschied besteht darin, dass Moltmann – im Gegensatz zu Goodall und Muadz – zwei Perspektiven auf einen drei-dimensionalen Baum erlaubt. Sie weist darauf hin, dass bei der kompositionellen Auswertung einer drei-dimensionalen Struktur zwei prinzipielle Vorgehensweisen möglich sind. Dies wird anhand der Struktur in (63) illustriert, wobei B ein ‘Splitting Node’ ist, d.h. die Elemente C und D befinden sich auf parallelen Ebenen.

(63)



Die erste Möglichkeit besteht darin, zuerst die Elemente C und D als Einheit zu interpretieren und zu evaluieren. Anschließend wird das Ergebnis mit der Interpretation des Elements E kombiniert (“cross-planare Prozedur”). Die zweite Möglichkeit besteht darin, C und E als Einheit zu interpretieren und parallel D und E auszuwerten. Anschließend werden die jeweiligen Ergebnisse (CE und DE) kombiniert. Im ersten Fall handelt es sich um eine lokale “cross-planare” Prozedur, die im Wesentlichen einer phrasalen Analyse entspricht. Es werden gewissermaßen Querschnitte betrachtet. Im zweiten Fall werden die einzelnen Längsschnitte betrachtet, was einer sententialen Analyse entspricht. In Goodall (1987) und Muadz (1991) wurden ausschließlich die Längsschnitte betrachtet. In Moltmann (1992) bilden sowohl Längs- als auch Querschnitte die Basis für syntaktische Bedingungen und die semantische Interpretation.

Das Interessante an diesem Ansatz ist, dass die Grundlage eine einzige drei-dimensionale Datenstruktur bildet. Je nach Perspektive auf die Datenstruktur ist es möglich, sowohl einer phrasalen als auch einer sententialen Analyse gerecht zu werden. So lassen sich beispielsweise auch kollektive Lesarten für (62a,b) ableiten. Hierfür müssen die Querschnitte betrachtet werden. Das Beispiel in (62b) ist ambig, es lässt sowohl eine kollektive als auch eine distributive Lesart zu. Der Satz kann so interpretiert werden, dass ‘Peter’ und ‘Bill’ gemeinsam

eine einzige Leiter tragen (kollektive Lesart) oder dass jeder eine eigene Leiter trägt (distributive Lesart). Für kollektive Lesarten müssen Querschnitte betrachtet werden, für distributive Lesarten müssen die einzelnen Längsschnitte ausgewertet werden.

Mehrdimensionale Modelle sind grundlegend symmetrisch. Die Konjunkte c-kommandieren sich nicht gegenseitig und beide Konjunkte müssen die Anforderungen der syntaktischen Umgebung erfüllen. Sie haben somit den gleichen syntaktischen Status. Symmetrische Phänomene wie ATB-Bewegung und symmetrische Subkategorisierung ergeben sich problemlos aus einem mehrdimensionalen Ansatz. Die symmetrische Kongruenz bei koordinierten Subjekten lässt sich nur ableiten, wenn die Querschnitte in einer drei-dimensionalen Struktur betrachtet werden. Die Möglichkeit hierfür wird in Moltmann (1992) geschaffen. Aus Goodall (1987) und Muadz (1991) lässt sich die symmetrische Kongruenz nicht ableiten.

Gegen einen mehrdimensionalen Ansatz sprechen Asymmetrien, bei denen eines der Konjunkte einen besonderen Stellenwert einnimmt. Das betrifft Phänomene wie asymmetrische Subkategorisierung (Abschnitt 2.2.3), asymmetrische Extraktion (Abschnitt 2.2.4) und m.E. asymmetrische Kongruenz (Abschnitt 2.2.5). Auch die Beobachtung, dass die Konjunktion mit dem zweiten Konjunkt eine Konstituente bildet (Ross, 1967), ist zumindest für die Ansätze in Goodall (1987) und Muadz (1991) problematisch. In Moltmann (1992) wird versucht, diesem Phänomen Rechnung zu tragen. Moltmann ordnet die Konjunktion explizit einer der Ebenen zu. Sie behandelt Konjunktionen als Kopf einer Konjunktionsphrase, welche wiederum ein Adjunkt zum nachfolgenden Konjunkt darstellt. Nach De Vries (2003) spannen Konjunktionen eine eigene funktionale Projektion auf, welche er in eine drei-dimensionale Grammatik integriert. Allerdings führt die asymmetrische Strukturierung der Konjunkte dazu, dass weniger syntaktische Knoten geteilt werden können.

Cross-kategoriale Koordination, wie z.B. in (64), ist hingegen unproblematisch für einen mehrdimensionalen Ansatz, zumindest solange die beteiligten zwei-dimensionalen Schnitte wohlgeformt sind. In einem mehrdimensionalen Ansatz müssen die Konjunkte nicht zwingend strukturgleich sein. Sie müssen lediglich jeweils die Anforderungen der syntaktischen

Umgebung erfüllen. Unter Ökonomie-Gesichtspunkten macht es jedoch tatsächlich einen Unterschied, ob die Konjunkte strukturgleich sind oder nicht. Je größer die strukturelle Parallelität der Konjunkte, desto mehr syntaktische Knoten können geteilt werden.

(64) John walked $\left\langle \begin{array}{l} \textit{and} \textit{ slowly} \\ \textit{with great care} \end{array} \right\rangle$

Ein Vorteil ist, dass sowohl phrasale Koordination als auch Koordinationsellipsen im Rahmen eines mehrdimensionalen Ansatzes behandelt werden können. Dies ist auch im Rahmen des Reduktionsansatzes möglich, jedoch kommt ein mehrdimensionaler Ansatz ohne Reduktion und ohne leere Strukturen aus, was aus Ökonomie-Gründen vorzuziehen ist.

2.4 Zusammenfassung und Motivation

In diesem Kapitel wurde der Phänomenbereich der Koordination vorgestellt und die symmetrischen bzw. asymmetrischen Eigenschaften koordinierter Strukturen beschrieben. Wir haben gesehen, dass sich der Begriff "Symmetrie" bzw. "Asymmetrie" unter vielerlei Gesichtspunkten betrachten lässt. Koordinierte Strukturen unterliegen grundlegenden Symmetriebedingungen unterschiedlicher Art. Auf der anderen Seite sind jedoch auch jeweils die entsprechenden Asymmetrien zu beobachten. Intuitiv scheint die jeweilige symmetrische Variante der Normalfall zu sein, d.h. sie wird vermutlich häufiger verwendet. Eine quantitative Untersuchung steht jedoch bislang aus. Desweiteren muss auch den jeweiligen asymmetrischen Phänomenen Rechnung getragen werden. Eine große Auswahl an Vorschlägen zur syntaktischen Analyse von Koordination reflektiert dieses Spannungsfeld zwischen Symmetrie und Asymmetrie. Abhängig davon, ob bei der Modellbildung den symmetrischen oder eher den asymmetrischen Eigenschaften den Vorzug gegeben wurde, führte das in der Vergangenheit zu symmetrischen bzw. zu asymmetrischen Modellen. Es scheint sich hierbei um einen unauflösbaren Widerspruch zu handeln.

Um diesen Widerspruch möglicherweise auflösen oder doch zumindest relativieren zu können, soll in dieser Arbeit das Spannungsfeld zwischen Symmetrie und Asymmetrie aus der Perspektive der Sprachverarbeitung beleuchtet werden. Es ist aus verschiedenen Gründen sinnvoll, hierfür die Sprachverarbeitung heranzuziehen. Das wird im Folgenden begründet.

- i. Durch die Ermittlung von Präferenzen bei der Sprachverarbeitung (Parsingpräferenzen, Korpusfrequenzen) lässt sich herausfinden, ob die symmetrische oder die asymmetrische Variante einer Eigenschaft den Normalfall darstellt. Ein syntaktisches Modell, das in der Lage ist, den Normalfall zu beschreiben, ist meines Erachtens einem Konkurrenz-Modell vorzuziehen, das nur die Ausnahmefälle beschreibt.
- ii. Mit Hilfe von Lesezeit-Effekten und Präferenzen können Verarbeitungsmodelle evaluiert werden, die ihrerseits auf syntaktischen Theorien basieren. Wenn verschiedene syntaktische Analysen zu unterschiedlichen Vorhersagen für die Verarbeitung führen, kann indirekt zwischen syntaktischen Theorien unterschieden werden.
- iii. Sollte es möglich sein, bestimmte Symmetrien oder Asymmetrien außerhalb der Syntax, also beispielsweise im Rahmen von Verarbeitungsmechanismen, zu erklären, könnte auf diese Weise obiger Widerspruch aufgelöst werden.

Alle drei oben geschilderten Möglichkeiten zur Gewinnung von Evidenz aus der Sprachverarbeitung werden in der vorliegenden Arbeit ausgeschöpft. Für meine Untersuchungen unterscheide ich zwei Arten von Symmetrie bzw. Asymmetrie. Ich nehme an, dass sich die beiden folgenden Symmetrie-Arten hinsichtlich der involvierten Verarbeitungsprozesse unterscheiden.

1. *Symmetrie zwischen den Konjunkten* (“Konjunkt-Symmetrie”): Hierbei geht es um die Relation zwischen den Konjunkten. Wie parallel müssen oder sollten die Konjunkte sein?

Beispiele:

- Kategoriale Symmetrie vs. cross-kategoriale Koordination
- (A)Symmetrie des semantischen Typs

2. *Symmetrie zur syntaktischen Umgebung* (“Umgebungs-Symmetrie”): Hierbei geht es um das Verhältnis der Konjunkte zu ihrer syntaktischen Umgebung. Haben die Konjunkte den gleichen syntaktischen Status, oder hat ein Konjunkt einen besonderen Stellenwert?

Beispiele:

- Symmetrische vs. asymmetrische Subkategorisierung
- Symmetrische vs. asymmetrische Extraktion
- Symmetrische vs. asymmetrische Kongruenz

In der vorliegenden Arbeit soll die Verarbeitung beider Symmetrie-Arten beleuchtet werden. Dabei konzentriere ich mich auf die syntaktischen Aspekte der Verarbeitung. Neben den jeweiligen Präferenzen für Symmetrie oder Asymmetrie sollen insbesondere die zugrundeliegenden Verarbeitungsmechanismen beleuchtet werden. Im Zusammenhang mit der *Konjunkt-Symmetrie* stellt sich die Frage, ob koordinierte Strukturen nach allgemeinen Parsingprinzipien verarbeitet werden, oder ob koordinationspezifische Mechanismen angenommen werden müssen. Zur Verarbeitung der Konjunkt-Symmetrie existieren bereits experimentelle und theoretische Arbeiten zum Englischen. Der Forschungsstand hierzu wird in Kapitel 3 vorgestellt. Die Verarbeitungseffekte zur Konjunkt-Symmetrie aus der Literatur dienen als Basis für die eigenen Untersuchungen (theoretische Modellierung, Korpusuntersuchung, Simulationen mit neuronalen Netzen) in den Kapiteln 4 – 6.

Hinsichtlich der *Umgebungs-Symmetrie* ist meine Ausgangshypothese, dass die beobachteten Asymmetrien keinen Reflex der Syntax darstellen, sondern vielmehr auf die Inkrementalität der Satzverarbeitung von links nach rechts zurückzuführen sind. Bei dieser Symmetrie-Art scheint das erste Konjunkt eine Sonderstellung einzunehmen. Das könnte damit zusammenhängen, dass das erste Konjunkt einen geringeren Abstand zu der zuvor verarbeiteten syntaktischen Umgebung hat als das zweite Konjunkt. Demnach wäre die Sonderstellung des ersten Konjunks lediglich eine Folge davon, dass das erste Konjunkt *zeitlich vor* dem zweiten Konjunkt verarbeitet wird. Die Umgebungs-Symmetrie wird exemplarisch anhand der asymmetrischen Kongruenz im Deutschen in Kapitel 7 untersucht. Hierzu werden eigene Experimente und Korpusuntersuchungen durchgeführt. Für beide Symmetrie-Arten soll jeweils untersucht werden, welche Implikationen sich für die Syntax ergeben.

Kapitel 3

Die Verarbeitung koordinierter Strukturen

In diesem Kapitel werden Verarbeitungseffekte aus der Literatur präsentiert, die beim Sprachverstehen von koordinierten Strukturen im Englischen beobachtet wurden. Diese Effekte betreffen die oben genannte *Konjunkt-Symmetrie*, d.h. die Relation zwischen den Konjunkten, und dienen als Grundlage für die nachfolgenden Kapitel. In Kapitel 4 werden die Effekte modelliert, in Kapitel 5 wird der Frage nachgegangen, ob diese auch in der Sprachproduktion auftreten. In Kapitel 6 schließlich werden künstliche neuronale Netze trainiert, die im Hinblick auf die genannten Verarbeitungseffekte ausgewertet werden.

Bevor die Verarbeitungseffekte im Einzelnen präsentiert werden, gehe ich im nächsten Abschnitt 3.1 auf allgemeine Mechanismen und Strategien bei der syntaktischen Verarbeitung im Sprachverstehen ein. In den Abschnitten 3.2 und 3.3 werden die Effekte bei der Verarbeitung koordinierter Strukturen vorgestellt (Parallelitätseffekte, Präferenzen bei Ambiguitäten). Es wird jeweils geprüft, inwiefern diese Effekte aus allgemeinen von Koordination unabhängigen Parsingmechanismen abgeleitet werden können. In Abschnitt 3.4 diskutiere ich das Copy- α -Modell (Frazier und Clifton, 2001), das speziell für die Verarbeitung koordinierter Strukturen und Ellipsen entwickelt wurde, und stelle im Anschluss in Abschnitt 3.5 die Ziele der eigenen Untersuchungen vor.

3.1 Grundlegendes zur syntaktischen Verarbeitung

Beim Verstehen eines Satzes muss der Hörer oder Leser die einzelnen Wörter des Satzes identifizieren und zu komplexeren Einheiten (Phrasen) kombinieren. Dieser Strukturierungsprozess erfolgt nach den syntaktischen Regeln der einzelsprachlichen Grammatik, d.h. die für einen Satz aufgebaute Struktur spiegelt wider, welche Regeln dabei angewandt wurden und welche Relationen zwischen Satzteilen bestehen. Durch die hierarchische Strukturierung des Satzes können die relevanten Informationen besser im Arbeitsgedächtnis behalten werden. Die syntaktische Struktur bildet wiederum die Basis dafür, die Bedeutung des Satzes berechnen zu können. Kleine Änderungen in der Wortabfolge können bereits zu einer völlig anderen Satzstruktur und somit auch zu einer anderen Satzbedeutung führen (siehe hierzu Beispiel (1) aus Frazier und Clifton, 1996, S. 3).

- (1) a. He showed [her baby] [the pictures].
 b. He showed [her] [the baby pictures].

Lesezeit-Experimente erlauben Rückschlüsse auf den Ablauf des Strukturierungs- oder Parsingprozesses. Die Annahme, die dabei zugrunde liegt, ist, dass Lesezeiten von der Komplexität der Struktur abhängen. Dieser Sachverhalt ist in der folgenden "Strukturhypothese" formuliert, die im Verlauf des Kapitels noch relevant sein wird.

Strukturhypothese: "Structure building is a matter of syntactic inference, and building more structure requires more inferences and is more costly." (Frazier und Clifton, 2001, S. 1)

Die Strukturhypothese besagt, dass die Lesezeiten für ein Wort oder für eine Phrase umso länger werden, je mehr Struktur für dieses Wort oder diese Phrase aufgebaut werden muss. Als Basis für die mentalen Repräsentationen syntaktischer Strukturen werden im Allgemeinen die in der theoretischen Linguistik entwickelten Datenstrukturen verwendet.

Aufschluss über den Parsingprozess erhält man insbesondere durch die gezielte Untersuchung von syntaktischen Ambiguitäten. Es handelt sich dabei um

Sätze, die dieselbe Abfolge von Wörtern enthalten, denen aber unterschiedliche Bedeutungen zugeordnet werden können. Dabei liegt jeder Satzbedeutung eine andere syntaktische Struktur zugrunde. Generell muss zwischen lokalen bzw. temporären Ambiguitäten und globalen Ambiguitäten unterschieden werden.

3.1.1 Lokale vs. globale Ambiguitäten

Eine *lokale Ambiguität* liegt vor, wenn es bei der Verarbeitung eines Satzes mehrere Möglichkeiten gibt, ein Wort in die schon bestehende syntaktische Struktur einzubinden, und sich diese Ambiguität im Verlauf des Satzes wieder auflöst. Dies wird anhand von Beispiel (2) illustriert, welches aus Frazier und Clifton (1996, S. 4) stammt. In (2a) ist die Nominalphrase [*NP* the answer] ein Objekt zu ‘knew’, in (2b) ist sie jedoch Subjekt des eingebetteten Satzes. In (2a,b) muss die ambige Phrase unterschiedlich in die bestehende Struktur eingebunden werden. Durch die nachfolgenden Wörter (‘*very well*’ bzw. ‘*was wrong*’) wird die Ambiguität wieder aufgelöst. Die Auflösung erfolgt durch syntaktische Information, da es zu einem Widerspruch innerhalb des grammatischen Regelsystems kommt.

- (2) a. John [*VP* knew [*NP* the answer] *very well*].
 b. John [*VP* knew [[*NP* the answer] *was wrong*]].

Globale Ambiguitäten unterscheiden sich von lokalen Ambiguitäten dadurch, dass sich die Ambiguität im Verlauf des Satzes nicht auflöst. Am Ende des Satzes können ihm immer noch zwei mögliche Strukturen zugewiesen werden, die zwei Interpretationsmöglichkeiten zulassen (siehe z.B. (3)). In der ersten Lesart besitzt der Mann das Fernglas. Die dazugehörige Struktur ist in (3a) skizziert. In der zweiten Lesart schaut ‘John’ durch das Fernglas hindurch (siehe (3b)). Bei globalen Ambiguitäten kann es durch die Plausibilität der Sätze oder durch den Kontext zu einer Auflösung der Ambiguität kommen.

- (3) a. John [*VP* saw [*NP* the man with the telescope]].
 b. John [*VP* saw [*NP* the man] with the telescope].

3.1.2 Der Parsingprozess

Es gibt prinzipiell mehrere Möglichkeiten, wie der menschliche Parser mit Ambiguitäten umgeht. Eine Möglichkeit besteht darin, mehrere konkurrierende Strukturen parallel aufzubauen und an der desambiguierenden Stelle die passende Struktur auszuwählen. Es wäre ebenso denkbar, dass der menschliche Parser abwartet bis die Ambiguität aufgelöst ist, um nur die eine korrekte Struktur aufbauen zu müssen. Die dritte Möglichkeit besteht darin, dass der Parser eine Analyse berechnet und diese erst verwirft, wenn er auf Gegenevidenz stößt.

Um Aufschluss über den Parsingprozess zu bekommen, wurde in der Vergangenheit hauptsächlich die Verarbeitung von Sätzen mit *lokalen* Ambiguitäten untersucht. Eine Fülle von Evidenzen aus Lesezeit-Experimenten zeigt, dass der Parser an der desambiguierenden Stelle Verarbeitungsschwierigkeiten hat. Diese sind jedoch asymmetrisch, d.h. sie treten bei einer bestimmten Interpretation auf, bei einer anderen wiederum nicht (siehe Frazier und Clifton, 1996, Kapitel 1). Beispielsweise treten in Satzkonstruktionen wie in (2) an der desambiguierenden Region in (2b) Verarbeitungsschwierigkeiten auf, nicht jedoch in (2a). Das heißt, es werden längere Lesezeiten für die Wortfolge ‘*was wrong*’ in (2b) als für ‘*very well*’ in (2a) beobachtet (zur Verarbeitung von Konstruktionen wie in Beispiel (2) siehe Rayner und Frazier, 1987; Frazier und Rayner, 1982).

Diese empirischen Evidenzen sprechen für die dritte oben genannte Möglichkeit des Parsingprozesses. Der Parser arbeitet einen Satz inkrementell Wort für Wort ab. Jedes Wort wird ohne Verzögerung in die bis dahin aufgebaute Struktur eingebunden. Im Falle einer strukturellen Ambiguität berechnet der Parser eine präferierte Struktur für die ambige Stelle im Satz und verwirft diese erst bei Gegenevidenz. Das hat den Vorteil, dass zumindest eine Analyse sobald wie möglich zur Verfügung steht. Stellt sich danach heraus, dass die präferierte Analyse inakzeptabel ist, wird ein zeitaufwändiger Reanalyseprozess eingeleitet.

Für die Sätze in (2) bedeutet das, dass die ambige NP [_{NP} the answer] zuerst als Objekt interpretiert und strukturell entsprechend eingebunden wird. Stößt der Parser an der desambiguierenden Stelle auf einen Widerspruch, muss die Struktur revidiert und neu aufgebaut werden. Dies ist der Fall, wenn der Parser in Beispiel (2b) die Wortfolge ‘*was wrong*’ verarbeitet. Die ambige Phrase, die

bereits als Objekt in die Struktur integriert wurde, muss an diesem Punkt als Subjekt des eingebetteten Objekt-Satzes strukturell neu eingebunden werden. Dieser Reanalyseprozess erfordert zusätzliche Verarbeitungszeit, was sich in den Lesezeiten niederschlägt. Entspricht die präferierte Struktur der korrekten Analyse, wie in (2a), sind keine Reanalyseprozesse erforderlich.

Studien zeigen, dass der Parser auch bei *globalen* Ambiguitäten Präferenzen hat und dem Satz erst einmal eine einzige Struktur zuweist. Erst bei Gegenevidenzen aus dem Kontext oder ähnlichem wird die nicht-präferierte Analyse berechnet (siehe z.B. Konieczny, 1996).

3.1.3 Präferenzen bei Ambiguitäten

Es stellt sich nun die Frage, welche Präferenzen der Parser bei Ambiguitäten hat bzw. ob es allgemeine Prinzipien gibt, nach denen der Parser vorgeht. Im *Garden-Path-Modell* (Frazier, 1987a, 1990) wird angenommen, dass der Parser bei der Berechnung der präferierten Struktur nach folgenden zwei Prinzipien vorgeht, welche durch eine Vielzahl von Evidenzen gestützt werden.

- *Minimal Attachment* (MA):
Do not postulate any potentially unnecessary nodes.
- *Late Closure* (LC):
If grammatically permissible, attach new items into the clause or phrase currently being processed (i.e., the clause or phrase postulated most recently).

Durch das Minimal-Attachment-Prinzip wird im Falle einer strukturellen Ambiguität die einfachste der möglichen Strukturen gewählt. Dabei ist die Anzahl der syntaktischen Knoten relevant. Durch '*Late Closure*' wird sichergestellt, dass eine neue Konstituente an die letztmögliche angebunden wird und nicht an eine weiter zurückliegende Konstituente. Dadurch wird beispielsweise auch das Arbeitsgedächtnis entlastet. Wenn beide Prinzipien angewandt werden können und zu einem Widerspruch führen, hat das Minimal-Attachment-Prinzip Vorrang.

'*Minimal Attachment*' und '*Late Closure*' sind Parsingprinzipien, die das menschliche Leseverhalten bei einer Vielzahl von Satzkonstruktionen erklären

können (siehe z.B. Frazier, 1979; Frazier und Rayner, 1982; Rayner und Frazier, 1987; Frazier und Clifton, 1996, S. 9–12). In Satzbeispiel (2) lässt sich die Präferenz für die Objekt-Lesart von [*NP* the answer] beispielsweise durch das Minimal-Attachment-Prinzip erklären. Die syntaktische Struktur erfordert eine größere Anzahl an Knoten, wenn die ambige Phrase [*NP* the answer] als Subjekt des Komplementsatzes anstatt als Objekt zu ‘know’ in die Struktur integriert wird. Die einfachere Struktur wird präferiert.

Diese beiden Kernprinzipien des Garden-Path-Modells spielen im aktuellen Nachfolgermodell ‘*Construal*’ immer noch eine wichtige Rolle. In der *Construal*-Theorie (Frazier und Clifton, 1996, 1997) wird davon ausgegangen, dass es unterschiedliche Parsingprinzipien für primäre und sekundäre Phrasen bzw. Relationen gibt. Unter primären Phrasen oder Relationen verstehen die Autoren das Prädikat eines Satzes, das Subjekt, die Komplemente sowie obligatorische Konstituenten von primären Phrasen. Sekundäre Phrasen sind im wesentlichen Modifikatoren. Handelt es sich bei allen Anbindungsalternativen um primäre Phrasen, sind die beiden Kernprinzipien des Garden-Path-Modells immer noch gültig. Die Präferenz für die Objekt-Lesart in (2) kann daher nach wie vor mit dem Minimal-Attachment-Prinzip erklärt werden, da es sich bei den beiden Analysemöglichkeiten um primäre Phrasen handelt (Subjekt vs. Objekt).

Die beiden Kern-Parsingprinzipien ‘*Minimal Attachment*’ und ‘*Late Closure*’ treten in ähnlichen Varianten immer wieder in Erscheinung. Eine Variante des Minimal-Attachment-Prinzips ist beispielsweise die ‘*Simple Structure Hypothesis*’ (Carlson, 2001).¹ Prinzipien dieser Art reflektieren eine generelle Präferenz für einfache ökonomische Strukturen. Eine Variante des Late-Closure-Prinzips ist die allgemeine ‘*Recency Preference*’ (Gibson, 1991; Gibson, Pearlmutter, Canseco-Gonzalez und Hickok, 1996).² Diese Prinzipien reflektieren eine generelle Präferenz für Nähe, d.h. neue Wörter oder Phrasen an (zeitlich) aktuelle Strukturen anzubinden.

¹ Die *Simple Structure Hypothesis* besagt: ‘The simplest legitimate syntactic analysis of an input is preferred.’ (Carlson, 2001, S. 4).

² Die *Recency Preference* besagt: ‘Preferentially attach structures for incoming lexical items to structures built more recently.’ (Gibson, 1991; Gibson, Pearlmutter, Canseco-Gonzalez und Hickok, 1996).

Zusammengefasst lässt sich sagen, dass in Modellen zur Satzverarbeitung im Allgemeinen Strukturierungsprozesse spezifiziert werden, woraus sich Vorhersagen für Lesezeiten ergeben. Meist werden Parsingprinzipien definiert (wie z.B. *Minimal Attachment*, *Late Closure*, etc.), nach denen Ambiguitäten aufgelöst werden, d.h. es kann daraus abgeleitet werden, welche der möglichen Anbindungsalternativen präferiert wird. Im Allgemeinen wird davon ausgegangen, dass der Parsingprozess inkrementell verläuft. Das heißt, nach jedem gehörten oder gelesenen Wort wird die dazu passende Struktur ohne Verzögerung nach den Regeln der Grammatik aufgebaut und in die bestehende Struktur integriert. Je mehr Struktur aufgebaut werden muss und je komplexer die Struktur ist, desto mehr Zeit benötigt der Parser für die Analyse (siehe 'Strukturhypothese'). In lokal ambigen Strukturen werden die Lesezeiten außerdem durch zeitaufwändige Reanalyseprozesse beeinflusst. Lesezeit-Experimente erlauben daher sowohl Rückschlüsse auf Parsingprozesse, und somit auf Verarbeitungsmodelle, als auch auf die Komplexität von syntaktischen Strukturen.

In den folgenden beiden Abschnitten 3.2 und 3.3 werden Effekte aus der Literatur vorgestellt, die bei der Verarbeitung koordinierter Strukturen beobachtet wurden. Es wird jeweils aufgezeigt, inwiefern sich diese Effekte aus den genannten allgemeinen Parsingmechanismen ableiten lassen.

3.2 Parallelitätseffekte bei Koordination

Im Folgenden werden syntaktische Parallelitätseffekte bei der Verarbeitung koordinierter Strukturen beschrieben. Es wird jeweils aufgezeigt, inwiefern sich diese Effekte aus den allgemeinen Parsingmechanismen ableiten lassen.

3.2.1 Parallelitätseffekte bei der Interpretation

Beim Sprachverstehen von Koordination werden strukturgleiche Konjunkte präferiert, d.h. ambige Konstruktionen werden bevorzugt parallel interpretiert. Die Sätze in (4) beispielsweise lassen unterschiedliche Interpretationsmöglichkeiten zu (siehe Frazier et al., 1984).

- (4) a. Joshua hit [_{NP} the girl with the book] and [_{NP} the boy with the bat].
- b. Joshua [_{VP} hit [_{NP} the girl] with the book] and [_{VP} ε [_{NP} the boy] with the bat].³
- c. Joshua hit [_{NP} the girl with the book] and [_{NP} the boy] with the bat.

Eine Möglichkeit besteht darin, dass ‘Joshua’ das Mädchen und den Jungen schlägt, wobei das Instrument, also womit geschlagen wird, nicht genannt wird. Das Mädchen hat ein Buch bei sich und der Junge einen Stock. Diese Interpretation ist in (4a) dargestellt. Die zweite Möglichkeit ist, dass das Buch das Instrument ist, womit das Mädchen geschlagen wird, und der Stock ist das Instrument, womit der Junge geschlagen wird. Diese Interpretation ist in (4b) dargestellt. Eine dritte Möglichkeit besteht darin, dass der Stock das Instrument ist, mit dem sowohl das Mädchen als auch der Junge geschlagen werden. Das Mädchen hat außerdem ein Buch bei sich. Dies ist in (4c) dargestellt. Bei der Variante in (4a,b) sind die beiden Konjunkte strukturgleich, in (4c) hingegen sind sie nicht strukturgleich. Bei der Interpretation werden die strukturgleichen Varianten (4a,b) bevorzugt gegenüber der nicht-parallelen Variante in (4c).

Ein weiteres (eigenes) Beispiel aus dem Deutschen ist in (5) dargestellt. Die einzelnen Konjunkte sind eingeklammert. In der ersten Lesart werden Mitarbeiter beiderlei Geschlechts mit Führerschein gesucht (5a). In der zweiten Lesart benötigen Frauen einen Führerschein, Männer jedoch nicht (5b). Die strukturell parallele Variante in (5a) wird im Allgemeinen bevorzugt (meine eigene Untersuchung).

- (5) a. Die neue Firma sucht [Männer] und [Frauen] mit Führerschein.
- b. Die neue Firma sucht [Männer] und [Frauen mit Führerschein].

³ Bei dieser syntaktischen Analyse werden zwei Verbalphrasen koordiniert mit anschließender Tilgung des Verbs ‘hit’ im zweiten Konjunkt. Das fehlende Verb wird durch ε dargestellt.

Die Präferenz, koordinierte Strukturen parallel zu interpretieren, ist auch bei der Anbindung von Relativsätzen zu beobachten (siehe Henstra, 1996; Schepman, 1997). In (6a) bezieht sich der Relativsatz [*CP* who was/were walking to the courtroom] auf beide Konjunkte, in (6b) nur auf das zweite Konjunkt.

- (6) a. The lawyer greeted [[the powerful barrister] and [the old judge]] who were walking to the courtroom.
- b. The lawyer greeted [[the powerful barrister] and [the old judge who was walking to the courtroom]].

Bei der Verarbeitung des Relativpronomens ‘*who*’ sind beide Sätze noch ambig. Diese Ambiguität löst sich beim darauffolgenden finiten Verb (‘*was*’/‘*were*’) auf. Nur in (6a) sind die Konjunkte parallel strukturiert. In Satzvervollständigungs-Tests sollte die finite Verbform (‘*was*’/‘*were*’) von den Probanden ergänzt werden. Dabei hat sich herausgestellt, dass die parallele Variante in (6a) unabhängig von der Reihenfolge der Konjunkte bei der Interpretation bevorzugt wird.

Diese Präferenz für Parallelität lässt sich nicht aus den bestehenden Parsingprinzipien ableiten. Diese würden für das letzte Beispiel in (6) sogar eine Präferenz für Nicht-Parallelität vorhersagen. Vergleicht man die Komplexität der beiden möglichen Strukturen (6a) vs. (6b), stellt man fest, dass beide Varianten die gleiche Anzahl an syntaktischen Knoten erfordern. Keine der beiden Varianten ist ökonomischer, so dass weder das Minimal-Attachment-Prinzip noch die Simple-Structure-Hypothese eine Vorhersage liefern kann. Letztlich würde das Late-Closure-Prinzip oder das allgemeine Recency-Prinzip zur Anwendung kommen, was fälschlicherweise eine Präferenz für die Anbindung an das letzte Konjunkt vorhersagen würde.

3.2.2 Parallel-Struktur-Effekte auf sententialer Ebene

Die Präferenz für parallele Konjunkte hat auch Einfluss auf die Geschwindigkeit beim Lesen. In Lesezeit-Experimenten (*Self-paced Reading*⁴) zum Englischen wurde beobachtet, dass das zweite Konjunkt in koordinierten Sätzen schneller gelesen wird, wenn es mit dem ersten Konjunkt strukturgleich ist (*Parallel-Struktur-Effekt*). Frazier et al. (1984) haben diesen Effekt in einer Reihe von Satzkonstruktionen gefunden, wie z.B. in Aktiv-Passiv-Konstruktionen (siehe (7)), in lokal ambigen Sätzen (direktes Objekt vs. Satzkomplement, siehe (8)) sowie in Konstruktionen mit einer bewegten ('shifted') bzw. nicht-bewegten ('nonshifted') "schweren" NP (siehe (9)).

In Aktiv-Passiv-Konstruktionen (siehe (7)) wurde ein Parallel-Struktur-Effekt für die zweiten Konjunkte im Aktiv gefunden, d.h. das zweite Konjunkt [*the king hugged the queen*] wird in (7a) schneller gelesen als in (7d). In (7a) sind die beiden Konjunkte strukturell parallel, beide Konjunkte enthalten eine Aktiv-Konstruktion. Die Konjunkte in (7d) sind strukturell nicht parallel (Passiv-plus Aktiv-Konstruktion).

- (7) a. [The prince kissed the princess]
and [the king hugged the queen]. (Aktiv + Aktiv)
- b. [The princess was kissed by the prince]
and [the queen was hugged by the king]. (Passiv + Passiv)
- c. [The prince kissed the princess]
and [the queen was hugged by the king]. (Aktiv + Passiv)
- d. [The princess was kissed by the prince]
and [the king hugged the queen]. (Passiv + Aktiv)

⁴ 'Self-paced Reading' ist Lesen in selbstgesteuerter Geschwindigkeit am Bildschirm. Die Präsentation erfolgt üblicherweise wort- oder phrasenweise, wobei sich die Probanden mit jedem Tastendruck das nächste Segment auf den Bildschirm holen. Es wird die Zeit gemessen, die die Probanden auf einem Segment verbringen.

Die Sätze in (8) enthalten eine lokale Ambiguität (Garden-Path-Sätze). Die NP [*all Tom's stories*] kann entweder als direktes Objekt von 'believe' ('minimal attachment') oder als Subjekt eines Objektsatzes ('nonminimal attachment') interpretiert werden. Die Interpretation als direktes Objekt wird vom menschlichen Parser im Allgemeinen bevorzugt, weil für diese Lesart weniger syntaktische Knoten aufgebaut werden müssen (Frazier, 1979). Diese Präferenz führt zu längeren Lesezeiten in Konstruktionen mit Satzkomplement (Garden-Path-Effekt).

- (8) a. [Jim believed all Tom's stories]
and [Sue believed Jim's stories].
(minimal + minimal)
- b. [Jim believed all Tom's stories were literally true]
and [Sue believed Jim's stories were fictitious].
(nonminimal + nonminimal)
- c. [Jim believed all Tom's stories]
and [Sue believed Jim's stories were fictitious].
(minimal + nonminimal)
- d. [Jim believed all Tom's stories were literally true]
and [Sue believed Jim's stories].
(nonminimal + minimal)

Frazier et al. (1984) fanden in diesen Sätzen einen Parallel-Struktur-Effekt für Konstruktionen mit Satzkomplement. Das zweite Konjunkt [*Sue believed Jim's stories were fictitious*] wird in (8b) schneller gelesen als in (8c). In (8b) sind die beiden Konjunkte strukturell parallel, beide Konjunkte enthalten ein Satzkomplement. In (8c) sind die Konjunkte strukturell nicht parallel.

In Sätzen mit "schwerer" NP (siehe (9)) wurde der Parallel-Struktur-Effekt für die zweiten Konjunkte mit bewegter NP gefunden. Das zweite Konjunkt [*Sue wired to her father a telegram requesting more money*], in dem die Objekt-NP aus stilistischen Gründen ans Satzende bewegt wurde, wird in (9b) schneller gelesen als in (9c). In (9b) sind die beiden Konjunkte strukturell parallel,

beide Konjunkte enthalten eine bewegte “schwere” NP. Die Konjunkte in (9c) sind strukturell nicht parallel.

- (9) a. [Mary wrote a long note about her predicament to her mother]
and [Sue wired a telegram requesting more money to her father].
(nonshifted + nonshifted)
- b. [Mary wrote to her mother a long note about her predicament]
and [Sue wired to her father a telegram requesting more money].
(shifted + shifted)
- c. [Mary wrote a long note about her predicament to her mother]
and [Sue wired to her father a telegram requesting more money].
(nonshifted + shifted)
- d. [Mary wrote to her mother a long note about her predicament]
and [Sue wired a telegram requesting more money to her father].
(shifted + nonshifted)

Insgesamt fanden Frazier und Kollegen einen robusten Parallel-Struktur-Effekt in allen untersuchten Satzkonstruktionen. Der menschliche Parser profitiert offensichtlich von der strukturellen Übereinstimmung der Konjunkte. Der Effekt ist nicht nur auf ambige Strukturen, wie in (8), beschränkt und kann daher nicht ausschließlich darauf zurückgeführt werden, dass der Parser bei lokalen Ambiguitäten im zweiten Konjunkt diejenige syntaktische Regel wählt, die auch im ersten Konjunkt zur Anwendung kam.

Nach allgemeinen Parsingmechanismen wird für das erste und zweite Konjunkt jeweils die passende Struktur nach den Regeln der Grammatik aufgebaut. Dies geschieht unabhängig davon, ob die beiden Konjunkte parallel sind oder nicht. So muss z.B. für das zweite Konjunkt [*the king hugged the queen*] in (7a) und in (7d) jeweils die gleiche Struktur aufgebaut werden, obwohl es sich in (7a) um parallele und in (7d) um nicht-parallele Konjunkte handelt. Die Form des ersten Konjunks hat keinen Einfluss auf den Parsingprozess des zweiten Konjunks. Unter Beibehaltung der etablierten *Strukturhypothese* (siehe

Abschnitt 3.1) dürften sich die Lesezeiten für [*the king hugged the queen*] in (7a) und (7d) nicht unterscheiden, da in beiden Fällen die gleiche Struktur aufgebaut werden muss. Die Verarbeitungserleichterung bei parallelen Strukturen kann aus den allgemeinen Parsingmechanismen nicht abgeleitet werden.

3.2.3 Parallel-Struktur-Effekte auf phrasaler Ebene

3.2.3.1 Cross-kategoriale Koordination

Der Parallel-Struktur-Effekt wurde auch beim Lesen von koordinierten Phrasen gefunden. Frazier et al. (2000) haben in einem Lesezeit-Experiment zum Englischen ('Eye Tracking') beobachtet, dass die Verknüpfung von gleichen syntaktischen Kategorien, wie in (10a), schneller verarbeitet wird als die Koordination von ungleichen syntaktischen Kategorien (cross-kategoriale Koordination), wie in (10b).

- (10) a. John walked [*AdvP* slowly] and [*AdvP* carefully], avoiding the broken glass.
- b. John walked [*AdvP* slowly] and [*PP* with great care], avoiding the broken glass.
- c. John walked [*AdvP* carefully], avoiding the broken glass.
- d. John walked [*PP* with great care], avoiding the broken glass.

Die Adverbphrase [*AdvP* carefully] wird schneller gelesen, wenn sie mit einer weiteren Adverbphrase verknüpft ist (10a), als wenn sie unabhängig von Koordination vorkommt (10c). Im Gegensatz dazu wird die Präpositionalphrase [*PP* with great care] nicht schneller gelesen, wenn sie mit einer Adverbphrase verknüpft ist (10b), als wenn sie unabhängig von Koordination vorkommt (10d).

Mit einer zusätzlichen Fragebogen-Studie hatten die Autoren sichergestellt, dass die verwendeten Testsätze als voll grammatisch beurteilt wurden. Die Präferenz für gleiche syntaktische Kategorien können somit nicht auf die Ungrammatikalität von cross-kategorialer Koordination zurückgeführt

werden. Gleiche syntaktische Kategorien bei der Koordination scheinen eine Verarbeitungserleichterung im Parsingprozess zu bewirken.

Diese Verarbeitungserleichterung bei gleichen syntaktischen Kategorien lässt sich jedoch nicht aus allgemeinen Parsingmechanismen ableiten. Wie auch bereits für den Parallel-Struktur-Effekt auf sententialer Ebene ausgeführt, muss die Struktur für beide Konjunkte unabhängig vom vorangehenden Kontext nach den Regeln der Grammatik aufgebaut werden. Es ist also unerheblich, ob eine Adverb-Phrase wie *'carefully'* mit einer anderen Adverb-Phrase koordiniert wird wie in (10a) oder ob sie unabhängig von Koordination wie in (10c) vorkommt. In beiden Fällen muss für die AdvP *'carefully'* genau die gleiche Struktur aufgebaut werden (siehe Abbildungen 3.1 und 3.2).⁵ Die syntaktischen Knoten, die jeweils für das zweite Konjunkt aufgebaut werden müssen, sind in beiden Abbildungen eingekreist.⁶ Nach der etablierten *Strukturhypothese* dürften sich daher die Lesezeiten für *'carefully'* in (10a) und (10c) nicht unterscheiden. Die Verarbeitungserleichterung bei gleichen syntaktischen Kategorien kann mit allgemeinen Parsingmechanismen nicht erklärt werden.

3.2.3.2 Parallel-Struktur-Effekte im Phrasenaufbau

In Frazier et al. (2000) wurde gezeigt, dass der Parallel-Struktur-Effekt sogar den inneren Phrasenaufbau der Konjunkte betrifft. Es wurden Sätze wie in (11) verglichen, in denen gleiche syntaktische Kategorien (hier Nominalphrasen) verknüpft sind.

(11) a. Hilda noticed [_{NP} a strange man] and [_{NP} a tall woman].

b. Hilda noticed [_{NP} a man] and [_{NP} a tall woman].

⁵ Das flektierte Verb wird beim Parsing im Allgemeinen in der Form repräsentiert, wie es im Input vorgegeben ist. Die Flexionsendung wird nicht abgespalten. In weiteren Verarbeitungsschritten erfolgt ein Abgleich der Flexionsmerkmale mit dem Subjekt (siehe z.B. Crocker, 1996).

⁶ Die genaue Analyse der koordinierten Struktur ist an dieser Stelle unerheblich. Relevant ist die Symmetrie zwischen den Konjunkten.

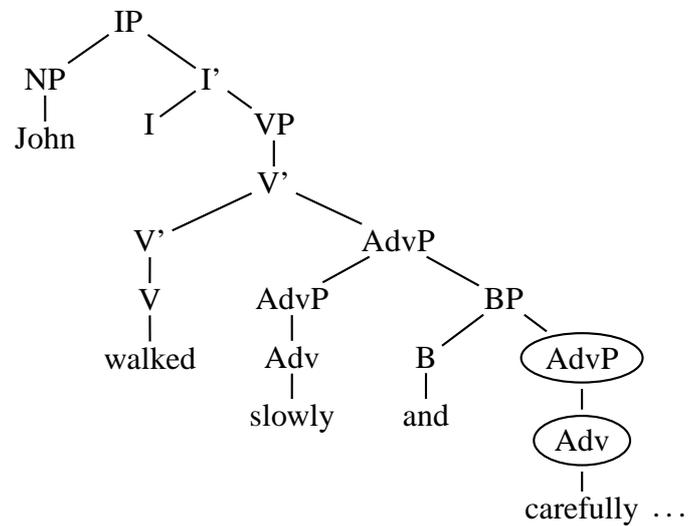


Abbildung 3.1: Parsing von (10a). Syntaktische Analyse der Koordination nach Munn (2000): Die Boolean Phrase (BP) enthält das zweite Konjunkt und ist Adjunkt zum ersten Konjunkt.

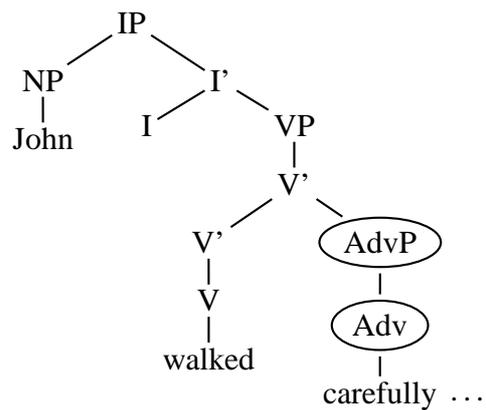


Abbildung 3.2: Parsing von (10c)

Die Ergebnisse zeigen, dass der menschliche Parser eine Verarbeitungserleichterung hat, wenn die beiden verknüpften Phrasen strukturell identisch aufgebaut sind. So wird die Nominalphrase [*NP* a tall woman] in (11a) schneller gelesen als in (11b), da in (11a) die beiden verknüpften Nominalphrasen im inneren Aufbau strukturgleich sind, nicht jedoch in (11b).

Die Autoren haben weiterhin untersucht, ob es sich beim Parallel-Struktur-Effekt um ein rein syntaktisches Priming handeln könnte. Syntaktisches Priming ist eine generelle Verarbeitungserleichterung, wenn eine strukturgleiche Phrase oder Phrasensequenz wiederholt wird (Branigan et al., 1995). Dieser Priming-Effekt ist unabhängig von Koordination. Frazier und Kollegen haben deshalb ein weiteres Lesezeit-Experiment (*Self-paced Reading*) durchgeführt, in denen Sätze ohne Koordination untersucht wurden. Sätze, in denen die Subjekt- und Objekt-Nominalphrase strukturgleich waren wie in (12a), wurden mit Sätzen verglichen, in denen dies nicht der Fall war (12b).

- (12) a. [*NP* A strange man] noticed [*NP* a tall woman].
 b. [*NP* A man] noticed [*NP* a tall woman].

Würde es sich beim Parallel-Struktur-Effekt um ein syntaktisches Priming handeln, müsste die Objekt-NP [*NP* a tall woman] in (12a) schneller verarbeitet werden als in (12b), da in (12a) die Subjekt- und Objekt-NP strukturell parallel sind, nicht jedoch in (12b).

Die parallelen Strukturen wurden in diesem Experiment nicht schneller gelesen als die nicht-parallelen. Die Verarbeitungserleichterung bei koordinierten Phrasen tritt nicht auf, wenn dieselben Phrasen als Subjekt und Objekt in einem Satz ohne Koordination auftreten. Frazier et al. (2000) schließen aus diesen Ergebnissen, dass der beobachtete Parallel-Struktur-Effekt nicht auf ein rein syntaktisches Priming zurückzuführen ist, sondern dass es sich um einen Effekt handelt, der spezifisch für Koordination ist.⁷

⁷ Ein Kritikpunkt der in diesem Zusammenhang geäußert werden kann, ist, dass in Frazier et al. (2000) der Parallel-Struktur-Effekt bei koordinierten Strukturen mit 'Eye Tracking' beobachtet wurde, während der fehlende Parallelitätseffekt in nicht-koordinierten

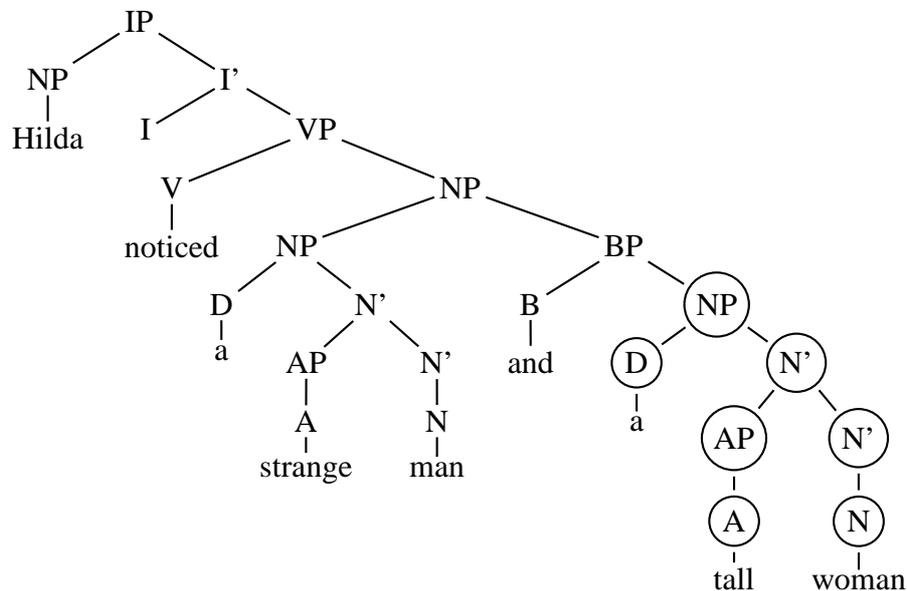


Abbildung 3.3: Syntaktische Repräsentation von (11a), Analyse der Koordination nach Munn (2000).

Die Verarbeitungserleichterung bei Parallelität im Phrasenaufbau der Konjunkte lässt sich aus den allgemeinen Parsingmechanismen nicht ableiten. Wie bereits beim Parallel-Struktur-Effekt auf sententialer Ebene und bei cross-kategorialer Koordination ausgeführt, muss die Struktur für beide Konjunkte unabhängig vom vorangehenden Kontext aufgebaut werden. So muss z.B. für das zweite Konjunkt [_{NP} a tall woman] in (11a) und in (11b) jeweils die gleiche Struktur aufgebaut werden, obwohl in (11a) beide Konjunkte im Phrasenaufbau strukturgleich sind, nicht aber in (11b). In Abbildung 3.3 ist die syntaktische Repräsentation der parallelen Variante dargestellt, in Abbildung 3.4 die nicht-parallele Variante. Die syntaktischen Knoten, die jeweils für das zweite

Strukturen mit ‘*Self-paced Reading*’ untersucht wurde, welches die weniger subtile Methode ist. Apel et al. (2007) konnten jedoch die Ergebnisse von Frazier et al. (2000) mit Eye-Tracking-Studien zum Deutschen replizieren und schließen ebenfalls, dass es sich beim Parallel-Struktur-Effekt um einen koordinationspezifischen Effekt handeln muss.

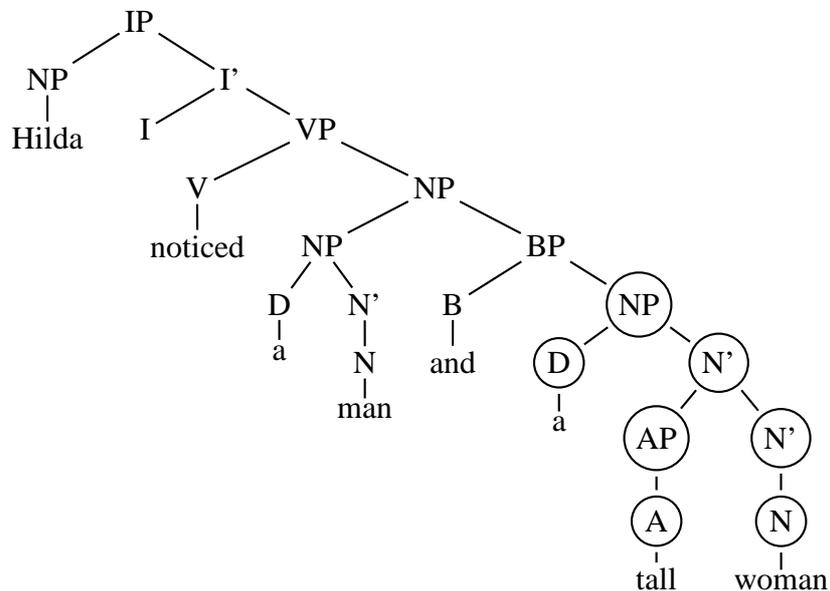


Abbildung 3.4: Syntaktische Repräsentation von (11b), Analyse der Koordination nach Munn (2000).

Konjunkt aufgebaut werden müssen, sind eingekreist. Unter Beibehaltung der etablierten *Strukturhypothese* dürften sich die Lesezeiten für [_{NP} a tall woman] in (11a) und (11b) nicht unterscheiden, da jeweils gleich viel Struktur aufgebaut werden muss.

Zusammenfassend lässt sich sagen, dass die Präferenz für strukturell parallele Konjunkte sowohl die Interpretation von Sätzen als auch die Geschwindigkeit beim Lesen beeinflusst. Die Verarbeitungserleichterung beim Lesen lässt sich auf sententialer und auf phrasaler Ebene beobachten und betrifft sogar den inneren Phrasenaufbau der Konjunkte. Der Parallel-Struktur-Effekt kann nicht durch ein syntaktisches Priming erklärt werden, sondern scheint spezifisch für Koordination zu sein. Wie ich jeweils aufgezeigt habe, können die genannten Effekte nicht aus allgemeinen Parsingmechanismen abgeleitet werden. Inwiefern sich diese Effekte aus dem koordinationspezifischen Copy- α -Modell von Frazier und Clifton (2001) ableiten lassen, wird in Abschnitt 3.4 aufgezeigt.

3.3 Präferenzen bei Ambiguitäten

Koordinierte Strukturen, die eine einfache Konjunktion wie z.B. ‘*and*’ enthalten, sind im Allgemeinen syntaktisch ambig. Die Konjunktion ist dabei die ambige Stelle. Wenn der Parser bei der inkrementellen Sprachverarbeitung von links nach rechts auf die Konjunktion stößt, gibt es mehrere Möglichkeiten, wie diese in die bestehende Struktur integriert werden kann. Die Integration der Konjunktion hängt davon ab, welcher Teil des Satzes koordiniert werden soll. In ‘*Peter kissed Mary and...*’ könnte die NP, die VP oder die IP koordiniert werden. Auch bei koordinierten Strukturen findet man sowohl lokale Ambiguitäten, die sich im Verlauf des Satzes wieder auflösen, als auch globale Ambiguitäten, die am Satzende immer noch bestehen. Der menschliche Parser hat jedoch in beiden Fällen klare Präferenzen, die in den folgenden beiden Abschnitten beschrieben werden. Es wird jeweils aufgezeigt, inwiefern sich diese Präferenzen aus allgemeinen Parsingprinzipien ableiten lassen.

3.3.1 Lokale Ambiguitäten

3.3.1.1 NP- vs. IP-Koordination

Frazier (1979) verglich die Lesezeiten von Sätzen, die zwei koordinierte NPs wie in (13a) bzw. zwei koordinierte Sätze wie in (13b) enthielten. Bei der Anbindung der Konjunktion in (13a,b) tritt eine lokale Ambiguität auf. Es kann die Objekt-NP, die Verbalphrase oder der ganze Satz koordiniert werden.

(13) a. Peter kissed [_{NP} Mary] and [_{NP} her sister] *too*.

b. [_{IP} Peter kissed Mary] and [_{IP} her sister *laughed*].

Auch die darauffolgende Phrase ist noch ambig. Die NP [*her sister*] kann entweder als Teil des Objekts von ‘*kissed*’ interpretiert werden (NP-Koordination) oder als Subjekt eines neuen Satzes (IP-Koordination). Diesen beiden Interpretationsmöglichkeiten liegen unterschiedliche syntaktische Strukturen zugrunde, welche in Abbildung 3.5 und 3.6 dargestellt sind. Erst bei der Verarbeitung des letzten Wortes kann die Ambiguität aufgelöst werden.

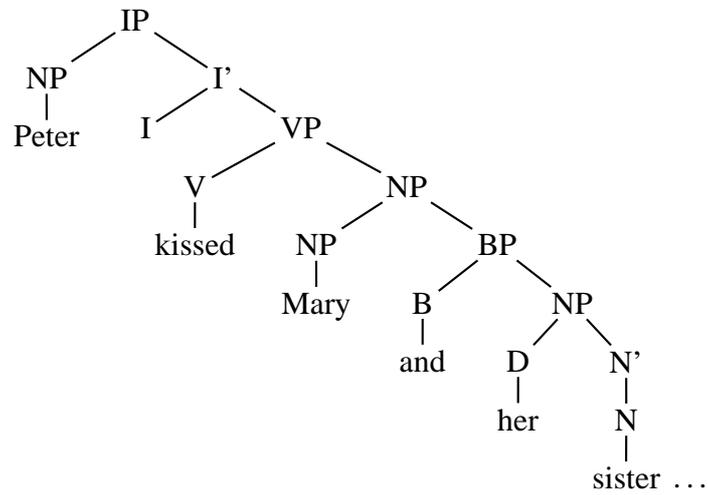


Abbildung 3.5: Analyse als NP-Koordination

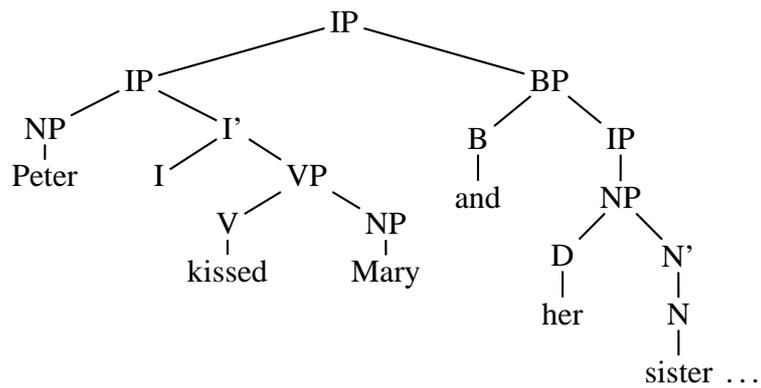


Abbildung 3.6: Analyse als IP-Koordination

Frazier fand einen Garden-Path-Effekt, d.h. signifikant längere Lesezeiten, auf dem letzten Wort *'laughed'* in (13b) im Vergleich zum letzten Wort *'too'* in (13a). Die Verarbeitungsschwierigkeiten bei der Satz-Koordination in (13b) zeigen, dass der Parser eine Präferenz hat, die Nominalphrase [*her sister*] als Teil einer koordinierten Objekt-NP wie in Abb. 3.5 zu interpretieren und nicht als Subjekt eines neuen Satzes wie in Abb. 3.6 (siehe Frazier (1987b) und Hoeks et al. (2002, 2006) zu vergleichbaren Ergebnissen im Holländischen).⁸

Wenn der Parser auf das letzte Wort *'laughed'* in (13b) trifft und sich die Ambiguität auflöst, muss die bereits aufgebaute präferierte Struktur (siehe Abbildung 3.5) reanalysiert werden in die Struktur, die in Abbildung 3.6 dargestellt ist. Diese Reanalyseprozesse kosten Zeit und führen zu längeren Lesezeiten, d.h. zu einem Garden-Path-Effekt. Trifft der Parser hingegen auf das letzte Wort *'too'* in (13a) sind keine Reanalyseprozesse erforderlich. Das letzte Wort kann direkt in die bis dahin aufgebaute Struktur integriert werden.

Die Präferenz für die Koordination der Nominalphrase kann durch allgemeine Parsingprinzipien, d.h. hier durch das Minimal-Attachment-Prinzip, erklärt werden (siehe Abschnitt 3.1.3). Dieses Parsingprinzip besagt, dass die Analyse bevorzugt wird, die weniger syntaktische Knoten erfordert. In (13) ist die Interpretation von [*her sister*] als Teil einer koordinierten Objekt-NP die ökonomischere Variante. Diese Analyse erfordert weniger syntaktische Knoten als die Analyse von [*her sister*] als Subjekt eines neuen Satzes (vgl. die Abbildungen 3.5 und 3.6). Das Minimal-Attachment-Prinzip macht daher die richtige Vorhersage für die Präferenz bei NP- vs. IP-Koordination.

⁸ Die Präferenz für NP-Koordination gegenüber Satz-Koordination wurde außerdem bei Imperativ-Konstruktionen im Englischen gefunden (Engelhardt und Ferreira, 2007).

3.3.1.2 NP- vs. VP-Koordination

In einem Lesezeit-Experiment (*Stop-making-sense-Paradigma*) haben Blodgett und Boland (1998) Sätze verglichen, die zwei koordinierte NPs wie in (14a) bzw. zwei koordinierte VPs wie in (14b) enthielten.

- (14) a. Mary sold [_{NP} vegetables] and [_{NP} seeds].
 b. Mary [_{VP} sold vegetables] and [_{VP} bought seeds].

Die lokale Ambiguität tritt bei der Konjunktion ‘*and*’ auf. Es kann die Objekt-NP, die Verbalphrase oder der ganze Satz koordiniert werden. Je nach Interpretation muss die Konjunktion ‘*and*’ auf andere Weise in die syntaktische Struktur integriert werden. Beim darauffolgenden Wort löst sich die Ambiguität auf.

Die Ergebnisse zeigen einen Garden-Path-Effekt nach der Konjunktion. Beim Lesen des Wortes ‘*bought*’ in (14b) traten größere Verarbeitungsschwierigkeiten auf als beim Lesen des Wortes ‘*seeds*’ in (14a), das sich an der gleichen Satzposition befindet. Der menschliche Parser hat offensichtlich eine Präferenz die Objekt-NP zu koordinieren im Vergleich zur Koordination der Verbalphrasen.

Wenn der Parser auf das Verb ‘*bought*’ in (14b) stößt, muss die bereits aufgebaute präferierte Struktur reanalysiert werden, d.h. die Konjunktion muss von der zuletzt verarbeiteten Phrase, der Nominalphrase, gelöst werden und höher im Baum an die Verbalphrase angebunden werden. Dieser Reanalyseprozess wird als Ursache für die längeren Lesezeiten (Garden-Path-Effekt) in (14b) betrachtet. Trifft der Parser hingegen auf das Wort ‘*seeds*’ in (14a), kann dieses direkt in die bis dahin aufgebaute Struktur integriert werden. Es sind keine Reanalyseprozesse erforderlich.

Auch diese Präferenz für NP-Koordination kann durch allgemeine Parsingprinzipien, d.h. hier durch das Late-Closure-Prinzip, erklärt werden. Das Late-Closure-Prinzip besagt, dass jedes neue Wort der Eingabesequenz in die aktuelle Phrase integriert wird. Die Konjunktion ‘*and*’ sollte demnach an die Nominalphrase angebunden werden. Dies erklärt die Verarbeitungsschwierigkeit, wenn das zweite Konjunkt eine Verbalphrase ist. Die Präferenz für NP-Koordination in (14) kann nicht durch das

Minimal-Attachment-Prinzip erklärt werden, da beide konkurrierende Analysen die gleiche Anzahl an syntaktischen Knoten erfordert. In diesem Fall greift das Late-Closure-Prinzip.

3.3.2 Globale Ambiguitäten

Carlson (2001) hat im Rahmen einer Fragebogenstudie ambige Gapping-Konstruktionen wie in (15) untersucht. Die strukturelle Ambiguität in diesen Sätzen beginnt bei der Konjunktion und löst sich bis zum Satzende nicht auf (globale Ambiguität). Auch hier gibt es eine Vielzahl an Möglichkeiten, welcher Teil des vorangehenden Satzes koordiniert werden kann.

- (15) a. [*NP* Josh] visited [*NP* the office] during the vacation and [*NP* Sarah] during the week.
- b. [*NP* Josh] visited [*NP* Marjorie] during the vacation and [*NP* Sarah] during the week.

Auch die der Konjunktion folgende Nominalphrase ‘Sarah’ ist bis zum Satzende ambig. Die ambige NP kann entweder als Subjekt (Subjekt-Lesart) oder als Objekt (Objekt-Lesart) interpretiert werden, wobei den beiden Lesarten unterschiedliche syntaktische Strukturen zugrundeliegen. Zusätzlich wurde der Einfluss paralleler Strukturen auf die Interpretation der Sätze untersucht. In manchen Sätzen, wie in (15a), war die ambige NP (‘Sarah’) nur parallel zur Subjekt-NP des ersten Satzes (‘Josh’) aufgebaut, nicht aber zur Objekt-NP (‘the office’). In den Sätzen wie in (15b) war die ambige NP sowohl zur Subjekt-NP (‘Josh’) als auch zur Objekt-NP (‘Marjorie’) parallel. Die gleichzeitige Parallelität zu Subjekt- und Objekt-NP des ersten Satzes (15b) sollte die Interpretation des Satzes nicht beeinflussen. Die ausschließliche Parallelität zur Subjekt-NP (15a) sollte die Präferenz zu Gunsten der Subjekt-Lesart beeinflussen.

Um die Präferenzen bei der Interpretation dieser Sätze zu untersuchen, wurden Probanden in einer schriftlichen Fragebogenstudie angehalten, die beste Umschreibung für einen Satz zu wählen, die Schwierigkeit beim Verstehen des Satzes zu beurteilen und weitere Bedeutungen des Satzes anzugeben. Die

Ergebnisse zeigen, dass die Objekt-Lesart in beiden Satztypen (15a,b) präferiert wird. Ist die ambige NP ausschließlich zur Subjekt-NP des ersten Satzes parallel wie in (15a), wird signifikant häufiger die Subjekt-Lesart gewählt. Dennoch bleibt die Präferenz für die Objekt-Lesart bestehen. Die Subjekt-Lesart wurde von den Probanden außerdem als schwieriger und häufiger ambig eingestuft als die Objekt-Lesart.

Weiterhin wurde in Carlson (2001) der Einfluss von Prosodie bei der Interpretation der Gapping-Konstruktionen untersucht. In einer auditorischen Fragebogenstudie wurde die Parallelität durch die Betonung der entsprechenden NPs manipuliert (siehe (16)).

(16) a. BOB insulted the guests during dinner and SAM during the dance.

b. Bob insulted the GUESTS during dinner and SAM during the dance.

In (16a) ist das Subjekt des ersten Satzes (BOB) betont, in (16b) das Objekt des ersten Satzes (GUESTS). Die ambige NP (SAM) war in beiden Satztypen betont. Wie in der schriftlichen Studie wurde die Objekt-Lesart in beiden Satztypen präferiert. Die parallele Betonung zum Subjekt des ersten Satzes erhöhte die Anzahl der Subjekt-Antworten signifikant. Dennoch blieb die Präferenz für die Objekt-Lesart bestehen. Carlson kann somit zeigen, dass die Interpretation der Sätze im Wesentlichen durch strukturelle Faktoren beeinflusst wurden.

Diese Ergebnisse werden mit der allgemein beobachteten Präferenz für ökonomische bzw. minimale Strukturen erklärt. Carlson (2001) nimmt an, dass der Subjekt-Lesart in den untersuchten Konstruktionen eine elliptische Struktur (IP-Koordination) wie in Abbildung 3.7 zugrunde liegt. Für die Objekt-Lesart wird eine VP-Koordination wie in Abbildung 3.8 angenommen. In Carlson (2001) wird die Präferenz für die Objekt-Lesart aus der Simple-Structure-Hypothese abgeleitet (siehe Abschnitt 3.1.3), das grob dem Minimal-Attachment-Prinzip entspricht. Wir erhalten einen Komplexitätseffekt, da für die VP-Koordination (Objekt-Lesart) weniger syntaktische Knoten erforderlich sind als für die elliptische IP-Koordination (Subjekt-Lesart). Nach Carlson lassen sich diese Ergebnisse aus allgemeinen Parsingprinzipien ableiten und machen keine besonderen Verarbeitungsmechanismen für Gapping erforderlich.

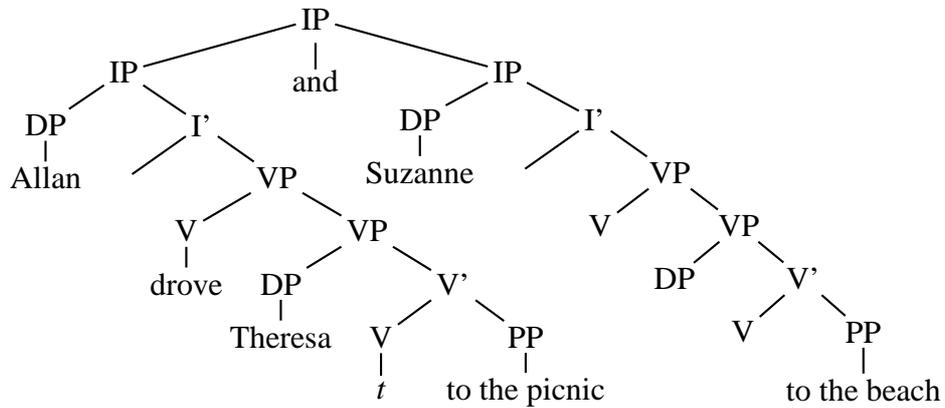


Abbildung 3.7: Subjekt-Lesart (IP-Koordination) nach Carlson (2001)

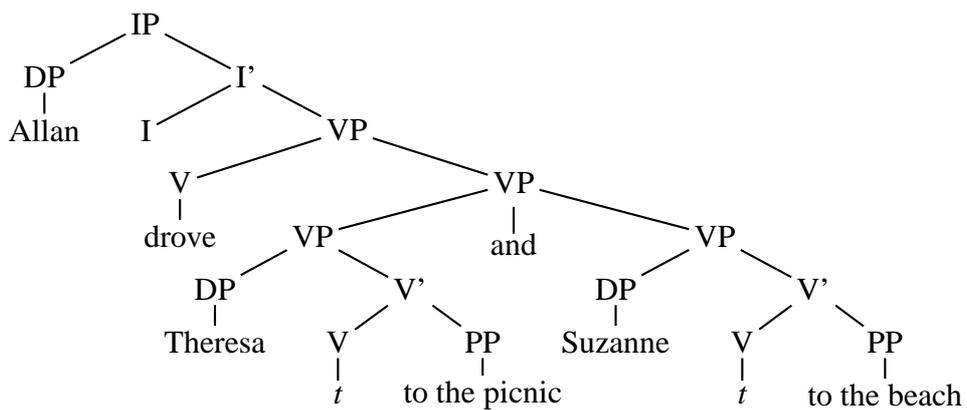


Abbildung 3.8: Objekt-Lesart (VP-Koordination) nach Carlson (2001)

In diesem Unterkapitel wurde aufgezeigt, dass sich die Präferenzen bei lokal und global ambigen koordinierten Strukturen aus diversen allgemeinen Parsingstrategien ableiten lassen. Im folgenden Abschnitt wird das *Copy- α -Modell* (Frazier und Clifton, 2001) vorgestellt, das speziell zur Verarbeitung koordinierter Strukturen und Ellipsen entwickelt wurde. Es wird aufgezeigt, inwiefern dieses Modell in der Lage ist, die in diesem Kapitel beschriebenen Effekte zu modellieren.

3.4 Das Copy- α -Modell

Das Copy- α -Modell (Frazier und Clifton, 2001) wurde für die Verarbeitung koordinierter Strukturen und Ellipsen entwickelt und basiert auf konventionellen zwei-dimensionalen syntaktischen Strukturen. Für Koordination wird eine phrasale Analyse zugrundegelegt. Frazier und Clifton (2001) nehmen an, dass der menschliche Parser zwei verschiedene Mechanismen für die Strukturbildung zur Verfügung hat. Der erste Mechanismus betrifft die inkrementelle schrittweise Verarbeitung, wie sie in Abschnitt 3.1 beschrieben wurde: Je mehr Struktur aufgebaut werden muss, desto länger ist die Verarbeitungszeit (siehe *Strukturhypothese*). Der zweite Mechanismus ist ein kostenfreier Kopiermechanismus, der bei Ellipsen und koordinierten Strukturen zum Einsatz kommt. Das Kopieren soll keinen Verarbeitungsaufwand erfordern, und zwar unabhängig davon, wieviel Struktur kopiert wird.

Allerdings schränken Frazier und Clifton die Verfügbarkeit des Kopiermechanismus auf nicht-ambige Fälle ein, d.h. wenn der linke Rand der Koordination oder Ellipse eindeutig markiert ist. Der linke Rand kann beispielsweise durch *'either'* in *either-or*-Konstruktionen markiert sein, durch eine Interrogativ-Phrase bei Fragen oder durch *'did'* bei VP-Ellipsen (Frazier und Clifton, 2001, S. 15). In diesen nicht-ambigen Fällen wird die Struktur zwischen der genannten Markierung und der Konjunktion kopiert.

Die Einschränkung des Kopiermechanismus auf nicht-ambige Konstruktionen wird begründet mit Komplexitätseffekten bei ambigen Strukturen (Gapping) und fehlenden Komplexitätseffekten bei nicht-ambigen Strukturen (z.B. *'either or'* bzw. VP-Ellipsen). Auf beide Effekte gehe ich im Folgenden ein.

Erklärungsansatz bei Gapping-Konstruktionen und gehen davon aus, dass der Kopiermechanismus bei Gapping nicht verfügbar ist.

3.4.2 Fehlende Komplexitätseffekte bei ‘*either or*’

Desweiteren wird die Einschränkung des Kopiermechanismus auf nicht-ambige Fälle durch weitere Experimente begründet. Dabei wurden Konstruktionen untersucht, bei denen der linke Rand der Koordination oder Ellipse eindeutig markiert ist, wie z.B. in ‘*either or*’ (Frazier und Clifton, 2001) bzw. VP-Ellipsen (Frazier und Clifton, 2000). Auf die Untersuchung zu ‘*either or*’ gehe ich im Folgenden ein.

Frazier und Clifton (2001) untersuchten die Verarbeitung von *either-or*-Konstruktionen wie in (19) und (20) im Rahmen einer Lesezeit-Studie (Self-paced Reading). Die Schrägstriche in (19) und (20) markieren die Segmente bei der experimentellen Darbietung. Die Testsätze unterscheiden sich durch die Position von ‘*either*’ in den (a)- und (b)-Sätzen, (c) enthält die Variante ohne ‘*either*’. Da man davon ausgeht, dass nur gleiche syntaktische Kategorien verknüpft werden, kann man die Konstruktionen in (b) als elliptisch betrachten (Reduktionsansatz nach Schwarz (1999)). Alternativ kann auch angenommen werden, dass ‘*either*’ in den (b)-Sätzen nach links bewegt wurde (Movement-Ansatz, siehe z.B. Larson, 1985; den Dikken, 2006). Frazier und Clifton legen ihrer Untersuchung den Reduktionsansatz nach Schwarz (1999) zugrunde, der Movement-Ansatz wird nicht diskutiert.

- (19) a. Mary / is looking / for either a maid / *or a cook*. (NP or NP)
 b. Mary / is looking / either for a maid / *or a cook*. (PP or PP)
 c. Mary / is looking / for a maid / *or a cook*. (lokal ambig)
- (20) a. Sam / wants either his mother / *or his father*. (NP or NP)
 b. Sam / either wants his mother / *or his father*. (VP or VP)
 c. Sam / wants his mother / *or his father*. (lokal ambig)

Nach Schwarz (1999) markiert *'either'* den linken Rand der Disjunktion. Dementsprechend enthalten die Beispiele in (19a) und (20a) zwei koordinierte NPs (*[a maid]*, *[a cook]* bzw. *[his mother]*, *[his father]*) und das Beispiel in (19b) zwei koordinierte PPs (*[for a maid]*, *[for a cook]*), wobei die Präposition *'for'* im zweiten Konjunkt getilgt wurde. In (20b) werden zwei Verbalphrasen koordiniert, wobei das Verb *'wants'* im zweiten Konjunkt getilgt ist. Die Beispiele in (19c) und (20c) enthalten kein *'either'* und sind daher bei der Verarbeitung der Konjunktion *'or'* lokal ambig. Der Vergleich von *either-or*-Konstruktionen mit lokal ambigen *or*-Konstruktionen soll zeigen, ob für ambige und nicht-ambige Strukturen verschiedene Mechanismen beim Aufbau des zweiten Konjunks verwendet werden.

Ist die Annahme der zugrundeliegenden Struktur nach Schwarz (1999) korrekt, dann müsste der Parser in (19b) mehr Struktur für das zweite Konjunkt aufbauen als in (19a). Bei inkrementellem Strukturaufbau würde daher das letzte Segment *[or a cook]* in (19b) längere Lesezeiten erfordern als in (19a). Entsprechendes gilt für (20): Das letzte Segment *[or his father]* würde in (20b) längere Lesezeiten erfordern als in (20a), da in (20b) eine Verbalphrase aufgebaut werden müsste, in (20a) dagegen nur eine NP.

Es wurden jedoch keine Verarbeitungsunterschiede in (19) und (20) zwischen den (a)- und (b)-Sätzen gefunden. Die Autoren schließen aus diesen fehlenden Komplexitätseffekten, dass in den *either-or*-Konstruktionen das zweite Konjunkt nicht auf inkrementelle Weise aufgebaut wurde, sondern, dass die Struktur des ersten Konjunks kostenlos kopiert wurde. Weiterhin schließen die Autoren, dass es für den Kopiermechanismus unerheblich ist, wieviel Struktur kopiert wird. Das Kopieren einer VP erfordert keinen höheren Verarbeitungsaufwand als das Kopieren einer NP.⁹

⁹ Die fehlenden Komplexitätseffekte bei *'either or'* können jedoch auch im Rahmen eines Movement-Ansatzes mit Hilfe der Strukturhypothese abgeleitet werden. In einem Movement-Ansatz (Larson, 1985; den Dikken, 2006) wird angenommen, dass *'either'* nach links bewegt werden kann. Es werden keine elliptischen Strukturen aufgebaut. Der Strukturaufbau des zweiten Konjunks ist unabhängig davon, ob *'either'* bewegt wurde oder nicht. Auf Basis der Strukturhypothese werden daher keine Verarbeitungsunterschiede erwartet.

In den lokal ambigen Strukturen (19c, 20c) wurden auf dem letzten Segment längere Lesezeiten gemessen als in den entsprechenden (a)- und (b)-Sätzen. Dieses Ergebnis werten Frazier und Clifton als weiteren Hinweis dafür, dass in ambigen Konstruktionen das zweite Konjunkt inkrementell aufgebaut wird, während in nicht-ambigen Konstruktionen ein kostenloser Kopiermechanismus zur Verfügung steht.

3.4.3 Probleme des Copy- α -Modells

Copy- α ist kein genereller Mechanismus für koordinierte Strukturen. Der Kopiermechanismus im Copy- α -Modell ist nur für nicht-ambige Konstruktionen verfügbar und betrifft im Wesentlichen mehrteilige Konjunktionen. Die grundlegende Strukturhypothese (je mehr Struktur, desto mehr Zeitaufwand) wird durch den Kopiermechanismus für einzelne Konstruktionen, d.h. nicht-ambige koordinierte Strukturen und Ellipsen, aufgehoben.

Durch die Einschränkung auf nicht-ambige Strukturen gibt es keinen Widerspruch zwischen den Vorhersagen des Copy- α -Modells und dem Komplexitätseffekt bei Gapping. Diese Einschränkung, wonach der linke Rand der Koordination markiert sein muss, hat jedoch auch zur Folge, dass alle in Abschnitt 3.2 beschriebenen Satzkonstruktionen nach dieser Definition ambig sind und der Kopiermechanismus nicht zur Verfügung steht. Das heißt, die beobachteten Parallelitätseffekte bei der Interpretation und alle Parallel-Struktur-Effekte auf phrasaler und sententialer Ebene können im Rahmen des Copy- α -Modells nicht erklärt werden. Desweiteren macht das Copy- α -Modell durch die Einschränkung auf nicht-ambige Konstruktionen generell keine Vorhersagen für die Präferenzen bei lokalen und globalen Ambiguitäten.

3.5 Zusammenfassung und Ziele der eigenen Untersuchungen

In diesem Kapitel wurde gezeigt, dass koordinierte Strukturen auch eine Herausforderung für die Modelle der menschlichen Satzverarbeitung darstellen. Eine Reihe von Parallelitätseffekten bei der Interpretation und bei der Leseverarbeitung auf sententialer und phrasaler Ebene können im Rahmen der bisherigen Verarbeitungsmechanismen und -modelle nicht erklärt werden, weder mit allgemeinen Parsingprinzipien noch im Rahmen des Copy- α -Modells. Präferenzen bei ambigen koordinierten Strukturen (lokale und globale Ambiguitäten) können aus verschiedenen allgemeinen Parsingprinzipien abgeleitet werden, die jedoch nur bedingt die fehlenden Komplexitätseffekte bei *'either or'* vorhersagen können. Diese fehlenden Komplexitätseffekte lassen sich aus dem Copy- α -Modell ableiten, das jedoch wiederum keine Vorhersagen für ambige Strukturen macht. Die in diesem Kapitel vorgestellten Verarbeitungseffekte und Präferenzen bilden die Basis für die eigenen Untersuchungen zur Konjunkt-Symmetrie (theoretische Modellierung, Korpusuntersuchungen, Simulationen) in den Kapiteln 4–6. Die Umgebungs-Symmetrie wird in Kapitel 7 untersucht. Im Folgenden werden die Ziele der Arbeit im Einzelnen aufgeführt.

In *Kapitel 4* soll ein Modell zur syntaktischen Verarbeitung koordinierter Strukturen beim Sprachverstehen entwickelt werden. Idealerweise sollte ein solches Modell korrekte Vorhersagen für die Lesezeiten und Präferenzen bei der Verarbeitung koordinierter Strukturen machen. Das heißt, auch die oben genannten Verarbeitungseffekte sollten sich aus dem Modell ableiten lassen. Außerdem soll das Modell die Generierung neuer testbarer Hypothesen für Lesezeiten und Präferenzen bei Koordination erlauben.

In *Kapitel 5* wird der Frage nachgegangen, ob die in diesem Kapitel beschriebenen Verarbeitungseffekte beim Sprachverstehen auch in der Sprachproduktion in Form von Korpushäufigkeiten auftreten. Dazu werden englische spontansprachliche Korpusdaten statistisch ausgewertet. Diese Ergebnisse erlauben Rückschlüsse auf das Verhältnis von Sprachverstehens- und Sprachproduktionsmechanismen. Die Korpusdaten, die in Kapitel 5

ausgewertet werden, verwende ich außerdem als Trainingsdaten für die Simulationen in Kapitel 6. Die quantitative Verteilung der koordinierten Strukturen in den Trainingsdaten hat Implikationen für die Vorhersagbarkeit von bestimmten Strukturen, für ihre Lernbarkeit und für die Interpretation der Simulationsergebnisse.

In *Kapitel 6* soll die syntaktische Verarbeitung koordinierter Strukturen im Rahmen eines kognitiven Systems simuliert werden. Dazu werden künstliche neuronale Netze trainiert, selbständig strukturelle Regularitäten und Parsingmechanismen aus Korpora zu lernen. Dem Netz wird hierfür keinerlei linguistisches Wissen beigebracht, es wird ausschließlich mit Sequenzen von Wortarten trainiert. Da entsprechende Simulationen in der Vergangenheit bislang nur auf Basis von stark vereinfachten oder künstlichen Trainingsdaten durchgeführt wurden, soll geprüft werden, ob dies auch auf Basis von echten spontansprachlichen Korpusdaten möglich ist. Nach abgeschlossener Trainingsphase soll das Verhalten des Netzes im Hinblick auf die beschriebenen Verarbeitungseffekte mit der menschlichen Leseverarbeitung verglichen werden.

In *Kapitel 7* wird die Relation der Konjunkte zu ihrer syntaktischen Umgebung (Umgebungs-Symmetrie) exemplarisch anhand der asymmetrischen Kongruenz im Deutschen untersucht. Es soll herausgefunden werden, ob die asymmetrische Kongruenz möglicherweise ein Verarbeitungseffekt ist und überhaupt nicht mit syntaktischen Mechanismen erklärt werden muss. Zur Klärung dieser Frage werden vier verschiedene Datentypen untersucht: geschriebene und gesprochene Korpusdaten sowie inkrementelle Grammatikalitätsurteile und Lesezeiten.

Kapitel 4

Das Iterationsmodell

In diesem Kapitel wird ein eigenes Modell – das *Iterationsmodell* - zur syntaktischen Verarbeitung koordinierter Strukturen beim Sprachverstehen vorgestellt. Das Modell macht Vorhersagen für Lesezeiten und Präferenzen und ist, im Gegensatz zu bestehenden Verarbeitungsmodellen, in der Lage, alle in Kapitel 3 beschriebenen Verarbeitungseffekte zu erklären. Das Kapitel ist wie folgt strukturiert: Nachdem das Iterationsmodell in Abschnitt 4.1 beschrieben wird, zeige ich in Abschnitt 4.2, wie sich die Verarbeitungseffekte aus der Literatur (siehe Kapitel 3) aus dem Iterationsmodell ableiten lassen. In Abschnitt 4.3 wird gezeigt, dass sich die Mechanismen des Iterationsmodells auch auf die Sprachverständnisleistung bei agrammatischer Aphasie übertragen lassen. In der abschließenden Zusammenfassung (Abschnitt 4.4) wird ein Vergleich mit dem Copy- α -Modell präsentiert sowie die Implikationen für die Syntax der Koordination aufgezeigt.

4.1 Das Modell

Das Iterationsmodell basiert auf einem neuen Parsingmechanismus, der die strukturellen Redundanzen in den Konjunkten gezielt ausnutzt. Dies wird an folgendem Beispielsatz in (1) illustriert, der strukturgleiche Konjunkte enthält. Die syntaktische Repräsentation des Satzes in (1) ist in Abbildung 4.1 dargestellt,¹

¹ Das flektierte Verb wird beim Parsing im Allgemeinen in der Form repräsentiert, wie es im Input vorgegeben ist. Die Flexionsendung wird nicht abgespalten. In weiteren

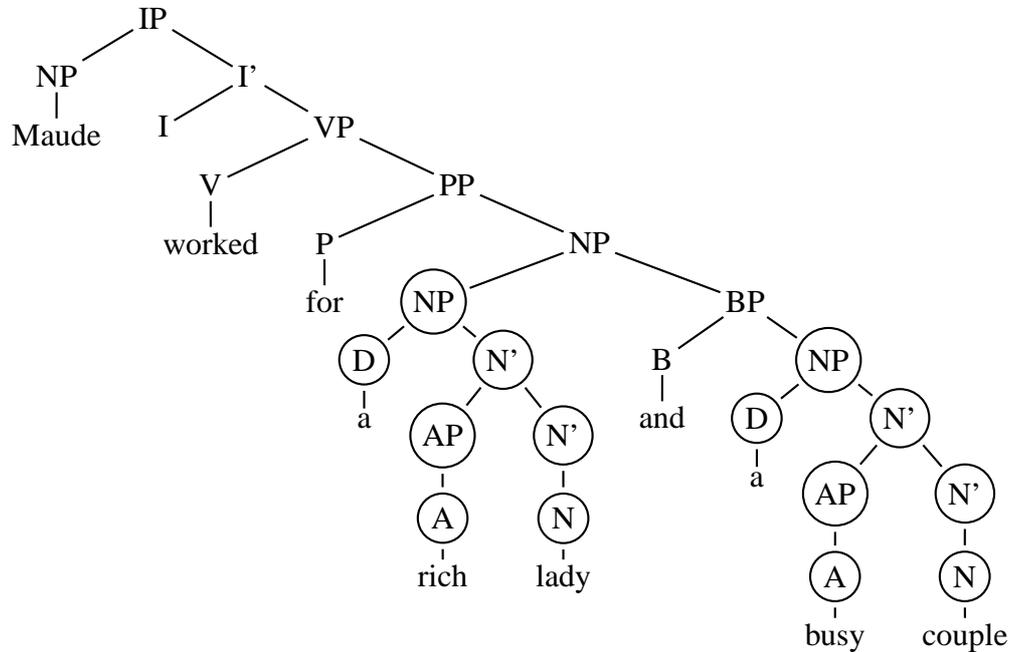


Abbildung 4.1: Syntaktische Repräsentation von (1). Analyse der Koordination nach Munn (2000): Die Boolean Phrase (*BP*) enthält das zweite Konjunkt und ist Adjunkt zum ersten Konjunkt.

wobei die übereinstimmenden Knoten der Konjunkte eingekreist sind.² Wie können diese Redundanzen während des Parsingprozesses genutzt werden?

(1) Maude worked for [*NP* a rich lady] and [*NP* a busy couple].

Im Iterationsmodell wird davon ausgegangen, dass während des Sprachverstehens eine strukturelle Repräsentation des Satzes nach allgemeinen syntaktischen Parsingprinzipien inkrementell aufgebaut wird, bis der Parser auf

Verarbeitungsschritten erfolgt ein Abgleich der Flexionsmerkmale mit dem Subjekt (siehe z.B. Crocker, 1996).

² Die genaue Analyse der koordinierten Struktur ist an dieser Stelle unerheblich. Relevant ist die Symmetrie zwischen den Konjunkten.

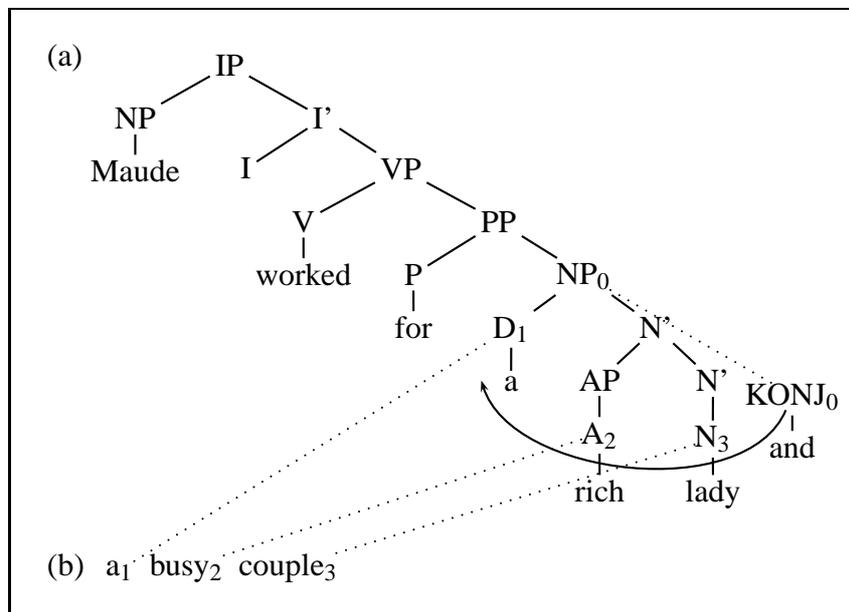


Abbildung 4.2: Parsing von (1) im Iterationsmodell

eine koordinierende Konjunktion stößt. Die Verarbeitung der Konjunktion löst einen Iterationsmechanismus aus, wobei der Parser an den Beginn des ersten Konjunks zurückspringt. Bei der Verarbeitung des Satzes (1) beispielsweise wird die Struktur für das Satzfragment ‘*Maude worked for a rich lady*’ inkrementell aufgebaut. Wenn der Parser auf die Konjunktion ‘*and*’ stößt, springt er an den Beginn der Nominalphrase [*NP a rich lady*] zurück. Wie der Parser den Beginn des ersten Konjunks lokalisiert, wird am Ende des Abschnitts beschrieben (siehe *Iterationsprinzip*).

Nach diesem Feedback-Prozess geht der Parser wieder inkrementell von links nach rechts vor und nutzt dabei die Struktur, die bereits aufgebaut wurde (siehe Abbildung 4.2). Da in (1) die Struktur der beiden Konjunkte identisch ist, muss für das zweite Konjunkt keine neue Struktur aufgebaut werden. Stattdessen wird die Struktur des ersten Konjunks wiederverwendet. Die Redundanz in den Konjunkten erstreckt sich jedoch nur auf die Struktur und

nicht auf die lexikalischen Elemente. Das heißt, bei der Verarbeitung des zweiten Konjunks muss zwar keine neue Struktur aufgebaut werden, dennoch müssen die lexikalischen Elemente in die bestehende Struktur eingefügt werden. Die resultierende strukturelle Repräsentation ist in Abbildung 4.2 dargestellt, wobei die Indizes an den Knoten angeben, wo die Wörter des zweiten Konjunks in (b) an die Struktur in (a) angebunden werden. So wird beispielsweise das Adjektiv ‘*busy*’ im zweiten Konjunkt an denselben A-Knoten angebunden wie das Adjektiv ‘*rich*’ im ersten Konjunkt.

Im Gegensatz zu den zwei-dimensionalen Baumstrukturen, wie z.B. in Abbildung 4.1 dargestellt, handelt es sich bei den Datenstrukturen, die sich aus dem vorgeschlagenen Iterationsmechanismus ergeben, um drei-dimensionale Bäume, in denen sich die Konjunkte redundante Knoten teilen. Jedes weitere Konjunkt, d.h. jede weitere Iterationsschleife, ergibt eine weitere Ebene innerhalb der dritten Dimension. In unserem Beispiel stellt Abbildung 4.2 (b) eine weitere Ebene in der dritten Dimension dar. Die Indizes an den Knoten zeigen, wie die beiden Ebenen 4.2 (a) und (b) miteinander verbunden sind. Zur Veranschaulichung sind die Verbindungen der beiden Ebenen in der vorliegenden Abbildung durch gepunktete Linien gekennzeichnet. In den nachfolgenden Abbildungen wird nur noch die Markierung durch Indizes verwendet.

Genaugenommen ist ein drei-dimensionaler Baum ein gerichteter Graph, in dem es neben einer Präzedenz-Relation (links-rechts) und einer Dominanz-Relation (oben-unten) noch zusätzlich eine ‘Behindance’-Relation (vorne-hinten) gibt (zu ‘Behindance’ siehe De Vries, 2003). Aus Gründen der Übersichtlichkeit ist in den Abbildungen die Ebene (b) graphisch unter der Ebene (a) dargestellt, obwohl sich Ebene (b) eigentlich *hinter* der Ebene (a) befindet. Sätze, die keine koordinierte Struktur enthalten, benötigen keine dritte Dimension und die strukturelle Repräsentation bleibt zwei-dimensional.

Die beschriebene Prozedur führt zu einem Parsingmodell für koordinierte Strukturen, das vier verschiedene Phasen bei der Verarbeitung unterscheidet (siehe Steiner, 2003a,b, 2005). Diese werden im Folgenden beschrieben.

Das Iterationsmodell

Phase 1:

Inkrementeller Strukturaufbau nach allgemeinen Parsingprinzipien bis der Parser auf eine koordinierende Konjunktion stößt.

Phase 2:

Feedbackprozess: Die Verarbeitung der Konjunktion löst eine Iterationsschleife aus. Der Parser springt innerhalb der bereits aufgebauten Struktur zurück, die Zielposition ist der Beginn des ersten Konjunks (siehe *Iterationsprinzip*).

Phase 3:

Iterationsphase: Die redundanten Anteile der existierenden Struktur werden wiederverwendet, die nicht-redundanten Anteile der Struktur werden neu aufgebaut.

Phase 4:

Nach Abschluss der Iterationsphase: Strukturaufbau nach allgemeinen Parsingprinzipien.

Im Iterationsmodell wird angenommen, dass der Parser bei der Verarbeitung der Konjunktion immer eine Iterationsschleife beginnt (siehe *Phase 2*). Wie es bei der inkrementellen Verarbeitung der Fall ist, wird die Konjunktion sofort verarbeitet, auch wenn es zu diesem Zeitpunkt noch mehrere Anbindungsmöglichkeiten geben kann (lokale Ambiguität). Im ambigen Falle, d.h. wenn die bis dahin aufgebaute Struktur aus mehreren Phrasen besteht, an die der Parser springen könnte, hat der Parser eine Präferenz für die kleinste Iterationsschleife. Dies ist im folgenden *Iterationsprinzip* formuliert.³

Iterationsprinzip: “Springe zurück an den Beginn einer Phrase innerhalb der bisher aufgebauten Struktur. Präferiere im ambigen Fall die kleinere Iterationsschleife.”

³ Das Iterationsprinzip entspricht grob einer allgemeinen Präferenz für ‘Recency’ (Gibson, Pearlmutter, Canseco-Gonzalez und Hickok, 1996) bzw. dem Late-Closure-Prinzip im Garden-Path-Modell (Frazier, 1979). Das Iterationsprinzip ist gewissermaßen eine Übertragung des Recency-Prinzips auf drei-dimensionale Datenstrukturen.

Die Konjunktion wird innerhalb der bis dahin aufgebauten Struktur an die Phrase angebunden, an die der Parser gesprungen ist.⁴ Wenn sich die Ambiguität auflöst, muss die Iterationsschleife und die Anbindung der Konjunktion im Falle einer nicht-präferierten Fortsetzung des Satzes korrigiert werden (Reanalyse).

Das in diesem Abschnitt beschriebene Iterationsmodell für koordinierte Strukturen basiert auf drei-dimensionalen syntaktischen Strukturen, wie sie im Rahmen von mehrdimensionalen Modelle vorgeschlagen werden (siehe Kapitel 2, Abschnitt 2.3.3). Im Vergleich zu zwei-dimensionalen Strukturen, wie sie im Rahmen einer phrasalen Analyse oder im Reduktionsansatz angenommen werden, sind die hier verwendeten drei-dimensionalen Strukturen maximal ökonomisch. Das heißt, sie kommen mit der kleinstmöglichen Anzahl an syntaktischen Knoten aus. Das Iterationsmodell macht Vorhersagen für Lesezeiten und Präferenzen auf Basis der etablierten Strukturhypothese (siehe Abschnitt 3.1), d.h. je mehr Struktur aufgebaut wird, desto länger sind die Lesezeiten. Im Folgenden wird beschrieben, wie die Verarbeitungseffekte aus der Literatur (siehe Kapitel 3) im Rahmen des Iterationsmodells erklärt werden können.

4.2 Modellierung der Verarbeitungseffekte

4.2.1 Parallel-Struktur-Effekt auf phrasaler Ebene

4.2.1.1 Parallel-Struktur-Effekt im Phrasenaufbau

Im Folgenden wird beschrieben, wie sich der Parallel-Struktur-Effekt im Phrasenaufbau aus dem Iterationsmodell ergibt. Die entsprechenden Beispielsätze aus Kapitel 3 (Abschnitt 3.2.3.2) sind in (2) nochmal wiederholt. Frazier et al. (2000) haben beobachtet, dass strukturgleiche Konjunkte zu einer Verarbeitungserleichterung für das zweite Konjunkt führen. Das zweite Konjunkt [*NP a tall woman*] in (2a) wird beispielsweise schneller gelesen als in (2b).

⁴ Genau genommen wird die Konjunktion an den sog. ‘*Splitting Node*’ angebunden. Das ist der Knoten, der weitere Knoten auf verschiedenen hintereinanderliegenden Ebenen dominiert. Falls es mehrere ‘*Splitting Nodes*’ gibt, wird die Konjunktion an den Phrasenknoten angebunden, der die verschiedenen ‘*Splitting Nodes*’ dominiert.

- (2) a. Hilda noticed [_{NP} a strange man] and [_{NP} a tall woman].
b. Hilda noticed [_{NP} a man] and [_{NP} a tall woman].

Nach dem Iterationsmodell werden die Sätze in (2) bis zur Konjunktion ‘and’ inkrementell aufgebaut, der Parser springt zurück an den Beginn der Objekt-NP und versucht, die bis dahin aufgebaute Struktur wiederzuverwenden.⁵ Sind die Konjunkte strukturgleich, wie in (2a), müssen bei der Verarbeitung des zweiten Konjunks lediglich die Wörter an die bestehende Struktur angebunden werden (siehe Abbildung 4.3). Die komplette NP-Struktur kann wiederverwendet werden. Sind die Konjunkte nicht strukturgleich, wie in (2b), müssen die nicht-redundanten Anteile der Struktur neu aufgebaut werden. Die resultierende Struktur für (2b) ist in Abbildung 4.4 dargestellt. In beiden Abbildungen ist die Ebene (b) eine weitere Ebene in der dritten Dimension und befindet sich eigentlich *hinter* der Ebene (a). Die Indizes markieren, wie die beiden Ebenen verbunden sind.

Während das Wort ‘tall’ in der parallelen Variante direkt an die existierende Struktur angebunden werden kann (siehe Abb. 4.3), muss in der nicht-parallelen Variante für dieses Wort eine Adjektiv-Phrase aufgebaut werden (siehe Abb. 4.4). Der Parser muss also in (2a) weniger Struktur für das zweite Konjunkt aufbauen als in (2b), was auf Basis der Strukturhypothese die schnellere Verarbeitung des parallel-strukturierten zweiten Konjunks erklärt.

An dieser Stelle sei angemerkt, dass sich erst durch die Verwendung der drei-dimensionalen Strukturen Unterschiede im Strukturaufbau des parallelen und nicht-parallelen zweiten Konjunks [_{NP} a tall woman] in (2a) vs. (2b) ergeben. Wie in Kapitel 3 (Abschnitt 3.2.3.2) ausgeführt, würde die Strukturhypothese auf Basis von zwei-dimensionalen Baumstrukturen keine Verarbeitungsunterschiede zwischen parallelen und nicht-parallelen zweiten Konjunkten vorhersagen. Erst durch die Verwendung der drei-dimensionalen Strukturen macht die Strukturhypothese die korrekte Vorhersage.

⁵ Der Beginn der Objekt-NP ist nach dem Iterationsprinzip die präferierte Zielposition des Parsers.

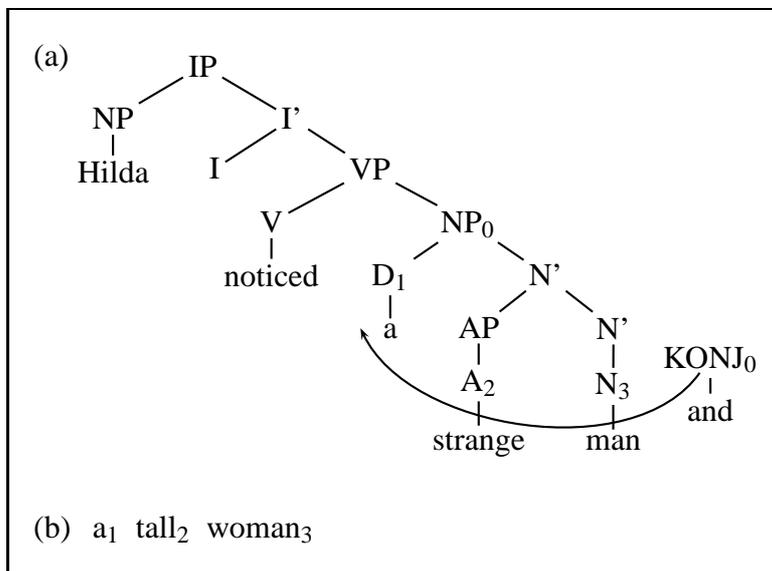


Abbildung 4.3: Parsing von (2a) im Iterationsmodell

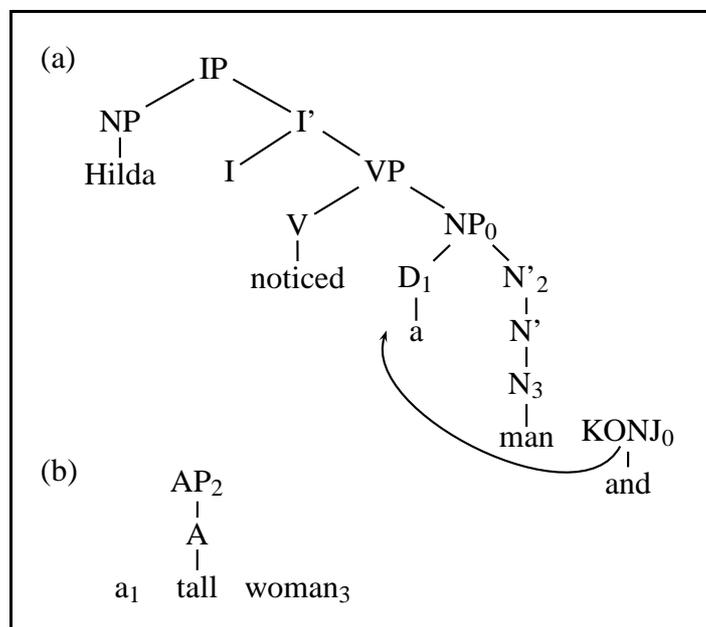


Abbildung 4.4: Parsing von (2b) im Iterationsmodell

Das Copy- α -Modell kann die Verarbeitung der Sätze in (2) ebenfalls nicht erklären. Da es sich dabei um ambige Strukturen⁶ handelt, ist der Kopiermechanismus für diese Sätze nicht verfügbar und es wird keine Verarbeitungserleichterung für parallele Konjunkte vorhergesagt.

4.2.1.2 Cross-kategoriale Koordination

Auch die Präferenz für die Verknüpfung von gleichen syntaktischen Kategorien gegenüber cross-kategorialer Koordination lässt sich aus dem Iterationsmodell ableiten. Die Beispielsätze aus Kapitel 3 (Abschnitt 3.2.3.1) werden in (3) wiederholt. Frazier et al. (2000) haben beobachtet, dass die Adverbphrase [*AdvP* carefully] in (3a) schneller gelesen wird als in (3c). Diese Verarbeitungserleichterung trifft jedoch nicht auf die Präpositionalphrase [*PP* with great care] in (3b) vs. (3d) zu.

- (3) a. John walked [*AdvP* slowly] and [*AdvP* carefully], avoiding the broken glass.
- b. John walked [*AdvP* slowly] and [*PP* with great care], avoiding the broken glass.
- c. John walked [*AdvP* carefully], avoiding the broken glass.
- d. John walked [*PP* with great care], avoiding the broken glass.

Im Iterationsmodell werden die Strukturen für die Sätze in (3a,b) inkrementell bis zur Konjunktion ‘and’ aufgebaut. Der Parser springt zurück an den Beginn des ersten Konjunks [*AdvP* slowly] und versucht, die bestehende Struktur soweit als möglich wiederzuverwenden (siehe Abbildungen 4.5 und 4.7).⁷ Die Strukturen für die Sätze in (3c,d) werden komplett inkrementell aufgebaut, da diese Sätze keine Konjunktion enthalten. Es wird daher keine Iterationsschleife ausgelöst (siehe Abbildungen 4.6 und 4.8) und die Strukturen bleiben zwei-dimensional.

⁶ Durch die Verwendung der Konjunktion ‘and’ ist der linke Rand der Koordination nicht markiert. Die Ambiguität bezieht sich auf die Anbindung der Konjunktion.

⁷ Der Beginn der AdvP ‘slowly’ ist nach dem Iterationsprinzip die präferierte Zielposition des Parsers.

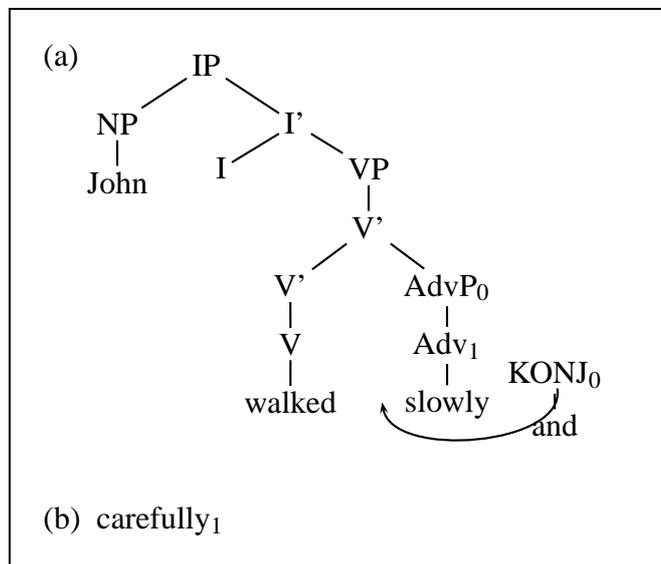


Abbildung 4.5: Parsing von (3a) im Iterationsmodell

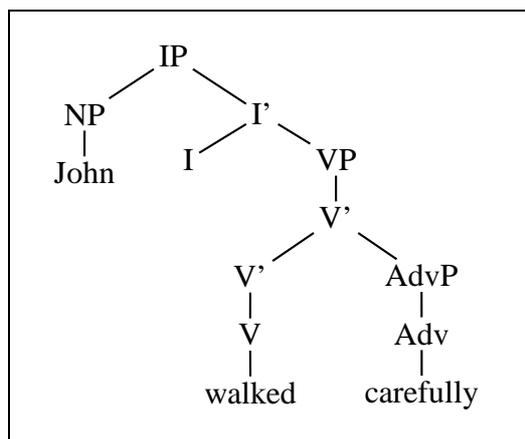


Abbildung 4.6: Parsing von (3c) im Iterationsmodell

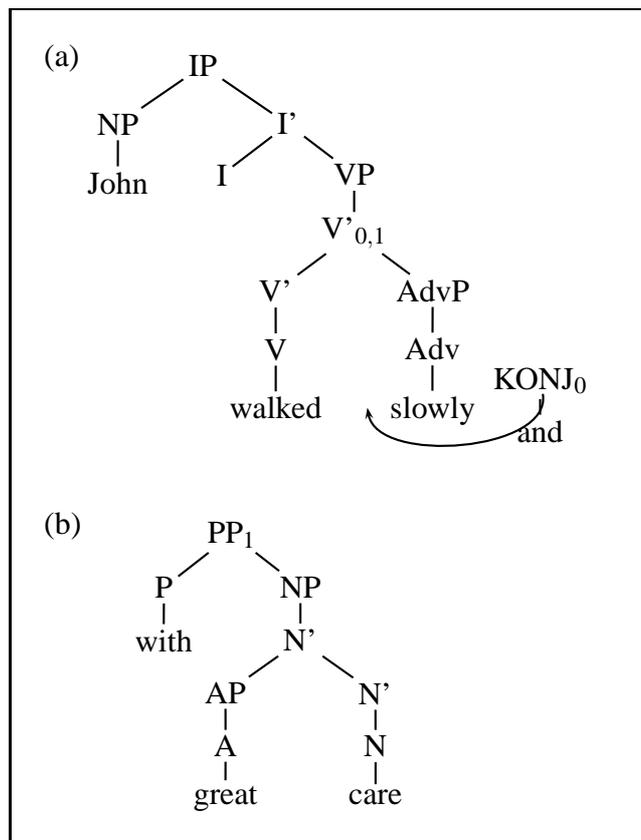


Abbildung 4.7: Parsing von (3b) im Iterationsmodell

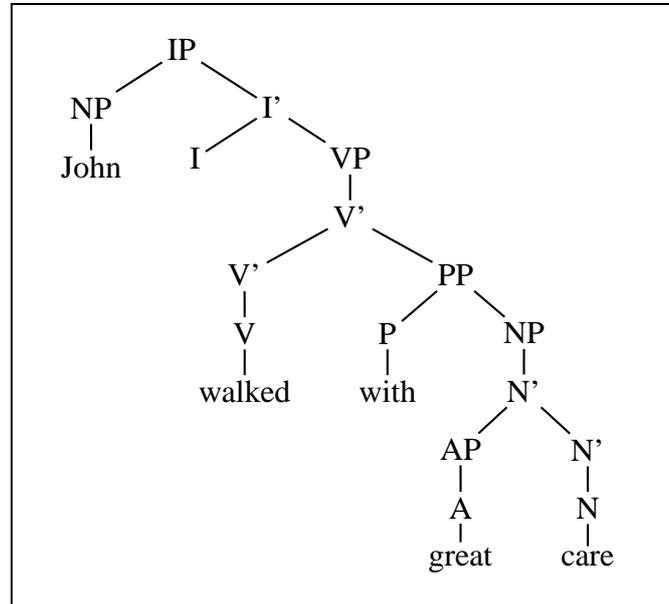


Abbildung 4.8: Parsing von (3d) im Iterationsmodell

Wie in Abbildung 4.5 zu sehen ist, muss das Wort ‘*carefully*’ in (3a) lediglich an die bestehende Struktur angebunden werden, d.h. an den *Adv*-Knoten. Die Struktur der Adverb-Phrase kann wiederverwendet werden. In (3c) muss für ‘*carefully*’ hingegen die Adverbphrase neu aufgebaut werden (siehe Abbildung 4.6). Die schnellere Verarbeitung von [*AdvP* *carefully*] in (3a) im Vergleich zu (3c) folgt somit aus der Strukturhypothese, da in (3a) weniger Struktur für das Wort ‘*carefully*’ aufgebaut werden muss als in (3c).

Die Verarbeitung der Präpositionalphrase [*PP* *with great care*] in (3b) hingegen profitiert nicht von der bereits aufgebauten Struktur des ersten Konjunks, da die beiden Konjunkte keinerlei syntaktische Redundanzen enthalten. Der Parser kann die Struktur des ersten Konjunks nicht wiederverwenden, die Präpositionalphrase muss komplett neu aufgebaut werden (siehe Abbildung 4.7). Dabei bildet die *PP*-Struktur in Abb. 4.7 (b) eine weitere Ebene in der dritten Dimension. Die *PP* befindet sich sozusagen *hinter* der *AdvP*-Struktur und wird, wie auch die *AdvP*, an denselben *V'*-Knoten innerhalb der bestehenden Struktur angebunden.

Die Präpositionalphrase [*PP* *with great care*] muss auch in (3d) vollständig inkrementell aufgebaut werden, weil dieser Satz gar keine Koordination enthält

(siehe Abbildung 4.8). Somit lässt sich auch erklären, warum die *PP* in (3b) nicht schneller verarbeitet wird als in (3d). In beiden Sätzen muss für die Präpositionalphrase [*PP* with great care] die komplette Struktur aufgebaut werden.

Auch hier ergeben sich die Unterschiede im Strukturaufbau erst durch die Verwendung der drei-dimensionalen Baumstrukturen. Wie bereits in Kapitel 3 (Abschnitt 3.2.3.1) ausgeführt, würde die Strukturhypothese auf Basis der zwei-dimensionalen Datenstrukturen keinen Verarbeitungsunterschied für die Adverbphrase [*AdvP* carefully] in (3a) vs. (3c) vorhersagen. Erst durch die Verwendung der drei-dimensionalen Strukturen macht die Strukturhypothese die korrekte Vorhersage.

Das Copy- α -Modell kann die Verarbeitung der Sätze in (3) ebenfalls nicht erklären. Da es sich dabei um ambige Strukturen handelt, ist der Kopiermechanismus für diese Sätze nicht verfügbar und es wird keine Verarbeitungserleichterung für die parallelen Konjunkte vorhergesagt.

4.2.2 Parallel-Struktur-Effekt auf sententialer Ebene

Die Präferenz für parallele Strukturen in koordinierten Sätzen ergibt sich aus dem Iterationsmodell analog zum Parallel-Struktur-Effekt auf phrasaler Ebene. Je größer die strukturelle Redundanz in den Konjunkten, desto mehr Struktur kann bei der Verarbeitung des zweiten Konjunks wiederverwendet werden. Das gilt für die Koordination von Phrasen ebenso wie für die Koordination von Sätzen.

Zwei relevante Satzkonstruktionen aus Kapitel 3 (Abschnitt 3.2.2) werden in (4) und (5) wiederholt. Frazier et al. (1984) haben beobachtet, dass das zweite Konjunkt [*IP* the king hugged the queen] in (4a) schneller gelesen wird als in (4b). Und das zweite Konjunkt [*IP* Sue believed Jim's stories were fictitious] wird in (5a) schneller gelesen als in (5b).

- (4) a. [*IP* The prince kissed the princess] and
 [*IP* the king hugged the queen]. (Aktiv + Aktiv)
- b. [*IP* The princess was kissed by the prince] and
 [*IP* the king hugged the queen]. (Passiv + Aktiv)

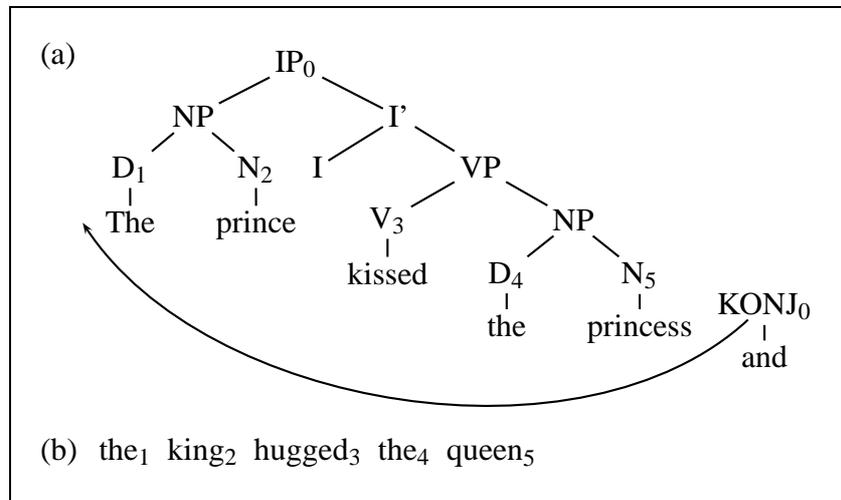


Abbildung 4.9: Parsing von (4a) im Iterationsmodell

- (5) a. [_{IP} Jim believed all Tom's stories were literally true] and
 [_{IP} Sue believed Jim's stories were fictitious].
 (nonminimal + nonminimal)
- b. [_{IP} Jim believed all Tom's stories] and
 [_{IP} Sue believed Jim's stories were fictitious].
 (minimal + nonminimal)

Wie die Sätze in (4) im Iterationsmodell verarbeitet werden, ist in den Abbildungen 4.9 und 4.10 dargestellt. Im Iterationsmodell werden die Sätze bis zur Konjunktion 'and' inkrementell aufgebaut, der Parser springt zurück an den Beginn des gesamten Satzes und versucht, die bis dahin aufgebaute Struktur wiederzuverwenden.⁸ Sind die Konjunkte strukturgleich, wie in (4a), müssen bei der Verarbeitung des zweiten Konjunks lediglich die Wörter an die bestehende Struktur angebunden werden (siehe Abbildung 4.9). Die komplette Satzstruktur

⁸ Der Beginn des Gesamtsatzes ist nach dem Iterationsprinzip nicht die präferierte Zielposition. In (4) beispielsweise springt der Parser nach dem Iterationsprinzip zuerst an den Beginn der letzten Phrase. Bei der Verarbeitung der folgenden Subjekt-NP [*the king*] muss die Iterationsschleife jedoch korrigiert werden und der Parser springt daraufhin an den Beginn des Gesamtsatzes (Reanalyse).

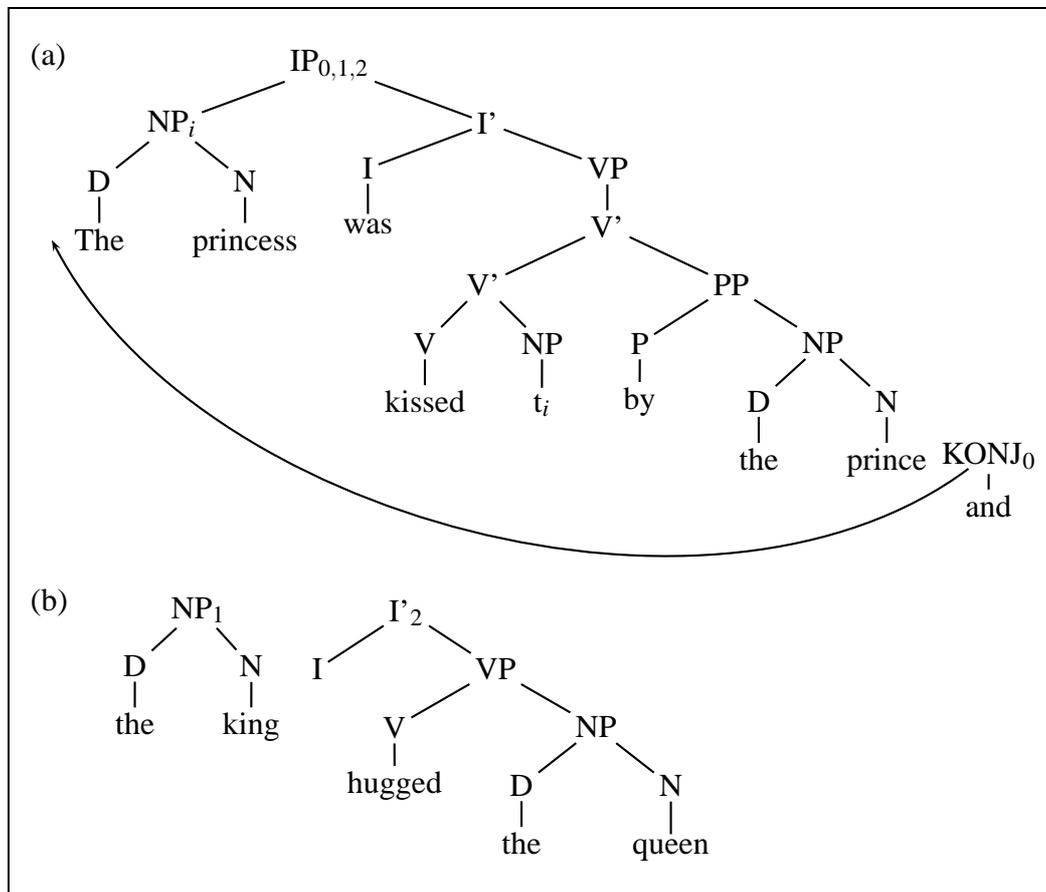


Abbildung 4.10: Parsing von (4b) im Iterationsmodell

kann wiederverwendet werden. Sind die Konjunkte nicht strukturgleich, wie in (4b), müssen die nicht-redundanten Anteile der Struktur neu aufgebaut werden. Das heißt, in (4b) muss die Struktur für die Subjekt-NP [NP the king] und für das Segment [I' hugged the queen] neu aufgebaut werden (siehe Abbildung 4.10). Der einzige syntaktische Knoten, der von beiden Konjunkten geteilt werden kann, ist der IP-Knoten. Die schnelleren Lesezeiten für das zweite Konjunkt [IP the king hugged the queen] in (4a) im Vergleich zu (4b) folgen aus der Strukturhypothese, da für das zweite Konjunkt in (4a) weniger Struktur aufgebaut werden muss als in (4b) (siehe Abbildung 4.9 vs. 4.10).

Der Parallel-Struktur-Effekt für die Satzkonstruktionen mit lokalen Ambiguitäten in (5) kann auf ähnliche Weise erklärt werden. In den Abbildungen 4.11 und 4.12 ist dargestellt, wie die Sätze in (5) im Iterationsmodell verarbeitet werden. Es muss wesentlich weniger Struktur für das zweite Konjunkt [*IP* Sue believed Jim's stories were fictitious] in (5a) aufgebaut werden als in (5b) (siehe Abbildung 4.11 vs. 4.12). Dies erklärt die schnelleren Lesezeiten für das zweite Konjunkt in (5a) versus (5b).

In diesen Satzkonstruktionen kommt jedoch hinzu, dass das zweite Konjunkt jeweils eine lokale Ambiguität enthält, wodurch der Parallel-Struktur-Effekt bei nicht-präferierten zweiten Konjunkten ('nonminimal attachment') wie in (5) vermutlich verstärkt wird. Wenn der Strukturaufbau des zweiten Konjunkts von vornherein schon lange dauert, weil dies beispielsweise Reanalyseprozesse erfordert, dann ist die zeitliche Differenz zum Recycling-Prozess besonders groß und somit der Parallel-Struktur-Effekt besonders deutlich.

Auch auf sententialer Ebene ergeben sich die Unterschiede im Strukturaufbau von parallelen und nicht-parallelen zweiten Konjunkten erst durch die Verwendung der drei-dimensionalen Datenstrukturen. Wie in Kapitel 3 (Abschnitt 3.2.2) ausgeführt, würde die Strukturhypothese auf Basis von zwei-dimensionalen Baumstrukturen keinen Verarbeitungsunterschied zwischen parallelen und nicht-parallelen zweiten Konjunkten vorhersagen. Diese Vorhersage ist unabhängig davon, ob das zweite Konjunkt eine lokale Ambiguität enthält oder nicht. Erst durch die Verwendung der drei-dimensionalen Strukturen macht die Strukturhypothese die korrekte Vorhersage.

Das Copy- α -Modell kann die Verarbeitung der Sätze in (4) und (5) ebenfalls nicht erklären. Da es sich dabei um ambige Strukturen handelt, ist der Kopiermechanismus für diese Sätze nicht verfügbar und es wird keine Verarbeitungserleichterung für die parallelen Konjunkte erwartet.

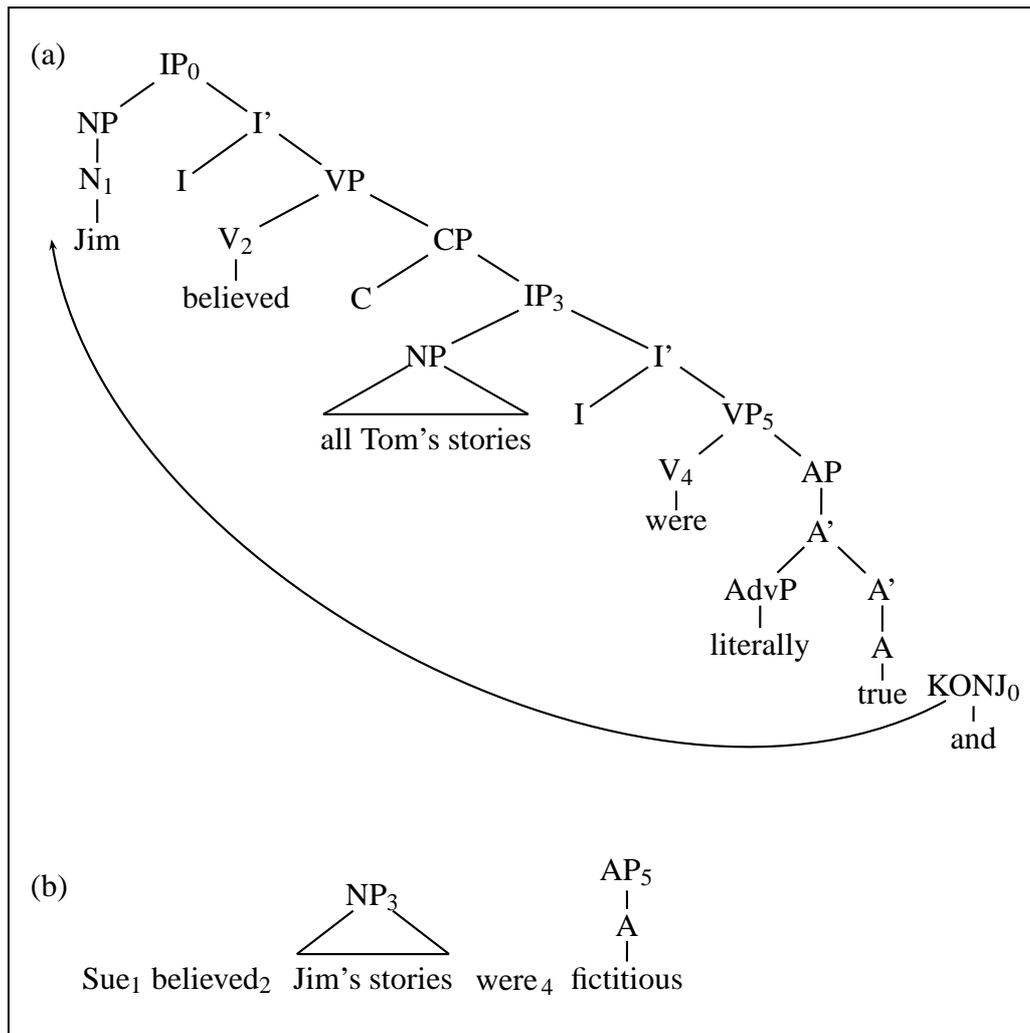


Abbildung 4.11: Parsing von (5a) im Iterationsmodell

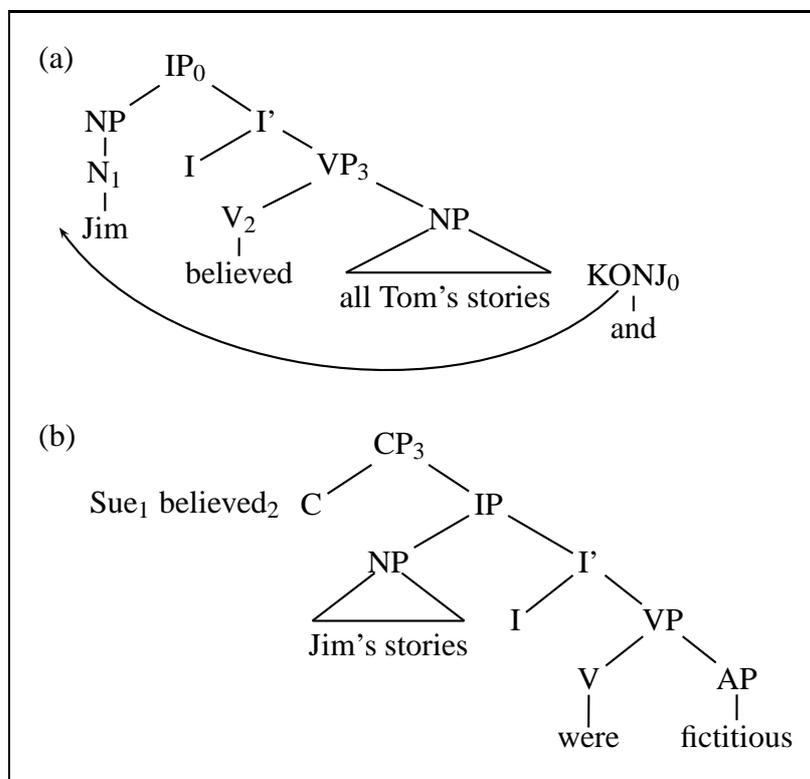


Abbildung 4.12: Parsing von (5b) im Iterationsmodell

4.2.3 Präferenzen bei Ambiguitäten

Im Folgenden wird die Frage beleuchtet, was passiert, wenn der Parser auf eine koordinierende Konjunktion stößt. Der Satz ist an dieser Stelle ambig. Der menschliche Parser hat jedoch eindeutige Präferenzen, die bereits in Kapitel 3 (Abschnitt 3.3) beschrieben wurden. Die Ambiguität kann sich im Verlauf des Satzes wieder auflösen (lokale Ambiguität). Sie kann aber auch am Satzende noch bestehen (globale Ambiguität).

In den folgenden beiden Abschnitten wiederhole ich die Präferenzen des Parsers bei lokalen und globalen Ambiguitäten und zeige, wie sich alle Präferenzen aus dem *Iterationsprinzip* (siehe Abschnitt 4.1) im Rahmen des Iterationsmodells ableiten lassen.

4.2.3.1 Lokale Ambiguitäten

Die experimentellen Ergebnisse aus der Literatur (siehe Kapitel 3, Abschnitt 3.3.1) zeigen, dass der menschliche Parser bei der Verarbeitung der Konjunktion in (6) eine Präferenz hat, die Objekt-NP [*NP* Mary] anstelle des gesamten Satzes (IP) zu koordinieren. Frazier (1979) fand einen Garden-Path-Effekt, d.h. längere Lesezeiten, bei der Verarbeitung des letzten Wortes ‘*laughed*’ in (6b) im Vergleich zu ‘*too*’ in (6a).

(6) a. Peter kissed [*NP* Mary] and [*NP* her sister] *too*.

b. [*IP* Peter kissed Mary] and [*IP* her sister *laughed*].

(7) a. Mary sold [*NP* vegetables] and [*NP* seeds].

b. Mary [*VP* sold vegetables] and [*VP* *bought* seeds].

In Sätzen wie in (7) hat der Parser eine Präferenz, die Objekt-NP [*NP* vegetables] anstelle der Verbalphrase zu koordinieren (Blodgett und Boland, 1998). Hier tritt ein Garden-Path-Effekt bei der Verarbeitung des Verbs ‘*bought*’ in (7b) im Vergleich zu ‘*seeds*’ in (7a) auf. In beiden Fällen sind die längeren Lesezeiten darauf zurückzuführen, dass die bereits aufgebaute Struktur bei der nicht-präferierten Fortsetzung des Satzes in (6b) und (7b) reanalysiert werden muss.

Beide Präferenzen lassen sich direkt aus dem Iterationsprinzip ableiten. In (6) beispielsweise springt der Parser nach dem Iterationsprinzip bevorzugt an den Beginn der Objekt-NP [*NP* Mary], da es sich hierbei um die kleinere Iterationsschleife handelt (siehe Abbildung 4.13). Der Parser versucht anschließend, die Struktur dieser NP für das nachfolgende zweite Konjunkt wiederzuverwenden. Bei der Verarbeitung des letzten Wortes (‘*too*’ bzw. ‘*laughed*’) löst sich die Ambiguität auf. Die Verarbeitung des Satzes (6a) ist dementsprechend unproblematisch. In (6b) dagegen führt die Verarbeitung des letzten Wortes ‘*laughed*’ dazu, dass die Iterationsschleife korrigiert werden muss. Außerdem muss die Anbindung von [*NP* her sister] umstrukturiert werden. Diese Reanalyseprozesse führen zu längeren Lesezeiten. Die Präferenz, die

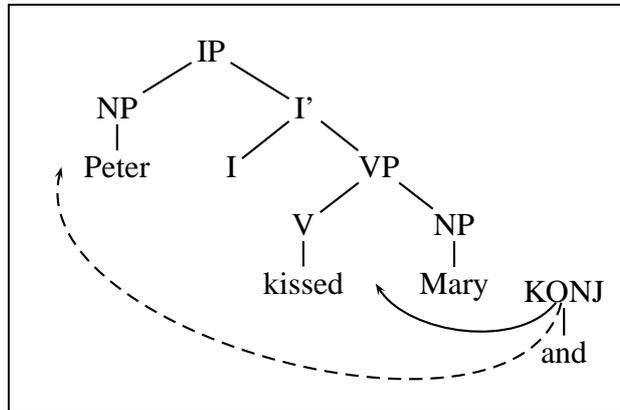


Abbildung 4.13: Präferenzen des Parsers: NP- vs. IP-Koordination

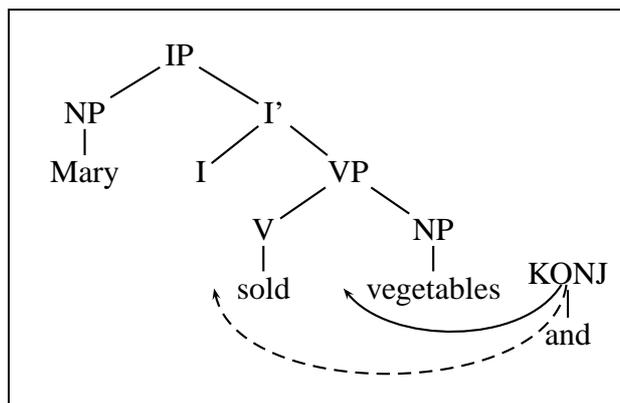


Abbildung 4.14: Präferenzen des Parsers: NP- vs. VP-Koordination

Objekt-NP anstelle des gesamten Satzes zu koordinieren, folgt somit aus dem Iterationsprinzip und erklärt auch den Garden-Path-Effekt in (6b).

In (7) springt der Parser nach dem Iterationsprinzip bevorzugt an den Beginn der Objekt-NP [*NP* vegetables], da es sich hierbei um die kleinere Iterationsschleife handelt (siehe Abbildung 4.14). Der Parser versucht auch hier, die Struktur dieser NP für das nachfolgende zweite Konjunkt wiederzuverwenden. In (7) löst sich die Ambiguität allerdings bereits beim darauffolgenden Wort (*'seeds'* bzw. *'bought'*) auf. Die Verarbeitung von (7a) ist entsprechend unproblematisch, in (7b) muss die Iterationsschleife korrigiert werden, was zu einem Garden-Path-Effekt, d.h. zu längeren Lesezeiten, führt. Auch hier folgen die Präferenz für die Koordination der Objekt-NP und der Garden-Path-Effekt in (7b) aus dem Iterationsprinzip.

4.2.3.2 Globale Ambiguitäten

Wie in Kapitel 3 (Abschnitt 3.3.2) beschrieben, hat Carlson (2001) im Rahmen einer Fragebogenstudie Gapping-Konstruktionen, wie in (8), untersucht, wobei die NP *'Sarah'* entweder als Subjekt oder als Objekt interpretiert werden kann. In beiden Fällen löst sich die Ambiguität bis zum Satzende nicht auf. Es handelt sich daher um globale Ambiguitäten.

- (8) a. [*NP* Josh] visited [*NP* the office] during the vacation and [*NP* Sarah] during the week.
- b. [*NP* Josh] visited [*NP* Marjorie] during the vacation and [*NP* Sarah] during the week.

Eine schriftliche und eine auditorische Fragebogenstudie ergaben, dass die Objekt-Lesart generell präferiert wird. Diese Präferenz bleibt auch dann bestehen, wenn die NP *'Sarah'* nur zum Subjekt des ersten Satzes (*'Josh'*) eine syntaktische Parallelität aufweist, wie in (8a), oder bei paralleler Betonung zum Subjekt des ersten Satzes. Carlson schließt daraus, dass den beiden Lesarten unterschiedliche syntaktische Strukturen zugrunde liegen (IP- vs. VP-Koordination, siehe Kapitel 3, Abbildungen 3.7 und 3.8). Der beobachtete Komplexitätseffekt

zwischen der Subjekt- und der Objekt-Lesart wird mit einer Präferenz für einfache Strukturen (*'Simple Structure Hypothesis'*) erklärt.

Dieser Komplexitätseffekt bei globalen Ambiguitäten kann jedoch auch alternativ aus dem Iterationsprinzip im Rahmen des Iterationsmodells abgeleitet werden. Wie bei den lokalen Ambiguitäten in Abschnitt 4.2.3.1 ist auch in den Gapping-Konstruktionen in (8) die Konjunktion *'and'* die ambige Stelle. Es gibt eine Vielzahl von Möglichkeiten wie der Satz fortgesetzt werden kann, d.h. welcher Teil des vorangehenden Satzes koordiniert werden soll. Im Iterationsmodell erfordert die Objekt-Lesart ein Zurückspringen des Parsers an den Beginn der Objekt-NP *'the office'* bzw. *'Marjorie'*. In diesem Falle würden die Strukturen, die für die Konstituenten *'the office'* bzw. *'Marjorie during the vacation'* aufgebaut wurden, für die Konstituenten im zweiten Konjunkt *'Sarah during the week'* wiederverwendet werden.

Die Subjekt-Lesart erfordert ein Zurückspringen zur Subjekt-NP *'Josh'* an den Satzanfang. Hier würden die Strukturen, die für die Phrasen *'Josh'* bzw. *'during the vacation'* im ersten Konjunkt aufgebaut wurden, für die Konstituenten *'Sarah'* bzw. *'during the week'* im zweiten Konjunkt wiederverwendet werden. Aus dem Iterationsprinzip folgt, dass die kleinere Iterationsschleife präferiert wird. Das Zurückspringen an den Beginn der Objekt-NP ist die kleinere Schleife und wird daher präferiert.

In diesem Abschnitt zu ambigen Strukturen habe ich gezeigt, dass sich die beobachteten Präferenzen bei lokalen und globalen Ambiguitäten aus einem einzigen Iterationsprinzip im Rahmen des Iterationsmodells ableiten lassen. Interessanterweise ist dies erst durch das Abbilden der Konjunkte auf drei-dimensionale Strukturen möglich.

Das Copy- α -Modell kann die Komplexitätseffekte bei Gapping-Konstruktionen in (8) und die Garden-Path-Effekte bei lokalen Ambiguitäten in (6) und (7) nicht erklären. Copy- α macht generell keine Vorhersagen zu Präferenzen bei Ambiguitäten.

4.2.4 Fehlende Komplexitätseffekte bei ‘*either or*’

In diesem Abschnitt wiederhole ich die fehlenden Komplexitätseffekte in *either-or*-Konstruktionen, die in Kapitel 3 (Abschnitt 3.4.2) beschrieben wurden. Anschließend zeige ich, dass sich diese Effekte auch alternativ aus dem Iterationsmodell ableiten lassen.

Frazier und Clifton (2001) untersuchten Konstruktionen mit ‘*either or*’ wie in (9) und (10). Die Autoren fanden keine Verarbeitungsunterschiede zwischen den (a)- und (b)-Sätzen. Das Segment [*or a cook*] wurde beispielsweise in (9a) und in (9b) gleich schnell gelesen. Entsprechendes gilt für das Segment [*or his father*] in (10).

- (9) a. Mary / is looking / for either a maid / *or a cook*. (NP or NP)
 b. Mary / is looking / either for a maid / *or a cook*. (PP or PP)
- (10) a. Sam / wants either his mother / *or his father*. (NP or NP)
 b. Sam / either wants his mother / *or his father*. (VP or VP)

Legt man den Reduktionsansatz von Schwarz (1999) zugrunde, würde man jedoch auf Basis der Strukturhypothese unterschiedliche Lesezeiten in den (a)- und (b)-Sätzen erwarten. Das liegt daran, dass in den (b)-Sätzen mehr Struktur aufgebaut werden müsste als in den (a)-Sätzen. Nach Schwarz (1999) werden in (9a) zwei NPs koordiniert, in (9b) jedoch zwei Präpositionalphrasen mit anschließender Tilgung der Präposition. Demnach müsste [*or a cook*] in (9b) langsamer verarbeitet werden als in (9a), da der Aufbau einer PP länger dauern sollte als der Aufbau einer NP. Außerdem würden nach diesem Ansatz in (10a) zwei NPs koordiniert werden, in (10b) jedoch zwei Verbalphrasen mit anschließender Tilgung des Verbs. Auch hier müsste [*or his father*] in (10b) langsamer verarbeitet werden als in (10a), da der Aufbau einer VP länger dauern sollte, als der Aufbau einer NP.

Aus den fehlenden Verarbeitungsunterschieden bzw. den fehlenden Komplexitätseffekten schließen Frazier und Clifton (2001), dass in Konstruktionen mit ‘*either or*’ das zweite Konjunkt nicht auf inkrementelle

Weise aufgebaut wurde, sondern, dass die Struktur des ersten Konjunks kostenlos kopiert wurde. Da *either-or*-Konstruktionen nicht ambig sind, d.h. der linke Rand der koordinierten Struktur durch 'either' eindeutig markiert ist, steht der Kopiermechanismus im Rahmen des Copy- α -Modells zur Verfügung. Weiterhin schließen die Autoren, dass es für den Kopiermechanismus unerheblich ist, wieviel Struktur kopiert wird. Das Kopieren einer VP erfordert beispielsweise keinen höheren Verarbeitungsaufwand als das Kopieren einer NP.

Die fehlenden Komplexitätseffekte können jedoch auch alternativ im Rahmen des Iterationsmodells erklärt werden. Im Iterationsmodell werden drei-dimensionale Strukturen für die *either-or*-Konstruktionen aufgebaut. Drei-dimensionale Strukturen sind dadurch gekennzeichnet, dass leere Strukturen grundsätzlich vermieden werden. Ohne die Annahme von leeren Strukturen ist das Segment [*or a cook*] in (9b) nicht komplexer als in (9a), es werden jeweils zwei Nominalphrasen koordiniert. Entsprechendes gilt für das Segment [*or his father*] in (10).

Da in (9) und (10) die beiden Konjunkte strukturgleich sind, kann im Iterationsmodell jeweils die Struktur des ersten Konjunks für das zweite Konjunkt wiederverwendet werden. Bei der Verarbeitung von (9a) und (9b) wird die komplette NP-Struktur des ersten Konjunks [*a maid*] für das zweite Konjunkt [*a cook*] wiederverwendet. Lediglich die lexikalischen Elemente des zweiten Konjunks ('*a cook*') müssen an die bereits aufgebaute Struktur angebunden werden. Im Iterationsmodell wird in beiden Fällen gleich viel Struktur aufgebaut, und es werden daher auf Basis der Strukturhypothese auch keine Unterschiede in den Lesezeiten erwartet. Auch in (10a) und (10b) wird die komplette NP-Struktur des ersten Konjunks [*his mother*] für das zweite Konjunkt [*his father*] wiederverwendet. Wieder wird in beiden Fällen gleich viel Struktur aufgebaut, und es werden keine Unterschiede in den Lesezeiten erwartet.

In diesem Abschnitt habe ich gezeigt, dass die fehlenden Komplexitätseffekte bei 'either or' auch problemlos im Rahmen des Iterationsmodells erklärt werden können. Der Vorteil des Iterationsmodells ist dabei, dass die aufgebauten Strukturen maximal ökonomisch sind und die etablierte Strukturhypothese beibehalten werden kann.

4.2.5 Parallelitätseffekte bei der Interpretation

Wie in Kapitel 3 (Abschnitt 3.2.1) beschrieben, hat der Parser eine Präferenz, koordinierte Strukturen beim Sprachverstehen parallel zu interpretieren. In Satzvervollständigungs-Tests (Henstra, 1996; Schepman, 1997) wird beispielsweise die Variante in (11a) gegenüber (11b) präferiert.

- (11) a. The lawyer greeted [[*NP* the powerful barrister] and [*NP* the old judge]] [*CP* who *were* walking to the courtroom].
- b. The lawyer greeted [*NP* the powerful barrister] and [[*NP* the old judge [*CP* who *was* walking to the courtroom]].

Der Relativsatz [*CP* who was/were walking to the courtroom] bezieht sich in (11a) auf beide Konjunkte, d.h. die Konjunkte sind parallel strukturiert. In (11b) bezieht sich der Relativsatz nur auf das zweite Konjunkt, was ungleiche Konjunkte zur Folge hat. Bei der Verarbeitung des Relativpronomens ‘*who*’ liegt eine strukturelle Ambiguität bei der Anbindung des Relativsatzes vor. Diese Ambiguität löst sich beim darauffolgenden finiten Verb (‘*was*’ bzw. ‘*were*’) auf. Die Präferenz für parallele Konjunkte lässt sich elegant aus dem Iterationsmodell und den zugrundeliegenden drei-dimensionalen Datenstrukturen ableiten. Das wird im Folgenden illustriert.

Im Iterationsmodell werden die Sätze in (11) bis zur Konjunktion ‘*and*’ inkrementell aufgebaut. Die Verarbeitung der Konjunktion löst eine Iterationsschleife aus. Nach dem Iterationsprinzip springt der Parser zurück an den Beginn der Objekt-NP [*NP* the powerful barrister] und verwendet die bereits aufgebaute Struktur für die nachfolgende NP [*NP* the old judge]. Wenn der Parser also auf das Relativpronomen ‘*who*’ stößt, wurde im Iterationsmodell bereits eine ökonomische drei-dimensionale Struktur für die beiden Nominalphrasen aufgebaut (siehe Abbildung 4.15).

Das Relativpronomen ‘*who*’ wird nach dem Recency-Prinzip an die letzte NP angebunden. Das ist in diesem Fall der NP-Mutterknoten der Koordination. Dieser NP-Mutterknoten, wie auch die anderen redundanten Knoten, existieren in einer drei-dimensionalen Struktur nur einmal und werden von beiden Konjunkten geteilt (siehe Abbildung 4.15). Das heißt, die Präferenz für die Anbindung an

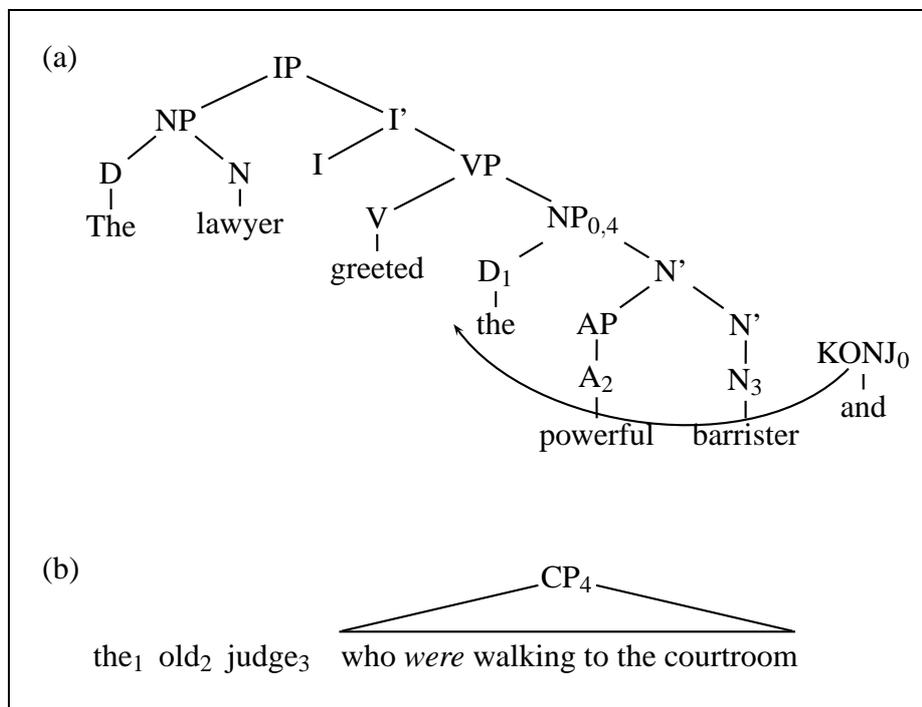


Abbildung 4.15: Parsing von (11a) im Iterationsmodell

den NP-Mutterknoten der Koordination ergibt sich aus den drei-dimensionalen Strukturen und dem Recency-Prinzip.

Anschließend trifft der menschliche Parser auf das finite Verb des Relativsatzes ('*was*' bzw. '*were*'). Im Falle der Verbform '*were*' ist die bis dahin aufgebaute Struktur kompatibel mit dem finiten Verb und der Relativsatz kann ohne Reanalyse integriert werden. Dabei bezieht sich der Relativsatz auf beide Konjunkte, d.h. die Konjunkte sind parallel strukturiert. Die resultierende Struktur ist in Abbildung 4.15 dargestellt.

Im Falle der Verbform '*was*' ist die bis dahin aufgebaute Struktur nicht kompatibel mit dem finiten Verb. Der Relativsatz bezieht sich hierbei auf das zweite Konjunkt, d.h. die bis dahin aufgebaute Struktur muss reanalysiert werden. Dabei muss der Relativsatz an das zweite Konjunkt angebunden werden. In einer drei-dimensionalen Struktur kann in diesem Fall der NP-Mutterknoten nicht mehr geteilt werden. Jedes Konjunkt benötigt einen eigenen NP-Knoten, wobei

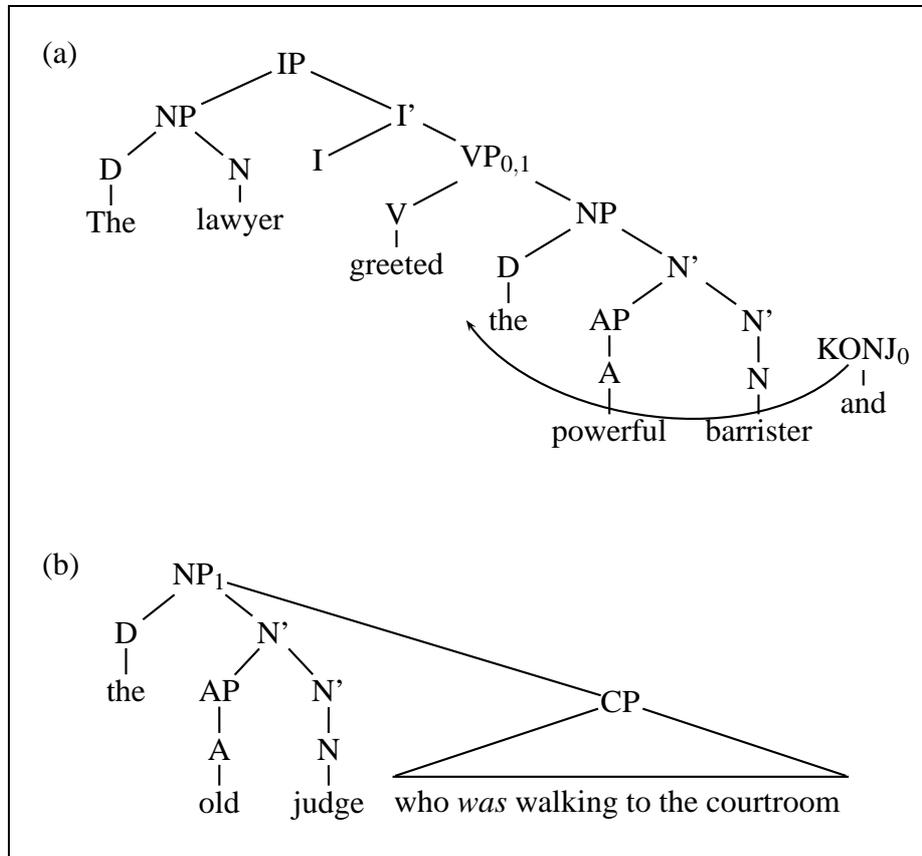


Abbildung 4.16: Parsing von (11b) im Iterationsmodell

der nachfolgende Relativsatz nur an den NP-Knoten des zweiten Konjunks angebunden wird. Die resultierende Struktur ist in Abbildung 4.16 dargestellt. Bei dieser Analyse sind die Konjunkte nicht mehr parallel strukturiert.

Die Anbindung des Relativsatzes an den gemeinsamen NP-Mutterknoten ergibt die ökonomischere Struktur (siehe Abbildung 4.15) und erfordert keine Reanalyseprozesse. Bei dieser Analyse bezieht sich der Relativsatz auf beide Konjunkte, d.h. die Konjunkte sind parallel strukturiert. Bezieht sich der Relativsatz nur auf das zweite Konjunkt, muss mehr Struktur aufgebaut werden (siehe Abbildung 4.16) und die bereits aufgebaute ökonomische Struktur muss reanalysiert werden. Die Analyse in Abbildung 4.15 mit parallel strukturierten

Konjunkten erfordert im Iterationsmodell einen geringeren Verarbeitungsaufwand und wird daher präferiert.

Die Präferenz, koordinierte Strukturen beim Sprachverstehen parallel zu interpretieren, ergibt sich jedoch nur durch die Verwendung der drei-dimensionalen Datenstrukturen. Das Copy- α -Modell macht keine Vorhersage für die Verarbeitung der Ambiguität in (11). Da es sich um ambige Strukturen handelt, ist der Kopiermechanismus für diese Sätze nicht verfügbar und der ganze Satz wird inkrementell nach allgemeinen Parsingmechanismen aufgebaut. Verwendet man ein allgemeines Verarbeitungsmodell auf Basis von zwei-dimensionalen Strukturen, ist die Anbindung des Relativsatzes an die koordinierte NP nicht ökonomischer. Im zwei-dimensionalen Fall lässt sich keine Präferenz für eine der beiden syntaktischen Strukturen aus allgemeinen Parsingprinzipien, wie z.B. dem Minimal-Attachment-Prinzip, ableiten, da beide Varianten die gleiche Anzahl an syntaktischen Knoten erfordern. Da Minimal-Attachment nicht greift, würde letztlich das Late-Closure-Prinzip oder das allgemeine Recency-Prinzip zur Anwendung kommen, was fälschlicherweise eine Präferenz für die Anbindung des Relativsatzes an das letzte Konjunkt vorhersagen würde.

4.3 Sprachverständnis bei agrammatischer Aphasie

Sprachstörungen im Bereich der Syntax, wie z.B. die agrammatische Aphasie, sind ein interessanter Testfall für die vorgeschlagenen Parsingmechanismen im Iterationsmodell und die zugrundeliegenden drei-dimensionalen Datenstrukturen. Im folgenden Abschnitt werden die agrammatische Sprache sowie agrammatische Verstehensdaten zu Koordination im Deutschen präsentiert. Anschließend zeige ich, wie diese Daten im Rahmen des Iterationsmodells erklärt werden können.

4.3.1 Die agrammatische Sprache

Der Agrammatismus tritt typischerweise bei Broca-Aphasie auf. Die Patienten sprechen mühevoll und verwenden einen telegrammartigen Sprachstil. Die folgende Definition für agrammatische Sprache ist aus Menn und Obler (1990).

Agrammatic speech:

- a. "... reduction of grammar through the omission of free or bound grammatical morphemes,
- b. use of unmarked forms (e.g., nominatives of nouns, infinitive for verbs),
- c. limited syntactic variety
- d. near-absence of syntactic constructions indicating subordination."

Das Defizit ist nicht auf die Sprachproduktion beschränkt. Auch das Sprachverstehen ist betroffen, wenn für die Interpretation eines Satzes syntaktische Information notwendig ist. Verschiedene Defizit-Theorien zum Agrammatismus gestehen diesen Patienten ein unterschiedliches Maß an syntaktischen Fertigkeiten zu. Es wird in der Regel angenommen, dass Agrammatiker keine vollständigen syntaktischen Strukturen, aufbauen können (Ouhalla, 1993; Grodzinsky, 1995; Friedmann und Grodzinsky, 1997; Friedmann, 2001). Diese Defizit-Theorien basieren im Allgemeinen auf einer syntaktischen Theorie und spezifizieren welche Komponenten des Modells beeinträchtigt sind.

Ouhalla (1993) geht beispielsweise davon aus, dass Agrammatiker nicht mehr in der Lage sind, Strukturen für funktionale Kategorien, wie z.B. I(nflection), C(omplementizer) oder D(eterminierer), aufzubauen. Dieser Ansatz basiert auf Chomskys Prinzipien- und Parameter-Theorie (z.B. Chomsky, 1986; Abney, 1987). Nach Ouhalla werden diese Kategorien bei der Verarbeitung "ignoriert", in dem Sinne, dass dafür keine funktionalen Projektionen aufgebaut werden können. Dies führt zu unvollständigen strukturellen Repräsentationen. Einem Satz kann bestenfalls eine VP-Struktur mit eingebetteten NPs zugewiesen werden. Bewegung und Spuren sind nicht mehr möglich.

Interessanterweise sind Agrammatiker trotz dieser syntaktischen Beeinträchtigungen in der Lage, koordinierte Nominalphrasen⁹ im Deutschen zu verarbeiten. In Steiner (2005) habe ich gezeigt, dass agrammatische Patienten,

⁹ Ob Agrammatiker eine NP- oder eine DP-Struktur aufbauen, ist unklar. Im Folgenden gehe ich nach Ouhalla (1993) davon aus, dass Agrammatiker nur NPs aufbauen können. Dies ist jedoch für die weitere Argumentation unerheblich.

im Rahmen einer Satz-Bild-Zuordnungsaufgabe, Sätze wie in (12) verstehen können.¹⁰ Bei dieser Aufgabe wird dem Patienten ein Satz mündlich dargeboten und er soll auf einem Testbogen auf das passende Bild zeigen. Ein Testbogen für den Satz in (12c) ist in Abbildung 4.17 dargestellt. Dieselben Patienten hatten jedoch Probleme, W-Fragen wie in (13) korrekt zu interpretieren, was zeigt, dass diese Patienten keine vollständigen syntaktischen Repräsentationen aufbauen können.

- (12) a. Das Mädchen und der Junge tauchen.
 b. Der Mann und die Frau gießen die Blumen.
 c. Der Mann trägt die Leiter und den Eimer.

- (13) a. Wer wäscht die Bären?
 b. Wen waschen die Bären?
 c. Welcher Affe wäscht die Bären?
 d. Welchen Affen waschen die Bären?

Wie in Steiner (2005) ausgeführt, liefern die Verstehensschwierigkeiten bei W-Fragen Evidenz für die Defizit-Theorie von Ouhalla (1993). Die Vorhersage für die koordinierten Strukturen in (12) hängt jedoch davon ab, welche syntaktische Struktur für die Koordination von NPs bzw. DPs angenommen wird. Die agrammatischen Daten sind beispielsweise nicht kompatibel mit einem Reduktionsansatz für koordinierte NPs (/DPs). Bei einem Reduktionsansatz würde die koordinierte NP auf die Koordination von vollständigen Sätzen (CPs) mit anschließender Tilgung des redundanten sprachlichen Materials zurückgehen (*Der Mann trägt die Leiter und ~~der Mann trägt~~ den Eimer*). Dies erfordert den Aufbau von funktionalen Projektionen.

Die Verstehensdaten zu den W-Fragen zeigen außerdem, dass die Patienten keine vollständigen Satzstrukturen aufbauen können. Die agrammatischen

¹⁰ Um die Sätze in (12) verstehen zu können, müssen zwei Nominalphrasen strukturell verknüpft und als Einheit für die Weiterverarbeitung betrachtet werden. Nur dann kann die thematische Rolle AGENS oder THEMA korrekt zugewiesen werden.

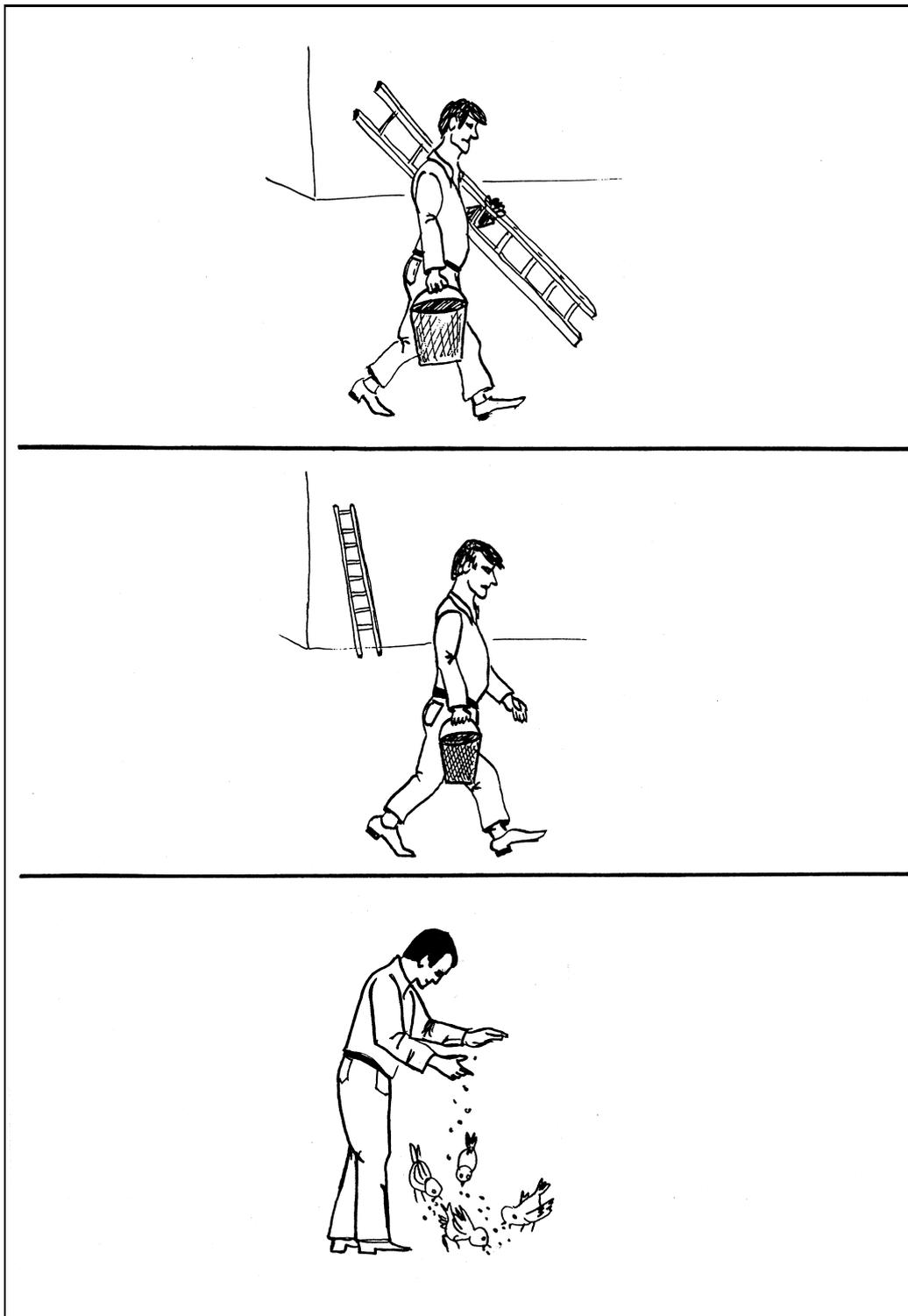


Abbildung 4.17: Testbogen für den Satz “Der Mann trägt die Leiter und den Eimer.” (Ziel: oberes Bild)

Daten sind jedoch kompatibel mit einem phrasalen Ansatz für koordinierte NPs (/DPs), in dem keine funktionale Projektion für die Konjunktion, d.h. keine Konjunktionsphrase, angenommen wird. Die Verstehensdaten sind außerdem kompatibel mit einem drei-dimensionalen Ansatz für koordinierte NPs (/DPs), da für diese Phrasen keine vollständigen Satzstrukturen aufgebaut werden müssen. Weiterhin werden in den meisten mehrdimensionalen Ansätzen keine Konjunktionsphrasen angenommen.

Im Verlaufe dieses Kapitels habe ich gezeigt, dass eine Vielzahl von experimentellen Ergebnissen bei der Satzverarbeitung für eine drei-dimensionale Analyse der Koordination sprechen. Es stellt sich nun die Frage, ob es möglich ist, dass Agrammatiker trotz des syntaktischen Defizits drei-dimensionale syntaktische Strukturen für die koordinierten NPs aufgebaut haben. Im Folgenden möchte ich zeigen, dass die agrammatischen Daten sehr gut mit dem Iterationsmodell und den drei-dimensionalen Strukturen harmonieren.

4.3.2 Agrammatische Verarbeitung im Iterationsmodell

Die Divergenz aus syntaktischem Defizit einerseits und intakter NP-Koordination andererseits lässt sich problemlos aus dem Iterationsmodell ableiten. Ich gehe davon aus, dass die syntaktischen Probleme beim Agrammatismus im hierarchischen Strukturaufbau liegen. Das betrifft die Phasen 1 und 4 des Iterationsmodells. Die Feedback- und Iterationsphase (Phase 2, 3) dagegen ist intakt. Das bedeutet, dass die in Phase 1 aufgebaute (defizitäre) Struktur auch wiederverwendet werden kann. Möglicherweise können in Phase 1 nur einfachste Phrasen aufgebaut werden. Diese können jedoch wiederverwendet, d.h. koordiniert werden. Kann ein Agrammatiker also eine Struktur für Nominalphrasen aufbauen, dann kann er auch Nominalphrasen koordinieren, etc.

Die Abbildungen 4.18 und 4.19 zeigen, wie die agrammatische Verarbeitung der Sätze in (12b,c) im Iterationsmodell abläuft. Dabei lege ich die Defizit-Theorie von Ouhalla (1993) zugrunde. In Abbildung 4.18 baut der Parser eine defizitäre Struktur für das erste Konjunkt (*der Mann*) auf, die nur aus einer NP-Konstruktion besteht (Phase 1). Da der Determinierer *D* eine funktionale Kategorie ist, kann nach Ouhalla (1993) keine DP-Struktur aufgebaut werden. Im Iterationsmodell springt der Parser bei der Verarbeitung der Konjunktion

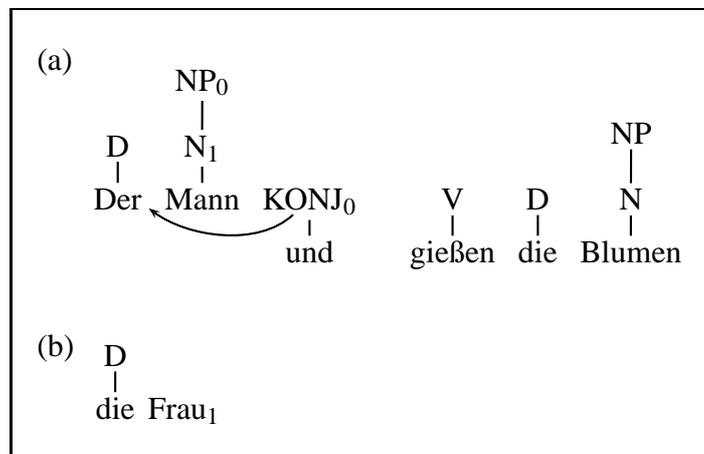


Abbildung 4.18: Parsing von (12b) im Iterationsmodell

an den Beginn einer Phrase zurück, wobei die kleinste Iterationsschleife bevorzugt wird. In (12b) kommt hierfür nur der Beginn der Nominalphrase (*Mann*) in Frage (Phase 2). Anschließend versucht der Parser, die bereits aufgebaute Struktur für das zweite Konjunkt (*die Frau*) wiederzuverwenden. Da die beiden Konjunkte strukturidentisch sind, kann die komplette NP-Struktur wiederverwendet werden (Phase 3). Nach abgeschlossener Iterationsphase erfolgt der (defizitäre) Strukturaufbau für den restlichen Satz (Phase 4). Dabei können nur Nominalphrasen aufgebaut werden. Aufgrund der Verb-Zweitstellung des Satzes kann für das Verb keine weitere Struktur aufgebaut werden. Dies würde den Aufbau von funktionalen Projektionen erforderlich machen.

In Abbildung 4.19 baut der Parser ebenfalls eine defizitäre Struktur auf, bis er auf die Konjunktion trifft. Der Parser springt zurück an den Beginn des letzten Nominalphrase (*Leiter*). Dieser Phrasenbeginn wird bevorzugt, da es sich dabei um die kleinste Iterationsschleife handelt. Anschließend versucht der Parser, die bereits aufgebaute Struktur für das zweite Konjunkt (*den Eimer*) wiederzuverwenden. Auch hier kann die komplette NP-Struktur wiederverwendet werden. Die Konjunktion wird in beiden Sätzen (12b,c) an die jeweilige koordinierte Nominalphrase angebunden. Das ist problemlos möglich, obwohl die Konjunktion eine funktionale Kategorie ist. Das liegt daran, dass in drei-dimensionalen Strukturen keine funktionale Projektion für die Konjunktion, d.h. keine Konjunktionsphrase, aufgebaut werden muss.

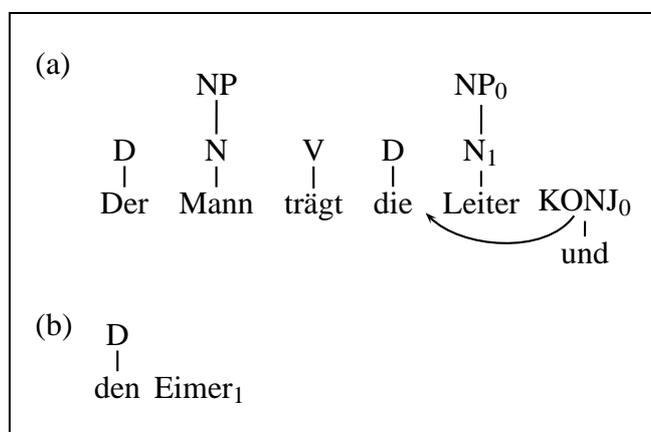


Abbildung 4.19: Parsing von (12c) im Iterationsmodell

Während in bisherigen Verarbeitungsmodellen Strukturen grundsätzlich hierarchisch aufgebaut werden, nimmt das vorgeschlagene Iterationsmodell zwei unterschiedliche Prozesse an: hierarchischer Strukturaufbau für hypotaktische Konstruktionen vs. Iterationsmechanismus für Koordination. Die agrammatischen Daten liefern Evidenz dafür, dass einer dieser Prozesse unabhängig vom anderen beeinträchtigt sein kann. Das Defizit betrifft den hierarchischen Strukturaufbau (Phase 1, 4), der Iterationsmechanismus scheint jedoch intakt (Phase 2, 3).

Die syntaktischen Strukturen, die im Agrammatismus aufgebaut werden können, sind zwar defizitär, aber bei Koordination dennoch drei-dimensional. Auf struktureller Ebene betrifft das Defizit folglich den Strukturaufbau *innerhalb* einer Ebene, die Verknüpfung *zwischen* den drei-dimensional angeordneten Ebenen (Behindance-Relation) ist intakt.

In diesem Abschnitt wurde gezeigt, dass sich die Parsingmechanismen des Iterationsmodells und die drei-dimensionalen syntaktischen Strukturen auch auf die agrammatische Sprache übertragen lassen. Die drei-dimensionalen Datenstrukturen scheinen nur auf den ersten Blick sehr komplex zu sein, die Verarbeitung dieser Strukturen ist jedoch einfacher und ökonomischer als die Verarbeitung zwei-dimensionaler Strukturen. Aus dem Iterationsmodell lassen sich letztlich Vorhersagen für die Verarbeitung und das Sprachverstehen koordinierter Strukturen im Agrammatismus ableiten.

4.4 Zusammenfassung und Diskussion

In diesem Kapitel wurde das *Iterationsmodell*, ein Modell zur syntaktischen Verarbeitung koordinierter Strukturen, vorgestellt. Das Iterationsmodell basiert auf einem alternativen Parsingmechanismus, wobei bereits aufgebaute Strukturen bei Koordination wiederverwendet werden. Die syntaktischen Strukturen, die sich aus diesem Parsingmechanismus ergeben, sind drei-dimensionale Baumstrukturen, wie sie im Rahmen von mehrdimensionalen Modellen vorgeschlagen werden. Die Verarbeitungsmechanismen im Iterationsmodell und die daraus resultierenden drei-dimensionalen Strukturen sind maximal ökonomisch.

Das Iterationsmodell macht Vorhersagen für Lesezeiten und Präferenzen bei koordinierten Strukturen auf Basis der etablierten Strukturhypothese: Je mehr Struktur aufgebaut wird, desto länger sind die Lesezeiten. In diesem Kapitel wurde gezeigt, dass sich eine Vielzahl von Verarbeitungseffekten bei Koordination aus dem Iterationsmodell ableiten lassen. Viele dieser Effekte lassen sich nicht oder nur eingeschränkt aus allgemeinen Parsingmechanismen bzw. dem Copy- α -Modell ableiten. Allgemeine Parsingmechanismen sind unabhängig von Koordination und basieren auf zwei-dimensionalen Datenstrukturen. Das Copy- α -Modell ist koordinationspezifisch und verwendet ebenfalls zwei-dimensionale syntaktische Strukturen. Das Iterationsmodell ist auch koordinationspezifisch, basiert aber auf drei-dimensionalen Datenstrukturen. Tabelle 4.1 zeigt eine Übersicht, welche der in Kapitel 3 beschriebenen Effekte durch welche Theorien erklärt werden können. Desweiteren lässt sich das Iterationsmodell auch auf die Verarbeitung koordinierter Strukturen im Agrammatismus übertragen.

Der Vorteil des Iterationsmechanismus ist, dass dieser unabhängig vom hierarchischen Strukturaufbau beschrieben werden kann. Es gibt eine definierte Schnittstelle zwischen dem hierarchischen Strukturaufbau innerhalb einer Ebene (Phase 1, 4) und dem Iterationsmechanismus für Koordination (Phase 2, 3). Für das Iterationsmodell ist es nicht entscheidend, welche syntaktischen Strukturen in Phase 1 und 4 zugrundegelegt werden und nach welchen Parsingmechanismen diese verarbeitet werden. Die Strukturen, die in Phase 1 aufgebaut werden, werden in Phase 3 wiederverwendet.

	Allgemeine Mechanismen		
	Copy-α	Iterationsmodell	
Parallelitätseffekt (Interpretation)	nein	nein	ja
Parallel-Struktur-Effekt (sentential)	nein	nein	ja
Parallel-Struktur-Effekt (cross-kategorial)	nein	nein	ja
Parallel-Struktur-Effekt (Phrasenaufbau)	nein	nein	ja
Lokale Ambiguitäten	ja	kV	ja
Globale Ambiguitäten	ja	kV	ja
Fehlender Komplexitäts- effekt (<i>'either or'</i>)	ja?	ja?	ja

Tabelle 4.1: Welcher Effekt kann durch welches Modell erklärt werden?
(kV: keine Vorhersage, ja?: abhängig von Syntax-Theorie zu *'either or'*)

Im Folgenden möchte ich auf die Unterschiede zwischen dem beschriebenen Iterationsmodell und dem Copy- α -Modell (Frazier und Clifton, 2001) eingehen. Beide Modelle sind speziell für die Verarbeitung koordinierter Strukturen entwickelt worden.

Iterationsmodell vs. Copy- α -Modell: Im Copy- α -Modell werden die konventionellen zwei-dimensionalen Datenstrukturen beibehalten, dafür wird die etablierte Strukturhypothese bei der Verarbeitung von speziellen (nicht-ambigen) koordinierten Strukturen und Ellipsen aufgegeben. Stattdessen wird ein kostenloser Kopierprozess angenommen. Das Iterationsmodell basiert auf drei-dimensionalen Datenstrukturen, dafür kann die etablierte Strukturhypothese generell beibehalten werden. Im Iterationsmodell werden Strukturen nicht kopiert, sondern wiederverwendet (Recycling). Während im Copy- α -Modell der Kopiermechanismus auf nicht-ambige Strukturen eingeschränkt wird, steht der Recycling-Mechanismus im Iterationsmodell grundsätzlich bei Koordination zur Verfügung. In ambigen Strukturen werden die Präferenzen durch das Iterationsprinzip geregelt.

Aus den beschriebenen Mechanismen der beiden Modelle ergeben sich unterschiedliche Vorhersagen für Lesezeiten bzw. Präferenzen. Einerseits machen die beiden Modelle unterschiedliche Vorhersagen für die Verarbeitung ambiger koordinierter Strukturen, d.h. Konstruktionen, bei denen die syntaktische Grenze des ersten Konjunks nicht eindeutig markiert ist. Das betrifft alle in Kapitel 3 (Abschnitt 3.2) beschriebenen Satzkonstruktionen. Im Copy- α -Modell steht der Kopiermechanismus für diese Konstruktionen nicht zur Verfügung, d.h. es wird keine Verarbeitungserleichterung für parallele Konjunkte erwartet. Im Iterationsmodell werden parallele Strukturen grundsätzlich wiederverwendet. Eine Verarbeitungserleichterung ist demnach auch in ambigen Strukturen zu erwarten. Wie in Kapitel 3 beschrieben, wurde eine Verarbeitungserleichterung für ambige koordinierte Strukturen gefunden. Diese sind im Rahmen des Iterationsmodells erklärbar, nicht jedoch im Copy- α -Modell (siehe auch Tabelle 4.1).

Desweiteren machen die beiden Modelle unterschiedliche Vorhersagen bezüglich Konjunkten, die zwar strukturähnlich, aber nicht exakt strukturidentisch sind. Im Copy- α -Modell wird eine Verarbeitungserleichterung nur bei

strukturidentischen Konjunkten erwartet, d.h. wenn das zweite Konjunkt strukturell tatsächlich eine Kopie des ersten Konjunktes ist. Im Iterationsmodell dagegen wird ein *gradueller* Parallel-Struktur-Effekt erwartet. Je größer die strukturelle Übereinstimmung, desto mehr Struktur kann wiederverwendet werden und desto größer ist die Verarbeitungserleichterung.

Schließlich ergeben sich aus den beiden Modellen unterschiedliche Vorhersagen zur Länge der Konjunkte. Im Copy- α -Modell wird ein kostenloser Kopierprozess angenommen. Der Verarbeitungsaufwand ist dabei unabhängig von der Größe der zu kopierenden Struktur. Daraus ergeben sich keinerlei Präferenzen für die Konjunktlänge. Im Iterationsmodell dagegen ergibt sich aus dem Iterationsprinzip eine Präferenz für kurze Konjunkte. Diese Präferenz wird in Kapitel 5 anhand von Korpusdaten evaluiert.

Implikationen für die Syntax: In Kapitel 2.3 wurden verschiedene Ansätze zur syntaktischen Analyse koordinierter Strukturen vorgestellt: der Reduktionsansatz, die phrasale Analyse und mehrdimensionale Modelle. Die agrammatischen Daten aus Abschnitt 4.3.1 und die Modellierung der Lesezeiteffekte bzw. Präferenzen im Rahmen des Iterationsmodells liefern auf indirekte Weise Evidenz für die syntaktische Analyse koordinierter Strukturen.

Die Sprachverstehensdaten bei agrammatischen Aphasiepatienten können für die Evaluation einer Defizit-Theorie verwendet werden. Wenn im Rahmen dieser Defizit-Theorie verschiedene syntaktische Analysen zu unterschiedlichen Vorhersagen für das agrammatische Sprachverständnis führen, kann indirekt zwischen diesen Analysen unterschieden werden. In Abschnitt 4.3.1 wurde Evidenz für die Defizit-Theorie von Ouhalla (1993) auf der Basis von zwei Experimenten gefunden. Aufgrund der Ergebnisse von beiden Untersuchungen kann der Reduktionsansatz für die Analyse von NP- (/DP)-Koordination ausgeschlossen werden. Zwischen einer phrasalen und einer drei-dimensionalen Analyse kann mit diesen Daten nicht unterschieden werden. Diese Unterscheidung wird allerdings durch die Modellierung der Verarbeitungseffekte möglich.

Mit Hilfe von Lesezeiten und Präferenzen können wiederum Verarbeitungsmodelle evaluiert werden, die ebenfalls auf syntaktischen Theorien basieren. Wenn verschiedene syntaktische Analysen zu unterschiedlichen Vorhersagen z.B. für Lesezeiten führen, kann indirekt auch zwischen diesen Analysen unterschieden werden. In diesem Kapitel habe ich gezeigt, dass eine Vielzahl von Verarbeitungseffekten im Iterationsmodell auf Basis von drei-dimensionalen syntaktischen Strukturen erklärt werden kann. Viele dieser Effekte lassen sich nicht oder nur eingeschränkt aus Verarbeitungsmodellen ableiten, die auf zwei-dimensionalen Strukturen basieren. Die Modellierung der Verarbeitungseffekte bei Koordination liefert somit indirekte Evidenz für eine drei-dimensionale syntaktische Analyse der Koordination.

Durch die Kombination dieser beiden Datentypen ergibt sich ein interessantes Bild. Die agrammatischen Daten alleine liefern Evidenz gegen den Reduktionsansatz, erlauben es aber nicht, zwischen der phrasalen und der drei-dimensionalen Analyse zu unterscheiden. Dies wird jedoch durch die Modellierung der Verarbeitungseffekte möglich. Andererseits ist es nicht möglich, durch die beschriebene Modellierung den Reduktionsansatz auszuschließen. Erst die Kombination der beiden Datentypen liefert Evidenz für die drei-dimensionale Analyse von Koordination. Wie im Agrammatismus drei-dimensionale Strukturen aufgebaut werden können, wurde in Abschnitt 4.3.2 beschrieben.

Nachdem in diesem Kapitel eine Vielzahl an Verarbeitungseffekten beim Sprachverstehen betrachtet und modelliert wurde, stellt sich nun die Frage, warum der menschliche Parser die beschriebenen Präferenzen hat. Es könnte beispielsweise sein, dass der menschliche Parser diese Präferenzen deshalb hat, weil die entsprechenden Strukturen, wie z.B. parallele Konjunkte, besonders häufig in der Sprache vorkommen. Diese Frage betrifft das Verhältnis zwischen Sprachverstehen und Sprachproduktion, welches im folgenden Kapitel beleuchtet wird. Es wird untersucht, ob die in diesem Kapitel beschriebenen Parsingpräferenzen in Form von Häufigkeiten in Korpusdaten zu finden sind.

Kapitel 5

Das Verhältnis von Parsingpräferenzen und Korpusfrequenzen

In Kapitel 3 wurden Verarbeitungseffekte präsentiert, die man in Sprachverstehensexperimenten, insbesondere in Lesezeit-Studien zum Englischen, bei koordinierten Strukturen gefunden hat. Diese Effekte wurden in Kapitel 4 modelliert. In diesem Kapitel soll der Frage nachgegangen werden, ob die Effekte auch in der Sprachproduktion zu finden sind. Dazu wurden englische spontansprachliche Korpusdaten statistisch ausgewertet. Diese Korpusdaten werden außerdem als Trainingsdaten für die Simulationen mit neuronalen Netzen in Kapitel 6 verwendet.

Im vorliegenden Kapitel können somit gleichzeitig zwei Fragestellungen beleuchtet werden. Einerseits erfahren wir, wie koordinierte Strukturen quantitativ in den Trainingsdaten verteilt sind. Diese Verteilung hat Implikationen für die Vorhersagbarkeit von bestimmten Strukturen, für ihre Lernbarkeit und für die Interpretation der Simulationsergebnisse in Kapitel 6. Andererseits bekommen wir Aufschluss über das Verhältnis von Sprachverstehens- und Sprachproduktionsmechanismen.

5.1 Einleitung

Ein wichtiger Gegenstand der Psycholinguistik betrifft das Verhältnis von Sprachverstehens- und Sprachproduktions-Mechanismen. Eine Möglichkeit, diese Frage zu beleuchten, besteht darin, Parsingpräferenzen (Sprachverstehensdaten) und Korpusfrequenzen (Sprachproduktionsdaten) miteinander zu vergleichen. Nach der *Tuning-Hypothese* (Cuetos et al., 1996; Mitchell et al., 1995) sind Parsingpräferenzen in syntaktisch ambigen Sätzen erfahrungsbasiert. Das heißt, sie sind abhängig davon, wie häufig der menschliche Parser ähnlichen Strukturen im sprachlichen Input ausgesetzt war und welche Auflösung einer Ambiguität am häufigsten vorkam. Dabei wird die statistische Verteilung in der Sprachproduktion als Ursache für die Ausbildung von Parsingpräferenzen betrachtet. Nach diesem Ansatz sollten Parsingpräferenzen und Korpusfrequenzen korrelieren, d.h. die Konstruktion, die während des Sprachverstehens präferiert wird, sollte auch in Korpusdaten häufiger vorkommen als die nicht-präferierte Konstruktion. Der Status dieser Hypothese ist jedoch bislang ungeklärt.

Eine Untersuchung, die gegen die Tuning-Hypothese spricht, betrifft die Ambiguität bei der Relativsatz-Anbindung in Konstruktionen der Form “*NP₁ Prep NP₂ Relative Clause*”, wie z.B. in (1). Der Relativsatz (RC) kann sich dabei entweder auf *NP₁* (hohe Anbindung) oder auf *NP₂* (tiefe Anbindung) beziehen. Im Englischen besteht eine Präferenz, den Relativsatz tief anzubinden, d.h. in (1) wird der Relativsatz bevorzugt auf [*NP₂ the actress*] bezogen (Cuetos und Mitchell, 1988). Diese Präferenz kann jedoch von Sprache zu Sprache variieren.

- (1) Someone shot [*NP₁ the servant*] of [*NP₂ the actress*] [*RC who was on the balcony*].

Im Niederländischen beispielsweise wurde in Lesezeit-Studien gezeigt, dass die Versuchspersonen eine Präferenz hatten, in obigen Konstruktionen den Relativsatz hoch an *NP₁* anzubinden (Brybaert und Mitchell, 1996; Mitchell et al., 2000). Es wäre nun denkbar, dass die unterschiedlichen Präferenzen mit einer unterschiedlichen Häufigkeitsverteilung in der jeweiligen Sprache

einhergehen. In vergleichenden Korpusstudien zum Holländischen haben Mitchell und Brysbaert (1998) jedoch beobachtet, dass in Korpora, entgegen der Parsingpräferenzen, tief angebundene Relativsätze häufiger vorkamen als hoch angebundene Relativsätze. Dies wurde von Mitchell und Brysbaert als Evidenz gegen die Tuning-Hypothese gewertet.

Auf der anderen Seite konnten Desmet et al. (2002) jedoch zeigen, dass die experimentellen Stimuli in den Lesezeit-Experimenten zum Holländischen nicht direkt vergleichbar waren mit den Sätzen, die in die Korpusauswertung eingingen. Die Testsätze enthielten hauptsächlich Nominalphrasen, die sich auf menschliche Entitäten bezogen, in den Korpusätzen dagegen referierten die meisten Nominalphrasen auf nicht-menschliche Entitäten. Eine Reanalyse des Korpus von Mitchell und Brysbaert (1998) ergab, dass die Anbindungspräferenzen von den Eigenschaften der NP_1 abhängen. Bezog sich die NP_1 auf menschliche Entitäten, waren die Relativsätze häufiger hoch angebinden. Dies korreliert mit den experimentellen Ergebnissen. Bezog sich NP_1 auf nicht-menschliche Entitäten, waren die Relativsätze häufiger tief angebinden. Somit konnte gezeigt werden, dass die Diskrepanz zwischen den Sprachverstehens- und Sprachproduktionsergebnissen nicht mehr vorliegt, sobald die Korpusdaten sorgfältig kontrolliert werden.

Die Ergebnisse weiterer Studien zum Englischen wurden ebenfalls als Evidenz gegen die Tuning-Hypothese betrachtet. Gibson und Schütze (1999) und Gibson, Schütze und Salomon (1996) untersuchten Präferenzen bei koordinierten Nominalphrasen der Form “ NP_1 Prep NP_2 Prep NP_3 and NP_4 ”, wobei es für NP_4 drei Anbindungsmöglichkeiten gibt. Ein Beispiel ist in (2) illustriert.

- (2) The salesman ignored [NP_1 a customer] with [NP_2 a child] with [NP_3 a dirty face] and ...
- a. [NP_4 a wet diaper] (tiefe Anbindung)
 - b. [NP_4 one with a wet diaper] (mittlere Anbindung)
 - c. [NP_4 one with a baby with a wet diaper] (hohe Anbindung)

In einer Offline-Studie, bei der die Verständlichkeit der verschiedenen Strukturen beurteilt werden sollte (Gibson, Schütze und Salomon, 1996), sowie in einer Lesezeit-Studie (Gibson und Schütze, 1999) wurde gezeigt, dass hohe Anbindung verständlicher ist als mittlere Anbindung (hoch < mittel). Detaillierte Korpusuntersuchungen ergaben jedoch, dass mittlere Anbindung in diesen Konstruktionen signifikant häufiger vorkommt als hohe Anbindung (mittel < hoch). Da in den Studien keine Korrelation zu den Korpusfrequenzen gefunden wurde, kamen Gibson und Schütze zu folgendem Schluss: "...the sentence comprehension mechanism is not using corpus frequencies in arriving at its preference in this ambiguity and hence the decision principles of sentence comprehension and sentence production must be partially distinct" (Gibson und Schütze, 1999, S. 263).

Desmet und Gibson (2003) konnten jedoch zeigen, dass die experimentellen Sätze, die in Gibson, Schütze und Salomon (1996) bzw. in Gibson und Schütze (1999) verwendet wurden, nicht direkt vergleichbar waren mit den Sätzen, die in die Korpusauswertung eingingen. Die Testsätze enthielten das Pronomen 'one' als Kopf der koordinierten Nominalphrase (NP_4), wobei die koordinierten NPs im Korpus in den meisten Fällen kein Pronomen enthielt. Desmet und Gibson (2003) konnten in einer Korpusstudie und in einem Lesezeit-Experiment zeigen, dass in Konstruktionen wie in (2) das Vorkommen des Pronomens 'one' in einer Nominalphrase die Anbindungspräferenzen beeinflusst. Die Anwesenheit des Pronomens 'one' führte zu einer Präferenz für hohe Anbindung, die Abwesenheit des Pronomens ergab eine Tendenz zu mittlerer Anbindung. Dadurch konnte gezeigt werden, dass die Präferenz für hohe Anbindung beim Sprachverstehen in Gibson und Schütze (1999) auf die Verwendung des Pronomens 'one' in den Testsätzen zurückzuführen ist. Desmet und Gibson (2003) argumentieren aufgrund dieser Ergebnisse, dass Gibson und Schützes Hypothese, dass bei dieser Ambiguität Sprachverstehen und Produktion auf unterschiedliche Mechanismen zurückgreifen, angezweifelt werden muss.

Die beschriebene Diskussion zeigt, dass es nach wie vor unklar ist, welches Verhältnis zwischen Sprachverstehen und Sprachproduktion besteht. Es ist beispielsweise nicht geklärt, ob Parsingpräferenzen und Korpusfrequenzen grundsätzlich korrelieren und inwiefern Korpusfrequenzen verwendet werden

können, um Aufschluss über die Mechanismen der Satzverarbeitung zu bekommen. Um weitere Evidenz zu erhalten, wird in diesem Kapitel das Verhältnis von Parsingpräferenzen und Korpusfrequenzen anhand von koordinierten Strukturen im Englischen untersucht. Ich konzentriere mich in diesem Kapitel auf drei Verarbeitungseffekte aus der Literatur, die in Lesezeit-Experimenten zum Englischen gefunden wurden: 1. Präferenz für die Koordination gleicher Kategorien, 2. Parallel-Struktur-Effekt im Phrasenaufbau, 3. Präferenz für NP-Koordination (gegenüber S-Koordination) bei lokalen Ambiguitäten. Alle drei Effekte wurden in Kapitel 3 beschrieben. Diese Präferenzen, die sich durch schnellere Lesezeiten ausdrücken, werden mit den entsprechenden Korpusdaten in der Englischen Verbmobil Baumbank (TÜBA-E, Hinrichs et al., 2000) verglichen (siehe hierzu Steiner, 2007a).

Die Baumbank TÜBA-E besteht aus spontansprachlichen Dialogen zu Terminvereinbarungen und Hotelreservierungen. Die Annotation der Baumbank umfasst morpho-syntaktische Kategorien (Wortarten), syntaktische Phrasenstrukturen und grammatische Funktionen. Ich habe mich für ein spontansprachliches Korpus entschieden, das aus uneditierten natürlichen Dialogen besteht, weil diese Daten viel eher Mechanismen der Sprachproduktion reflektieren als editierte Zeitungstexte.¹ Weiterhin zeige ich in Kapitel 7, dass Online-Datentypen, wie z.B. Lesezeiten, mit gesprochenen Korpusdaten korrelieren, während Offline-Datentypen, wie Grammatikalitätsurteile, eher mit geschriebenen Texten korrelieren.

¹ In Gibson, Schütze und Salomon (1996) wird betont, dass es für die Untersuchung von Sprachproduktions-Mechanismen relevant ist, uneditierte spontansprachliche Korpusdaten zu verwenden. In den bisherigen Korrelationsstudien wurden schriftsprachliche Korpora als Grundlage verwendet.

5.2 Erstellung des Koordinations-Datensatzes

Als Grundlage für die Korpusauswertungen in diesem Kapitel dient ein Korpusfragment der TÜBA-E-Baumbank, das 2906 Sätze (CD13) umfasst. Dasselbe Korpusfragment wird in Kapitel 6 als Trainingsdaten für die Simulationen mit den neuronalen Netzen verwendet. Aus diesem Korpusfragment wurden alle koordinierten Strukturen extrahiert und zu einem Koordinations-Datensatz zusammengestellt, der im Hinblick auf die verschiedenen Präferenzen ausgewertet werden soll.

Um eine Vergleichbarkeit mit den experimentellen Stimuli zu gewährleisten, wurden nur koordinierte Strukturen mit zwei Konjunkten berücksichtigt, die durch die Konjunktion *‘and’* verbunden sind. Desweiteren wurden nur Koordinationen extrahiert, die innerhalb von Sätzen vorkommen. Das letztere Kriterium ist notwendig, da es sich bei diesen Daten um gesprochene Sprache handelt, die eine Vielzahl fragmentarischer Äußerungen enthält. Das Korpusfragment enthält insgesamt 242 koordinierte Strukturen, die die genannten Kriterien erfüllen (Koordinations-Datensatz). Für jede dieser koordinierten Strukturen wurden zusätzlich die folgenden Informationen aus dem Korpus extrahiert:

- die syntaktische Kategorie der Mutter- und Tochterknoten,
- die grammatische Funktion der Mutter,
- die Länge der Konjunkte (Anzahl Wörter),
- der Grad der strukturellen Übereinstimmung in den Konjunkten.

Bei der Erstellung des Koordinations-Datensatzes wurde die syntaktische Annotation der Baumbank zugrunde gelegt. Der Datensatz wurde automatisch erstellt, die Werkzeuge dazu habe ich in der Programmiersprache “Java” programmiert. Lediglich der Grad der strukturellen Übereinstimmung in den Konjunkten wurde manuell ermittelt (zum Vorgehen siehe Abschnitt 5.4.1). Tabelle 5.1 enthält eine Liste aller syntaktischen Kategorien, die in der TÜBA-E-Baumbank koordiniert sind, sowie ein Korpusbeispiel und die Häufigkeit dieser Kategorie als Mutterknoten der Koordination.

Kategorie	Beschreibung	Korpusbeispiel	Häufigkeit
S	Satz	<i>[I am free that afternoon] and [the seventeenth I am completely free]</i>	77
VP	Verbalphrase	<i>if I can [run home] and [get a shower]</i>	19
NP	Nominal-phrase	<i>I have [a early morning meeting] and [a lunch meeting]</i>	113
PP	Präpositional-phrase	<i>[on the twenty sixth] and [on the twenty seventh] I am busy all day</i>	7
AP	Adjektiv-phrase	<i>but early on next week is actually [pretty free] and [late]</i>	1
ADVP	Adverbphrase	<i>[where] and [when] would like to do it</i>	1
CNUM	Kardinalzahl	<i>it sounds like between [twelve] and [one] o'clock would be the best time</i>	1
ONUM	Ordinalzahl	<i>I am going out of town on the [sixth] and [seventh]</i>	23

Tabelle 5.1: Syntaktische Kategorien, die in the TÜBA-E-Baumbank koordiniert sind.

5.3 Cross-kategoriale Koordination

Der erste Verarbeitungseffekt, den ich untersuche, ist die Präferenz für die Koordination gleicher syntaktischer Kategorien. Frazier et al. (2000) haben in einem Lesezeit-Experiment zum Englischen beobachtet, dass die Verknüpfung von gleichen syntaktischen Kategorien, wie in (3a), schneller verarbeitet wird als die Koordination von ungleichen syntaktischen Kategorien (cross-kategoriale Koordination), wie in (3b).

- (3) a. John walked [*AdvP* slowly] and [*AdvP* carefully], avoiding the broken glass.
- b. John walked [*AdvP* slowly] and [*PP* with great care], avoiding the broken glass.
- c. John walked [*AdvP* carefully], avoiding the broken glass.
- d. John walked [*PP* with great care], avoiding the broken glass.

Das zweite Konjunkt [*AdvP* carefully] in (3a) wird schneller gelesen als die gleiche Phrase in (3c), die unabhängig von Koordination verarbeitet werden muss. Im Gegensatz dazu wird das cross-kategoriale zweite Konjunkt [*PP* with great care] in (3b) nicht schneller gelesen als in (3d) (siehe Kapitel 3, Abschnitt 3.2.3.1). Der menschliche Parser präferiert offensichtlich Konjunkte der gleichen syntaktischen Kategorie.

Um herauszufinden, ob diese Präferenz auch in den Korpusdaten zu finden ist, wird der Koordinations-Datensatz nach gleichen vs. ungleichen syntaktischen Kategorien ausgewertet.

5.3.1 Korpusanalyse I:

Cross-kategoriale vs. gleiche syntaktische Kategorien

Die Korpusauswertung ergab, dass der Koordinations-Datensatz sechs Vorkommen enthält, die als cross-kategoriale Koordination annotiert sind: zwei Verknüpfungen von *NP* und *PP*, eine Verknüpfung von *VP* und *NP*, eine Verknüpfung von *PP* und *VP* sowie zwei Verknüpfungen von *NP* und *S*.

Bei genauerer Betrachtung der zuletzt genannten Konstruktion (*NP + S*) hat sich herausgestellt, dass diese keine echten cross-kategorialen Koordinationen darstellen. Vielmehr bezieht sich das erste "Konjunkt", d.h. die *NP*, auf die vorangehende Äußerung, der nachfolgende Satz beginnt genau genommen mit einem initialen 'and'. Diese äußerungsübergreifenden Relationen wurden in der Baubank nicht annotiert. Es bleiben letztlich vier Vorkommen von cross-kategorialer Koordination (siehe z.B. (4a,b)). In allen anderen Konstruktionen werden gleiche Kategorien verknüpft.

- (4) a. I would like to get together [*VP* talk] and [*NP* stuff] (*VP + NP*)
 b. I am [*PP* out of town] and [*VP* tied up the rest of the week] (*PP + VP*)

Insgesamt sind im Koordinations-Datensatz zu 98% ($n = 238$) gleiche Kategorien verknüpft und nur in 2% der Fälle ($n = 4$) sind es ungleiche syntaktische Kategorien. Diese Präferenz für die Verknüpfung von gleichen syntaktischen Kategorien ist so stark, dass sich eine statistische Analyse an dieser Stelle erübrigt. Wie man aufgrund der Tuning-Hypothese erwarten würde, ist die Präferenz, die in Lesezeit-Experimenten beobachtet werden kann, auch in Form von relativen Häufigkeiten in gesprochenen Korpora zu finden.

5.4 Parallel-Struktur-Effekt im Phrasenaufbau

Der zweite Verarbeitungseffekt, den ich in diesem Kapitel untersuche, ist der Parallel-Struktur-Effekt im Phrasenaufbau, der in Frazier et al. (2000) beobachtet wurde. In Lesezeit-Studien haben die Autoren gezeigt, dass das zweite Konjunkt in einer koordinierten Struktur schneller gelesen wird, wenn es mit dem ersten Konjunkt strukturgleich ist. Die Nominalphrase [*a tall woman*] wird in (5a) schneller verarbeitet als in (5b), da die beiden koordinierten NPs in (5a) strukturgleich sind, nicht jedoch in (5b) (siehe Kapitel 3, Abschnitt 3.2.3.2).

- (5) a. Hilda noticed [*NP a strange man*] and [*NP a tall woman*].
 b. Hilda noticed [*NP a man*] and [*NP a tall woman*].

Um herauszufinden, ob dieser Effekt auch in den Korpusdaten zu finden ist, wird der Koordinations-Datensatz ausgewertet. Dabei ist der Grad der strukturellen Übereinstimmung in den Konjunkten relevant. Wie ich diese Information ermittelt habe, wird im folgenden Abschnitt beschrieben.

5.4.1 Korpusanalyse II: Strukturelle Ähnlichkeit der Konjunkte

Für jedes Vorkommen im Koordinations-Datensatz habe ich den Grad der strukturellen Übereinstimmung von 0 bis 100% in den Konjunkten bestimmt. 100% Übereinstimmung bedeutet, dass beide Konjunkte exakt dieselbe Struktur einschließlich der lexikalischen Kategorien (Wortarten) haben. 0% Übereinstimmung bedeutet, dass keiner der syntaktischen Knoten identisch ist. Der Satz in (6a) ist ein Beispiel aus dem Koordinations-Datensatz mit einer strukturellen Übereinstimmung von 36%. Die Konjunkte in Beispiel (6b) sind dagegen strukturidentisch, d.h. 100% Übereinstimmung. Die entsprechenden Einträge in der Baumbank sind in den Abbildungen 5.1 und 5.2 dargestellt.

- (6) a. [*NP Monday*] and [*NP Tuesday the twenty seventh*] are bad for me
 b. [*NP twenty second*] and [*NP twenty fourth*] are pretty bad

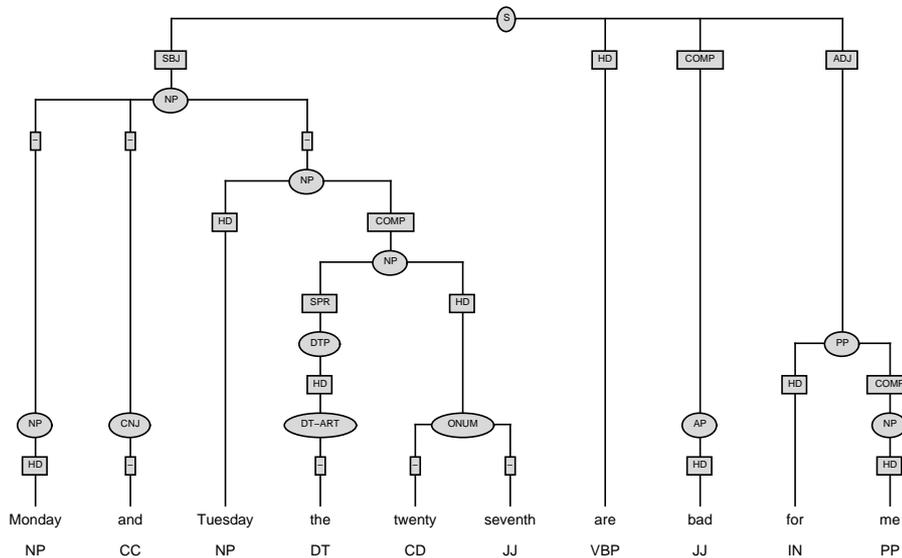


Abbildung 5.1: Korpusbeispiel mit 36% struktureller Übereinstimmung in den Konjunkten (siehe (6a))

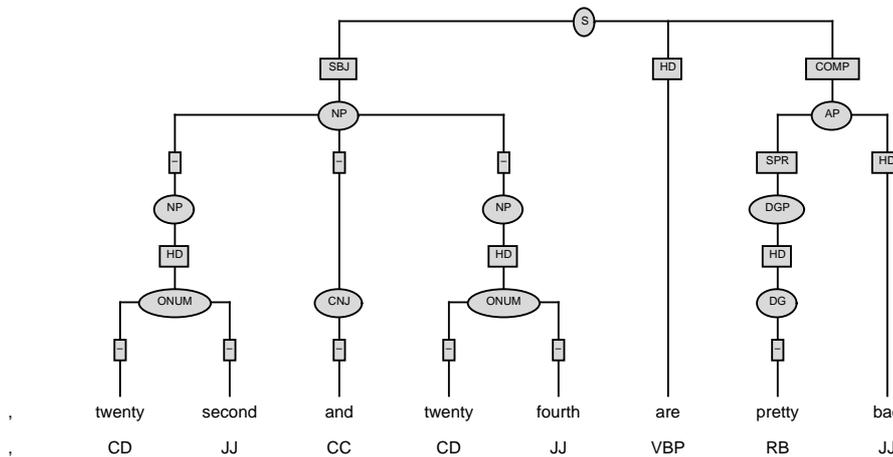


Abbildung 5.2: Korpusbeispiel mit 100% struktureller Übereinstimmung in den Konjunkten (siehe (6b))

In Abbildung 5.1 ist zu sehen, dass sich die beiden Konjunkte [*NP* Monday] und [*NP* Tuesday the twenty seventh] deutlich in ihrer Struktur unterscheiden. Der Grad der strukturellen Übereinstimmung wird bestimmt, indem die syntaktischen Knoten der beiden Konjunkte gezählt werden (einschließlich der Terminalknoten mit den lexikalischen Kategorien). Anschließend wird im Top-Down-Verfahren berechnet, wieviele dieser Knoten in beiden Konjunkten vorkommen. In Abbildung 5.1 enthalten beide Konjunkte zusammen elf syntaktische Knoten, wobei in jedem Konjunkt jeweils zwei Knoten identisch sind. Es handelt sich dabei um die *NP*-Mutterknoten der beiden Konjunkte und um die beiden Terminalknoten mit lexikalischer Kategorie *NP*, die den Wörtern ‘Monday’ und ‘Tuesday’ entsprechen. Vier der insgesamt elf Knoten kommen also parallel vor, woraus eine strukturelle Übereinstimmung von 36% resultiert. Im Gegensatz dazu haben die Konjunkte in Abbildung 5.2 eine identische Struktur. Acht von acht Knoten kommen parallel vor, was eine Übereinstimmung von 100% bedeutet.

In Abbildung 5.3 ist dargestellt, wie die koordinierten Strukturen nach Grad der strukturellen Übereinstimmung von 0% bis 100% im Koordinations-Datensatz quantitativ verteilt sind. Die erste Säule in Abbildung 5.3 repräsentiert die Anzahl der koordinierten Strukturen mit einer strukturellen Übereinstimmung zwischen 0% und 10% (exklusiv). Die zweite Säule repräsentiert die Anzahl der koordinierten Strukturen mit einer strukturellen Übereinstimmung zwischen 10% und 20% (exklusiv), etc. Die vorletzte Säule betrifft den Übereinstimmungsgrad zwischen 90% und 100% (exklusiv), während die letzte Säule ausschließlich die Vorkommen mit 100% Übereinstimmung repräsentiert. Das Diagramm zeigt, dass die unterschiedlichen Übereinstimmungsgrade gleichmäßig im Koordinations-Datensatz verteilt sind, mit Ausnahme der 100%-Säule. Diese letzte Säule unterscheidet sich deutlich von den anderen Säulen und macht 32,6% aller Vorkommen im Koordinations-Datensatz aus. Das heißt, fast ein Drittel aller koordinierter Strukturen enthält strukturidentische Konjunkte, wie z.B. der Satz in (6b).

Um dieses Ergebnis interpretieren zu können, stellt sich die Frage, wie häufig eine strukturelle Übereinstimmung von 100% zwischen zwei Phrasen zufällig vorkommt. Zur Beantwortung dieser Frage habe ich einen Zufalls-Datensatz erstellt. Als Basis wurde dasselbe Korpusfragment verwendet, das bereits Grundlage für die Erstellung des Koordinations-Datensatzes war. Für jedes

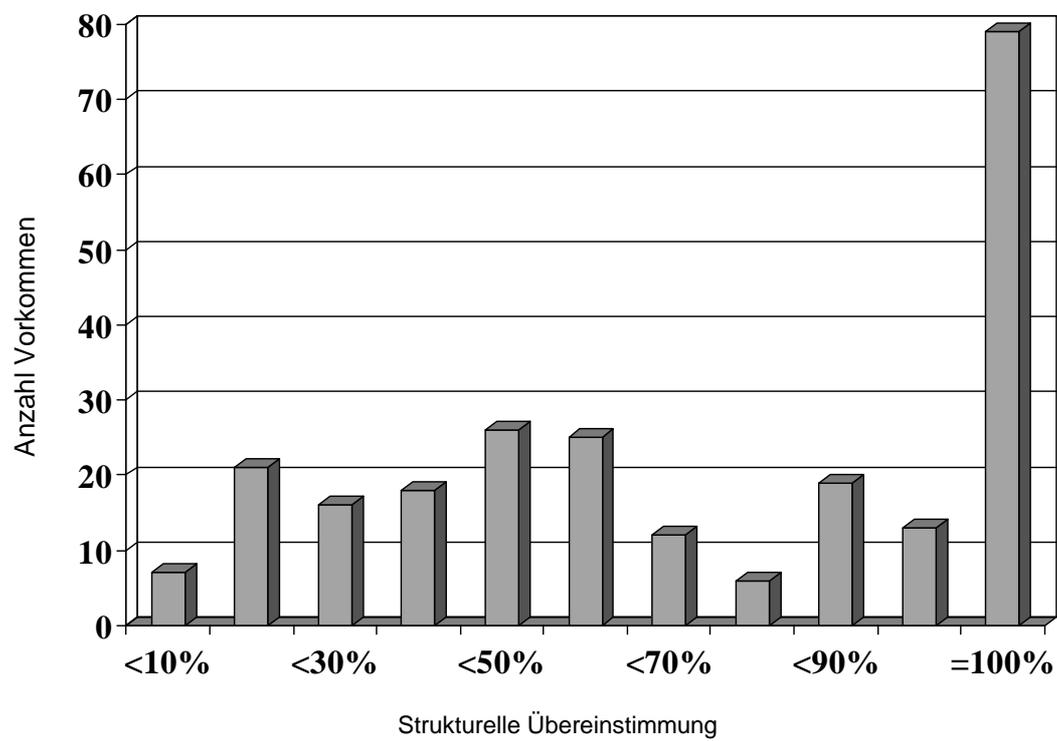


Abbildung 5.3: Strukturelle Übereinstimmung von 0% bis 100% im Koordinationsdatensatz

erste Konjunkt aus dem Koordinations-Datensatz habe ich ein zufälliges zweites “Konjunkt” aus dem genannten Korpusfragment extrahiert. Es handelt sich dabei um eine zufällige Phrase, die jedoch gewisse Kriterien erfüllen muss. Die Zufalls-Phrase darf nicht Teil einer koordinierten Struktur sein und sie muss mit dem ursprünglichen zweiten Konjunkt bezüglich der syntaktischen Kategorie übereinstimmen, sowie bezüglich der grammatischen Funktion und der Länge (Anzahl Wörter). In (7c) ist eine Zufalls-Phrase für die koordinierte Struktur in (6b) dargestellt, die in (7a) wiederholt wird. Die Paare aus ursprünglichem ersten Konjunkt (7b) und zufälligem zweiten “Konjunkt” (7c) bilden den Zufalls-Datensatz.

- (7) a. [*NP* twenty second] and [*NP* twenty fourth] are pretty bad
- b. [*NP* twenty second] (erstes Konjunkt)
- c. [*NP* the first] (Zufalls-Phrase)

Die Zufalls-Phrase [*NP* the first] in (7c) stimmt mit dem ursprünglichen zweiten Konjunkt [*NP* twenty fourth] hinsichtlich der syntaktischen Kategorie überein (*NP*), sowie hinsichtlich der grammatischen Funktion (“Subjekt”) und der Länge (2 Wörter). Es muss nun berechnet werden, wieviele Vorkommen im Zufalls-Datensatz strukturidentisch sind.

Die Auswertung des Zufalls-Datensatzes ergab, dass 10,7% der Vorkommen, d.h. Paare bestehend aus ursprünglichem ersten Konjunkt und zufälligem zweiten “Konjunkt”, eine strukturelle Übereinstimmung von 100% aufweisen. Im Koordinations-Datensatz waren 32,6% der Vorkommen strukturidentisch. Die Differenz zwischen dem Koordinations- und dem Zufallsdatensatz hinsichtlich strukturidentischer Vorkommen (32,6% vs. 10,7%) ist signifikant ($\chi^2(1) = 60,1$; $p < 0,01$). Vollständige strukturelle Übereinstimmung kommt in unserem Korpus signifikant häufiger in koordinierten Strukturen vor, als zwischen zwei Phrasen, die nicht koordiniert sind. Der Parallel-Struktur-Effekt im Phrasenaufbau von Konjunkten ist also auch in Form von relativen Häufigkeiten in gesprochenen Korpusdaten zu finden. Diese Korrelation zwischen Verarbeitung und Korpusfrequenzen ist nach der Tuning-Hypothese zu erwarten.

Syntaktische Kategorie	Koordinierte Strukturen	Parallele Vorkommen		Mittlere Länge (erstes Konjunkt)
		Koordination	Zufall	
ADVP	1	100%	100%	1,0
CNUM	1	100%	100%	1,0
ONUM	23	78,3%	78,3%	1,5
NP	113	50,0%	5,3%	2,4
AP	1	0%	0%	2,0
PP	7	28,6%	0%	4,4
S	77	1,3%	0%	8,1
VP	19	0%	0%	3,5
Alle Kategorien	242	32,6%	10,7%	4,3

Tabelle 5.2: Verteilung der parallelen Vorkommen nach syntaktischer Kategorie

5.4.2 Verteilung der parallelen Vorkommen

Um sicherzugehen, dass es sich bei dem beobachteten Parallel-Struktur-Effekt im Korpus um einen allgemeinen Effekt handelt und nicht um ein Artefakt, das bestimmten Eigenschaften der koordinierten Struktur zuzuschreiben ist, wurde die Verteilung der parallelen Vorkommen hinsichtlich syntaktischer Kategorie und Länge genauer untersucht. Ich betrachte zunächst die Verteilung nach syntaktischen Kategorien.

5.4.2.1 Verteilung nach syntaktischen Kategorien

In Tabelle 5.2 ist für jede syntaktische Kategorie die Anzahl der koordinierten Strukturen und der Anteil der parallelen Vorkommen (in Prozent) im Koordinations- und im Zufalls-Datensatz dargestellt, sowie die durchschnittliche Länge der ersten Konjunkte (Anzahl Wörter). Die Anzahl der koordinierten Strukturen für jede Kategorie (zweite Spalte) und die mittlere Länge des ersten Konjunks (letzte Spalte) sind im Koordinations- und im Zufalls-Datensatz identisch.

Wie man in Tabelle 5.2 sehen kann, kommen in fast allen Kategorien auch parallele Konjunkte vor (mit Ausnahme der Kategorien *AP* und *VP*). Vergleicht man jedoch die parallelen Vorkommen im Koordinations- und im Zufalls-Datensatz, stellt sich heraus, dass nur die Kategorien *NP* und *PP* einen deutlichen Unterschied zwischen koordinierten und zufälligen parallelen Vorkommen aufweisen (*NP*: 50,0% vs. 5,3%; *PP*: 28,6% vs. 0%). Einige Kategorien (*ADVP*, *CNUM* und *ONUM*) haben einen hohen Anteil an parallelen Vorkommen in beiden Datensätzen. Auf der anderen Seite gibt es auch Kategorien, die (fast) keine parallelen Vorkommen aufweisen (*AP*, *VP* und *S*), weder im Koordinations- noch im Zufalls-Datensatz. Im Folgenden gehe ich auf diese Verteilungen näher ein.

Kategorien *ADVP*, *CNUM*, *ONUM*: Diese Kategorien haben einen hohen Anteil an parallelen Vorkommen in beiden Datensätzen (*ADVP*: 100% vs. 100%; *CNUM*: 100% vs. 100%; *ONUM*: 78,3% vs. 78,3%). Die Ursache für diese Verteilung ist die eingeschränkte Variationsbreite in der Form dieser Kategorien. Im verwendeten Korpusfragment kommt beispielsweise die Kategorie *ONUM* in nur zwei strukturellen Varianten vor: a) als Einwort-Ordinalzahl, wie z.B. ‘*sixth*’, oder b) als Zweiwort-Ordinalzahl, wie z.B. ‘*twenty sixth*’.² In dieser Kategorie ist die Struktur der Konjunkte daher vollständig durch ihre Länge bestimmt. Da die Zufalls-Konjunkte mit den ursprünglichen zweiten Konjunkten bezüglich ihrer Länge übereinstimmen müssen, entspricht jedem parallelen Vorkommen im Koordinations-Datensatz zwangsläufig auch ein paralleles Vorkommen im Zufalls-Datensatz. Die gleiche Erklärung trifft auch auf die Kategorie *CNUM* zu. Auch in dieser Kategorie ist die Struktur der Konjunkte vollständig durch ihre Länge determiniert. Bei der Kategorie *ADVP* ist die Situation etwas anders. Die Struktur einer *ADVP* ist nicht durch die Länge festgelegt, das einzige Vorkommen von *ADVP*-Koordination hat jedoch eine grammatische Funktion, die in nur einer strukturellen Variante im Korpus vorkommt. Da die Zufalls-Konjunkte mit

² Ordinalzahlen, wie z.B. ‘*twenty sixth*’ sind in der Baumbank als Zweiwort-Konstruktionen annotiert. Daraus folgt, dass eine koordinierte Struktur wie z.B. [*the fifth*] and [*the twenty first*] nach den verwendeten Kriterien nicht vollständig parallel sind. Ein anderes Annotationsschema, bei dem solche Zweiwort-Ordinalzahlen als ein Wort behandelt werden, hätte eine höhere Anzahl an parallelen Vorkommen im Koordinations-Datensatz zur Folge.

den ursprünglichen zweiten Konjunkten auch bezüglich der grammatischen Funktion übereinstimmen müssen, entspricht dem parallelen Vorkommen im Koordinations-Datensatz zwangsläufig auch ein paralleles Vorkommen im Zufalls-Datensatz. Die Kategorien *ONUM*, *CNUM* und *ADVP* sind aufgrund ihrer eingeschränkten strukturellen Variationsbreite im verwendeten Korpusfragment weniger gut geeignet, um den Parallel-Struktur-Effekt zu untersuchen.

Kategorien *NP*, *PP*: Die Kategorien *NP* und *PP* scheinen bezüglich der strukturellen Variationsbreite im mittleren Bereich zu liegen. In beiden Kategorien gibt es einen deutlichen Unterschied zwischen der Anzahl der parallelen Vorkommen im Koordinations- und im Zufalls-Datensatz. Diese Unterschiede zeigen, dass der Parallel-Struktur-Effekt nicht auf die kombinatorischen Eigenschaften einer spezifischen syntaktischen Kategorie zurückzuführen ist.

Kategorien *AP*, *VP*, *S*: Diese Kategorien weisen keine parallelen Vorkommen auf, weder im Koordinations-Datensatz noch in den Zufallsdaten (mit Ausnahme eines parallelen Vorkommens bei *S*-Koordination). Über die *AP*-Kategorie kann nicht viel gesagt werden, da *AP*-Koordination nur einmal im Korpus vorkommt. Es gibt jedoch zwei mögliche Erklärungen für die fehlenden parallelen Vorkommen bei den Kategorien *VP* und *S*. Eine Ursache könnte in der hohen strukturellen Variationsbreite dieser Kategorien liegen. Vermutlich würde eine größere Datenbasis zu einigen parallelen Vorkommen in den Kategorien *VP* und *S* führen. Es wäre jedoch auch möglich, dass die Länge der Konjunkte den Parallel-Struktur-Effekt beeinflusst, da die durchschnittliche Länge des ersten Konjunks in der Kategorie *VP* (3,5 Wörter) und *S* (8,1 Wörter) größer ist, als beispielsweise in der Kategorie *NP* (2,4 Wörter).

Insgesamt lässt sich sagen, dass der Parallel-Struktur-Effekt im Korpus ein allgemeiner Effekt zu sein scheint, der nicht nur innerhalb einer bestimmten syntaktischen Kategorie auftritt. Ein gewisses Maß an struktureller Variationsbreite ist jedoch Voraussetzung dafür, den Effekt beobachten zu können. Im nächsten Abschnitt wird untersucht, ob die Länge der Konjunkte ein relevanter Faktor für den Parallel-Struktur-Effekt ist.

5.4.2.2 Verteilung nach Konjunktlänge

Es gibt zwei Gründe, die Verteilung der parallelen Vorkommen nach Konjunktlänge genauer zu untersuchen. Erstens soll herausgefunden werden, ob, aufgrund der kombinatorischen Eigenschaften der Konjunkte, der Parallel-Struktur-Effekt im Korpus nur bei bestimmten Konjunktlängen auftritt. Oben wurde argumentiert, dass die strukturelle Variationsbreite einer Kategorie beim Parallel-Struktur-Effekt eine Rolle spielt. Die Variationsbreite ist jedoch auch von der Länge abhängig, d.h. je länger eine Phrase, desto größer ist grundsätzlich die mögliche strukturelle Variation. Es wäre beispielsweise denkbar, dass der Effekt bei kurzen Konjunkten grundsätzlich nicht auftritt, weil die Variationsbreite dieser Konjunkte zu klein ist (wie oben für die Kategorien *ONUM*, *CNUM* und *ADVP* gezeigt).

Zweitens soll herausgefunden werden, ob in den Korpusdaten ein allgemeiner Längeneffekt zu finden ist, d.h. ein kognitiver Effekt, der auf Beschränkungen des Arbeitsgedächtnisses zurückzuführen ist. In diesem Fall ist zu erwarten, dass der Parallel-Struktur-Effekt bei kürzeren Konjunkten stärker ist als bei längeren Konjunkten, da für die Verarbeitung des zweiten Konjunks die komplette Struktur des ersten Konjunks im Arbeitsspeicher gehalten werden muss. Das wird zunehmend schwieriger, je länger das erste Konjunkt ist.

Die Verteilung der parallelen Vorkommen ist in Tabelle 5.3 in Abhängigkeit der Länge der ersten Konjunkte (Anzahl Wörter) dargestellt. Für jede Länge ist die Anzahl der koordinierten Strukturen und der Anteil der parallelen Vorkommen (in Prozent) im Koordinations- und im Zufalls-Datensatz angegeben. Wie Tabelle 5.3 entnommen werden kann, kommen parallele Strukturen im Koordinations-Datensatz bei Konjunkten der Länge 1 bis 5 vor. Konjunkte, die länger sind als fünf Wörter, sind im Korpus nicht mehr strukturidentisch (mit Ausnahme einer *S*-Koordination der Länge 9). Im Zufalls-Datensatz kommen strukturidentische Phrasen nur bei Länge 1 und 2 vor. Für die Längen 1 bis 4 ergeben sich quantitative Unterschiede zwischen den parallelen Vorkommen im Koordinations- und im Zufalls-Datensatz. Der Parallel-Struktur-Effekt scheint also für verschiedene Konjunktlängen (1–4) zu existieren.

Im Folgenden werden die statistischen Ergebnisse für die Längen 1 und 2 präsentiert. Für die Längen 3 und 4 kann keine Statistik gerechnet werden,

Länge (erstes Konjunkt)	Koordinierte Strukturen	Parallele Vorkommen Koordination	Zufall
1	72	66,7%	26,4%
2	43	44,2%	16,3%
3	20	30,0%	0%
4	10	40,0%	0%
5	19	5,3%	0%
6	11	0%	0%
7	13	0%	0%
8	17	0%	0%
9	16	6,3%	0%
10	9	0%	0%
11	2	0%	0%
12	3	0%	0%
13	4	0%	0%
14	1	0%	0%
15	1	0%	0%
16	1	0%	0%
Alle Längen	242	32,6%	10,7%

Tabelle 5.3: Verteilung der parallelen Vorkommen nach Länge

da jeweils kein paralleles Vorkommen im Zufalls-Datensatz existiert und somit eine Vorbedingung für den Chi-Quadrat-Test verletzt ist. Die Differenz zwischen der Anzahl der parallelen Vorkommen im Koordinations- und im Zufalls-Datensatz ist signifikant für Länge 1 ($\chi^2(1) = 12,9$; $p < 0,01$) und für Länge 2 ($\chi^2(1) = 10,6$; $p < 0,01$). Dieses Ergebnis zeigt, dass die große Anzahl an parallelen Vorkommen bei Konjunktlänge 1 nicht ausschließlich auf ihre begrenzten kombinatorischen Eigenschaften zurückzuführen ist. Entsprechendes gilt für die Konjunktlängen 2 bis 4.

Schließlich sollte überprüft werden, ob ein genereller Längeneffekt in den Korpusdaten zu finden ist, d.h. ob der Effekt bei kürzeren Konjunkten stärker ist als bei längeren Konjunkten. Aus diesem Grund wurde für jede Länge ein normalisiertes Parallelitätsmaß berechnet, d.h. die Differenz zwischen der Anzahl der parallelen Vorkommen im Koordinations-Datensatz und der erwarteten zufälligen Verteilung (Anzahl paralleler Vorkommen im Zufalls-Datensatz) für die jeweilige Länge. In Abbildung 5.4 ist die normalisierte Parallelität (in Prozent) für jede Länge (1–9) dargestellt.

Wie in Abbildung 5.4 zu sehen ist, weisen die Längen 1 bis 4 ein relativ hohes Maß an Parallelität auf. Ab Länge 5 nimmt die Parallelität drastisch ab. Statistisch unterscheiden sich die Parallelitätswerte der Länge 1 bis 4 nicht voneinander ($\chi^2(3) = 2,2$; $p = 0,5$), d.h. es gibt keinen Längeneffekt für Länge 1 bis 4. Vergleicht man jedoch die Längen 1–4 mit den Längen 5–9, ist die Differenz signifikant ($\chi^2(1) = 29,0$; $p < 0,01$). Das heißt, es gibt einen Längeneffekt für kurze (Länge 1–4) vs. lange Konjunkte (Länge 5–9). Lange Konjunkte sind selten strukturidentisch, was aufgrund der Eigenschaften des Arbeitsspeichers zu erwarten ist. Dadurch ergibt sich möglicherweise auch eine Erklärung für die fehlenden parallelen Vorkommen bei *S*-Koordination, da die Konjunkte dieser Kategorie im Allgemeinen länger sind (im Durchschnitt 8,1 Wörter). Ich vermute, dass die fehlende Parallelität bei *VP*-Koordination auf eine Kombination aus hoher struktureller Variationsbreite und Länge zurückzuführen ist.

In diesem Abschnitt wurde gezeigt, dass der Parallel-Struktur-Effekt im Phrasenaufbau nicht nur bei der Verarbeitung sondern auch in Korpusdaten zu finden ist. Im Korpus kommt strukturelle Parallelität signifikant häufiger in koordinierten Strukturen vor als zwischen zwei Phrasen, die nicht koordiniert

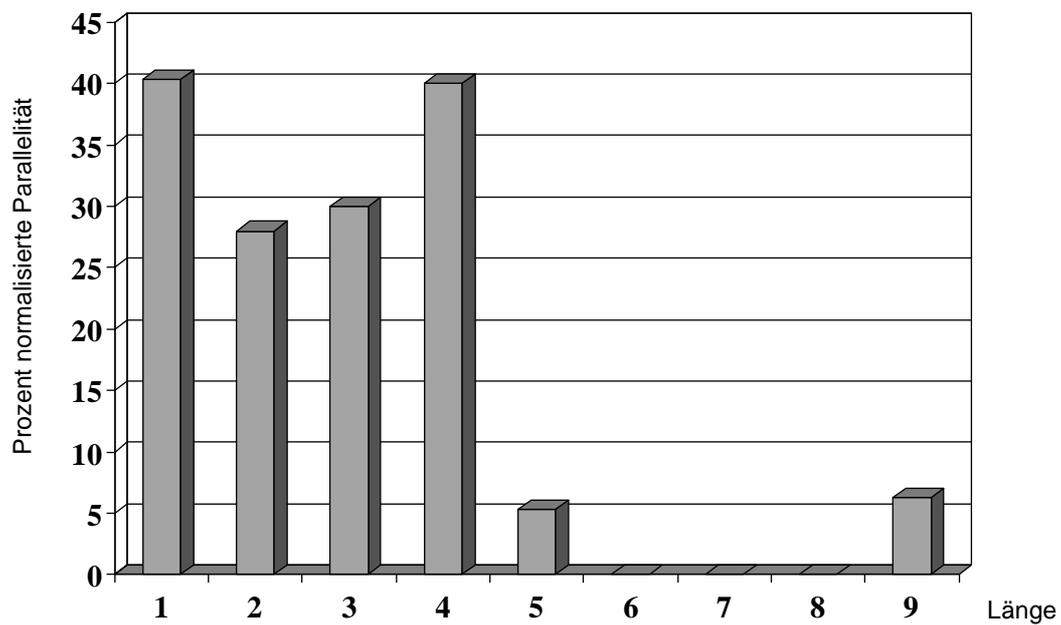


Abbildung 5.4: Normalisierte Parallelität für Länge 1–9 (*Normalisierte Parallelität*: Differenz zwischen der Anzahl der parallelen Vorkommen im Koordinations- und im Zufalls-Datensatz; *Länge*: Anzahl Wörter des ersten Konjunks)

sind. Diese Korrelation zwischen Verarbeitungspräferenz und Korpusfrequenzen war nach der Tuning-Hypothese zu erwarten. Desweiteren wurde gezeigt, dass der Parallel-Struktur-Effekt im Phrasenaufbau generell im Korpus bei koordinierten Strukturen auftritt und nicht nur bei einer bestimmten syntaktischen Kategorie oder Konjunktlänge. Der Parallel-Struktur-Effekt existiert beispielsweise für unterschiedliche Konjunktlängen, d.h. der Effekt ist nicht nur auf die kombinatorischen Eigenschaften kurzer Konjunkte zurückzuführen. Außerdem wurde ein genereller Längeneffekt für kurze (Länge 1–4) vs. lange Konjunkte (Länge 5–9) gefunden. Im untersuchten Korpusfragment kommen strukturidentische Konjunkte hauptsächlich bis zu einer Konjunktlänge von vier Wörtern vor. Dies ist vermutlich der begrenzten Kapazität des Arbeitsspeichers zuzuschreiben. Die Implikationen des Längeneffekts werden am Ende des Kapitels in Abschnitt 5.7 diskutiert.

5.5 Präferenzen bei lokalen Ambiguitäten

Der dritte Verarbeitungseffekt, der in diesem Kapitel untersucht werden soll, betrifft die Präferenz für NP-Koordination (gegenüber S-Koordination) in lokal ambigen Strukturen. Frazier (1979) verglich die Lesezeiten von Sätzen, die zwei koordinierte Nominalphrasen (8a) bzw. zwei koordinierte Sätze (8b) enthielten. Bei der Anbindung der Konjunktion in (8a,b) tritt eine lokale Ambiguität auf, die sich erst bei der Verarbeitung des letzten Wortes auflöst (siehe Kapitel 3, Abschnitt 3.3.1.1).

- (8) a. Peter kissed [_{NP} Mary] and [_{NP} her sister] too.
 b. [_S Peter kissed Mary] and [_S her sister laughed].

Frazier fand einen Garden-Path-Effekt, d.h. signifikant längere Lesezeiten, auf dem letzten Wort *'laughed'* in (8b) im Vergleich zum letzten Wort *'too'* in (8a). Diese Ergebnisse zeigen, dass der Parser eine Präferenz hat, die Nominalphrase [*her sister*] als Teil einer koordinierten Objekt-NP zu interpretieren und nicht als Beginn eines neuen Satzes.

5.5.1 Korpusanalyse III: NP- vs. S-Koordination

Um herauszufinden, ob die Präferenz für NP-Koordination (im Vergleich zur Satz-Koordination) auch in den Korpusdaten zu finden ist, wurde ein größerer Anteil der TÜBA-E-Baumbank ausgewertet. Zu dem bisher verwendeten Korpusfragment (CD13: 2906 Sätze) wurde ein weiteres Fragment (CD 6: 2554 Sätze) hinzugenommen, um eine größere Datenbasis für die relevante Konstruktion zu erhalten.

Aus dem Korpus wurden alle Vorkommen der Form “...*NP_{Subj}*...*Verb*...*NP_{Obj}* and...” (Basis-Konstruktion) extrahiert, die außerdem eine semantisch mögliche Weiterführung mit einer Nominalphrase (NP-Koordination) oder einem Satz (S-Koordination) erlauben. In (9a) ist eine relevante Konstruktion aus den Korpusdaten dargestellt, die mit einer Nominalphrase fortgesetzt wird. Das Beispiel in (9b) dagegen wird mit einem Satz weitergeführt. In (9c) ist ein Beispiel dargestellt, das von der Auswertung ausgeschlossen wurde, weil eine Weiterführung als S-Koordination aufgrund der Semantik des Verbs ‘*mix*’ nicht möglich ist.

- (9) a. and I will bring [*NP* the doughnuts] and [*NP* coffee] I guess
- b. and [*S* Friday I have a nine to ten meeting] and [*S* I also have a meeting in the early afternoon]
- c. I hate to mix [*NP* business] and [*NP* weekends]

Die Korpusauswertung ergab, dass die Basis-Konstruktion in 69,2% der Fälle mit einer Nominalphrase fortgesetzt wird (NP-Koordination, $n = 27$). Nur in 30,8% der Fälle wird die Basis-Konstruktion mit einem Satz weitergeführt (S-Koordination, $n = 12$). Unter der Annahme, dass eine zufällige Verteilung bei 50% liegt, ist die Differenz zwischen NP- und S-Fortsetzungen statistisch signifikant ($t(38) = 2,57$; $p < 0,05$). Die Verarbeitungspräferenz für NP-Koordination (gegenüber S-Koordination) in lokal ambigen Strukturen ist also auch in den Korpusdaten zu finden. Diese Korrelation zwischen Verarbeitungspräferenz und Korpusfrequenzen war nach der Tuning-Hypothese zu erwarten.

5.6 Methodische Überlegungen

In diesem Kapitel wurden spontansprachliche Korpusdaten hinsichtlich verschiedener Verarbeitungspräferenzen statistisch ausgewertet. Es muss festgehalten werden, dass diese Ergebnisse vollständig auf der syntaktischen Annotation der TÜBA-E-Baumbank basieren. Die Baumbank wurde von Linguisten manuell annotiert und nicht etwa automatisch durch computationelle Parsingsysteme. Der Vorteil einer manuellen Annotation ist, dass die Fehlerrate in den Annotationen vergleichsweise gering ist. Andererseits wird die Annotation in kritischen Fällen durch die Präferenzen des menschlichen Annotators beeinflusst, und letztlich reflektiert die Annotation natürlich auch das für die Baumbank entwickelte Annotationsschema.

Ein Beispiel für den Einfluss des TÜBA-E-Annotationsschemas auf die in diesem Kapitel vorgestellten Korpusauswertungen ist das Annotationsprinzip ‘*Longest Match Principle*’ (LMP). Nach diesem Prinzip werden möglichst viele Konstituenten in eine einzige syntaktische Struktur integriert, solange die komplette Struktur syntaktisch und semantisch wohlgeformt bleibt (Kordoni, 2000). Wenn eine Äußerung in der Baumbank beispielsweise aus mehreren Sätzen besteht, führt das LMP dazu, dass so viele Sätze wie möglich koordiniert werden. Die Baumbank enthält häufig koordinierte Sätze, die ich als separate Sätze annotiert hätte. Ohne das LMP, würde die Baumbank weniger Satz-Koordinationen enthalten und die beobachtete Präferenz für NP-Koordination (gegenüber Satz-Koordination) wäre vermutlich noch stärker.

5.7 Zusammenfassung und Diskussion

In diesem Kapitel habe ich das Verhältnis von Parsingpräferenzen und Korpusfrequenzen anhand von drei Verarbeitungseffekten untersucht: 1. Präferenz für die Koordination gleicher Kategorien, 2. Parallel-Struktur-Effekt im Phrasenaufbau, 3. Präferenz für NP-Koordination (gegenüber S-Koordination) in lokal ambigen Strukturen. Die Korpusanalysen ergaben, dass alle drei Verarbeitungseffekte auch in spontansprachlichen Korpusdaten zu finden sind. Es gibt prinzipiell zwei Möglichkeiten diese Korrelationen zu erklären.

Nach der Tuning-Hypothese ist es einerseits denkbar, dass die statistischen Muster in natürlicher Sprache, wie z.B. relative Häufigkeiten, den menschlichen Parser beeinflussen. In diesem Fall wird die statistische Verteilung in der Sprachproduktion als Ursache für die Ausbildung von Parsingpräferenzen betrachtet.

Eine weitere Möglichkeit, diese Korrelationen zu erklären, ist die Annahme von gemeinsamen zugrundeliegenden Mechanismen für die Sprachproduktion und für das Sprachverstehen. Wie in Kapitel 3 und 4 ausgeführt, können die untersuchten Verarbeitungseffekte im Rahmen von syntaktischen Verarbeitungsmodellen erklärt werden.

Die Präferenz für die Koordination von gleichen Kategorien sowie der Parallel-Struktur-Effekt im Phrasenaufbau können durch einen Recycling-Mechanismus im Rahmen des Iterationsmodells erklärt werden (siehe Kapitel 4, Abschnitt 4.2.1). Nach diesem Ansatz wird die Struktur, die für das erste Konjunkt aufgebaut wurde, bei der Verarbeitung des zweiten Konjunks wiederverwendet. Die Präferenz für NP-Koordination (gegenüber S-Koordination) in lokal ambigen Strukturen kann beispielsweise aus dem Iterationsprinzip im Rahmen des Iterationsmodells abgeleitet werden (siehe Kapitel 4, Abschnitt 4.2.3.1). Dabei wird die kleinere Iterationsschleife präferiert, was zu einer Präferenz für kurze Konjunkte führt. Wie in Kapitel 3 ausgeführt, kann die Präferenz für NP-Koordination jedoch auch alternativ aus dem Minimal-Attachment-Prinzip (Frazier, 1979; Frazier und Clifton, 1996) abgeleitet werden, wonach die Struktur mit den wenigsten syntaktischen Knoten bevorzugt wird, oder aus der allgemeinen Recency-Präferenz (Gibson, Pearlmutter, Canseco-Gonzalez und Hickok, 1996).

In diesen syntaktischen Verarbeitungsmodellen ist die präferierte Konstruktion im Allgemeinen ökonomischer und daher leichter und schneller zu verarbeiten. Unter der Annahme, dass dieselben Mechanismen auch für die Sprachproduktion zur Verfügung stehen, würden die Konstruktionen, die leichter verarbeitet werden, auch leichter und somit vermutlich häufiger produziert werden. Gemeinsame Mechanismen für Sprachproduktion und Sprachverstehen würde zu schnelleren Lesezeiten bei der Verarbeitung und zu höheren Frequenzen bei der Produktion führen.

Im Rahmen diese Studie ist es nicht möglich zwischen der Tuning-Hypothese auf der einen Seite und den gemeinsamen zugrundeliegenden Mechanismen auf der anderen Seite zu unterscheiden. Es konnte jedoch gezeigt werden, dass das Verhältnis zwischen Sprachproduktion und Sprachverstehen enger ist, als erwartet. Sollte eine generelle Korrelation zwischen den Präferenzen in diesen beiden Datentypen bestehen, würde das zu interessanten Schlussfolgerungen führen. Dann könnten beispielsweise Produktionsdaten aus Korpora verwendet werden, um Sprachverstehensmodelle, zumindest qualitativ, zu evaluieren.

Neben den Korrelationen für die untersuchten Verarbeitungseffekte wurde in diesem Kapitel außerdem ein Längeneffekt für kurze vs. lange Konjunkte im Korpus gefunden. Die Länge des ersten Konjunks scheint ein relevanter Faktor für den Parallel-Struktur-Effekt im Phrasenaufbau zu sein, d.h. dieser Effekt tritt, zumindest in Korpora, nur bei kurzen Konjunkten auf. Sollte tatsächlich eine generelle Korrelation zwischen den Präferenzen bei Produktion und Verstehen bestehen, dann könnte der Längeneffekt verwendet werden, um Verarbeitungsmodelle zu evaluieren. Dies wird im Folgenden ausgeführt.

Im Iterationsmodell ergibt sich der Parallel-Struktur-Effekt durch die Wiederverwendung von parallelen syntaktischen Strukturen. Die Voraussetzung dafür ist jedoch, dass während der Verarbeitung des zweiten Konjunks die Struktur des ersten Konjunks vollständig im Arbeitsspeicher gehalten werden kann. Dies wird zunehmend schwieriger, je länger das erste Konjunkt ist. Der beobachtete Längeneffekt lässt sich letztlich aus dem Iterationsprinzip (siehe Kapitel 4.1) ableiten. Dieses Prinzip für kleine Schleifen im Iterationsmodell ist durch die begrenzte Kapazität des Arbeitsspeichers motiviert und sagt eine Präferenz für kurze Konjunkte voraus. Der beobachtete Längeneffekt lässt sich

jedoch nicht aus dem Copy- α -Modell ableiten, da dort der Verarbeitungsaufwand beim Kopieren unabhängig von der Größe der zu kopierenden Struktur, d.h. der Konjunkte, ist. Aus dem Copy- α -Modell ergeben sich hinsichtlich der Konjunktlänge keinerlei Präferenzen (zu den unterschiedlichen Vorhersagen der beiden Modelle siehe auch Kapitel 4.4). Der beobachtete Längeneffekt würde folglich Evidenz für das Iterationsmodell und gegen das Copy- α -Modell liefern.

Die in diesem Kapitel untersuchten Korpusdaten sind gleichzeitig auch die Trainingsdaten für die Simulationen mit neuronalen Netzen, welche im folgenden Kapitel beschrieben werden. Die statistische Verteilung der koordinierten Strukturen in den Trainingsdaten ist relevant für die Interpretation der Simulationsergebnisse.

Kapitel 6

Simulationen mit künstlichen neuronalen Netzen

In diesem Kapitel soll die syntaktische Verarbeitung koordinierter Strukturen im Rahmen eines kognitiven Systems simuliert werden. Die Verarbeitungsmechanismen für koordinierte Strukturen wurden bereits in Kapitel 4 untersucht. In Kapitel 5 wurde gezeigt, dass die Verarbeitungseffekte, die beim Sprachverstehen auftreten, auch in der Sprachproduktion in Form von Korpshäufigkeiten zu finden sind. In diesem Zusammenhang ergibt sich die interessante Fragestellung, ob ein kognitives System, das über keinerlei linguistisches Wissen verfügt, in der Lage ist, entsprechende Verarbeitungspräferenzen aus natürlich-sprachlichem Input zu entwickeln. Dieser Fragestellung gehe ich in diesem Kapitel nach.

Für die Simulationen werden künstliche neuronale Netze verwendet. Ein neuronales Netzwerk ist ein kognitives System, das in der Lage ist, selbständig aus Erfahrung zu lernen, ohne dass es hierfür explizit programmiert werden muss (siehe z.B. Zell, 2000). Motiviert sind neuronale Netze durch ihre grobe Analogie zur biologischen Verarbeitung im Gehirn. Neben seiner Lernfähigkeit hat ein neuronales Netz viele positive Eigenschaften, wie z.B. die ökonomische parallele Verarbeitung, die Möglichkeit einer verteilten Wissensrepräsentation, was wiederum eine höhere Fehlertoleranz des Gesamtsystems zur Folge hat, die assoziative Speicherung von Informationen sowie die Generalisierungsfähigkeit auf ähnliche unbekannte Muster. Die große Herausforderung besteht darin,

ein Netzwerk zu entwickeln, das selbständig lernt, eigene strukturelle Repräsentationen aus Trainingsbeispielen aufzubauen. Ein solches Netz verfügt vor dem Lernprozess über keinerlei linguistisches Wissen, es lernt letztlich, statistische Zusammenhänge in den Trainingsdaten zu erkennen.

Das Ziel dieses Kapitels ist es, herauszufinden, ob neuronale Netze in der Lage sind, selbständig grammatische Regularitäten zu lernen und anschließend ein ähnliches Verhalten, d.h. die gleichen Effekte bei der Verarbeitung von Koordination, aufweisen wie Erwachsene beim Lesen. Im Gegensatz zu den Trainingsdaten in der Literatur sollen in dieser Arbeit echte spontansprachliche Korpusdaten verwendet werden, die bereits in Kapitel 5 quantitativ ausgewertet worden sind. Dieses Kapitel ist wie folgt strukturiert: In Abschnitt 6.1 wird beschrieben, wie künstliche neuronale Netze aufgebaut sind und in Abschnitt 6.2 wie diese Netze lernen. Der Forschungsstand zu Simulationen der syntaktischen Verarbeitung wird in Abschnitt 6.3 aufgezeigt. Schließlich stelle ich in Abschnitt 6.4 eigene Simulationen zum Lernen grammatischer Regeln vor und teste die Netze auf ihr Verhalten bei der Verarbeitung koordinierter Strukturen.

6.1 Künstliche neuronale Netze

Ein künstliches neuronales Netz besteht aus einer Vielzahl einfacher Einheiten oder Zellen, die miteinander verbunden sind. Die Informationsverarbeitung in einem neuronalen Netz erfolgt über die Aktivierung der Zellen. Im Folgenden wird das klassische Feedforward-Netzwerk beschrieben (siehe z.B. Hornik et al., 1989).

Ein Feedforward-Netz besteht aus mehreren Ebenen bzw. Schichten: einer Eingabeschicht, einer oder mehreren verdeckten Schicht(en) und einer Ausgabeschicht. In einem Feedforward-Netz breitet sich die Aktivierung nur in *eine* Richtung durch das Netzwerk aus, von der Eingabeschicht über die verdeckten Schichten bis hin zur Ausgabeschicht. Es gibt keine Rückkopplungsverbindungen. Alle Verbindungen von der Eingabeschicht bis zur Ausgabeschicht sind trainierbar, d.h. die Stärke (Gewichte) der Verbindungen wird im Laufe des Lernprozesses verändert.

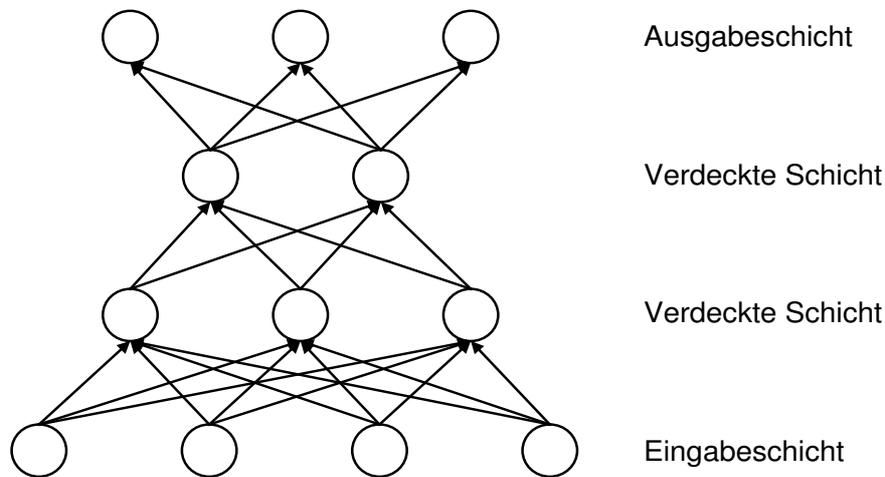


Abbildung 6.1: Ebenenweise verbundenes Feedforward-Netz mit 4 Zellschichten (eine Eingabeschicht, zwei verdeckte Schichten, eine Ausgabeschicht)

Die Zellen der Eingabeschicht nennt man *Eingabeneurone* ('input units'), die Zellen der Ausgabeschicht werden als *Ausgabeneurone* ('output units') bezeichnet. Entsprechend werden die Zellen der verdeckten Schichten *verdeckte Neurone* ('hidden units') genannt.

Ein *ebenenweise verbundenes* Feedforward-Netzwerk ist ein Feedforward-Netz, bei dem jede Ebene *vollständig* mit der nächsten Ebene verbunden ist, d.h. jede Zelle der Eingabeschicht ist mit jeder Zelle der ersten verdeckten Schicht verbunden. Im Falle mehrerer verdeckten Schichten ist jede Zelle der ersten verdeckten Schicht mit jeder Zelle der zweiten verdeckten Schicht verbunden usw. Außerdem ist jede Zelle der letzten verdeckten Schicht mit jeder Zelle der Ausgabeschicht verbunden (siehe Abb. 6.1).

Die einzelnen Zellen in einem Netzwerk haben im Allgemeinen die folgenden Eigenschaften (siehe Rumelhart, Hinton und McClelland, 1986): Jede Zelle j besitzt einen Aktivierungszustand a_j , der den Grad der Aktivierung angibt. Der Aktivierungswert einer Zelle kann dabei z.B. Werte zwischen 0 und 1 annehmen. Der Aktivierungszustand wird aus der Eingabe in die Zelle j (Netzeingabe net_j) mit Hilfe einer Aktivierungsfunktion berechnet, wie

z.B. der logistischen Aktivierungsfunktion¹. Die Aktivierung einer Zelle a_j entspricht nicht automatisch der Ausgabe o_j dieser Zelle. Dies wird durch die Ausgabefunktion festgelegt. Häufig wird jedoch die identische Abbildung als Ausgabefunktion verwendet, dann ist $a_j = o_j$. Die Verbindungen zwischen den Zellen haben eine bestimmte Stärke (Gewicht). Das Gewicht der Verbindung von Zelle i nach Zelle j wird mit w_{ij} bezeichnet.

Um für ein Eingabemuster die entsprechende Ausgabe zu erzielen, wird die Aktivierung der Zellen Schicht für Schicht berechnet (*Vorwärtspropagierung*). Die Präsentation eines Eingabemusters erfolgt durch die entsprechende Aktivierung der Eingabezellen. Die Aktivierung der verdeckten Zellen und der Ausgabezellen wird schichtweise mit Hilfe einer Propagierungsfunktion (6.1) berechnet.

$$net_j(t) = \sum_i o_i(t) w_{ij} \quad (6.1)$$

Die Standard-Propagierungsfunktion in (6.1) gibt an, wie sich die Netzeingabe net_j einer Zelle j aus den Ausgaben der Vorgängerzellen o_i und den Verbindungsgewichten w_{ij} von Zelle i nach Zelle j berechnet. Die resultierende Aktivierung der Ausgabezellen repräsentiert das Ausgabemuster zum Zeitpunkt t .

6.2 Lernen in neuronalen Netzen

Ein zentraler Bestandteil von künstlichen neuronalen Netzen ist ihre Lernfähigkeit, was sie aus kognitiver Perspektive besonders interessant macht. Lernen in neuronalen Netzen bedeutet, dass die Stärke (oder das Gewicht) der Zellverbindungen verändert werden. Dieser Lernprozess, d.h. die Modifikation der Zellverbindungen, folgt einem allgemeinen Lernalgorithmus bzw. Lernverfahren. Während des Lernens werden dem Netz wiederholt Trainingsbeispiele präsentiert. Nach dem Training sind die Zellverbindungen angepasst an die Aufgabenstellung und an die Trainingsdaten. Im Idealfall ist das trainierte Netz in der Lage, zu jedem Eingabemuster aus den Trainingsdaten

¹ Die logistische Aktivierungsfunktion: $f(x) = \frac{1}{1+e^{-x}}$

auch die korrekte Ausgabe zu liefern. Hat das Netz die Regularitäten in den Trainingsdaten erfolgreich gelernt, kann es auch für unbekannte ähnliche Muster die richtige Ausgabe produzieren (Generalisierung).

Im Folgenden wird die Hebbsche Lernregel beschrieben (Abschnitt 6.2.1), die die Grundlage für komplexere Lernverfahren darstellt. In Abschnitt 6.2.2 beschreibe ich den Unterschied zwischen überwachtem und unüberwachtem Lernen und zeige, dass es für beide Lernarten natürliche Beispiele gibt. Anschließend wird das in dieser Arbeit verwendete überwachte Lernszenario beschrieben und begründet.

6.2.1 Die Hebbsche Lernregel

Donald Hebb beschrieb bereits 1949 einen Mechanismus zur synaptischen Plastizität, welche die neurophysiologische Grundlage für Lernen und Gedächtnis darstellt. Dieser Lernmechanismus wird als *Hebbsche Lernregel* bezeichnet.

“When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.”

(Hebb, 1949, S. 62)

In künstlichen neuronalen Netzen wird diese Veränderung der synaptischen Übertragung als Änderung der Verbindungsstärke zwischen Neuronen abgebildet. Je häufiger also ein Neuron A ein Neuron B aktiviert, desto stärker wird die Verbindung zwischen den beiden Neuronen.

Sollen wie bei der klassischen Konditionierung zwei Muster (repräsentiert durch Neuron A und B) miteinander assoziiert werden, so gelingt es A vor dem Lernen jedoch im Allgemeinen nicht, B zu aktivieren. Nehmen wir als Beispiel das Experiment mit dem Pawlowschen Hund (Pawlow, 1905), bei dem ein Glockenton (Muster A: neutraler Reiz) und der Speichelfluss (Muster B: Reaktion) miteinander assoziiert wurde. Vor dem Lernen wird der Glockenton alleine den Speichelfluss nicht auslösen können. Ein externer (unkonditionierter) Stimulus durch die Außenwelt ist erforderlich, hier die Gabe des Futters, der

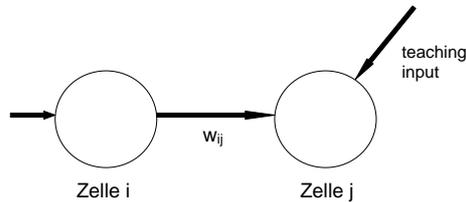


Abbildung 6.2: *Zelle i*: Muster A (z.B. Glockenton), *Zelle j*: Muster B (z.B. Speichelfluss), *teaching input* (optional): z.B. Futtergabe, w_{ij} : Stärke der Verbindung zwischen Zelle i und Zelle j

dann Muster B (den Speichelfluss) aktiviert. Dieser externe Stimulus wird auch als ‘Lehrer’ (*‘teaching input’*) bezeichnet, da durch diesen Stimulus vorgegeben wird, wie die erwünschte Reaktion aussehen soll. Erst nach wiederholter zeitgleicher Präsentation des Glockentons und des Futters (externer Stimulus), wird der Glockenton schließlich alleine den Speichelfluss auslösen können (siehe Abbildung 6.2).

Die Hebbsche Lernregel wird daher meist in etwas veränderter Form formuliert:

“When unit A and B are simultaneously excited, increase the strength of the connection between them.”
(McClelland et al., 1986, S. 36)

Diese allgemeinere Formulierung der Hebbschen Lernregel wird auch umschrieben als *‘Cells that fire together, wire together’*. Dabei bleibt es offen, ob Neuron B durch Neuron A oder durch einen externen Stimulus (“Lehrer”) aktiviert wurde.

Die mathematische Formulierung der Hebbschen Lernregel in ihrer einfachsten Form ist in (6.2) dargestellt, wobei Neuron A der Zelle *i* entspricht und Neuron B der Zelle *j* (siehe Rumelhart, Hinton und McClelland, 1986, S. 53).

$$\Delta w_{ij} = \eta \cdot o_i \cdot a_j \quad (6.2)$$

Hierbei ist Δw_{ij} die Veränderung des Gewichts (Verbindungsstärke) von Neuron i zu Neuron j . η ist eine konstante Lernrate, welche die Lerngeschwindigkeit modelliert, o_i ist die Ausgabe der Vorgängerzelle i und a_j die Aktivierung der Nachfolgerzelle j . Sind die Zellen i und j gleichzeitig aktiv, so wird die Verbindung zwischen Zelle i und j proportional zu ihren Aktivierungen verstärkt.

Eine Variante der Hebbschen Lernregel ist die *Delta-Regel* in (6.3), in der die erwünschte Aktivierung der Nachfolgerzelle j (Zielreiz) durch einen Lehrer explizit berücksichtigt wird (siehe Rumelhart, Hinton und McClelland, 1986, S. 53).

$$\Delta w_{ij} = \eta \cdot o_i \cdot (t_j - a_j) = \eta \cdot o_i \cdot \delta_j \quad (6.3)$$

Hierbei ist t_j die erwünschte Aktivierung (*teaching input*) und δ_j die Differenz zwischen erwünschter und tatsächlicher Aktivierung der Zelle j . Die tatsächliche Aktivierung a_j beschreibt, inwieweit die Vorgängerzelle i in der Lage ist, die Nachfolgerzelle j alleine zu aktivieren. Die Differenz zur erwünschten Aktivierung wird als Fehler betrachtet. Bei der Delta-Regel wird die Verbindungsstärke w_{ij} proportional zu diesem Fehler δ_j verändert.

Für assoziatives Lernen kann die einfache Hebbsche Lernregel in (6.2) oder die Delta-Regel in (6.3) verwendet werden. Bei beiden Lernregeln wird für die Gewichts Anpassung nur lokale Information benötigt, d.h. die Aktivierungen der Vorgänger- und Nachfolgerzelle, wobei die Nachfolgerzelle unter Umständen zusätzlich zur Eingabe aus der Vorgängerzelle auch noch Aktivierung durch den "Lehrer" erhält (*teaching input*). Für das Erlernen komplexerer Aufgaben werden jedoch größere, meist mehrschichtige, Netzwerke benötigt, die mit komplexeren Lernverfahren, wie z.B. *Backpropagation* (Rumelhart, Hinton und Williams, 1986), trainiert werden müssen. Die meisten dieser Lernregeln gehen jedoch im Kern auf die Hebbsche Lernregel und ihre Varianten zurück.

6.2.2 Überwachtes vs. unüberwachtes Lernen

Generell unterscheidet man überwachtes und unüberwachtes Lernen (siehe z.B. Sathya und Abraham, 2013). Grob kann man überwachtes Lernen als “zielgerichtet” (mit Zielreiz) umschreiben und unüberwachtes Lernen als “ungerichtet” (ohne Zielreiz). Es handelt sich dabei um die Art der zu lernenden Aufgabe.

6.2.2.1 Überwachtes Lernen

Beim überwachten Lernen werden dem neuronalen Netz Paare von Eingabe- und Ausgabemustern präsentiert, die miteinander assoziiert werden sollen. Durch die vorgegebenen Ausgabemuster ist ein Zielreiz vorhanden. Beispiele für überwachtes Lernen sind meiner Ansicht nach die klassische Konditionierung (siehe Abschnitt 6.2.1) oder das Lernen von Objektbegriffen im Spracherwerb. Beim Erwerb von Objektbegriffen wird ein visueller Stimulus, wie z.B. ein Auto, mit der dazugehörigen Wortform (auditorischer Stimulus) assoziiert. Der visuelle Stimulus ist dabei das Eingabemuster, und die Lautrepräsentation des Wortes (auditorischer Stimulus) ist das Ausgabemuster, also der vorgegebene Zielreiz. Ohne Zielreiz wäre ein Kind nicht in der Lage, Objektbegriffe zu lernen. Es würde beim Sehen eines Autos nie von alleine, also z.B. durch Raten, darauf kommen, dass dieses Objekt *Auto* heißt.

Beim überwachten Lernen wird dem neuronalen Netz während des Trainings zu jedem Eingabemuster aus den Trainingsdaten das korrekte Ausgabemuster (der Zielreiz) extern durch eine Art “Lehrer” vorgegeben. Aktivierungen aus anderen (z.B. sensorischen) Systemen können dabei die Rolle des sogenannten “Lehrers” übernehmen. Der Lernprozess besteht darin, die Gewichte (Stärke) der Zellverbindungen im Netz so zu verändern, dass das Netz die Assoziation zwischen Eingabe und Ausgabe selbständig durchführen kann. Dieser Lernvorgang folgt einem allgemeinen Lernalgorithmus bzw. Lernverfahren für überwachtes Lernen, wie z.B. *Backpropagation* (Rumelhart, Hinton und Williams, 1986)². Hat das Netz die Regularitäten in den Trainingsdaten erfolgreich gelernt,

² Das Lernverfahren ‘*Backpropagation*’ stellt eine Verallgemeinerung der Delta-Regel (siehe Abschnitt 6.2.1) für mehrschichtige Feedforward-Netze dar.

sollte es auch für unbekannte ähnliche Eingabemuster die korrekte Ausgabe produzieren können (Generalisierungsfähigkeit).

Eine bekannte Aufgabe, bei der neuronale Netze überwacht trainiert wurden, ist der Erwerb der *Past-tense*-Formen für Verben im Englischen (Rumelhart und McClelland, 1986; Plunkett und Marchman, 1991). Als Ausgangspunkt für diese Simulationen dient der typische Erwerbsverlauf des *Past-tense* bei Kindern, der in drei Phasen unterteilt ist (Kuczaj, 1977; Marcus et al., 1992):

1. Phase:

Korrektur Gebrauch weniger regelhafter und unregelmäßiger *Past-tense*-Verbformen,

2. Phase:

Gebrauch einer großen Anzahl meist regelhafter *Past-tense*-Verbformen; Übergeneralisierung von regelhaften *Past-tense*-Formen (z.B. 'comed' oder 'camed' statt 'came'), selbst bei unregelmäßigen Verben, die in der ersten Phase korrekt produziert wurden,

3. Phase:

Korrektur Gebrauch regelhafter und unregelmäßiger *Past-tense*-Formen.

Die Aufgabe, mit der die neuronalen Netze trainiert wurden, besteht darin, einem Verbstamm die passende *Past-tense*-Form zuzuordnen (*sleep* → *slept*, *walk* → *walked*). Die Eingabe besteht aus der phonologischen Repräsentation des Verbstamms, die Ausgabe aus der phonologischen Repräsentation der dazugehörigen *Past-tense*-Form. Das Modell von Rumelhart und McClelland (1986) enthält als Kern ein assoziatives Netzwerk (Feedforward-Netzwerk ohne verdeckte Schicht), das mit der Perzeptron-Lernregel (Rosenblatt, 1962) trainiert wurde. Plunkett und Marchman (1991) trainierten ein Feedforward-Netzwerk mit einer verdeckten Schicht unter Verwendung des Backpropagation-Lernalgorithmus (Rumelhart, Hinton und Williams, 1986). Beide Simulationen zeigen den oben dargestellten typischen U-förmigen Lernverlauf, den Kinder beim Erwerb des *Past-tense* aufweisen.

6.2.2.2 Unüberwachtes Lernen

Beim unüberwachten Lernen werden dem neuronalen Netz ausschließlich Eingabemuster präsentiert. Es werden keine Ausgabemuster vorgegeben, d.h. es gibt keinen Zielreiz und daher auch keinen "Lehrer". Ein Beispiel für unüberwachtes Lernen ist die Kategorisierung von Objekten, wie z.B. die visuelle Unterscheidung von *Baum* und *Auto*. Bei dieser Aufgabe geht es darum, ein visuell dargebotenes Objekt als Exemplar einer bestimmten Objektklasse zu identifizieren. Diese Aufgabe kann gelöst werden, indem die Objekte nach Ähnlichkeiten gruppiert werden. Der visuelle Stimulus ist hierbei das Eingabemuster, ein Zielreiz ist nicht erforderlich.

Beim unüberwachten Lernen organisiert sich ein neuronales Netz selbst. Dabei werden ähnliche Eingabemuster, wie z.B. ein visueller Stimulus, in ähnliche Kategorien gruppiert. Ein bekanntes Verfahren für unüberwachtes Lernen in neuronalen Netzen sind die selbstorganisierenden Karten von Kohonen (Kohonen, 1984, 2001). Die Netzwerkarchitektur besteht aus einem einschichtigen Netzwerk (Kohonenschicht) mit vorgeschalteten Eingabeneuronen. Der Lernprozess besteht darin, ähnliche Eingabemuster auf räumlich benachbarte Neurone der Kohonenschicht abzubilden. Nach erfolgreichem Lernprozess sollte das Netz auch neue Eingabemuster korrekt klassifizieren können (Generalisierungsfähigkeit).

Eine Aufgabe, bei der neuronale Netze unüberwacht trainiert wurden, ist die Erkennung von Wortgrenzen (Anderson, 1999). Diese Simulation basiert auf einem Experiment von Saffran et al. (1996), bei dem acht Monate alte Säuglinge gelernt haben, Wortgrenzen in gesprochener Sprache zu erkennen. In der Trainingsphase des Experiments wurden den Säuglingen zwei Minuten lang zufällig aneinandergereihte dreisilbige Fantasie-Wörter verbal präsentiert (z.B. *tilado*). Die Präsentation enthielt keinerlei Pausen zwischen den Wörtern, d.h. die Säuglinge hatten keine akustischen Hinweise zur Erkennung der Wortgrenzen. Die einzige Information zur Identifikation von Wörtern waren die statistischen Übergangswahrscheinlichkeiten zwischen den Silben. Die Fantasie-Wörter waren so konzipiert, dass jede Silbe innerhalb eines Wortes die nachfolgende Silbe vollständig determinierte ($p = 1$), während nach der letzten Silbe eines Wortes drei verschiedene Silben mit gleicher Wahrscheinlichkeit

folgen konnten ($p = \frac{1}{3}$). Nach der 2-minütigen Präsentation waren die Säuglinge in der Lage, die gehörten Fantasie-Wörter von zuvor nicht präsentierten Fantasie-Wörtern zu unterscheiden.

In Anderson (1999) wird ein Kohonen-Netzwerk präsentiert, das mit den gleichen Daten wie in dem eben beschriebenen Experiment trainiert wurde, d.h. mit Wortlisten, die aus zufällig aneinandergereihten dreisilbigen Fantasie-Wörtern bestanden. Die Eingabemuster bestehen aus aufeinanderfolgenden Silbenpaaren: Silbe 1+2 (Wort 1), Silbe 2+3 (Wort 1), Silbe 3 (Wort 1) + Silbe 1 (Wort 2), etc. Vor dem Training sind die verschiedenen Silbenpaare (insgesamt 20 Paare) auf der Kohonen-Karte zufällig verteilt. Nach dem Lernprozess bilden die Silbenpaare, die innerhalb von Wörtern vorkommen, und die Silbenpaare, die den Wortübergang markieren, jeweils eine kohärente Region auf der Kohonen-Karte. Die beiden Regionen sind jedoch ungleich proportioniert. Silbenpaare innerhalb von Wörtern (8 Silbenpaare von 20 möglichen) besetzen einen größeren Bereich auf der Kohonen-Karte als Silbenpaare am Wortübergang (12 Silbenpaare von 20 möglichen). Das Kohonen-Netzwerk hat sich während des Lernvorgangs die statistischen Eigenschaften der Eingabemuster, d.h. der Silbenpaare, zu Nutze gemacht und ähnliche Eingabemuster auf topologisch benachbarte Neurone abgebildet. Die Silbenpaare mit Übergangswahrscheinlichkeit $p = 1$ (innerhalb von Wörtern) und die Silbenpaare mit Übergangswahrscheinlichkeit $p = \frac{1}{3}$ (am Wortübergang) wurden unterschiedlich auf der Kohonen-Karte repräsentiert. Das Netzwerk ist daraufhin, ebenso wie die Säuglinge, in der Lage, die Fantasie-Wörter in den Trainingsdaten von anderen Wörtern zu unterscheiden.

Generell lässt sich sagen, dass beim überwachten Lernen durch die vorgegebene Ausgabe ein Algorithmus gelernt wird, der die Eingabemuster auf die vorgegebene Ausgabe abbildet. Das Netz hat sozusagen eine konkrete Aufgabe zu lernen. Beim unüberwachten Lernen hat das Netz keine konkrete Aufgabe. Das Netz lernt zwar die Eingabemenge nach Ähnlichkeiten zu klassifizieren, man hat aber keine Kontrolle darüber, nach welchen Eigenschaften klassifiziert wird. Es ist also möglich, dass das Netz einen Algorithmus lernt, der überhaupt nicht der intendierten Aufgabe entspricht. Das ist ein Nachteil

von unüberwachten Lernverfahren, was dazu führte, dass bisher meist überwachte Lernverfahren für die Simulation sprachlicher Fähigkeiten eingesetzt wurden.³

Ein Einwand, der im Zusammenhang mit überwachtem Lernen oft geäußert wird, ist, dass ein externer Lehrer biologisch unplausibel sei und beispielsweise der Spracherwerbsprozess nicht durch ein Feedback von außen erfolgen würde. Da ich in dieser Arbeit ein überwachtes Lernszenario verwende, möchte ich im Folgenden auf diesen Einwand eingehen:

a. Wie oben bereits ausgeführt, gibt es viele Aufgaben, die nur überwacht, das heißt mit vorgegebenem Zielreiz, gelernt werden können. Darunter fallen die Formen des assoziativen Lernens, wie z.B. der Worterwerb oder auch das Lesen und Schreiben (Schriftspracherwerb). Aktivierungen aus sensorischen Systeme können dabei die Rolle des "Lehrers" übernehmen.

b. Zugegebenermaßen mag die technische Umsetzung dieser überwachten Lernprozesse in komplexe Lernalgorithmen, wie z.B. Backpropagation bei mehrschichtigen Netzen, teilweise biologisch nicht mehr plausibel sein. Zumindest gibt es keine direkte Entsprechung zwischen den mathematischen Schritten des Lernalgorithmus und den neurobiologischen Prozessen auf Zellebene. Auf einer etwas abstrakteren Ebene können die überwachten Lernprozesse jedoch recht gut simuliert werden.

c. Das Argument, dass der Spracherwerbsprozess nicht durch ein Feedback von außen erfolgt, kann damit begründet werden, dass Kinder auch nach mehrmaliger Korrektur, d.h. Vorgabe der korrekten Äußerung, immer noch einen falschen Output produzieren. Dies ist jedoch auch beim überwachten Lernen mit neuronalen Netzen der Fall. Lernen ist ein langer Prozess, bei dem erst nach vielen Durchgängen durch ein Vielzahl von Trainingsdaten der korrekte Zielreiz produziert wird. Sicherlich erfolgt der Spracherwerb in der Regel nicht durch explizite Korrektur von außen. Das Kind bekommt jedoch ein *implizites Feedback*. Erwachsene reagieren entweder auf den Inhalt der kindlichen Äußerung in Form

³ An dieser Stelle sei angemerkt, dass die Aufgabe, Wortgrenzen zu erkennen, nicht unbedingt ein unüberwachtes Lernszenario, wie in Anderson (1999), erfordert. Die gleiche Aufgabe kann genauso gut überwacht gelernt werden. Anstelle der Silbenpaare als Eingabemuster würde man eine Silbe als Eingabemuster und die folgende Silbe als Ausgabemuster verwenden.

von erweiterten Wiederholungen und klärenden Nachfragen. Oder das Feedback bzw. der Zielreiz besteht in den Äußerungen der Eltern, die immer wieder in einem bestimmten Kontext produziert werden, wie z.B. das Äußern eines Wortes während zeitgleich auf ein Objekt gezeigt wird.

Es stellt sich nun die Frage, ob auch syntaktische Fähigkeiten durch implizites Feedback überwacht gelernt werden. In Elman (1990, 1991) wird speziell für das Lernen von Sequenzen und insbesondere auch für den Syntax-Erwerb ein überwacht Lernszenario präsentiert, welches auch in dieser Arbeit verwendet wird. Dabei soll für jedes Wort in einem Satz das nachfolgende Wort vorhergesagt bzw. antizipiert werden ('sentence prediction task', im Folgenden *Satzantizipations-Aufgabe*). Das Folgewort wird als Zielreiz betrachtet. Die Sätze werden somit auf inkrementelle Weise verarbeitet wie es auch bei der menschlichen Satzverarbeitung der Fall ist.

Bei dieser Aufgabe ist die Rolle des Lehrers minimiert, da der Zielreiz im nächsten Moment durch die Umgebung gegeben wird. Das Feedback erfolgt implizit durch die Erfüllung (positive Evidenz) oder Verletzung (negative Evidenz) einer Erwartungshaltung, die für das nachfolgende Wort aufgebaut wurde. Durch die indirekte negative Evidenz kann außerdem das *Baker's Paradox*⁴ (Baker, 1979; Pinker, 1989) aufgelöst werden, d.h. das Lernen einer natürlich-sprachlichen Grammatik ist auch ohne die Annahme von angeborenen grammatischen Prinzipien möglich. Es entsteht ein natürliches und psychologisch plausibles Lernszenario.

In der Literatur wurden syntaktische Fähigkeiten bisher überwacht gelernt, entweder durch explizites oder durch implizites Feedback. Im Folgenden gehe

⁴ Die sprachlichen Daten, denen Kinder im Spracherwerb ausgesetzt sind, reichen möglicherweise nicht aus, um eine Grammatik zu erschließen bzw. zu lernen (*Poverty-of-the-Stimulus-Argument*). Ist ein Hörer ausschließlich positiven Daten ausgesetzt, können nur reguläre Sprachen gelernt werden (Gold, 1967), natürliche Sprachen sind jedoch komplexer (Chomsky, 1957). Weiterhin wird angenommen, dass Kinder nicht über negative Evidenz lernen. Das Paradoxe ist, dass Kinder trotzdem Sprache lernen. Daraus wird im Allgemeinen geschlussfolgert, dass eine Universalgrammatik angeboren sein muss (Chomsky, 1980). An dieser Argumentationskette kann in vielerlei Hinsicht Kritik geübt werden. Eine Möglichkeit zur Auflösung des Paradoxons besteht beispielsweise darin, zu zeigen, dass beim Spracherwerb indirekte negative Evidenz zur Verfügung steht.

ich auf die Simulationen zur syntaktischen Verarbeitung in der Literatur ein und konzentriere mich dabei auf die Simulationen unter Verwendung der oben genannten Satzantizipations-Aufgabe ('sentence prediction task').

6.3 Simulationen zur syntaktischen Verarbeitung

In diesem Kapitel werden Simulationen aus der Literatur vorgestellt, bei denen Netzwerke überwacht lernen, eigenständig strukturelle Repräsentationen für Sätze aufzubauen. Modelle dieser Art können in zwei Klassen eingeteilt werden.

In der ersten Klasse werden syntaktische Strukturen **explizit** gelernt. Dabei werden Netzwerke trainiert, einem Satz eine bestimmte grammatische Struktur zuzuordnen. Anschließend soll das Netz in der Lage sein, dies auch für unbekannte (neue) Sätze zu tun (Chalmers, 1990; Pollack, 1990). Der Nachteil dieser Modelle ist, dass die grammatischen Strukturen dem Netz als erwünschte Ausgabe präsentiert werden und vor dem Trainingsprozess bereits feststehen muss, wie die syntaktischen (Ziel-)Strukturen im Einzelnen auszusehen haben. Diese Modelle haben große Ähnlichkeit mit statistischen symbolischen Ansätzen zum Erlernen von Grammatik aus geparsten Korpora (Brill et al., 1990; Charniak, 1996, 1997). Auch hier sind die syntaktischen Zielstrukturen und damit die grammatischen Regeln durch die geparsten Korpora bereits vorgegeben.

In der zweiten Klasse von Modellen werden syntaktische Strukturen **implizit** gelernt. Die (Ziel-)Strukturen werden dem Netz *nicht* als erwünschte Ausgabe präsentiert, sondern sie werden implizit aus der Sequenz von Wörtern gelernt (Elman, 1990, 1991; Christiansen und Chater, 1999, 2001; Towsey et al., 1998). Die Aufgabe besteht darin, für eine gegebene Sequenz von Wörtern (Kontext) das nächste Wort vorherzusagen bzw. zu antizipieren (Satzantizipations-Aufgabe). Diese Aufgabe kann nur dann korrekt ausgeführt werden, wenn das Netz die grammatischen Regularitäten gelernt hat, die den Trainingssätzen zugrunde liegen. Ohne grammatisches Wissen kann ein Netz beispielsweise nicht antizipieren, dass nach einem Determinierer (*der, die, das, ... etc.*) ein Nomen (*Haus, Mann, ... etc.*) oder ein Adjektiv (*schön, blau, ... etc.*) folgen kann, aber niemals ein Verb (*gehen, essen, ... etc.*). Auf diese Weise lernt das Netz grammatische Regeln indirekt aus den

Wortsequenzen. Die Satzantizipations-Aufgabe wurde erstmals von Elman (1990) präsentiert, gemeinsam mit einer neuen Netzwerkarchitektur (Elman-Netzwerk), die für das Erlernen von Sequenzen besonders gut geeignet ist.

Ich verfolge in dieser Arbeit den zweiten Ansatz des *impliziten* Lernens von syntaktischen Strukturen, da dies ein wesentlich plausibleres Lernszenario darstellt (siehe oben). Die Trainingsdaten bestehen ganz einfach aus Sätzen, d.h. Sequenzen von Wörtern, die inkrementell abgearbeitet werden. Der einzige Zielreiz ('teaching input'), der bei dieser Aufgabe für das Erlernen von grammatischen Regeln benötigt wird, ist das nächste Wort im Satz, ein Zielreiz, der bei der Sprachrezeption automatisch zur Verfügung steht.

Im Folgenden werden die Elman-Netzwerk-Architektur und das dazugehörige überwachte Lernverfahren vorgestellt, welche auch in dieser Arbeit verwendet werden. Anschließend präsentiere ich die Simulationen, die den oben genannten impliziten Lernansatz unter Verwendung der Satzantizipations-Aufgabe und der Elman-Netzwerkarchitektur verfolgen.

6.3.1 Elman-Netzwerk

Ein Elman-Netzwerk ist eine *partiell rekurrente* Netzwerkarchitektur, die von Elman (1990) entwickelt wurde. Während sich reine Feedforward-Netze (ohne Rückkopplungsverbindungen) recht gut für die Klassifikation von einzelnen, meist statischen, Mustern eignen, sind rekurrente Netze (mit Rückkopplungsverbindungen) wesentlich besser geeignet, zeitveränderliche Sequenzen, wie z.B. syntaktische Muster, zu verarbeiten. In einer zeitveränderlichen Sequenz darf ein Muster nicht isoliert betrachtet werden, wie es in einem Feedforward-Netz der Fall wäre, die Klassifikation eines solchen Musters hängt vielmehr vom (zeitlichen) Kontext ab. Für eine korrekte Ausgabe benötigt das Netz nicht nur die aktuelle Eingabe sondern auch zuvor verarbeitete Eingabemuster.

Elman-Netze sind speziell für das Lernen von zeitveränderlichen Sequenzen entwickelt worden und sind in der Lage, Informationen aus vorangegangenen Verarbeitungsschritten zu verwenden. Diese Architektur findet eine Vielzahl von Anwendungen in der Sprache und wurde bisher beispielsweise im Rahmen der Sprach- und Worterkennung eingesetzt (Gaskell und Marslen-Wilson, 1997,

2001), bei der Satzverarbeitung (Elman, 1991, 1993; Christiansen, 1994; Christiansen und Chater, 2001) sowie der Sprachproduktion (Dell et al., 1993, 2001).

6.3.1.1 Netzwerkkonstruktion

Ein Elman-Netz ist ein ebenenweise verbundenes Feedforward-Netz, das mit einer zusätzlichen Kontextschicht ausgestattet ist (siehe Elman, 1990, 1991). Das Netz besteht aus mehreren Ebenen bzw. Schichten: einer Eingabeschicht, einer verdeckten Schicht, einer Ausgabeschicht und einer Kontextschicht (siehe Abb. 6.3). Dabei ist die Eingabeschicht vollständig mit der verdeckten Schicht verbunden, diese ist wiederum vollständig mit der Ausgabeschicht verbunden. Bei den Kontextzellen handelt es sich um spezielle verdeckte Zellen (zu Elman-Netzen siehe auch Zell, 2000, Kap. 11).

Über die zusätzliche Kontextschicht findet eine indirekte Rückkopplung statt, wobei die Rückkopplungsverbindungen von der verdeckten Schicht zur Kontextschicht verlaufen. Jede verdeckte Zelle besitzt genau eine Rückkopplungsverbindung zu einer Kontextzelle mit einem festem Gewicht von 1,0. In einem Elman-Netzwerk ist daher die Anzahl der Kontextzellen gleich der Anzahl der verdeckten Zellen. Außerdem wird die identische Abbildung als Aktivierungsfunktion für die Kontextzellen verwendet. Das führt dazu, dass die Aktivierungen der verdeckten Zellen auf die Kontextzellen kopiert werden.

Die Kontextschicht ist wiederum vollständig mit der verdeckten Schicht verbunden, d.h. jede Zelle der Kontextschicht mit jeder Zelle der verdeckten Schicht. Dadurch wird die Ausgabe der Kontextzellen im nächsten Schritt (zusammen mit der Ausgabe der Eingabeschicht) wieder an die verdeckte Schicht zurückgeleitet. Durch die Kontextschicht entsteht ein Speichermechanismus (ein dynamisches Gedächtnis), der es erlaubt, Informationen aus vorangegangenen Verarbeitungsschritten zu verwenden. Alle Zellverbindungen in einem Elman-Netz (mit Ausnahme der Rückkopplungsverbindungen) sind trainierbar, d.h. die Stärke (Gewicht) dieser Verbindungen werden im Laufe des Lernprozesses verändert.

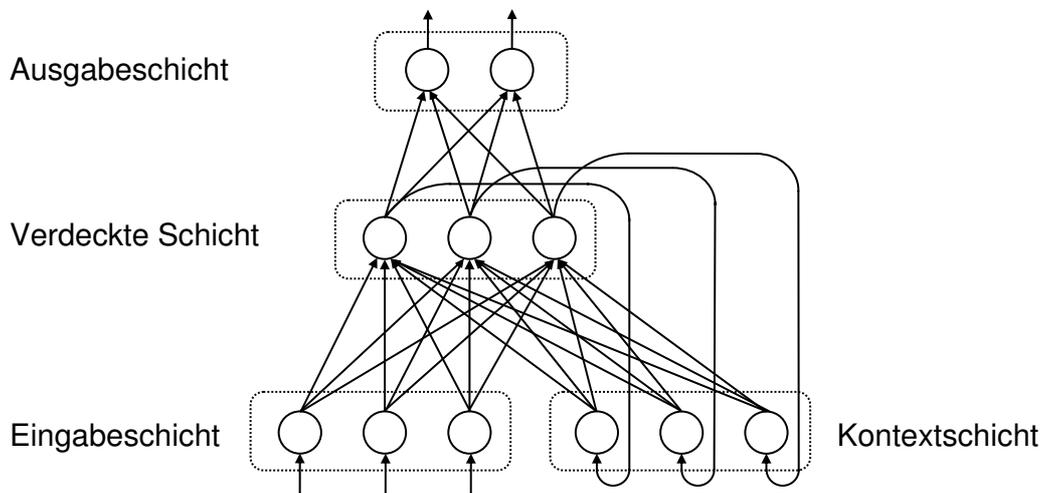


Abbildung 6.3: Partiiell rekurrente Netzwerk-Architektur (Elman-Netz)

6.3.1.2 Aktivierungsausbreitung (Vorwärtspropagierung)

Um für ein Eingabemuster die entsprechende Ausgabe zu erzielen (Vorwärtspropagierung), werden die Aktivierungen der Zellen in einem Elman-Netz wie in einem Feedforward-Netzwerk Schicht für Schicht mit Hilfe der Standard-Propagierungsfunktion⁵ berechnet. Die Aktivierung einer Zelle kann dabei Werte z.B. zwischen 0 und 1 annehmen (zur Berechnung in einem Feedforward-Netzwerk siehe Abschnitt 6.1).

Die Präsentation eines Eingabemusters erfolgt durch die entsprechende Aktivierung der Eingabezellen. Desweiteren werden die Kontextzellen vor dem ersten Durchgang initialisiert. In einem Elman-Netz werden sowohl die Aktivierungen der Eingabezellen als auch die Aktivierungen der Kontextzellen an die verdeckte Schicht geleitet. Zudem werden die Aktivierungen der verdeckten Schicht nicht nur an die Ausgabeschicht weitergeleitet, sondern sie werden zusätzlich an die Kontextzellen weitergeleitet (Rückkopplung) und stehen somit für den nächsten Verarbeitungsschritt, d.h. beim nächsten Eingabemuster, zur Verfügung. Die resultierende Aktivierung der Ausgabezellen repräsentiert das Ausgabemuster.

⁵ Die Standard-Propagierungsfunktion: $net_j(t) = \sum_i o_i(t) w_{ij}$

Beim nachfolgenden Durchgang enthalten die Kontextzellen die verdeckten Aktivierungen des zuvor verarbeiteten Eingabemusters. Durch die Kontextschicht stehen dem Netzwerk Informationen aus vorangegangenen Verarbeitungsschritten bzw. Eingabemustern zur Verfügung.

6.3.1.3 Lernverfahren

Ein untrainiertes Netz wird in der Regel für ein bestimmtes Eingabemuster nicht die erwünschte Ausgabe liefern. Die Verbindungsgewichte sind vor dem Training auf zufällige Werte gesetzt (Initialisierung der Gewichte). Das Ziel des Lernvorgangs ist es, die Stärke der Zellverbindungen (Verbindungsgewichte) so zu verändern, dass das Netz für jedes Eingabemuster die korrekte Ausgabe produziert bzw. die Abweichung von der erwünschten Ausgabe (d.h. der Fehler) möglichst minimal ist. Eine möglichst optimale Konfiguration der Verbindungsgewichte für eine bestimmte Aufgabe lässt sich nicht von Hand einstellen, sondern muss mit Hilfe eines Lernalgorithmus gefunden werden.

Beim überwachten Lernen werden einem Netz nacheinander die einzelnen Trainingsmuster aus dem Trainingskorpus präsentiert. Ein Trainingsmuster besteht dabei aus einem Eingabemuster und der erwünschten Ausgabe (Ausgabemuster). Die Anpassung der Verbindungsgewichte auf Basis dieser Trainingsmuster kann in einem Elman-Netzwerk mit einer leichten Modifikation des normalen Backpropagation-Lernalgorithmus⁶ erreicht werden.

Dabei werden Elman-Netze wie folgt betrachtet (siehe Zell, 2000, S. 143): Lässt man die Rückkopplungsverbindungen zu den Kontextzellen weg, reduziert sich ein Elman-Netz auf ein reines Feedforward-Netz, bei dem die Kontextzellen zusätzliche Eingabezellen sind. Die erweiterte Eingabe besteht dann aus den Aktivierungen der Eingabezellen sowie den Aktivierungen der Kontextzellen.

⁶ Das Lernverfahren 'Backpropagation' stellt eine Verallgemeinerung der Delta-Regel (siehe Abschnitt 6.2.1) für mehrschichtige Feedforward-Netze dar (Rumelhart, Hinton und Williams, 1986).

Der Kern des Backpropagation-Lernalgorithmus für Elman-Netze wird im Folgenden beschrieben (aus Zell, 2000, S. 143).⁷

1. Initialisierung der Kontextzellen
2. Für jedes Trainingsmuster wird folgendes durchgeführt:
 - (a) Anlegen des Eingabemusters und Vorwärtspropagierung bis zur Ausgabe (ohne Beachtung der rekurrenten Verbindungen)
 - (b) Vergleich der tatsächlichen Ausgabe mit der erwünschten Ausgabe und Berechnung des Fehlersignals für jede Ausgabezelle
 - (c) Rückwärtspropagierung der Fehlersignale von den Ausgabezellen bis zu den Eingabezellen (ohne Beachtung der rekurrenten Verbindungen)
 - (d) Berechnung der Gewichtsänderungen mit Hilfe der Fehlersignale (z.B. nach der Backpropagation-Lernregel in (6.4) bzw. (6.5))
 - (e) Adaption der Gewichte
 - (f) Berechnung des Folgezustands der Kontextzellen gemäß ihrer Eingangsverbindungen. Dies ist der einzige Schritt, bei dem die rekurrenten Verbindungen beachtet werden.

In Schritt (2d) wird die Berechnung der Gewichtsänderung $\Delta_p w_{ij}$ für ein Muster p mit Hilfe der Backpropagation-Lernregel oder verwandter lokaler Fehlerminimierungsverfahren durchgeführt. In (6.4) bzw. (6.5) ist die Standard-Backpropagation-Lernregel aufgeführt (Zell, 2000, S. 110), wobei η die Lernrate (Stärke der Gewichts Anpassung), o_{pi} die Ausgabe der Zelle i für ein Muster p , und δ_{pj} das Fehlersignal für Zelle j und Muster p darstellt. Das Fehlersignal δ_{pj} wird nach Formel (6.5) berechnet, wobei t_{pj} den 'teaching input' für Zelle j darstellt.

⁷ Bei dieser Beschreibung handelt es sich um die Online-Trainingsvariante, bei der die Gewichte nach jedem Trainingsmuster angepasst werden. Es ist auch möglich, Elman-Netze mit dem Offline-Verfahren zu trainieren, bei dem die Gewichte erst nach einem vollständigen Durchgang durch das Trainingskorpus verändert werden. In dieser Arbeit wird jedoch die Online-Variante verwendet.

$$\Delta_p w_{ij} = \eta o_{pi} \delta_{pj} \quad (6.4)$$

mit

$$\delta_{pj} = \left\{ \begin{array}{ll} o_{pj} (1 - o_{pj}) (t_{pj} - o_{pj}) & \text{falls } j \text{ Ausgabezelle ist} \\ o_{pj} (1 - o_{pj}) \sum_k \delta_{pk} w_{jk} & \text{falls } j \text{ verdeckte Zelle ist} \end{array} \right\} \quad (6.5)$$

Diese Beschreibung betrifft einen Durchgang durch das Trainingskorpus. Im Allgemeinen wird das Korpus mehrfach durchlaufen, bis der Fehler des Netzes möglichst klein ist.

6.3.2 Erste Simulationen (Elman, 1990, 1991)

6.3.2.1 Elman (1990)

Wie oben bereits erwähnt, präsentierte Elman in diesem Aufsatz erstmals eine neue Netzwerk-Architektur (Elman-Netz, siehe Abschnitt 6.3.1) für das Erlernen von strukturellen Zusammenhängen aus Sequenzen, deren Elemente nicht zeitgleich präsentiert werden (temporale Sequenzen). Elman (1990) geht auf eine sequentielle Version des XOR-Problems ein, auf das Erkennen von Wortgrenzen aus Buchstabensequenzen ('letter prediction task') sowie auf das Erlernen von grammatischen Strukturen aus Wortsequenzen ('sentence prediction task').

Die Grundidee bei diesen Antizipationsaufgaben ist immer, das nächste Element in einer Sequenz vorherzusagen bzw. zu antizipieren. Elman (1990) versucht mit seinem Ansatz, die zeitliche Dimension, die bei vielen kognitiven und insbesondere sprachlichen Aufgaben relevant ist, in ein konnektionistisches Modell zu integrieren. Im Folgenden gehe ich auf die Simulationen zur Satzantizipation ein.

Trainingsdaten und Lernaufgabe: Die Trainingsdaten, die in Elman (1990) für die Satzantizipation verwendet wurden, waren einfachste 2- und 3-Wort-Sätze der Form “Nomen - Verb” (N - V), “Nomen - Verb - Nomen” (N - V - N), wobei das erste Nomen das Subjekt und das zweite Nomen das direkte Objekt darstellte (siehe die Beispiele in (1)).

- | | |
|----------------------|-------------|
| (1) a. woman sleep | (N - V) |
| b. boy move | (N - V) |
| c. dog move mouse | (N - V - N) |
| d. woman smash plate | (N - V - N) |
| e. girl eat cookie | (N - V - N) |

Es wurden insgesamt 13 verschiedene syntaktische und semantische Kategorien für Nomen und Verben verwendet (z.B. Nomen_*menschlich*, Nomen_*Nahrung*, Verb_*transitiv*, Verb_*intransitiv*, Verb_*Perzeption*, Verb_*essen*, etc.). Das Vokabular bestand insgesamt aus 29 Wörtern. Basierend auf diesen Kategorien für Nomen und Verben wurden 15 verschiedene Satz-Schablonen erstellt. Zwei dieser Schablonen sind in (2) illustriert. Mit diesen Schablonen sollte ausgeschlossen werden, dass Sätze generiert werden, in denen z.B. unbelebte Subjekte wie Bücher oder ähnliches Nahrung zu sich nehmen, oder Sätze, in denen Menschen Bücher essen, etc.

- (2) a. **Wort 1:** Nomen_*belebt*, **Wort 2:** Verb_*essen*, **Wort 3:** Nomen_*Nahrung*
 b. **Wort 1:** Nomen_*menschlich*, **Wort 2:** Verb_*intransitiv*⁸

Mit den beiden Schablonen in (2) wurden z.B. die Sätze in (1a, e) erstellt. Insgesamt wurden auf diese Weise 10.000 Sätze zufällig generiert. Alle Sätze wurden ohne Unterbrechung aneinandergereiht, wodurch eine Eingabesequenz aus 27.354 Wörtern entstand.

⁸ Beispiele für intransitive Verben in dem hier verwendeten Datensatz sind ‘*think*’ oder ‘*sleep*’.

Die Trainingsaufgabe bestand darin, für jedes Wort aus der Eingabesequenz das nachfolgende Wort zu antizipieren. In (3) ist illustriert, wie die ersten beiden Trainingsmuster für den Satz ‘*woman smash plate*’ (1d) aussehen würden. Das erste Trainingsmuster besteht aus dem ersten Wort der Eingabesequenz (Eingabe) und dem zweiten Wort der Sequenz (Ausgabe), das zweite Trainingsmuster besteht aus dem zweiten Wort der Sequenz (Eingabe) und dem dritte Wort der Sequenz (Ausgabe), usw. Insgesamt bestand der Trainingsdatensatz aus 27.354 Trainingsmustern.

- | | |
|---|-------------|
| (3) Eingabe: <i>woman</i> , Ausgabe: <i>smash</i> | (1. Muster) |
| Eingabe: <i>smash</i> , Ausgabe: <i>plate</i> | (2. Muster) |

Um die Satzantizipations-Aufgabe erfolgreich lösen zu können, musste das Netz die folgenden Teilaufgaben lernen:

- Kategorisierung der Wörter in Nomen und Verben
- Einteilung der Nomen und Verben in verschiedene syntaktisch oder semantisch motivierte Unterkategorien: Welche Nomen besitzen die Merkmale “belebt”, “menschlich” oder “Nahrung”? Welche Verben besitzen das Merkmal “transitiv”? Usw.
- Strukturelle Zusammenhänge zwischen Wortklassen: Jeder Satz beginnt mit einem Nomen (Subjekt) gefolgt von einem Verb. Nach intransitiven Verben, wie z.B. ‘*sleep*’, folgt kein Objekt. Nach transitiven Verben folgt ein Objekt (je nach Verb optional oder zwingend). Vor einem Verb wie ‘*eat*’ stehen nur belebte Subjekte, usw.

Simulationen: Für diese Simulation wurde ein Elman-Netzwerk verwendet, dessen Ein- und Ausgabeschicht jeweils aus 31 Zellen bestand. Jede dieser Zellen repräsentierte ein Wort. In jedem Trainingsmuster war eine Zelle der Eingabeschicht und eine Zelle der Ausgabeschicht aktiviert, wobei die beiden aktivierten Zellen jeweils dem Eingabe- bzw. dem Ausgabewort entsprachen. Die anderen Zellen waren nicht aktiviert. Die verdeckte Schicht und die Kontextschicht bestand jeweils aus 150 Zellen. Das Netzwerk wurde mit dem

Backpropagation-Lernalgorithmus für Elman-Netze (siehe Abschnitt 6.3.1.3) trainiert. Das Training umfasste sechs Durchgänge (Zyklen) durch den kompletten Trainingsdatensatz.

Evaluation: Die Evaluation des Netzes bei der Satzantizipations-Aufgabe ist komplexer als bei normalen Klassifikationsaufgaben, da die Satzantizipation nicht deterministisch ist. Zu einer Eingabe gibt es mehrere grammatisch mögliche Ausgaben. Nach dem ersten Eingabewort (Subjekt) kann z.B. eine Vielzahl von verschiedenen Verben (= Wort 2) folgen. Nach einem Verb wie *'move'* (= Wort 2) kann ein Objekt folgen (= Wort 3), das nachfolgende Nomen könnte aber genauso gut das Subjekt eines neuen Satzes sein (= Wort 1). Die verschiedenen Ausgabe-Möglichkeiten für ein Eingabewort sind in den meisten Fällen nicht gleich wahrscheinlich, je nachdem wie häufig sie im Trainingskorpus vorkommen. Das Netzwerk lernt vermutlich die relative Wahrscheinlichkeit für jedes der möglichen Folgewörter und aktiviert die entsprechenden Ausgabezellen proportional zu diesen Wahrscheinlichkeiten.

Elman schlägt daher vor, bei der Evaluation des Netzes die Ausgabe nicht mit dem tatsächlichen Folgewort im Satz (= erwünschte Ausgabe beim Training) zu vergleichen, sondern mit den grammatisch möglichen Nachfolgern des Wortes und den damit verbundenen relativen Wahrscheinlichkeiten. Legt man den Wahrscheinlichkeiten Vorkommenshäufigkeiten zugrunde, können diese Werte aus dem Trainingskorpus extrahiert werden. Elman berechnet für jedes Trainingsmuster die absolute bedingte Wahrscheinlichkeitsverteilung, d.h. jedes Wort in einem Satz wird mit allen anderen Sätzen verglichen, die vom Satzbeginn bis zu diesem Wort identisch sind. Auf Basis dieser (Teil-)Sätze werden die Vorkommenswahrscheinlichkeiten für die möglichen Folgewörter berechnet.

Nehmen wir als Beispiel den Satz *'mouse move book'*. Während des Trainings wird dem Netz für das Eingabewort *'move'* das Folgewort *'book'* als erwünschte Ausgabe präsentiert. Für die Berechnung der absoluten bedingten Wahrscheinlichkeitsverteilung werden als Basis alle Sätze genommen, die mit der Sequenz *'mouse move'* beginnen. Dann wird berechnet, welche Wörter mit welcher Häufigkeit folgen (z.B. *book* 10%, *girl* 0%, *plate* 50%, etc.). Während im ursprünglich erwünschten Ausgabevektor nur eine Position aktiviert ist, nämlich diejenige, die das Wort *'book'* repräsentiert (siehe (4a)),

sind im Wahrscheinlichkeitsvektor mehrere Positionen proportional zu ihren Vorkommenshäufigkeiten aktiviert (siehe (4b)).

- (4) a. [1.0, 0.0, 0.0, ...] (erwünschte Ausgabe beim Training)
 b. [0.1, 0.0, 0.5, ...] (Wahrscheinlichkeitsvektor beim Testen)
book, girl, plate, ...

Um die Performanz des Netzwerks zu testen, wurden die Gewichte des Netzes eingefroren und dem Netz jedes Eingabemuster aus dem Trainingskorpus nochmals präsentiert. Die Ausgabe des trainierten Netzes wurde nicht mit den (erwünschten) Ausgabevektoren aus dem Trainingskorpus (4a) verglichen, sondern mit den berechneten Wahrscheinlichkeitsvektoren (4b). Die Wahrscheinlichkeitsverteilung wurde nur für die Testphase verwendet. Das Training fand mit den tatsächlichen Folgewörtern statt, da das Netz ja aus den Sätzen die Wahrscheinlichkeitsverteilung lernen sollte.

Ergebnisse: Während ein Vergleich der Netzausgabe mit den tatsächlichen Folgewörtern (erwünschte Ausgabe beim Training) einen RMS-Fehler ('root mean square error')⁹ von 0,88 ergab, waren die Ergebnisse bei einem Vergleich mit der Wahrscheinlichkeitsverteilung wesentlich besser (RMS = 0,053). Das Netz scheint eine Annäherung an die Wahrscheinlichkeitsverteilung gelernt zu haben.

Eine hierarchische Clusteranalyse ergab Aufschluss über die interne Repräsentation des Netzwerks. Dabei wurden die Aktivierungen der verdeckten Zellen für jedes Wort bestimmt und diese nach Ähnlichkeiten gruppiert. Die Analyse ergab, dass das Netz die Eingabewörter in syntaktisch und semantisch relevante Kategorien eingeteilt hat: je eine Gruppe, die Nomen bzw. Verben

⁹ Berechnet wird der Abstand zweier Vektoren: 1) tatsächliche Netzausgabe vs. 2) erwünschte Ausgabe bzw. Wahrscheinlichkeitsvektor. Der RMS-Fehler (RMSE: 'root mean square error') ist der durchschnittliche Fehler für ein Muster und berechnet sich wie folgt:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{SSE}{n}} = \sqrt{\frac{\sum_n \sum_j (t_{nj} - o_{nj})^2}{n}}$$

wobei n die Anzahl der Testmuster ist, j die Anzahl der Ausgabezellen, MSE ('mean square error'), SSE ('sum square error'). t_{nj} ist die erwünschte Ausgabe ('teaching input') für die Ausgabezelle j in Testmuster n , o_{nj} ist die tatsächliche Ausgabe.

entsprechen sowie jeweils weitere Untergruppen. Verben wurden unterteilt in solche, die ein direktes Objekt erfordern, intransitive Verben und solche mit optionalem direktem Objekt. Nomen wurden weiter unterteilt in belebte und unbelebte Nomen, und diese wiederum in weitere Untergruppen.

6.3.2.2 Elman (1991)

In Elman (1990) waren die strukturellen Zusammenhänge zwischen den Wortklassen recht einfach. Bei der Satzantizipation stand die Kategorisierung der Eingabewörter nach syntaktischen und semantischen Gesichtspunkten im Vordergrund. In Elman (1991) liegt der Schwerpunkt auf dem Erlernen syntaktischer Strukturen. Es ging dabei ausschließlich um syntaktische Kategorien und strukturelle Zusammenhänge. Semantische Aspekte spielten keine Rolle.

Trainingsdaten und Lernaufgabe: Die verwendeten Trainingsdaten bestanden aus einfachen und komplexen Sätzen, welche durch eine vorher festgelegte Phrasen-Struktur-Grammatik (PSG) erzeugt wurden. Die Wortarten, die in der Grammatik verwendet wurden, waren Nomen, Verben und das Relativpronomen 'who'. Nomen und Verben konnten im Singular oder im Plural stehen. Außerdem gab es drei verschiedene Verbklassen: a) Verben, die ein direktes Objekt erfordern (z.B. *hit, feed*), b) Verben, die optional ein direktes Objekt erlauben (z.B. *see, hear*), und c) Verben, die ein direktes Objekt ausschließen (z.B. *walk, live*).

Alle Sätze enthielten ein Subjekt, ein Verb und optional ein direktes Objekt (siehe die Beispiele in (5)), wobei Subjekt und Verb bezüglich des Merkmals "Numerus" übereinstimmen mussten (Subjekt-Verb-Kongruenz). Subjekt und Objekt konnten außerdem um einen Relativsatz erweitert werden (5c-g). Dabei war auch Rekursion erlaubt (5f). Das Vokabular bestand insgesamt aus 23 Wörtern (Nomen, Verben, und das Relativpronomen 'who').

- (5) a. Girls see.
b. Girls see boys.
c. Dog [_{RC} who chases cat] sees girl.
d. Dog [_{RC} who cat chases] sees girl.
e. Dog [_{RC} who boys feed] sees girl.
f. Boys [_{RC} who girls [_{RC} who dogs chase] see] walk.
g. Mary sees boys [_{RC} who chase dogs].

Mit Hilfe der Phrasen-Struktur-Grammatik wurden insgesamt vier Trainingskorpora zu je 10.000 Sätzen generiert. Die vier Korpora unterschieden sich in ihrer Komplexität.

- KORPUS 1: nur einfache Sätze (ohne Relativsätze), durchschnittliche Satzlänge: 3,5 Wörter
- KORPUS 2: 75% einfache Sätze, 25% komplexe Sätze (mit Relativsätzen), durchschnittliche Satzlänge: 3,9 Wörter
- KORPUS 3: je 50% einfache und komplexe Sätze, durchschnittliche Satzlänge: 4,4 Wörter
- KORPUS 4: je 25% einfache und 75% komplexe Sätze, durchschnittliche Satzlänge: 6 Wörter

Das Netz wurde zuerst mit einfachen Sätzen trainiert (Korpus 1), danach wurde schrittweise die Komplexität erhöht (Korpus 2-4). Jedes Korpus wurde dem Netz fünfmal präsentiert. In jedem Trainingskorpus waren alle 10.000 Sätze ohne Unterbrechung aneinandergereiht und bildeten eine Eingabesequenz. Wie in Elman (1990), bestand die Aufgabe darin, für jedes Wort aus der Eingabesequenz das nachfolgende Wort zu antizipieren. Elman konnte in früheren Pilotstudien sowie in Elman (1993) zeigen, dass mit Hilfe des inkrementellen Lernens das Netz komplexe Daten besser lernt, wenn zuerst die einfachen Daten präsentiert werden und erst danach die komplexeren Muster.

Um die Satzantizipations-Aufgabe erfolgreich lösen zu können, musste das Netz die folgenden Teilaufgaben lernen:

- Kategorisierung der Wörter in verschiedene syntaktische Wortklassen: Nomen, Verben, Relativpronomen.
- Einteilung der Nomen und Verben in verschiedene syntaktisch relevante Unterkategorien: Welche Wörter stehen im Singular bzw. Plural? Welche Verben sind transitiv? Usw.
- Strukturelle Zusammenhänge zwischen Wortklassen:
 - Welches Nomen ist Subjekt bzw. Objekt des Satzes (Verb-Argument-Struktur)?
 - Welches Nomen muss mit welchem Verb bezüglich des Merkmals “Numerus” übereinstimmen (Subjekt-Verb-Kongruenz)?
 - Interaktion mit Relativsätzen: Das Relativpronomen ‘*who*’ kann sich auf das Subjekt (wie in (5c)) oder das Objekt (wie in (5d)) des eingebetteten Satzes beziehen. Der Erwerb der korrekten Verb-Argument-Struktur bzw. der Subjekt-Verb-Kongruenz wird dadurch komplexer.
 - Mehrfach-Einbettung: Durch die erlaubte Rekursion bei der Einbettung von Relativsätzen (wie in (5f)) wird der Lernprozess zusätzlich erschwert.

Simulationen: Für die Simulationen wurde ein Elman-Netzwerk verwendet, dessen Ein- und Ausgabeschicht jeweils aus 26 Zellen bestanden. Wie auch in Elman (1990) repräsentierte jede dieser Zellen ein Wort. Die verdeckte Schicht und die Kontextschicht bestanden jeweils aus 70 Zellen. Das Netzwerk wurde mit dem Backpropagation-Lernalgorithmus für Elman-Netze trainiert. Das Training umfasste fünf Durchgänge (Zyklen) durch jedes der vier Trainingskorpora.

Ergebnisse: Getestet wurde das Netz mit neuen Sätzen, die, wie die Trainingssätze, mit Hilfe der Phrasen-Struktur-Grammatik (PSG) erzeugt wurden. Die Ausgabe des Netzes für die neuen Testsätze wurde mit der bedingten Wahrscheinlichkeitsverteilung für die grammatisch möglichen Nachfolger verglichen (siehe Elman, 1990). Die Abweichung zwischen Netzausgabe und Wahrscheinlichkeitsvektoren war relativ gering (Fehler: 0,177).¹⁰ Der Fehler vor dem Lernvorgang betrug 12,45, der minimale Fehler bei einer Strategie, alle Zellen gleich stark zu aktivieren, lag bei 1,92. Insgesamt zeigen diese Werte, dass das Netz eine hohe Performanz bei dieser Aufgabe erreicht hat.

Eine inhaltliche Untersuchung des Antizipationsverhaltens bei verschiedenen Satztypen ergab, dass das Netz Verb-Argument-Struktur und Subjekt-Verb-Kongruenz in einfachen und komplexen Sätzen gelernt hat. Die Autoren schließen, dass aufgrund der komplexen Interaktion von Argument-Struktur und Kongruenz mit eingebetteten Relativsätzen das Netz abstrakte Repräsentationen für Konstituentenstrukturen, Argumentstrukturen, grammatische Kategorien, grammatische Relationen und das Merkmal "Numerus" entwickelt haben muss.

6.3.3 Vergleich mit probabilistischen Verfahren

(Towsey et al., 1998)

Das Ziel dieser Studie war es, die Performanz von Elman-Netzen bei der Satzantizipations-Aufgabe mit probabilistischen Ansätzen, insbesondere den sogenannten n -Gramm-Modellen¹¹, zu vergleichen.

¹⁰ Das Fehlermaß zur Berechnung des Abstands zweier Vektoren wird jedoch nicht explizit genannt.

¹¹ Ein n -Gramm ist eine Kette von n Elementen innerhalb einer gegebenen Sequenz. In der Towsey-Studie handelt es sich um Wortsequenzen. Ein 3-Gramm (oder Trigramm) ist beispielsweise eine Kette aus 3 Elementen bzw. Wörtern in einem Text. Ein n -Gramm-Modell macht Vorhersagen zur Vorkommenswahrscheinlichkeit des n -ten Elements in einer Sequenz basierend auf den $n - 1$ Vorgänger-Elementen. Ein Trigramm-Modell sagt z.B. vorher, mit welcher Wahrscheinlichkeit das dritte Element in einem Trigramm vorkommt, basierend auf den beiden Vorgänger-Elementen. Die Übergangswahrscheinlichkeiten werden im Allgemeinen aus den relativen Häufigkeiten in einem Trainingskorpora berechnet (siehe z.B. Brown et al., 1992).

6.3.3.1 Trainingsdaten und Lernaufgabe

Im Gegensatz zu den künstlich generierten Trainingsdaten in den Elman-Simulationen verwendeten Towsey und Kollegen echte Texte (englische Schulbuchtexte der 1. Klasse). Das bedeutet auf der einen Seite, dass in den Sätzen wesentlich mehr Wortarten vorkamen (z.B. Artikel, Präpositionen, etc.) und somit eine größere Vielfalt an Phrasen. Andererseits waren die Sätze jedoch stark vereinfacht, d.h. Sätze mit eingebetteten Strukturen (Relativsätze) wurden ausgefiltert. Lange Sätze (> 8 Wörter) wurden ebenfalls ausgefiltert, wodurch jedoch eine vergleichbare Länge zu den Trainingssätzen in Elman (1991) entstand. Das resultierende Trainingskorpus bestand aus 106 Sätzen mit einer Länge von 3 bis 8 Wörtern (durchschnittlich 5,06 Wörter). In Towsey et al. (1998) werden keine Beispiele für Trainingssätze gegeben. Aufgrund der angegebenen Wortarten und Satzlängen habe ich in (6) Beispiele konstruiert.

- (6) a. The boy runs fast .
b. The boy runs in the garden .

Ein wesentlicher Unterschied zu den Elman-Simulationen besteht darin, dass die Satzantizipations-Aufgabe nicht mit Wörtern, sondern mit lexikalischen Kategorien durchgeführt wurde. Die Aufgabe bestand also darin, nicht das nachfolgende Wort, sondern die nachfolgende Kategorie zu antizipieren. Dieser Wechsel von der Wortebene zur Kategorienebene ist notwendig, sobald man mit echten Texten arbeitet. Echte Texte enthalten zu viele unterschiedliche Wörter, um allgemeine syntaktische Regeln für einzelne Wörter extrahieren zu können. Es existieren sozusagen zu wenig Datenpunkte für ein Wort ('sparse data problem'). Die Satzantizipation auf Wortebene funktioniert nur bei einem sehr eingeschränkten Vokabular, wie z.B. in den Elman-Simulationen mit 29 bzw. 23 Wörtern.¹²

¹² Genau genommen ist die Antizipation von lexikalischen Kategorien viel plausibler als die Antizipation von konkreten Wörtern. In der Regel erwartet man an einer Stelle im Satz eher eine bestimmte Wortart, dass z.B. auf einen Artikel ein Nomen folgt. Es wird nicht etwa ein spezielles Nomen erwartet (mit Ausnahme von sehr eingeschränkten Kontexten).

Verwendet man die Satzantizipations-Aufgabe auf Kategorienebene, müssen die Wörter des Trainingskorpus vor dem Training in lexikalische Kategorien konvertiert werden. In der Towsey-Studie wurden hierfür 10 verschiedene Kategorien verwendet: Artikel (AR), Konjunktion (CC), Präposition (IN), Adjektiv (JJ), Nomen (NN), Pronomen (PR), Possessiv (PS), Adverb (RB), Verb (VB) und Satzgrenze (/S).¹³ In (7) ist die resultierende Kategorien-Sequenz für den Satz in (6a) dargestellt.

(7) The boy runs fast .
AR NN VB RB /S

Die Konvertierung der 106 Sätze im Trainingskorpus ergab eine Eingabesequenz von 643 Kategorien (inkl. Satzgrenzen). Die Aufgabe bestand darin, für jede Kategorie der Eingabesequenz die nachfolgende Kategorie zu antizipieren. Da jede Kategorie der Eingabesequenz einmal antizipiert werden musste, ergaben sich daraus 643 Trainingsmuster. Das erste Trainingsmuster begann mit der Kategorie /S (Satzgrenze) als Eingabe und hatte als Ausgabe die erste Kategorie des ersten Satzes. In (8) ist illustriert, wie die ersten beiden Trainingsmuster für den Satz in (7) aussehen würden.

(8) Eingabe: /S Ausgabe: AR (1. Muster)
Eingabe: AR Ausgabe: NN (2. Muster)

Um die Satzantizipations-Aufgabe erfolgreich lösen zu können, musste das Netz die folgenden Teilaufgaben lernen:

- Erwerb von komplexen Phrasen: Welche Sequenzen bilden eine Phrase?
- Kategorisierung der Phrasen (Nominalphrasen, Präpositionalphrasen, etc.)
- Strukturelle Zusammenhänge zwischen Phrasen: einfache Argumentstruktur

¹³ Es wurden wesentlich mehr Wortarten verwendet als in den Elman-Simulationen. Der Nachteil diese Kategorisierung ist jedoch, dass morphologische Information verloren geht. Das Merkmal "Numerus" stand hier nicht zur Verfügung, so dass z.B. die Subjekt-Verb-Kongruenz in der Towsey-Studie keine Rolle spielte.

Da in dieser Studie die Eingabe aus Kategorien-Sequenzen bestand, musste das Netz keine Kategorisierung der Wörter in syntaktische Wortklassen lernen. Im Gegensatz zu den Elman-Simulationen fand jedoch eine Kategorisierung auf einer abstrakteren Ebene statt, nämlich die Kategorisierung von komplexen Phrasen, soweit sie für die Antizipationsaufgabe relevant waren. In den Elman-Simulationen musste das Netz keine komplexen Phrasen erwerben, da die einzelnen (nicht-verbalen) Phrasen jeweils nur aus einem einzigen Wort bestanden. Es musste lediglich eine Repräsentation für einfache Verbalphrasen und Sätze aufgebaut werden. Die Komplexität in Elman (1991) konzentrierte sich auf die rekursive Einbettung von Relativsätzen verbunden mit Argumentstruktur und Subjekt-Verb-Kongruenz. Eingebettete Strukturen und Subjekt-Verb-Kongruenz wurden wiederum in der Towsey-Studie ausgeklammert. Man kann also sagen, dass in den verschiedenen Simulationen unterschiedliche Aspekte der Grammatik im Vordergrund standen.

6.3.3.2 Simulationen

Für die Simulationen wurde ein Elman-Netzwerk verwendet, dessen Ein- und Ausgabeschicht jeweils aus 10 Zellen bestanden. Jede dieser Zellen repräsentierte eine Kategorie. In jedem Trainingsmuster war daher genau eine Zelle der Eingabe- und eine Zelle der Ausgabeschicht aktiviert, welche der jeweiligen Eingabe- bzw. Ausgabekategorie entsprachen. Die verdeckte Schicht und die Kontextschicht variierten zwischen 1 und 50 Zellen. Das Netzwerk wurde mit einer Modifikation des Backpropagation-Lernalgorithmus ("Backpropagation mit Momentum-Term") für Elman-Netze trainiert. Das Training umfasste 400 Durchgänge (Zyklen) durch das Trainingskorpus bei schneller Lerngeschwindigkeit (200 Zyklen mit Lernrate $\eta = 0,01$ und 200 Zyklen mit $\eta = 0,001$) bzw. alternativ 100.000 Zyklen bei langsamer Lerngeschwindigkeit (Lernrate $\eta = 0,0001$).

6.3.3.3 Ergebnisse

Die besten Ergebnisse erzielte ein Elman-Netz mit 9 verdeckten Zellen. Es lieferte 72% korrekte Vorhersagen auf dem kompletten Trainingsdatensatz. Dabei wurde die Netzausgabe mit der Kategorie des Folgeworts im Satz verglichen (erwünschte

Ausgabe beim Training).¹⁴ Mit Hilfe einer 106-fachen Kreuzauswertung wurde die Generalisierungsfähigkeit des Netzes auf unbekanntem Daten getestet. Hierfür wurde der Trainingsdatensatz (106 Sätze) verwendet, ein Satz ausgelassen und das Netz mit den restlichen 105 Sätzen trainiert. Der ausgelassene Satz war dann für das Netz unbekannt und wurde für den Test genutzt. Diese Prozedur wurde 106 Mal wiederholt bis jeder Satz einmal ausgelassen wurde und als Testsatz zur Verfügung stand. Ein Elman-Netz mit 9 verdeckten Zellen, das auf diese Weise trainiert wurde, lieferte 64% korrekte Vorhersagen auf den unbekanntem Daten.

Towsey und Kollegen verglichen die Performanz des Elman-Netzes mit der von n -Gramm-Modellen. Hierfür wurde ermittelt, welche Vorhersagen Bigramm- ($n = 2$), Trigramm- ($n = 3$), 4-Gramm und 5-Gramm-Modelle für den Trainingsdatensatz machen würden. Ein 4-Gramm-Modell liefert, wie auch das Elman-Netz, 72% Prozent korrekte Vorhersagen auf den verwendeten Trainingsdaten. Das Elman-Netz hat mit *einem* Kontextelement (eine Kategorie) als Eingabe eine Performanz wie ein 4-Gramm-Modell, das die Ergebnisse auf Basis von drei Kontextelementen berechnet. Das Netz hat also einen größeren Kontext gelernt, den es als Basis für die Vorhersage nutzen kann, als durch die Eingabe vorgegeben wurde.

Die Autoren finden außerdem Hinweise, dass das Elman-Netz n -Gramm-Wahrscheinlichkeiten lernt, wobei sich n während des Lernvorgangs erhöht. Aufschluss ergibt eine Untersuchung der Performanz des Elman-Netzes in verschiedenen Stadien des Lernprozesses. Dabei wird die Netzausgabe direkt (Muster für Muster) mit der Wahrscheinlichkeitsverteilung verglichen, die die n -Gramm-Modelle liefern. Zu Beginn des Lernvorgangs ist der RMS-Fehler des Netzes am kleinsten, wenn die Netzausgabe mit Bigramm-Wahrscheinlichkeiten verglichen wird. Nach 8.000 Durchgängen ist die Abweichung zu Trigrammen am geringsten, und nach Trainingsende (ab ca. 80.000 Zyklen) hat das Netz die geringste Abweichung zu 4-Gramm-Wahrscheinlichkeiten (RMS-Fehler = 0,068).¹⁵ Der RMS-Fehler ist am größten, wenn die Netzausgabe

¹⁴ An dieser Stelle sei nochmals angemerkt, dass die Satzantizipations-Aufgabe nicht-deterministisch ist und ein Vergleich mit der Nachfolger-Kategorie im Satz problematisch ist.

¹⁵ Eine direkte Vergleichbarkeit der Ergebnisse mit den Elman-Simulationen ist nicht möglich, da in Elman (1990, 1991) die Netzausgabe mit absoluten bedingten Wahrscheinlichkeiten

(anstelle von Wahrscheinlichkeiten) mit der Folgekategorie im Satz verglichen wird.

Es deutet daher alles darauf hin, dass das Elman-Netz eine Wahrscheinlichkeitsverteilung lernt. Um diese Vermutung zu untermauern, untersuchten die Autoren die Aktivierungen der Ausgabezellen. Sie konnten zeigen, dass die Summe der Ausgabeaktivierungen während des Lernens gegen 1,0 tendierte, was für eine Wahrscheinlichkeitsverteilung charakteristisch ist.

Diese Studie hat gezeigt, dass die Satzantizipations-Aufgabe auf Kategorienebene recht gut funktioniert. Die Ergebnisse legen nahe, dass das Elman-Netz eine Wahrscheinlichkeitsverteilung lernt und es gibt Evidenzen, dass das Netz einen größeren Kontext als Basis für die Vorhersagen nutzen kann als durch die Eingabe vorgegeben wird.

6.3.4 Vergleich mit menschlicher Performanz

(Christiansen und Chater, 1999, 2001)

Die Autoren möchten mit dieser Studie demonstrieren, dass ein neuronales Netz in der Lage ist, rekursive sprachliche Strukturen zu lernen, zumindest bis zu einem Grad, wie er auch der menschlichen Performanz entspricht. In (9) sind Beispiele für rekursiv eingebettete Relativsätze (Rekursionstyp: Zentrale Einbettung) mit Rekursionstiefe 0–3 dargestellt.

- (9) a. The mouse ran away. (ohne Einbettung)
- b. The mouse that the cat bit ran away. (1-fache Einbettung)
- c. The mouse that the cat that the dog chased bit ran away. (2-fache Einbettung)
- d. The mouse that the cat that the dog that the man frightened chased bit ran away. (3-fache Einbettung)

verglichen wurde. Das würde in der Towsey-Studie einem Vergleich mit einem 8-Gramm-Modell entsprechen, da hier 8 Wörter die maximale Satzlänge ist.

Es wird im Allgemeinen angenommen, dass die Grammatik des Menschen eine beliebige Einbettungstiefe rekursiver Strukturen erlaubt. Psycholinguistische Studien haben jedoch gezeigt, dass Menschen Schwierigkeiten haben, Sätze zu verstehen, die mehr als eine Einbettung enthalten (Marks, 1968; Bach et al., 1986). Dies wird auf außergrammatische Faktoren, wie z.B. die Beschränkung des Arbeitsgedächtnisses, zurückgeführt.

Das Ziel von Christiansen und Chater war es, die menschliche Performanz bei der Verarbeitung rekursiver Strukturen modellieren, wobei sie auf konkrete psycholinguistische Studien eingehen. Interessant ist hierbei vor allem auch der Vergleich zwischen der Performanz von neuronalen Netzen und menschlichem Verhalten (Lesezeiten, Beurteilung von Grammatikalität).

6.3.4.1 Trainingsdaten und Lernaufgabe

Als Basis für ihre Trainingsdaten verwendeten Christiansen und Chater drei künstliche Sprachen, die in Chomsky (1957) vorgestellt wurden. Die drei Sprachen sind komplexer als reguläre Sprachen und können daher nicht mit Hilfe von endlichen Automaten ('finite-state'-Mechanismen) verarbeitet werden. Sie enthalten jedoch, laut Chomsky, rekursive Muster, die in natürlicher Sprache vorkommen (siehe unten: Zählrekursion, Zentrale Einbettung, Kreuz-Abhängigkeiten).

Die Trainingsdaten enthielten diese drei nicht-regulären Rekursionstypen sowie einen weiteren regulären Rekursionstyp (Rechtsverzweigung), der mit endlichen Automaten verarbeitet werden kann und daher weniger komplex ist. Das verwendete Vokabular bestand aus 16 Wörtern: vier Nomen im Singular ($n_1 - n_4$), vier Nomen im Plural ($N_1 - N_4$), vier Verben im Singular ($v_1 - v_4$) und vier Verben im Plural ($V_1 - V_4$). Ein zusätzliches Zeichen markierte das Satzende (#). Wie man aus dem Vokabular ersehen kann, wurden nur zwei Wortarten verwendet (Nomen und Verben), die jeweils im Singular oder im Plural stehen konnten. Die Trainingsdaten enthielten also keine natürlichen Sätze, die Konzentration bei diesen Simulationen liegt vielmehr auf den rekursiven Grundmustern, die in den Daten enthalten waren.

Im Folgenden werden die verwendeten Rekursionstypen beschrieben:

- ZÄHLREKURSION ('counting recursion'): Diesen Rekursionstyp findet man in Ausdrücken der Form $a^n b^n$, wie z.B. *ab*, *aabb*, *aaabbb*, etc. Bei diesen Ausdrücken wird ein Zähler benötigt, der die Anzahl der 'a's zählt und anschließend die gleiche Anzahl an 'b's produziert. Dieser Rekursionstyp ist inspiriert durch "if S1, then S2"-Konstruktionen, die man ineinander schachteln kann ("if if S1, then S2, then S3"). Auch hier muss die Anzahl der 'if's und 'then's übereinstimmen. Das folgende Beispiel stammt aus Christiansen und Chater (2001).

(10) **If if** the cat is in, **then** the dog cannot come in, **then** the cat and dog dislike each other.

In den Trainingsdaten wurden Ausdrücke dieser Art durch Nomen-Verb-Sequenzen dargestellt (z.B. *NNVV*), wobei eine Numerus-Übereinstimmung ignoriert wurde. Entscheidend war, dass genauso viele Verben wie Nomen generiert wurden.

- ZENTRALE EINBETTUNG ('center-embedding recursion'): Es handelt sich um Ausdrücke wie z.B. *abba*, *babbbab*, wobei das erste Element mit dem letzten übereinstimmen muss, das zweite Element mit dem vorletzten usw. Die zweite Hälfte des Ausdrucks ist sozusagen das Spiegelbild der ersten Hälfte. Eine Entsprechung in der natürlichen Sprache sind beispielsweise eingeschachtelte Relativsätze (*the boy_a girls_b like_b runs_a*). Hier muss das erste Nomen 'boy' mit dem letzten Verb 'runs' hinsichtlich des Merkmals Numerus (Singular, Plural) übereinstimmen und das zweite Nomen 'girls' mit dem vorletzten Verb 'like'. In den Trainingsdaten wurde der Satz "the boy girls like runs" z.B. durch die Sequenz "nNVv" dargestellt.

- **KREUZ-ABHÄNGIGKEITEN** ('cross-dependency recursion'): Ausdrücke dieser Art enthalten einen String X gefolgt von einer identischen Kopie von X wie z.B. in *abab*, *bbbb*, *abaab*, etc. Im Schweizerdeutschen und im Holländischen findet man solche Konstruktionen, bei denen sich die Abhängigkeiten der Nomen (Objekt-NPs) und Verben überkreuzen. Das erste Nomen stellt das Objekt des ersten Verbs dar, das zweite Nomen das Objekt des zweiten Verbs, usw. In (11) ist ein Beispiel aus dem Holländischen dargestellt mit einer wörtlichen Übersetzung ins Englische (aus Bach et al., 1986). In den Trainingsdaten wurde eine Konstruktion wie in (11) z.B. durch die Sequenz "*nNvV*" dargestellt.

(11) Jantje heeft de lerares₁ de knickers₂ helpen₁ opruimen₂.
 Jantje has the teacher₁ the marbles₂ help₁ collect up₂.
 "Jantje helped the teacher collect up the marbles."

- **RECHTSVERZWEIGUNG** ('right-branching recursion'): Es handelt sich um Ausdrücke, bei denen immer zwei aufeinanderfolgende Elemente übereinstimmen müssen, wie z.B. *aabb*, *aabbbb*, etc. Diesen Rekursionstyp findet man in Konstruktionen, bei denen Relativsätze rechts an den Satz angebunden werden, wie z.B. in "*girls₁ like₁ the boy₂ that runs₂*". Dabei muss das erste Nomen mit dem ersten Verb hinsichtlich des Merkmals Numerus übereinstimmen, das zweite Nomen mit dem zweiten Verb, usw. In den Trainingsdaten wird dieser Satz z.B. durch die Sequenz "*NVnv*" dargestellt.

Insgesamt wurden drei Trainingskorpora mit jeweils 5000 Sätzen generiert, die zu 50% aus komplexen rekursiven Strukturen und zu 50% aus rechtsverzweigenden Strukturen bestanden. Das erste Trainingskorpus (Sprache 1: Zählrekursion) bestand zu 50% aus Konstruktionen mit Zählrekursion und zu 50% aus rechts-verzweigenden Strukturen. Das zweite Trainingskorpus (Sprache 2: Zentrale Einbettung) bestand zu 50% aus Konstruktionen mit zentraler Einbettung und zu 50% aus rechts-verzweigenden Strukturen. Das dritte Trainingskorpus (Sprache 3: Kreuz-Abhängigkeiten) bestand zu 50% aus

Konstruktionen mit Kreuz-Abhängigkeiten und zu 50% aus rechts-verzweigenden Strukturen. Zu jedem Trainingskorpus wurde außerdem ein Testkorpus mit 500 neuen Sätzen generiert, welches ebenfalls zu 50% aus komplexen rekursiven Strukturen des jeweiligen Rekursionstyps und zu 50% aus rechts-verzweigenden Strukturen bestand.

Die durchschnittliche Satzlänge in den Trainings- und Testkorpora betrug ca. 4,7 Wörter. Die Einbettungstiefe der rekursiven Strukturen variierte von 0 bis 3. In jedem Korpus waren alle Sätze ohne Unterbrechung aneinandergereiht und bildeten eine Eingabesequenz. Dies ergab eine Sequenz von 23.500 Wörtern bzw. Trainingsmustern für jedes Trainingskorpus und eine Sequenz von 2350 Wörtern bzw. Trainingsmustern für jedes Testkorpus. In (12) sind zwei Beispielsätze dargestellt mit einer Übersetzung in das Vokabular der Trainingsdaten.

(12) The girl runs . The boy girls like runs .
 $n_1 \ v_3 \ \# \quad n_3 \ N_1 \ V_4 \ v_3 \ \#$

Christiansen und Chater verwendeten die Satzantizipations-Aufgabe auf Wortebene, wobei in den Trainingsdaten lediglich 16 Wörter vorkamen. Die Aufgabe bestand also darin, für ein gegebenes Wort das nachfolgende Wort zu antizipieren. In (13) ist illustriert wie die ersten beiden Trainingsmuster für die Sequenz in (12) aussehen.

(13) Eingabe: n_1 Ausgabe: v_3 (1. Muster)
 Eingabe: v_3 Ausgabe: $\#$ (2. Muster)

Um die Satzantizipations-Aufgabe erfolgreich lösen zu können, musste das Netz die folgenden Teilaufgaben lernen:

- Kategorisierung der Wörter in Nomen und Verben
- Einteilung der Nomen und Verben bezüglich Numerus (Singular, Plural)
- Strukturelle Abhängigkeiten zwischen diesen Kategorien für den jeweiligen Rekursionstyp (siehe oben)

6.3.4.2 Simulationen

Für die Simulationen wurden Elman-Netzwerke verwendet, deren Ein- und Ausgabeschicht jeweils aus 17 Zellen bestanden. Jede dieser Zellen repräsentierte ein Wort oder das Satzende. In jedem Trainingsmuster war daher genau eine Zelle der Eingabe- und eine Zelle der Ausgabeschicht aktiviert, welche dem jeweiligen Eingabe- bzw. Ausgabewort entsprach. Die verdeckte Schicht und die Kontextschicht variieren zwischen 2 und 100 Zellen. Das Netzwerk wurde mit dem Backpropagation-Lernalgorithmus für Elman-Netze (ohne Momentum) trainiert. Das Training umfasste 100 Durchgänge (Zyklen) durch das Trainingskorpus mit einer Lernrate von 0,1.

6.3.4.3 Evaluation

Die Elman-Netze wurden jeweils mit einer der drei Sprachen trainiert und mit neuen rekursiven Strukturen unterschiedlicher Länge, d.h. mit den Testkorpora, in verschiedenen Lernstadien getestet. Dabei wurde die tatsächliche Ausgabe des Netzes mit der absoluten bedingten Wahrscheinlichkeitsverteilung im Trainingskorpus verglichen. Wie die Autoren bemerken, ist dieses Vorgehen auf Basis von Wortsequenzen problematisch, wenn das Netz mit neuen Sequenzen getestet wurde, die im Trainingskorpus nicht vorkamen. Christiansen und Chater verwendeten daher bei der Berechnung der Wahrscheinlichkeitsverteilung einen Kontext auf Basis lexikalischer Kategorien (anstelle von Wörtern). Die Wörter $n_1 - n_4$ wurden hierfür auf eine Kategorie n abgebildet, entsprechend für die anderen Wörter. Die Wahrscheinlichkeit eines Folgeworts hing demnach nicht mehr von den vorausgehenden Wörtern (z.B. n_1, v_3, V_2, \dots) ab, sondern von der Abfolge der vorausgehenden Kategorien (n, N, v, V).

6.3.4.4 Ergebnisse

Die Ergebnisse zeigen, dass sich alle Netze ab einer bestimmten Mindestanzahl an verdeckten Neuronen (ab 5 – 10 Zellen) während der ersten 40 Lernzyklen auf ein ähnliches, relativ kleines, Fehlerniveau einpendelten. Die Netze, die mit dem passenden Testkorpus des gleichen Rekursionstyps getestet wurden, wiesen nach 100 Lernzyklen einen MSE-Wert (*'mean square error'*)¹⁶ kleiner 0,05 auf. Wurden dieselben Netze ausschließlich mit rechts-verzweigende Strukturen getestet, lag der MSE-Wert unter 0,08. Ein Vergleich mit der Performanz von Bigram- und Trigram-Modellen auf denselben Testdaten zeigte, dass die Elman-Netze ab einer Mindestanzahl von 5 – 10 verdeckten Zellen deutlich weniger Fehler machten als die Bigram- und Trigram-Modelle.

Die Autoren haben gezeigt, dass Elman-Netze in der Lage sind, rekursive Strukturen bis zu einer Einbettungstiefe von 3 zu lernen. Betrachtete man die Performanz der Netze in Abhängigkeit von der Rekursionstiefe, so ließ sich feststellen, dass die Performanz mit steigender Rekursionstiefe abnahm. Die Netze waren außerdem in der Lage, auf Rekursionstiefe 4 zu generalisieren, was dann zu einer etwas höheren Fehlerrate führte. Im folgenden Abschnitt wird beschrieben, wie die Ergebnisse mit menschlichem Verhalten verglichen wurden.

6.3.4.5 Vergleich mit menschlicher Performanz

Die Autoren verglichen die Performanz der trainierten Netze mit psycholinguistischen Studien zur Beurteilung von rekursiven Strukturen. Bach et al. (1986) zeigten in Beurteilungsaufgaben zur Verständlichkeit, dass bei Einbettungstiefe 2 Kreuz-Abhängigkeiten im Niederländischen verständlicher sind als Sätze mit zentraler Einbettung im Deutschen. Bei Einbettungstiefe 1

¹⁶ Berechnet wird der Abstand zweier Vektoren: 1) tatsächliche Netzausgabe vs. 2) erwünschte Ausgabe bzw. Wahrscheinlichkeitsvektor. Der MSE (mittlerer quadratischer Fehler) ist der durchschnittliche Fehler für ein Muster und berechnet sich wie folgt:

$$MSE = \frac{SSE}{n} = \frac{\sum_n \sum_j (t_{nj} - o_{nj})^2}{n}$$

wobei n die Anzahl der Testmuster, j die Anzahl der Ausgabezellen, SSE (*'sum square error'*). t_{nj} ist die erwünschte Ausgabe (*'teaching input'*) für die Ausgabezelle j in Testmuster n , o_{nj} ist die tatsächliche Ausgabe.

wurden beide Rekursionstypen ähnlich gut verstanden. Dieses Ergebnis ist interessant, da Kreuz-Abhängigkeiten in der formalen Sprachentheorie als komplexer eingestuft werden im Vergleich zu Konstruktionen mit zentraler Einbettung.¹⁷

Verglich man die Elman-Netze, die mit den Sprachen *Kreuz-Abhängigkeiten* (Sprache 3) bzw. *Zentrale Einbettung* (Sprache 2) trainiert wurden, erhielt man das gleiche Muster. Bei Rekursionstiefe 2 hatte das Netz, das mit *Zentraler Einbettung* trainiert wurde, signifikant mehr Schwierigkeiten, die entsprechenden Wortsequenzen zu antizipieren als das Netz, das mit *Kreuz-Abhängigkeiten* trainiert wurde. Bei Rekursionstiefe 1 gab es keinen signifikanten Unterschied. Dabei wurde jeweils der durchschnittliche Fehler für die Antizipation des ganzen Satzes berechnet. Interessanterweise zeigten die Ergebnisse der Bi- und Trigramme genau den umgekehrten Effekt. Bi- und Trigramme hatten größere Schwierigkeiten mit *Kreuz-Abhängigkeiten* als mit *Zentraler Einbettung* und stimmen, im Gegensatz zu den Elman-Netzen, nicht mit menschlichem Verhalten überein.

In einer weiteren Studie sollte die Grammatikalität von Sätzen beurteilt werden, die zweifach eingebettete Relativsätze (zentrale Einbettung) wie in (14) enthielten (Christiansen und MacDonald, 1999).

- (14) a. The chef who the waiter who the busboy offended appreciated admired the musicians.
- b. *The chef who the waiter who the busboy offended frequently admired the musicians.

Bei (14a) handelt es sich um eine grammatisch korrekte Zweifach-Einbettung mit drei Verbalphrasen (VPs), (14b) ist dagegen eine ungrammatische Zweifach-Einbettung mit nur zwei VPs. Interessanterweise wurde der grammatisch korrekte Satz mit drei Verbalphrasen in (14a) signifikant schlechter beurteilt als die ungrammatische Konstruktion mit zwei VPs in (14b).

¹⁷ *Kreuz-Abhängigkeiten* können (im Gegensatz zu “Zentraler Einbettung”) nicht mit kontextfreien Phrasenstrukturregeln abgeleitet werden. *Kreuz-Abhängigkeiten* erfordern einen komplexeren Grammatik-Typ innerhalb der Chomsky-Hierarchie (siehe Chomsky, 1956).

Das Elman-Netz, das mit *Zentraler Einbettung* trainiert wurde, zeigte ein vergleichbares Muster. Wurde dem Netz eine Sequenz mit drei Nomen und zwei Verben (*nnnvv*) als Input vorgegeben, so hatte das Netz eine Präferenz für das Satzende-Zeichen als Folgeelement (*nnnvv#*), was der ungrammatischen Konstruktion mit nur zwei Verben entspricht, verglichen mit einem dritten Verb als Folgeelement (*nnnvvv*), was der grammatischen Version entsprechen würde. Mittelte man den Fehler über die ersten sechs Wörter (*nnnvvv* vs. *nnnvv#*), war der Wert bei der grammatischen Konstruktion (*nnnvvv*) signifikant höher, d.h. das Netz hatte größere Verarbeitungsschwierigkeiten, als bei der ungrammatischen Konstruktion (*nnnvv#*).

Die Autoren zeigten für beide psycholinguistische Studien eine Korrelation zwischen der Performanz von Elman-Netzen und menschlicher Performanz in Beurteilungsaufgaben (Verständnis, Grammatikalität). In MacDonald und Christiansen (2002) wurde außerdem gezeigt, dass eine Korrelation zu Lesezeiten hergestellt werden kann. Der Fehler des Netzes für ein einzelnes Wort soll die Schwierigkeit des Netzes reflektieren, dieses Wort in einem bestimmten Kontext vorherzusagen.¹⁸ Die Autoren haben gezeigt, dass der Fehler für einzelne Wörter qualitativ auf Wort-Lesezeiten abgebildet werden kann, wobei ein niedriger Fehler-Wert schnelle Lesezeiten und ein hoher Fehler-Wert langsame Lesezeiten reflektiert (MacDonald und Christiansen, 2002). Der durchschnittliche Fehler für den ganzen Satz reflektiert die Schwierigkeit, den Satz als Ganzes zu verarbeiten. Dieser Wert kann auf Grammatikalitätsurteile für ganze Sätze abgebildet werden, wobei ein niedriger Durchschnittswert einer hohen Akzeptabilität des Satzes entspricht, ein hoher Wert reflektiert eine niedrige Akzeptabilität (Christiansen und Chater, 1999, 2001).

Diese Korrelation mit menschlicher Performanz ist ein interessanter Ausgangspunkt für die vorliegende Arbeit. Dabei stellt sich insbesondere die Frage, wie ein Elman-Netz mit koordinierten Strukturen umgeht bzw. ob sich eine Korrelation zu den entsprechenden Verarbeitungseffekten in Kapitel 3 herstellen lässt. Diese Fragestellungen werde ich im Folgenden untersuchen.

¹⁸ Bei der Berechnung des Fehlers wurde die Netzausgabe mit einer Wahrscheinlichkeitsverteilung verglichen. Um den Abstand der beiden Vektoren zu messen, wurde ein spezielles Fehlermaß, der *'Grammatical Prediction Error'* (GPE), verwendet.

6.4 Eigene Simulationen zur Verarbeitung koordinierter Strukturen

In diesem Kapitel werden die von mir durchgeführten Arbeiten vorgestellt. Dabei wird die Frage beleuchtet, ob Elman-Netze in der Lage sind, koordinierte Strukturen zu verarbeiten. Angesichts der Problematik, dass koordinierte Strukturen eine der größten Fehlerquellen für die maschinelle Sprachverarbeitung darstellen (siehe z.B. Argamon et al., 1998; Collins, 1999), stellt sich die Frage, ob neuronale Netze ähnliche Schwierigkeiten mit diesen Konstruktionen haben oder ob sie ohne Probleme verarbeitet werden, wie es im menschlichen Sprachverarbeitungssystem der Fall ist, bzw. ob die Verarbeitung bei parallelen Strukturen vielleicht sogar erleichtert ist (Parallel-Struktur-Effekt beim Lesen). Nachdem Christiansen und Chater (1999, 2001) eine Korrelation zwischen der Performanz von Elman-Netzen und menschlichem Verhalten für rekursive Strukturen gefunden haben, ist es denkbar, dass auch eine entsprechende Korrelation für koordinierte Strukturen existiert.

Desweiteren sollen die Netze mit echten spontansprachlichen Korpusdaten trainiert werden. In den bisherigen Simulationen bestanden die Trainingsdaten entweder aus stark vereinfachten Sätzen (Elman, 1990, 1991; Towsey et al., 1998) oder aus künstlichen Sprachen (Christiansen und Chater, 1999, 2001). Eine weitere Fragestellung dieser Arbeit betrifft daher die Machbarkeit der Satzantizipations-Aufgabe für echte Korpusdaten. Im Einzelnen werden die folgenden Fragestellungen beleuchtet.

1. Ist ein Elman-Netz in der Lage, mittels der Satzantizipations-Aufgabe grammatische Regularitäten selbständig aus echten spontansprachlichen Korpusdaten zu lernen?
2. Wie verhält sich das Netz bei koordinierten Strukturen? Gibt es eine Korrelation zwischen der Performanz des Netzwerks und menschlichem Verhalten bei der Verarbeitung von Koordination (siehe Kapitel 3)?

6.4.1 Trainingsdaten

6.4.1.1 Verwendetes Korpus

Als Basis für die Trainingsdaten dient die “*Englische Verbmobil Baumbank*” (TÜBA-E, Hinrichs et al., 2000), die bereits in Kapitel 5 beschrieben wurde. Die Baumbank enthält spontansprachliche englische Dialoge zwischen Geschäftspartnern zur Vereinbarung von Terminen oder Hotelreservierungen. Für das Training der Netze wird ein Korpusfragment (CD13) verwendet, das 2906 Äußerungen bzw. 37.563 Wörter (Tokens) umfasst. Dieses Korpusfragment ist bereits in Kapitel 5 im Hinblick auf koordinierte Strukturen statistisch ausgewertet worden.

Für die Simulationen werden *englische* Spontansprache-Daten verwendet, weil die Simulations-Ergebnisse mit menschlichen Daten verglichen werden sollen und die entsprechenden Verarbeitungseffekte bisher zum Englischen gefunden wurden (siehe Kapitel 3). Desweiteren sind gesprochene Daten für die vorliegende Arbeit als Trainingsdaten besser geeignet als geschriebene Texte, weil die gesprochene Sprache viel direkter Verarbeitungsprozesse widerspiegelt und daher auch eher mit menschlichen Verarbeitungseffekten beim Lesen korreliert (siehe Steiner, 2007a, 2009).¹⁹

Es ist zudem erforderlich, dass die Korpusdaten morpho-syntaktisch und syntaktisch aufbereitet sind, wie es in der *Englischen Verbmobil Baumbank* (TÜBA-E) der Fall ist. Die morpho-syntaktische Annotation der Baumbank (Wortarten) wird als Basis für die Satzantizipation benötigt, da diese Aufgabe auf Kategorienebene (Wortarten) und nicht auf Wortebene durchgeführt wird. Eine genaue Beschreibung der verwendeten Lernaufgabe findet der Leser in Abschnitt 6.4.2. Die syntaktische Annotation der TÜBA-E-Baumbank wird zwar nicht für die Satzantizipations-Aufgabe benötigt, ist aber für die automatische Vorverarbeitung der Trainingsdaten unerlässlich (siehe folgender Abschnitt).

¹⁹ In Kapitel 5 wurde eine Korrelation zwischen Präferenzen in Lesezeit-Experimenten und Vorkommenshäufigkeiten in den hier verwendeten Trainingsdaten (gesprochene Korpora) gefunden (siehe auch Steiner, 2007a). Außerdem wird in Kapitel 7 gezeigt, dass gesprochene Daten mit Online-Datentypen wie z.B. Lesezeiten korrelieren, während geschriebene Texte eher mit Offline-Datentypen korrelieren, wie z.B. den Grammatikalitätsurteilen (siehe auch Steiner, 2009).

6.4.1.2 Vorverarbeitung der Daten

Die *Englische Verbmobil Baumbank* enthält die typischen Charakteristika der Spontansprache, d.h. fragmentarische Äußerungen, Einschübe (Parenthesen), Wiederholungen, Korrekturen, etc. Dies macht es unumgänglich, die Korpusdaten entsprechend vorzuverarbeiten. Da das neuronale Netz, wie auch in den bisherigen Simulationen, grammatische Regeln auf Satzebene lernen soll, verwende ich für die Trainingsdaten nur die *Sätze* aus der Baumbank. Fragmentarische Äußerungen, Einschübe, etc. werden vorher ausgefiltert. Für die Trainingsdaten habe ich die folgenden Kriterien festgelegt:

- Die Trainingsdaten enthalten nur Sätze, die mit *S* (sentence) in der Baumbank annotiert sind²⁰, d.h. keine Fragmente, Wiederholungen oder Korrekturen.

In einer Konstruktion, wie in (15a, siehe Abb. 6.4), wird nur der ganze Satz (*'six thirty sounds good'*) für die Trainingsdaten verwendet, die fragmentarischen Äußerungen davor und danach werden ausgefiltert. Das Ergebnis der Bearbeitung ist in (15b) dargestellt. In (16a, siehe Abb. 6.5), wird der wiederholte Determinierer (*'the'*), der nicht am Satzknoden angebunden ist, ausgefiltert. (16b) illustriert das Ergebnis.

(15) a. yeah , six , [*S* six thirty sounds good] , sounds good (Original)

b. [*S* six thirty sounds good] (bearbeitet)

(16) a. [*S* I was on the , the wrong scene] (Original)

b. [*S* I was on the wrong scene] (bearbeitet)

²⁰ Eine Äußerung wird dann als Satz (*S*) in der TüBa-E annotiert, wenn sie mindestens ein Subjekt und ein finites Verb enthält (Kordoni, 2000). Oft werden auch Äußerungen als Sätze annotiert, die zwar ein Subjekt und ein Verb enthalten, aber dennoch unvollständig sind, weil ein Objekt oder ähnliches fehlt. Das ist z.B. der Fall, wenn ein Satz abgebrochen wurde. Diese Konstruktionen wurden dennoch berücksichtigt.

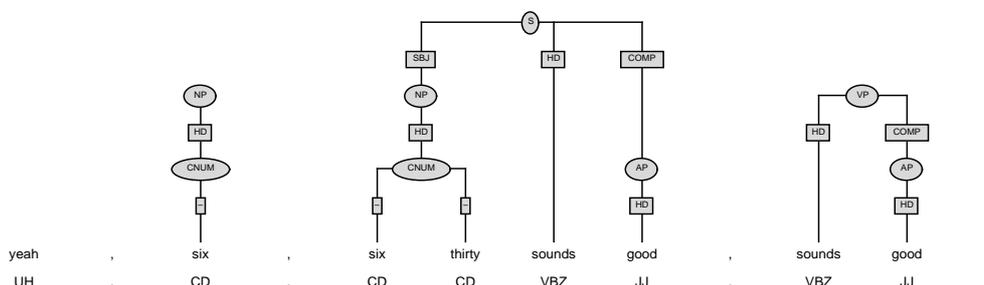


Abbildung 6.4: Satz mit fragmentarischen Äußerungen (Nr. 1591, CD 13)

- Jeder Satz wird separat präsentiert. Dies betrifft z.B. Äußerungen, die mehrere Sätze enthalten, sowie unangebundene Satz-Parenthesen. Die Satz-Parenthese ('*you know*') in (17a, siehe Abb. 6.6) wird beispielsweise als separater Satz betrachtet und nach dem Trägersatz ('*if I can run home and get a shower*') präsentiert. Das Ergebnis der Bearbeitung ist in (17b) dargestellt.

(17) a. [_S if I can [_S you know] , run home and get a shower] (Original)

b. Satz 1: [_S if I can run home and get a shower] (bearbeitet)

Satz 2: [_S you know]

- Die maximale Satzlänge wird auf 25 Wörter festgelegt.²¹

Die automatische Vorverarbeitung der Trainingsdaten umfasst das Ausfiltern von Nicht-Sätzen, die Linearisierung der Sätze und das Ausfiltern von Sätzen, die länger als 25 Wörter sind. Diese Prozedur kann nur deshalb automatisch durchgeführt werden, weil die *Englische Verbmobil Baumbank* syntaktisch annotiert ist. Ohne syntaktische Annotation ist es nicht möglich, automatisch eindeutig zu erkennen, wo ein Satz beginnt, wo ein Satz aufhört bzw. welche Wörter zu einem Satz gehören und welche nur einen Einschub darstellen. An dieser Stelle sei angemerkt, dass die Trainingssätze durch die Vorverarbeitung

²¹ Der längste Satz in CD13 besteht aus 69 Wörtern. Die Annotation dieser sehr langen Konstruktionen ist jedoch umstritten. Durch die Längenbegrenzung werden 1,4% der Sätze ausgefiltert. Die maximale Satzlänge in den bisher beschriebenen Simulationen liegt zwischen 3 und 16 Wörtern (siehe Abschnitt 6.3).

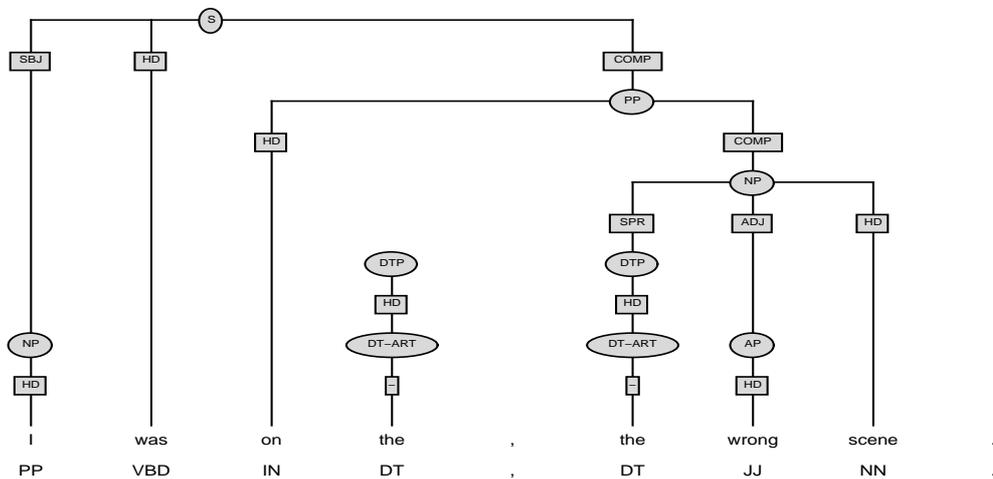


Abbildung 6.5: Satz enthält Wiederholung (Nr. 34, CD 13)

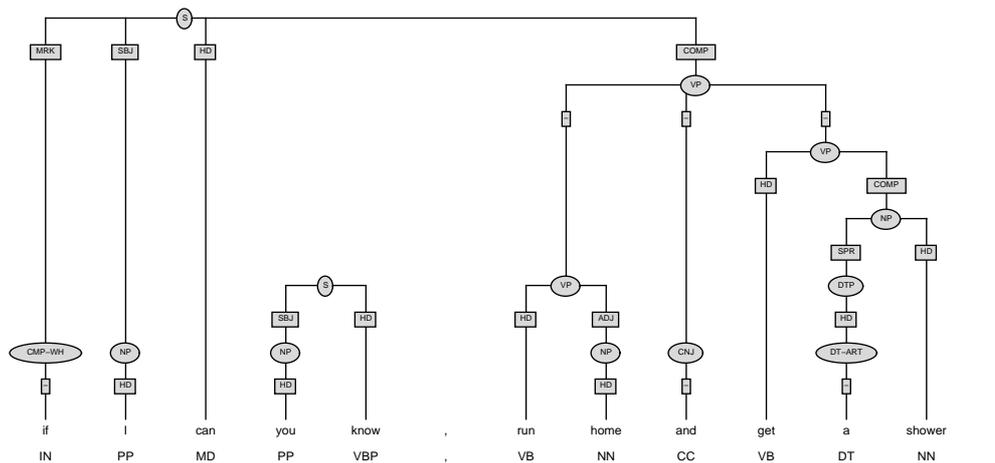


Abbildung 6.6: Satz mit satzwertigem Einschub (Nr. 1086, CD 13)

strukturell nicht vereinfacht werden (im Gegensatz zu den Schulbuchtexten in der Towsey-Studie, siehe Abschnitt 6.3.3.1).

Nach der Vorverarbeitung des Korpusfragments (CD13) enthalten die Daten noch 2678 Sätze bzw. 23.556 Wörter. Diese bilden das eigentliche Trainingskorpus. Einige Beispielsätze daraus sind in (18) dargestellt. Die Sätze in (18c-g) enthalten rekursiv eingebettete Nominalphrasen (*NP*) bzw. rekursiv eingebettete Sätze (*S*). Die Sätze in (18h-j) sind unvollständig bzw. wurden abgebrochen. Sie sind aber dennoch in den Trainingsdaten enthalten, weil sie in der Baumbank als “Sätze” annotiert sind (siehe dazu auch Fußnote 20).

- (18) a. [*S* that sounds good]
- b. [*S* for the next two weeks I will be busy]
- c. [*S* I think [*S* [*NP* Wednesday [*NP* the third] at three] is fine for me]]
- d. [*S* the fifth I have [*NP* [*NP* a seminar from ten to two thirty] and [*NP* a lecture from three to five]]]
- e. [*S* now you said you were free on [*NP* the mornings of [*NP* the twenty ninth] and [*NP* the thirtieth]]
- f. [*S* so [*S* if you could tell me some times around that [*S* when you are free]] that would be great]
- g. [*S* I still could not hear you very well] but [*S* I think you said two o'clock on Saturday] and [*S* that sounds good with me]
- h. [*S* Friday is]
- i. [*S* I am a little]
- j. [*S* I am starting to get]

39 Sätze (= 1,4%) sind bei der Vorverarbeitung aufgrund einer Satzlänge von über 25 Wörtern ausgefiltert worden. Trotz der Längenbegrenzung sind die hier verwendeten Trainingssätze deutlich länger als die Sätze in den bisherigen

Simulationen. Dies betrifft sowohl die durchschnittliche Satzlänge (8,8 Wörter) als auch die maximale Satzlänge (25 Wörter).²²

Nach der Vorverarbeitung werden aus den Sätzen im Trainingskorpus die Trainingsmuster generiert und diese anschließend in Binärcode übersetzt. Im folgenden Abschnitt wird die Kodierung der Trainingsmuster beschrieben. In Abschnitt 6.4.3 wird auf die binäre Repräsentation der Trainingsmuster eingegangen. Die Prozeduren für die Vorverarbeitung der Trainingsdaten und die Generierung der Trainingsmuster habe ich in der Programmiersprache “Java” implementiert.

6.4.2 Lernaufgabe und Kodierung der Trainingsmuster

Da für das Training echte Korpora verwendet werden, wird die Satzantizipations-Aufgabe, wie in Towsey et al. (1998), auf Kategorienebene (Wortarten) anstatt auf Wortebene durchgeführt.²³ Das Netz hat dabei die Aufgabe, für eine oder mehrere vorgegebene morpho-syntaktische Kategorien die nachfolgende Kategorie vorherzusagen. Das Lernszenario auf Kategorienebene ist außerdem psychologisch plausibler, da man an einer Stelle im Satz eher eine bestimmte Wortart, wie z.B. ein Nomen, erwartet und nicht etwa ein spezielles Wort (mit Ausnahme von sehr eingeschränkten Kontexten).²⁴

Für die Satzantizipation wird die morpho-syntaktische Annotation der TÜBA-E verwendet, die auf dem standardisierten Tagset der ‘*Penn Treebank*’ (Marcus et al., 1993) basiert. Die Trainingsdaten enthalten somit 36 morpho-syntaktische Kategorien (Tagset der ‘*Penn Treebank*’ ohne Satzzeichen) sowie eine zusätzliche Kategorie für das “Leerzeichen”. In

²² Ein Vergleich der vorliegenden Arbeit mit den in Abschnitt 6.3 beschriebenen Simulationen ist in Tabelle 6.5 dargestellt.

²³ Wie bereits in Abschnitt 6.3.3.1 begründet, funktioniert die Satzantizipations-Aufgabe *auf Wortebene* nur bei einem sehr kleinen Vokabular und nicht unter Verwendung von echten Texten. In echten Texten kommen die einzelnen Wörter zu selten im Trainingskorpus vor (‘sparse data problem’).

²⁴ Bei der menschlichen Satzverarbeitung spielt die Antizipation von morpho-syntaktischen Kategorien eine Rolle. Wird die Erwartungshaltung für eine bestimmte Kategorie (Wortart) verletzt, lässt sich das in EEG-Experimenten messen (Hahne und Friederici, 1999; Friederici, 2002).

Tabelle 6.1 sind die verwendeten Kategorien beschrieben und ihre Positionen im Netzwerk dargestellt.

Durch den Einsatz echter Korpusdaten sind die hier verwendeten Trainingsdaten wesentlich komplexer als die in den bisher beschriebenen Simulationen. Das Netz muss eine größere Anzahl an Kategorien verarbeiten, entsprechend findet sich eine größere Variation in den syntaktischen Phrasen und die Sätze sind länger, was wiederum längere Abhängigkeiten zwischen den Phrasen nach sich zieht. Ich habe mich daher entschieden, dem Netz einen größeren Kontext als Input vorzugeben. Anstelle des üblichen 1-Kategorien-Kontexts verwende ich sieben Kategorien als Kontext.²⁵

Während des Trainings wird jeder Satz aus dem Trainingskorpus einzeln präsentiert, wobei die Sätze als Kategorien-Sequenzen vorliegen. In (19) ist ein Satz aus dem Trainingskorpus (siehe (16b)) mit der dazugehörigen Kategorien-Sequenz dargestellt. Jeder Satz wird inkrementell von links nach rechts verarbeitet (Technik des beweglichen Fensters), d.h. jede Kategorie im Trainingskorpus muss bei gegebenem Kontext einmal antizipiert werden. Der Kontext besteht aus den sieben Vorgängerkategorien, was insbesondere für die Verarbeitung längerer Sätze von Bedeutung ist. Bei kurzen Sätzen und am Satzanfang werden die nicht besetzten Kontextelemente mit Leerstellen aufgefüllt. Die sieben Kontextelemente stellen die Eingabe für das Netz dar, die zu antizipierende Nachfolge-Kategorie ist die Ausgabe des Netzes. Eingabe und Ausgabe bilden zusammen ein Trainingsmuster.

(19) I	was	on the	wrong	scene							(Wörter)
	<i>PP</i>	<i>VBD</i>	<i>IN</i>	<i>DT</i>	<i>JJ</i>	<i>NN</i>					(Kategorien)

In Tabelle 6.2 ist die Trainingsmuster-Sequenz für die Verarbeitung des Satzes ‘*I was on the wrong scene*’ in (19) dargestellt. Bei der Antizipation der ersten Kategorie *PP* besteht der Kontext aus sieben Leerstellen. Das erste Trainingsmuster für den Satz in (19) besteht also aus sieben Leerstellen als Eingabe und der ersten Kategorie *PP* (‘personal pronoun’) als Ausgabe.

²⁵ Inspiriert wurde diese Kontextgröße durch die magische Zahl “Sieben” aus Miller (1956). Miller zeigte in diesem Aufsatz, dass Menschen ca. sieben Elemente plus/minus zwei (z.B. Ziffern, Buchstaben, Wörter) im Arbeitsgedächtnis behalten können.

Zelle	Wortart	Beschreibung
1	–	<empty>
2	CC	Coordinating conjunction
3	CD	Cardinal number
4	DT	Determiner
5	EX	Existential <i>there</i>
6	FW	Foreign word
7	IN	Preposition / subordinating conjunction
8	JJ	Adjective
9	JJR	Adjective, comparative
10	JJS	Adjective, superlative
11	LS	List item marker
12	MD	Modal
13	NN	Noun, singular or mass
14	NNS	Noun, plural
15	NP	Proper noun, singular
16	NPS	Proper noun, plural
17	PDT	Predeterminer
18	POS	Possessive ending
19	PP	Personal pronoun
20	PP\$	Possessive pronoun
21	RB	Adverb
22	RBR	Adverb, comparative
23	RBS	Adverb, superlative
24	RP	Particle
25	SYM	Symbol (mathematical or scientific)
26	TO	<i>to</i>
27	UH	Interjection
28	VB	Verb, base form
29	VBD	Verb, past tense
30	VBG	Verb, gerund / present participle
31	VBN	Verb, past participle
32	VBP	Verb, non-3rd person singular present
33	VBZ	Verb, 3rd person singular present
34	WDT	<i>wh</i> -determiner
35	WP	<i>wh</i> -pronoun
36	WP\$	Possessive <i>wh</i> -pronoun
37	WRB	<i>wh</i> -adverb

Tabelle 6.1: Verwendete morpho-syntaktische Kategorien (Wortarten) und ihre Positionen im Netzwerk

Trainings- muster	Eingabe (7 Kategorien)	Ausgabe (Folge-Kategorie)	
Nr. 1	-----	<i>PP</i>	personal pronoun
Nr. 2	----- <i>PP</i>	<i>VBD</i>	verb, past tense
Nr. 3	----- <i>PP VBD</i>	<i>IN</i>	preposition
Nr. 4	----- <i>PP VBD IN</i>	<i>DT</i>	determiner
Nr. 5	----- <i>PP VBD IN DT</i>	<i>JJ</i>	adjective
Nr. 6	----- <i>PP VBD IN DT JJ</i>	<i>NN</i>	noun, singular

Tabelle 6.2: Trainingsmuster-Sequenz, die der inkrementellen Verarbeitung des Satzes ‘*I was on the wrong scene*’ (siehe (19)) entspricht.

Bei der Antizipation der zweiten Kategorie besteht der Kontext aus sechs Leerstellen und der Vorgängerkategorie *PP* (Eingabe). Die zweite Kategorie *VBD* (‘verb, past tense’) stellt die Ausgabe dar. Das dritte Trainingsmuster besteht aus fünf Leerstellen plus zwei Vorgängerkategorien *PP* und *VBD* als Eingabe und der dritten Kategorie *IN* (‘preposition’) als Ausgabe, usw. Da im Trainingskorpus jede Kategorie einmal antizipiert wird, entspricht die Anzahl der Trainingsmuster genau der Anzahl der Wörter (bzw. Kategorien). Die hier verwendeten Trainingsdaten bestehen daher aus insgesamt 23.556 Trainingsmustern.

Wie aus Tabelle 6.1 zu entnehmen ist, verwende ich – im Gegensatz zu den in Abschnitt 6.3 beschriebenen Arbeiten – kein Symbol für das Satzende. Dafür gibt es zwei Gründe:

1. Ein Satzende kann nie eindeutig vorhergesagt werden, da ein Satz prinzipiell immer fortgesetzt werden kann, wie z.B. durch Relativsätze, koordinative Verknüpfungen, etc.
2. In den vorliegenden spontansprachlichen Daten sind die Äußerungen oft unvollständig oder werden abgebrochen. Bei solchen Äußerungen würde es keinen Sinn machen, ein Satzende vorherzusagen.

Nachdem die letzte Kategorie eines Satzes antizipiert wurde, folgt anstelle eines Satzende-Zeichens der nächste Satz, der wieder mit sieben Leerstellen als Kontext beginnt. Um die Satzantizipations-Aufgabe erfolgreich lösen zu können, muss das Netz die folgenden Teilaufgaben lernen:

- Erwerb von komplexen Phrasen: Welche Kategorien-Sequenzen bilden eine Phrase?
- Kategorisierung der Phrasen (Nominalphrasen, Präpositionalphrasen, etc.)
- Strukturelle Zusammenhänge zwischen Phrasen: grammatische Funktionen, Argumentstruktur, Subjekt-Verb-Kongruenz.

Diese Teilaufgaben sind denen in der Towsey-Studie recht ähnlich, was zum großen Teil daran liegt, dass die Satzantizipations-Aufgabe in Towsey et al. (1998) ebenfalls auf Kategorienebene durchgeführt wurde. Wie bereits erwähnt, sind die hier verwendeten Trainingsdaten jedoch wesentlich komplexer als in der Towsey-Studie (siehe auch Tabelle 6.5).

Durch die größere Anzahl an Wortarten (37 anstatt 10 Kategorien) ergibt sich eine größere Vielfalt an Phrasen und vor allem wesentlich mehr Variation im Phrasen- und Satzaufbau. Die Phrasen und Sätze sind außerdem länger und komplexer und können rekursiv eingebettet sein (siehe (18c-g)). So kann z.B. eine Nominalphrase weitere modifizierende Nominalphrasen enthalten wie in [*NP Wednesday [NP the third] at three*] (18c), Sätze können weitere eingebettete Sätze enthalten (siehe (18f,g)). Dadurch ergeben sich komplexe Abhängigkeitsrelationen zwischen Phrasen über größere Distanzen. Die Trainingsdaten in der Towsey-Studie enthalten nur einfache, relativ kurze Hauptsätze, eingebettete Nebensätze wurden vorher ausgefiltert.

Im Gegensatz zu den Towsey-Simulationen muss das Netz in dieser Arbeit (zumindest in eingeschränkter Form) auch die Kongruenz zwischen Subjekt und Verb erlernen, was daran liegt, dass das verwendete Tagset der *Penn Treebank* unterschiedliche Kategorien für Nomen im Singular bzw. Plural hat. Bei Verben wird im Präsens immerhin zwischen der Verbform für die dritte Person Singular (z.B. 'goes') und der Verbform für die anderen Personen (z.B. 'go') unterschieden. So muss das Netz lernen, dass bei einem Subjekt im Plural, wie z.B. 'boys', die

Verb-Kategorie für die dritte Person Singular (z.B. ‘goes’) für die Antizipation nicht in Frage kommt.

Der Erwerb der Subjekt-Verb-Kongruenz setzt allerdings voraus, dass das Netz “weiß”, welche Phrase als Subjekt fungiert und ggf. welches Element Kopf der Phrase ist. In einem Satz muss das Netz erkennen, dass die Sequenz ‘*the mornings of the twenty ninth and the thirtieth*’ eine komplexe Nominalphrase ist und ob diese Phrase Subjekt des Satzes ist. Falls ja, ist es für die Antizipations-Aufgabe relevant, ob diese Phrase im Singular oder Plural steht, da ein Abgleich mit dem zugehörigen Verb stattfinden muss. Das Plural-Merkmal kann jedoch nur dann korrekt erkannt werden, wenn das Netz “weiß”, dass ‘*mornings*’ das Element ist, das den Numerus bestimmt, d.h. der Kopf der Phrase ist (und nicht etwa ‘*thirtieth*’). Dieses Beispiel zeigt, dass das Netz für eine korrekte Satzantizipation nicht nur komplexe Phrasen erwerben und kategorisieren muss, sondern den Phrasen auch grammatische Funktionen (wie Subjekt, Objekt, etc.) zuweisen muss, wobei der Kopfbegriff eine Rolle spielt.

In den Towsey-Simulationen muss die Subjekt-Verb-Kongruenz nicht erlernt werden, da das dort verwendete Tagset jeweils nur eine Kategorie für Nomen und Verben enthält, die hinsichtlich Numerus nicht spezifiziert sind. Das vereinfacht die Antizipations-Aufgabe natürlich.

Schließlich müssen die Netze bei der Satzantizipation immer auch eine Form der Argument-Struktur erwerben. Je nachdem welche Verben in den Trainingssätzen vorkommen, ist eine andere Zahl und Art von Ergänzungen erforderlich. Während ein Verb wie ‘*sound*’ neben dem Subjekt noch eine adjektivische Ergänzung benötigt (‘*that sounds good*’), ist für das Verb ‘*find*’ ein direktes Objekt erforderlich (‘*we will find a solution*’). Da weder das in der Towsey-Studie verwendete Tagset noch das hier verwendete Tagset der *Penn Treebank* eine Unterscheidung von Verben hinsichtlich ihrer Argumentstruktur vornimmt, kann das Netz in beiden Arbeiten keine explizite Assoziation zwischen einem Verb und seiner Argumentstruktur lernen. Das Netz kann jedoch lernen, welche Argumentstrukturen in den Trainingssätzen vorkommen, und Präferenzen aufgrund allgemeiner Häufigkeiten entwickeln.

Nach diesen Ausführungen wird deutlich, dass die Komplexität und die genaue Art der Klassifikationsaufgabe sowohl von der Länge, Komplexität und Art der Trainingssätze abhängt, als auch von der Auswahl der morpho-syntaktischen Kategorien, d.h. dem Tagset (siehe Tabelle 6.1). Durch das Tagset wird vorgegeben, welche Informationen das Netz für die Satzantizipations-Aufgabe zur Verfügung hat. Wenngleich die Satzantizipation der hier verwendeten spontansprachlichen Korpora wesentlich komplexer ist, als in den bisher beschriebenen Simulationen²⁶, so ist dennoch zu erwarten, dass durch die Verwendung des größeren Kontexts (sieben Kategorien) die Aufgabe lösbar wird.

6.4.3 Simulationen

Für die Simulation wird ein Elman-Netzwerk verwendet, dessen Ausgabeschicht aus 37 Zellen besteht, wobei jede Zelle eine morpho-syntaktische Kategorie repräsentiert (siehe Abbildung 6.7). In jedem Trainingsmuster ist daher genau eine Zelle der Ausgabeschicht aktiviert, welche der jeweiligen Zielkategorie entspricht. Da eine Sequenz von sieben Kategorien die Eingabe für das Netz bildet, besteht die Eingabeschicht aus $7 \times 37 = 259$ Zellen. Die ersten 37 Zellen (erster Zellkomplex) repräsentieren die erste Kategorie, die nächsten 37 Zellen (zweiter Zellkomplex) repräsentieren die zweite Kategorie, usw. (siehe Tabelle 6.3). In jedem Zellkomplex ist genau eine Zelle aktiviert, die der jeweiligen morpho-syntaktischen Kategorie an dieser Kontextposition entspricht. Welche der 37 Zellen welche Kategorie innerhalb eines Zellkomplexes repräsentiert, kann Tabelle 6.1 entnommen werden. Insgesamt sind in jedem Trainingsmuster genau sieben Zellen der Eingabeschicht aktiviert, die den sieben Kontext-Kategorien entsprechen. In Tabelle 6.4 ist die binäre Repräsentation eines Trainingsmusters veranschaulicht (Trainingsmuster Nr. 4 aus Tabelle 6.2). Die verdeckte Schicht und die Kontextschicht bestehen aus je 100 Zellen.

²⁶ Ein Vergleich der vorliegenden Arbeit mit den in Abschnitt 6.3 beschriebenen Simulationen ist in Tabelle 6.5 dargestellt.

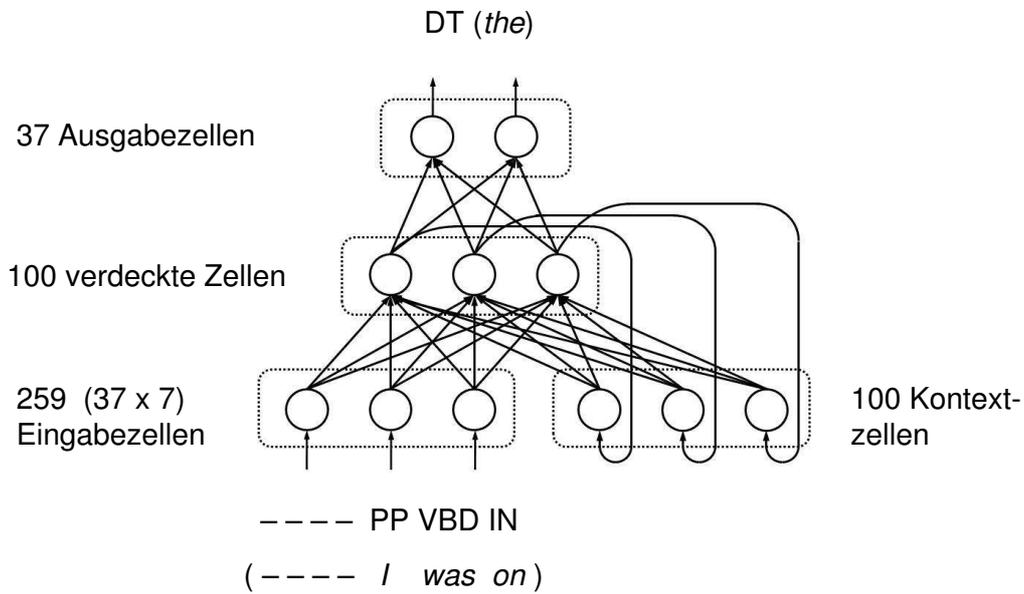


Abbildung 6.7: Verwendetes Elman-Netzwerk mit 259 (37 x 7) Eingabezellen, 100 Kontextzellen, 100 verdeckten Zellen und 37 Ausgabezellen.

Zellkomplex	Eingabezellen	Kontextelement
1	Zelle 1 – 37	1. Kategorie
2	Zelle 38 – 74	2. Kategorie
3	Zelle 75 – 111	3. Kategorie
4	Zelle 112 – 148	4. Kategorie
5	Zelle 149 – 185	5. Kategorie
6	Zelle 186 – 222	6. Kategorie
7	Zelle 223 – 259	7. Kategorie

Tabelle 6.3: Repräsentation der sieben Kontextelemente im Eingabemuster

Das Netzwerk wurde mit dem Lernalgorithmus *Standard Backpropagation* für Elman-Netze²⁷ (siehe Abschnitt 6.3.1.3) trainiert. Das Training umfasste 400 Durchgänge (Zyklen) durch das Trainingskorpus mit einer Lernrate von $\eta = 0,2$. Die Verbindungsgewichte wurden vor dem Training mit zufälligen Werten zwischen $-0,3$ und $0,3$ initialisiert, die Kontextzellen wurden auf einen Anfangswert von $0,2$ gesetzt.²⁸ Die Simulationen wurden mit dem *Stuttgarter Neuronale Netze Simulator* (SNNS, Zell et al., 1998) auf folgendem PC-System durchgeführt: Dual AMD Athlon MP 2200, 1,8 GHz, 2 GB RAM. Eine Simulation (400 Zyklen) dauerte 12 Stunden. Ein Vergleich mit den in Abschnitt 6.3 beschriebenen Simulationen ist in Tabelle 6.5 dargestellt.

6.4.4 Evaluation des Netzes

Nach dem Training wird das Elman-Netz auf zweierlei Arten evaluiert. Einerseits teste ich, wie gut das Netz die Satzantizipations-Aufgabe auf den Trainingsdaten gelernt hat, d.h. wie gut die zugrundeliegenden grammatischen Regeln gelernt wurden. Weiterhin vergleiche ich die Performanz des Netzes mit menschlicher Leseverarbeitung. Dadurch wird gleichzeitig auch die Generalisierungsfähigkeit des Netzes getestet. Es wird untersucht, ob das Netz ähnliche Präferenzen und Verarbeitungseffekte aufweist wie Menschen beim Lesen. Das sind Informationen, die dem Netz nicht explizit beigebracht wurden. Eine gut gelernte Grammatik ist die Voraussetzung dafür, das Netz mit menschlichen Verarbeitungseffekten vergleichen zu können. Im Folgenden wird darauf eingegangen, wie die Evaluation im Einzelnen durchgeführt wird.

²⁷ Es handelt sich hierbei um ein Online-Trainingsverfahren, d.h. die Verbindungsgewichte wurden nach jedem Trainingsmuster angepasst. Die Aktivierungsfunktionen, die dieser Simulation zugrunde lagen, war die identische Abbildung für Eingabe- und Kontextzellen und die logistische Funktion für verdeckte Zellen und Ausgabezellen. Als Ausgabefunktion wurde die identische Abbildung für alle Zellen verwendet.

²⁸ Vor dem Training wurden einige Testdurchläufe mit verschiedenen Parametern durchgeführt (Lernrate, Initialisierung). Die Unterschiede hinsichtlich der Fehlerrate des Netzes (MSE) waren gering. Die besten Resultate wurden mit einer Lernrate von $\eta = 0,2$ und einer Initialisierung der Gewichte zwischen $[-0,3; 0,3]$ erreicht. Mit diesen Parametern wurde daher das eigentliche Training durchgeführt.

	Elman (1990)	Elman (1991)	Towsey et al. (1998)	Christiansen & Chater (1999, 2001)	Steiner (2007b)
<i>Trainingsdaten</i>	2- und 3-Wort- Sätze	Durch PSG erzeugte Sätze	Schulbuch- texte 1. Klasse	Künstliche formale Sprachen	Spontan- sprachliche Korpora
<i>Lernaufgabe: Antizipation von ...</i>	Folge- Wort	Folge- Wort	Folge- Kategorie	Folge-Wort	Folge- Kategorie
<i>Kontext (Input)</i>	1 Wort	1 Wort	1 Kategorie	1 Wort	7 Kategorien
<i>Anzahl Sätze</i>	10.000	4 x 10.000	80	5000	2678
<i>Anzahl Trainingsmuster</i>	27.354	178.000	486	23.500	23.556
<i>Durchschnittl. Satzlänge</i>	2,7	4,5	6,1	4,7	8,8
<i>Max. Satzlänge</i>	3	16	8	9	25
<i>Anzahl Ausgabeklassen</i>	29	23	10	17	37
<i>Elman-Netz (Zellen: Input/ Hidden/Output)</i>	31/150/ 31	26/70/ 26	10/[1-50]/ 10	17/[2-100]/ 17	259/100/ 37
<i>Lern- algorithmus</i>	Backprop.	Backprop.	Backprop. Momentum	Standard Backprop.	Standard Backprop.
<i>Lernzyklen</i>	6	5	400/ 100.000	100	400

Tabelle 6.5: Vergleich mit den in Abschnitt 6.3 beschriebenen Simulationen

In der Testphase werden die Verbindungsgewichte eingefroren. Um die Performanz des Netzes zu testen, werden dem Netz ausschließlich Eingabemuster vorgegeben und durch das trainierte Netz propagiert. Die Ausgabe des Netzes wird anschließend mit der erwünschten Ausgabe verglichen. Die Abweichung zwischen tatsächlicher und erwünschter Ausgabe stellt den Fehler des Netzes dar. Grundsätzlich stellen sich bei dieser Berechnung zwei Fragen.

1. *Frage nach dem Gold-Standard* Was ist eigentlich die erwünschte Ausgabe? Womit wird die Netzausgabe verglichen?
2. *Frage nach dem Fehlermaß*: Wie wird die Differenz (Fehler) zwischen Netzausgabe und erwünschter Ausgabe berechnet?

Diese Fragen werden in den beiden folgenden Abschnitten 6.4.4.1 und 6.4.4.2 beantwortet. In Tabelle 6.6 sind die verwendeten Gold-Standards und Fehlermaße zusammengefasst.

6.4.4.1 Der verwendete Gold-Standard

Die Fragen nach dem Gold-Standard muss unterschiedlich beantwortet werden, je nachdem, ob es sich um die erwünschte Ausgabe beim Training handelt, um die erwünschte Ausgabe beim Testen der grammatischen Fähigkeiten oder um die erwünschte Ausgabe beim Vergleich mit der menschlichen Verarbeitung.

Die Frage nach dem Gold-Standard ist deshalb so kompliziert, weil die Satzantizipations-Aufgabe nicht deterministisch ist (siehe Elman (1990) in Abschnitt 6.3.2.1). Die Nachfolger-Kategorie an einer bestimmten Position im Satz kann nicht eindeutig vorhergesagt werden. Es ist jedoch möglich, die grammatisch *möglichen Nachfolger* und die damit verbundenen Übergangswahrscheinlichkeiten aus den Trainingsdaten zu lernen. Genau das macht das Netz auch, wenn ihm während des Trainings die Nachfolger-Kategorie im Satz als erwünschte Ausgabe präsentiert wird. Dies wurde in Towsey et al. (1998) gezeigt (siehe Abschnitt 6.3.3). Das Netz lernt gewissermaßen auf indirekte Weise aus den Kategorien-Sequenzen eine probabilistische Grammatik.

Es ist daher sinnvoll, beim Testen der grammatischen Fähigkeiten einen anderen Gold-Standard zu verwenden als beim Training. Bei der Wahl des Gold-Standards muss meiner Ansicht nach weiterhin unterschieden werden, ob das Netz auf grammatische Fähigkeiten hin getestet wird oder ob das Netz mit menschlichen Lesezeiten verglichen wird. Welchen Vergleichsmaßstab ich jeweils verwende, wird in den folgenden Unterabschnitten beschrieben.

Beim Training: Während des Trainings wird, wie auch in den beschriebenen Simulationen aus der Literatur, die Nachfolger-Kategorie im Satz als erwünschte Ausgabe verwendet. Das ergibt sich bereits aus der Lernaufgabe, die in Abschnitt 6.4.2 beschrieben wird. Zur Berechnung des Fehlers während des Trainingsprozesses wird für jedes Trainingsmuster der Ausgabevektor des Netzes mit dem Vektor verglichen, der die Nachfolger-Kategorie im Satz kodiert. Die beiden zu vergleichenden Vektoren für das Trainingsmuster aus Tabelle 6.4 sind in (20) illustriert. In diesem Trainingsmuster ist die Kategorie *DT* ('determiner') die Nachfolger-Kategorie und somit die erwünschte Ausgabe.

- (20) a. [0,0 0,0 0,0 1,0 ... 0,0] (Vektor mit Nachfolger-Kategorie *DT*)
 b. [0,0 0,0 0,1 0,8 ... 0,2] (Ausgabevektor des Netzes)
 – *CC CD DT ... PP*

Der Vektor in (20a) kodiert die Nachfolger- bzw. Ausgabe-Kategorie und enthält genau eine mit 1,0 aktivierte Zelle. Das ist die Zelle, die die Ziel-Kategorie *DT* repräsentiert. Alle anderen Zellen sind nicht aktiviert. In (20b) ist der Vektor dargestellt, der die Netzausgabe für dieses Trainingsmuster repräsentiert. In der Netzausgabe können prinzipiell alle Zellen mehr oder weniger stark aktiviert sein, d.h. Werte zwischen 0 und 1 annehmen. In (20b) ist die korrekte Kategorie *DT* mit 0,8 am stärksten aktiviert, weitere Kategorien, wie z.B. *CD* ('cardinal number') oder *PP* ('personal pronoun'), sind schwach aktiviert (0,1 bzw. 0,2).

Beim Testen der grammatischen Fähigkeiten: Hier wird im Allgemeinen eine Wahrscheinlichkeitsverteilung als erwünschte Ausgabe verwendet. Dabei wird untersucht, ob das Netz die grammatisch *möglichen Nachfolger* und ihre Gewichtung antizipieren kann. Elman (1990, 1991) sowie Christiansen und Chater (1999, 2001) verwenden die absolute bedingte Wahrscheinlichkeitsverteilung als erwünschte Ausgabe, Towsey et al. (1998) vergleichen die Netzausgabe mit der Nachfolger-Kategorie im Satz sowie mit einer Wahrscheinlichkeitsverteilung auf Basis eines kleineren Kontexts (n -Gramme mit $n = 2-4$).

Ich verwende die *absolute bedingte* Wahrscheinlichkeitsverteilung als erwünschte Ausgabe (Gold-Standard) beim Testen. Ein Vergleich mit der Nachfolger-Kategorie im Satz ist im Grunde nicht sinnvoll, da diese Vorhersage nicht deterministisch ist. Um jedoch einen Vergleich mit den Ergebnissen aus der Towsey-Studie durchführen zu können, gebe ich zusätzlich die Performanz des Netzes für die Vorhersage der Nachfolger-Kategorie an.

Während die Nachfolger-Kategorie für jedes Trainingsmuster bereits aus der Trainingsphase vorliegt, muss die absolute bedingte Wahrscheinlichkeitsverteilung, d.h. der Gold-Standard, vor der Testphase erst einmal berechnet werden. Bei der Berechnung werden den Wahrscheinlichkeiten die Vorkommenshäufigkeiten im Trainingskorpus zugrunde gelegt. Für die Auftretenswahrscheinlichkeit einer Kategorie c_p sind alle Vorgänger-Kategorien in diesem Satz c_1, c_2, \dots, c_{p-1} relevant. Auf Basis der Häufigkeiten der Vorgänger-Kategorien c_1, c_2, \dots, c_{p-1} sowie der Nachfolger-Kategorien c_p im Trainingskorpus berechne ich die Wahrscheinlichkeit $P(c_p)$ für jede der 37 möglichen Folge-Kategorien in jedem Trainingsmuster nach Formel (6.6). Die Summe der Wahrscheinlichkeiten für die 37 Folge-Kategorien in einem Trainingsmuster ist immer 1,0. Auf diese Weise werden die Wahrscheinlichkeitsvektoren generiert, die als Vergleichsmaß, d.h. als erwünschte Ausgabe, fungieren.

$$P(c_p | c_1, c_2, \dots, c_{p-1}) = \frac{\text{Freq}(c_1, c_2, \dots, c_{p-1}, c_p)}{\text{Freq}(c_1, c_2, \dots, c_{p-1})} \quad (6.6)$$

In der Testphase vergleiche ich für jedes Trainingsmuster die Netzausgabe mit den berechneten Wahrscheinlichkeitsvektoren, d.h. mit den Übergangswahrscheinlichkeiten für die grammatisch möglichen Nachfolger. In (21) sind die beiden zu vergleichenden Vektoren für das Trainingsmuster aus Tabelle 6.4 dargestellt.

$$\begin{array}{ll}
 (21) \text{ a. } [0,0 \ 0,0 \ 0,0 \ 0,7 \ \dots \ 0,3] & \text{(Wahrscheinlichkeitsvektor)} \\
 \text{b. } [0,0 \ 0,0 \ 0,1 \ 0,8 \ \dots \ 0,2] & \text{(Ausgabevektor des Netzes)} \\
 \text{– } CC \ CD \ DT \ \dots \ PP &
 \end{array}$$

In (21a) ist der Wahrscheinlichkeitsvektor abgebildet, der die erwünschte Ausgabe für dieses Trainingsmuster beim Testen darstellt. Im Gegensatz zur erwünschten Ausgabe während des Trainings, wie in (20a), können hier mehrere Zellen aktiviert sein. Die aktivierten Zellen repräsentieren die grammatisch möglichen Nachfolger-Kategorien (hier: Kategorie *DT* und *PP*), die Stärke der Aktivierung kodiert die jeweilige Übergangswahrscheinlichkeit. Die Summe der Aktivierungen ergibt immer 1,0. Der Vektor in (21b) kodiert die Ausgabe-Aktivierungen des Netzes. Man sieht, dass die Netzausgabe (21b) dicht an der Wahrscheinlichkeitsverteilung im Korpus (21a) liegt: eine stark aktivierte Kategorie *DT* und eine schwach aktivierte Kategorie *PP*.

Da ich die maximale Satzlänge auf 25 Wörter festgelegt habe und die nicht besetzten Kontextelemente mit Leerstellen (Kategorie: –) aufgefüllt wurden, entspricht die absolute bedingte Wahrscheinlichkeitsverteilung in unserem Trainingskorpus der Vorhersage eines n -Gramm-Modells mit $n = 25$. Das Netz wird also mit 7 Kontextelementen trainiert, aber auf Basis von 24 Kontextelementen auf grammatische Fähigkeiten getestet.

Beim Vergleich mit der menschlichen Verarbeitung: Christiansen und Chater (1999, 2001) verwenden für den Vergleich mit der menschlichen Performanz ein spezielles Fehlermaß, den ‘*Grammatical Prediction Error*’ (GPE), der jedoch letztlich auch auf einer Wahrscheinlichkeitsverteilung basiert. Bei der Wahl des Gold-Standards unterscheiden die Autoren nicht zwischen dem Testen der grammatischen Fähigkeiten und dem Vergleich mit

menschlicher Performanz. Sie vergleichen die Netzausgabe in beiden Fällen mit einer Wahrscheinlichkeitsverteilung und verwenden lediglich unterschiedliche Fehlermaße (MSE vs. GPE).

Meiner Ansicht nach ist es jedoch notwendig, beim Vergleich mit menschlichen Lesezeiten die Netzausgabe mit der Nachfolger-Kategorie im Satz zu vergleichen und nicht mit einer Wahrscheinlichkeitsverteilung. Das hat damit zu tun, dass die Lesezeiten nicht davon abhängen, wie gut ein Sprecher die Grammatik beherrscht, sondern davon, ob die Erwartungshaltung des Sprechers erfüllt oder verletzt wurde. Eine gut gelernte Grammatik kann über den Vergleich mit einer Wahrscheinlichkeitsverteilung evaluiert werden. Das ist aber nur die Voraussetzung dafür, eine Erwartungshaltung beim Lesen zu entwickeln.

Beim Vergleich mit Lesezeiten ist es daher relevant, ob die Vorhersage des Netzes mit der tatsächlichen Nachfolger-Kategorie im Satz übereinstimmt oder nicht. Stimmt sie mit der Nachfolger-Kategorie überein, dann wurde die Erwartungshaltung erfüllt und der Fehler des Netzes (Abstand zwischen der Netzausgabe und dem Vektor, der die Nachfolger-Kategorie kodiert) ist gering. Diese Situation ist in (22) illustriert. Die tatsächliche Nachfolger-Kategorie stimmt mit der Präferenz des Netzes überein (jeweils Kategorie *DT*). Dies entspricht einer schnellen Lesezeit. Stimmt die Vorhersage nicht mit der Nachfolger-Kategorie überein, wurde die Erwartungshaltung verletzt und der Fehler ist dementsprechend größer. Dies ist in (23) illustriert. Die tatsächliche Nachfolger-Kategorie *PP* stimmt nicht mit der Präferenz des Netzes überein. Das entspricht einer langsameren Lesezeit.

(22) a. [0,0 0,0 0,0 1,0 ... 0,0] (Vektor mit Nachfolger-Kategorie *DT*)
 b. [0,0 0,0 0,1 0,8 ... 0,2] (Ausgabevektor des Netzes)
 – *CC CD DT ... PP*

(23) a. [0,0 0,0 0,0 0,0 ... 1,0] (Vektor mit Nachfolger-Kategorie *PP*)
 b. [0,0 0,0 0,1 0,8 ... 0,2] (Ausgabevektor des Netzes)
 – *CC CD DT ... PP*

Im Gegensatz zu Christiansen und Chater (1999, 2001) verwende ich daher beim Vergleich mit menschlicher Performanz die Nachfolger-Kategorie im Satz als Gold-Standard. Auf die verwendeten Fehlermaße gehe ich im folgenden Abschnitt ein.

6.4.4.2 Die verwendeten Fehlermaße

Es stellt sich nun die Frage, wie der Fehler des Netzes, d.h. der Abstand zwischen Netzausgabe und erwünschter Ausgabe, berechnet wird. Während des Trainings wird standardmäßig der ‘*Mean square error*’ (MSE) als Fehlermaß verwendet. Um das Netz auf grammatische Fähigkeiten zu testen und mit menschlicher Verarbeitung zu vergleichen, verwende ich geeignetere Fehlermaße: a. korrekte Vorhersagen in Prozent, b. die halbe L_1 -Norm. Im Folgenden werden die verwendeten Fehlermaße beschrieben.

Mean square error (MSE): Beim MSE handelt es sich um den mittleren quadratischen Abstand zweier Vektoren, der nach folgender Formel berechnet wird.

$$MSE = \frac{SSE}{n} = \frac{\sum_n \sum_j (t_{nj} - o_{nj})^2}{n} \quad (6.7)$$

wobei n die Anzahl der Testmuster, j die Anzahl der Ausgabezellen und SSE der ‘sum square error’ ist. t_{nj} ist die erwünschte Ausgabe (‘teaching input’) für die Ausgabezelle j in Testmuster n und o_{nj} die tatsächliche Ausgabe, d.h. die Netzausgabe. Der MSE-Fehler wird über alle n Testmuster gemittelt.

Der MSE bzw. eine Variante davon (RMSE: ‘root mean square error’) wird in Elman (1990, 1991), in Towsey et al. (1998) und in Christiansen und Chater (1999, 2001) als Fehlermaß für das Training und für die Testphase verwendet. Ich verwende den ‘*Mean square error*’ ausschließlich für das Training. Für das Testen der grammatischen Fähigkeiten existieren geeignetere Fehlermaße (siehe unten).

Korrekte Vorhersagen in Prozent: Dieses Maß gibt an, wie viele Testmuster korrekt vorhergesagt werden. Die Vorhersage eines Testmusters wird als korrekt gewertet, wenn in der Netzausgabe die Zelle mit dem höchsten Aktivierungslevel mit der am stärksten aktivierten Komponente in der erwünschten Ausgabe übereinstimmt. Das Trainingsmuster in (21) und in (22) wird beispielsweise korrekt vorhergesagt, da sowohl in der Netzausgabe als auch in der erwünschten Ausgabe die Kategorie *DT* (4. Zelle) am stärksten aktiviert ist. Das Trainingsmuster in (23) hingegen wird nicht korrekt vorhergesagt: In der erwünschten Ausgabe ist die Kategorie *PP* am stärksten aktiviert, in der Netzausgabe ist es die Kategorie *DT*.

Dieses Fehlermaß wird in Towsey et al. (1998) in der Testphase verwendet. Ich verwende es ebenfalls in der Testphase sowie beim Vergleich mit der menschlichen Performanz.

Halbe L_1 -Norm: Dieses Maß ist besonders gut geeignet, um Wahrscheinlichkeitsverteilungen zu vergleichen (zur L_1 -Norm siehe Manning und Schütze, 1999, S. 304). Die Ausgabe des trainierten Netzes wird vor der Evaluation normiert, d.h. die Summe der Aktivierungen ergibt 1,0 (siehe Abschnitt 6.4.4.3). Der Abstand zwischen tatsächlicher und erwünschter Ausgabe wird nach folgender Formel berechnet.

$$\text{Halbe } L_1\text{-Norm} = \frac{\sum_n \frac{\sum_j (|t_{nj} - o_{nj}|)}{2}}{n} \quad (6.8)$$

wobei n die Anzahl der Testmuster ist und j die Anzahl der Ausgabezellen. t_{nj} ist die erwünschte Ausgabe ('teaching input') für die Ausgabezelle j in Testmuster n und o_{nj} ist die tatsächliche Ausgabe. Die halbe L_1 -Norm wird über alle n Testmuster gemittelt.²⁹

²⁹ Im Unterschied zum 'Mean square error' werden bei der L_1 -Norm die Differenzen zwischen tatsächlicher und erwünschter Ausgabe nicht quadriert, sondern es wird der Betrag genommen. In beiden Fällen kann so verhindert werden, dass die Differenzen negativ werden. Durch die Quadrierung beim MSE werden die Fehler jedoch unterschiedlich stark gewichtet: große Abweichungen werden durch die Quadrierung weniger stark verkleinert als kleine Abweichungen. Dadurch ist der MSE für den Vergleich von Wahrscheinlichkeitsverteilungen weniger geeignet.

	Gold-Standard	Fehlermaß
Training	Nachfolger	MSE
Grammatik-Erwerb	Nachfolger	Prozent, L_1 -Norm/2
	Bedingte Wahrscheinlichkeits- verteilung (absolut)	Prozent, L_1 -Norm/2
Vergleich mit Lesezeiten	Nachfolger	Prozent, L_1 -Norm/2

Tabelle 6.6: Verwendete Gold-Standards und Fehlermaße (MSE: Mean square error, Prozent: Korrekte Vorhersagen in Prozent, L_1 -Norm/2: Halbe L_1 -Norm)

Weder die L_1 -Norm noch die halbe L_1 -Norm wurde von anderen Autoren bislang zur Evaluation der Satzantizipations-Aufgabe verwendet. Ich verwende die halbe L_1 -Norm für das Testen der grammatischen Fähigkeiten sowie beim Vergleich mit der menschlichen Verarbeitung, da dieses Maß – im Gegensatz zum MSE – besonders gut geeignet ist, um Wahrscheinlichkeitsverteilungen zu vergleichen (siehe auch Fußnote 29) und es außerdem ein differenzierteres Bild liefert als die Anzahl der korrekten Vorhersagen. Auch für den Vergleich mit der menschlichen Performanz, d.h. mit der Nachfolger-Kategorie im Satz, ist die (halbe) L_1 -Norm besser geeignet als das MSE-Fehlermaß (siehe Tabelle 6.6).³⁰

Bei der L_1 -Norm werden die Fehler gleich stark gewichtet. Die L_1 -Norm wird halbiert (halbe L_1 -Norm), weil ein Fehler in der Aktivierung ansonsten doppelt gezählt werden würde. Eine zu schwach aktivierte Zelle hat grundsätzlich an anderer Stelle eine zu stark aktivierte Zelle zur Folge, da die Summe der Aktivierungen immer 1,0 ergibt.

³⁰ Vergleicht man die Netzausgabe mit einem Einheitsvektor (Nachfolger-Kategorie), ist bei der Berechnung durch die (halbe) L_1 -Norm nur die Abweichung zwischen der am stärksten aktivierten Kategorie in der Netzausgabe (erwartete Kategorie) und der tatsächlichen Nachfolger-Kategorie relevant. Wie die übrige abweichende Aktivierung auf die anderen Kategorien verteilt ist, spielt keine Rolle. Genau dies würde man auch für Lesezeiten erwarten. Beim MSE spielt die Verteilung der abweichenden Aktivierung auf andere Zellen durchaus eine Rolle, was für Lesezeiten jedoch unerheblich sein sollte.

6.4.4.3 Auswertungsprozedur

Nach dem Training liegt die Netzausgabe für jedes Eingabemuster aus dem Trainingskorpus vor. Von diesen Daten ausgehend führe ich die folgenden Schritte durch:

- Normierung der Netzausgabe, so dass die Summe der Aktivierungen in jedem Testmuster 1,0 ergibt.
- Vergleich der Netzausgabe mit der Nachfolger-Kategorie im Satz. Es wird jeweils die Übereinstimmung in Prozent (siehe Fehlermaß: “Korrekte Vorhersagen in Prozent”) sowie die halbe L_1 -Norm (siehe Fehlermaß: “Halbe L_1 -Norm”) berechnet.
- Generierung der absoluten bedingten Wahrscheinlichkeit für jedes Testmuster aus dem Trainingskorpus nach Formel (6.6).
- Vergleich der Netzausgabe mit der berechneten Wahrscheinlichkeitsverteilung. Es wird jeweils die Übereinstimmung in Prozent sowie die halbe L_1 -Norm berechnet.

Für die Evaluation der grammatischen Fähigkeiten wurde der Fehler über alle Muster aus dem Trainingskorpus gemittelt. Für den Vergleich mit der menschlicher Verarbeitung wurde der Fehler nur über die relevanten Muster aus dem Trainingskorpus gemittelt. Die Auswertungsprozeduren und die Berechnung der Gold-Standards habe ich in der Programmiersprache “Java” implementiert.

6.4.5 Performanz des Netzes

Bevor das Netz inhaltlich ausgewertet wird, illustriere ich in Abbildung 6.8 den Lernverlauf des Netzes, das ist die Performanz des Netzes zu verschiedenen Stadien des Lernprozesses. Während des Trainings wird die Netzausgabe mit der Nachfolger-Kategorie im Satz (Gold-Standard) verglichen, mit dem Ziel, den mittleren quadratischen Fehler (MSE) des Netzes zu minimieren. Diese MSE-Werte sind in Abbildung 6.8 dargestellt.

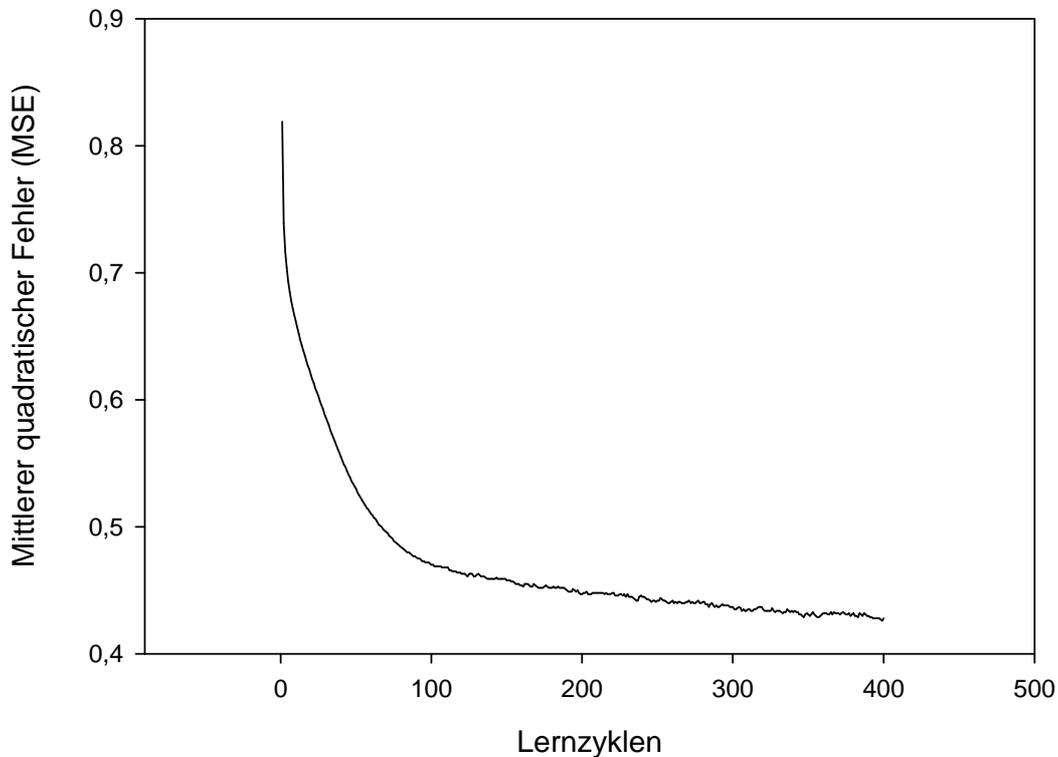


Abbildung 6.8: Lernverlauf des Netzes

Das Training umfasste 400 Zyklen durch das Trainingskorpus. Die Abbildung zeigt den MSE-Wert nach jedem Durchgang. Bei der verwendeten Aufgabenstellung und Netzwerk-Architektur kann der MSE prinzipiell Werte zwischen 0 und 37 annehmen. Eine Strategie, alle Ausgabezellen gleich stark zu aktivieren, um so den Fehler zu minimieren, würde im Idealfall einen MSE von 1,0 ergeben. Bereits nach einem Lernzyklus ist der MSE-Wert bei 0,82, nach 400 Lernzyklen hat das Netz noch einen Fehler von 0,43. Dies entspricht einem MSE-Fehler von 0,012 pro Ausgabezelle. Man sieht, dass das Elman-Netz einen kontinuierlichen Lernverlauf ohne nennenswerte Oszillationen aufweist.³¹

³¹ Eine direkte Vergleichbarkeit mit MSE- oder RMSE-Werten aus anderen Simulationen ist nicht möglich, da der MSE bzw. RMSE von der genauen Netzwerk-Architektur, z.B. der Anzahl der Ausgabezellen, und der Komplexität und Kodierung der Trainingsdaten abhängig ist.

Nach 400 Lernzyklen bei einem MSE-Wert von 0,43 werden die Verbindungsgewichte nicht mehr verändert und das Netz wird inhaltlich ausgewertet. An dieser Stelle wiederhole ich noch einmal die Fragestellungen, die in diesem Kapitel beleuchtet werden sollen. In den beiden folgenden Unterabschnitten werden die entsprechenden Ergebnisse präsentiert.

1. Ist ein Elman-Netz in der Lage, mittels der Satzantizipations-Aufgabe grammatische Regularitäten selbständig aus echten spontansprachlichen Korpusdaten zu lernen?
2. Wie verhält sich das Netz bei koordinierten Strukturen? Gibt es eine Korrelation zwischen der Performanz des Netzwerks und menschlichem Verhalten bei der Verarbeitung von Koordination?

6.4.5.1 Erwerb grammatischer Regeln

In diesem Abschnitt wird gezeigt, wie gut das Netz die Satzantizipations-Aufgabe für die Trainingsdaten gelernt hat. Diese Aufgabe kann nur dann korrekt ausgeführt werden, wenn das Netz die zugrundeliegenden grammatischen Regularitäten gelernt hat.

Die Ergebnisse werden auf unterschiedliche Arten präsentiert: Ich vergleiche die Netzausgabe mit zwei verschiedenen Gold-Standards und wende jeweils zwei verschiedene Fehlermaße an. Die Netzausgabe wird verglichen mit 1. der Nachfolger-Kategorie im Satz (erwünschte Ausgabe beim Training), und 2. der absoluten bedingten Wahrscheinlichkeitsverteilung im Trainingskorpus (Kontextgröße = 24). Der Abstand zwischen Netzausgabe und erwünschter Ausgabe wird jeweils angegeben a) in korrekte Vorhersagen in Prozent, und b) durch die halbe L_1 -Norm. Da die Satzantizipations-Aufgabe nicht deterministisch ist, ist ein Vergleich mit der tatsächlichen Nachfolger-Kategorie im Grunde nicht besonders sinnvoll. Dieser Vergleich wurde dennoch durchgeführt, um die Ergebnisse mit den Towsey-Ergebnissen vergleichen zu können.

Vergleicht man die Netzausgabe mit der *Nachfolger-Kategorie* im Satz, werden 65,12% der Trainingsmuster korrekt vorhergesagt. Eine zufällige Aktivierung der Ausgabeneurone würden 2,7% korrekte Vorhersagen liefern. Das Netz in der Towsey-Studie lieferte 72% korrekte Vorhersagen auf dem kompletten Trainingskorpus. Das ist zwar etwas besser als die Performanz unseres Netzes, man muss aber bedenken, dass die Trainingsdaten in Towsey et al. (1998) aus stark vereinfachten Schulbuchttexten bestanden und nur in zehn verschiedene Ausgabeklassen unterteilt werden mussten, während das Elman-Netz in der vorliegenden Arbeit Vorhersagen auf echten spontansprachlichen Korpusdaten macht und 37 Ausgabeklassen unterscheidet. Die Berechnung des Abstands zwischen tatsächlicher und erwünschter Ausgabe durch die halbe L_1 -Norm ergibt einen Fehler von 0,407.

Vergleicht man die Ausgabe unseres Elman-Netzes mit der *absoluten bedingten Wahrscheinlichkeitsverteilung* werden 74,41% der Trainingsmuster korrekt vorhergesagt. Hier wird nicht der tatsächliche Nachfolger vorhergesagt, sondern der wahrscheinlichste Nachfolger. Der Zufallswert liegt wieder bei 2,7%. Berechnet man den Abstand zwischen Netzausgabe und Wahrscheinlichkeitsverteilung durch die halbe L_1 -Norm, liegt der Fehler bei 0,298. Dieser Wert ist ein Maß dafür, wie gut das Netz die Auftretenswahrscheinlichkeiten für die *möglichen* Nachfolger gelernt hat. Man sieht, dass der Fehler deutlich kleiner ist, wenn man die Netzausgabe mit einer Wahrscheinlichkeitsverteilung vergleicht anstatt mit den tatsächlichen Nachfolgern. Ein entsprechender Vergleich mit der Towsey-Studie ist nicht möglich, da nicht berichtet wird, wie häufig das Towsey-Netzwerk den wahrscheinlichsten Nachfolger vorhersagt.

Die Ergebnisse zeigen, dass ein Elman-Netz durchaus in der Lage ist, grammatische Regularitäten selbständig aus echten spontansprachlichen Korpora zu lernen. Die Performanz des Netzes bei 74% korrekten Vorhersagen ist sehr ordentlich angesichts des Umstandes, dass das Netz mit 7 Kontextkategorien trainiert, aber mit einer Kontextgröße von 24 Kategorien (absolute bedingte Wahrscheinlichkeitsverteilung) evaluiert wurde. Weitere Verbesserungen der Performanz ließen sich voraussichtlich durch die Kodierung zusätzlicher Informationen, wie z.B. Valenzeigenschaften oder semantischer Merkmale, in den Eingabemustern erzielen.

In diesem Abschnitt wurde die allgemeine Performanz des Netzes auf den Trainingsdaten untersucht, d.h. wie gut das Netz grammatische Regeln erworben hat. Der erfolgreiche Grammatikerwerb ist die Voraussetzung dafür, das Netz hinsichtlich der Verarbeitung koordinierter Strukturen auswerten bzw. mit menschlicher Performanz beim Lesen vergleichen zu können. Ich gehe davon aus, dass das Netz die grammatischen Regularitäten der Trainingsdaten in ausreichendem Maße erworben hat und beschreibe im folgenden Abschnitt die Performanz bezüglich koordinierter Strukturen.

6.4.5.2 Verarbeitung koordinierter Strukturen

Besonders interessant ist der Vergleich von Netzwerkperformanz und menschlicher Verarbeitung. Der Erwerb von grammatischen Regularitäten ist die Voraussetzung dafür. Wie in Kapitel 6.3.4 beschrieben, haben Christiansen und Chater (1999, 2001) eine Korrelation zwischen der Performanz von Elman-Netzen und menschlichem Verhalten für rekursive Strukturen gefunden. Eine höhere Fehlerrate für die Antizipation eines Wortes entspricht dabei einer längeren Lesezeit. In diesem Abschnitt soll untersucht werden, ob eine entsprechende Korrelation für koordinierte Strukturen existiert.

Das Korpusfragment, das die Basis für die verwendeten Trainingssätze darstellt, ist in Kapitel 5 statistisch ausgewertet worden. Die Auswertung zeigt, dass die in Kapitel 3 aufgeführten Präferenzen für gleiche Kategorien (vs. cross-kategorial), für strukturgleiche Konjunkte (Parallel-Struktur-Effekt) sowie die Präferenz bei Ambiguitäten (NP- vs. Satz-Koordination) in den Trainingsdaten in Form von relativen Häufigkeiten zu finden sind. Es stellt sich nun die Frage, ob das Netz auf Basis dieser Eingabedaten ebenfalls die entsprechenden Präferenzen in Form von Fehlerraten entwickelt hat.

In den folgenden Unterabschnitten gehe ich auf die einzelnen Verarbeitungseffekte ein. Es werden jeweils die relevanten Muster aus den Trainingsdaten ausgewertet. Als Vergleichsmaß (Gold-Standard) fungiert die Nachfolger-Kategorie im Satz (siehe Abschnitt 6.4.4.1).

Cross-kategoriale Koordination: Frazier et al. (2000) haben in einem Lesezeit-Experiment zum Englischen beobachtet, dass die Verknüpfung von gleichen syntaktischen Kategorien, wie in (24a), schneller verarbeitet wird als die Koordination von ungleichen syntaktischen Kategorien (cross-kategoriale Koordination), wie in (24b). Entsprechend wird das zweite Konjunkt [*AdvP* carefully] in (24a) schneller gelesen als die gleiche Phrase in (24c), welche unabhängig von Koordination verarbeitet werden muss. Im Gegensatz dazu wird das cross-kategoriale zweite Konjunkt [*PP* with great care] in (24b) nicht schneller gelesen als in (24d) (siehe Kapitel 3, Abschnitt 3.2.3.1).

- (24) a. John walked [*AdvP* slowly] and [*AdvP* carefully], avoiding the broken glass.
- b. John walked [*AdvP* slowly] and [*PP* with great care], avoiding the broken glass.
- c. John walked [*AdvP* carefully], avoiding the broken glass.
- d. John walked [*PP* with great care], avoiding the broken glass.

In Kapitel 5 dieser Arbeit wurde gezeigt, dass die Präferenz für die Verknüpfung gleicher Kategorien auch in englischen spontansprachlichen Korpusdaten zu finden ist. Im untersuchten Korpusfragment werden zu 98% gleiche Kategorien verknüpft und nur in 2% der Fälle sind es ungleiche syntaktische Kategorien. Diese Präferenz ist so stark, dass das Korpus insgesamt nur vier Vorkommen von cross-kategorialer Koordination enthält. Da das Netz mit genau diesen Korpusdaten trainiert wurde, ist eine Auswertung hinsichtlich der Satzantizipation von cross-kategorialer Koordination praktisch unmöglich. Mit der Auswertung von vier Vorkommen lassen sich keine statistisch aussagekräftigen Ergebnisse erzielen. Ich konzentriere mich daher im Folgenden auf die Auswertung des allgemeineren Parallel-Struktur-Effekts.

Parallel-Struktur-Effekt: In Lesezeit-Experimenten zum Englischen wurde ein Parallel-Struktur-Effekt gefunden, welcher besagt, dass das zweite Konjunkt in einer koordinierten Struktur schneller verarbeitet wird, wenn es zum ersten Konjunkt strukturgleich ist (Frazier et al., 2000). Dies betrifft auch den inneren Aufbau einer Phrase (siehe Kapitel 3, Abschnitt 3.2.3.2). Beispielsweise wird in (25a) das zweite Konjunkt [*NP a tall woman*] schneller gelesen als in (25b).

(25) a. Hilda noticed [*NP a strange man*] and [*NP a tall woman*].

b. Hilda noticed [*NP a man*] and [*NP a tall woman*].

In Kapitel 5 dieser Arbeit konnte ich zeigen, dass dieser Effekt auch in englischen spontansprachlichen Korpora in Form von relativen Häufigkeiten zu finden ist. Das untersuchte Korpus enthält 33% parallele koordinierte Strukturen, deren Konjunkte im inneren Phrasenaufbau einschließlich der Wortarten strukturidentisch sind. Der Vergleich mit einem Zufallsdatensatz zeigt, dass zwei koordinierte Phrasen signifikant häufiger strukturgleich sind als zwei Phrasen, die nicht miteinander koordiniert sind.

Da das Netz mit genau diesen Korpusdaten trainiert wurde, stellt sich nun die Frage, ob das Netz auf Basis dieser Eingabedaten ebenfalls eine Präferenz für parallele koordinierte Strukturen entwickelt. In diesem Falle sollte das Netz bei der Antizipation eines parallelen zweiten Konjunks eine geringere Fehlerrate aufweisen als bei der Antizipation eines nicht-parallelen zweiten Konjunks.

Bei der Evaluation dieser Hypothese gibt es das Problem, dass das Korpus im Gegensatz zum experimentellen Setup keine Satzpaare enthält, die bis auf eine kleine Abweichung im ersten Konjunkt identisch sind, wie z.B. in (25a,b). Es ist daher nicht möglich, die Netzwerkperformanz für parallele und nicht-parallele zweite Konjunkte *direkt* miteinander zu vergleichen.

Die Lösung besteht darin, in parallelen Konstruktionen die Netzwerkperformanz für das erste und für das zweite Konjunkt miteinander zu vergleichen. Beide Konjunkte bestehen aus identischen Kategorien-Sequenzen und befinden sich in derselben syntaktischen Umgebung. Der wesentliche Unterschied in der Verarbeitung der beiden Konjunkte besteht darin, dass das

erste Konjunkt wie jede andere Phrase konventionell vorhergesagt werden muss, während das zweite Konjunkt von der Redundanz mit dem ersten Konjunkt profitieren könnte. Ein Vergleich dieser Art entspricht dem Vorgehen im Lesezeit-Experiment zur cross-kategorialen Koordination, wie in (24) dargestellt (Frazier et al., 2000). Dort wird ebenfalls eine Phrase, die als zweites Konjunkt fungiert, mit der gleichen Phrase verglichen, die unabhängig von Koordination, aber in derselben syntaktischen Umgebung vorkommt.

Hypothese (Parallel-Struktur-Effekt): In parallelen koordinierten Strukturen sollte das Netz bei der Antizipation des zweiten Konjunks eine geringere Fehlerrate aufweisen als bei der Antizipation des ersten Konjunks.

Für die Evaluation des Parallel-Struktur-Effekts wurden alle parallelen koordinierten Konstruktionen aus dem Trainingskorpus mit zwei Konjunkten ausgewertet. Zwei Korpusbeispiele sind in (26a) und (26b) dargestellt. Dabei wurde die Netzausgabe mit der Nachfolger-Kategorie im Satz verglichen. Als Fehlermaß diente die halbe L_1 -Norm. Die Fehlerrate für ein Konjunkt wird als Durchschnittswert über die Wörter des Konjunks berechnet.

- (26) a. I just realized I have got whole bunch of [NP meetings] and [NP seminars] through the week
- b. [PP on the twenty sixth] and [PP on the twenty seventh] I am busy all day

Die Auswertung ergab bei der Antizipation der ersten Konjunkte eine durchschnittliche Fehlerrate von 0,47. Die durchschnittliche Fehlerrate bei der Antizipation der zweiten Konjunkte beträgt 0,17. Dieser Unterschied wurde statistisch ausgewertet. Ein gepaarter T-Test zeigt, dass dieser Unterschied signifikant ist ($t(77) = 6,030$; $p < 0,001$). Die obige Hypothese zum Parallel-Struktur-Effekt kann somit bestätigt werden.

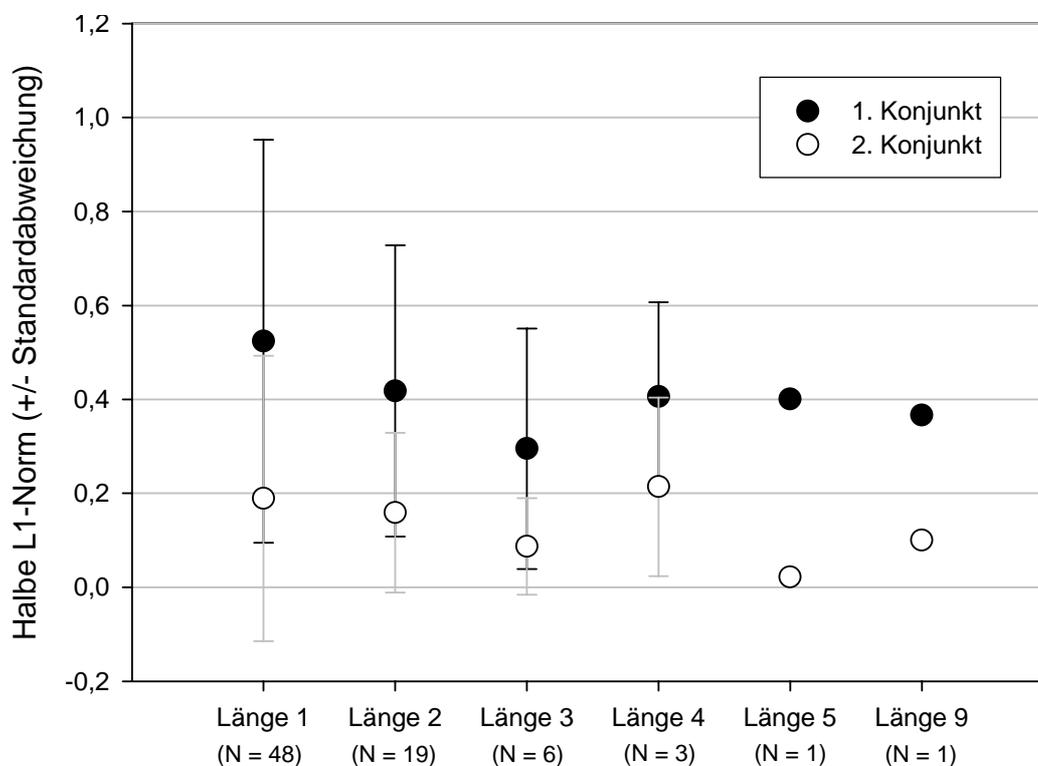


Abbildung 6.9: Fehler des Netzes für parallele koordinierte Strukturen

Einflussfaktor Konjunktlänge: In Kapitel 5 haben wir gesehen, dass der Parallel-Struktur-Effekt, zumindest in Korpora, längenabhängig ist. Parallele Strukturen kommen im Korpus bis zu einer Konjunktlänge von 4 Wörtern vor. Längere Konjunkte sind nur selten parallel.

Es stellt sich nun die Frage, ob der beobachtete Parallel-Struktur-Effekt bei der Satzantizipation auf bestimmte Längen beschränkt ist. Es wäre denkbar, dass sich dieser Effekt beispielweise nur für Konjunkte der Länge 1 beobachten lässt und nicht mehr bei längeren Konjunkten, da die Struktur des ersten Konjunks nicht lange genug aktiviert gehalten werden kann. Um dies zu untersuchen, wurde der Fehler des Netzes bei parallelen Strukturen für die verschiedenen Konjunktlängen ausgewertet. Das Ergebnis ist in Abbildung 6.9 dargestellt.

Für alle Konjunktlängen, die bei parallelen Strukturen im Korpus vorkommen, ist die Fehlerrate für das zweite Konjunkt geringer als für das erste Konjunkt. Für Konjunkte der Länge 1 liegt der durchschnittliche Fehler für die ersten Konjunkte bei 0,52 bzw. bei 0,19 für die zweiten Konjunkte. Dieser Unterschied ist signifikant ($t(47) = 4,533$; $p < 0,001$). Für Konjunkte der Länge 2 beträgt der durchschnittliche Fehler für die ersten Konjunkte 0,42 bzw. 0,16 für die zweiten Konjunkte. Dieser Unterschied ist ebenfalls signifikant ($t(18) = 3,236$; $p = 0,005$). Für die längeren Konjunkte wurden keine separaten T-Tests berechnet, da die Anzahl der Vorkommen zu klein ist. Man sieht jedoch, dass auch die anderen Konjunktlängen (sogar das Vorkommen mit Länge 9) dasselbe Muster aufweisen. Der beobachtete Parallel-Struktur-Effekt bei der Satzantizipation ist also nicht auf bestimmte Längen beschränkt.

Einflussfaktor Satzkontext: Es ist ein bekanntes Phänomen, dass der Satzkontext die Lesegeschwindigkeit beeinflusst. Wörter, die durch den vorausgehenden Kontext stark vorhersagbar sind, werden schneller erkannt als Wörter, die in einem gegebenen Kontext schlechter vorhersagbar sind ('predictability effect', Ehrlich und Rayner, 1981). Tulving et al. (1964) haben herausgefunden, dass die Worterkennung mit zunehmender Größe des Satzkontexts immer leichter wird. Das wird im Allgemeinen darauf zurückgeführt, dass beim Satzlesen mit jedem Wort die Anzahl der möglichen Fortsetzungen immer weiter eingeschränkt wird und so die Vorhersagbarkeit gegen Ende des Satzes immer größer wird.

Es wäre daher denkbar, dass sich der Parallel-Struktur-Effekt anders verhält, wenn die koordinierte Struktur am Satzanfang steht, d.h. wenn sie ohne oder mit wenig einschränkendem Kontext vorkommt, als wenn sie am Satzende steht. In diesem Zusammenhang stellt sich die Frage, ob das trainierte Elman-Netz ebenfalls einen Satzkontext-Effekt ('predictability effect') aufweist, d.h. bei der Antizipation am Satzende weniger Fehler macht als am Satzanfang. Für die Antizipation von Wörtern kann der Einfluss des Kontexts sowohl semantischer als auch syntaktischer Natur sein. Bei der Antizipation von Kategorien (Wortarten) ist jedoch im Wesentlichen der vorangehende syntaktische Kontext für die Vorhersagbarkeit relevant. Dieser syntaktische Kontext steht dem hier verwendeten Elman-Netz auf zweierlei Arten zur Verfügung: (a) durch die 7 Kontext-Kategorien im Eingabevektor, (b) durch die Kontextschicht der

Elman-Architektur. Sollte das Netz einen syntaktischen Kontexteffekt zeigen, ist der beobachtete Parallel-Struktur-Effekt bei der Satzantizipation möglicherweise durch die Größe des Kontexts, die der Koordination im Satz vorangeht, beeinflusst.

Um diesen möglichen Einflussfaktor zu untersuchen, habe ich für jedes parallele Vorkommen, das im Trainingskorpus vorkommt, die Kontextgröße kodiert, d.h. die Anzahl der Wörter vor der Koordination. Im Beispielsatz (26a) ist das ein Kontext von 9 Wörtern, in (26b) ein Kontext von 0 Wörtern. Der Datensatz wurde aufgeteilt in alle Vorkommen mit kleinem Kontext, d.h. ein Kontext ≤ 4 ($n = 39$), und in Vorkommen mit großem Kontext, d.h. ein Kontext > 4 ($n = 39$). Bei kleinem Kontext liegt der durchschnittliche Fehler für die ersten Konjunkte bei 0,63 bzw. bei 0,22 für die zweiten Konjunkte. Dieser Unterschied ist signifikant ($t(38) = 5,667$; $p < 0,001$). Bei großem Kontext beträgt der durchschnittliche Fehler für die ersten Konjunkte 0,32 bzw. 0,13 für die zweiten Konjunkte. Dieser Unterschied ist ebenfalls signifikant ($t(38) = 2,955$; $p = 0,005$).

Generell lässt sich feststellen, dass das Netz bei kleinem Kontext, also beispielsweise am Satzanfang, mehr Fehler bei der Antizipation macht als gegen Ende des Satzes, d.h. bei großem Kontext (0,63 versus 0,32 für die ersten Konjunkte; 0,22 versus 0,13 für die zweiten Konjunkte). Das deutet daraufhin, dass das trainierte Elman-Netzwerk ebenfalls einen syntaktischen Kontexteffekt aufweist, wie man ihn auch beim menschlichen Lesen beobachtet. Dennoch ist sowohl bei kleinem als auch bei großem Kontext der Fehler des Netzes für die zweiten Konjunkte signifikant geringer als der Fehler für die erste Konjunkte. Der Parallel-Struktur-Effekt bei der Satzantizipation ist demnach nicht nur bei einer bestimmten Kontextgröße zu beobachten.

Einflussfaktor Satzposition: Wie oben bereits erwähnt, wird die Lesegeschwindigkeit unter anderem durch die Größe des Satzkontexts beeinflusst. Je größer der Kontext, desto leichter fällt die Worterkennung. Wir haben im letzten Abschnitt gesehen, dass auch das trainierte Elman-Netzwerk einen syntaktischen Kontexteffekt zeigt, nämlich größere Fehlerraten für koordinierte Strukturen am Satzanfang im Vergleich zum Satzende. Es ist daher nicht völlig auszuschließen, dass die geringeren Fehlerraten für die zweiten Konjunkte unter

Umständen darauf zurückzuführen sind, dass das zweite Konjunkt in der linearen Abfolge des Satzes grundsätzlich nach dem ersten Konjunkt steht. In diesem Falle wäre der beobachtete Parallel-Struktur-Effekt bei der Satzantizipation nichts anderes als ein syntaktischer Kontexteffekt.

Um diesen Einflussfaktor genauer zu untersuchen, habe ich eine *lineare Regressionsanalyse* durchgeführt, um den Zusammenhang zwischen der Fehlerrate des Netzes und der Satzposition bzw. der Konjunkt-Information (erstes vs. zweites Konjunkt) zu untersuchen. Hierfür wurde die mittlere Position im Satz für jedes Konjunkt in allen parallelen Konstruktionen bestimmt. In (26a) beispielsweise steht das erste Konjunkt [*meetings*] an Position 10 im Satz, das zweite Konjunkt [*seminars*] an Position 12. In (26b) bestehen die Konjunkte aus mehreren Wörtern. Daher wurde ein Durchschnittswert für die Position der Konjunkte berechnet. Das erste Konjunkt [*on the twenty sixth*] in (26b) besetzt die Positionen 1–4, was einen Durchschnittswert von 2,5 ergibt. Das zweite Konjunkt besetzt die Positionen 6–9, woraus ein mittlerer Wert von 7,5 resultiert.

In der linearen Regressionsanalyse ist die abhängige Variable L_1 die Fehlerrate des Netzes, die aus den Werten der unabhängigen Variablen vorhergesagt werden soll. Die Variable L_1 kodiert den durchschnittlichen Fehler, d.h. genau genommen die halbe L_1 -Norm, für jedes Konjunkt in allen parallelen koordinierten Strukturen aus dem Trainingskorpus. Die Variable L_1 kann Werte zwischen 0 und 1 annehmen. Die unabhängigen Variablen sind: (a) die Satzposition *POS* und (b) die Konjunkt-Information *KONJ* (erstes vs. zweites Konjunkt). Die Variable *POS* kodiert, wie oben ausgeführt, die mittlere Satzposition für jedes Konjunkt in allen parallelen Strukturen und hat Werte zwischen 1 und 20. Die Variable *KONJ* kodiert, ob es sich um das erste oder um das zweite Konjunkt handelt und hat entweder den Wert 1 oder 2.

Die Einbeziehung der unabhängigen Variablen *POS* und *KONJ* wurde schrittweise durchgeführt. In einem ersten Schritt wurde der Einfluss der Satzposition *POS* auf die Fehlerrate L_1 untersucht. Die Aufgabe der linearen Regression ist dabei, die Parameter a und b in der Gleichung (6.9) abzuschätzen, wobei a ein konstanter Term ist und b den Gewichtungsfaktor der Variable *POS* darstellt. Das Ergebnis der Parameterabschätzung (Regressionsmodell 1) ist in (6.10) dargestellt.

$$L_1 = a + b \cdot POS \quad (6.9)$$

$$L_1 = 0,578 - 0,033 \cdot POS \quad (6.10)$$

Wie man in der Regressionsgleichung (6.10) sehen kann, wird die Fehlerrate L_1 immer kleiner, je größer der Wert für die Variable POS wird, d.h. je größer der Satzkontext ist. Das Modell 1 in (6.10) hat eine Anpassungsgüte von $R^2 = 0,192$ und stellt eine signifikante Verbesserung gegenüber einem Modell dar, das nur aus einem konstanten Term (= Mittelwert der Fehlerrate L_1) besteht (Änderung in $F(1, 154) = 36,686$; $p < 0,001$). Der Faktor *Satzposition* hat somit einen signifikanten Einfluss auf die Fehlerraten des Netzes.

In einem zweiten Schritt wurde die Konjunkt-Information *KONJ* (erstes vs. zweites Konjunkt) als zweite unabhängige Variable hinzugenommen. Bei der Regressionsanalyse mit zwei unabhängigen Variablen müssen die Parameter a , b_1 und b_2 in der Gleichung (6.11) geschätzt werden, wobei a wieder ein konstanter Term ist und b_1 und b_2 die Gewichtungsfaktoren der Variablen POS und $KONJ$ darstellen. Das Ergebnis der Parameterabschätzung (Regressionsmodell 2) ist in (6.12) dargestellt.

$$L_1 = a + b_1 \cdot POS + b_2 \cdot KONJ \quad (6.11)$$

$$L_1 = 0,871 - 0,026 \cdot POS - 0,231 \cdot KONJ \quad (6.12)$$

Man sieht in der Regressionsgleichung (6.12), dass die Fehlerrate L_1 mit zunehmendem POS immer kleiner wird. L_1 wird jedoch außerdem kleiner, wenn die Variable $KONJ$ den Wert 2 (für zweites Konjunkt) anstatt den Wert 1 (für erstes Konjunkt) annimmt. Das Modell 2 in (6.12) hat eine Anpassungsgüte von $R^2 = 0,289$ und stellt eine signifikante Verbesserung gegenüber dem Regressionsmodell 1 in (6.10) dar, in dem nur der Faktor *Satzposition* berücksichtigt wird (Änderung in $F(1, 153) = 20,792$; $p < 0,001$). Der Faktor

Konjunkt-Information hat somit einen zusätzlichen signifikanten Einfluss auf die Fehlerraten des Netzes, der unabhängig vom Einfluss der Satzposition ist.

Zusammenfassend kann man sagen, dass die Position im Satz einen signifikanten Einfluss auf die Fehlerraten des Netzes hat. Ähnlich wie beim menschlichen Lesen zeigt das Netz einen syntaktischen Kontexteffekt. Je größer der Kontext, desto geringer sind die Fehlerraten des Netzes. Weiterhin gibt es jedoch einen zusätzlichen signifikanten Einfluss der Koordination. Im zweiten Konjunkt sind die Fehlerraten geringer als im ersten Konjunkt. Wie die Regressionsanalyse gezeigt hat, ist dieser Effekt nicht ausschließlich durch die Position im Satz erklärbar. Es handelt sich um einen unabhängigen Parallel-Struktur-Effekt bei der Satzantizipation.

Präferenzen bei Ambiguitäten (NP- vs. Satz-Koordination): In Lesezeit-Experimenten zum Englischen wurde gezeigt, dass die NP-Koordination in ambigen Konstruktionen gegenüber der Satz-Koordination präferiert wird (Frazier, 1979). Die ambige Nominalphrase [*her sister*] in (27) wird als Teil einer koordinierten Objekt-NP interpretiert (27a) und nicht als Beginn (Subjekt) eines neuen Satzes (27b). Frazier fand einen Garden-Path-Effekt, d.h. längere Lesezeiten, auf dem desambiguierenden Segment '*laughed*' bei der Satzkoordination (siehe Kapitel 3, Abschnitt 3.3.1.1).

(27) a. Peter kissed [_{NP} Mary] and [_{NP} *her sister*] too.

b. [_S Peter kissed Mary] and [_S *her sister* laughed].

In Kapitel 5 dieser Arbeit wurde gezeigt, dass diese Präferenz für NP-Koordination auch in englischen spontansprachlichen Korpusdaten zu finden ist. Als Basis für die Korpusauswertung dienten CD-6 und CD-13 der TüBa-E-Baumbank. Sätze der Form "NP_{Subj}...Verb...NP_{Obj} and..." wurden zu 69% als NP-Koordination fortgeführt und nur zu 31% als Satz-Koordination. Das Netz wurde jedoch nur mit den Sätzen aus CD-13 trainiert, so dass für die Auswertung des Netzes die relativen Häufigkeiten in CD-13 relevant sind. Doch auch dort sind die Verhältnisse ganz ähnlich: 66% NP-Koordinationen ($n = 14$) und 33% Satz-Koordinationen ($n = 7$).

Es stellt sich nun die Frage, ob das Netz auf Basis der Daten in CD-13 mit 21 relevanten Vorkommen ebenfalls eine Präferenz für NP-Koordination entwickelt hat und bei Satz-Koordination an der desambiguierenden Stelle ins Stolpern gerät, d.h. mehr Fehler macht. Es wird die folgende Hypothese aufgestellt.

Hypothese 1 (NP- vs. Satz-Koordination): In ambigen Konstruktionen sollte das Netz an der desambiguierenden Stelle weniger Fehler bei der Antizipation der NP-Koordination machen als bei der Antizipation der Satz-Koordination.

Bevor die 21 relevanten Vorkommen in den Trainingsdaten ausgewertet werden können, muss erst einmal ermittelt werden, wo in diesen Sätzen die desambiguierende Stelle auf Basis von Kategoriensequenzen ist. Eine genauere Inspektion dieser Sätze zeigt, dass bereits die Kategorie nach der Konjunktion ‘and’ Aufschluss darüber gibt, ob der Satz als NP-Koordination oder als Satz-Koordination weitergeführt wird. Folgt auf das ‘and’ die Kategorie *NN* (Nomen_{sg}), *NNS* (Nomen_{pl}), *NP* (Eigennamen_{sg}), *JJ* (Adjektiv) oder *CD* (Kardinalzahl) wird der Satz als NP-Koordination weitergeführt (siehe z.B. (28a,b)). Folgt die Kategorie *PP* (Personalpronomen) oder *RB* (Adverb) wird der Satz als Satz-Koordination weitergeführt (siehe z.B. (28e)). Nur die Kategorie *DT* (Determinierer) kann in diesen Sätzen am Beginn beider Weiterführungen stehen (28c,d).

- (28) a. I just realized I have got whole bunch of [*NP* meetings] *and* [*NP* seminars_*NNS*] through the week
- b. that is [*NP* an hour for lunch] *and* [*NP* two_*CD* hours for the meeting]
- c. Monday looks pretty good except for I have [*NP* a early morning meeting] *and* [*NP* a_*DT* lunch meeting]
- d. [*S* that way we have got two hours of lunch break] *and* [*S* that_*DT* will be wonderful]
- e. [*S* it might turn out to be a fun meeting] *and* [*S* we_*PP* will see you on Saturday]

In 17 der 21 Vorkommen ist die Kategorie nach der Konjunktion die kritische (desambiguierende) Stelle, an der sich entscheidet, ob eine NP- oder eine Satz-Fortführung vorliegt. In den verbleibenden vier Fällen folgt auf die Konjunktion die Kategorie *DT* (Determinierer). Hier ist die kritische Stelle die Kategorie, die auf den Determinierer folgt, d.h. zwei Positionen nach der Konjunktion. Im Folgenden schauen wir uns die Vorhersagen des Netzes für die jeweilige desambiguierende Stelle an.

Für jede der 21 Konstruktionen wurde die Vorhersage des Netzes (präferierte Kategorie) für die desambiguierende Stelle ermittelt. Die Vorhersage wurde als korrekt gewertet, wenn die Netzausgabe (präferierte Kategorie) mit der entsprechenden Nachfolger-Kategorie im Satz übereinstimmt. Ansonsten wurde die Vorhersage als falsch gewertet. Die Auswertung ergibt bei der NP-Koordination 85,7% ($n = 12$) korrekt vorhergesagte Fälle und 14,3% ($n = 2$) falsche Vorhersagen. Bei der Satz-Koordination werden 57,1% ($n = 4$) der Fälle korrekt vorhergesagt, 42,9% ($n = 3$) der Fälle werden falsch vorhergesagt. Wir finden einen quantitativen Unterschied zwischen den Vorhersagen für NP- bzw. Satz-Koordination. Die NP-Koordination wird an der desambiguierenden Stelle häufiger korrekt vorhergesagt als die Satz-Koordination.

Für die statistische Auswertung ist es aufgrund der geringen Fallzahlen nicht möglich, den Chi-Quadrat-Test anzuwenden. Ich wende stattdessen den *Fisher Exact Probability Test* ("Exakter Test auf Unabhängigkeit nach Fisher") an, der auch bei einer geringen Anzahl von Beobachtungen verwendet werden kann. Obwohl die Tendenz, d.h. weniger Fehler bei der NP-Koordination, auf den ersten Blick recht deutlich scheint, ergibt der Fisher-Test keinen signifikanten Unterschied zwischen den Vorhersagen bei NP- und Satz-Koordination ($p = 0,18$). Bei den vorliegenden Verhältnissen ist vermutlich eine größere Anzahl an Fällen notwendig, um einen signifikanten Unterschied für die desambiguierende Stelle zu erhalten. Mit den vorliegenden Fallzahlen erhalten wir keine ausreichende Evidenz für Hypothese 1 (NP- vs. Satz-Koordination).

In einer weiteren Auswertung wird nicht nur die Vorhersage an der kritischen Stelle betrachtet, sondern auch die Vorhersage für die Kategorie nach der kritischen Stelle (Spillover-Segment). In Lesezeit-Experimenten kann man beobachten, dass in ambigen Sätzen (Garden-Path-Sätze) nicht nur auf dem

desambiguierenden Segment längere Lesezeiten auftreten, sondern dass sich diese auch noch auf das nachfolgende Segment auswirken (Spillover-Effekt). Hypothese 1 wird entsprechend modifiziert und als Hypothese 2 reformuliert.

Hypothese 2 (NP- vs. Satz-Koordination): In ambigen Konstruktionen sollte das Netz auf dem desambiguierenden sowie dem Spillover-Segment weniger Fehler bei der Antizipation der NP-Koordination machen als bei der Antizipation der Satz-Koordination.

In (29) sind die Beispielsätze aus (28) nochmals dargestellt. Die desambiguierende Stelle und das Spillover-Segment sind kursiv markiert. In (29a) ist *'seminars'* (Kategorie *NNS*) die desambiguierende Stelle, *'through'* (Kategorie *IN*) ist das Spillover-Segment. In (29c) ist *'lunch'* (Kategorie *NN*) die desambiguierende Stelle, da die Kategorie *DT* nach der Konjunktion ambig ist. Das Spillover-Segment ist *'meeting'* (Kategorie *NN*), etc.

- (29) a. ... and [*NP seminars_NNS*] *through_IN* the week
 b. ... and [*NP two_CD hours_NNS* for the meeting]
 c. ... and [*NP a_DT lunch_NN meeting_NN*]
 d. ... and [*S that_DT will_MD be_VB* wonderful]
 e. ... and [*S we_PP will_MD see you on Saturday*]

Für jede der 21 Konstruktionen wurde die Vorhersage des Netzes (präferierte Kategorie) für die desambiguierende Stelle sowie für die Spillover-Kategorie ermittelt. Die Vorhersage wurde als korrekt gewertet, wenn sowohl für die desambiguierende als auch für die Spillover-Kategorie die Netzausgabe (präferierte Kategorie) mit der entsprechenden Nachfolger-Kategorie im Satz übereinstimmt. Ansonsten wurde die Vorhersage als falsch gewertet. In drei Sätzen existiert keine Spillover-Kategorie, weil der Satz nach der desambiguierenden Stelle endet. In diesen Fällen war die desambiguierende Kategorie für die Korrektheit entscheidend.

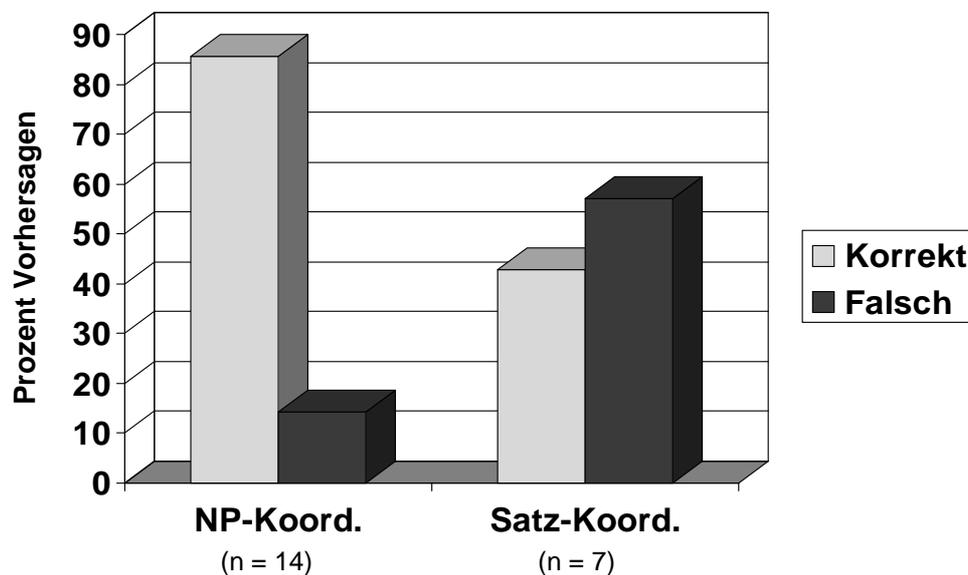


Abbildung 6.10: Vorhersagen des Netzes für NP- und Satz-Koordination: Desambiguierende plus Spillover-Kategorie.

Die Auswertung ergibt bei der NP-Koordination 85,7% ($n = 12$) korrekt vorhergesagte Fälle und 14,3% ($n = 2$) falsche Vorhersagen. Bei der Satz-Koordination werden 42,9% ($n = 3$) der Fälle korrekt vorhergesagt, 57,1% ($n = 4$) der Fälle werden falsch vorhergesagt. Die Ergebnisse sind in Abbildung 6.10 dargestellt. Wir finden ein deutlicheres Ergebnis vor, wenn zur desambiguierenden Kategorie noch die Spillover-Kategorie hinzugenommen wird. Bei der NP-Koordination werden beide Kategorien häufiger korrekt vorhergesagt als bei der Satz-Koordination.

Der Fisher-Test ergibt einen Unterschied zwischen den Vorhersagen bei NP- und Satz-Koordination mit einer Irrtumswahrscheinlichkeit von 6% ($p = 0,064$). Zieht man die desambiguierende Stelle plus das Spillover-Segment für die Auswertung heran, erhält man einen Effekt, d.h. eine Präferenz für NP-Koordination, der als marginal signifikant bezeichnet werden kann. Wir erhalten somit Evidenz für Hypothese 2. Auch hier wäre es hilfreich, mehr

Daten zur Verfügung zu haben, um ein möglicherweise deutlicheres Ergebnis zu bekommen.

6.5 Zusammenfassung und Diskussion

In diesem Kapitel wurde die syntaktische Verarbeitung koordinierter Strukturen simuliert. Die Herausforderung bestand darin, ein neuronales Netzwerk zu trainieren, das selbständig strukturelle Repräsentationen und Parsingmechanismen für Sätze entwickelt. Neuronale Netze lernen aus Erfahrung, ohne die vorherige Integration von linguistischen Vorannahmen. Für das Training habe ich die Satzantizipations-Aufgabe auf Kategorienebene verwendet. Dabei soll für eine gegebene Sequenz von Kategorien (Wortarten) die nachfolgende Kategorie vorhergesagt werden. Syntaktische Strukturen und Mechanismen werden somit auf inkrementelle Weise implizit aus Kategoriensequenzen gelernt.

Es handelt sich um ein überwacht, aber dennoch natürliches Lernszenario, da der Zielreiz, d.h. das nächste Element im Satz, automatisch bei der Sprachrezeption zur Verfügung steht. Die Satzantizipations-Aufgabe liefert außerdem bei der Verletzung von Erwartungshaltungen indirekte negative Evidenz für den Spracherwerb, wodurch das Lernen einer natürlich-sprachlichen Grammatik auch ohne die Annahme von angeborenen Prinzipien möglich ist. Das *Baker's Paradox* könnte auf diese Weise aufgelöst werden (siehe Fußnote 4).

Für das Training wurde ein Elman-Netz verwendet, eine partiell rekurrente Netzwerkarchitektur, die aufgrund ihrer zusätzlichen Kontextschicht besonders gut für das Lernen von zeitveränderlichen Sequenzen geeignet ist. Diese Architektur ermöglicht es, strukturelle Zusammenhänge aus Sequenzen zu lernen, deren Elemente nicht zeitgleich präsentiert werden. Dadurch wird auch ein Lösungsansatz für das *Binding Problem* der Sprache (siehe Kapitel 1) bereitgestellt.

In diesem Kapitel habe ich zwei Fragestellungen untersucht. Einerseits ging es darum, herauszufinden, ob die Satzantizipations-Aufgabe auch mit echten spontansprachlichen Korpusdaten erfolgreich gelernt werden kann. Bislang war diese Aufgabe nur mit stark vereinfachten oder künstlichen Trainingsdaten durchgeführt worden. Die zweite Fragestellung betraf das Verhalten des Netzes

hinsichtlich koordinierter Strukturen. Es wurde untersucht, ob die Performanz des Netzes mit der menschlichen Leseverarbeitung korreliert. Eine solche Korrelation wurde bisher nur anhand rekursiver Strukturen in künstlichen Sprachen untersucht und gefunden. In diesem Zusammenhang ist es besonders sinnvoll, spontansprachliche Daten anstatt geschriebener Sprache für das Training zu verwenden, da die spontansprachlichen Daten wesentlich besser mit Verarbeitungseffekten korrelieren (siehe Steiner, 2007a, 2009).

Zur Klärung dieser beiden Fragestellungen wurde ein Elman-Netz mit einem Korpusfragment der *Englischen Verbmobil Baumbank* trainiert, welches bereits in Kapitel 5 statistisch ausgewertet wurde. Dieses Korpus enthält echte spontansprachliche Dialoge mit wesentlich längeren Sätzen und komplexeren Strukturen als bislang in der Literatur verwendet. Außerdem habe ich eine deutlich größere Anzahl an Ausgabeklassen verwendet. Durch diese Charakteristika wird die Satzantizipations-Aufgabe erheblich komplexer. Ich habe mich daher für ein größeres Kontextfenster in den Eingabedaten entschieden.

Die Ergebnisse zeigen, dass ein Elman-Netz durchaus in der Lage ist, grammatische Regularitäten selbständig aus echten spontansprachlichen Korpora zu lernen. Die Performanz des Netzes bei 74% korrekten Vorhersagen ist sehr ordentlich angesichts des Umstandes, dass das Netz mit sieben Kontextkategorien trainiert, aber mit einer Kontextgröße von 24 Kategorien (absolute bedingte Wahrscheinlichkeitsverteilung) evaluiert wurde. Weitere Verbesserungen der Performanz würden sich voraussichtlich durch die Kodierung zusätzlicher Informationen, wie z.B. Valenzeigenschaften oder semantischer Merkmale, in den Eingabemustern erzielen lassen. Desweiteren deuten die Ergebnisse darauf hin, dass das Netz eine Wahrscheinlichkeitsverteilung gelernt hat. Der erfolgreiche Grammatikerwerb ist die Voraussetzung dafür, das Netz hinsichtlich der Verarbeitung koordinierter Strukturen auswerten bzw. mit menschlicher Performanz beim Lesen vergleichen zu können. Aus den Ergebnissen schließe ich, dass das Netz die grammatischen Regularitäten der Trainingsdaten in ausreichendem Maße erworben hat.

Der Vergleich zwischen Netzwerkperformanz und der menschlichen Leseverarbeitung wurde anhand des Parallel-Struktur-Effekts sowie der Präferenz für NP-Koordination (vs. Satz-Koordination) bei Ambiguitäten durchgeführt.

Beide Effekte wurden in Lesezeit-Experimenten beobachtet. Die cross-kategoriale Koordination konnte mangels ausreichender Vorkommen nicht ausgewertet werden.

Die Auswertung des Parallel-Struktur-Effekts ergab eine Korrelation zwischen der Netzwerkperformanz und der menschlichen Leseverarbeitung. Beim Lesen wird ein strukturgleiches zweites Konjunkt schneller verarbeitet. Entsprechend macht das Elman-Netz in parallelen koordinierten Strukturen signifikant weniger Fehler bei der Antizipation des zweiten Konjunks als bei der Antizipation des ersten Konjunks. Das Netz zeigt diesen Effekt nicht nur bei Konjunkten einer bestimmten Länge, sondern bei allen im Trainingskorpus vorkommenden Konjunktlängen.

Weiterhin zeigt das Netz einen syntaktischen Kontexteffekt vergleichbar mit der menschlichen Leseverarbeitung. Beim Lesen nimmt die Geschwindigkeit mit zunehmendem Satzkontext zu. Auch im Netzwerk werden die Fehlerraten zum Satzende hin immer kleiner. Die Position im Satz hat einen signifikanten Einfluss auf die Fehlerraten. Es konnte jedoch gezeigt werden, dass die Koordination einen zusätzlichen signifikanten Einfluss auf die Fehlerraten hat, d.h. der Parallel-Struktur-Effekt bei der Satzantizipation ist ein unabhängiger Effekt.

Die Auswertung der Präferenz für NP-Koordination (vs. Satz-Koordination) ergab eine schwache Korrelation zwischen der Netzwerkperformanz und der Leseverarbeitung. Beim Lesen wird die NP-Koordination schneller verarbeitet als die Satz-Koordination. Entsprechend macht das Netzwerk weniger Fehler bei der Antizipation der NP-Koordination als bei der Antizipation der Satz-Koordination. Dieser Effekt ist marginal signifikant, wenn die desambiguierende Kategorie und die nachfolgende Spillover-Kategorie ausgewertet wird. Die Unterschiede zwischen den Fehlerraten bei NP- bzw. Satz-Koordination sind auf den ersten Blick eindeutig. Es sind vermutlich mehr Trainingsdaten erforderlich, um einen robusten Effekt zu erhalten.

Während koordinierte Strukturen eine der größten Fehlerquellen für die maschinelle Sprachverarbeitung darstellen, hat das neuronale Netz interessanterweise keine besonderen Schwierigkeiten bei der Satzantizipation dieser Strukturen. Das Netz zeigt sogar eine Verarbeitungserleichterung bei parallelen koordinierten Strukturen und verhält sich bezüglich der untersuchten

Effekte vergleichbar mit der menschlichen Leseverarbeitung. Interessant ist dabei, dass diese Effekte zu beobachten sind, obwohl das Netz ausschließlich mit Kategoriensequenzen trainiert wurde und keine weiteren linguistischen Vorannahmen integriert wurden. Das Netz lernt lediglich, das nächste Element im Satz zu antizipieren, die beobachteten Verarbeitungseffekte sind emergente Eigenschaften. Insgesamt konnte gezeigt werden, dass ein neuronales Netz in der Lage ist, ausschließlich aufgrund statistischer Muster im sprachlichen Input Verarbeitungspräferenzen zu entwickeln, wie sie auch in Lesezeit-Experimenten beobachtet werden.

Abschließend ist zu bemerken, dass sich auf die vorliegende Arbeit mit neuronalen Netzen in vielfältiger Weise aufbauen ließe. Weitere Simulationen mit einer größeren Datenmenge würden beispielsweise eine bessere Auswertung von linguistisch relevanten Konstruktionen erlauben. Die Variation der Kontextgröße in den Eingabedaten ermöglicht weitere Rückschlüsse auf die Performanz und Generalisierungsfähigkeit des Netzes. Die Integration von zusätzlichen Informationen in den Eingabedaten, wie z.B. Valenzinformation, würde vermutlich die allgemeine Performanz des Netzes verbessern. Besonders interessant wäre auch eine schrittweise Präsentation der Trainingsdaten mit steigender Komplexität der Daten (inkrementelles Lernen) wie in Elman (1991, 1993), da auf diese Weise komplexe Strukturen vermutlich leichter gelernt werden können.

Kapitel 7

Asymmetrische Kongruenz: Ein Verarbeitungseffekt?

In diesem Kapitel wird die Umgebungs-Symmetrie, d.h. das Verhältnis der Konjunkte zu ihrer syntaktischen Umgebung untersucht. Dabei stellt sich die Frage, ob alle Konjunkte den gleichen syntaktischen Status haben oder ob ein Konjunkt einen besonderen Stellenwert einnimmt. Unter diesen Symmetrie-Typ fallen Phänomene wie die (a)symmetrische Subkategorisierung, die (a)symmetrische Extraktion und die (a)symmetrische Kongruenz (siehe Kapitel 2.4). Die asymmetrischen Varianten dieser Phänomene scheinen eine Sonderstellung des ersten Konjunks nahelegen und werden häufig als Argument für eine asymmetrische syntaktische Struktur der Koordination verwendet.

Hinsichtlich der Umgebungs-Symmetrie ist meine Ausgangshypothese, dass die beobachteten Asymmetrien keinen Reflex der Syntax darstellen, sondern auf die Inkrementalität der Satzverarbeitung von links nach rechts zurückzuführen sind. Demnach wäre die Sonderstellung des ersten Konjunks lediglich eine Folge davon ist, dass das erste Konjunkt *zeitlich vor* dem zweiten Konjunkt verarbeitet wird und somit einen geringeren Abstand zu der zuvor verarbeiteten syntaktischen Umgebung aufweist (siehe Kapitel 2.4). In diesem Kapitel wird diese Ausgangshypothese exemplarisch anhand der (a)symmetrischen Kongruenz bei koordinierten Subjekten untersucht. Für eine systematische Untersuchung dieser Fragestellung ist es notwendig, eine Sprache mit einer relativ freien

Wortstellung zu verwenden, so dass ein koordiniertes Subjekt an verschiedenen Positionen im Satz getestet werden kann. Aus diesem Grund werden die folgenden Untersuchungen zum Deutschen durchgeführt (siehe Steiner, 2009).

7.1 Einleitung

Die Kongruenz zwischen einem Subjekt im Plural und dem Verb ist im Allgemeinen unproblematisch und führt zu Plural-Morphologie am Verb (siehe (1)). Das ist unabhängig davon, ob die Subjekt-NP [*drei Schüler*] vor dem finiten Verb im Vorfeld steht, wie in (1a,b), oder nach dem finiten Verb im Mittelfeld, wie in (1c,d).

- (1) a. [*Drei Schüler*]_{PL} stehen_{PL} vor dem Lehrerzimmer.
 b. [*Drei Schüler*]_{PL} *steht_{SG} vor dem Lehrerzimmer.
 c. Vor dem Lehrerzimmer stehen_{PL} [*drei Schüler*]_{PL}.
 d. Vor dem Lehrerzimmer *steht_{SG} [*drei Schüler*]_{PL}.

Weniger eindeutig hingegen ist die Situation bei koordinierten Subjekten. Besteht das Subjekt aus zwei koordinierten Nominalphrasen, wie in (2), werden im Allgemeinen beide Konjunkte für die Herstellung der Kongruenz herangezogen (symmetrische Kongruenz, wie in (2a,c)). In manchen Fällen ist jedoch nur ein Konjunkt für die Kongruenz relevant (asymmetrische Kongruenz, wie in (2d)).

- (2) a. [*Ein Vater*]_{SG} und [*eine Mutter*]_{SG} stehen_{PL} vor dem Lehrerzimmer.
 b. [*Ein Vater*]_{SG} und [*eine Mutter*]_{SG} *steht_{SG} vor dem Lehrerzimmer.
 c. Vor dem Lehrerzimmer stehen_{PL} [*ein Vater*]_{SG} und [*eine Mutter*]_{SG}.
 d. Vor dem Lehrerzimmer steht_{SG} [*ein Vater*]_{SG} und [*eine Mutter*]_{SG}.

Interessanterweise wird asymmetrische Kongruenz nur dann beobachtet, wenn das finite Verb dem Subjekt vorangestellt ist. Steht das Subjekt vor dem finiten Verb ist nur symmetrische Kongruenz möglich (2a), asymmetrische Kongruenz ist dagegen ungrammatisch (2b). Bei V-S-Wortstellung ist sowohl symmetrische als auch asymmetrische Kongruenz möglich (2c,d). Dieses Muster wurde auch für andere Sprachen dokumentiert, u.a. für das marokkanische und libanesische Arabisch (Aoun et al., 1994), für das Russische (Babyonyshev, 1997), für Swahili (Marten, 2005) und für das Polnische (Citko, 2004).

Asymmetrische Kongruenz wird in der Regel mit syntaktischen Mechanismen erklärt (einen Überblick bietet Lorimor, 2007). Munn (2000) beispielsweise zieht die asymmetrische Kongruenz als Evidenz für eine Adjunktionsanalyse heran, in der das erste Konjunkt der Kopf der koordinierten Phrase ist. Aoun et al. (1994) argumentieren, dass die asymmetrische Kongruenz im Arabischen Evidenz für eine sententiale Analyse der Koordination liefert (*Vor dem Lehrerzimmer steht ein Vater und ~~vor dem Lehrerzimmer steht~~ eine Mutter*). Syntaktische Ansätze zur asymmetrischen Kongruenz sind jedoch mit dem Problem konfrontiert, dass unterschiedliche Mechanismen abhängig von der syntaktischen Position des Subjekts angenommen werden müssen. Die Optionalität der asymmetrischen Kongruenz stellt ebenfalls eine Herausforderung für syntaktische Ansätze dar.

Ein alternativer Erklärungsansatz für asymmetrische Kongruenz ergibt sich aus der Sprachverarbeitung. Demnach wäre die Wortstellungsasymmetrie eine Folge der inkrementellen Sprachverarbeitung von links nach rechts. Der entscheidende Faktor ist dabei, welche Information bei der Verarbeitung des finiten Verbs zur Verfügung steht. Befindet sich das Subjekt in präverbaler Position, wie in (2a,b), ist die Pluralität des Subjekts bereits berechnet, wenn der Parser auf das finite Verb trifft. Dies sollte zu symmetrischer Kongruenz führen. Wenn der Parser jedoch in V-S-Wortstellung auf das finite Verb trifft, wie in (2c,d), hat er noch keine Information über das Subjekt zur Verfügung. In diesen Konstruktionen mit postverbalem Subjekt stellt sich die Frage, welche Faktoren einen Einfluss darauf haben, ob beide Konjunkte (2c) oder ob nur eines der Konjunkte (2d) für die Herstellung von Kongruenz herangezogen wird. Beide Optionen scheinen möglich zu sein.¹

¹ Im Rahmen der *Dynamischen Syntax* stellt Marten (2005) einen unabhängig entwickelten Ansatz zur asymmetrischen Kongruenz vor, der auf ähnlichen Grundannahmen basiert. Die

Ich nehme an, dass in der *Sprachproduktion* die Präferenz für eine der beiden Optionen (symmetrische vs. asymmetrische Kongruenz) davon abhängt, ob bei der Verarbeitung des finiten Verbs beide Konjunkte bereits geplant sind. Normalerweise sollten zu diesem Zeitpunkt beide Konjunkte geplant sein, was zu symmetrischer Kongruenz führt. Ist zu diesem Zeitpunkt jedoch nur das erste Konjunkt geplant, hat dies die asymmetrische Kongruenz zur Folge. Diese Situation könnte beispielsweise durch erhöhten Verarbeitungsdruck entstehen.

Ich nehme an, dass beim *Sprachverstehen* die Präferenz davon abhängt, wie leicht es ist, die Verb-Information bei der Verarbeitung des postverbalen Subjekts abzurufen. Dies kann beispielsweise durch die Distanz zwischen dem Verb und dem Subjekt beeinflusst werden. Folgt das Subjekt direkt auf das finite Verb, wie in (3a,c), sollte bei der Verarbeitung der beiden Konjunkte die Verb-Information im Allgemeinen noch zur Verfügung stehen. Dies sollte zu einer Präferenz für symmetrische Kongruenz führen. Befindet sich jedoch zusätzliches sprachliches Material zwischen dem Verb und dem Subjekt (*wieder mal*), wie in (3b,d), entsteht ein erhöhter Verarbeitungsaufwand. Die Verb-Information muss länger im Arbeitsspeicher gehalten werden und steht unter Umständen bei der Verarbeitung des zweiten Konjunks nicht mehr zur Verfügung. In diesen Fällen ist es denkbar, dass nur das erste Konjunkt für die Herstellung von Kongruenz herangezogen wird. Asymmetrische Kongruenz sollte dann akzeptabler oder wahrscheinlicher sein.

- (3) a. Vor dem Lehrerzimmer stehen_{PL} [ein Vater]_{SG} und [eine Mutter]_{SG}.
- b. Vor dem Lehrerzimmer stehen_{PL} *wieder mal* [ein Vater]_{SG} und [eine Mutter]_{SG}.
- c. Vor dem Lehrerzimmer steht_{SG} [ein Vater]_{SG} und [eine Mutter]_{SG}.
- d. Vor dem Lehrerzimmer steht_{SG} *wieder mal* [ein Vater]_{SG} und [eine Mutter]_{SG}.

Optionalität der asymmetrischen Kongruenz in V-S-Wortstellung und die Faktoren, die die Kongruenz beeinflussen, werden in diesem Ansatz jedoch nicht erklärt.

Sowohl bei der Sprachproduktion als auch beim Sprachverstehen sollte daher die Präferenz für symmetrische oder asymmetrische Kongruenz in postverbalen Konstruktionen in hohem Maße vom Verarbeitungsaufwand abhängen. Je höher der Verarbeitungsaufwand, desto lokaler arbeitet das Sprachsystem, d.h. die asymmetrische Kongruenz sollte häufiger auftreten. Entsprechend sollte die asymmetrische Kongruenz bevorzugt in Datentypen zu beobachten sein, die Verarbeitungsmechanismen unmittelbar reflektieren, wie z.B. in gesprochener Sprache oder in Lesezeiten (Online-Daten). In Offline-Datentypen, wie z.B. in Grammatikalitätsurteilen oder geschriebenen Texten, sollte die asymmetrische Kongruenz seltener auftreten. Aus dem beschriebenen Verarbeitungsansatz lassen sich die folgenden Hypothesen ableiten.

Hypothese 1: Die asymmetrische Kongruenz sollte häufiger vorkommen (oder leichter verarbeitet werden), wenn sich das Subjekt in postverbaler Position befindet.

Hypothese 2: Die asymmetrische Kongruenz sollte häufiger vorkommen (oder leichter verarbeitet werden), wenn die Distanz zwischen Verb und postverbalem Subjekt größer ist.

Hypothese 3: Die asymmetrische Kongruenz sollte in Online-Datentypen häufiger vorkommen (oder leichter verarbeitet werden) als in Offline-Datentypen.

Um diese Hypothesen zu testen, untersuche ich vier verschiedene Datentypen im Deutschen: gesprochene und geschriebene Korpusdaten (Sprachproduktion) sowie Grammatikalitätsurteile und Lesezeiten (Sprachverstehen). Gesprochene Korpusdaten und Lesezeiten zählen zu den Online-Datentypen, während geschriebene Korpusdaten und Grammatikalitätsurteile zu den Offline-Daten gerechnet werden können. Online- und Offline-Daten können selbstverständlich nicht streng in zwei Klassen unterteilt werden, es handelt sich dabei vermutlich eher um ein Kontinuum. Im Hinblick auf Hypothese 3 werden daher lokale Vergleiche zwischen den Online- und Offline-Datentypen durchgeführt (eher online vs. eher offline). Entsprechend vergleiche ich gesprochene mit geschriebenen Korpusdaten, sowie Lesezeiten mit Grammatikalitätsurteilen.

7.2 Korpusuntersuchungen

In diesem Abschnitt werden Korpusdaten aus Zeitungstexten (TÜBA-D/Z) und spontansprachlichen Dialogen (TÜBA-D/S) präsentiert. Die Baubank TÜBA-D/Z (Telljohann et al., 2006) umfasst 27.124 Sätze aus *‘die tageszeitung’* (taz). Die Baubank TÜBA-D/S (Hinrichs et al., 2000) besteht aus spontansprachlichen Dialogen zu Terminvereinbarungen und umfasst 38.196 Sätze. Beide Baubanken sind manuell mit morpho-syntaktischen Kategorien (Wortarten) annotiert, mit syntaktischer Phrasenstruktur und grammatischen Funktionen.

Für die Studie wurde ein großer Teil der TÜBA-D/Z-Baubank (14.940 Sätze) und die gesamte TÜBA-D/S-Baubank (38.196 Sätze) nach Subjekten durchsucht, die aus zwei mit *‘und’* koordinierten Nominalphrasen (je im Singular) bestehen. Die Korpusrecherche wurde mit dem Suchwerkzeug *TIGERSearch* (Lezius, 2002) durchgeführt. In die Analyse gingen nur die Fälle ein, bei denen die Kongruenz mit einem der Konjunkte Singular-Morphologie am Verb nach sich zieht und die Kongruenz mit beiden Konjunkten Plural-Morphologie zur Folge hat. In manchen Fällen ist aus semantischen Gründen nur die Singular-Form oder nur die Plural-Form lizenziert. Dies ist beispielsweise der Fall, wenn eines der Konjunkte eine Negation enthält (Singular-Verbform)² oder die Verbsemantik ein Subjekt im Plural erfordert (Plural-Verbform). Solche Vorkommen wurden ausgeschlossen.

Zunächst werden die Daten aus der geschriebenen TÜBA-D/Z-Baubank betrachtet. Einige Beispiele für asymmetrische Kongruenz aus der TÜBA-D/Z sind in (5) aufgeführt. Die Sätze in (5a,b) enthalten jeweils ein Subjekt in postverbaler Position. (5c) enthält ein präverbales Subjekt (siehe auch Zinsmeister (2006) für weitere Korpusbeispiele mit asymmetrischer Kongruenz aus der TÜBA-D/Z). Die einzelnen Konjunkte sind jeweils eingeklammert.

² Das folgende Korpusbeispiel wurde von der Studie ausgeschlossen, da aufgrund der Negation nur die Singular-Verbform möglich ist.

- (4) ... [die Wirtschaft] und [nicht die Politik] sei für den Mißbrauch bei sozialversicherungsfreien Beschäftigungsverhältnissen verantwortlich.

Subjektposition	Kongruenz	V2 + VEnd (ungefiltert)	V2 + VEnd (gefiltert)
Präverbal	symmetrisch	132 (98,51%)	111 (100%)
Präverbal	asymmetrisch	2 (1,49%)	0 (0%)
<i>Insgesamt</i>		134	111
Postverbal	symmetrisch	62 (84,93%)	53 (93,0%)
Postverbal	asymmetrisch	11 (15,07%)	4 (7,0%)
<i>Insgesamt</i>		73	57

Tabelle 7.1: Asymmetrische Kongruenz in der TÜBA-D/Z (geschrieben)

- (5) a. In jeder Pilotenweste **ist**_{SG} [ein Kompass]_{SG} und [ein kleiner Sender zur Bestimmung des Abschussorts]_{SG} integriert.
- b. Bei heiteren bis wolkigen Abschnitten um die 18 Grad **kann**_{SG} [Gemüse]_{SG} und [Obst]_{SG} auch in unseren Breiten so richtig gedeihen.
- c. [Ein Abbruch des Daches]_{SG} und [eine Neuerrichtung mit leicht veränderter Konstruktion]_{SG} **würde**_{SG} knapp 2,7 Millionen Mark kosten.

Die Korpusergebnisse für die TÜBA-D/Z-Baumbank sind in Tabelle 7.1 in der Spalte “ungefiltert” dargestellt (die gefilterten Daten in der letzten Spalte werden später diskutiert). Die präverbalen Konstruktionen weisen in der überwiegenden Mehrheit (98,5%) symmetrische und nur zu 1,5% asymmetrische Kongruenz auf. In den postverbalen Konstruktionen tritt in 84,9% der Fälle symmetrische Kongruenz auf und nur in 15,1% der Fälle asymmetrische Kongruenz. Der Unterschied zwischen der Verteilung in den präverbalen und postverbalen Konstruktionen ist signifikant ($\chi^2(1) = 9,7; p < 0,01$). Asymmetrische Kongruenz kommt mit postverbalen Subjekten signifikant häufiger vor als mit präverbalen

Subjektposition	Kongruenz	V2 + VEnd (ungefiltert)	V2 + VEnd (gefiltert)
Präverbal	symmetrisch	10 (41,67%)	7 (70%)
Präverbal	asymmetrisch	14 (58,33%)	3 (30%)
<i>Insgesamt</i>		24	10
Postverbal	symmetrisch	1 (6,25%)	1 (10%)
Postverbal	asymmetrisch	15 (93,75%)	9 (90%)
<i>Insgesamt</i>		16	10

Tabelle 7.2: Asymmetrische Kongruenz in der TÜBA-D/S (gesprochen)

Subjekten. Dennoch besteht generell eine Präferenz für symmetrische Kongruenz in präverbalen und in postverbalen Konstruktionen.

Eine entsprechende Suche in der spontansprachlichen TÜBA-D/S-Baumbank ergab, dass in gesprochener Sprache Subjekte wesentlich seltener koordiniert sind als in geschriebener Sprache (siehe Tabelle 7.2). In der TÜBA-D/Z-Baumbank (geschrieben) wurden 207 relevante Fälle in einer Datenbasis von 14.940 Sätzen (= 1,4%) gefunden. In der TÜBA-D/S-Baumbank (gesprochen) waren es nur 40 relevante Konstruktionen aus einer Datenbasis von 38.196 Sätzen (= 0,1%). Einige Beispiele zur asymmetrischen Kongruenz in der spontansprachlichen TÜBA-D/S-Baumbank sind in (6) aufgeführt. Die Sätze in (6a,b) enthalten Subjekte in postverbaler Position. (6c) ist ein Beispiel mit einem präverbalen Subjekt. Die einzelnen Konjunkte sind jeweils eingeklammert.

- (6) a. Dann **ist**_{SG} dort [ein Hallenbad]_{SG} und [ein Fitnessraum]_{SG}.
- b. Also **bleibt**_{SG} eigentlich nur noch [der Juni]_{SG} und [der Juli]_{SG}.
- c. [Dienstag]_{SG} und [Mittwoch]_{SG} **passt**_{SG} ausgezeichnet.

Die Korpusergebnisse für die TÜBA-D/S-Baumbank sind in Tabelle 7.2 in der Spalte “ungefiltert” dargestellt (die gefilterten Daten in der letzten Spalte werden später diskutiert). In präverbalen Konstruktionen gibt es keine eindeutige Präferenz für symmetrische oder asymmetrische Kongruenz: 41,7% der Sätze weisen symmetrische und 58,3% asymmetrische Kongruenz auf. In postverbalen Konstruktionen zeigt sich eine starke Präferenz für asymmetrische Kongruenz: 93,75% der Fälle weisen asymmetrische und 6,25% symmetrische Kongruenz auf. Der Unterschied zwischen der Verteilung in den präverbalen und postverbalen Konstruktionen ist signifikant ($\chi^2(1) = 6,04$; $p = 0,014$). Auch hier kommt asymmetrische Kongruenz signifikant häufiger mit postverbalen als mit präverbalen Subjekten vor.

Insgesamt weisen 6,3% der Vorkommen in geschriebenen Texten asymmetrische Kongruenz auf im Vergleich zu 72,5% in gesprochenen Daten. Dieser Unterschied ist ebenfalls signifikant ($\chi^2(1) = 85,6$; $p < 0,01$). Asymmetrische Kongruenz kommt signifikant häufiger in gesprochenen als in geschriebenen Daten vor.

An dieser Stelle stellt sich die Frage, ob der semantische Typ der Subjekt-NP möglicherweise ein zusätzlicher Faktor ist, der die Kongruenz beeinflussen und die beschriebenen Effekte überlagern könnte. Subjekte, die aus koordinierten *abstrakten* NPs bestehen, führen unter Umständen zu einer Verbform im Singular. Dies könnte eine semantische Ursache haben und darauf zurückgeführt werden, dass die Konjunkte leicht eine abstrakte Einheit bilden können (siehe Lorimor, 2007).³ In den hier verwendeten Korpora betrifft dies beispielsweise Massenterme wie in (5b), ereignisbezogene NPs wie in (5c) und andere Abstrakta wie in (6c). In diesen Fällen kann die Subjekt-NP im Deutschen durch das singularische Pronomen ‘*das*’ ersetzt werden, was den Singular erleichtert. Prinzipiell ist in solchen Fällen auch eine Verbform im Plural möglich, die Semantik dieser Subjekt-NPs erleichtert jedoch die Singular-Form. Um den semantischen

³ Lorimor (2007) hat gezeigt, dass die Kongruenz mit koordinierten Subjekten durch den semantischen Typ der Konjunkte beeinflusst wird. In englischen Korpora führten koordinierte Massenterme und von Verben abgeleitete Nomina häufiger zu einer Singular-Form (41% bei Massennomen, 86% bei deverbale Nomina) als koordinierte Pronomen oder belebte Entitäten (jeweils 0%). Die Korpusdaten stammen jedoch aus dem *World Wide Web*, d.h. es ist leider nicht bekannt, welcher Anteil davon gesprochene Daten sind.

Einfluss so weit wie möglich zu reduzieren, wurden die gesprochenen und die geschriebenen Korpusdaten gefiltert. Es wurden alle Sätze ausgeschlossen, die eine Subjekt-NP enthalten, welche durch das Pronomen 'das' ersetzt werden kann. Die Ergebnisse der gefilterten Korpusdaten sind in den Tabellen 7.1 und 7.2 jeweils in der letzten Spalte ("gefiltert") dargestellt.

Die Ergebnisse zeigen, dass die Präferenzen in den gefilterten und ungefilterten Daten beinahe identisch sind, mit Ausnahme der präverbalen Konstruktionen im gesprochenen Korpus. In den gesprochenen Daten besteht nun eine Präferenz für symmetrische Kongruenz in präverbalen Konstruktionen, in postverbalen Konstruktionen wird hingegen (nach wie vor) die asymmetrische Kongruenz präferiert. In den geschriebenen Daten besteht nach wie vor eine Präferenz für symmetrische Kongruenz in präverbalen und postverbalen Konstruktionen. Der Fisher-Test (*'Fisher Exact Probability Test'*) zeigt, dass auch in den gefilterten Daten die asymmetrische Kongruenz signifikant häufiger bei postverbalen Subjekten auftritt als bei präverbalen Subjekten. Dieser Effekt zeigt sich sowohl in den geschriebenen ($p = 0,012$) als auch in den gesprochenen Korpusdaten ($p < 0,01$).

Insgesamt finden sich in den gefilterten Daten 2,4% asymmetrische Kongruenz in geschriebenen Texten im Vergleich zu 60,0% in den spontansprachlichen Dialogen. Nach dem Fisher-Test ist dieser Unterschied signifikant ($p < 0,01$). Die gefilterten Daten bestätigen somit, dass asymmetrische Kongruenz signifikant häufiger in gesprochenen als in geschriebenen Daten auftritt.

Zusammenfassend zeigen die Korpusdaten folgendes Bild: In den geschriebenen und gesprochenen Daten tritt die asymmetrische Kongruenz signifikant häufiger mit postverbalen als mit präverbalen Subjekten auf. Dies gilt sowohl für die gefilterten als auch für die ungefilterten Daten. Die Ergebnisse liefern somit insgesamt Evidenz für Hypothese 1. Es war nicht möglich, Hypothese 2 auf Basis der Korpusdaten zu testen, d.h. ob eine erhöhte Distanz zwischen dem Verb und einem postverbalem Subjekt die Kongruenz beeinflusst. Für eine weitere Differenzierung war die Anzahl der postverbalen Konstruktionen zu gering. Die Ergebnisse zeigen desweiteren, dass die asymmetrische Kongruenz signifikant häufiger in gesprochener Sprache (Online-Daten) als in geschriebener Sprache (Offline-Daten) auftritt. Dies gilt sowohl für die gefilterten als auch

für die ungefilterten Daten und liefert somit Evidenz für Hypothese 3. Um alle drei Hypothesen systematisch zu testen, wurden zwei experimentelle Studien durchgeführt, die im nächsten Abschnitt beschrieben werden.

7.3 Experimentelle Studien

In den experimentellen Studien wird die asymmetrische Kongruenz mit Hilfe von inkrementellen Grammatikalitätsurteilen (Offline-Datentyp) und einem Lesezeit-Experiment (Online-Datentyp) untersucht. Die inkrementellen Urteile werden am Rechner unter Verwendung einer Moving-Window-Technik (nicht-kumulativ) durchgeführt. Bei jedem Satzsegment muss eine bewusste Entscheidung getroffen werden, ob der Satz noch grammatisch ist. Ansonsten wird der Satz abgebrochen. Diese Methode hat gegenüber einer Fragebogenstudie oder der Magnitude-Estimation-Methode den Vorteil, dass genau lokalisiert werden kann, an welcher Stelle im Satz Probleme auftreten. Ohne Moving-Window-Technik kann nur eine Beurteilung für den Gesamtsatz erhoben werden.

Zur Erhebung der Lesezeiten wird die Self-paced-Reading-Methode (nicht-kumulativ) verwendet. Dabei sollen Sätze Segment für Segment in natürlicher Geschwindigkeit gelesen werden. Mit Ausnahme der Entscheidungskomponente, ob ein Satz noch grammatisch ist, handelt es sich hierbei um exakt dieselbe Methode, die auch für die inkrementellen Grammatikalitätsurteile verwendet wurde. Auch hier kann genau lokalisiert werden, an welcher Stelle im Satz Verarbeitungsschwierigkeiten oder eine Verarbeitungserleichterung auftritt. Ich habe diese beiden Methoden ausgewählt, weil sie bis auf die oben genannte Entscheidungskomponente identisch sind. Auf diese Weise können zwei Datentypen miteinander verglichen werden, die sich nur in ihrem Offline- bzw. Online-Aspekt unterscheiden (bewusste Entscheidung vs. unbewusste Leseprozesse). Im Hinblick auf die Überprüfung von Hypothese 3 ist somit eine Vergleichbarkeit gewährleistet.⁴

⁴ Würde man beispielsweise Eye-Tracking zur Erhebung der Lesezeiten (Online-Daten) verwenden und Magnitude-Estimation zur Erhebung der Beurteilungsdaten (Offline-Daten),

Gegenüber den Korpusstudien hat das experimentelle Vorgehen entscheidende Vorteile. Zum einen kann die Interaktion mit semantischen Faktoren minimiert werden. So werden in den Experimenten beispielsweise ausschließlich Testsätze verwendet, in denen das Subjekt das semantische Merkmal [+ menschlich] besitzt. Solche Subjekt-Phrasen können nicht durch das Pronomen ‘*das*’ ersetzt werden. Zum anderen kann der Verarbeitungsaufwand systematisch variiert werden, indem z.B. die Distanz zwischen dem finiten Verb und dem Subjekt vergrößert wird. In beiden Experimenten verwende ich dasselbe Testmaterial, das im Folgenden beschrieben wird.

7.3.1 Testmaterial für Experiment 1 und 2

Das Testmaterial besteht aus V2-Deklarativsätzen in drei unterschiedlichen Wortstellungsvarianten: i. Subjekt in präverbaler Position (7a,b), ii. Subjekt in postverbaler Position (7c,d) und iii. Subjekt in postverbaler Position mit zusätzlichem Abstand zwischen Verb und Subjekt (7e,f). Die Subjektphrasen bestehen aus zwei mit ‘*und*’ koordinierten NPs im Singular. Der Faktor *Wortstellung* wurde mit zwei Kongruenztypen (symmetrisch vs. asymmetrisch) gekreuzt. Daraus ergeben sich die folgenden sechs Bedingungen (7a–f). Die Schrägstriche markieren die Segmentierung bei der experimentellen Darbietung.

- (7) a. [Eine gutmütige Lehrerin]_{SG} / und / [eine unerfahrene Referendarin]_{SG} /
gehen_{PL} / nach dem Unterricht / frustriert / aus dem Klassenzimmer /
 einer elften Klasse.
- b. [Eine gutmütige Lehrerin]_{SG} / und / [eine unerfahrene Referendarin]_{SG} /
geht_{SG} / nach dem Unterricht / frustriert / aus dem Klassenzimmer /
 einer elften Klasse.

könnten unterschiedliche Ergebnisse in den beiden Experimenten auf andere methodische Aspekte als den Online- bzw. Offline-Unterschied zurückgeführt werden.

- c. Frustriert / **gehen**_{PL} / [eine gutmütige Lehrerin]_{SG} / und / [eine unerfahrene Referendarin]_{SG} / nach dem Unterricht / aus dem Klassenzimmer / einer elften Klasse.
- d. Frustriert / **geht**_{SG} / [eine gutmütige Lehrerin]_{SG} / und / [eine unerfahrene Referendarin]_{SG} / nach dem Unterricht / aus dem Klassenzimmer / einer elften Klasse.
- e. Frustriert / **gehen**_{PL} / nach dem Unterricht / [eine gutmütige Lehrerin]_{SG} / und / [eine unerfahrene Referendarin]_{SG} / aus dem Klassenzimmer / einer elften Klasse.
- f. Frustriert / **geht**_{SG} / nach dem Unterricht / [eine gutmütige Lehrerin]_{SG} / und / [eine unerfahrene Referendarin]_{SG} / aus dem Klassenzimmer / einer elften Klasse.

Die Testsätze werden in beiden Experimenten in acht Segmenten (Regionen) präsentiert. Manche dieser Segmente (kritische Segmente) sind für diese Studie von besonderem Interesse. Das kritische Segment ist die erste Region, die bei der inkrementellen Verarbeitung von links nach rechts eine Entscheidung erlaubt, ob der Satz hinsichtlich Subjekt-Verb-Kongruenz grammatisch ist oder nicht. Dafür müssen sowohl das Verb als auch das Subjekt bereits verarbeitet worden sein.

In präverbale Konstruktionen ist das Verb das kritische Segment (Region 4). In postverbalen Konstruktionen hängen die kritischen Segmente davon ab, ob der Parser beide Konjunkte zur Herstellung der Kongruenz heranzieht (globaler Prozess) oder ob er nur eines der Konjunkte betrachtet (lokaler Prozess). Sollte der Parser global vorgehen, ist das kritische Segment das zweite Konjunkt (Region 5 in (7c,d) und Region 6 in (7e,f)). Sollte der Parser lokal vorgehen, ist das kritische Segment das erste Konjunkt (Region 3 in (7c,d) und Region 4 in (7e,f)).

Durch das experimentelle Design wird ausgeschlossen, dass die Versuchspersonen elliptische Strukturen (nach Aoun et al., 1994) für die Testsätze aufbauen. Die Testsätze sind so konstruiert, dass die Probanden

gezwungen sind, eine phrasale Struktur für die Subjekt-NP aufzubauen.⁵ Es ist prinzipiell nicht ausgeschlossen, dass eine phrasale Analyse am Satzende zugunsten einer sententialen Struktur reanalysiert wird, dies sollte sich jedoch in den Lesezeiten widerspiegeln. Dieser Punkt wird in der abschließenden Diskussion (Abschnitt 7.4) nochmals aufgegriffen.

Insgesamt wurden 36 Testsätze erstellt, wobei jeder Testsatz in sechs Versionen vorkommt (7a–f), 20 Kontrollitems in je 4 Versionen (8a–d) sowie 40 Ablenkensätze. Die Kontrollitems enthalten nicht-koordinierte pluralische Subjektphrasen in zwei Wortstellungsvarianten: i. Subjekt in präverbaler Position (8a,b), ii. Subjekt in postverbaler Position (8c,d). Der Faktor *Wortstellung* wurde mit zwei Numerusvarianten gekreuzt (Singular vs. Plural). Die Sätze

⁵ In Sprachverstehensexperimenten, die eine Moving-Window-Technik verwenden, kann genau kontrolliert werden, welche Teile des Satzes dem menschlichen Parser zu einem bestimmten Zeitpunkt zur Verfügung stehen. In Sprachproduktionsexperimenten ist dies nicht möglich. Ich habe die Testsätze in der vorliegenden Studie so konstruiert, dass der Parser bei der Verarbeitung der *Konjunktion* nicht genügend Informationen hat, um zu diesem Zeitpunkt eine vollständige sententiale Struktur für das erste Konjunkt aufzubauen (ich nenne dies: ‘lack of information’). Im Folgenden sind die verwendeten Testsätze mit asymmetrischer Kongruenz auf der Basis einer sententialen Analyse in den drei verwendeten Wortstellungsvarianten aufgeführt (präverbales Subjekt (i.), postverbales Subjekt (ii.), postverbales Subjekt mit Abstand (iii.)).

- i. [Eine gutmütige Lehrerin *geht nach dem Unterricht* *frustriert aus dem Klassenzimmer einer elften Klasse*] und [eine unerfahrene Referendarin *geht nach dem Unterricht* *frustriert aus dem Klassenzimmer einer elften Klasse*].
- ii. [Frustriert *geht eine gutmütige Lehrerin nach dem Unterricht aus dem Klassenzimmer einer elften Klasse*] und [frustriert *geht eine unerfahrene Referendarin nach dem Unterricht aus dem Klassenzimmer einer elften Klasse*].
- iii. [Frustriert *geht nach dem Unterricht eine gutmütige Lehrerin aus dem Klassenzimmer einer elften Klasse*] und [frustriert *geht nach dem Unterricht eine unerfahrene Referendarin aus dem Klassenzimmer einer elften Klasse*].

Für eine sententiale Analyse der Testsätze müsste der Parser eine syntaktische Struktur für die kursiv markierten Teile der Sätze in (i.–iii.) erzeugen. Wenn der Parser in den Testsätzen auf die Konjunktion trifft, stehen diese Teile des Satzes jedoch noch nicht zur Verfügung. Da diese Informationen erst am Satzende verfügbar sind, ist der Parser nicht in der Lage, inkrementell eine sententiale Struktur für die Testsätze aufzubauen. Das Sprachsystem ist gezwungen, eine phrasale Analyse für die Subjekt-NP zu erzeugen.

im Singular, d.h. 50% der Kontrollitems, sind ungrammatisch hinsichtlich der Subjekt-Verb-Kongruenz. Mit Hilfe der Kontrollitems kann festgestellt werden, ob sich die Verarbeitung von koordinierten Subjekten, bestehend aus zwei NPs im Singular, von nicht-koordinierten pluralischen Subjektphrasen hinsichtlich der Subjekt-Verb-Kongruenz unterscheidet.

- (8) a. [Drei motivierte Fußballspieler]_{PL} / **verlassen**_{PL} / nach dem Training / erschöpft / das Stadion.
- b. [Drei motivierte Fußballspieler]_{PL} / **verlässt**_{SG} / nach dem Training / erschöpft / das Stadion.
- c. Erschöpft / **verlassen**_{PL} / [drei motivierte Fußballspieler]_{PL} / nach dem Training / das Stadion.
- d. Erschöpft / **verlässt**_{SG} / [drei motivierte Fußballspieler]_{PL} / nach dem Training / das Stadion.

Die Ablenkersätze enthalten i. koordinierte Subjektphrasen mit der Konjunktion ‘*oder*’, ii. koordinierte Objektphrasen oder Adverbiale mit ‘*oder*’, iii. koordinierte Objektphrasen mit der Konjunktion ‘*und*’, iv. nicht-koordinierte Objektphrasen im Plural. Zehn der 40 Ablenkersätze enthalten Kongruenzverletzungen bezüglich Numerus, Genus und Kasus innerhalb einer Nominalphrase.

Das Testmaterial besteht insgesamt aus 216 Testsätzen, 80 Kontrollitems und 40 Ablenkersätzen, die gemäß dem *Latin-Square-Design* auf 12 Listen verteilt wurden. Jede Liste enthält jedes Test- und Kontrollitem in nur einer Version, d.h. insgesamt 36 Testitems und 20 Kontrollitems. Die 40 Ablenkersätze wurden jeder Liste hinzugefügt. Jede Versuchsperson bekommt eine Liste bestehend aus 96 Sätzen in randomisierter Abfolge dargeboten. Für die Hälfte der Sätze wurden Verständnisfragen erstellt, die mit ‘*Ja*’ oder ‘*Nein*’ beantwortet werden sollen. Die Verständnisfragen sind gleichmäßig über alle Bedingungen verteilt.

7.3.2 Experiment 1: Inkrementelle Grammatikalitätsurteile

7.3.2.1 Versuchspersonen

An diesem Experiment haben 48 Studenten der Universität Tübingen teilgenommen. Alle Probanden waren deutsche Muttersprachler und wurden für ihre Teilnahme bezahlt. Das Experiment dauerte ca. 30 Minuten.

7.3.2.2 Methoden

Die Versuchspersonen führten eine Beurteilungsaufgabe zur Grammatikalität von Sätzen durch, wobei die Segmente inkrementell unter Verwendung einer nicht-kumulativen Moving-Window-Technik präsentiert wurden. Zu Beginn eines Durchgangs wurde der Satz maskiert dargeboten, wobei jeder Buchstabe durch einen Bindestrich verdeckt war. Um das nächste Segment des Satzes aufzudecken, musste ein Knopf der Antwort-Box gedrückt werden. Bei jedem Segment mussten die Probanden eine Entscheidung treffen. Ist der Satz noch grammatisch, sollten die Probanden fortfahren und hierfür den rechten Knopf der Antwort-Box drücken. Ist der Satz ungrammatisch, sollten die Probanden den Durchgang abbrechen und den linken Knopf der Antwort-Box drücken. Die Verständnisfragen wurden nach dem letzten Segment des Satzes präsentiert und sollten mit 'Ja' oder 'Nein' beantwortet werden.

Vor dem eigentlichen Experiment wurde eine Übungseinheit mit neun Sätzen durchgeführt, wobei auf jeden Satz eine Verständnisfrage folgte. Drei der neun Übungssätze enthielten Kongruenzverletzungen. Nach jedem Übungsdurchgang erhielten die Teilnehmer eine Rückmeldung, ob ihre Antwort korrekt war bzw. ob sie einen ungrammatischen Satz akzeptiert hatten. Während des eigentlichen Experiments gab es keine Rückmeldung. Die Versuchspersonen sollten die Sätze in normaler Geschwindigkeit lesen, jedoch so aufmerksam, dass die Verständnisfragen beantwortet werden können.

Subjektposition	Symmetrische Kongruenz	Asymmetrische Kongruenz
Präverbal	1,04%	93,75%
Postverbal	3,82%	81,94%
Postverbal + Abstand	4,51%	79,51%

Tabelle 7.3: Abbruchraten für jede Bedingung (Testsätze)

7.3.2.3 Ergebnisse und Diskussion

Die Abbruchraten in den Testsätzen sind für jede Bedingung in Tabelle 7.3 dargestellt. Die Mehrheit der Probanden beurteilte die asymmetrische Kongruenz in den Testsätzen als ungrammatisch: 93,8% in präverbalen Konstruktionen, 81,9% in postverbalen Konstruktionen und 79,5% in postverbalen Konstruktionen mit zusätzlichem Abstand. Sätze mit symmetrischer Kongruenz wurden hingegen selten abgebrochen. Die ungrammatischen Kontrollitems (nicht-koordiniertes Subjekt im Plural und Verbform im Singular) wurden in 98 bis 100% der Fälle abgebrochen, während die grammatischen Kontrollitems (Verbform im Plural) nur in 1% der Fälle abgebrochen wurden. Die Verständnisfragen wurden zu 89% korrekt beantwortet.

Die Abbruchraten in den Testsätzen wurden mit Hilfe von Varianzanalysen mit den Faktoren *Wortstellung* (präverbal, postverbal, postverbal + Abstand) und *Kongruenz* (symmetrisch, asymmetrisch) ausgewertet. Die Analyse ergab einen signifikanten Haupteffekt des Faktors *Kongruenz* ($F_1(1,47) = 995,58$; $p < 0,01$; $F_2(1,35) = 2175,63$; $p < 0,01$). Das zeigt, dass die symmetrische gegenüber der asymmetrischen Kongruenz stark präferiert wird. Weiterhin wurde ein signifikanter Haupteffekt der *Wortstellung* gefunden ($F_1(1,47) = 4,48$; $p = 0,016$; $F_2(1,35) = 3,94$; $p < 0,05$) sowie eine signifikante Interaktion zwischen *Wortstellung* und *Kongruenz* ($F_1(1,47) = 10,23$; $p < 0,01$; $F_2(1,35) = 8,85$; $p < 0,01$). Dies zeigt, dass die drei Wortstellungsvarianten hinsichtlich Kongruenz unterschiedlich verarbeitet werden.

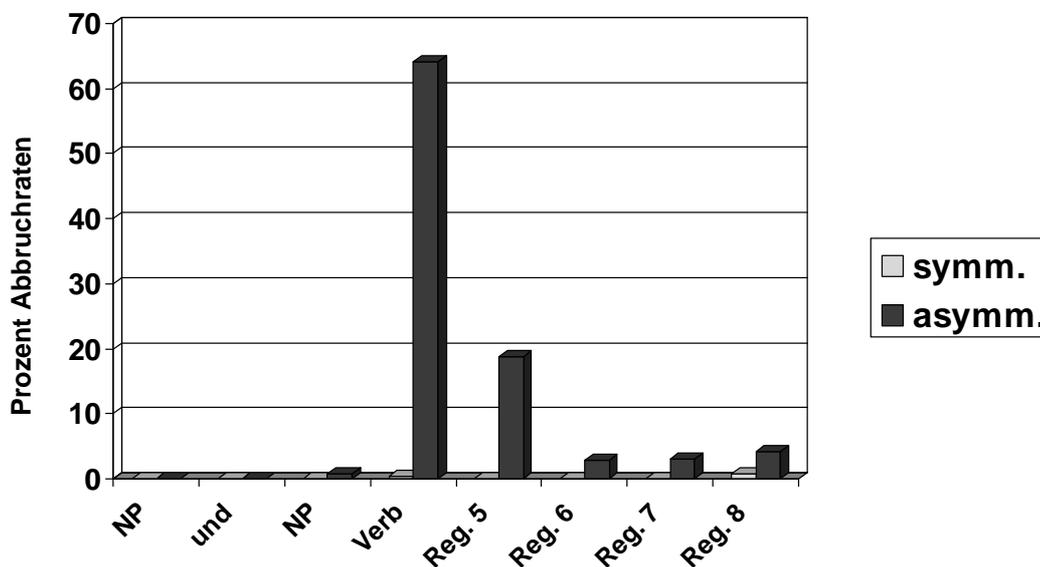


Abbildung 7.1: Abbruchraten pro Segment (Region) für präverbale Konstruktionen

Zusätzliche t-Tests (für gepaarte Stichproben) ergaben, dass die asymmetrische Kongruenz mit postverbalen Subjekten (7d) signifikant besser beurteilt wurde als mit präverbalen Subjekten (7b) ($t_1(47) = 3,02$; $p < 0,01$; $t_2(35) = 3,11$; $p < 0,01$). Dies liefert Evidenz für Hypothese 1. Der Unterschied zwischen der asymmetrischen Kongruenz bei postverbalen und entfernt stehenden postverbalen Subjekten war nicht signifikant ($t_{1,2} < 1$). Es wurde somit keine Evidenz für Hypothese 2 gefunden. Dies ist jedoch nicht verwunderlich, da es sich bei der Beurteilung von Grammatikalität eher um einen Offline-Datentyp handelt. Verarbeitungsunterschiede könnten durch Entscheidungsprozesse überdeckt worden sein. Obwohl die asymmetrische Kongruenz mit koordinierten Subjekten in postverbaler Position signifikant besser beurteilt wurde als mit Subjekten in präverbaler Position, zeigen die Probanden eine deutliche Präferenz für symmetrische Kongruenz in allen drei Wortstellungstypen.

Wie die Abbruchraten der Testsätze auf die einzelnen Segmente (Regionen) verteilt sind, wird für die präverbalen Konstruktionen (7a,b) in Abbildung 7.1 illustriert, für die postverbalen Konstruktionen (7c,d) in Abbildung 7.2 und für

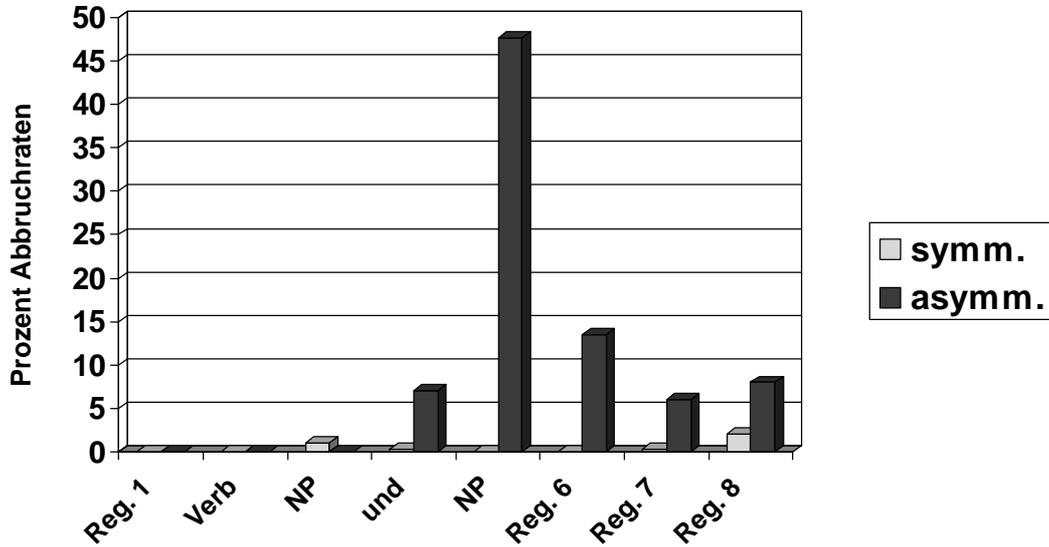


Abbildung 7.2: Abbruchraten pro Segment (Region) für postverbale Konstruktionen

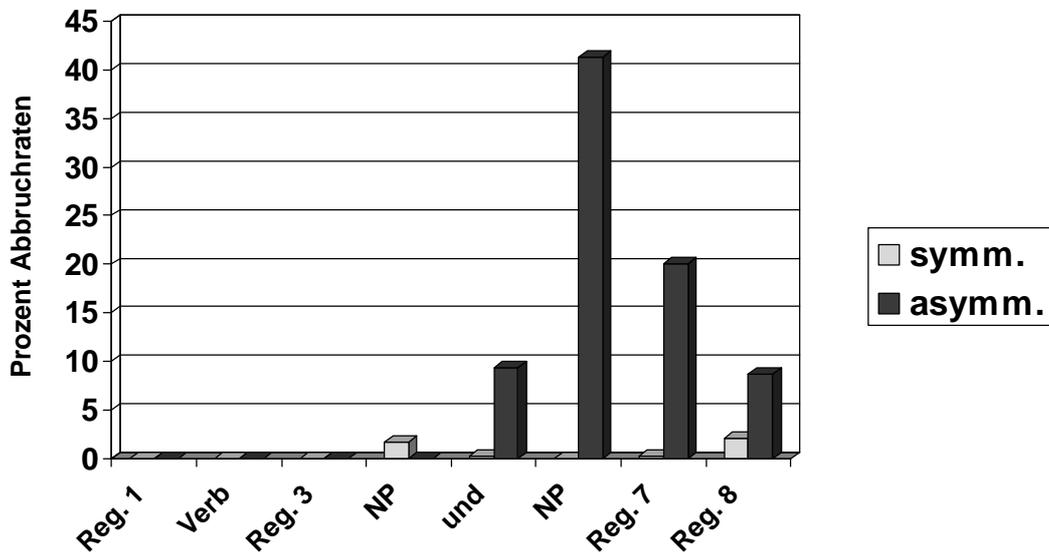


Abbildung 7.3: Abbruchraten pro Segment (Region) für postverbale Konstruktionen mit entfernt stehendem Subjekt

Sätze mit entfernt stehendem postverbalem Subjekt (7e,f) in Abbildung 7.3. Wie in den Abbildungen 7.1, 7.2 und 7.3 zu sehen ist, werden die Sätze mit präverbialem Subjekt in den meisten Fällen auf dem Verb-Segment abgebrochen und die Sätze in den beiden postverbalen Wortstellungstypen auf dem zweiten Konjunkt (*NP*). Das sind die jeweiligen kritischen Segmente, wenn das Sprachsystem beide Konjunkte für die Kongruenz heranzieht. Die Ergebnisse zeigen somit auch, dass der Parser bei der Verarbeitung der Kongruenz in dieser Beurteilungsaufgabe *global* vorgeht.

7.3.3 Experiment 2: Lesezeitexperiment

7.3.3.1 Versuchspersonen

Es wurden 48 Studenten der Universität Tübingen getestet, die nicht an Experiment 1 teilgenommen haben. Alle Probanden waren deutsche Muttersprachler und wurden für ihre Teilnahme bezahlt. Das Experiment dauerte ca. 30 Minuten.

7.3.3.2 Methoden

Die Teilnehmer führten eine Lesezeit-Aufgabe ('Self-paced Reading') mit dem oben beschriebenen Testmaterial durch, unter Verwendung einer nicht-kumulativen Moving-Window-Technik. Dabei sollen die Sätze Segment für Segment in natürlicher Geschwindigkeit gelesen werden. Zu Beginn eines Durchgangs wurde der Satz maskiert dargeboten, wobei jeder Buchstabe durch einen Bindestrich verdeckt war. Um das nächste Segment des Satzes aufzudecken, musste der mittlere Knopf der Antwort-Box gedrückt werden. Es wurde die Zeit gemessen, die die Probanden auf einem Segment verbringen. Mit Ausnahme der Entscheidungskomponente, d.h. der Möglichkeit, einen Durchgang abubrechen, waren Experiment 1 und 2 identisch.

7.3.3.3 Ergebnisse

Die Lesezeiten, die im Folgenden präsentiert werden, sind ausreißerkorrigiert. Lesezeiten unter 100 ms und Zeiten, die mehr als drei Standardabweichungen über dem Mittelwert lagen, wurden durch den Mittelwert ersetzt. Diese Prozedur wurde für jede Bedingung und für jedes Segment durchgeführt.⁶ Die durchschnittlichen Lesezeiten in ms sind für präverbale Konstruktionen in Abbildung 7.4 dargestellt, für postverbale Konstruktionen in Abbildung 7.5 und für Sätze mit entfernt stehendem postverbalen Subjekt in Abbildung 7.6. Die Verständnisfragen wurden zu 90% korrekt beantwortet.

In präverbalen Konstruktionen wurden längere Lesezeiten für die asymmetrische im Vergleich zur symmetrischen Kongruenz gefunden (siehe Abbildung 7.4). Dies betrifft insbesondere das Verb-Segment (Region 4) und das nachfolgende Spillover-Segment (Region 5). In postverbalen Konstruktionen wurde die asymmetrische Kongruenz ähnlich schnell wie die symmetrische Kongruenz verarbeitet (siehe Abbildung 7.5). In postverbalen Konstruktionen mit zusätzlichem Abstand zwischen Verb und Subjekt wurde die asymmetrische Kongruenz sogar schneller verarbeitet als die symmetrische Kongruenz (siehe Abbildung 7.6). Dies betrifft vor allem das erste Konjunkt (Region 4).

Um die drei verschiedenen Wortstellungstypen (präverbal, postverbal, postverbal + Abstand) miteinander vergleichen zu können, wurden die relativen Lesezeit-Unterschiede zwischen symmetrischer und asymmetrischer Kongruenz anstatt der absoluten Lesezeiten betrachtet. Hierfür gibt es zwei Gründe. Zum einen wurden bei der Analyse der *lokal* kritischen Segmente in den verschiedenen

⁶ Durch die Ausreißerkorrektur sollen Messfehler aussortiert werden. Bei Lesezeiten von unter 100 ms pro Segment kann es sich nicht um echte Lesezeiten handeln, da es überhaupt nicht möglich ist, ein Segment in dieser Geschwindigkeit zu lesen. Diese Werte könnten beispielsweise darauf zurückgeführt werden, dass der Knopf der Antwort-Box versehentlich zweimal hintereinander gedrückt wurde und dadurch ein Segment übersprungen wurde. Lesezeiten unter 100 ms pro Segment wurden daher als Messfehler betrachtet. Extrem lange Lesezeiten pro Segment wurden ebenfalls als Fehler betrachtet. Extrem lange Zeiten reflektieren keine unbewussten Leseprozesse sondern sind vermutlich eher darauf zurückzuführen, dass die Probanden in diesem Moment unkonzentriert waren und möglicherweise angefangen haben, über den Satz nachzudenken. Werte, die mehr als drei Standardabweichungen über dem Mittelwert lagen, wurden daher nicht als echte Lesezeiten gewertet.

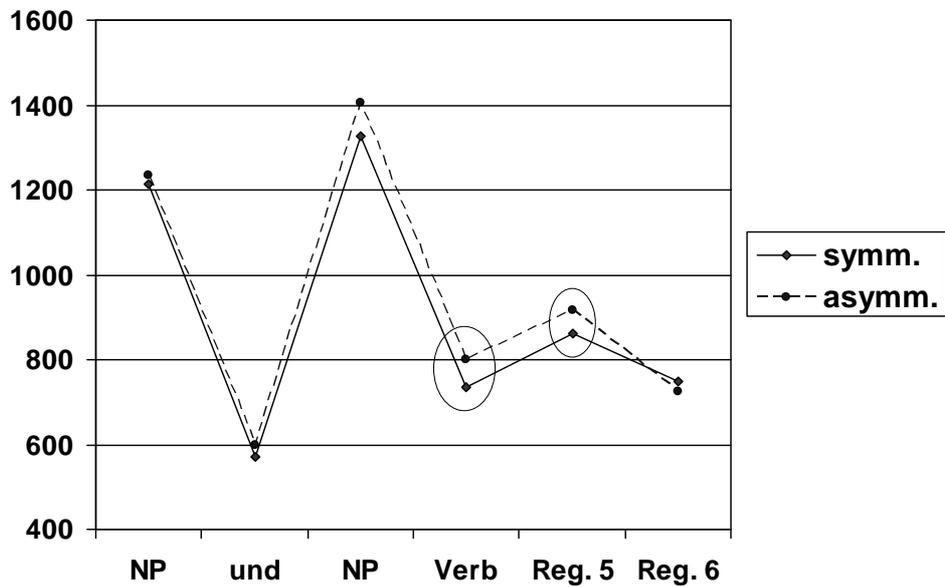


Abbildung 7.4: Mittlere Lesezeiten pro Segment (in ms) für präverbale Konstruktionen (Region 1–6). Signifikante Unterschiede sind durch Kreise markiert.

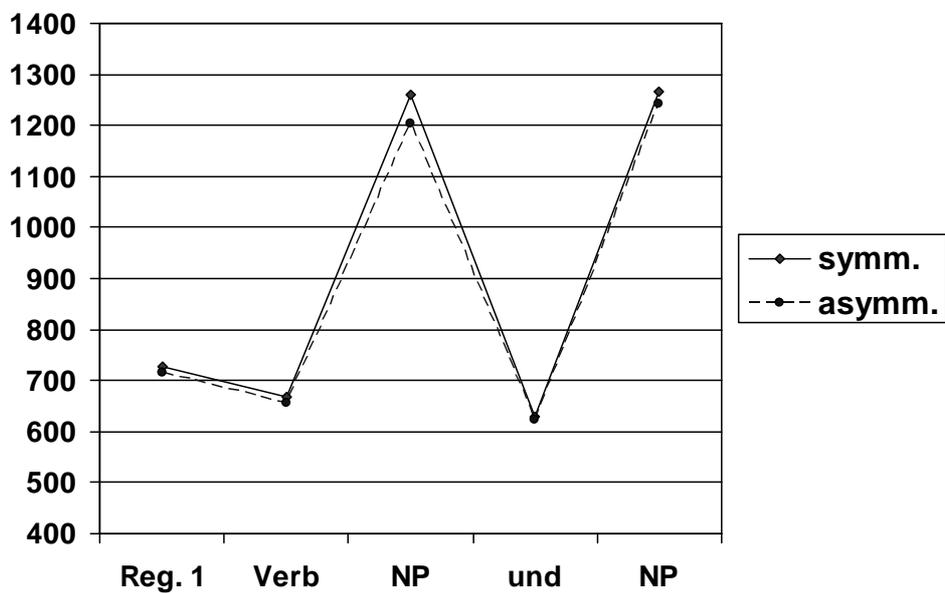


Abbildung 7.5: Mittlere Lesezeiten pro Segment (in ms) für postverbale Konstruktionen (Region 1–5)

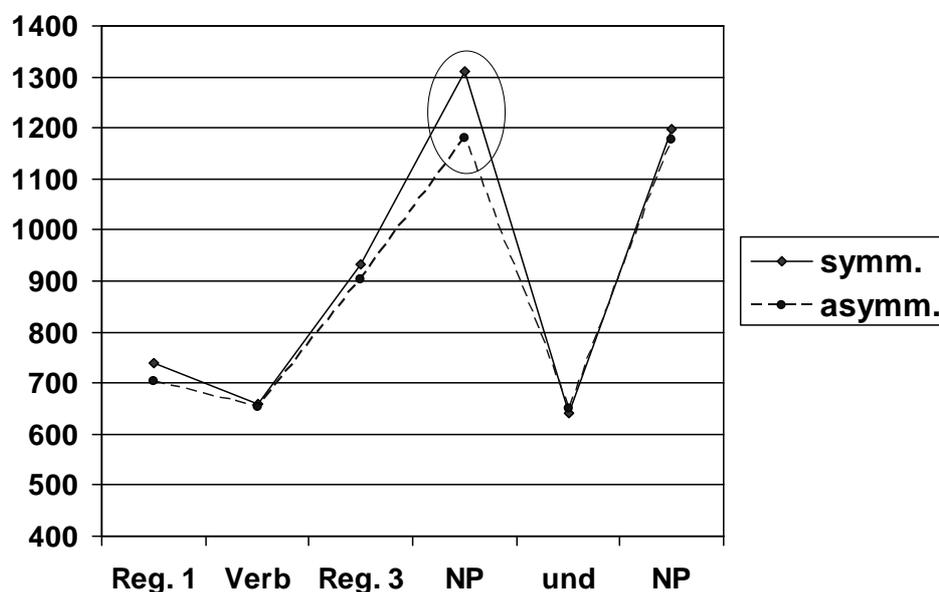


Abbildung 7.6: Mittlere Lesezeiten pro Segment (in ms) für postverbale Konstruktionen mit Abstand (Region 1–6).

Wortstellungstypen unterschiedliche Teile des Satzes miteinander verglichen. In den präverbale Konstruktionen war das Verb das lokal kritische Segment, in den postverbalen Konstruktionstypen war es das erste Konjunkt. Zum anderen kamen die kritischen Segmente in Sätzen mit unterschiedlicher Wortstellung vor, d.h. sie folgten auf unterschiedliche Segmente und konnten daher nicht direkt miteinander verglichen werden.

Entsprechend wurden die mittleren Differenzwerte zwischen asymmetrischer und symmetrischer Kongruenz für die relevanten Segmente innerhalb eines Wortstellungstyps pro Versuchsperson und pro Item berechnet. Die Lesezeiten für symmetrische Kongruenz wurden von den Lesezeiten für asymmetrische Kongruenz subtrahiert. Positive Werte bedeuten, dass die symmetrische Kongruenz schneller verarbeitet wurde als die asymmetrische Kongruenz. Negative Werte bedeuten, dass die asymmetrische Kongruenz schneller verarbeitet wurde. Ein Differenzwert von 0 bedeutet, dass es keinen Unterschied zwischen symmetrischer und asymmetrischer Kongruenz gab. Um herauszufinden, ob die verschiedenen Wortstellungstypen hinsichtlich Kongruenz

unterschiedlich verarbeitet wurden, habe ich die mittleren Lesezeit-Differenzen unter Verwendung von Varianzanalysen (mit Messwiederholung) und t-Tests (für gepaarte Stichproben) verglichen. Mit Hilfe von Einstichproben-t-Tests wurden die mittleren Differenzwerte auch gegen den Wert 0 getestet. Auf diese Weise können innerhalb eines Wortstellungstyps signifikante Unterschiede zwischen symmetrischer und asymmetrischer Kongruenz gefunden werden.

Zunächst wurden die drei verschiedenen Wortstellungstypen miteinander verglichen. Die mittleren Lesezeit-Differenzen der *lokal* kritischen Segmente wurden mit Hilfe von Varianzanalysen mit dem Faktor *Wortstellung* (präverbal, postverbal, postverbal + Abstand) untersucht. Die Analyse ergab einen signifikanten Haupteffekt der *Wortstellung* ($F_1(1,47) = 7,27$; $p < 0,01$; $F_2(1,35) = 9,06$; $p < 0,01$). Das zeigt, dass sich die mittleren Lesezeit-Differenzen zwischen asymmetrischer und symmetrischer Kongruenz in den verschiedenen Wortstellungstypen unterscheiden (präverbal: 64,95 ms, postverbal: -54,24 ms, postverbal + Abstand: -131,71 ms). Weitere t-Tests (für gepaarte Stichproben) ergaben einen signifikanten Unterschied für die Differenzwerte in den präverbalen (64,95 ms) und postverbalen (-54,24 ms) Konstruktionen ($t_1(47) = 2,70$; $p = 0,010$; $t_2(35) = 3,12$; $p < 0,01$). Asymmetrische Kongruenz wurde in den postverbalen Konstruktionen signifikant schneller verarbeitet als in den präverbalen Konstruktionen. Der Unterschied zwischen postverbalen Konstruktionen (-54,24 ms) und Konstruktionen mit entfernt stehendem postverbalen Subjekt (-131,71 ms) war nicht signifikant ($t_1(47) = 1,36$; $p = 0,18$; $t_2(35) = 1,43$; $p = 0,16$).

Weiterhin wurden die Unterschiede zwischen den beiden Kongruenztypen innerhalb eines Wortstellungstyps untersucht. Mit Hilfe von Einstichproben-t-Tests wurden die mittleren Lesezeit-Differenzen zwischen asymmetrischer und symmetrischer Kongruenz gegen den Wert 0 getestet. Positive t-Werte bedeuten, dass die Sätze mit symmetrischer Kongruenz schneller gelesen wurden, negative t-Werte bedeuten, dass Sätze mit asymmetrischer Kongruenz schneller waren.

In präverbalen Konstruktionen waren die Lesezeit-Unterschiede zwischen asymmetrischer und symmetrischer Kongruenz signifikant für das Verb-Segment (Region 4) und für das nachfolgende Spillover-Segment (Region 5) (Region 4: $t_1(47) = 2,42$; $p = 0,02$; $t_2(35) = 2,28$; $p = 0,03$; Region 5: $t_1(47) = 2,29$; $p = 0,03$; $t_2(35) = 2,40$; $p = 0,02$).⁷ Das zeigt, dass die Sätze mit symmetrischer Kongruenz schneller gelesen wurden. Die Analyse des vorangehenden Segments (Region 3) ergab keinen signifikanten Unterschied ($t_1(47) = 1,54$; $p = 0,13$; $t_2(35) = 1,62$; $p = 0,11$). In postverbalen Konstruktionen waren die Lesezeit-Differenzen für das erste Konjunkt (Region 3) nicht signifikant ($t_1(47) = -1,55$; $p = 0,13$; $t_2(35) = -1,31$; $p = 0,20$). Das heißt, dass Sätze mit symmetrischer und asymmetrischer Kongruenz gleich schnell gelesen wurden. In postverbalen Konstruktionen mit entfernt stehendem Subjekt ergab die Analyse für das erste Konjunkt einen signifikanten Unterschied ($t_1(47) = -3,02$; $p < 0,01$; $t_2(35) = -2,86$; $p < 0,01$). Sätze mit asymmetrischer Kongruenz wurden hier schneller gelesen. Die Analyse des vorangehenden Segments (Region 3) ergab keinen signifikanten Unterschied.

Schließlich wurden die *global* kritischen Segmente in beiden postverbalen Wortstellungstypen analysiert, d.h. die mittleren Lesezeiten für die jeweiligen zweiten Konjunkte. Es ist legitim, in diesem Fall die absoluten Lesezeiten, anstatt der Differenzwerte, zu vergleichen, da jeweils identische Segmente (die zweiten Konjunkte) miteinander verglichen wurden. Außerdem waren die beiden vorangehenden Segmente in den beiden Wortstellungstypen ebenfalls identisch. Varianzanalysen mit den Faktoren *Wortstellung* (präverbal, postverbal, postverbal + Abstand) und *Kongruenz* (symmetrisch, asymmetrisch) ergaben einen signifikanten Haupteffekt der *Wortstellung* in der Versuchspersonen-Analyse ($F_1(1,47) = 7,17$; $p = 0,01$; $F_2(1,35) = 3,95$; $p = 0,06$). Es wurde kein Haupteffekt für *Kongruenz* gefunden und auch keine Interaktion zwischen *Wortstellung* und *Kongruenz* (alle $F_s < 1$). Das zeigt, dass in beiden postverbalen Wortstellungstypen kein Lesezeit-Unterschied zwischen symmetrischer und asymmetrischer Kongruenz für das zweite Konjunkt besteht.

⁷ Das Verb-Segment ist das einzige Segment, das in den beiden Kongruenztypen nicht völlig identisch ist (3. Pers. Sg. / 3. Pers. Pl.). Allerdings wurden die längeren Lesezeiten für das kürzere Segment (Verform im Singular) gefunden, d.h. die Lesezeit-Unterschiede wären noch größer, wenn die Daten hinsichtlich Länge korrigiert worden wären.

7.3.3.4 Diskussion

In diesem Lesezeit-Experiment ('Self-paced Reading') wurde die asymmetrische Kongruenz in Sätzen mit postverbalem Subjekt schneller verarbeitet als in Sätzen mit präverbalem Subjekt. Das liefert Evidenz für Hypothese 1. In präverbalen Konstruktionen wurden signifikant längere Lesezeiten für die asymmetrische Kongruenz gegenüber symmetrischer Kongruenz gefunden. In postverbalen Konstruktionen wurde die asymmetrische Kongruenz gleich schnell verarbeitet wie die symmetrische Kongruenz. In postverbalen Konstruktionen mit entfernt stehendem Subjekt wurde die asymmetrische Kongruenz sogar signifikant schneller verarbeitet als symmetrische Kongruenz. Das deutet darauf hin, dass die Präferenz durch den höheren Verarbeitungsaufwand zur asymmetrischen Kongruenz hin verschoben wurde. Diese Ergebnisse stützen Hypothese 2.

Weiterhin geben die Ergebnisse Aufschluss über die Vorgehensweise des menschlichen Parsers (lokal vs. global) bei der Herstellung von Kongruenz. In der Beurteilungsaufgabe (Experiment 1) wurde die Entscheidung über Grammatikalität in den postverbalen Konstruktionstypen *global* beim zweiten Konjunkt getroffen. Im Lesezeit-Experiment (Experiment 2) hingegen traten die Unterschiede zwischen symmetrischer und asymmetrischer Kongruenz in postverbalen Konstruktionen mit entferntem Subjekt *lokal* beim ersten Konjunkt auf. Die längeren Lesezeiten für symmetrische Kongruenz sind vermutlich auf die lokale Inkompatibilität zwischen der Plural-Verbform und dem ersten Konjunkt im Singular zurückzuführen. Beim zweiten Konjunkt wurde kein Effekt gefunden. Daraus kann geschlossen werden, dass bei erhöhtem Abstand zwischen Verb und Subjekt nur das erste Konjunkt für die Kongruenz herangezogen wird.

In der Beurteilungsaufgabe (Offline-Datentyp) wurde die asymmetrische Kongruenz in postverbalen Konstruktionen als ungrammatisch bewertet. Dieselben Sätze wurden jedoch im Lesezeit-Experiment (Online-Datentyp) problemlos verarbeitet oder sogar präferiert. Das liefert Evidenz für Hypothese 3.

7.4 Zusammenfassung und Diskussion

In diesem Kapitel habe ich drei Hypothesen getestet, die aus einem Verarbeitungsansatz für asymmetrische Kongruenz abgeleitet werden können. Nach diesem Ansatz ist asymmetrische Kongruenz auf die inkrementelle Sprachverarbeitung von links nach rechts zurückzuführen. Steht das Subjekt in präverbaler Position, ist die Pluralität des Subjektausdrucks bereits berechnet, wenn der Parser auf das finite Verb trifft. Dies sollte eine Präferenz für symmetrische Kongruenz zur Folge haben. Besetzt das Subjekt eine postverbale Position wird die Präferenz für symmetrische oder asymmetrische Kongruenz durch den Verarbeitungsaufwand beeinflusst. Je höher der Verarbeitungsaufwand ist, desto lokaler operiert das Sprachsystem. Dies sollte ein vermehrtes Auftreten von asymmetrischer Kongruenz zur Folge haben. Um die aufgestellten Hypothesen zu testen, habe ich vier verschiedene Datentypen im Deutschen untersucht: geschriebene und gesprochene Korpusdaten sowie inkrementelle Grammatikalitätsurteile und Lesezeiten. Im Folgenden gehe ich darauf ein, welche Evidenzen ich für die jeweiligen Hypothesen gefunden habe.

In *Hypothese 1* wird angenommen, dass asymmetrische Kongruenz häufiger vorkommt oder leichter verarbeitet wird, wenn das Subjekt eine postverbale Position besetzt. Für diese Hypothese habe ich in allen Datentypen starke Evidenz gefunden. Asymmetrische Kongruenz kommt häufiger in postverbalen als in präverbalen Konstruktionen vor, sowohl in gesprochenen als auch in geschriebenen Korpusdaten. Die inkrementellen Grammatikalitätsurteile haben gezeigt, dass asymmetrische Kongruenz mit postverbalen Subjektphrasen akzeptabler ist als mit präverbalen Subjektphrasen. Im Lesezeitexperiment wurde asymmetrische Kongruenz in den postverbalen Wortstellungstypen signifikant schneller verarbeitet als in präverbalen Konstruktionen.

An dieser Stelle ist anzumerken, dass Lorimor (2007) ähnliche Befunde in Sprachproduktions-Experimenten mit Sprechern des libanesisch Arabischen und des Englischen erzielt hat. In einer Bildbeschreibungsaufgabe wurden Verben im Singular häufiger mit postverbalen als mit präverbalen Subjektphrasen produziert. Da die Wortstellung im Englischen relativ unflexibel ist, hat Lorimor (2007) in ihrer Studie Deklarativsätze (präverbale Konstruktionen)

und Fragesätze (postverbale Konstruktionen) miteinander verglichen. Lorimor diskutiert ihre Ergebnisse im Rahmen von syntaktischen Ansätzen. Ich bin jedoch der Auffassung, dass ihre Ergebnisse als zusätzliche sprachübergreifende Evidenz für Hypothese 1 betrachtet werden können und die Annahme stützen, dass die Wortstellungsasymmetrie bei Subjekt-Verb-Kongruenz auf Verarbeitungsmechanismen zurückzuführen ist.

Hypothese 2 besagt, dass die asymmetrische Kongruenz mit zunehmendem Abstand zwischen Verb und Subjekt häufiger auftritt bzw. schneller verarbeitet wird. Es war nicht möglich, diese Hypothese anhand von Korpusdaten zu überprüfen. Die Ergebnisse des Lesezeit-Experiments liefern jedoch entsprechende Evidenz. In postverbalen Konstruktionen gab es keinen signifikanten Unterschied zwischen symmetrischer und asymmetrischer Kongruenz. Bei erhöhter Distanz zwischen Verb und Subjekt wird asymmetrische Kongruenz jedoch signifikant schneller verarbeitet als symmetrische Kongruenz. Dies deutet darauf hin, dass sich die Präferenz durch den erhöhten Verarbeitungsaufwand zur asymmetrischen Kongruenz hin verschoben hat. In den Grammatikalitätsurteilen wurde kein signifikanter Unterschied zwischen der Kongruenz mit postverbalen Subjektphrasen und solchen mit größerer Distanz zum finiten Verb gefunden. Ich nehme an, dass dieser Unterschied zu subtil ist und möglicherweise nur in Online-Datentypen nachgewiesen werden kann.

In *Hypothese 3* wird angenommen, dass asymmetrische Kongruenz in Online-Datentypen häufiger auftritt oder leichter verarbeitet wird als in Offline-Datentypen. Für diese Hypothese habe ich sowohl in den Korpusdaten als auch in den experimentellen Daten Evidenz gefunden. In geschriebenen Korpora (Offline-Daten) kommt asymmetrische Kongruenz nur selten vor. In gesprochenen Korpora (Online-Daten) kommt asymmetrische Kongruenz häufig vor und wird bei postverbalen Subjektphrasen sogar bevorzugt. Es konnte gezeigt werden, dass asymmetrische Kongruenz in spontansprachlichen Dialogen signifikant häufiger auftritt als in geschriebenen Texten. In den Grammatikalitätsurteilen (Offline-Daten) wird asymmetrische Kongruenz in allen drei Wortstellungstypen meist als ungrammatisch eingestuft. Im Lesezeitexperiment (Online-Daten) wird asymmetrische Kongruenz in postverbalen Konstruktionen gleich schnell verarbeitet wie die symmetrische Kongruenz. Bei erhöhter Distanzstellung

zwischen Verb und Subjekt wird die asymmetrische Kongruenz sogar bevorzugt. Dies zeigt, dass asymmetrische Kongruenz in Lesezeitexperimenten leichter verarbeitet wird als in Beurteilungsaufgaben zur Grammatikalität.

In den experimentellen Untersuchungen wurden nicht nur unterschiedliche Präferenzen für Kongruenz in Online- und Offline-Daten gefunden, es konnten auch unterschiedliche Positionen im Satz lokalisiert werden, die für die Herstellung von Kongruenz relevant sind. Bei der Beurteilungsaufgabe zur Grammatikalität wurden von den Probanden beide Konjunkte für die Bewertung eines Satzes herangezogen. Sätze mit asymmetrischer Kongruenz in postverbalen Konstruktionen wurden in den meisten Fällen beim zweiten Konjunkt abgebrochen. Im Lesezeit-Experiment hingegen trat der Unterschied zwischen symmetrischer und asymmetrischer Kongruenz in Konstruktionen mit entfernt stehendem postverbalem Subjekt lokal am ersten Konjunkt auf. Beim zweiten Konjunkt wurde kein Effekt gefunden. Dies zeigt, dass bei erhöhter Distanz zwischen Verb und Subjekt, d.h. mit erhöhtem Verarbeitungsaufwand, das Sprachsystem lokaler arbeitet, was vermehrt zu asymmetrischer Kongruenz führt.

Darüber hinaus konnte die Annahme zurückgewiesen werden, dass die Probanden elliptische Strukturen für postverbale Konstruktionen mit asymmetrischer Kongruenz aufbauen und phrasale Strukturen für die Subjekt-NP bei symmetrischer Kongruenz. Nach diesem Ansatz werden in postverbalen Konstruktionen längere Lesezeiten für asymmetrische gegenüber symmetrischer Kongruenz erwartet, da es zeitaufwändiger sein sollte, sententiale Strukturen aufzubauen als phrasale Strukturen. Das Lesezeit-Experiment ergab jedoch den gegenteiligen Befund. Außerdem waren die Testsätze so konstruiert, dass es nicht möglich war, eine sententiale Struktur auf inkrementelle Weise aufzubauen (siehe Fußnote 5). Bei der Verarbeitung der Konjunktion verfügten die Probanden nicht über ausreichende Informationen, um eine vollständige sententiale Struktur für das erste Konjunkt aufzubauen. Sie waren daher gezwungen, eine phrasale Struktur für die Subjekt-NP zu konstruieren. Erst am Ende des Satzes verfügt der Parser über die relevanten Informationen. Es wäre zwar sehr unökonomisch, aber dennoch denkbar, dass der Parser am Satzende die phrasale Struktur zugunsten einer sententialen Struktur reanalysiert. Dieser Reanalyseprozess sollte jedoch zu

längeren Lesezeiten für das letzte Segment des Satzes führen. In postverbalen Konstruktionen sollten längere Lesezeiten bei asymmetrischer im Vergleich zu symmetrischer Kongruenz gefunden werden. Dieser Effekt ist im Experiment jedoch nicht aufgetreten. Daraus kann geschlossen werden, dass keine elliptischen Strukturen aufgebaut wurden.

Insgesamt zeigen die in diesem Kapitel beschriebenen Ergebnisse, dass die asymmetrische Kongruenz durch Verarbeitungsmechanismen und den Verarbeitungsaufwand erheblich beeinflusst wird. Diese Annahme wird insbesondere durch den Vergleich von Online- und Offline-Daten gestützt. Das in der Einleitung beschriebene Muster der asymmetrischen Kongruenz kann meiner Auffassung nach elegant durch einen Verarbeitungsansatz erklärt werden, ohne auf syntaktische Asymmetrie-Ansätze rekurrieren zu müssen.

In dem vorgeschlagenen Verarbeitungsansatz sind die unterschiedlichen Muster in präverbalen und postverbalen Konstruktionen auf die Inkrementalität der Sprachverarbeitung von links nach rechts zurückzuführen. Das heißt, die Reihenfolge der Informationen im sprachlichen Input ist ausschlaggebend und führt dazu, dass asymmetrische Kongruenz nur in V-S-Worstellung akzeptiert wird. In postverbalen Konstruktionen handelt es sich um eine scheinbare Optionalität für symmetrische bzw. asymmetrische Kongruenz. Die symmetrische Kongruenz scheint der grammatisch verankerte Normalfall zu sein, was durch die Ergebnisse der Offline-Daten (Grammatikalitätsurteile, Zeitungstexte) gestützt wird. Die asymmetrische Kongruenz hingegen entsteht durch Verarbeitungsdruck. Die Sonderstellung des ersten Konjunks ist lediglich darauf zurückzuführen, dass es näher am Verb steht und zeitlich vor dem zweiten Konjunkt verarbeitet wird. Wird die asymmetrische Kongruenz als Verarbeitungseffekt betrachtet, fällt ein wichtiges Argument für asymmetrische Modelle zur syntaktischen Beschreibung von Koordination weg, wie z.B. für die Adjunktionsanalyse nach Munn (2000) oder den Reduktionsansatz von Aoun et al. (1994).

Kapitel 8

Allgemeine Diskussion

Der Ausgangspunkt dieser Arbeit waren koordinierte Strukturen, die für eine Vielzahl sprachlicher Bereiche eine große Herausforderung darstellen. In der theoretischen Syntax besteht keinerlei Konsens darüber, wie die der Koordination zugrundeliegenden syntaktischen Strukturen beschaffen sind. Weiterhin sind bestehende syntaktische Verarbeitungsmodelle nicht oder nur sehr eingeschränkt in der Lage, die Effekte und Präferenzen beim Sprachverstehen koordinierter Strukturen zu beschreiben. Schließlich sind koordinierte Strukturen eine der größten Fehlerquellen für maschinelle Parsingsysteme. Das Ziel dieser Arbeit bestand darin, die syntaktischen Prozesse bei der Verarbeitung koordinierter Strukturen unter verschiedenen Gesichtspunkten zu untersuchen und auf diese Weise zusätzliche Evidenz für die syntaktische Analyse der Koordination zu erhalten.

Im Mittelpunkt der Untersuchungen stand das Spannungsfeld zwischen den symmetrischen und asymmetrischen Eigenschaften koordinierter Strukturen. Wie in *Kapitel 2* aufgezeigt wurde, unterliegen koordinierte Strukturen grundlegenden Symmetriebedingungen unterschiedlicher Art. Auf der anderen Seite sind jedoch auch jeweils die entsprechenden Asymmetrien zu beobachten. Abhängig davon, ob bei der Modellbildung den symmetrischen oder eher den asymmetrischen Eigenschaften den Vorzug gegeben wurde, führte das in der Vergangenheit zu symmetrischen bzw. zu asymmetrischen Modellen. Dieser scheinbar unauflösbare Widerspruch sollte in dieser Arbeit aus einer neuen Perspektive, nämlich aus der Verarbeitungsperspektive, beleuchtet werden.

Hierfür wurden die symmetrischen und asymmetrischen Eigenschaften in zwei unterschiedliche (A)Symmetrie-Arten unterteilt: a. *Konjunkt-Symmetrie* und b. *Umgebungs-Symmetrie*. Die Konjunkt-Symmetrie betrifft die Relation zwischen den Konjunkten und die Frage, wie parallel die Konjunkte sein müssen. Die Umgebungs-Symmetrie betrifft die Relation der Konjunkte zu ihrer syntaktischen Umgebung und die Frage, ob alle Konjunkte gleichermaßen die Anforderungen der syntaktischen Umgebung erfüllen müssen. Die Ausgangshypothese war, dass sich die beiden (A)Symmetrie-Arten in ihren Verarbeitungsmechanismen unterscheiden.

Zur Konjunkt-Symmetrie: Hinsichtlich der Relation zwischen den Konjunkten (Konjunkt-Symmetrie) gab es bereits experimentelle und theoretische Arbeiten zum Englischen, welche in *Kapitel 3* vorgestellt wurden. Beim Sprachverstehen wurden eine Reihe von Parallelitätseffekten auf phrasaler und sententialer Ebene beobachtet, sowie Präferenzen in ambigen koordinierten Strukturen und Komplexitätseffekte in *either-or*-Konstruktionen. Ich habe aufgezeigt, dass sich die Parallelitätseffekte weder mit allgemeinen, von Koordination unabhängigen, Parsingmechanismen erklären lassen, noch im Rahmen des Copy- α -Modells (Frazier und Clifton, 2001), das speziell für die Verarbeitung koordinierter Strukturen und Ellipsen entwickelt wurde. Die beobachteten Präferenzen in ambigen koordinierten Strukturen können aus verschiedenen allgemeinen Parsingprinzipien abgeleitet werden, die jedoch nur bedingt die fehlenden Komplexitätseffekte bei *'either or'* vorhersagen können. Die fehlenden Komplexitätseffekte lassen sich aus dem Copy- α -Modell ableiten, das wiederum keine Vorhersagen für ambige Strukturen macht. Insgesamt können die beobachteten Verarbeitungseffekte und Präferenzen nicht oder nur sehr eingeschränkt mit den bestehenden Verarbeitungsmodellen erklärt werden. Die beschriebenen Effekte bildeten die Basis für die eigenen Untersuchungen zur Konjunkt-Symmetrie (theoretische Modellierung, Korpusuntersuchung, Simulationen) in den Kapiteln 4–6.

In *Kapitel 4* habe ich ein eigenes Modell – das Iterationsmodell – entwickelt, das Vorhersagen für Lesezeiten und Präferenzen bei der syntaktischen Verarbeitung koordinierter Strukturen macht. Das Modell basiert auf einem neuen Parsingmechanismus, bei dem bereits aufgebaute Strukturen des ersten Konjunks bei der Verarbeitung der weiteren Konjunkte wiederverwendet werden. Die syntaktischen Strukturen, die sich aus diesem Parsingmechanismus ergeben, sind drei-dimensionale Baumstrukturen, wie sie im Rahmen von mehrdimensionalen Modellen vorgeschlagen werden. Die Verarbeitungsmechanismen im Iterationsmodell und die daraus resultierenden Datenstrukturen sind besonders ökonomisch. Das Iterationsmodell macht die korrekten Vorhersagen für alle in Kapitel 3 beschriebenen Verarbeitungseffekte und Präferenzen. Desweiteren habe ich gezeigt, dass sich das Iterationsmodell auch auf die Verarbeitung koordinierter Strukturen bei agrammatischer Aphasie übertragen lässt. Das Iterationsmodell erlaubt die Generierung neuer testbarer Hypothesen für Lesezeiten und Präferenzen bei koordinierten Strukturen.

Eine weitere Fragestellung der vorliegenden Arbeit betraf das Verhältnis von Sprachverstehens- und Sprachproduktions-Mechanismen, das in *Kapitel 5* in Bezug auf koordinierte Strukturen beleuchtet wurde. Nach der Tuning-Hypothese entwickelt der menschliche Parser bestimmte Verarbeitungspräferenzen, weil diese Präferenzen in Form von relativen Häufigkeiten im sprachlichen Input vorhanden sind. Der Status dieser Hypothese ist jedoch bislang ungeklärt. Um dieser Frage nachzugehen, habe ich spontansprachliche Korpusdaten in Bezug auf drei Verarbeitungseffekte statistisch ausgewertet: 1. Präferenz für die Koordination gleicher Kategorien, 2. Parallel-Struktur-Effekt im Phrasenaufbau, 3. Präferenz für NP-Koordination (gegenüber Satz-Koordination). Die Korpusanalysen ergaben, dass alle drei Verarbeitungseffekte in spontansprachlichen Korpora zu finden sind.

Prinzipiell gibt es zwei Möglichkeiten, diese Korrelationen zu erklären. Auf der einen Seite ist es denkbar, dass die Präferenzen beim Sprachverstehen gemäß der Tuning-Hypothese durch die statistischen Muster in natürlicher Sprache entstehen. Auf der anderen Seite könnten die Korrelationen jedoch auch auf gemeinsame zugrundeliegende Mechanismen für die Sprachproduktion und das Sprachverstehen zurückgeführt werden. Gemeinsame Mechanismen würden dazu führen, dass Strukturen, die leichter verarbeitet werden, auch

leichter und damit häufiger produziert werden. Im Rahmen dieser Studie war es nicht möglich, zwischen der Tuning-Hypothese und der Annahme von gemeinsamen zugrundeliegenden Mechanismen zu unterscheiden. Es konnte jedoch gezeigt werden, dass das Verhältnis zwischen Sprachproduktion und Sprachverstehen enger ist als erwartet. Sollte eine generelle Korrelation zwischen Parsingpräferenzen und Korpusfrequenzen bestehen, würde das zu der interessanten Schlussfolgerung führen, dass Korpusdaten verwendet werden können, um Sprachverstehensmodelle zu evaluieren.

Neben den Korrelationen für die oben aufgeführten Effekte habe ich in den Korpusdaten bei der Auswertung des Parallel-Struktur-Effekts außerdem einen Längeneffekt für kurze vs. lange Konjunkte gefunden. Dieser war bislang nicht bekannt. Die Länge des ersten Konjunks scheint ein relevanter Faktor für den Parallel-Struktur-Effekt zu sein, d.h. der Parallel-Struktur-Effekt tritt zumindest in Korpora nur bei kurzen Konjunkten auf. Unter der Voraussetzung, dass Korpusfrequenzen tatsächlich zur Evaluation von Verarbeitungsmodellen herangezogen werden können, liefert der Längeneffekt Evidenz für das Iterationsmodell und gegen das Copy- α -Modell.

In *Kapitel 6* wurde die syntaktische Verarbeitung koordinierter Strukturen im Rahmen eines kognitiven Systems simuliert. Hierfür wurden künstliche neuronale Netze trainiert, selbständig grammatische Regularitäten und Parsingmechanismen aus spontansprachlichen Korpusdaten zu lernen. Syntaktische Regularitäten und Mechanismen wurden auf indirekte Weise mittels der Satzantizipations-Aufgabe erworben. Dabei hatte das Netz die Aufgabe, für eine Sequenz von Wortarten die nachfolgende Wortart vorherzusagen. Es handelt sich hierbei um ein überwacht, aber dennoch natürliches Lernszenario, da der Zielreiz, d.h. das nächste Element im Satz, automatisch bei der Sprachrezeption zur Verfügung steht. Die Satzantizipations-Aufgabe liefert außerdem bei der Verletzung von Erwartungshaltungen indirekte negative Evidenz beim Spracherwerb, wodurch das Lernen einer Grammatik auch ohne die Annahme von angeborenen Prinzipien möglich ist. Für das Training wurde die Elman-Netzwerk-Architektur verwendet, die besonders gut für das Lernen von zeitveränderlichen Sequenzen geeignet ist. Die Trainingsdaten bestanden aus echten spontansprachlichen Korpusdaten. In der Literatur wurde bislang nur von entsprechenden Simulationen mit stark vereinfachten oder künstlichen Trainingsdaten berichtet.

Die Simulationen ergaben, dass ein neuronales Netz durchaus in der Lage ist, grammatische Regularitäten aus echten spontansprachlichen Korpusdaten zu lernen. Auch wenn die Performanz des Netzes auf verschiedene Weise noch verbessert werden kann, habe ich aus den Ergebnissen geschlossen, dass das Netz die Grammatik in ausreichendem Maße erworben hat, um es mit der menschlichen Leseverarbeitung vergleichen zu können. Der Vergleich zwischen Netzwerkperformanz und der menschlichen Leseverarbeitung wurde anhand des Parallel-Struktur-Effekts sowie der Präferenz für NP-Koordination (vs. Satz-Koordination) bei Ambiguitäten durchgeführt. Dabei wurde untersucht, ob Satzsegmente, die in Lesezeit-Experimenten schneller verarbeitet wurden, auch zu geringeren Fehlerraten des Netzes führen. Die Auswertung hinsichtlich des Parallel-Struktur-Effekts ergab eine solche Korrelation. Das Netz zeigte eine Verarbeitungserleichterung bei parallelen koordinierten Strukturen für alle Konjunktlängen. Weiterhin zeigte das Netz einen syntaktischen Kontexteffekt vergleichbar mit der menschlichen Leseverarbeitung. Beim Lesen nimmt die Geschwindigkeit mit zunehmendem Satzkontext zu. Entsprechend nahmen die Fehlerraten im Netzwerk zum Satzende hin ab. Die Auswertung der Präferenz für NP-Koordination (vs. Satz-Koordination) ergab einen marginal signifikanten Effekt. Die Unterschiede zwischen den Fehlerraten bei NP- bzw. Satz-Koordination sind auf den ersten Blick eindeutig. Für einen robusten Effekt sind jedoch vermutlich mehr Trainingsdaten erforderlich.

Während koordinierte Strukturen eine der größten Fehlerquellen für die maschinelle Sprachverarbeitung darstellen, hatte das neuronale Netz interessanterweise keine besonderen Schwierigkeiten bei der Satzantizipation dieser Strukturen. Das Netz zeigte sogar eine Verarbeitungserleichterung bei parallelen koordinierten Strukturen und verhielt sich bezüglich der untersuchten Effekte vergleichbar mit der menschlichen Leseverarbeitung. Interessant ist dabei, dass diese Effekte zu beobachten sind, obwohl das Netz ausschließlich mit Kategoriensequenzen trainiert wurde und keine weiteren linguistischen Vorannahmen integriert wurden. Die beobachteten Verarbeitungseffekte sind emergente Eigenschaften. Insgesamt wurde gezeigt, dass ein neuronales Netz in der Lage ist, ausschließlich aufgrund statistischer Muster im sprachlichen Input Verarbeitungspräferenzen zu entwickeln, wie sie auch in Lesezeit-Experimenten beobachtet werden.

Nachdem die syntaktischen Prozesse in Bezug auf die Konjunkt-Symmetrie unter verschiedenen Gesichtspunkten beleuchtet wurden, möchte ich im Folgenden auf die Umgebungs-Symmetrie eingehen.

Zur Umgebungs-Symmetrie: Die Ausgangshypothese war, dass die asymmetrischen Phänomene im Bereich der Umgebungs-Symmetrie keinen Reflex der Syntax darstellen, sondern eine Folge der inkrementellen Sprachverarbeitung von links nach rechts sind. Entscheidend ist dabei, in welcher Reihenfolge der Parser die relevanten Informationen erhält sowie die jeweiligen Abstände der Konjunkte zu ihrer syntaktischen Umgebung. Möglicherweise ist die Sonderstellung des ersten Konjunks lediglich darauf zurückzuführen, dass das erste Konjunkt zeitlich vor dem zweiten Konjunkt verarbeitet wird und einen geringeren Abstand zu der zuvor verarbeiteten Umgebung aufweist.

Diese Ausgangshypothese wurde in *Kapitel 7* exemplarisch anhand der asymmetrischen Kongruenz im Deutschen untersucht. Hierfür habe ich vier verschiedene Datentypen miteinander verglichen: geschriebene und gesprochene Korpusdaten (Sprachproduktion) sowie Grammatikalitätsurteile und Lesezeiten (Sprachverstehen). Geschriebene Sprache und Grammatikalitätsurteile sind Offline-Datentypen und reflektieren in erster Linie grammatisches Normwissen. Lesezeiten und gesprochene Sprache sind Online-Datentypen und zeichnen sich durch einen höheren Verarbeitungs- und Zeitdruck aus. Verarbeitungseffekte können daher eher in Online-Datentypen beobachtet werden.

Die Untersuchungen ergaben, dass die asymmetrische Kongruenz in geschriebener Sprache (Zeitungstexten) nur selten auftritt. Weiterhin wurde die asymmetrische Kongruenz bei der Grammatikalitätsbeurteilung in den meisten Fällen als ungrammatisch eingestuft. In diesen Offline-Datentypen bestand eine eindeutige Präferenz für symmetrische Kongruenz. Die Untersuchung der Online-Datentypen (Lesezeiten, spontansprachliche Korpusdaten) ergab ein anderes Bild. War das Subjekt im Vorfeld vor dem finiten Verb positioniert, ließ sich eine Präferenz für symmetrische Kongruenz beobachten. In V-S-Wortstellung wurde jedoch die asymmetrische Kongruenz präferiert. Dies wurde insbesondere in Konstruktionen mit zusätzlichem Abstand

zwischen dem finitem Verb und dem postverbalen Subjekt deutlich, d.h. bei erhöhtem Verarbeitungsaufwand. Insgesamt zeigen die Ergebnisse, dass die symmetrische Kongruenz vermutlich der grammatisch verankerte Normalfall ist. Asymmetrische Kongruenz entsteht bei erhöhtem Verarbeitungsdruck und lässt sich daher vor allem in Online-Datentypen beobachten. Ich schließe aus diesen Ergebnissen, dass die asymmetrische Kongruenz keinen Reflex der Syntax darstellt, sondern vielmehr als Verarbeitungseffekt betrachtet werden kann.

Nachdem die syntaktischen Verarbeitungsprozesse sowohl in Bezug auf die Konjunkt-Symmetrie als auch hinsichtlich der Umgebungs-Symmetrie untersucht wurden, möchte ich im Folgenden auf die Implikationen für die syntaktische Analyse koordinierter Strukturen eingehen.

Implikationen für die syntaktische Analyse: Aus der Sprachverarbeitung kann auf verschiedene Weise Evidenz für syntaktische Strukturen gewonnen werden. Einerseits ergibt die Untersuchung von Präferenzen bei der Sprachverarbeitung Aufschluss über den Normalfall einer bestimmten symmetrischen oder asymmetrischen Eigenschaft. Werden die experimentellen Ergebnisse aus der Literatur sowie die eigenen Korpusuntersuchungen zusammengefasst, ergibt sich hinsichtlich der Konjunkt-Symmetrie das folgende Bild: Es besteht eine generelle Präferenz für Symmetrie hinsichtlich der Koordination von gleichen Kategorien vs. cross-kategorialer Koordination sowie bezüglich der Strukturgleichheit der Konjunkte im Phrasenaufbau. Diese Präferenz zeigt sich sowohl beim Sprachverstehen (Interpretation, Lesezeiten) als auch in der Sprachproduktion (spontansprachliche Korpusdaten) auf phrasaler und sententialer Ebene. Hinsichtlich der Umgebungs-Symmetrie zeigen die eigenen Experimente und Korpusauswertungen ebenfalls eine grundlegende Präferenz für symmetrische Kongruenz in allen Datentypen. Die asymmetrische Kongruenz wird nur unter erhöhtem Verarbeitungsdruck bei postverbaler Stellung des Subjekts präferiert. Diese Ergebnisse lassen darauf schließen, dass die symmetrischen Eigenschaften bei koordinierten Strukturen den Normalfall darstellen. Diese Präferenzen liefern Hinweise dafür, ein symmetrisches syntaktisches Modell vorzuziehen.

Weitere Evidenz ergibt sich auf indirekte Weise aus den agrammatischen Daten und durch die Modellierung der Lesezeit-Effekte und Präferenzen (siehe Kapitel 4). Die Sprachverstehensdaten bei agrammatischen Aphasiepatienten können für die Evaluation einer Defizit-Theorie verwendet werden. Wenn im Rahmen dieser Defizit-Theorie verschiedene syntaktische Analysen zu unterschiedlichen Vorhersagen für das agrammatische Sprachverständnis führen, kann indirekt zwischen diesen Analysen unterschieden werden. In Kapitel 4 wurde Evidenz für die Defizit-Theorie von Ouhalla (1993) auf der Basis von zwei Experimenten gefunden. Aufgrund der Ergebnisse von beiden Untersuchungen kann der Reduktionsansatz für die Analyse von NP- (/DP)-Koordination ausgeschlossen werden. Zwischen einer phrasalen und einer drei-dimensionalen Analyse kann mit diesen Daten nicht unterschieden werden. Diese Unterscheidung wird allerdings durch die Modellierung der Verarbeitungseffekte möglich.

Mit Hilfe von Lesezeiten und Präferenzen können wiederum Verarbeitungsmodelle evaluiert werden, die ebenfalls auf syntaktischen Theorien basieren. Wenn verschiedene syntaktische Analysen zu unterschiedlichen Vorhersagen z.B. für Lesezeiten führen, kann indirekt zwischen diesen Analysen unterschieden werden. In Kapitel 4 habe ich gezeigt, dass eine Vielzahl von Verarbeitungseffekten im Iterationsmodell auf Basis von drei-dimensionalen syntaktischen Strukturen erklärt werden kann. Viele dieser Effekte lassen sich nicht oder nur eingeschränkt aus Verarbeitungsmodellen ableiten, die auf zwei-dimensionalen Strukturen basieren. Die Modellierung der Verarbeitungseffekte bei Koordination liefert somit indirekte Evidenz für eine drei-dimensionale syntaktische Analyse der Koordination.

Schließlich konnte exemplarisch anhand der asymmetrischen Kongruenz Evidenz für meine Ausgangshypothese gewonnen werden, dass die beobachteten Asymmetrien aus dem Bereich der Umgebungs-Symmetrie vermutlich Verarbeitungseffekte darstellen. Damit fällt ein wichtiges Argument für asymmetrische Modelle zur syntaktischen Beschreibung von Koordination weg. Sollte es möglich sein, auch die anderen asymmetrischen Phänomene der Umgebungs-Symmetrie außerhalb der Syntax zu beschreiben, könnte der scheinbare Widerspruch zwischen Symmetrie und Asymmetrie aufgelöst werden.

Insgesamt sprechen die Ergebnisse für eine symmetrische und vermutlich mehrdimensionale Analyse koordinierter Strukturen. Drei-dimensionale Baumstrukturen in Kombination mit dem Iterationsmodell machen die richtigen Vorhersagen für Lesezeiten und Präferenzen bei der syntaktischen Verarbeitung. In einem drei-dimensionalen Modell müssen die Konjunkte nicht zwingend strukturgleich sein. Im parallelen Fall kann jedoch eine größere Anzahl von syntaktischen Knoten geteilt werden, was zu einer schnelleren Verarbeitung führt. Dies erklärt die Symmetrie-Präferenz bezüglich der Konjunkt-Symmetrie. In einem mehrdimensionalen Modell haben alle Konjunkte den gleichen syntaktischen Status. Das heißt, die Asymmetrien bezüglich der Umgebungs-Symmetrie werden nicht vorhergesagt. Dies ist jedoch unproblematisch, da ich zumindest für die asymmetrische Kongruenz gezeigt habe, dass diese Asymmetrien gar keinen Reflex der Syntax darstellen.

Literaturverzeichnis

- Abney, Steven (1987): *The English Noun Phrase in its Sentential Aspect*, Dissertation, MIT, Cambridge.
- Anderson, Britt (1999): „Kohonen Neural Networks and Language“, *Brain and Language* **70**, S. 86–94.
- Aoun, Joseph, Elabbas Benmamoun und Dominique Sportiche (1994): „Agreement, word order and conjunction in some varieties of Arabic“, *Linguistic Inquiry* **25** (2), S. 195–220.
- Apel, Jens, Pia Knoeferle und Matthew W. Crocker (2007): „Processing Parallel Structure: Evidence from Eye-Tracking and a Computational Model“, in S. Vosniadou, D. Kayser und A. Protopapas (Hrsg.), *Proceedings of the 2nd European Cognitive Science Conference*, London: Taylor and Francis, S. 125–131.
- Argamon, Shlomo, Ido Dagan und Yuval Krymolowski (1998): „A memory-based approach to learning shallow natural language patterns“, in *Proceedings of the 36th Annual Meeting of the ACL and 17th International Conference on Computational linguistics*, Association for Computational Linguistics, S. 67–73.
- Babyonyshev, Maria (1997): *Structural connection in syntax and processing : Studies in Russian and Japanese*, Dissertation, MIT.
- Bach, Emmon, Colin Brown und William Marslen-Wilson (1986): „Crossed and nested dependencies in German and Dutch: A psycholinguistic study“, *Language and Cognitive Processes* **1**, S. 249–262.

- Baker, Carl L. (1979): „Syntactic theory and the projection problem“, *Linguistic Inquiry* **10**, S. 533–581.
- Blodgett, Allison und Julie E. Boland (1998): „Understanding BOTH in Coordinate Structures: The Osu Subject Pool both Surprised Allison and Julie!“, in *Poster session at the 11th CUNY Conference on Human Sentence Processing*, New Brunswick, NJ.
- Branigan, Holly P., Martin J. Pickering, Simon P. Liversedge, Andrew J. Stewart und Thomas P. Urbach (1995): „Syntactic priming: Investigating the mental representation of language“, *Journal of Psycholinguistic Research* **24** (6), S. 489–506.
- Brill, Eric, David Magerman, Mitchell Marcus und Beatrice Santorini (1990): „Deducing linguistic structure from the statistics of large corpora“, in *DARPA Speech and Natural Language Workshop*, Morgan Kaufmann.
- Brown, Peter F., Peter V. Desouza, Robert L. Mercer, Vincent J. Della Pietra und Jenifer C. Lai (1992): „Class-based n-gram models of natural language“, *Computational linguistics* **18** (4), S. 467–479.
- Brysbaert, Marc und Don C. Mitchell (1996): „Modifier attachment in sentence processing: Evidence from Dutch“, *The Quarterly Journal of Experimental Psychology* **49A**, S. 664–695.
- Carlson, Katy (2001): „The Effects of Parallelism and Prosody in the Processing of Gapping Structures“, *Language and Speech* **44** (1), S. 1–26.
- Chalmers, David J. (1990): „Syntactic transformations on distributed representations“, *Connection Science* **2**, S. 53–62.
- Charniak, Eugene (1996): „Tree-Bank Grammars“, in *Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference 1996*, S. 1031–1036.
- Charniak, Eugene (1997): „Statistical Parsing with a context-free grammar and word statistics“, in *Proceedings of the Fourteenth National Conference on Artificial Intelligence 1997*, AAAI Press/MIT Press.

- Chomsky, Noam (1956): „Three models for the description of language“, *IRE Transactions on Information Theory* **2**, S. 113–124.
- Chomsky, Noam (1957): *Syntactic Structures*, The Hague: Mouton.
- Chomsky, Noam (1980): *Rules and Representations*, New York: Columbia University Press.
- Chomsky, Noam (1981): *Lectures on Government and Binding*, Foris: Dordrecht.
- Chomsky, Noam (1986): *Barriers*, Cambridge: MIT Press.
- Christiansen, Morten H. (1994): *Infinite languages, finite minds: Connectionism, learning and linguistic structure*, Dissertation, University of Edinburgh.
- Christiansen, Morten H. und Maryellen C. MacDonald (1999): *Processing of Recursive Sentence Structure: Testing Predictions from a Connectionist Model*. Manuskript.
- Christiansen, Morten H. und Nick Chater (1999): „Toward a Connectionist Model of Recursion in Human Linguistic Performance“, *Cognitive Science* **23** (2), S. 157–205.
- Christiansen, Morten H. und Nick Chater (2001): „Finite Models of Infinite Language: A Connectionist Approach to Recursion“, in M. H. Christiansen und N. Chater (Hrsg.), *Connectionist Psycholinguistics*, Westport, Connecticut: Ablex Publishing, S. 138–176.
- Citko, Barbara (2004): „Agreement Asymmetries in Coordinate Structures“, in *Formal Approaches to Slavic Linguistics: The Ottawa Meeting*, Ann Arbor, MI: Michigan Slavic Publications.
- Collins, Michael (1999): *Head-Driven Statistical Models for Natural Language Parsing*, Dissertation, University of Pennsylvania.
- Crocker, Matthew W. (1996): *Computational Psycholinguistics: An Interdisciplinary Approach to the Study of Language*, Band 20 in *Studies in Theoretical Psycholinguistics*, Dordrecht: Kluwer.

- Cuetos, Fernando und Don C. Mitchell (1988): „Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish“, *Cognition* **30**, S. 73–105.
- Cuetos, Fernando, Don C. Mitchell und Martin M. B. Corley (1996): „Parsing in different languages“, in M. Carreiras, J. E. Garcia-Albea und N. Sebastian-Galles (Hrsg.), *Language Processing in Spanish*, Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., S. 145–187.
- De Vries, Mark (2003): „Three-Dimensional Grammar“, *Linguistics in the Netherlands* **20**, S. 201–212.
- Dell, Gary S., Cornell Juliano und Anita Govindjee (1993): „Structure and content in language production: A theory of frame constraints in phonological speech errors“, *Cognitive Science* **17**, S. 149–195.
- Dell, Gary S., Franklin Chang und Zenzi M. Griffin (2001): „Connectionist Models of Language Production: Lexical Access and Grammatical Encoding“, in M. H. Christiansen und N. Chater (Hrsg.), *Connectionist Psycholinguistics*, Westport, Connecticut: Ablex Publishing, S. 212–243.
- den Dikken, Marcel (2006): „Either-float and the syntax of co-or-dination“, *Natural Language and Linguistic Theory* **24**, S. 689–749.
- Desmet, Timothy und Edward Gibson (2003): „Disambiguation preferences and corpus frequencies in noun phrase conjunction“, *Journal of Memory and Language* **49**, S. 353–374.
- Desmet, Timothy, Marc Brysbaert und Constantijn De Baecke (2002): „The correspondence between sentence production and corpus frequencies in modifier attachment“, *The Quarterly Journal of Experimental Psychology* **55A** (3), S. 879–896.
- Ehrich, Veronika (2007): „Der Bloße Singular in koordinativen Verknüpfungen“, *Neue Beiträge zur Germanistik. Internationale Ausgabe von „Doitsu Bungaku“* **6** (3), S. 9–30.

- Ehrlich, Susan F. und Keith Rayner (1981): „Contextual Effects on Word Perception and Eye Movements during Reading“, *Journal of Verbal Learning and Verbal Behavior* **20**, S. 641–655.
- Elman, Jeffrey L. (1990): „Finding structure in time“, *Cognitive Science* **14**, S. 179–211.
- Elman, Jeffrey L. (1991): „Distributed representation, simple recurrent networks, and grammatical structure“, *Machine Learning* **7**, S. 195–225.
- Elman, Jeffrey L. (1993): „Learning and development in neural networks: The importance of starting small“, *Cognition* **48**, S. 71–99.
- Engelhardt, Paul E. und Fernanda Ferreira (2007): „Processing Coordination Ambiguity: Frequency and context do not override the initial preference for syntactic simplicity“, in *Proceedings of the AMLaP Conference 2007*, Turku, Finland, S. 193.
- Frazier, Lyn (1979): *On Comprehending Sentences: Syntactic Parsing Strategies*, Dissertation, University of Connecticut, Storrs.
- Frazier, Lyn (1987a): „Sentence processing: A tutorial review“, in M. Coltheart (Hrsg.), *Attention and performance XII: The psychology of reading*, Hove / London / Hillsdale: Lawrence Erlbaum, S. 559–586.
- Frazier, Lyn (1987b): „Syntactic Processing: Evidence from Dutch“, *Natural Language and Linguistic Theory* **5**, S. 519–559.
- Frazier, Lyn (1990): „Exploring the architecture of the language-processing system“, in G. Altmann (Hrsg.), *Cognitive Models of Speech Processing*, Cambridge, Mass.: MIT Press, S. 409–433.
- Frazier, Lyn, Alan Munn und Charles Clifton (2000): „Processing Coordinate Structures“, *Journal of Psycholinguistic Research* **29** (4), S. 343–370.
- Frazier, Lyn und Charles Clifton (1996): *Construal*, MA: MIT Press.
- Frazier, Lyn und Charles Clifton (1997): „Construal: Overview, motivation, and some new evidence“, *Journal of Psycholinguistic Research* **26**, S. 277–295.

- Frazier, Lyn und Charles Clifton (2000): „On bound variable interpretations: The LF-only hypothesis“, *Journal of Psycholinguistic Research* **29**, S. 125–140.
- Frazier, Lyn und Charles Clifton (2001): „Parsing Coordinates and Ellipsis: Copy α “, *Syntax* **4** (1), S. 1–22.
- Frazier, Lyn und Keith Rayner (1982): „Making and Correcting Errors during Sentence Comprehension: Eye Movements in the Analysis of Structurally Ambiguous Sentences“, *Cognitive Psychology* **14**, S. 178–210.
- Frazier, Lyn, Lori Taft, Tom Roeper, Charles Clifton und Kate Ehrlich (1984): „Parallel Structure: A Source of Facilitation in Sentence Comprehension“, *Memory and Cognition* **12** (5), S. 421–430.
- Friederici, Angela D. (2002): „Towards a neural basis of auditory sentence processing“, *TRENDS in Cognitive Sciences* **6** (2), S. 78–84.
- Friederici, Angela D., Erdmut Pfeifer und Anja Hahne (1993): „Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations“, *Cognitive Brain Research* **1**, S. 183–192.
- Friedmann, Na'ama (2001): „Agrammatism and Psychological Reality of the Syntactic Tree“, *Journal of Psycholinguistic Research* **30** (1), S. 71–90.
- Friedmann, Na'ama und Yosef Grodzinsky (1997): „Tense and Agreement in Agrammatic Production: Pruning the Syntactic Tree“, *Brain and Language* **56**, S. 397–425.
- Gaskell, M. Gareth und William D. Marslen-Wilson (1997): „Integrating Form and Meaning: A Distributed Model of Speech Perception“, *Language and Cognitive Processes* **12** (5/6), S. 613–656.
- Gaskell, M. Gareth und William D. Marslen-Wilson (2001): „Simulating Parallel Activation in Spoken Word Recognition“, in M. H. Christiansen und N. Chater (Hrsg.), *Connectionist Psycholinguistics*, Westport, Connecticut: Ablex Publishing, S. 76–105.

- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum und Ivan Sag (1985): *Generalized Phrase Structure Grammar*, Oxford: Basil Blackwell.
- Gibson, Edward (1991): *A computational theory of human linguistic processing: Memory limitations and processing breakdown*, Dissertation, Carnegie Mellon University, Pittsburgh, PA.
- Gibson, Edward und Carson T. Schütze (1999): „Disambiguation Preferences in Noun Phrase Conjunction Do Not Mirror Corpus Frequency“, *Journal of Memory and Language* **40**, S. 263–279.
- Gibson, Edward, Carson T. Schütze und Ariel Salomon (1996): „The Relationship Between the Frequency and Processing Complexity of Linguistic Structure“, *Journal of Psycholinguistic Research* **25** (1), S. 59–92.
- Gibson, Edward, Neal Pearlmutter, Enriqueta Canseco-Gonzalez und Gregory Hickok (1996): „Recency preference in the human sentence processing mechanism“, *Cognition* **59**, S. 23–59.
- Gold, E. Mark (1967): „Language identification in the limit“, *Information and control* **10** (5), S. 447–474.
- Goldsmith, John (1985): „A principled exception to the Coordinate Structure Constraint“, *Chicago Linguistic Society* **21**, S. 133–143.
- Goodall, Grant (1987): *Parallel Structures in Syntax: Coordination, Causatives, and Restructuring*, Cambridge University Press.
- Grodzinsky, Yosef (1995): „A restrictive theory of trace deletion in agrammatism“, *Brain and Language* **51**, S. 26–51.
- Hagoort, Peter (2003): „How the brain solves the binding problem for language: A neurocomputational model of syntactic processing“, *NeuroImage* **20**, S. 18–29.
- Hagoort, Peter (2006): „On Broca, Brain, and Binding“, in Y. Grodzinsky und K. Amunts (Hrsg.), *Broca's Region*, Oxford University Press.
- Hahne, Anja und Angela D. Friederici (1999): „Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes“, *Journal of Cognitive Neuroscience* **11**, S. 194–205.

- Hartmann, Katharina (2002): *Right Node Raising and Gapping. Interface Conditions on Prosodic Deletion*, Philadelphia/Amsterdam: Benjamins.
- Hebb, Donald O. (1949): *The Organization of Behavior: A Neurophysiological Theory*, Wiley, New York.
- Henstra, Judith-Ann (1996): *On parsing of syntactically ambiguous sentences: Coordination and relative clause attachment*, Dissertation, University of Sussex.
- Heycock, Caroline und Anthony Kroch (1994): „Verb Movement and Coordination in a Dynamic Theory of Licensing“, *The Linguistic Review* **11**, S. 257–284.
- Heycock, Caroline und Roberto Zamparelli (2003): „Coordinated Bare Definites“, *Linguistic Inquiry* **34** (3), S. 443–469.
- Hinrichs, Erhard W., Julia Bartels, Yasuhiro Kawata, Valia Kordoni und Heike Telljohann (2000): „The Tübingen Treebanks for Spoken German, English, and Japanese“, in *Verbmobil: Foundations of Speech-to-Speech Translation, Artificial Intelligence*, Berlin, Heidelberg, New York: Springer Verlag, S. 552–576.
- Hoeks, John C. J., Petra Hendriks, Wietske Vonk, Colin M. Brown und Peter Hagoort (2006): „Processing the noun phrase versus sentence coordination ambiguity: Thematic information does not completely eliminate processing difficulty“, *The Quarterly Journal of Experimental Psychology* **59** (9), S. 1581–1599.
- Hoeks, John C. J., Wietske Vonk und Herbert Schriefers (2002): „Processing Coordinated Structures in Context: The Effect of Topic-Structure on Ambiguity Resolution“, *Journal of Memory and Language* **46**, S. 99–119.
- Höhle, Tilman N. (1983): *Subjektlücken in Koordinationen*. Manuskript, Tübingen.
- Höhle, Tilman N. (1990): „Assumptions about Asymmetric Coordination in German“, in J. Mascaró und M. Nespó (Hrsg.), *Grammar in Progress. Glow Essays for Henk van Riemsdijk*, Dordrecht: Foris, S. 221–235.

- Höhle, Tilman N. (1991): „On Reconstruction and Coordination“, in H. Haider und K. Netter (Hrsg.), *Representation and Derivation in the Theory of Grammar*, Dordrecht: Kluwer, S. 139–197.
- Hornik, Kurt, Maxwell Stinchcombe und Halbert White (1989): „Multilayer Feedforward Networks are Universal Approximators“, *Neural Networks* **2**, S. 359–366.
- Jackendoff, Ray (1977): *X' Syntax*, Cambridge, MA: MIT Press.
- Johannessen, Janne Bondi (1998): *Coordination*, Oxford: Oxford University Press.
- Kaan, Edith, Anthony Harris, Edward Gibson und Phillip Holcomb (2000): „The P600 as an index of syntactic integration difficulty“, *Language and Cognitive Processes* **15** (2), S. 159–201.
- Kayne, Richard (1994): *The Antisymmetry of Syntax*, Cambridge, Mass.: MIT Press.
- Klein, Wolfgang (1993): „Ellipse“, in J. Jacobs, A. von Stechow, W. Sternefeld und T. Vennemann (Hrsg.), *Syntax Ein Handbuch zeitgeössischer Forschung. Syntax An International Handbook of Contemporary Research*, Berlin: De Gruyter, S. 763–799.
- Kohonen, Teuvo (1984): *Self-Organization and Associative Memory*, Springer Series in Information Sciences, Springer-Verlag.
- Kohonen, Teuvo (2001): *Self-Organizing Maps*, Band 30 in *Springer Series in Information Sciences*, Springer-Verlag.
- Konieczny, Lars (1996): *Human Sentence Processing: A Semantics-Oriented Parsing Approach*, Dissertation, Universität Freiburg. IIG-Berichte 3/96.
- Kordoni, Valia (2000): *Stylebook for the English Treebank in VERBMOBIL*, Verbmobil Report No. 241, Seminar für Sprachwissenschaft, Universität Tübingen.

- Kuczaj, Stan A. (1977): „The acquisition of regular and irregular past tense forms“, *Journal of Verbal Learning and Verbal Behavior* **16**, S. 589–600.
- Lakoff, George (1986): „Frame Semantic Control of the Coordinate Structure Constraint“, in A. M. Farley, P. T. Farley und K.-E. McCoullough (Hrsg.), *Papers from the Chicago Linguistic Society*, Band 22, S. 152–167.
- Lakoff, George und Stanley Peters (1969): „Phrasal Conjunction and Symmetric Predicates“, in D. A. Reibel und S. A. Schane (Hrsg.), *Modern studies in English: Readings in Transformational Grammar*, Prentice-Hall, Englewood Cliffs, New Jersey, S. 113–142.
- Lang, Ewald (1977): *Semantik der koordinativen Verknüpfung*, Akademie-Verlag, Berlin.
- Larson, Richard K. (1985): „On the syntax of disjunction scope“, *Natural Language and Linguistic Theory* **3**, S. 217–264.
- Lezius, Wolfgang (2002): „TIGERSearch - Ein Suchwerkzeug für Baumbanken“, in S. Busemann (Hrsg.), *Proceedings der 6. Konferenz zur Verarbeitung natürlicher Sprache (KONVENS 2002)*, Saarbrücken.
- Lorimor, Heidi (2007): *Conjunctions and grammatical agreement*, Dissertation, University of Illinois at Urbana-Champaign.
- MacDonald, Maryellen C. und Morten H. Christiansen (2002): „Reassessing Working Memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996)“, *Psychological Review* **109** (1), S. 35–54.
- Manning, Christopher D. und Hinrich Schütze (1999): *Foundations of statistical natural language processing*, MIT Press: Cambridge, MA.
- Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen und Fei Xu (1992): *Overregularization in language acquisition*, Band 57 [Serial No. 228] in *Monographs of the Society for Research in Child Development*.

- Marcus, Mitchell P., Beatrice Santorini und Mary Ann Marcinkiewicz (1993): „Building a Large Annotated Corpus of English: The Penn Treebank“, *Computational Linguistics* **19** (2).
- Marks, Lawrence E. (1968): „Scaling of grammaticalness of self-embedded English sentences“, *Journal of Verbal Learning and Verbal Behavior* **7**, S. 965–967.
- Marten, Lutz (2005): „The dynamics of agreement and conjunction“, *Lingua* **115**, S. 527–547.
- McClelland, James L., David E. Rumelhart und Geoffrey E. Hinton (1986): „The Appeal of Parallel Distributed Processing“, in D. E. Rumelhart und J. L. McClelland (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Band 1: Foundations, Cambridge, MA: MIT Press, S. 3–44.
- Menn, Lise und Loraine K. Obler (1990): „Methodology: Data collection, presentation and guide to interpretation“, in L. Menn und L. K. Obler (Hrsg.), *Agrammatic Aphasia: A Cross-Language Narrative Sourcebook*, Band I, Amsterdam: Benjamins, S. 13–35.
- Miller, George A. (1956): „The Magical Number Seven, Plus or Minus Two: Some limits on our capacity for processing information“, *The Psychological Review* **63**, S. 81–97.
- Mitchell, Don C., Fernando Cuetos, Martin M. B. Corley und Marc Brysbaert (1995): „Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records“, *Journal of Psycholinguistic Research* **24**, S. 469–488.
- Mitchell, Don C. und Marc Brysbaert (1998): „Challenges to recent theories of crosslinguistic variation in parsing: Evidence from Dutch“, in D. Hillert (Hrsg.), *Syntax and semantics: A crosslinguistic perspective*, San Diego, CA: Academic Press, S. 313–335.

- Mitchell, Don C., Marc Brysbaert, Stefan Grondelaers und Piet Swanepoel (2000): „Modifier attachment in Dutch: Testing aspects of Construal Theory“, in A. Kennedy, R. Radach, D. Heller und J. Pynte (Hrsg.), *Reading as a perceptual process*, Oxford: Elsevier, S. 493–516.
- Moltmann, Friederike (1992): *Coordination and Comparatives*, Dissertation, MIT, Cambridge Mass.
- Muadz, H. (1991): *A Planar Theory of Coordination*, Dissertation, University of Arizona, Tucson.
- Munn, Alan (1993): *Topics in the syntax and semantics of coordinate structures*, Dissertation, University of Maryland, College Park.
- Munn, Alan (1999): „First conjunct agreement: against a biclausal analysis“, *Linguistic Inquiry* **30**, S. 643–668.
- Munn, Alan (2000): „Three Types of Coordination Asymmetries“, in K. Schwabe und N. Zhang (Hrsg.), *Ellipsis in Conjunction*, Tübingen: Niemayer Verlag.
- Ouhalla, Jamal (1993): „Functional Categories, Agrammatism and Language Acquisition“, *Linguistische Berichte* **143**, S. 3–36.
- Pawlow, Iwan (1905): *Klassische Konditionierung nach Pawlow*. Lehrtextsammlung der Abteilung für Pädagogik & Pädagogische Psychologie am Institut für Pädagogik und Psychologie der Johannes Kepler Universität Linz.
- Pinker, Steven (1989): *Learnability and cognition*, Cambridge, MA: MIT Press.
- Plunkett, Kim und Virginia Marchman (1991): „U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition“, *Cognition* **38**, S. 43–102.
- Pollack, Jordan B. (1990): „Recursive distributed representations“, *Artificial Intelligence* **46**, S. 77–105.
- Postal, Paul (1974): *On Raising*, Mass.: MIT Press.

- Progovac, Ljiljana (1998a): „Structure for coordination – Part I“, *Glott International* **3** (7), S. 3–6.
- Progovac, Ljiljana (1998b): „Structure for coordination – Part II“, *Glott International* **3** (8), S. 3–9.
- Rayner, Keith und Lyn Frazier (1987): „Parsing Temporarily Ambiguous Complements“, *The Quarterly Journal of Experimental Psychology* **39A**, S. 657–673.
- Reich, Ingo (2009): *Asymmetrische Koordination im Deutschen*, Band 75 in *Studien zur deutschen Grammatik*, Stauffenburg Verlag.
- Rosenblatt, Frank (1962): *Principles of neurodynamics*, New York: Spartan Books.
- Ross, John R. (1967): *Constraints in Variables in Syntax*, Dissertation, MIT.
- Ross, John R. (1970): „Gapping and the Order of Constituents“, in M. Bierwisch und M. Heidolph (Hrsg.), *Progress in Linguistics*, The Hague: Mouton, S. 249–259.
- Rumelhart, David E., Geoffrey E. Hinton und James L. McClelland (1986): „A General Framework for Parallel Distributed Processing“, in D. E. Rumelhart und J. L. McClelland (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Band 1: Foundations, Cambridge, MA: MIT Press, S. 45–76.
- Rumelhart, David E., Geoffrey E. Hinton und Ronald J. Williams (1986): „Learning internal representations by error propagation“, in D. E. Rumelhart und J. L. McClelland (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Band 1: Foundations, Cambridge, MA: MIT Press, S. 318–362.
- Rumelhart, David E. und James L. McClelland (1986): „On learning of past tenses of English verbs“, in J. L. McClelland und D. E. Rumelhart (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Band 2: Psychological and biological models, Cambridge, MA: MIT Press, S. 216–271.

- Saffran, Jenny R., Richard N. Aslin und Elissa L. Newport (1996): „Statistical Learning by 8-Month-Old Infants“, *Science* **274**, S. 1926–1928.
- Sag, Ivan, Gerald Gazdar, Thomas Wasow und Steven Weisler (1985): „Coordination and How to Distinguish Categories“, *Natural Language and Linguistic Theory* **3**, S. 117–171.
- Sathya, R. und Annamma Abraham (2013): „Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification“, *International Journal of Advanced Research in Artificial Intelligence* **2** (2), S. 34–38.
- Schachter, Paul (1977): „Constraints on Coordination“, *Language* **53**, S. 86–103.
- Schepman, Astrid (1997): *Prosody and on-line parsing*, Dissertation, University of Sussex.
- Schwarz, Bernhard (1999): „On the Syntax of *Either...Or*“, *Natural Language and Linguistic Theory* **17** (2), S. 339–370.
- Steiner, Ilona (2003a): „Parsing Syntactic Redundancies in Coordinate Structures“, in F. Schmalhofer, R. Young und G. Katz (Hrsg.), *Proceedings of EuroCogSci 03 – The European Cognitive Science Conference 2003*, Lawrence Erlbaum, S. 443.
- Steiner, Ilona (2003b): „Recycling Structure: A New Dimension in the Processing of Coordination“, in *Proceedings of the 9th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP-2003)*, S. 107–108.
- Steiner, Ilona (2005): „On the syntax of DP-coordination: Combining evidence from reading-time studies and agrammatic comprehension“, in S. Kepser und M. Reis (Hrsg.), *Linguistic Evidence – Empirical, Theoretical and Computational Perspectives*, Band 85 in *Studies in Generative Grammar*, Berlin: Mouton de Gruyter, S. 507–527.
- Steiner, Ilona (2007a): „Coordinate Structures: On the Relationship between Parsing Preferences and Corpus Frequencies“, in S. Featherston und W. Sternefeld (Hrsg.), *Roots: Linguistics in Search of its Evidential Base*, Band 96 in *Studies in Generative Grammar*, Berlin: De Gruyter, S. 341–360.

- Steiner, Ilona (2007b): *Simulationen mit künstlichen neuronalen Netzen*. Manuskript, Universität Tübingen.
- Steiner, Ilona (2009): „Partial agreement in German: A processing issue?“, in S. Winkler und S. Featherston (Hrsg.), *The fruits of empirical linguistics*, Band 2: Product, Berlin: De Gruyter, S. 239–260.
- Telljohann, Heike, Erhard W. Hinrichs, Sandra Kübler und Heike Zinsmeister (2006): *Stylebook for the Tübingen Treebank of Written German (TüBa-D/Z)*, Technischer Bericht, Seminar für Sprachwissenschaft, Universität Tübingen.
- Towsey, Michael, Joachim Diederich, Ingo Schellhammer, Stephan Chalup und Claudia Brugman (1998): „Natural Language Learning by Recurrent Neural Networks: A Comparison with Probabilistic Approaches“, in D. M. W. Powers (Hrsg.), *NeMLaP3/CoNLL98: New Methods in Language Processing and Computational Natural Language Learning*, ACL, S. 3–10.
- Tulving, Endel, George Mandler und Ruth Bauml (1964): „Interaction of two sources of information in tachistoscopic word recognition“, *Canadian Journal of Psychology* **18** (1), S. 62–71.
- van Oirsouw, Robert (1987): *The Syntax of Coordination*, London: Croom Helm.
- van Oirsouw, Robert (1993): „Coordination“, in J. Jacobs, A. von Stechow, W. Sternefeld und T. Vennemann (Hrsg.), *Syntax – Ein internationales Handbuch zeitgenössischer Forschung. Syntax - An International Handbook of Contemporary Research*, Berlin / New York: De Gruyter, S. 748–763.
- Wesche, Birgit (1995): *Symmetric Coordination - An Alternative Theory of Phrase Structure*, Tübingen: Niemeyer.
- Wilder, Chris (1994): „Coordination, ATB and Ellipsis“, in J.-W. Zwart (Hrsg.), *Minimalism and Kayne's Antisymmetry Hypothesis*, Groninger Arbeiten zur germanistischen Linguistik (GAGL) 37, Rijksuniversiteit Groningen, German Department, S. 291–331.

- Wilder, Chris (1999): „Right Node Raising and the LCA“, in S. Bird, A. Carnie, J. Haugen und P. Norquest (Hrsg.), *Proceedings of the 18th West Coast Conference on Formal Linguistics (WCCFL)*, Somerville, MA: Cascadilla Press, S. 586–598.
- Williams, Edwin (1978): „Across the Board Rule Application“, *Linguistic Inquiry* **9**, S. 31–43.
- Wunderlich, Dieter (1988): „Some Problems of Coordination in German“, in U. Reyle und C. Rohrer (Hrsg.), *Natural Language Parsing and Linguistic Theories*, Band 35 in *Studies in Linguistics and Philosophy*, Dordrecht: Reidel, S. 289–316.
- Zell, Andreas (2000): *Simulation Neuronaler Netze*, Oldenbourg Verlag.
- Zell, Andreas, G. Mamier, M. Vogt, N. Mache, R. Hübner, S. Döring, K.-U. Herrmann, T. Soye, M. Schmalzl, T. Sommer, A. Hatzigeorgiou, D. Posselt, T. Schreiner, B. Kett, G. Clemente, J. Wieland und J. Gatter (1998): *SNNS – Stuttgart Neural Network Simulator, User Manual, Version 4.2*, University of Tübingen, Wilhelm-Schickard-Institute for Computer Science.
- Zinsmeister, Heike (2006): „Treebank data as linguistic evidence – Coordination in TüBa-D/Z“, in *Pre-Proceedings of the International Conference on Linguistic Evidence 2006*, S. 210–212.
- Zoerner, Ed (1995): „Conjunction as a Case Feature-Checker“, *Berkeley Linguistics Society (BLS)* **21**.