

***Ereigniskonzeptualisierung durch sprachliche Struktur: Eine
empirische Untersuchung im Chinesischen und Deutschen***

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

der Neuphilologischen Fakultät der Ruprecht-Karls-Universität Heidelberg

vorgelegt von

Lai Wei (魏莱)

aus Hubei, VR China

Erstgutachter: PD Dr. Johannes Gerwien

Zweitgutachterin: Prof. Dr. Christiane von Stutterheim

Heidelberg, im August 2025

Danksagung

Diese Dissertation wäre ohne die Unterstützung und Ermutigung vieler Menschen nicht möglich gewesen.

Zuallererst möchte ich mich herzlich bei meinem Betreuer, Herrn PD Dr. Johannes Gerwien, bedanken. Er hat mich in die Psycholinguistik eingeführt – mein allererstes Seminar in diesem Fachgebiet während des Masterstudiums wurde von ihm geleitet. Seitdem hat er mich von meinen ersten Schritten in der Psycholinguistik bis zum Abschluss der Promotion begleitet. Jedes Mal, wenn ich auf Schwierigkeiten oder Fragen stieß, erhielt ich von ihm hilfreiche, klare und ermutigende Ratschläge. Für seine wertvolle Unterstützung und seinen beständigen Beistand bin ich ihm zutiefst dankbar.

Mein herzlicher Dank gilt ebenso Frau Prof. Dr. Christiane von Stutterheim. Bereits im Masterstudium hatte ich die Gelegenheit, eine ihrer Lehrveranstaltungen zu besuchen, von der ich nachhaltig profitiert habe. Für ihre hilfreichen Anmerkungen und die konstruktive Beratung zu meinem Promotions-Exposé sowie für ihre detaillierten Kommentare zu dieser Dissertation bin ich ihr in besonderem Maße verbunden.

Ein besonderer Dank gilt allen Teilnehmerinnen und Teilnehmern meiner Experimente. Ohne ihre Bereitschaft und ihren Einsatz wäre die Durchführung der Studien nicht möglich gewesen.

Des Weiteren danke ich von Herzen all meinen Freunden, insbesondere Mengxue Zhao, Di Zhao, Fan Fei und Lan Tao, die in den vergangenen Jahren an meiner Seite standen, mir zuhörten, mich ermutigten und mich in schwierigen Momenten unterstützten. Gemeinsame Gespräche, Essen, Ausflüge, Reisen und viele kleine Alltagsmomente haben mir geholfen, den Stress des Studiums zu bewältigen und meine psychische Gesundheit zu erhalten. Lili Ji und Ruofan Zhang danke ich zudem ganz besonders für ihre wertvollen Ratschläge und ihre Unterstützung in der intensiven Phase der Disputationsvorbereitung.

Mein tief empfundener Dank gilt meinen Eltern, Wenlian Yao und Xinping Wei, die mich über all die Jahre hinweg in jeder Hinsicht unterstützt und mir stets Vertrauen entgegengebracht haben.

Die vorliegende Arbeit wurde durch ein Promotionsstipendium des China Scholarship Council (CSC) finanziell gefördert, wofür ich ebenfalls meinen herzlichen Dank ausspreche.

Inhaltsverzeichnis

1.	EINLEITUNG	1
2.	EREIGNISKOGNITION	6
2.1.	EREIGNISKOGNITION UND SPRACHVERARBEITUNG	9
2.2.	THEMATISCHE ROLLEN	14
2.3.	EREIGNISSEGMENTIERUNG	18
2.4.	OBJEKTZUSTÄNDE	23
3.	EREIGNISKOGNITION UND SPRACHSTRUKTUREN	28
3.1.	KOGNITIVE MODELLE UND IMAGE SCHEMA	33
3.1.1.	<i>Kognitive Modelle</i>	34
3.1.2.	<i>Image Schema</i>	38
3.2.	CONSTRUAL	43
3.2.1.	<i>Profiling</i>	45
3.2.2.	<i>Mentales Scanning</i>	56
3.2.3.	<i>Specificity</i>	62
4.	SPRACHSTRUKTUREN IM CHINESISCHEN UND DEUTSCHEN	65
4.1.	HAUPTSATZ	66
4.1.1.	<i>Chinesischer Hauptsatz</i>	68
4.1.2.	<i>Deutscher Hauptsatz</i>	70
4.2.	SUBJEKT-RELATIVSATZ	72
4.2.1.	<i>Chinesischer Subjekt-Relativsatz</i>	73
4.2.2.	<i>Deutscher Subjekt-Relativsatz</i>	75
4.3.	NOMINALISIERUNG	77
4.3.1.	<i>Chinesische Nominalisierung</i>	78
4.3.2.	<i>Deutsche Nominalisierung</i>	83
5.	FRAGESTELLUNG UND FORSCHUNGSDESIGN	86
5.1.	FRAGESTELLUNG	86
5.2.	FORSCHUNGSDESIGN	91
5.2.1.	<i>Linguistische Stimuli</i>	91
5.2.2.	<i>Visuelle Stimuli</i>	93
5.2.3.	<i>Experimente</i>	94
6.	EXPERIMENT 1	102
6.1.	HYPOTHESEN	102
6.2.	METHODISCHES VORGEHEN	104
6.2.1.	<i>Versuchspersonen</i>	104
6.2.2.	<i>Material</i>	104
6.2.3.	<i>Durchführung</i>	108
6.2.4.	<i>Datenkodierung und -analyse</i>	109
6.3.	ERGEBNISSE	111

6.4.	DISKUSSION	113
7.	EXPERIMENT 2	116
7.1.	HYPOTHESEN	116
7.2.	METHODISCHES VORGEHEN.....	118
7.2.1.	<i>Versuchspersonen</i>	118
7.2.2.	<i>Material</i>	118
7.2.3.	<i>Durchführung</i>	122
7.2.4.	<i>Datenkodierung und -analyse</i>	124
7.3.	ERGEBNISSE	127
7.4.	DISKUSSION	129
8.	EXPERIMENT 3	135
8.1.	HYPOTHESEN	136
8.2.	METHODISCHES VORGEHEN.....	139
8.2.1.	<i>Versuchspersonen</i>	139
8.2.2.	<i>Apparatur</i>	139
8.2.3.	<i>Material</i>	140
8.2.4.	<i>Durchführung</i>	143
8.2.5.	<i>Datenkodierung</i>	145
8.2.6.	<i>Datenanalyse</i>	146
8.3.	ERGEBNISSE	151
8.3.1.	<i>Ergebnisse der ersten Fixation</i>	151
8.3.2.	<i>Ergebnisse der Fixationslatenz</i>	153
8.3.3.	<i>Ergebnisse der Gazedauer</i>	156
8.4.	DISKUSSION	158
9.	EXPERIMENT 4	166
9.1.	HYPOTHESEN	168
9.2.	METHODISCHES VORGEHEN.....	169
9.2.1.	<i>Versuchspersonen</i>	169
9.2.2.	<i>Apparatur</i>	170
9.2.3.	<i>Material</i>	170
9.2.4.	<i>Durchführung</i>	170
9.2.5.	<i>Datenkodierung</i>	171
9.2.6.	<i>Datenanalyse</i>	171
9.3.	ERGEBNISSE	173
9.3.1.	<i>Ergebnisse zu den deutschen Daten</i>	173
9.3.2.	<i>Ergebnisse zum Sprachvergleich</i>	179
9.4.	DISKUSSION	189
10.	ALLGEMEINE DISKUSSION	200
10.1.	ZUSAMMENFASSUNG	200
10.2.	REFLEXION	206
10.3.	AUSBLICK.....	208

APPENDIX I : KRITISCHE STIMULI IN DEN CHINESISCHEN EXPERIMENTEN	212
APPENDIX II : KRITISCHE STIMULI IM DEUTSCHEN EXPERIMENT	230
APPENDIX III : VISUELLE STIMULI	236
LITERATUR	248

Abbildungsverzeichnis

Abbildung 2.1: Modell zur Segmentierung von Ereignissen (zit. nach Zacks et al., 2007, S. 274).....	20
Abbildung 3.1: Abstrakte Darstellungen von Ereignissen in Hauptsätzen, Relativsätzen und Nominalisierungen.....	30
Abbildung 3.2: Die Handlungskette (angepasste Darstellung nach Langacker, 1991, S. 283).	35
Abbildung 3.3: Das kanonische Ereignismodell mit den zentralen Rollen Agens (Ag) und Patiens (Pat) (angepasste Darstellung nach Langacker, 1991, S. 285).	37
Abbildung 3.4: Das <i>Link</i> -Schema (zit. nach Johnson, 1987, S. 118).	39
Abbildung 3.5: Das <i>Center-Periphery</i> -Schema (angepasste Darstellung nach Johnson, 1989, S. 112).	41
Abbildung 3.6: Unterschiedliche Profile.	46
Abbildung 3.7: Schematische Charakterisierung von Nomen und Verben.....	50
Abbildung 3.8: Schemata der Profilverschiebung (angepasste Darstellung nach Li & Yang, 2020, S. 838).	53
Abbildung 3.9: Schemata für Verben und episodische Nominalisierungen.	54
Abbildung 3.10: Schema des sequentiellen Scannings.	59
Abbildung 3.11: Schema des holistischen Scannings.	61
Abbildung 4.1: Abstrakte Darstellung eines finiten Satzes (zit. nach Langacker, 1991, S. 37).	68
Abbildung 4.2: Abstrakte Darstellung des chinesischen Hauptsatzes „ <i>jīxièshī xiūlǐ qìchē</i> “ (der Mechaniker repariert das Auto).....	70
Abbildung 4.3: Abstrakte Darstellung des deutschen Hauptsatzes „ <i>Der Mechaniker repariert das Auto</i> “	72
Abbildung 4.4: Abstrakte Darstellung der Nominalisierung „ <i>qìchē de xiūlǐ</i> “ (die Autoreparatur).	82
Abbildung 4.5: Abstrakte Darstellung des Rektionskompositums „ <i>die Autoreparatur</i> “	85
Abbildung 5.1: Ein Beispiel der visuellen Stimuli. Es entspricht den linguistischen Stimuli in Bsp. 5-1.	94
Abbildung 6.1: Ein Beispiel der vier Varianten eines kritischen visuellen Stimulus in Experiment 1.	103
Abbildung 6.2: Ein Beispiel der Bildvarianten der Filler-Bedingung in Experiment 1.	107
Abbildung 6.3: Ein Beispiel-Trial im Fragebogen in Experiment 1.	109
Abbildung 6.4: Proportion der Bildwahl nach sprachlicher Struktur (Experiment 1).....	112
Abbildung 7.1: Ein Beispiel für die kritischen Bedingungen in Experiment 2.....	120
Abbildung 7.2: Beispiele für die Filler-Bedingungen in Experiment 2	121
Abbildung 7.3: Beispielablauf eines Trials (Experiment 2).....	124
Abbildung 7.4: Geschätzte Reaktionszeiten nach Übereinstimmungsbedingung (Experiment 2).	128
Abbildung 7.5: Geschätzte Reaktionszeiten nach Sprachstruktur und Übereinstimmungsbedingung (Experiment 2).....	129
Abbildung 8.1: Ein Beispiel für die kritischen visuellen Stimuli in Experiment 3.	142

Abbildung 8.2: Beispielablauf eines Trials (Experiment 3).....	145
Abbildung 8.3: Ein Beispiel für AOs der kritischen Bedingung (Experiment 3).	146
Abbildung 8.4: Posterior-Wahrscheinlichkeit für die Erstfixation auf die Patiens- /Handlungsregion (Experiment 3).....	153
Abbildung 8.5: Geschätzte Fixationslatenz in die Agens-Region (Experiment 3).....	154
Abbildung 8.6: Geschätzte Fixationslatenz in die Patiens-/Handlungsregion (Experiment 3).	155
Abbildung 8.7: Geschätzte gesamte Gazedauer in der Agens-Region (Experiment 3)....	156
Abbildung 8.8: Geschätzte gesamte Gazedauer in der Patiens-/Handlungsregion (Experiment 3).....	157
Abbildung 9.1: Posterior-Wahrscheinlichkeit für die Erstfixation auf die Patiens- /Handlungsregion (Experiment 4).....	174
Abbildung 9.2: Geschätzte Fixationslatenz in die Agens-Region (Experiment 4).....	175
Abbildung 9.3: Geschätzte Fixationslatenz in die Patiens-/Handlungsregion (Experiment 4).	176
Abbildung 9.4: Geschätzte gesamte Gazedauer in der Agens-Region (Experiment 4)....	178
Abbildung 9.5: Geschätzte gesamte Gazedauer in der Patiens-/Handlungsregion (Experiment 4).....	179
Abbildung 9.6: Posterior-Wahrscheinlichkeit für die Erstfixation auf die Patiens- /Handlungsregion im Sprachvergleich.	181
Abbildung 9.7: Geschätzte Fixationslatenz in die Agens-Region im Sprachvergleich.	183
Abbildung 9.8: Geschätzte Fixationslatenz in die Patiens-/Handlungsregion im Sprachvergleich.	185
Abbildung 9.9: Geschätzte gesamte Gazedauer in der Agens-Region im Sprachvergleich.	187
Abbildung 9.10: Geschätzte gesamte Gazedauer in der Patiens-/Handlungsregion im Sprachvergleich.	189

Tabellenverzeichnis

Tabelle 6.1: Ergebnisse der <i>Bayesian Proportion Tests</i> (Experiment 1).....	113
Tabelle 7.1: Deskriptive Statistiken und Ergebnisse des Bayes'schen Modells zu den Reaktionszeiten (Experiment 2).	129
Tabelle 8.1: Ergebnisse der posterioren paarweisen Vergleiche zur Erstfixation (Experiment 3).	153
Tabelle 8.2: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz (Experiment 3).....	155
Tabelle 8.3: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer (Experiment 3).	158
Tabelle 9.1: Ergebnisse der posterioren paarweisen Vergleiche zur Erstfixation (Experiment 4).	174
Tabelle 9.2: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz (Experiment 4).....	177
Tabelle 9.3: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer (Experiment 4).	179
Tabelle 9.4: Ergebnisse der posterioren paarweisen Vergleiche zur ersten Fixation auf der Patiens-/Handlungsregion im Sprachvergleich.....	181
Tabelle 9.5: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz in der Agens-Region im Sprachvergleich.....	184
Tabelle 9.6: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz in der Patiens-/Handlungsregion im Sprachvergleich.....	185
Tabelle 9.7: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer in der Agens-Region im Sprachvergleich.....	187
Tabelle 9.8: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer in der Patiens-/Handlungsregion im Sprachvergleich.....	189
Tabelle 9.9: Vergleich der linearen und semantischen Struktur von Hauptsätzen im Deutschen und Chinesischen.....	191
Tabelle 9.10: Vergleich der linearen und semantischen Struktur von Nominalisierungen im Deutschen und Chinesischen.....	191
Tabelle 9.11: Vergleich der linearen und semantischen Struktur von Relativsätzen im Deutschen und Chinesischen.....	192

Abkürzungen

AOI	Area of Interest
Bsp.	Beispiel
CI	Konfidenzintervall
DemPron.	Demonstrativpronomen
DO	Doppelobjekt
EMMs	Estimated Marginal Means
ERP	Event-Related Potential
EST	Event Segmentation Theory
HDI	Höchstsdichteintervall
IOH	Intersecting Object Histories Model
lm	Landmark
M	Mean (Mittelwert)
ms	Millisekunde
N	Nomen
NP	Nominalphrase
fMRI	funktionelle Magnetresonanztomographie
PEF	Perfekt
pd	probability of direction
PO	Präpositionalobjekt
ROPE	Region of Practical Equivalence
SD	Standard Deviation
SVO	Subjekt-Verb-Objekt
tr	Trajektor
V	Verb
VP	Verbalphrase

1. Einleitung

Menschen sind täglich einer Vielzahl von Informationen aus der Außenwelt ausgesetzt, sowohl durch direkte persönliche Erfahrungen als auch durch indirekte sprachliche Vermittlung. Um diesen kontinuierlichen Informationsfluss zu verstehen, wird er durch das kognitive System verarbeitet und in sinnvolle Einheiten organisiert. Zu diesen Einheiten zählen auch diskrete und individuell unterschiedliche Ereignisse, die es ermöglichen, Erfahrungen zu strukturieren und mit anderen darüber zu kommunizieren.

Ereignisse und Kommunikationssituationen sind so vielfältig, dass sie nicht auf einheitliche Weise ausgedrückt werden können. Aus der Vielfalt sprachlicher Kodierungsmöglichkeiten konzentriert sich diese Dissertation insbesondere auf drei sprachliche Strukturen: den einfachen aktiven transitiven Hauptsatz, den Subjekt-Relativsatz sowie die deverbale Nominalisierung. Diese sprachlichen Strukturen stellen keine rein stilistischen Varianten dar, sondern reflektieren unterschiedliche Perspektiven oder Fokussetzungen der Sprachbenutzer, die durch ihre jeweilige kommunikative Absicht motiviert werden. Während der einfache aktive Hauptsatz eine prototypische Satzstruktur zur Kodierung eines kausativen Ereignisses darstellt (vgl. Langacker, 1991), werden Relativsätze und Nominalisierungen typischerweise in einen übergeordneten Satz eingebettet und verknüpfen das enkodierte Ereignis mit einem größeren sprachlichen Kontext. Obwohl beide letztgenannten Strukturen zur Kontextverknüpfung dienen, unterscheiden sie sich wesentlich. Relativsätze liefern typischerweise ergänzende Informationen, um einen Referenten genauer zu identifizieren oder näher zu beschreiben, z.B. „*Der Mechaniker, der das Auto reparierte, ist heute krank.*“ Nominalisierungen kodieren das Ereignis dagegen kompakter und gröber, wie in „*Die Autoreparatur kostet viel Zeit.*“ Diese drei sprachlichen Strukturen erfüllen somit jeweils spezifische kommunikative Funktionen.

Die Fähigkeit, ein und dasselbe Ereignis unterschiedlich zu konstruieren und sprachlich zu kodieren, wird in der kognitiven Linguistik unter dem Begriff *Construal* gefasst

(Langacker, 1991, 2013, 2019). Laut Langacker (1991, 2019) besteht die Bedeutung eines sprachlichen Ausdrucks nicht ausschließlich in dessen konzeptuellem Inhalt, sondern auch maßgeblich darin, wie dieser Inhalt mental konstruiert und sprachlich dargestellt wird. *Construal* steuert dabei, wie Erfahrungen kognitiv gerahmt werden, auf welche Komponenten eines Ereignisses die Aufmerksamkeit gelenkt wird und welche Aspekte selektiv ausgeblendet werden (vgl. Divjak, Milin & Medimorec, 2020). Durch die Wahl zwischen sprachlichen Strukturen können unterschiedliche Bedeutungen und Fokussierungen in Bezug auf dasselbe Ereignis entstehen. Sprache funktioniert in diesem Zusammenhang wie ein Spotlight, das bestimmte Informationen salient macht, während andere abgeschwächt werden und damit die Aufmerksamkeit moduliert.

Obwohl *Construal* einen zentralen Begriff der kognitiven Linguistik bildet, wurde sein Einfluss auf die Ereignisrepräsentation bisher nur eingeschränkt empirisch untersucht. Frühere Studien zu diesem Thema konzentrierten sich überwiegend auf einfache sprachliche Alternationen in Hauptsätzen, wie Aktiv- und Passivsätzen, und blieben auf eine einzelne Sprache beschränkt (vgl. Divjak et al., 2020; Sauppe & Flecken, 2021).

Diese Dissertation erweitert die bisherigen Untersuchungen zum *Construal* von Ereignissen, indem sie zunächst den Einfluss komplexerer Sprachstrukturen auf die durch Sprachverstehen generierte Ereignisrepräsentation im Chinesischen berücksichtigt. Dabei wird insbesondere analysiert, welche Ereigniskomponenten durch die untersuchten sprachlichen Strukturen jeweils hervorgehoben werden und wie dadurch die Aufmerksamkeit moduliert wird.

Zudem variieren Sprachstrukturen typischerweise über Sprachen hinweg erheblich in ihrer oberflächlichen Realisierung. Viele psycholinguistische Studien legen nahe, dass sprachspezifische Eigenschaften, z.B. Präferenzen bei der Kodierung von Bewegungsereignissen („*path*“ oder „*manner*“), die Ereignisrepräsentationen bei der Sprachproduktion beeinflussen können (vgl. Gerwien & von Stutterheim, 2018; Papafragou et al., 2008; Tınaz & Ünal, 2024).

Um weiter zu untersuchen, ob die im Chinesischen beobachteten Effekte auch in anderen Sprachen auftreten, insbesondere im Hinblick auf Unterschiede in der sprachlichen Oberflächenstruktur, wird in der vorliegenden Dissertation noch ein sprachübergreifender Vergleich zwischen Chinesisch und Deutsch vorgenommen. Diese beiden Sprachen verfügen zwar über die gleichen grundlegenden Strukturen (Hauptsatz, Subjekt-Relativsatz, Nominalisierung), weisen jedoch bedeutende Unterschiede in ihren Oberflächenstrukturen auf, insbesondere bei Subjekt-Relativsätzen. Während das Kopfnomen in chinesischen Subjekt-Relativsätzen am Satzende steht, nimmt es in deutschen Subjekt-Relativsätzen die initiale Position ein. Diese typologischen Unterschiede bieten die Gelegenheit, gezielt zu untersuchen, ob und inwiefern die lineare Wortstellung sprachlicher Konstruktionen das *Construal* von Ereignissen beeinflusst.

Die vorliegende Dissertation gliedert sich in zehn Kapitel. Zu Beginn werden in Kapitel 2 grundlegende Theorien zur Ereigniskognition und deren Beziehung zur Sprache eingeführt. Dabei steht zunächst die Darstellung der Ereignisrepräsentation in konzeptuellen und sprachlichen Domänen im Vordergrund. Im Mittelpunkt stehen die Fragen, wie Ereignisse durch die Sprache wahrgenommen werden und welchen Einfluss Sprache auf die kognitive Verarbeitung von Ereignisrepräsentationen ausübt. Im Anschluss daran werden Theorien und Verarbeitungsmodelle vorgestellt, die verschiedene zentrale Facetten der Ereignisrepräsentation beleuchten. Besonderes Augenmerk liegt hierbei auf den thematischen Rollen der Ereignisteilnehmer (vgl. Dowty, 1989; Rissman & Majid, 2019; Ünal et al., 2021), der Theorie zur Ereignissegmentierung (vgl. Zacks & Tversky, 2001; Zacks et al., 2007) sowie auf dem „*Intersecting Object Histories Model*“, das dynamische Zustandsrepräsentationen von Objekten innerhalb eines Ereignisses beschreibt (vgl. Altmann & Ekves, 2019).

In Kapitel 3 werden die kognitiven Mechanismen und Verarbeitungsprozesse beleuchtet, die den drei untersuchten sprachlichen Strukturen zugrunde liegen. Aufbauend auf theoretischen Konzepten der kognitiven Linguistik wird erläutert, dass unterschiedliche sprachliche Kodierungen spezifischen kognitiven Mechanismen entsprechen. Kapitel

3.1 beschäftigt sich insbesondere mit kognitiven Modellen nach Langacker (1991) sowie mit *Image*-Schemata nach Johnson (1987) und Lakoff (1987). Diese Modelle und Schemata verdeutlichen, wie sensorische und körperliche Erfahrungen metaphorisch auf abstrakte sprachliche Strukturen übertragen werden und somit zur Konzeptualisierung sprachlicher Konstruktionen beitragen. Kapitel 3.2 widmet sich ausführlich dem Begriff „*Construal*“, der mehrere konzeptuelle Dimensionen umfasst, die die Interpretation kodierter Inhalte beeinflussen. Anhand der Faktoren *Profiling*, *mentales Scanning* und *Specificity* werden die drei untersuchten Sprachstrukturen detailliert analysiert und hinsichtlich ihrer konzeptuellen Unterschiede betrachtet.

Kapitel 4 erläutert im Detail die konkreten Oberflächenstrukturen und Bildungsstrategien der untersuchten sprachlichen Strukturen im Chinesischen und Deutschen. Aufbauend auf den in Kapitel 3 eingeführten theoretischen Konzepten werden Gemeinsamkeiten und Unterschiede in den zugrunde liegenden kognitiven Mechanismen und Verarbeitungsprozessen in beiden Sprachen diskutiert und gegenübergestellt.

In Kapitel 5 werden die zentralen Fragestellungen der Dissertation ausführlich dargelegt und das experimentelle Design erläutert. Insgesamt wurden vier empirische Studien durchgeführt, die in den Kapiteln 6 bis 9 im Detail vorgestellt werden. Dabei kamen unterschiedliche methodische Ansätze zum Einsatz, um den Einfluss der sprachlichen Strukturen auf die Ereigniskonzeptualisierung sowohl im Chinesischen als auch im Deutschen aus verschiedenen Perspektiven systematisch zu erfassen.

Kapitel 6 präsentiert das erste Experiment, das durch eine *offline Picture-Selection*-Aufgabe zunächst die theoretische Annahme überprüft, dass die drei untersuchten sprachlichen Strukturen (Hauptsatz, Subjekt-Relativsatz und Nominalisierung) jeweils spezifische Ereigniskomponenten hervorheben. Dabei wird durch eine farbliche Hervorhebung der linguistisch betonten Komponenten die Assoziation zwischen sprachlicher und visueller Salienz explizit veranschaulicht. Zudem dient dieses Experiment der Normierung von kritischen Stimuli für die weiteren Experimente.

In Kapitel 7 wird das zweite Experiment vorgestellt, das auf den Befunden von Experiment 1 sowie auf theoretischen Annahmen zur unterschiedlichen Hervorhebung von Ereigniskomponenten bei Subjekt-Relativsätzen und Nominalisierungen basiert. Mithilfe einer *Picture-Sentence-Matching*-Aufgabe wird überprüft, ob sprachlich evozierte Hervorhebungen tatsächlich zu unterschiedlichen Reaktionszeiten bei der Echtzeitverarbeitung von Ereignisrepräsentationen führen. Konkret wird untersucht, ob Probanden visuelle Stimuli schneller verarbeiten, wenn die visuell und linguistisch salienten Komponenten übereinstimmen als wenn sie nicht übereinstimmen.

Kapitel 8 widmet sich dem dritten Experiment, in dem mittels Eyetracking die genaue Verteilung der visuellen Aufmerksamkeit während der Ereignisverarbeitung erfasst wird. Ziel ist es zu ermitteln, welche Ereigniskomponenten je nach verwendeter Sprachstruktur priorisiert verarbeitet werden. Im Zentrum steht hierbei die Frage, ob die unterschiedlichen, linguistisch hervorgehobenen Komponenten auch zu unterschiedlichen Blickbewegungsmustern führen.

Kapitel 9 stellt das vierte Experiment vor, das in analoger Weise wie Experiment 3, jedoch in deutscher Sprache durchgeführt wurde, um eine sprachvergleichende Perspektive einzunehmen. Es wird untersucht, ob die im Chinesischen nachgewiesenen Aufmerksamkeitsmuster der Sprachstrukturen auch im Deutschen auftreten oder ob sprachspezifische Unterschiede feststellbar sind. Ein besonderes Augenmerk liegt auf den Subjekt-Relativsätzen, deren Wortstellung sich zwischen beiden Sprachen unterscheidet, während ihre semantische Struktur theoretisch vergleichbar bleibt. Ziel dieses Experiments ist es, zu überprüfen, inwieweit syntaktische Wortstellungsunterschiede die visuelle Ereignisverarbeitung beeinflussen.

In Kapitel 10 werden abschließend die zentralen Erkenntnisse der vorliegenden Arbeit zusammengefasst und Perspektiven sowie methodische und theoretische Implikationen für zukünftige Forschung aufgezeigt.

2. Ereigniskognition

Menschen nehmen kontinuierlich eine Vielzahl von Informationen aus der physischen Welt wahr. Mithilfe ihrer kognitiven Fähigkeiten wird dieser Informationsfluss in sinnvolle Einheiten, sogenannte Ereignisse, konstruiert und organisiert (vgl. Ünal, Ji & Papafragou, 2021; Zacks & Tversky, 2001). Ereignisse können dabei unterschiedlich ausfallen. Sie variieren in Bezug auf Zeit, Raum, die Anzahl und Art der beteiligten Entitäten sowie deren Beziehung zueinander (vgl. Elman, 2009; Zwaan, Langston & Graesser, 1995). Ein Ereignis kann beispielsweise wenige Sekunden andauern, wie ein Pfeilschuss, oder sich über Jahre erstrecken, wie ein Krieg.

Ereigniskognition wird als grundlegende menschliche Fähigkeit verstanden, um komplexe, oft multidimensionale und miteinander verknüpfte Informationen der Welt mental zu strukturieren und zu verarbeiten. Die Ereignisrepräsentationen spiegeln diese Komplexität wider und sind insofern wie Bilder, als sie isomorph zu den repräsentierten Ereignissen sind. Allerdings bilden Ereignisrepräsentationen nicht alle Details ab, sondern erfassen vor allem die für das Verständnis zentralen Merkmale (vgl. Radvansky & Zacks, 2011). In diesem Zusammenhang sind Ereignisse komplexe Repräsentationen. Ereignisrepräsentationen dienen demnach der mentalen Verarbeitung, Reaktion auf, Vorhersage und Interaktion mit der dynamischen Welt.

Ereignisrepräsentationen können auf zwei grundlegende Weisen entstehen. Zum einen werden sie aus direkten sensomotorischen Erfahrungen generiert, z.B. aus visueller oder auditiver Wahrnehmung. Die sensorischen Inputsignale werden vom kognitiven System verarbeitet und mit vorhandenen Ereignisschemata, d.h. aus früheren Erfahrungen abstrahiertem, im Langzeitgedächtnis gespeichertem generalisiertem Wissen, integriert. Auf diese Weise entstehen mentale Repräsentationen, die eine Interpretation der aktuellen Sinneseindrücke widerspiegeln. Zum anderen können Ereignisrepräsentationen aus der Verarbeitung sprachlicher Äußerungen hervorgehen. Beim Sprachverstehen, bei dem Verarbeitungsprozesse wie lexikalische Erkennung, syntaktische Analyse sowie semantische und pragmatische Integration involviert sind,

werden aus sprachlich vermittelten Informationen mentale Modelle und Repräsentationen des Ereignisses konstruiert (vgl. Dietrich, 2002; Rickheit, Sichelschmidt & Strohnert, 2002). Diese Ereignisrepräsentationen können zudem als Ausgangspunkt für Sprachproduktion dienen, indem sie als Input in den Sprachproduktionsprozess eingehen und eventuell in Sprache umgesetzt werden.

Die Ereignisrepräsentation ist ein zentrales Forschungsthema in der Kognitionswissenschaft und Psycholinguistik. Im Mittelpunkt stehen dabei Fragen wie: Wie werden Ereignisse im kognitiven System repräsentiert und verarbeitet? Welche kognitiven Prozesse und Wissensbestände sind daran beteiligt? Welche Aspekte sind für die Ereignisrepräsentation entscheidend? Zur Untersuchung dieser Fragen wurden verschiedene theoretische Ansätze und Verarbeitungsmodelle entwickelt, die jeweils wichtige Facetten der Ereignisrepräsentation in den Fokus rücken. In diesem Kapitel werden drei grundlegende Theorien zur Charakterisierung von Ereignisrepräsentationen vorgestellt, nämlich thematische Rollen, Ereignissegmentierung und das *Intersecting Object Histories Model* (IOH-Modell; Altmann & Ekves, 2019). Diese Ansätze richten den Blick auf den Aufbau und die Verarbeitung von Ereignisrepräsentationen, auf die Konstruktion mentaler Modelle sowie auf die Aktivierung und Veränderung der Repräsentationen. Sie tragen zu einem differenzierten Verständnis der Konturierung und Dynamik von Ereignisrepräsentationen bei und erlauben es, genauer zu bestimmen, welche Veränderungen in sprachlichen oder nicht-sprachlichen Inputs die mentale Repräsentation eines Ereignisses beeinflussen.

Ereignisse können verschiedene Teilnehmer beinhalten, die unterschiedliche thematische Rollen einnehmen, z.B. Agens, Patiens oder Instrument. Die spezifischen Eigenschaften dieser Teilnehmer sowie ihre Beziehungen zueinander tragen zur Spezifizierung und Interpretation von Ereignissen bei (vgl. Radvansky & Zacks, 2011). Beispielsweise umfassen die Ereignisse „*Der Hund jagt die Katze*“ und „*Die Katze jagt den Hund*“ zwar die gleichen Teilnehmer, jedoch stehen diese in unterschiedlichen Beziehungen zueinander. Im ersten Fall ist *der Hund* das Agens und *die Katze* das

Patiens, im zweiten Fall sind die Rollen umgekehrt. Diese Rollenzuweisung führt zu unterschiedlichen konzeptuellen Repräsentationen und Interpretationen des zugrunde liegenden Ereignisses.

Zudem ist die Segmentierung des kontinuierlichen Informationsflusses in sinnvolle Einheiten eine grundlegende Voraussetzung für die Wahrnehmung und Verarbeitung von Ereignissen. Zacks et al. (2007) entwickelten in diesem Zusammenhang einen Mechanismus. Demgemäß werden aus der Interaktion zwischen dem aktuellen sensorischen bzw. konzeptuellen Input und den im Langzeitgedächtnis gespeicherten Ereignisschemata Ereignismodelle im Arbeitsgedächtnis aufgebaut. Diese Arbeitsmodelle dienen zugleich als Basis für Vorhersagen darüber, wie sich ein Ereignis in der unmittelbaren Zukunft fortsetzen wird.

Ein weiteres wesentliches Merkmal von Ereignissen ist ihre Dynamik. Ein Ereignis umfasst vielfältige Objekte, die ständigen Veränderungen über die Zeit unterliegen (vgl. Altmann & Ekves, 2019; Radvansky, 2012). So verändern sich im Ereignis „*Der Bauer erntet das Getreide*“ die Körperhaltung des Bauers, die Position des Instrumentes und die physische Form des Getreides im Laufe der Zeit. Um ein solches Ereignis zu verstehen, müssen die Zustandsänderungen der Ereignisteilnehmer, ihre räumlich-zeitlichen Beziehungen und ihre Interaktionen in den Ereignisrepräsentationen kodiert und als Ganzes miteinander in Beziehung gesetzt werden. Das IOH-Modell (Altmann & Ekves, 2019) bietet hierfür eine theoretische Grundlage. Demnach definieren sich Ereignisse durch die Überschneidung der Verläufe mehrerer Objekte in Raum und Zeit. Dabei spielt die Zustandsrepräsentation der Objekte, die als Ereignisteilnehmer interpretiert werden, eine grundlegende Rolle.

Eine zentrale, bislang offene Forschungsfrage im Bereich der Ereigniskognition besteht darin, ob und inwiefern Sprache die nicht-sprachliche kognitive Verarbeitung beeinflusst. Zu dieser Frage gibt es unterschiedliche theoretische Auffassungen. Während einige Wissenschaftler von einer klaren Trennung zwischen Sprache und nicht-sprachlicher Kognition ausgehen, betrachten andere Sprache als einen integralen

Bestandteil allgemeiner kognitiver Prozesse (vgl. Jackendoff, 1990; Thierry, 2016; Whorf, 1956).

Die vorliegende Dissertation geht von dieser Frage aus und untersucht, wie unterschiedliche sprachliche Strukturen mit kognitiven Systemen interagieren. Dabei liegt der Fokus insbesondere darauf, ob und wie spezifische linguistische Strukturen die konzeptuelle Ereignisrepräsentation und die Aufmerksamkeitsverteilung beeinflussen können. Dies soll zu einem tieferen Verständnis der Beziehung zwischen Sprachstruktur und kognitiver Verarbeitung beitragen.

In diesem Kapitel wird zunächst die Beziehung zwischen Sprache und Kognition diskutiert (Kapitel 2.1). Anschließend werden, inspiriert von Gerwien, Marberg und Nicolaisen (2024), Theorien und Verarbeitungsmodelle zu drei zentralen Aspekten der Ereignisrepräsentation vorgestellt, nämlich thematische Rollen (Kapitel 2.2), die Theorie der Ereignissegmentierung (Kapitel 2.3) und das IOH-Modell (Kapitel 2.4), das den zeitlichen Verlauf von Ereignissen sowie die Zustandsveränderungen der daran beteiligten Objekte beschreibt.

2.1. Ereigniskognition und Sprachverarbeitung

Eine zentrale Fragestellung in der Kognitionswissenschaft und Linguistik betrifft das Verhältnis zwischen Sprache und Kognition. Konkret geht es um die Frage, inwiefern Sprache die kognitiven Prozesse beeinflusst.

Diese Frage wird kontrovers diskutiert. Vertreter des universalistischen Ansatzes argumentieren, dass sprachliche und konzeptuelle Repräsentationen grundsätzlich unabhängig voneinander sind (vgl. Jackendoff, 1990; Talmy, 1985). *Der universalistische Ansatz* postuliert, dass Kernaspekte der menschlichen Wahrnehmung und Kognition universell sind. Die grundlegenden kognitiven Konzepte zu Objekten und Ereignissen bleiben demnach bei allen Menschen weitgehend gleich, unabhängig davon, welche Sprache sie sprechen. Sprachliche Unterschiede in Grammatik und Lexikon spiegeln lediglich oberflächliche Variationen in der Enkodierung wider und

werden als rein sprachliche Phänomene betrachtet, die keine tiefgreifenden Auswirkungen auf die Ereigniskognition haben (vgl. Gennari, Sloman, Malt & Fitch, 2002; Jackendoff, 1990; Talmy, 1985).

Studien zur räumlichen Kognition, Farb- und Objektkategorisierung unterstützen diesen Ansatz. Beispielsweise bezieht sich das Wort *mouton* im Französischen sowohl auf das Tier als auch auf sein Fleisch. Diese beiden Konzepte werden im Englischen durch zwei verschiedene Wörter ausgedrückt, und zwar *sheep* und *mutton*. Trotz dieser sprachlichen Unterschiede können französische Muttersprachler problemlos zwischen Tier und Fleisch unterscheiden. Dies zeigt, dass die Sprache nicht unbedingt die Art und Weise beeinflusst, wie Menschen Kategorien bilden (vgl. Thierry, 2016).

Demgegenüber haben viele Psycho- und Neurolinguisten den universalistischen Ansatz kritisiert. Neurophysiologische Befunde zeigen, dass die an der Sprachverarbeitung beteiligten Hirnareale auch bei nicht-linguistischen Aufgaben aktiv sind (vgl. Thierry, Giraud, & Price, 2003; Thierry, 2016). Außerdem interagiert Sprache mit anderen wahrnehmungsbezogenen kognitiven Prozessen wie Emotionen (z.B. Wu & Thierry, 2012), was auf eine enge Verbindung zwischen Sprache und Kognition hinweist. Sprache ist demnach kein isoliertes Modul, sondern integraler Bestandteil vernetzter kognitiver Systeme und untrennbar mit der Kognition verbunden (vgl. Thierry, 2016).

Diese Sichtweise wird als *sprachliche Relativität* bezeichnet. Sie besagt, dass die Sprache, die wir sprechen, unsere Wahrnehmung und Konzeptualisierung von Ereignissen beeinflusst (vgl. Whorf, 1956). Demnach konzeptualisieren Menschen mit verschiedenen Muttersprachen die Welt auf unterschiedliche Weise (vgl. Whorf, 1956; Levinson, 2003).

Es lassen sich zwei mögliche Mechanismen unterscheiden, wie sich Sprache auf die mentale Repräsentation von Ereignissen auswirken könnte. Zum einen beeinflusst Sprache die Wahrnehmung von Ereignissen nur kurzfristig während des Sprechens oder der expliziten Sprachwendung. Zum anderen kann Sprache nicht-linguistische kognitive Prozesse langfristig prägen (vgl. Gennari et al., 2002; Thierry, 2016; Wolff &

Holmes, 2011).

Ein zentraler Ansatz bezüglich des kurzfristigen Einflusses ist das von Slobin (1996) vorgeschlagene „*thinking for speaking*“, das besagt, dass verschiedene Sprachen Ereignisse hinsichtlich vielfältiger Dimensionen wie Lexikon und Syntax unterschiedlich kodieren. Sprachbenutzer erwerben durch sprachliche Erfahrungen intuitiv das Wissen darüber, welche spezifischen syntaktischen und semantischen Informationen in ihrer Sprache erforderlich sind und wie diese in sprachlichen Strukturen kodiert werden müssen (vgl. Papafragou et al., 2008; Wolff & Holmes, 2011). Diese sprachspezifischen Anforderungen können die Aufmerksamkeitslenkung bereits früh während der Vorbereitung des Sprechens beeinflussen (vgl. Levelt, 1989; Slobin, 1996), da Sprachbenutzer dazu neigen, sich beim Sprechen auf bestimmte Aspekte eines Ereignisses zu konzentrieren, die in ihrer Sprache grammatisch oder lexikalisch ausgedrückt werden müssen (vgl. Altmann & Kamide, 2004; Griffin, 2001; Papafragou et al., 2008).

Papafragou et al. (2008) haben beispielsweise mittels Eyetracking herausgefunden, dass die Englischsprecher bei der Beschreibung von Bewegungsereignissen überwiegend *Manner*-Verben verwendeten und ihren Blick bevorzugt auf *Manner* der Bewegung richteten, während Griechischsprecher *Path*-Informationen vorrangig im Hauptverb kodierten und sich bevorzugt auf *Path* der Bewegung fokussierten. Diese sprachlichen Effekte zeigten sich jedoch nicht in der nicht-sprachlichen Gedächtnisaufgabe, in der die Versuchspersonen die Bewegungsereignisse nur betrachten und memorieren sollten. Die Aufmerksamkeit war dabei unabhängig von der Muttersprache ähnlich verteilt, bevor die Videos stoppten.

Die Ergebnisse deuten darauf hin, dass Sprache die Aufmerksamkeit bei der Sprachproduktion beeinflusst. Sprecher verschiedener Sprachen konzentrieren sich bei der Planung einer Äußerung auf diejenigen Aspekte eines Ereignisses, die in ihrer Sprache relevant sind. Jedoch tritt der Einfluss sprachspezifischer Unterschiede auf die Aufmerksamkeit und die Wahrnehmung von Ereignissen nur kurzfristig auf, wenn eine

sprachliche Kodierung erforderlich ist. Ohne explizite Sprachverarbeitung nehmen Menschen Bewegungsereignisse unabhängig von ihrer Muttersprache ähnlich wahr.

Eine stärkere Hypothese zu sprachlicher Relativität erweitert die Auswirkungen von Sprache auf die Konzeptualisierung der Ereignisse und die Aufmerksamkeitslenkung ihrer Sprecher über den sprachlichen Kontext hinaus. Demnach beeinflusst Sprache die Kognition nicht nur kurzfristig bei der Sprachproduktion, sondern hat auch langfristige und tiefgreifende Auswirkungen auf nicht-linguistische kognitive Prozesse (vgl. Levinson, 2003; Majid et al., 2004; Whorf, 1956). Durch den dauerhaften Gebrauch einer bestimmten Sprache können diejenigen Aspekte eines Ereignisses, die für die Sprachproduktion zentral sind, auch potenziell in anderen kognitiven Prozessen mehr Aufmerksamkeit auf sich ziehen, während weniger relevante Aspekte an Salienz verlieren (vgl. Gennari et al., 2002; Wolff & Holmes, 2011). Dies hat zur Folge, dass Sprache die Denkgewohnheiten ihrer Sprecher nachhaltig prägt und deren Wahrnehmung, Aufmerksamkeit und Gedächtnis beeinflusst (vgl. Levinson, 2003; Majid et al., 2004; Papafragou et al., 2008; Whorf, 1956).

Empirische Studien zu dieser Hypothese nutzen typischerweise nicht-linguistische Aufgaben. Obwohl die implizite Sprachverwendung in solchen Experimenten nicht vollständig ausgeschlossen werden kann, beleuchten diese Studien dennoch, dass Sprache die Kognition beeinflusst. Gerwien und von Stutterheim (2018) haben beispielsweise eine sprachübergreifende Studie zur Ereignissegmentierung zwischen Französisch und Deutsch durchgeführt. Diese Sprachen unterscheiden sich typologisch in der Kodierung von Bewegungsereignissen. Während im Französischen die *Path*-Information primär durch das Hauptverb ausgedrückt wird, und jede Richtungsänderung ein neues Verb erfordert, kodiert das Deutsche vorrangig *Manner* der Bewegung im Hauptverb und ergänzt die Richtungsangaben durch Präpositionen oder Partikeln. Dies ermöglicht es im Deutschen, mehrere Richtungsänderungen in einer einzigen sprachlichen Einheit zu beschreiben, während im Französischen jede Richtungsänderung eine separate sprachliche Einheit darstellt.

Die Ergebnisse dieser Studie unterstützen die Hypothese, dass sprachliche Strukturen die Segmentierung von Bewegungsereignissen beeinflussen. In der linguistischen Aufgabe, in der die Versuchspersonen die Bewegungsereignisse in den Videoclips beschrieben, unterteilten die französischen Muttersprachler die Bewegungen mit Richtungsveränderungen in mehrere separate sprachliche Einheiten, während die deutschen Muttersprachler diese eher als ein Ereignis beschrieben. Bemerkenswert ist, dass dieser sprachspezifische Effekt auch in der nicht-linguistischen Aufgabe beobachtet wurde. Französische Muttersprachler segmentierten die Bewegungsereignisse während der freien Betrachtung signifikant häufiger an den Stellen, an denen sich die Bewegungsrichtung änderte, als deutsche Muttersprachler.

Einen ähnlichen Effekt haben Flecken, Athanasopoulos, Kuipers und Thierry (2015) mithilfe von *Event-Related Potentials* (ERPs) gefunden. In dieser Studie nahmen englische und deutsche Muttersprachler an einem Bewegungsereignis-*Matching*-Test teil. Gemessen wurde die P3-Komponente, die mit Aufmerksamkeit und Stimulusverarbeitung assoziiert ist. Es zeigte sich, dass deutsche Muttersprachler eine stärkere P3-Reaktion auf den Endpunkt der Bewegung als auf die Trajektorie zeigten, während englische Muttersprachler keinen signifikanten Unterschied zwischen den beiden zeigten. Diese Ergebnisse stimmen mit den sprachspezifischen Eigenschaften überein. In Englisch, einer Aspektsprache, wird sowohl der Trajektorie als auch dem Endpunkt der Bewegung Aufmerksamkeit geschenkt, während Deutsch keine Aspektsprache ist und dem Endpunkt mehr Aufmerksamkeit schenkt als der Trajektorie.

Die Studie deutet darauf hin, dass selbst in einem nonverbalen Kontext grammatische Eigenschaften einer Sprache beeinflussen, worauf Sprecher ihre Aufmerksamkeit bei der Wahrnehmung von Bewegungsereignissen richten. Da die Unterschiede in den frühen elektroenzephalographischen Signalen, nicht aber in der behavioralen Aufgabe sichtbar wurden, legen diese Befunde nahe, dass Sprachstrukturen bereits unbewusste, frühe Wahrnehmungsprozesse prägen. Dieser Effekt ist wahrscheinlich nicht auf eine sprachliche Enkodierung zurückzuführen.

Diese Studien weisen insgesamt darauf hin, dass sich die Informationen, die in der Konzeptualisierungsphase für die Sprachproduktion notwendig sind, auch in nicht-linguistischen kognitiven Prozessen widerspiegeln. Sprachlich geprägte Muster manifestieren sich daher nicht nur in der Sprachproduktion, sondern auch langfristig in der allgemeinen Wahrnehmung von Ereignissen.

2.2. Thematische Rollen

Ereignisteilnehmer bilden eine grundlegende Komponente mentaler Ereignisrepräsentationen. Ihre spezifischen Rollen strukturieren das Ereignis, indem sie die relationalen Informationen zwischen den Beteiligten festlegen. Diese Ereignisrollen, Elemente also, die man aus der Semantik kennt, sind entscheidend für das Verständnis der Ereignisstruktur und beeinflussen maßgeblich, wie Ereignisse kognitiv konstruiert werden (vgl. Rissman & Majid, 2019; Ünal, Ji & Papafragou, 2021). In diesem Zusammenhang wird das Konzept der „thematischen Rolle“ eingeführt, das die Beziehungen zwischen den Ereignisteilnehmern definiert (vgl. Dowty, 1989). In der Linguistik beschreiben thematische Rollen, in welcher Beziehung Satzteile mit einem Verb stehen. Zu den thematischen Rollen zählen beispielsweise Agens, Patiens und Instrument.

In psycholinguistischen Studien wurde die psychologische Realität thematischer Rollen untersucht. In der Studie von Griffin und Bock (2000) wurden den Versuchspersonen Bilder von einfachen kausativen Ereignissen präsentiert. Sie sollten das Patiens so schnell wie möglich visuell fixieren. Die Ergebnisse zeigten eine Präferenz für die Fixation auf das Patiens bereits durchschnittlich 288 ms nach dem Bild-Onset. Dies deutet darauf hin, dass die Versuchspersonen die Kohärenz der Ereignisstruktur auf einen Blick erfassten und in der Lage waren, die Patiens-Rolle effizient zu extrahieren und zu identifizieren. Die schnelle Identifikation impliziert das Verständnis der zugrunde liegenden semantischen Struktur des Ereignisses.

Im Einklang mit dieser Studie gingen Dobel et al. (2007) einen Schritt weiter, indem sie die Erkennung von Szenen mit komplexeren Ereignissen und mehreren

thematischen Rollen untersuchten, um festzustellen, inwieweit und wie schnell die Versuchspersonen solche Informationen verarbeiten können. In ihrer Studie wurden Szenen von Schieß- oder Transferereignissen mit jeweils zwei Teilnehmern gezeigt, und zwar einem Agens und einem Rezipienten (z.B. „*Ein Jäger schießt auf einen Tiger*“), wobei die Darbietungsdauer zwischen 100 und 300 ms betrug. Die Versuchspersonen wurden aufgefordert, die Teilnehmer in den Ereignissen nicht nur zu identifizieren, sondern sie auch konkret zu benennen. Es zeigte sich, dass selbst bei der kürzesten Darbietungsdauer die Kohärenz der Szene korrekt beurteilt werden konnte. Ab einer Darbietungsdauer von 200 ms wurden mehr als 60% der Agens und Rezipienten korrekt benannt. Darüber hinaus bestand bei allen Darbietungsdauern eine signifikante Korrelation zwischen der Identifizierung von Agens und Rezipient, was darauf hinweist, dass die Versuchspersonen nicht nur einen Teilnehmer des Ereignisses identifizierten und die anderen ignorierten, sondern eine Gesamtvorstellung sowohl der semantischen als auch der strukturellen Ereignisrepräsentation generierten.

Die Extraktion von Ereignisrollen erfolgt nicht nur unter bestimmten Anweisungen. Hafri et al. (2018) untersuchten beispielsweise, ob das mentale System Rolleninformationen (in dieser Studie speziell *Agens* und *Patiens*) schnell und spontan aus einfachen Ereignissen extrahieren kann. In ihrem Experiment sahen die Versuchspersonen die Bilder, die kausative Ereignisse mit zwei Teilnehmern darstellten, z.B. „*Die Frau sieht den Mann an*“, und sollten dann angeben, auf welcher Seite des Bildschirms sich das Target befand. Obwohl die Aufgabe in keinem direkten Zusammenhang mit den thematischen Rollen stand, reagierten die Versuchspersonen langsamer, wenn die thematische Rolle des Targets im Vergleich zum vorherigen Trial anders war. Diese Ergebnisse belegen, dass Menschen in der Lage sind, Agens und Patiens in kausativen Ereignissen schnell und unbewusst zu identifizieren, selbst wenn ihre Aufmerksamkeit durch eine nicht verwandte Aufgabe beansprucht wird. Dies spricht für eine automatische Identifikation von Ereignisrollen.

Diese Studien legen nahe, dass Sekundenbruchteile ausreichen, um auf *high-level*-Repräsentationen visueller Szenen zuzugreifen. Dies ermöglicht ein schnelles und

effizientes Erfassen der Kernstruktur eines Ereignisses. Thematische Rollen werden dabei als Ereigniskomponenten erkannt und im Wahrnehmungsprozess als Elemente einer holistischen Szene als relevante Bestandteile routiniert und mit großer Geschwindigkeit extrahiert.

Rolleninformationen können nicht nur bei der Verarbeitung visueller Reize, sondern auch bei linguistischen Inputs extrahiert werden (z.B. Kim & Osterhout, 2005). Ziegler und Snedeker (2018) fanden in einem Priming-Experiment heraus, dass die Versuchspersonen mehr Präpositionalobjektsätze (PO-Sätze) produzierten, wenn der Prime-Satz dieselbe Abfolge thematischer Rollen aufwies, nämlich *Agens-Thema-Rezipient*. Beispielsweise führte der Satz „*The woman threw the ball to the bird*“ häufiger zur Produktion von PO-Sätzen im Vergleich zu einem Prime-Satz mit derselben syntaktischen Struktur, jedoch einer anderen Abfolge der thematischen Rollen, wie „*The woman raised the ball above the bird*“, bei dem die thematischen Rollen *Agens-Thema-Lokation* sind.

Ein vergleichbarer Effekt wurde auch im Chinesischen nachgewiesen. Cai, Pickering und Branigan (2012) verwendeten das strukturelle Priming-Paradigma und stellten fest, dass chinesische *Bǎ*-Doppelobjektsätze (*Bǎ*-DO-Sätze) zwar die gleiche syntaktische Struktur wie reguläre DO-Sätze aufweisen, *Bǎ*-DO-Sätze jedoch häufiger PO-Sätze statt DO-Sätze primten. Dies war darauf zurückzuführen, dass *Bǎ*-DO-Sätze die gleiche Abfolge der thematischen Rollen wie PO-Sätze aufwiesen, nämlich *Agens-Thema-Rezipient*.

Diese Befunde deuten darauf hin, dass während der linguistischen Verarbeitung die thematischen Rollen der Ereignisteilnehmer extrahiert werden. Die beobachteten Priming-Effekte beruhen demnach nicht auf der syntaktischen Reihenfolge, sondern auf der Reihenfolge der thematischen Rollen. Dies liefert einen Beweis dafür, dass die Sprachverarbeitung auch von der zugrunde liegenden Repräsentation abhängt und nicht ausschließlich von der linearen Wortstellung.

Studien zur Verarbeitung thematischer Rollen haben eine Asymmetrie in deren

Prominenz festgestellt. Das bedeutet, dass nicht alle thematischen Rollen in gleichem Maße verarbeitet werden. Einige Rollen tragen mehr zur Ereignisstruktur bei als andere. Die Rangfolge der thematischen Rollen wird als *thematische Hierarchie* bezeichnet (vgl. Fillmore, 1968; Levin & Rappaport Hovav, 2005). Die thematische Hierarchie bestimmt in linguistischer Hinsicht, wie häufig und in welchen syntaktischen Positionen verschiedene Ereignisrollen kodiert werden (vgl. Ünal, Ji & Papafragou, 2021). Empirische Studien zeigen, dass prominentere Rollen häufiger erwähnt, bevorzugt in syntaktisch hervorgehobenen Positionen (z.B. am Satzanfang) platziert und eher höheren grammatischen Funktionen (z.B. Subjekt) zugeordnet werden (vgl. Cai et al., 2012; Sridhar, 1988; Ünal, Ji & Papafragou, 2021). Darüber hinaus zeigt sich die thematische Hierarchie auch in nicht-sprachlichen kognitiven Aufgaben, denn Objekte, die mit prominenteren Rollen kompatibel sind, werden von Versuchspersonen schneller und akkurat identifiziert (vgl. Wilson et al., 2011, 2014) und besser erinnert (vgl. Lakusta & Landau, 2005).

Obwohl keine allgemeingültige Rangfolge für thematische Rollen existiert, besteht ein breiter Konsens darüber, dass dem *Agens* eine prominentere Rolle zukommt, gefolgt vom *Patiens* (vgl. Baker, 1997; Fillmore, 1968; Haspelmath, 2011; Jackendoff 1972). Fillmore (1968) positioniert beispielsweise das *Agens* an erster Stelle in der kanonischen Reihenfolge der Argumente. Dies impliziert, dass das Agens, wenn es in einem Satz vorkommt, typischerweise als Subjekt kodiert wird, während das *Patiens* häufig als Objekt erscheint. In Übereinstimmung damit argumentiert Dowty (1991), dass das Argument mit den meisten Proto-Agens-Eigenschaften (z.B. Intentionalität und Belebtheit) als Subjekt und das Argument mit den meisten Proto-Patiens-Eigenschaften (z.B. von der Handlung betroffen) als Objekt realisiert wird.

Die Annahme der besonderen Prominenz des Agens wird durch zahlreiche empirische Befunde gestützt. In einem Experiment von Dobel et al. (2007) zeigte sich beispielsweise, dass das Agens bei Darbietungsdauern zwischen 200 und 300 ms besser identifiziert und benannt wurde als der Rezipient. In der Eyetracking-Studie von Wilson et al. (2011) betrachteten die Versuchspersonen Bilder von Bewegungsereignissen, in

denen eine Person (*Agens*) ein Objekt (*Patiens*) mithilfe eines Werkzeugs oder eines Körperteils (*Instrument*) zu einem Ziel bewegte, wie z.B. „*ein Mann harkt mit einem Rechen Blätter in einen Korb*“. Die Versuchspersonen sollten das gefragte Target so schnell wie möglich identifizieren. Es zeigte sich ebenfalls, dass die Versuchspersonen das Agens schneller fixierten als andere thematische Rollen.

2.3. Ereignissegmentierung

Die physische Außenwelt bietet Menschen täglich eine Fülle von Informationen und Aktivitäten. Um diese Informationen zu verstehen, zu strukturieren, im Gedächtnis zu speichern und darüber zu kommunizieren, müssen sie in sinnvolle Einheiten gegliedert werden. Diese Einheiten werden von einigen Autoren als Ereignisse bezeichnet.

Zacks und Tversky (2001, S. 3) definieren ein Ereignis als „*a segment of time at a given location that is conceived by an observer to have a beginning and an end*“. Diese Definition betont einen räumlich-zeitlichen Rahmen, der Dauer, Grenzen und räumliche Struktur als Kernmerkmale eines Ereignisses einschließt (vgl. Ünal, Ji & Papafragou, 2021). Daraus ergeben sich die Fragen, was als Beginn oder Ende gilt und wie der kontinuierliche Zeitverlauf in sinnvolle Einheiten segmentiert wird. Um diese Fragen zu analysieren, stellten Zacks et al. (2007) die *Event Segmentation Theory* (EST) vor, die sich darauf konzentriert, wie kontinuierliche Aktivitäten in diskrete Ereignisse segmentiert werden können, und auf der Grundlage aktuell aktiver Ereignismodelle und im Langzeitgedächtnis verankerter Ereignisschemata Vorhersagen über den weiteren Verlauf der Ereignisrepräsentationen generiert.

Nach dem EST-Mechanismus (Zacks et al., 2007) werden die *bottom-up* sensorischen Inputs jedes Zeitschritts durch die Wahrnehmungsverarbeitung zu den Outputs weitergeleitet. Der Output ist die Vorhersage des möglichen Inputs im nächsten Zeitschritt. Die Ereignismodelle konzentrieren sich auf „das, was gerade passiert“. In den meisten Fällen behalten die aktuellen Ereignismodelle ihre inhaltliche Stabilität über die Zeit bei, was bedeutet, dass sie unempfindlich gegenüber unmittelbaren sensorischen Inputs sind. Diese Stabilität dient dazu, eine konsistente und zuverlässige

Repräsentation des aktuellen Ereignisses bereitzustellen und die fortlaufende Wahrnehmungsverarbeitung zu steuern.

Darüber hinaus werden laufende Ereignismodelle unter Rückgriff auf Ereignisschemata aufgebaut. Ereignisschemata erfassen gemeinsame Eigenschaften von Ereignistypen, wie z.B. die typische sequentielle Struktur der Aktivität oder die typischen Rollen der Beteiligten. Während die Ereignismodelle aktive und zugängliche Repräsentationen aktuell ablaufender Ereignisse sind, speichern die Ereignisschemata Informationen, die Menschen aus früheren Erfahrungen gelernt haben, im Langzeitgedächtnis. Sie fungieren als Gerüst für den Aufbau und die Aktualisierung der jeweils aktuellen Ereignismodelle im Arbeitsgedächtnis. Auf dieser Grundlage wird beim Wahrnehmen der aktuellen Inputs passendes Schemawissen herangezogen. Die Ereignisschemata beeinflussen die Ereignismodelle mit semantischem Wissen und Weltwissen aus dem Gedächtnis in einer *top-down*-Weise. Wenn die auf Basis des aktuellen Ereignismodells generierte Vorhersage (Output) mit dem tatsächlichen Input im nächsten Zeitschritt übereinstimmt, bedeutet dies, dass die Vorhersage von guter Qualität ist und die aktuellen Ereignismodelle gut zum tatsächlichen Ereignis passen. In diesem Fall bleibt der Mechanismus stabil.

Allerdings werden Vorhersagen insbesondere dann ungenauer, wenn sich zentrale perzeptuelle Merkmale wie neue Bewegungen oder zeitliche bzw. räumliche Veränderungen ergeben (vgl. Zacks & Tversky, 2001; Zacks, 2004) oder wenn sich konzeptuelle Informationen wie Ziele und Intentionen verändern (vgl. Kuperberg, 2021; Radvansky & Zacks, 2014; Ünal, Ji & Papafragou, 2021). In solchen Fällen steigt die Diskrepanz zwischen vorhergesagtem und tatsächlichem Input, sodass das aktuelle Ereignismodell nicht mehr hinreichend zu den perzeptuellen Informationen passt.

Nach der EST wird angenommen, dass erhöhte Vorhersagefehler ein Signal für die Aktualisierung des laufenden Ereignismodells sind. Wenn das Modell die eingehenden Informationen nicht mehr in geeigneter Weise erfassen kann, wird ein neues Ereignismodell aufgebaut, das besser an die aktuellen Inputs angepasst ist. Dieses

aktualisierte Ereignismodell bildet wiederum die Grundlage für weitere Vorhersagen über die nahe Zukunft.

Zu diesem Zweck schlägt EST (Zacks et al., 2007) den *Gating*-Mechanismus vor, der aktiviert wird, wenn der Vorhersagefehler zunimmt. Dabei werden die aktuellen Repräsentationen zurückgesetzt und der Einfluss der sensorischen Inputs vorübergehend erhöht. Dadurch wird ein neues, besser passendes Ereignisschema identifiziert und aktiviert, um ein neues Ereignismodell aufzubauen. Auf diese Weise wird das Ereignismodell aktualisiert und kehrt in einen stabilen Zustand zurück. Die stabile Periode des Ereignismodells wird als Ereignis wahrgenommen, während die Periode der Aktualisierung als Ereignisgrenze wahrgenommen wird (vgl. Zacks et al., 2007). Die schematische Darstellung von EST ist in Abbildung 2.1 dargestellt.

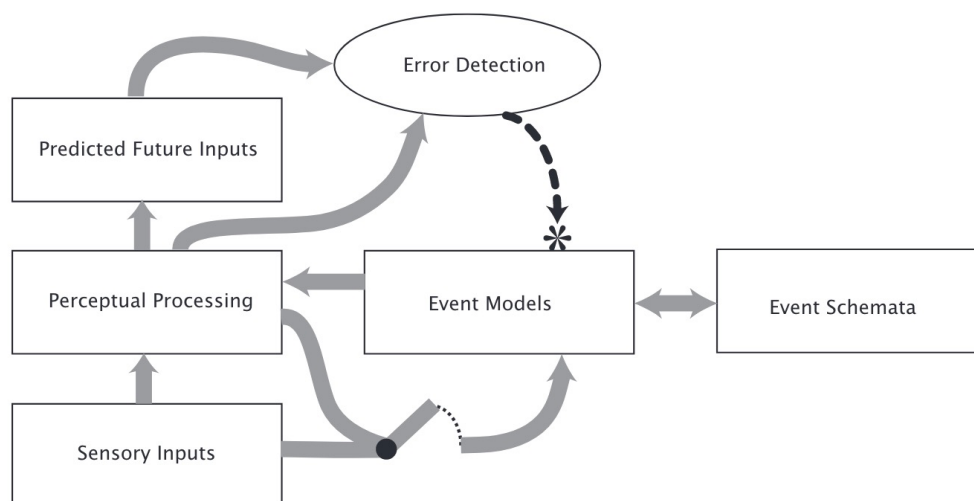


Abbildung 2.1: Modell zur Segmentierung von Ereignissen. Die grauen Pfeile zeigen den Informationsfluss zwischen den Verarbeitungsgebieten an. Die gestrichelten Linien stellen Vorhersagen dar, die zu einer Neuausrichtung der Ereignismodelle führen. Die Verbindung von sensorischen Inputs zu den Ereignismodellen wird durch den *Gating*-Mechanismus gesteuert, der bewirkt, dass die Ereignismodelle nur während der *Reset*-Phase sensorische Inputs erhalten (zit. nach Zacks et al., 2007, S. 274).

Zahlreiche empirische Untersuchungen haben gezeigt, dass wesentliche Veränderungen zu einer Aktualisierung der Ereignismodelle führen können. Newton, Engquist und Bois (1977) analysierten in ihrer Studie, wie die Grenzen von Aktivitätseinheiten definiert werden. Sie zeigten den Versuchspersonen filmische Sequenzen menschlicher

Aktivitäten und ließen diese in sinnvolle Einheiten segmentieren. Dabei wurden die Körperpositionen des Schauspielers in den Filmen kodiert und verglichen. Es stellte sich heraus, dass eine Ereignisgrenze markiert wurde, wenn sich die Körperposition des Schauspielers signifikant änderte.

Diese Ergebnisse wurden von Hard, Tversky und Lang (2006) unterstützt. In dieser Studie wurden Animationsfilme gezeigt, in denen geometrische Figuren jagen oder sich verstecken. Es zeigte sich, dass die Versuchspersonen dazu neigten, Ereignisse bei auffälligen Bewegungs- oder Positionsänderungen zu segmentieren. Diese Studien deuten darauf hin, dass Ereignisse häufig segmentiert werden, wenn sich die aufgenommenen Inputs stärker verändern und somit eine größere Menge neuer Informationen verarbeitet werden muss. Diese neuen Informationen führten zu einer Aktualisierung des Ereignismodells.

Dass Ereignissegmentierung bei Veränderungen auftritt, zeigt sich sowohl in der visuellen Wahrnehmung als auch in der Sprachverarbeitung. In der Studie von Zacks, Speer und Reynolds (2009) wurden den Versuchspersonen sowohl narrative Texte als auch Filme präsentiert, die Veränderungen in Bezug auf Kausalbeziehungen, Raum, Zeit, Charaktere, Objekte und Ziele umfassten. Die Versuchspersonen sollten diese in bedeutungsvolle Ereignisse segmentieren. Die Ergebnisse waren in beiden Modi ähnlich. Es wurde herausgefunden, dass all diese Veränderungen mit einer deutlichen Erhöhung der Wahrscheinlichkeit einer Segmentierung einhergingen. Je mehr Veränderungen ein Satz enthielt, desto wahrscheinlicher wurde er als Ereignisgrenze bewertet. Dies spricht für eine gemeinsame, konzeptuelle Grundlage der Ereignisverarbeitung, die sowohl sprachlich als auch visuell wirksam ist.

Der EST-Mechanismus ist hierarchisch in dem Sinne, dass er auf mehreren Zeitskalen analysiert werden kann. Mit anderen Worten, es gibt verschiedene Granularitätsstufen für die Segmentierung eines Ereignisses. Ein Ereignis kann in kleinere Unterereignisse unterteilt werden, und diese kleineren Ereignisse können wiederum hierarchisch in größere Ereignisse eingebettet werden. Zum Beispiel: Eine Mutter legt Gemüse auf

einen Teller, gießt Salatsoße darüber, brät ein Steak in einer Pfanne, legt das Steak auf einen Teller und stellt alle Teller auf den Tisch. Auf einer gröberen Zeitskala kann die gesamte Szene als ein einziges Ereignis betrachtet werden, nämlich *die Zubereitung einer Mahlzeit*. Auf einer feineren Zeitskala kann jedoch die Zubereitung jeder Speise als separates Ereignis segmentiert werden. Das bedeutet, dass Ereignisse eine partonomische Hierarchie aufweisen, die sich auf die hierarchische Beziehung zwischen Teilen und Unterteilen bezieht (vgl. Hard et al., 2006; Zacks und Tversky, 2001).

Zur empirischen Untersuchung dieses Phänomens wurde häufig die Aufgabe der Ereignissegmentierung verwendet. Dabei wurden den Versuchspersonen Videos von alltäglichen Aktivitäten gezeigt und sie wurden aufgefordert, einen Knopf zu drücken, wenn sie glaubten, dass eine sinnvolle Handlung endete und eine andere begann. In einem derartigen Experiment von Newtonson (1973) wurden die Versuchspersonen angewiesen, die fortlaufenden Aktivitäten entweder in kleinste oder größte Einheiten einzuteilen, die ihnen natürlich und sinnvoll erschienen. Die Ergebnisse zeigten ein hohes Maß an Übereinstimmung hinsichtlich der Einheitengrenzen. Die Versuchspersonen in der feinen Einheit-Bedingung tendierten dazu, einige der Einheiten, die von den Versuchspersonen in der groben Einheit-Bedingung identifiziert wurden, weiter in kleinere Unterteilungen zu unterteilen, wobei die feinen Einheiten Teilmengen der großen Einheiten bildeten. Ähnliche Ergebnisse wurden auch in der Studie von Hanson und Hirst (1989) beobachtet. Diese Studien legen nahe, dass Ereignisse hierarchisch strukturiert sind und dass der EST-Mechanismus gleichzeitig auf verschiedenen Zeitskalen funktioniert. Die Beobachter können je nach Bedarf zwischen den Granularitätsstufen wechseln und selbst wählen, welche Zeitskala sie verwenden möchten.

Es gibt experimentell Hinweise darauf, dass die Ereignissegmentierung tatsächlich einen fundamentalen Prozess der kognitiven Verarbeitung darstellt. Doch geschieht dies nur durch entsprechende Instruktionen oder ist die Segmentierung ein automatisiert ablaufender Prozess? Studien mit funktioneller Magnetresonanztomographie (fMRI)

haben gezeigt, dass es vorübergehende Anstiege der neuronalen Aktivität an den Stellen gab, die die Versuchspersonen später explizit als Grenzen zwischen Ereignissen identifizierten, während sie passiv Videos betrachteten (vgl. Zacks et al., 2001), oder kurze Erzählungen lasen (vgl. Speer, Zacks & Reynolds, 2007). Die neuronalen Aktivitätsmuster bei passiven und aktiven Ereignissegmentierungsaufgaben ähneln sich. Diese Befunde legen nahe, dass die Versuchspersonen ohne explizite Anweisungen in der Lage waren, die kontinuierlichen Aktivitäten während der Wahrnehmung zu segmentieren. Sie unterstützen die Ansicht, dass die Ereignissegmentierung automatisch erfolgt.

2.4. Objektzustände

Um zu untersuchen, wie sich ein Ereignis von einem anderen unterscheiden lässt und wie individuelle Objekttoken von Objekttypen, oder episodische Instanzen eines Ereignisses von generalisiertem Ereigniswissen differenziert werden können, stellen Altmann und Ekves (2019) das *Intersecting Object Histories Model* (IOH-Modell) vor, das sich auf den Inhalt und die Kodierung von Ereignisrepräsentationen fokussiert. Das Modell ist sowohl auf die direkte Wahrnehmung eines Ereignisses als auch auf die indirekte, sprachliche Vermittlung desselben anwendbar.

Ereignisse entfalten sich über Zeit und Raum, wobei sich die daran beteiligten Objekte verändern. Die Ereignisrepräsentationen spiegeln die dynamische Eigenschaft wider, indem sie sich ebenfalls verändern. Das IOH-Modell legt den Fokus auf dynamische Objektrepräsentationen, in denen alle zeitlichen und räumlichen Zustandsänderungen der beteiligten Objekte kodiert sind. Die Trajektorie dieser Objektzustände über Zeit und Raum bildet die einzigartige Geschichte des jeweiligen Objekts ab. Da ein Ereignis viele Teilnehmer umfassen kann, von denen jeder im Laufe der Zeit Zustandsänderungen erfährt und somit seine eigene Trajektorie hat, genügt die Trajektorie eines einzelnen Objekts allein nicht, um ein Ereignis vollständig zu verstehen. Dem IOH-Modell gemäß überschneiden sich alle Trajektorien, wenn sie gemeinsam in Zeit und Raum auftreten. Das Ensemble dieser sich überschneidenden

Trajektorien wird als Ereignis betrachtet (vgl. Altmann & Ekves, 2019).

Ein Beispiel soll dies verdeutlichen: Im Ereignis „*Der Koch hackt die Zwiebel*“ ist jeder Teilnehmer Veränderungen im Laufe der Zeit unterworfen und weist unterschiedliche Zustände während des Fortschreitens des Ereignisses auf. Die Hände des Kochs bewegen sich und die Form der Zwiebel verändert sich. Die Trajektorien dieser Teilnehmer überschneiden sich, weil sie zur gleichen Zeit am gleichen Ort koexistieren. Um das Ereignis richtig zu verstehen, muss man erkennen, dass sich die Objekte von einem vorherigen Zustand zu einem neuen Zustand verändert haben. Dies erfordert die Repräsentation und Aktivierung ihrer jeweiligen vollständigen Trajektorien, einschließlich ihrer initialen und finalen Zustände sowie aller Zwischenzustände. Beispielsweise war die Zwiebel zu Beginn eine vollständige Zwiebel, aber am Ende eine gehackte Zwiebel.

Empirische Studien zeigen, dass beim Ereignisverstehen mehrere Zustandsrepräsentationen desselben Objekts gleichzeitig aktiviert werden und miteinander konkurrieren können. In einem fMRI-Experiment von Hindy et al. (2012) wurden den Versuchspersonen beispielsweise Sätze präsentiert, in denen das Objekt entweder substantiell oder minimal verändert wurde, z.B. „*The squirrel will crack the acorn*“ oder „*The squirrel will sniff the acorn*“. Darauf folgte ein zweiter Satz, in dem das Objekt immer minimal verändert wurde, der aber einen anderen zeitlichen Kontext beschrieb, z. B. „*And then, it will lick the acorn*“ oder „*But first, it will lick the acorn*“. Wenn das Objekt im ersten Satz substantiell verändert wurde, unterschieden sich der initiale Zustand, in diesem Beispiel eine intakte Eichel, und der finale Zustand des Objekts, nämlich eine zerbrochene Eichel, deutlich voneinander. Umgekehrt blieben der initiale und der finale Zustand gleich, wenn das Objekt in beiden Sätzen nur minimal verändert wurde. Die Ergebnisse zeigten, dass der linke posteriore ventrolaterale präfrontale Kortex, der auf semantische Konflikte reagiert, bei den substantiellen Zustandsänderungen aktiviert wurde. Dieser Effekt wurde bei den minimalen Zustandsänderungen nicht beobachtet.

Menschen reagieren nicht nur passiv auf Ereignisse, sondern können auch aktiv vorhersagen, was als Nächstes geschieht. Das IOH-Modell reflektiert diese Fähigkeit. Es übernimmt wichtige Prinzipien des *Simple Recurrent Network Systems*, einschließlich der rekurrenten Eigenschaft, das für die Vorhersage der nächsten Inputs zur Verfügung steht (vgl. Elman, 1990). In diesem System fließen Informationen vom Input zum Output. Die externen Inputs zu jedem Zeitpunkt werden mit den zuvor gespeicherten internen Zuständen des Systems kombiniert und gemeinsam auf den Output abgebildet. Aufgrund dieser rekurrenten Verarbeitung, bei der aktuelle Zustände mit früheren Zuständen über die Zeit hinweg verbunden sind, können Objekttoken mit ihrem früheren Selbst sowie miteinander und mit dem Kontext assoziiert werden. Die Integration aktueller Inputs mit gespeicherten internen Zuständen und Kontextinformationen erlaubt es somit dem System, die Inputs zum nächsten Zeitpunkt vorherzusagen (vgl. Altmann & Mirković, 2009; Elman, 1991).

Im Fall der Sprachverarbeitung könnte das Modell auf Basis des aktuellen und vorherigen Kontexts vorhersagen, wie sich die Sprache zu einem späteren Zeitpunkt entfalten wird. Ebenso könnte es voraussagen, welche Veränderungen in den Zuständen der realen Welt durch die von der Sprache beschriebenen Ereignisstrukturen entstehen werden (vgl. Altmann & Ekves, 2019; Altmann & Mirković, 2009; Elman, 1991).

Ein Ereignis tritt mit Kontext auf. Im IOH-Modell (Altmann & Ekves, 2019) kodieren Ereignisrepräsentationen sowohl die systematische Überschneidung der Objekttrajektorien als auch die Trajektorien anderer zufällig begleitender Elemente, die als Kontext zur Spezifizierung eines Ereignisses dienen. Solche Verbindungen bilden den einzigartigen Verlauf des Objekts. Diesen Mechanismus nennen Altmann und Ekves (2019) „*relational binding*“. Er hilft bei der Tokenisierung eines individuellen Objekts und Ereignisses von generalisierten Typen und bei der Enkodierung im episodischen Gedächtnis.

Zum Beispiel könnten in dem spezifischen Ereignis „*Der Koch hackt Zwiebeln*“ neben dem *Koch* und den *Zwiebeln* auch zufällig eine *Schürze* auftauchen. Die Trajektorien

dieser zufälligen begleitenden Elemente überschneiden sich im Laufe der Zeit sowohl untereinander als auch mit den Trajektorien der systematisch verbundenen Teilnehmer, um das Ereignis im Kontext zu verankern. Es ist diese zufällige Überschneidung aller Trajektorien der Objektrepräsentationen, die das einzigartige Ereignis definiert. Wenn sich diese zufälligen Elemente ändern, beispielsweise wenn statt einer *Schürze* ein *geöffneter Wasserhahn* beteiligt ist, ändert sich auch das Ereignis.

Die Frage ist nun, welche Teilnehmer systematisch und welche zufällige Begleitelemente sind? Diese Unterscheidung entwickelt sich mit zunehmender Erfahrung im IOH-Modell. Wie bereits erwähnt, ist das Modell rekurrent. Zu Beginn erfasst das Modell alle Inputs, die während der Rekurrenz jedoch unterschiedlich behandelt werden. Um die Diskrepanz zwischen den tatsächlichen Inputs und den Vorhersagen zu minimieren, werden in diesem Prozess nur die systematisch verbundenen Inputs verstärkt und salienter als andere, da sie mehr Informationen über die weitere Entwicklung des Ereignisses liefern. Im Gegensatz dazu sind unsystematische Inputs weniger relevant für die Strukturierung des Ereignisses und werden daher in ihrer Aktivierung unterdrückt, um die Genauigkeit der Vorhersage zu erhöhen.

Welche Objektrepräsentationen als informativer wahrgenommen werden, hängt von Faktoren wie dem Fokus, der Salienz und vorhandenem semantischem Wissen ab (vgl. Altmann & Ekves, 2019). Beispielsweise werden in dem Ereignis „*Der Koch hackt ...*“ Lebensmittel wie „*die Zwiebel*“ aufgrund des semantischen Kontexts als informativer wahrgenommen als beispielsweise eine *Schürze*. Daher werden die Objektrepräsentationen der systematisch verbundenen Teilnehmer wie der *Zwiebel* salienter und stärker aktiviert, während ereignis-irrelevante Elemente unterdrückt werden.

Des Weiteren können sich die Objekttrajektorien eines Ereignisses mit anderen Ereignissen überschneiden, indem sie mit früheren Inputs kombiniert werden. Das bedeutet, dass Ereignisrepräsentationen noch das Ensemble der Objekttrajektorien über

mehrere Erfahrungen hinweg kodieren. Aus dieser räumlich-zeitlichen Überlappung von Ereignissen können übergeordnete Schemata abstrahiert werden, die das Wissen über den Ereignistyp und die typischen Teilnehmer vermitteln (vgl. Altmann & Ekves, 2019).

Da alle Inputs miteinander, mit ihren jeweiligen vorherigen Zuständen und mit dem Kontext durch den *relational binding* Mechanismus assoziiert sind, führt die Aktivierung einer Repräsentation dazu, dass auch andere überlappende Repräsentationen sowie das entsprechende semantische Wissen aktiviert werden können. Konkret bedeutet dies: Wenn die Repräsentation eines einzelnen Teilnehmers oder eines Ereignisses aktiviert wird, werden auch die im semantischen Gedächtnis gespeicherten Repräsentationen der relevanten potenziellen Teilnehmer, entsprechender Ereignistypen oder Schemata aktiviert, selbst wenn diese nicht explizit erwähnt werden.

Kamide, Altmann und Haywood (2003) schlussfolgerten durch das *Visual-World*-Paradigma, dass das semantische Wissen und Ereigniswissen zur Antizipation der kommenden Wörter beitragen. In ihrem Experiment wurde den Versuchspersonen eine Kirmes-Szene mit einem Mann, einem kleinen Mädchen, einem Motorrad und einem Karussell präsentiert, während sie entweder „*The man will ride the motorbike*“ oder „*The girl will ride the carousel*“ hörten. Es wurde beobachtet, dass die Versuchspersonen mehr auf das Motorrad blickten, wenn sie „*The man will ride ...*“ hörten, und mehr auf das Karussell blickten, wenn sie „*The girl will ride ...*“ hörten. Der mögliche Grund dafür lag darin, dass Agens in den beiden Sätzen unterschiedliche semantische Hinweise für die Antizipation boten. Gemäß dem Weltwissen sowie der semantischen Information ist es plausibel, dass ein Mann eher ein Motorrad als ein Karussell fahren würde, während ein kleines Mädchen eher ein Karussell als ein Motorrad fahren würde.

3. Ereigniskognition und Sprachstrukturen

Menschen sind in der Lage, Ereignisse wahrzunehmen, mental zu repräsentieren, zu verarbeiten und mit anderen darüber zu kommunizieren. Die sprachliche Bezugnahme auf Ereignisse kann über verschiedene sprachliche Strukturen erfolgen, wobei jede mögliche Struktur dazu dient, unterschiedliche Kommunikationsziele zu erreichen. Wie wir sehen werden, geht mit jeder möglichen Struktur außerdem eine Vorhebung verschiedener Aspekte einer Ereignisrepräsentation einher. Bevor wir die kognitiven Grundlagen der im Folgenden behandelten Sprachstrukturen analysieren, betrachten wir die folgenden Beispiele:

Bsp. 3-1:

- a) Hauptsatz: [*Der Mechaniker repariert das Auto.*]
- b) Subjekt-Relativsatz: [*Der Mechaniker, der das Auto repariert*], hat dem Unternehmen mehr als 100.000 Euro gestohlen.
- c) Nominalisierung: [*Die Autoreparatur*] dauert einen Monat.

Wie die Beispiele zeigen, erscheinen Aspekte derselben Ereignisrepräsentation – *X-repariert-Y* – in verschiedenen sprachlichen Strukturen. Dieses Ereignis wird im Weiteren als Referenz-Ereignis bezeichnet. Während 3-1 a) ausschließlich auf die Repräsentation des Referenz-Ereignisses Bezug nimmt, nehmen die beiden anderen Beispiele zusätzlich Bezug auf weitere Ereignisrepräsentationen. Das Referenz-Ereignis ist Teil einer übergeordneten Struktur. Im Folgenden geht es um die Frage, wie sich der unterschiedliche Bezug auf das Referenz-Ereignis aus der Perspektive der kognitiven Linguistik beschreiben lässt.

Hauptsätze sind eigenständige, grammatisch unabhängige Sätze (vgl. Bloomfield, 1933). Die in der vorliegenden Arbeit betrachteten Hauptsätze beschränken sich auf einfache aktive transitive Deklarativsätze wie in Beispiel 3-1 a). Diese Sätze können verwendet werden, um dynamische Ereignisse mit klar definierten Teilnehmern (z.B. *der Mechaniker*; *das Auto*) und deren Beziehungen oder Handlungen (z.B. *reparieren*)

zu beschreiben.

Relativsätze (wir beschränken uns hier auf Subjekt-Relativsätze) werden verwendet, um zusätzliche Informationen zu einem Teilnehmer oder Element im Matrixsatz bereitzustellen (vgl. Holler, 2013; Lehmann, 1986). Dabei wird der betreffende Teilnehmer in den Kontext eines ergänzenden Ereignisses gestellt, auf das der Relativsatz Bezug nimmt. Relativsätze kombinieren somit zwei Ereignisse, und zwar das Hauptereignis im Matrixsatz und das Referenz-Ereignis im Relativsatz. Zum Beispiel beschreibt der Matrixsatz in 3-1 b) die Hauptinformation („*Der Mechaniker hat dem Unternehmen mehr als 100.000 Euro gestohlen*“). Der Relativsatz („*Der Mechaniker, der das Auto repariert*“) liefert die ergänzende Information zum Subjekt „*der Mechaniker*“, indem er ihn durch ein zweites Ereignis näher beschreibt oder einschränkt. Dies macht den Teilnehmer für den Hörer klar identifizierbar und unterscheidbar von weiteren möglichen Referenten.

Ereignisrepräsentationen können auch in deverbalen Nominalisierungen zum Vorschein kommen, wie in Beispiel 3-1 c). Nominalisierungen ermöglichen es, auf ein zuvor erwähntes oder zumindest dem Hörer bekanntes Ereignis in kompakter Form Bezug zu nehmen, um es in einem neuen Kontext wieder aufzugreifen, weiter zu spezifizieren oder weitere Informationen darüber zu geben (vgl. Chen, 2010; Zhan, 1998a). Nominalisierungen können als Subjekt oder Objekt in einem Hauptsatz, Relativsatz oder subordinierten Satz verwendet werden (vgl. Zhan, 1998a). Damit können sie, ähnlich wie Relativsätze, auf Ereignisse Bezug nehmen, die im Kontext übergeordneter Strukturen eine Rolle spielen. Im Unterschied zu Relativsätzen, die einen Teilnehmer eines Ereignisses fokussieren und weitere Informationen über diesen Teilnehmer liefern, liegt der Fokus bei Nominalisierungen jedoch auf dem gesamten Ereignis selbst. Im Satz „*Der Mechaniker hat das Auto sorgfältig repariert. ... Die Autoreparatur dauert einen Monat.*“ wird beispielsweise das Ereignis im ersten Hauptsatz in der Nominalisierung „*Autoreparatur*“ zusammengefasst. Dies erlaubt es dem Sprecher, an das bereits geäußerte Ereignis anzuknüpfen, ohne den ursprünglichen Prozess erneut ausführlich zu beschreiben. Durch die Verwendung dieser Nominalisierung wird eine

Wiederholung einer vollständigen Bezugnahme per Hauptsatz vermieden.

Zusammengefasst können Ereignisrepräsentationen in Hauptsätzen, Subjekt-Relativsätzen und Nominalisierungen auftauchen, wobei jede Struktur unterschiedliche Aspekte eines Ereignisses betont. Hauptsätze können selbständig ein dynamisches Ereignis ausdrücken, wie in Abbildung 3.1a) abstrakt dargestellt. Subjekt-Relativsätze und Nominalisierungen können zwar auf dasselbe Referenz-Ereignis Bezug nehmen, aber sie müssen in einen übergeordneten Kontext eingebettet sein. Subjekt-Relativsätze spezifizieren einen Teilnehmer eines Ereignisses, indem sie ihn mit einem ergänzenden Ereignis in Beziehung setzen, wie Abbildung 3.1 b) zeigt. Nominalisierungen erlauben es, ein Ereignis als Referenzpunkt für zusätzliche Informationen zu verwenden, wie Abbildung 3.1 c) zeigt.

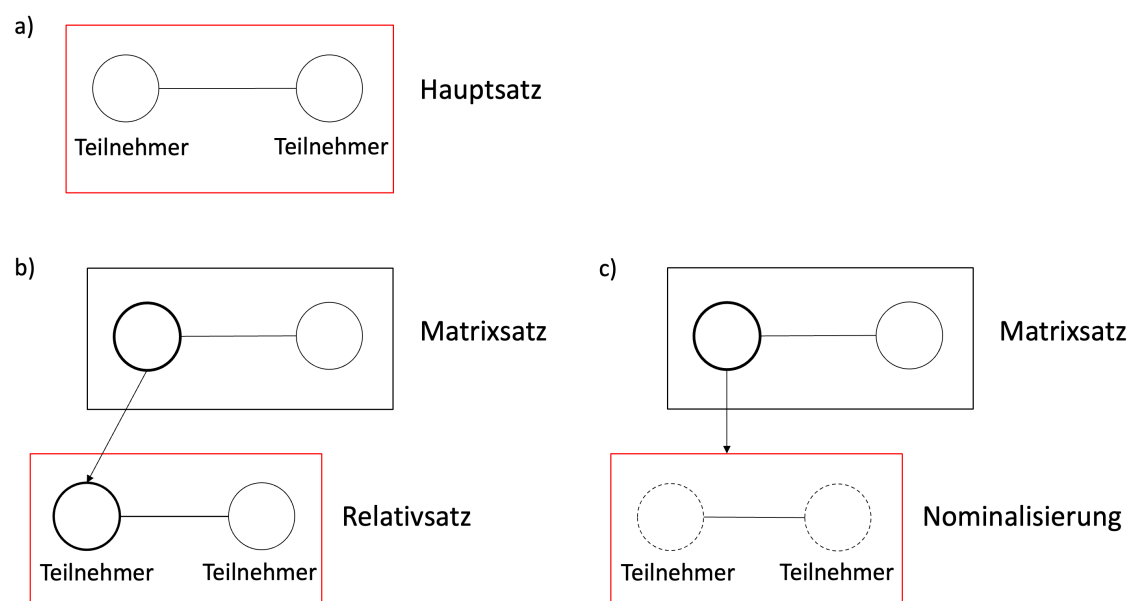


Abbildung 3.1: Abstrakte Darstellungen von Ereignissen in Hauptsätzen, Relativsätzen und Nominalisierungen. Die roten Rechtecke symbolisieren jeweils dasselbe Referenz-Ereignis. Die Kreise stellen die Ereignisteilnehmer dar, die Linie zwischen diesen die Beziehung zwischen den Teilnehmern. Die Pfeile repräsentieren die Co-Referenz.

Diese drei sprachlichen Strukturen weisen deutliche Unterschiede in ihrer oberflächlichen Struktur auf. Die jeweiligen spezifischen Formen im Chinesischen und Deutschen werden in Kapitel 4 detailliert behandelt. Das vorliegende Kapitel konzentriert sich auf folgende Fragen:

- Welche spezifische Konzeptualisierung äußert sich in der konkreten sprachlichen Kodierung eines Referenz-Ereignisses im Hauptsatz, Relativsatz und Nominalisierung?
- In welcher Weise aktivieren diese sprachlichen Strukturen unterschiedliche Aspekte einer Repräsentation eines Referenz-Ereignisses beim Sprachverstehen?

Im vorhergehenden Kapitel wurden kognitionswissenschaftliche Ansätze zur Ereigniskognition dargestellt, die vor allem den Prozess des Aufbaus von Ereignisrepräsentationen in den Blick nehmen. Sie modellieren, wie Informationen in Echtzeit verarbeitet, segmentiert, im Arbeitsgedächtnis organisiert und mit Ereignisschemata im Langzeitgedächtnis integriert werden.

Das vorliegende Kapitel knüpft daran an, richtet den Fokus jedoch stärker auf die konzeptuelle Konturierung der Repräsentationen, die bei der Verarbeitung sprachlicher Äußerungen aufgebaut werden, und auf die Frage, wie verschiedene sprachliche Strukturen unterschiedliche Weisen der Konzeptualisierung desselben Ereignisses nahelegen. Die Konzepte der kognitiven Linguistik bieten hierfür den theoretischen Rahmen, um die strukturierte Form solcher Repräsentationen und ihre Abbildung in sprachliche Muster zu charakterisieren.

Laut Langacker beeinflussen sich *Construal* und sprachliche Struktur wechselseitig. Einerseits wählen Sprecher, je nachdem, wie sie einen konzeptuellen Inhalt konzeptualisieren, z.B. welche Perspektive sie einnehmen oder welche Aspekte sie hervorheben, eine passende sprachliche Struktur. Andererseits beeinflusst die konkrete Formulierung einer Äußerung wiederum, wie Hörer den Sachverhalt konstruieren (vgl. Langacker, 1987, 1991, 2019). Unterschiedliche sprachliche Kodierungen von Ereignissen beruhen auf unterschiedlichen kognitiven Mechanismen und Verarbeitungsprozessen und führen zu unterschiedlichen konzeptuellen Repräsentationen (vgl. Johnson, 1987; Lakoff, 1987; Langacker, 1987, 1991; Talmy, 2000).

In Kapitel 3.1 werden zunächst die konzeptuellen Modelle wie das kanonische Ereignismodell von Langacker (1991) und die *Image-Schemata* von Johnson (1987) und Lakoff (1987) eingeführt. Diese Modelle erfassen abstrakte kognitive Muster, die aus körperlichen und sensomotorischen Erfahrungen in der physischen Welt abgeleitet werden. *Image-Schemata* werden als dynamische, analoge Strukturen definiert, die aus Wahrnehmungen, Körperbewegungen, Manipulationen von Objekten und Erfahrungen physischer Kräfte abstrahiert werden (vgl. Gibbs & Colston, 1995; Johnson, 1987; Lakoff, 1987). In der vorliegenden Arbeit dienen sie dazu, unterschiedliche strukturelle Organisationsweisen von Hauptsätzen, Relativsätzen und Nominalisierungen mit grundlegenden Mustern aus der konkreten physischen Domäne in Beziehung zu setzen. Auf diese Weise lassen sich abstrakte syntaktische Strukturen über relativ grobgranulare, schematische Muster an konkrete Erfahrungen rückbinden und so in ihrer konzeptuellen Struktur erläutern.

Diese konzeptuellen Modelle bieten einen schematischen Rahmen, in dem unterschiedliche sprachliche Strukturen als verschiedene formale Realisierungen derselben grundlegenden kognitiven Szene aufgefasst werden können. Sie erfassen die strukturelle Organisation der sprachlichen Form auf einer relativ grobgranularen Ebene (z.B. als transitive Struktur oder als Kopf-Modifikator-Struktur) und machen so die zugrunde liegenden Unterschiede beim Konstruieren sichtbar.

Zudem erlaubt das Konzept *Construal* eine feiner aufgelöste Beschreibung der Perspektivierung und Aufmerksamkeitslenkung innerhalb derselben Ereignisstruktur. Nach Langacker (2013, 2019) setzt sich sprachliche Bedeutung aus dem konzeptuellen Inhalt einer Äußerung und der spezifischen Art und Weise zusammen, wie dieser Inhalt mental konstruiert wird (*Construal*). Hierbei bezieht sich der Inhalt auf die objektive Situation oder den Sachverhalt, den eine sprachliche Äußerung beschreibt. In einer konzeptualistischen Semantik (*conceptualist semantics*) ist der Inhalt gleichbedeutend mit der neutralen Vorstellung einer Situation. Allerdings kann keine Konzeption völlig objektiv oder neutral sein, da sich die Außenwelt nicht einfach direkt eins zu eins wie eine Kopie mental einprägt, sondern Wahrgenommenes aktiv durch kognitive Prozesse

konstruiert und interpretiert wird (vgl. Langacker, 2019). Daher lässt sich derselbe Inhalt durch unterschiedliche sprachliche Strukturen kodieren, die jeweils spezifische Perspektiven auf diesen Inhalt vermitteln und die kognitive Verarbeitung sowie die Konzeptualisierung des Ereignisses beeinflussen, indem sie die Aufmerksamkeit auf unterschiedliche Aspekte lenken. *Construal* verdeutlicht somit die Beziehung zwischen Sprecher bzw. Zuhörer und der konzeptualisierten Szene (vgl. Langacker, 1987, 1991; Radden & Dirven, 2007).

Construal umfasst mehrere konzeptuelle Dimensionen und Faktoren, die die Interpretation und sprachliche Kodierung von Inhalten beeinflussen und für die semantische Analyse relevant sind. In Kapitel 3.2 werden die drei Dimensionen bzw. Faktoren *Profiling* (welcher Teil eines Ereignisses wird hervorgehoben?), *mentales Scanning* (wie entwickelt sich eine Konzeptualisierung im Laufe der Verarbeitungszeit?) und *Specificity* (wird ein Ereignis allgemein oder detailliert beschrieben?) theoretisch eingeführt und anschließend hinsichtlich ihrer spezifischen Ausprägung in den untersuchten sprachlichen Strukturen analysiert.

Damit schlägt dieses Kapitel eine Brücke zwischen den prozessorientierten kognitionswissenschaftlichen Modellen der Ereigniskognition und der empirischen Untersuchung, wie konkrete sprachliche Strukturen unterschiedliche Ereignisrepräsentationen aktivieren.

3.1. Kognitive Modelle und Image Schema

Menschen interagieren ständig mit ihrer Umwelt durch körperliche Bewegungen, Wahrnehmungen und Handlungen. Sie besitzen die Fähigkeit, diese körperlichen und sensomotorischen Erfahrungen aus der konkreten, physischen Domäne durch konzeptuelle Metaphern auf abstrakte Domänen abzubilden, was zum Verständnis und zur Konstruktion von Konzepten und sprachlicher Bedeutung beiträgt. Dieses metaphorische Mapping ist nicht arbiträr, sondern basiert auf der systematischen Korrelation zwischen den beiden Domänen (vgl. Johnson, 1987).

In Kapitel 3.1.1 werden das *Billiard-Ball*-Modell, die Handlungskette und das kanonische Ereignismodell von Langacker (1991) vorgestellt. Diese Modelle beruhen auf physischen Erfahrungen und werden zur Beschreibung linguistischer Phänomene verwendet. Unter kanonischen Ereignissen versteht Langacker (1991) kausative Ereignisse. Mithilfe dieser Modelle können die konzeptuellen Repräsentationen beschrieben werden, die den drei untersuchten Sprachstrukturen zugrunde liegen. Insbesondere sind die Handlungskette und das kanonische Ereignismodell für die Charakterisierung der prototypischen aktiven transitiven Hauptsätze relevant.

Kapitel 3.1.2. konzentriert sich darauf, wie die drei untersuchten Sprachstrukturen mithilfe der unterschiedlichen *Image*-Schemata, nämlich des *Link*-Schemas und des *Center-Periphery*-Schemas, verstanden werden. Diese *Image*-Schemata dienen als kognitive Grundlage der Sprachstrukturen und helfen dabei, die abstrakten Konzepte sowie Strukturen konkret und präzise zu beschreiben.

3.1.1. Kognitive Modelle

Sprache spiegelt unser Verständnis der Welt wider. Basierend auf der alltäglichen Vorstellung der Welt stellt Langacker (1991) das *Billiard-Ball*-Modell dar, das von diskreten Objekten und deren Bewegungen sowie Interaktionen geprägt ist. Das Modell bietet eine konzeptuelle Grundlage zur Analyse grammatischer Konstrukte. Aus diesem Modell wird die Handlungskette entwickelt, die zusammen mit anderen Modellen das kanonische Ereignismodell bildet. Außerdem wird das Modell noch verwendet, um die grundlegenden Konzepte hinter Nomen und Verben zu erklären, was später in Kapitel 3.2.1 näher erläutert wird.

Das *Billard-Ball*-Modell basiert auf vier grundlegenden Komponenten, und zwar Raum, Zeit, materielle Substanz und Energie. Raum ist der Bereich, in dem materielle Substanz existiert, während Zeit den Bereich definiert, in dem Veränderungen und Energie wahrnehmbar sind. Raum und Zeit bilden damit gemeinsam den multidimensionalen Rahmen, in dem sich die beiden anderen Komponenten

manifestieren. Materielle Substanz können physische Objekte bilden, die als diskret, stabil und raumbesetzend wahrgenommen werden, während Energie durch Bewegung, Interaktion und Veränderung beobachtbar wird, die Zeit erfordert. Diese Elemente werden so konzipiert, dass sie eine Welt bilden, in der sich diskrete Objekte im Raum bewegen und miteinander in Kontakt treten. Diese Bewegung wird durch Energie gesteuert. Wenn diese Bewegungen zu einem starken Kontakt mit anderen Objekten führen, wird die Energie übertragen und weitere Interaktionen verursachen (vgl. Langacker 1991).

Aus den physischen Kontakten zwischen Objekten und dem Energietransfer wird die Handlungskette abstrahiert, um dynamische Interaktionen zwischen den an der Handlung beteiligten Entitäten und die entsprechende Satzstruktur zu beschreiben (vgl. Langacker, 1991; Ungerer & Schmid, 1996). Die Handlungskette beginnt mit einem initialen Objekt, das als Energiequelle gilt und kraftvoll mit dem nächsten Objekt in Kontakt tritt, was zu einem Energietransfer führt. Das zweite Objekt überträgt die Kraft in gleicher Weise dann auf das dritte Objekt. Dieser Prozess setzt sich fort, bis zum letzten Objekt oder bis die Kraft erschöpft ist. Das initiale Objekt in einer solchen Handlungskette wird als *Head* und das letzte Objekt als *Tail* bezeichnet (vgl. Langacker, 1991). Abbildung 3.2 illustriert schematisch, wie Energie entlang einer Handlungskette übertragen wird.

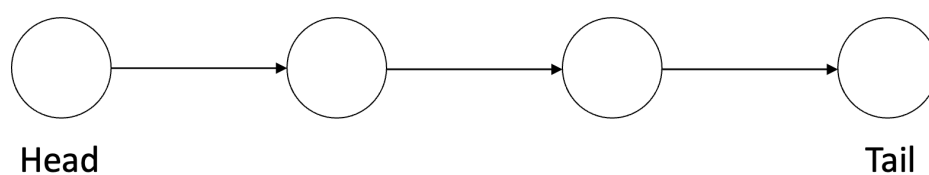


Abbildung 3.2: Die Handlungskette. Die Pfeile stellen die Energieübertragung dar (angepasste Darstellung nach Langacker, 1991, S. 283).

Die einfachste Form einer Handlungskette umfasst nur das initiale und das finale Objekt, ohne Zwischenobjekte. Dabei steht *Head* der Handlungskette konzeptuell dem prototypischen Agens nahe, da er die Handlung absichtsvoll einleitet und die Energie auf ein anderes Objekt überträgt. *Tail* steht in Zusammenhang mit dem prototypischen

Patiens, da er die Auswirkungen der Handlung empfängt und dadurch eine Zustandsänderung erfährt.

Auf das sprachliche Phänomen angewendet, lässt sich das Konzept der einfachsten Handlungskette in einfachen aktiven transitiven Hauptsätzen äußern. Das Subjekt eines einfachen aktiven Hauptsatzes enkodiert typischerweise *Head* der Handlungskette, fungiert als Energiequelle und übernimmt die semantische Rolle des Agens. Das direkte Objekt ist oft der Empfänger der Energie und erfüllt die Rolle des Patiens. Das Verb selbst spiegelt die energetische Interaktion wider und bringt die dynamische Natur der Handlung zwischen den Teilnehmern zum Ausdruck (vgl. Langacker, 1991).

Durch die Kombination der theoretischen Komponenten, und zwar der aus dem *Billiard-Ball-Modell* abgeleiteten einfachsten Handlungskette und des *Stage-Modells*, entwickelte Langacker (1991) das kanonische Ereignismodell. Dieses Modell dient als Grundlage, um prototypische kausative Ereignisse zu beschreiben und ihre Relevanz für die Struktur aktiver finiter transitiver Hauptsätze in der Sprache zu verdeutlichen.

Ein kanonisches Ereignis wird als eine Interaktion beschrieben, die innerhalb eines Rahmens (*Setting*) stattfindet und aus einer bestimmten Perspektive eines externen Betrachters beobachtet und konstruiert wird – vergleichbar mit der Betrachtung einer Bühne. Die zentrale Dynamik eines kanonischen Ereignisses besteht in der Übertragung von Energie durch physische Interaktion, wie sie in der Handlungskette modelliert wird. Dabei führt ein Agens (*Head* der Handlungskette) durch physischen Kontakt eine prototypische Handlung (Energieübertragung) aus, die ein Patiens (*Tail* der Handlungskette) manipuliert oder verändert. Ein kanonisches Ereignis findet typischerweise in einem bestimmten Raum und zu einer bestimmten Zeit statt, die *Setting* bilden. Das kanonische Ereignismodell stellt somit die normale Beobachtung einer prototypischen Handlung zwischen einem Agens und einem Patiens innerhalb eines Rahmens dar, wie in Abbildung 3.3 dargestellt (vgl. Langacker, 1991).

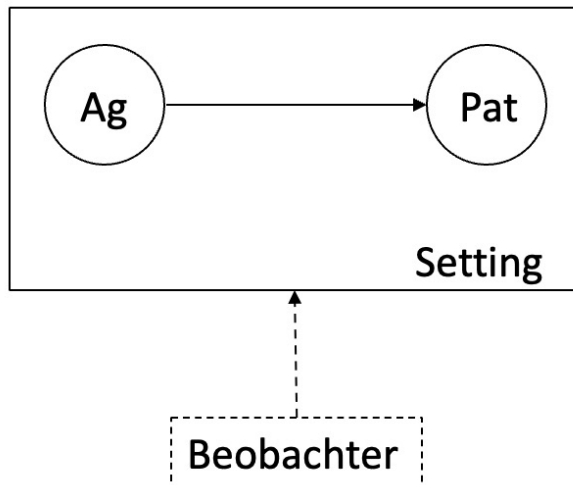


Abbildung 3.3: Das kanonische Ereignismodell mit den zentralen Rollen Agens (Ag) und Patiens (Pat) (angepasste Darstellung nach Langacker, 1991, S. 285).

Das kanonische Ereignismodell weist direkte Parallelen zur grammatischen Struktur prototypischer aktiver transitiver Hauptsätze auf, indem es deren semantische Charakterisierung widerspiegelt (vgl. Langacker, 1991). Ein Beispiel hierfür ist der Hauptsatz „*Der Mechaniker repariert das Auto*“. Dieser Satz beschreibt eine kanonische Handlung, bei der das Subjekt (*der Mechaniker*) und das direkte Objekt (*das Auto*) jeweils mit den prototypischen thematischen Rollen von Agens und Patiens übereinstimmen. Das Verb spezifiziert die Handlung zwischen diesen beiden zentralen Teilnehmern und verankert das Ereignis zeitlich mittels grammatischer Flexion. Somit entsteht eine einfache Handlungskette: *Der Mechaniker* → *Das Auto*.

Sprecher und Hörer entsprechen dem externen Beobachter im kanonischen Ereignismodell. Sie konzeptualisieren das Ereignis aus einer bestimmten Perspektive, was eine linguistische Analogie zur Betrachtung des Ereignisses aus einem externen Blickwinkel darstellt. In prototypischen transitiven Sätzen entspricht die Perspektive des externen Beobachters der sprachlichen Realisierung des Subjekts und Objekts in der dritten Person. Neben diesen Kernteilnehmern (Agens und Patiens) können räumliche und zeitliche Angaben als *Settings* fungieren. Diese sind jedoch nicht Teil der zentralen Handlungskette und können, wenn irrelevant, implizit bleiben (vgl. Langacker, 1991).

Die Sprache stellt jedoch vielfältige grammatische Mittel bereit, um Ereignisse aus unterschiedlichen Perspektiven zu kodieren. Dies ermöglicht es dem externen Beobachter (Sprecher bzw. Hörer), einen eigenen Standpunkt einzunehmen und ein und dasselbe Ereignis auf verschiedene Weise zu konstruieren. Während aktive transitive Hauptsätze das kanonische Ereignismodell unmarkiert widerspiegeln und die gesamte Handlungskette fokussieren, können andere Strukturen von diesem Modell abweichen und dadurch eine veränderte Konzeptualisierung des Ereignisses hervorrufen. Relativsätze wie „*Der Mechaniker, der das Auto repariert*“ zeigen beispielsweise eine perspektivische Verschiebung. Hier dient der Relativsatz als ergänzende oder restriktive Information zur näheren Bestimmung eines Referenten. Diese Vielfalt wird in Kapitel 3.2 detailliert analysiert, wo unterschiedliche *Construals* eines kanonischen Ereignisses in Bezug auf drei untersuchte sprachliche Strukturen analysiert werden.

3.1.2. Image Schema

Angesichts der Bedeutung konkreter sensorischer Erfahrungen für das Verständnis abstrakter sprachlicher Strukturen entwickelten Johnson (1987) und Lakoff (1987) das Konzept des „*Image-Schemas*“. Laut Johnson (1987) ist ein *Image-Schema* ein wiederkehrendes, dynamisches Muster unserer Wahrnehmungsinteraktionen und motorischen Programme, das menschlicher Erfahrung Kohärenz und Struktur verleiht. Lakoff (1987) betrachtet *Image-Schemata* als relativ einfache Strukturen, die stetig in der alltäglichen physischen Erfahrung wiederkehren. Gibbs und Colston (1995) definieren *Image-Schemata* als dynamische, analoge Repräsentationen räumlicher Beziehungen und Bewegungen.

Diese abstrahierten und vereinfachten Strukturen dienen dazu, komplexere Bedeutungen auf der Grundlage konkreter körperlicher Interaktionen und sensorischer Erfahrungen zu konstruieren und sie auf konzeptuelle und sprachliche Strukturen abzubilden. Es werden vielfältige *Image-Schemata* abgeleitet, wie z.B. das *Link-Schema*, das *Center-Periphery-Schema*, das *Source-Path-Goal-Schema*, das *Part-Whole-Schema* etc. (vgl. Johnson, 1987; Lakoff, 1987).

3.1.2.1. Das *Link*-Schema und der Hauptsatz

Das *Link*-Schema (Johnson, 1987; Lakoff, 1987) basiert auf sehr häufig auftretenden physischen und abstrakten Erfahrungen. Bereits früh im Leben erfährt der Mensch die physische Verbindung zwischen Fötus und Mutter über die Nabelschnur. Auch alltägliche Gegenstände wie eine Uhr, die mit einem Nagel oder einem Seil an der Wand befestigt ist, veranschaulichen physische Verbindungen im Raum.

Neben diesen konkreten räumlichen Verbindungen gibt es durch metaphorische Erweiterung auch abstrakte Verbindungen. Dazu zählen beispielsweise soziale Verbindungen zwischen Individuen, Familien, Freunden oder Gesellschaften, logische Verbindungen zwischen Prämissen und Schlussfolgerungen in einem Argument sowie kausale Verbindungen, z.B. wenn ein Ereignis A einem Ereignis B zeitlich vorausgeht und dieses verursacht (vgl. Johnson, 1987).

Aus diesen allgegenwärtigen Verbindungen lässt sich ein wiederkehrendes schematisches Muster abstrahieren. Im vereinfachten Schema gibt es zwei Entitäten, A und B, die durch eine Verbindung (*Link*) miteinander verbunden sind. Abbildung 3.4 zeigt die abstrakte Struktur des *Link*-Schemas. Die grundlegende Logik des Schemas besteht darin, dass, wenn A mit B verbunden ist, A durch B eingeschränkt und von B abhängig ist - und umgekehrt. Das Schema betont die Verbindungen zwischen Entitäten und ihre gegenseitigen Abhängigkeiten (vgl. Johnson, 1987; Lakoff, 1987).



Abbildung 3.4: Das *Link*-Schema (zit. nach Johnson, 1987, S. 118).

Wenn man sich der abstrakten sprachlichen Struktur zuwendet, repräsentiert das *Link*-Schema die grammatischen Beziehungen und die Koreferenzbeziehungen (vgl. Lakoff, 1987). Dieses Schema entspricht den prototypischen einfachen transitiven Hauptsätzen, die Subjekt, Verb und direktes Objekt umfassen, z.B. „*Der Mechaniker repariert das Auto*“. In diesem Zusammenhang wird das Subjekt als die Entität A und das Objekt als

Entität B angesehen. Das Verb fungiert als *Link* zwischen den beiden Entitäten, das die grammatischen Beziehungen von Subjekt und Objekt zuweist und sie semantisch miteinander verbindet.

Mithilfe des *Link*-Schemas, das eng mit der Handlungskette und dem kanonischen Ereignismodell zusammenarbeitet, lassen sich die Struktur und Bedeutung von prototypischen transitiven Hauptsätzen erklären. In der Handlungskette spezifiziert das Verb die dynamische Beziehung, indem es die Energieübertragung bzw. die Interaktion zwischen den Teilnehmern eines Ereignisses beschreibt. Das Verb fungiert somit als ein Bindeglied und verdeutlicht die abstrakte syntaktische und semantische Verbindung zwischen den Teilnehmern im Rahmen eines aktiven transitiven Hauptsatzes.

3.1.2.2. Das Center-Periphery-Schema und der Relativsatz sowie die Nominalisierung

Das *Center-Periphery*-Schema wurde ebenfalls aus wiederkehrenden Mustern menschlicher Wahrnehmung abstrahiert. Es interagiert mit der Aufmerksamkeitslenkung, wobei bestimmte Informationen in den Vordergrund (Zentrum) gestellt werden und andere Informationen in den Hintergrund (Peripherie) treten (vgl. Talmy, 2000).

Das Schema hat eine physiologische Basis und entwickelt sich früh in der menschlichen Wahrnehmung. Bereits Säuglinge lernen, ihre Aufmerksamkeit auf bestimmte Bereiche innerhalb ihres visuellen Feldes zu lenken und so eine Figur vor einem Hintergrund hervorzuheben, der an den Rändern in einen unbestimmten Wahrnehmungshorizont übergeht (vgl. Johnson, 1989).

Das *Center-Periphery*-Schema zeichnet sich durch eine Asymmetrie zwischen dem Zentrum und der Peripherie aus. Das Zentrum ist das wichtigste und relevanteste Element. Es definiert die Identität des Individuums in einer Weise, die die Peripherie nicht tut. Die Peripherie hängt vom Zentrum ab, aber nicht umgekehrt (vgl. Lakoff, 1987).

Das Schema betont die zentrale Bedeutung bestimmter Teile im Vergleich zu peripheren Teilen und hilft dabei, die Aufmerksamkeit zu fokussieren und wichtige von unwichtigen Informationen zu unterscheiden. Beispielsweise erleben Menschen den Kopf und den Rumpf als Zentrum und die Gliedmaßen und Haare als Peripherie. Verletzungen der zentralen Teile sind schwerwiegender als Verletzungen der peripheren Teile. Ebenso hat ein Baum den Stamm als Zentrum und die Blätter als Peripherie. Selbst wenn alle Blätter verloren gehen, bleibt immer noch derselbe Baum übrig (vgl. Lakoff, 1987).

Zentrum und Peripherie sind nicht immer unveränderlich, sondern kontextabhängig und von der Aufmerksamkeitslenkung beeinflusst. Diese dynamische Interaktion zwischen Zentrum und Peripherie lässt sich durch eine alltägliche Erfahrung erklären: Beim Lesen eines Buches in einem Café liegt die primäre Aufmerksamkeit auf dem Buch (Zentrum), während Geräusche, wie die leise Musik im Café, peripher und weniger bewusst im Hintergrund bleiben. Wird jedoch ein Teller laut fallend auf den Boden geworfen, verschiebt sich das Zentrum der Aufmerksamkeit abrupt von dem Buch zu dem Geräusch.

Abbildung 3.5 zeigt das *Center-Periphery*-Schema (vgl. Johnson, 1989). Der schwarze Punkt in der Mitte repräsentiert das Zentrum. Da die Peripherie nicht fest ist, sondern aus dem Zentrum verblasst, wird sie durch gestrichelte Linien dargestellt.

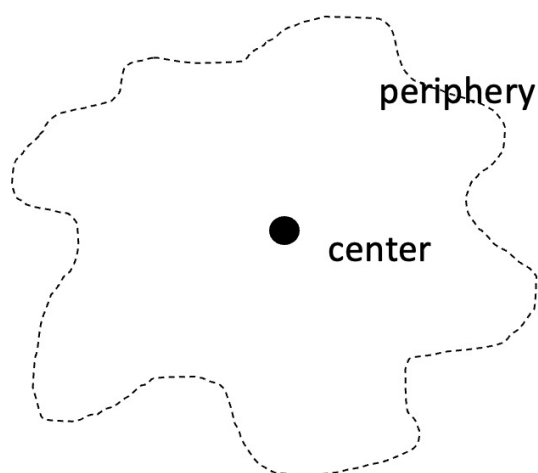


Abbildung 3.5: Das *Center-Periphery*-Schema (angepasste Darstellung nach Johnson, 1989, S. 112).

Durch das metaphorische Mapping lässt sich das *Center-Periphery*-Schema auf abstrakte Sprachdomänen anwenden, insbesondere auf die Kopf-Modifikator-Struktur (vgl. Lakoff, 1987). Dabei entspricht der Kopf dem Zentrum und der Modifikator der Peripherie. So wie in der Wahrnehmung zentrale Elemente deutlicher hervortreten als periphere, dominiert in einer Kopf-Modifikator-Struktur der Kopf, während der Modifikator unterstützende Informationen bereitstellt.

Dieser Eigenschaft entsprechen die Subjekt-Relativsätze. Ein Relativsatz besteht typischerweise aus einem Kopfnominal und einem Relativsyntagma, das den Kopf semantisch modifiziert (vgl. Lehmann, 1992). Die semantische Beziehung zwischen dem Kopfnominal (Nukleus) und dem Relativsyntagma ist asymmetrisch. Während das Kopfnominal der semantische Kern ist, wird das Relativsyntagma als Modifikator des Nukleus betrachtet (vgl. Zifonun, 2001). Das Kopfnominal kann syntaktisch ohne das Relativsyntagma in dem Matrixsatz existieren. Relativsätze sind jedoch subordiniert und bilden keinen eigenständigen Satz (vgl. De Vries, 2001).

Das *Center-Periphery*-Schema hilft hierbei, die Kohärenz und Struktur von Relativsätzen zu verstehen, indem es klar zwischen zentralen und peripheren Elementen unterscheidet und die Beziehung zwischen beiden verdeutlicht. Es spiegelt die Aufmerksamkeitslenkung zwischen zentralen und peripheren Elementen wider und ist eng mit der Fähigkeit verknüpft, bestimmte Objekte oder Informationen aus einem größeren Kontext hervorzuheben.

Bei Relativsätzen steht das Kopfnominal im Zentrum der Aufmerksamkeit und dient als kognitiver Ankerpunkt, um zusätzliche Informationen (Peripherie) aus dem Relativsatz zu integrieren und damit das Kopfnominal weiter zu beschreiben oder einzuschränken. Periphere Informationen erweitern somit das Verständnis des zentralen Elements, indem sie detailliertere und präzisere Bedeutungen generieren.

Beispiel 3-1: Zentrum: Das Buch

└─ Peripherie: das auf dem Tisch liegt

Im Satz „*Das Buch, das auf dem Tisch liegt, gehört mir*“ ist „*das Buch*“ das Zentrum, auf das sich der Satz konzentriert. Das Relativsyntagma „*das auf dem Tisch liegt*“ liefert ergänzende Informationen und schränkt das zentrale Kopfnomen näher ein. Der Matrixsatz bleibt jedoch auch ohne den Relativsatz ein vollständiger und syntaktisch korrekter Satz.

Anhand des kanonischen Ereignismodells ermöglichen Relativsätze es, Ereignisse aus einer spezifischen Perspektive zu betrachten und zusätzliche Details einzubeziehen. Der Sprecher bzw. Hörer richtet dabei den Fokus eher auf einen der beiden Teilnehmer, was zu einer anderen *Construal* eines Referenz-Ereignisses führen kann als bei den prototypischen aktiven transitiven Hauptsätzen.

Analog zu Relativsätzen lassen sich auch Nominalisierungen wie „*die Autoreparatur*“ durch das *Center-Periphery*-Schema erklären. Dieser Typ von Nominalisierung besteht grundsätzlich aus einem Kopf, der die nominalisierte Handlung selbst bezeichnet, sowie einem modifizierenden Element, das den Kopf näher bestimmt. Somit ergibt sich auch hier eine asymmetrische Beziehung zwischen den beiden Elementen.

Beispielsweise bezeichnet das Kopfwort „*Reparatur*“ die grundlegende Handlung, während das Wort „*Auto*“ diese nominalisierte Handlung weiter spezifiziert, indem es angibt, um welche Art von Reparatur es sich handelt. Wie beim Relativsatz bleibt auch hier die zentrale Bedeutung des Kopfs erhalten, selbst wenn das modifizierende Element weggelassen wird, sodass auch der Satz „*Die (Auto)Reparatur dauert sechs Wochen*“ weiterhin grammatisch und semantisch korrekt ist.

3.2. Construal

Construal bezeichnet die Art und Weise, wie ein bestimmter Inhalt mental konstruiert und sprachlich präsentiert wird (vgl. Langacker, 2019). Angesichts der Komplexität der mentalen Welt und der begrenzten Verarbeitungskapazität ist es nicht möglich, alle Aspekte gleichzeitig mental zu repräsentieren, und diese ggf. auf sprachliche Strukturen

abzubilden. Im *Construal* spiegelt sich die Fähigkeit von Sprachbenutzern wider, abhängig von Präferenz, Kommunikationszweck und kognitiver Kompetenz verschiedene mögliche Konzeptualisierungen aufzubauen, was die sprachliche Kodierung bestimmt und es einem Rezipienten erlaubt, ein Ereignis aus unterschiedlichen Perspektiven zu interpretieren. Diese verschiedenen Konzeptualisierungen führen zur Konstruktion ähnlicher, aber nicht identischer Bedeutungen (vgl. Langacker, 1987, 1991, 2013; Verhagen, 2007). In diesem Sinne sind sprachliche Strukturen als Resultat individueller Selektion und Konstruktion durch den Sprecher zu betrachten (vgl. Pütz, 2007; Radden & Dirven, 2007).

Construal umfasst mehrere Dimensionen und Faktoren, die die Kodierung und Konzeptualisierung eines Ereignisses beeinflussen. Im vorliegenden Kapitel werden drei davon behandelt, nämlich *Profiling*, *mentales Scanning* und *Specificity*. Diese Faktoren stehen in enger Wechselwirkung mit der Aufmerksamkeitslenkung, insofern der jeweils ausgewählte Aspekt gewissermaßen als Fokus verstanden werden kann.

Profiling bezeichnet, welcher Teil eines Konzepts als Profil hervorgehoben wird. Dieser Faktor trägt dazu bei, die Bedeutungen linguistischer Strukturen mit demselben konzeptuellen Inhalt zu unterscheiden, indem bestimmte Teile in den Fokus gerückt werden. Beispielsweise haben *Ehemann* und *Ehefrau* denselben konzeptuellen Inhalt, nämlich eine Person innerhalb einer Ehebeziehung, aber ihr Profil betont unterschiedliche Rollen, d.h. *Ehemann* fokussiert auf den männlichen Part und *Ehefrau* auf den weiblichen Part (vgl. Langacker, 2019). *Profiling* hilft zudem bei der Bestimmung der grammatischen Funktion von Nomen, Verben und Nominalisierungen, da jeweils unterschiedliche Aspekte als Profil hervorgehoben werden. Eine detaillierte Analyse erfolgt in Kapitel 3.2.1.

Ereignisse sind dynamisch und entfalten sich über die Zeit. Wenn ein Ereignis sprachlich beschrieben wird, entspricht die zeitliche Verarbeitung dieses Ereignisses jedoch nicht notwendigerweise seinem tatsächlichen Ablauf. An dieser Stelle setzt ein weiterer Faktor des *Construals* an, nämlich das *mentale Scanning*, das sich auf die

Konzeptualisierung komplexer Ereignisse in Bezug auf ihre zeitliche Domäne konzentriert. Es entscheidet darüber, ob ein Ereignis als Prozess oder als Einheit strukturiert und verstanden wird. Dabei werden zwei Modi des mentalen Scannings unterschieden, nämlich das *sequentielle Scanning* und das *holistische Scanning*. Diese Verarbeitungsmodi beeinflussen die Konzeptualisierung und sprachliche Kodierung von Ereignissen mit unterschiedlichen Profilen. Dies wird in Kapitel 3.2.2 näher erläutert.

Das in Kapitel 2.4 vorgestellte IOH-Modell beschreibt, wie Zustandsrepräsentationen von Objekten entlang der Zeitdimension organisiert, miteinander gebunden und für Vorhersagen genutzt werden. Das vorliegende Kapitel ergänzt diese prozessorientierte Perspektive um eine konzeptionsorientierte Sichtweise. Im Mittelpunkt steht die Frage, wie solche zeitlich verknüpften Zustandsrepräsentationen konzeptuell „gescannt“ werden können, beispielsweise als sukzessive Abfolge eines Prozesses oder als zusammengefasste, atemporale Gestalt.

Ein weiterer wesentlicher Faktor bei der Selektion der Inhalte ist *Specificity* und ihr Gegenbegriff *Schematicity*. Dabei geht es darum, wie detailliert und feingranular (*fine-grained*) oder allgemein und grobgranular (*coarse-grained*) ein Ereignis betrachtet und beschrieben wird (vgl. Langacker, 1991, 2013, 2019; Verhagen, 2007). In dieser Hinsicht unterscheiden sich die linguistischen Strukturen in der Granularität der *Specificity*, was sich auf die Aufmerksamkeitsverteilung und auf die spezifische bzw. schematische Verarbeitung von Ereignissen auswirkt. Dies wird in Kapitel 3.2.3 näher beschrieben.

3.2.1. Profiling

Profiling ist der Prozess, durch den mittels eines sprachlichen Ausdrucks Aufmerksamkeit auf einen spezifischen Teil einer komplexen Repräsentation gelenkt wird. Dies ist zentral für die Bedeutungserzeugung, da es bestimmt, welches Element einer konzeptuellen Basis fokussiert wird. Nach Langacker (1987, 1991) kann die

Bedeutung einer sprachlichen Struktur aus zwei Teilen bestehen, nämlich dem Profil und der Basis. Das Profil bezeichnet den hervorstechendsten Teil der konzeptuellen Struktur, der durch den sprachlichen Ausdruck in den Fokus der Aufmerksamkeit gerückt wird. Die Basis bezieht sich hingegen auf die umfassendere konzeptuelle Struktur, die als kognitiver Hintergrund oder Bezugsrahmen dient. Sie stellt die notwendige Grundlage bereit, aus der das Profil hervorgehoben wird und liefert somit den notwendigen Kontext für die Existenz des Profils (vgl. Langacker, 2019).

Derselbe konzeptuelle Inhalt kann durch verschiedene sprachliche Strukturen dargestellt werden, die jeweils ein spezifisches Profil erzeugen. Dies führt dazu, dass das enkodierte komplexe Ereignis angesichts unterschiedlicher Standpunkte und konzeptueller Perspektiven konstruiert wird, damit bestimmte Facetten aktiviert und hervorgehoben werden.

Zur Illustration zeigt Abbildung 3.6 ein Glas, das zur Hälfte mit Wasser gefüllt ist. Auf konzeptueller Ebene ist es wahrscheinlich möglich, diesen Inhalt neutral darzustellen. Bei der sprachlichen Kodierung muss jedoch eine bestimmte sprachliche Struktur gewählt werden, die eine spezifische Perspektivierung mit sich zieht. Wenn das Glas beispielsweise als „*halb voll*“ beschrieben wird, wird das Wasser, das sich noch in dem Glas befindet, profiliert. Wird das Glas hingegen als „*halb leer*“ beschrieben, liegt der Fokus auf dem leeren Raum, der die Hälfte des möglichen Volumens ausmacht (vgl. Langacker, 2013). In diesem Sinne fungieren sprachliche Strukturen als Realisierungsformen für die unterschiedlichen Profile einer Szene.

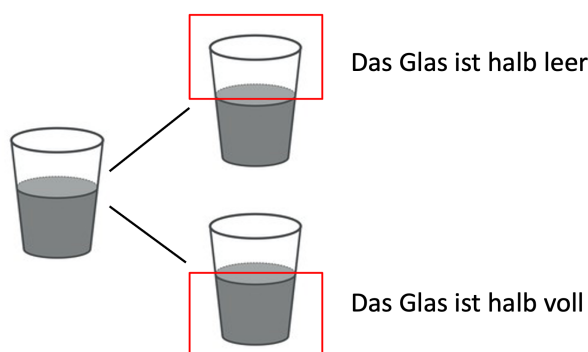


Abbildung 3.6: Unterschiedliche Profile. Die Profile sind rot markiert.

Laut Langacker (2019) hilft das Profil dabei, die grammatische Kategorie einer sprachlichen Einheit zu determinieren. Unterschiedliche Wortarten wie Nomen, Verben, Adjektive und Nominalisierungen zeichnen sich jeweils durch spezifische Profile aus. Da einfache transitive Hauptsätze und Subjekt-Relativsätze typischerweise aus Nomen und Verben bestehen, die jeweils eine zentrale Rolle im Satz spielen, ist ein Verständnis der Profile dieser Wortarten grundlegend für die Analyse satzstrukturbezogener Profile. Nominalisierungen können durch Verben oder verbale Prädikate erzeugt werden. Sie weisen gemischte Eigenschaften von Nomen und Verben auf. Im Folgenden werden zunächst die unterschiedlichen Profile von Nomen und Verben und anschließend die von Nominalisierungen dargestellt.

3.2.1.1. *Nomen*

Nomen und Verben zählen zu den grundlegenden Wortarten fast aller Sprachen. In Kapitel 3.1.1 wird das *Billiard-Ball-Modell* (vgl. Langacker, 1991) vorgestellt, das nicht nur zur Erklärung der Handlungskette dient, sondern auch eine konzeptuelle Grundlage für diese beiden grammatischen Konstrukte bietet.

Dem *Billiard-Ball-Modell* gemäß stehen physische Objekte und energetische Interaktionen konzeptuell im polaren Gegensatz zueinander. Physische Objekte bestehen aus materieller Substanz und sind in der Regel diskret, stabil und räumlich begrenzt. Ein Ball ist z.B. ein typisches Objekt. Sie manifestieren sich primär im Raum. Man kann beispielsweise fragen „*Wo ist der Ball?*“ anstatt „*Wann ist der Ball?*“. Physische Objekte sind konzeptuell autonom, d.h. ihre räumliche Existenz zu einem bestimmten Zeitpunkt hängt nicht von den Interaktionen mit anderen Objekten ab. Im Gegenteil dazu benötigen energetische Interaktionen Zeit, um sich durch Bewegung oder Veränderung zu manifestieren. Interaktionen sind konzeptuell abhängig, da sie mit anderen beteiligten Entitäten verbunden sind (vgl. Langacker, 1991).

Wenn diese Konzepte auf grammatische Strukturen abgebildet werden, repräsentieren Nomen und Verben ebenfalls eine polare Opposition in der Grammatik, die mit den

gegensätzlichen Eigenschaften von physischen Objekten und Interaktionen vergleichbar ist. Physische Objekte und energetische Interaktionen liefern die jeweiligen Prototypen für Nomen und Verben, z.B. „*Buch*“ und „*schlagen*“ (vgl. Langacker, 1991).

Die Klassen der Nomen und Verben sind sehr vielfältig. Viele Nomen und Verben weichen von den jeweiligen Prototypen ab. Beispielsweise sind abstrakte Begriffe wie „*Idee*“ und Kopulaverben wie „*sein*“ keine prototypischen Vertreter der jeweiligen Kategorie. Um dieser Vielfalt gerecht zu werden, schlägt Langacker (1991, S.15) eine allgemeinere und abstraktere Definition für Nomen vor:

„a noun profiles (i.e. designates) a region in some domain, where a region is defined abstractly as a set of interconnected entities.“

Die materielle Substanz physischer Objekte, die in der räumlichen Domäne instanziiert und kontinuierlich verteilt ist, entspricht in diesem abstrakteren Bezug des Schemas auf *entities* (Entitäten). Die Vorstellung der *interconnection* (Verbindung) spiegelt die Kontinuität wider. Die räumliche Domäne wird in diesem Fall durch eine allgemeine Domäne ersetzt, die abstrakte Konzepte wie Zeit oder metaphorische Räume umfasst (vgl. Langacker, 1987, 1991).

Nomen beziehen sich nicht ausschließlich auf Objekte in der räumlichen Domäne wie „*Haus*“, sondern auch auf andere kognitive Domänen. Beispielsweise profiliert das Nomen „*Moment*“ eine Region in der zeitlichen Domäne, und zwar einen bestimmten Moment in der Zeit. Das Nomen „*Grün*“ profiliert eine Region im Farbspektrum, das eine Domäne der Farbwahrnehmung ist. Das Nomen „*Vater*“ profiliert eine Region in der sozialen Domäne der Verwandtschaftsbeziehungen.

Der Begriff *Entität* wird hier in einem maximal allgemeinen Sinne verwendet. Er umfasst alles, was für analytische Zwecke referenziert werden kann, z.B. Objekte, Relationen, Orte, Empfindungen etc. Es ist zu beachten, dass eine Entität nicht diskret, individuell erkennbar oder kognitiv salient sein muss, sondern sowohl materiell als

auch immateriell sein kann, z.B. ein Geräusch oder eine Zeitspanne (vgl. Langacker, 1991).

Bezüglich der Definition steht jedoch nicht der Charakter einzelner Entitäten im Mittelpunkt, sondern die Region, die durch verbundene Entitäten gebildet wird. Auf der abstrakten Ebene sind diese Verbindungen keine objektiv messbare Beziehung, sondern beruhen auf *Construal* und kognitiver Verarbeitung. Zwei oder mehr Entitäten gelten als miteinander verbunden, wenn sie durch kognitive Prozesse oder Operationen als Teil einer integrierten, übergeordneten kognitiven Struktur konstruiert werden. Diese kognitiven Verbindungsoperationen können vielfältiger Art sein. Bereits die Vorstellung, dass mehrere Entitäten zu einer einzigen mentalen Erfahrung gehören, genügt, um sie als Region zu betrachten. Räumliche Nähe oder funktionale Beziehungen zwischen Entitäten führen zu deren Konzeptualisierung als Region. Auch ohne physische Verbindung können mehrere Entitäten als Region erfasst werden, sofern sie als Instanzen desselben Typs erkannt werden (vgl. Langacker, 1991).

Ein Beispiel hierfür ist das Konzept „Stern“. Ein Stern besteht aus Wasserstoff, Helium und Staubwolken. Wenn man über „Stern“ spricht, werden alle Facetten des Sterns gleichzeitig aktiviert, wobei der Fokus nicht auf der materiellen Substanz, sondern auf dem geformten physischen Objekt liegt. Der Stern kann wiederum als eine Entität gelten, die mit anderen Sternen mental verbunden und koaktiviert werden kann, um eine Konstellation zu bilden. Bei der „Konstellation“ liegt die Aufmerksamkeit auf der Gesamteinheit, wobei die einzelnen Sterne ihre individuelle Salienz verlieren. Diese kognitive Verarbeitung, durch die alle Entitäten mental koaktiviert und kollektiv als eine holistische Gestalt wahrgenommen werden, entspricht dem holistischen Scanning, das in Kapitel 3.2.2 detailliert diskutiert wird.

Ein Nomen profiliert eine Region in einer Domäne, wobei die Region sehr flexibel und abstrakt verwendet wird und nicht auf die räumliche Domäne beschränkt ist. Ein physisches Objekt wie ein Baum wird als Region konzeptualisiert, da sich seine Materials substanz kontinuierlich ausdehnt. Ähnlich erfordert auch die Wahrnehmung

einer Linie als einheitliche Region kognitive Operationen, die mehrere diskrete Punkte entlang ihrer räumlichen Ausdehnung mental integrieren.

Um Missverständnisse aufgrund der räumlichen Herkunft des Begriffs „Region“ zu vermeiden, stellte Langacker (1987, 1991) eine einfache und grobe Formulierung vor, nämlich ein Nomen profiliert ein *thing* (Ding). Das Wort *thing* kann sich auf ein konkretes physisches Objekt beziehen, aber in diesem Zusammenhang bezieht es sich auf ein sehr abstraktes Konzept, das fast alle Arten von Nomen umfasst – von konkreten Nomen wie „Tisch“ bis zu abstrakten Nomen wie „Idee“ und von zählbaren Nomen wie „Kuchen“ bis zu unzählbaren Nomen wie „Wasser“.

Ein *thing* wird hierbei als eine zusammenhängende Einheit wahrgenommen und in seiner Gesamtheit kollektiv profiliert, wobei alle beinhalteten Entitäten bzw. Bestandteile ihre individuelle Salienz verlieren. Abbildung 3.7 a) zeigt eine abstrakte Darstellung von Nomen.

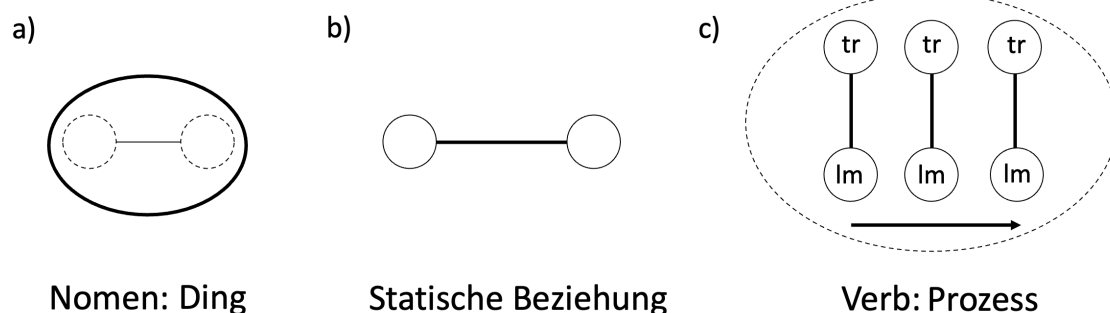


Abbildung 3.7: Schematische Charakterisierung von Nomen und Verben. Bei a) fungieren die beiden gestrichelten Kreise als Vertreter aller verbundenen Entitäten. Die Linie dazwischen stellt die Verbindung dar. Die Ellipse symbolisiert die gebildete Region. Nomen profilieren nicht die einzelne Entität, sondern die Region. Diese Profilierung wird durch die dicke Linie der Ellipse dargestellt. b) zeigt die Hervorhebung der Verbindung bzw. Beziehung zwischen den Entitäten. Da es keinen zeitlichen Ablauf gibt, ist diese Verbindung statisch. c) ist eine abstrakte Darstellung eines Verbs. Die kleinen Kreise mit *tr* (Trajektor) und *lm* (Landmark) repräsentieren zwei saliente Entitäten in einer Beziehung. Der Pfeil symbolisiert die zeitliche Abfolge. Bei Verben werden die Verbindungen bzw. Beziehungen zwischen Entitäten sowie die zeitliche Domäne profiliert, was durch dicke Linien visualisiert wird. Diese drei Zustände stehen exemplarisch für unendlich viele mögliche Zustände entlang der Zeitachse. Alle Zustände können auch eine abstrakte Region bilden, die aber im Fall von Verben nur latent bleibt und daher durch eine gestrichelte Ellipse dargestellt wird.

3.2.1.2. Verben

Im Unterschied zu Nomen profilieren Verben nicht die Region, die durch die verbundenen Entitäten etabliert wird, sondern die Verbindung bzw. Beziehung selbst. Dabei ist zu beachten, dass Verben nicht nur eine einzelne statische Verbindung oder Beziehung profilieren, sondern eine Abfolge von Verbindungen oder Beziehungen zwischen Entitäten, die kontinuierlich über die Zeit verteilt sind. Der zeitliche Rahmen, der sprachlich durch verschiedene Methoden wie Flexion oder Affixe realisiert werden kann, spielt dabei eine bedeutende Rolle. Entlang der Zeitachse gibt es zu jedem Zeitpunkt einen Zustand von Verbindungen oder Beziehungen. Eine kontinuierliche Serie dieser Zustände, die zeitlich geordnet sind, bildet einen Prozess. Verben profilieren also Prozesse, wie in Abbildung 3.7 c) dargestellt (vgl. Langacker, 1987, 1991).

Es ist jedoch nicht erforderlich, dass die Zustände, die entlang der zeitlichen Dimension eines Prozesses verteilt sind, verschieden sein müssen. Es besteht demnach die Möglichkeit, dass alle Zustände gleich bleiben, sodass keine Variabilität von einem Zustand zum nächsten vorliegt. Während sich manche Verben, wie z. B. „*fallen*“ oder „*brechen*“, durch eine Veränderung der Zustände im Laufe der Zeit auszeichnen, stellen andere Verben, wie z. B. „*sein*“ und „*ähneln*“, Zustände dar, die sich über die Zeit hinweg unverändert und konstant bleiben.

Im Gegensatz zu Nomen sind Verben konzeptuell abhängig. Verbindungen bzw. Beziehungen können nicht allein konzeptualisiert werden, ohne auch die beteiligten Entitäten zu konzeptualisieren (vgl. Langacker, 1987, 1991; McRae et al., 2005). In der kognitiven Grammatik werden die salienten inhärenten Entitäten eines Verbs als *Trajektor* (tr) und *Landmark* (lm) bezeichnet. Diese beiden Begriffe werden zwar am häufigsten für die Beschreibung von Bewegungsereignissen verwendet, aber sie haben in diesem Kontext eine breite Anwendbarkeit, um die Prominenz von Teilnehmern innerhalb eines profilierten Prozesses zu beschreiben. Der Trajektor ist der salienteste Teilnehmer in dem profilierten Prozess – normalerweise derjenige, der die Handlung

ausführt oder den Zustand erlebt. Er steht im primären Fokus der Aufmerksamkeit, unabhängig davon, ob er sich bewegt oder statisch ist, und wird sprachlich typischerweise als Subjekt kodiert. Die Landmark ist der sekundär saliente Teilnehmer, der als Referenzpunkt für die Verankerung des Trajektors dient und häufig als internes Objekt der semantischen Struktur betrachtet wird (vgl. Langacker, 1987, 1991).

Trajektor und Landmark sind nicht an spezifische semantische Rollen oder konzeptuelle Inhalte gebunden, sondern werden durch die Prominenz definiert. In einem prototypischen Prozess sind der Trajektor und die Landmark unabhängig, klar unterscheidbar und physisch eindeutig definiert. Dennoch weichen viele Prozesse von diesem Prototyp ab, indem sie lediglich einen Teilnehmer in den Fokus rücken. In solchen Fällen fungiert der einzige saliente Teilnehmer als Trajektor, während die anderen verbundenen Entitäten nicht explizit, individuell identifizierbar oder salient sein müssen (vgl. Langacker, 2013). Beispielsweise erfährt der einzige saliente Teilnehmer „*Baum*“ als Trajektor im Satz „*Der Baum fällt*“ eine Zustandsänderung.

3.2.1.3. Nominalisierung

Nominalisierungen sind sprachliche Konstruktionen, bei denen ein durch ein Verb profiliertes Konzept in eine nominale Struktur umgewandelt wird. Ihr Charakter kann in Bezug auf Nomen und Verben erklärt werden. Die vorliegende Analyse konzentriert sich ausschließlich auf deverbale Nominalisierungen.

Bei der Nominalisierung handelt es sich um eine konzeptuelle Verdinglichung des durch das Verb bezeichneten Prozesses (vgl. Langacker, 1991). Konkret wird ein Prozess, der durch ein Verb profiliert wird, durch die Nominalisierung als eine statische Einheit, also als Ding, hervorgehoben. Da die Nominalisierung mit dem Prozess der Entzeitlichung einhergeht, wird die zeitliche Domäne, die für Verben zentral ist, nicht mehr profiliert (vgl. Langacker, 1991).

Der einfachste Typ der Nominalisierung ist eine Profilverschiebung. Dabei verschiebt sich das Profil von der Beziehung zwischen den Teilnehmern hin zu den Teilnehmern,

die als Teil der inhärenten Struktur hervorgebracht werden. Ein Verb lizenziert in seiner Argumentstruktur bestimmte zentrale Teilnehmerrollen (Trajektor und Landmark), die in seiner semantischen Repräsentation kodiert sind. Darüber hinaus können weitere semantische Rollen wie Instrument (I), Produkt (P) oder Lokation (L) mit der durch das Verb ausgedrückten Ereignisstruktur assoziiert sein. All diese Teilnehmer können durch Nominalisierung profiliert werden. Bei der Nominalisierung wird entschieden, welche Facette des zugrunde liegenden Verbs zur Verdinglichung und Profilierung ausgewählt wird (vgl. Langacker, 1991). Die folgenden Beispiele illustrieren zentrale Typen der Profilverschiebung:

Durch die Nominalisierung kann das Profil von der dynamischen Beziehung auf das Agens bzw. den Trajektor der Handlung verschoben werden, z.B. *malen* – *Maler*, wie in Abbildung 3.8 a) dargestellt.

Ebenso kann das Profil auf das inhärente Patiens bzw. die Landmark der Handlung verschoben werden, z.B. *befragen* – *Befragter*, wie in Abbildung 3.8 b) dargestellt.

Darüber hinaus können auch weitere Teilnehmer profiliert werden, z.B. *scheren* – *Schere*, wobei der Fokus auf das Instrument der Handlung gerichtet wird. Die Nominalisierung *Gemälde* stellt das Produkt der Handlung „malen“ dar.

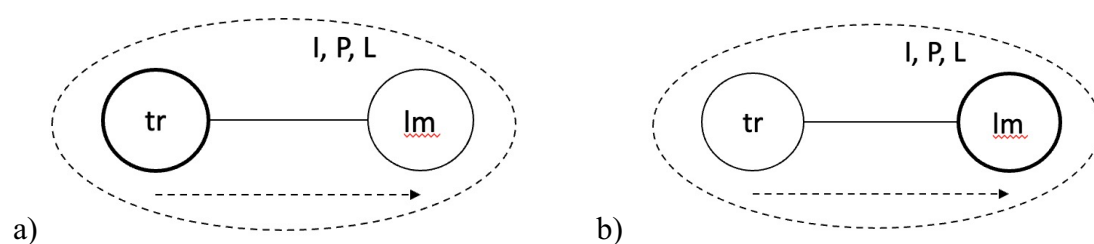


Abbildung 3.8: Schemata der Profilverschiebung (angepasste Darstellung nach Li & Yang, 2020, S. 838). Die Kreise mit dicker Linie markieren die hervorgehobenen Teilnehmer. Die gestrichelten Pfeile stellen die zeitliche Domäne dar, die das zugrunde liegende Verb profiliert. Diese wird bei den Nominalisierungen nicht mehr profiliert.

Ein weiterer Typ der Nominalisierung bezeichnet eine einzelne Episode eines durch ein Verb profilierten Prozesses. Die zugrunde liegenden Verben dieses Typs zeichnen sich entweder durch eine Veränderung von Zuständen im Zeitverlauf aus, wie „fallen“ oder

sie stellen einen Prozess mit klar definiertem Anfang und Ende dar, wie „*Schlafen*“. Wie Abbildung 3.9 zeigt, haben sowohl die zugrunde liegenden Verben als auch ihre Nominalisierungen dieses Typs eine Reihe von Komponentenzuständen. Der Unterschied zwischen beiden besteht darin, wie diese Komponentenzustände kognitiv verarbeitet werden, ob als ein Prozess oder als eine einheitliche Region (vgl. Langacker, 1987, 1991). Die kognitive Verarbeitung für diese Unterscheidung ist das mentale Scanning, auf das in Kapitel 3.2.2 näher eingegangen wird.

Ein Verb enthält eine Reihe zeitlich verteilter Komponentenzustände, von denen jeder eine Beziehung profiliert, wie in Abbildung 3.9 a) dargestellt. Diese Beziehungen können im weiteren Sinne als Entitäten betrachtet werden. Die kognitive Verarbeitung, durch die diese Zustände nacheinander in eine zeitliche Abfolge verarbeitet werden, kann als eine Verbindungsoperation betrachtet werden. So entsteht eine abstrakte Region, die aus diesen miteinander verbundenen Komponentenzuständen besteht. Diese Region bleibt latent bei Verben (vgl. Langacker, 1991).

Die Nominalisierung wandelt die Abfolge von Zuständen in eine einzige, abgeschlossene, statische Einheit um. Dabei liegt das Profil der Nominalisierung auf der Region, wie Abbildung 3.9 b) zeigt. Das Verb „*fallen*“ beschreibt beispielsweise den Prozess, bei dem eine Serie von Zuständen nacheinander in zeitlicher Abfolge verarbeitet wird, also von oben nach unten. Die Nominalisierung „*Fall*“ stellt eine abgeschlossene, spezifische Episode des durch das Verb profilierten Prozesses dar und bezeichnet einen einzelnen Untergang.

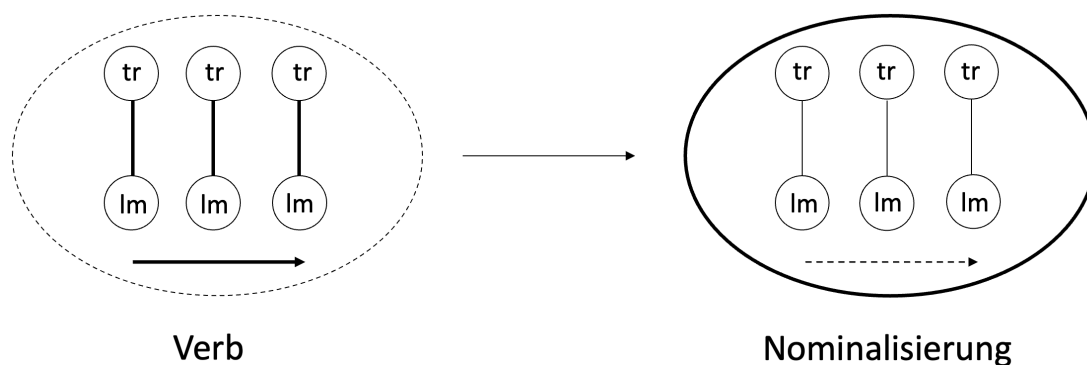


Abbildung 3.9: Schemata für Verben und episodische Nominalisierungen.

Zusammenfassend lässt sich sagen, dass sich Nomen und Verben in Bezug auf das *Profiling* unterscheiden. Nomen profilieren eine abstrakte Region, die aus miteinander verbundenen Entitäten besteht, während Verben die Verbindungen bzw. Beziehungen im Laufe der Zeit profilieren (vgl. Langacker, 1987, 1991, 2013; Tuggy, 2010).

Nun wenden wir uns den untersuchten Sprachstrukturen zu. In prototypischen einfachen transitiven Hauptsätzen nimmt das Verb die zentrale Rolle ein. Da das Profil einer sprachlichen Struktur typischerweise durch ihren Kopf bestimmt wird (vgl. Langacker, 1991, 2013), gibt das Verb das prozessuale Profil des gesamten Hauptsatzes vor. Das Subjekt (Agens) und das Objekt (Patiens) eines aktiven, transitiven Hauptsatzes, die durch das Verb in Beziehung gesetzt werden, spezifizieren die salienten Entitäten. Sie übernehmen die Rollen des Trajektors und der Landmark innerhalb des durch das Verb profilierten Prozesses.

Bei der Nominalisierung handelt es sich um eine Profilverschiebung von den Beziehungen zu einem statischen *thing*, das entweder ein Teilnehmer des ursprünglichen Prozesses oder eine konzeptuelle Verdinglichung dieses Prozesses selbst sein kann.

Ein besonderer Fall sind finite Relativsätze, z.B. „*Der Baum, der fällt*“. Subjekt-Relativsätze werden grundsätzlich als Modifikatoren des jeweils zugehörigen Kopfnomens betrachtet. Sie unterscheiden sich jedoch von anderen prototypischen Nominalmodifikatoren wie Adjektiven. Werden Adjektive als Modifikatoren in eine übergeordnete Nominalstruktur eingebettet, wie beispielsweise in „*der alte Baum*“, übernehmen sie das nominale Profil des Kopfnomens (vgl. Langacker, 1991).

Im Unterschied dazu sind finite Subjekt-Relativsätze prozessual, obwohl sie ebenfalls ein Kopfnomen modifizieren. Die Verben in den Relativsätzen profilieren die Prozesse, die durch zeitliche Markierungen wie Präsens oder Präteritum in Bezug auf die spezifische Sprechzeit und den Sprachkontext verankert werden. Diese Verankerung unterscheidet finite Relativsätze von anderen nominalen Modifikatoren und ermöglicht

es, dass der im Relativsatz dargestellte Prozess primär als gewisse eigenständig konzipiert wird (vgl. Langacker, 2013). Diese Verankerung wird in Kapitel 3.2.3 näher erläutert.

3.2.2. Mentales Scanning

Das mentale Scanning spielt eine bedeutende Rolle bei der Konzeptualisierung und Strukturierung komplexer Szenen und Ereignisse sowie bei der sprachlichen Verarbeitung. Beim mentalen Scanning handelt es sich um einen kognitiven Verarbeitungsprozess, bei dem konzeptuelle Inhalte wie Objekte, Szenen oder Ereignisse auf eine bestimmte Weise mental erfasst oder durchlaufen werden. Ereignisse, die in verschiedenen Sprachstrukturen auftauchen, können durch unterschiedliche mentale Zugangswege verarbeitet werden, was zu verschiedenen *Construals* führen kann (vgl. Langacker, 2013, 2019).

Die Idee des mentalen Scannings knüpft an die im vorhergehenden Kapitel diskutierten Modelle zur Ereigniskognition an, insbesondere an das IOH-Modell von Altmann und Ekves (2019). Das IOH-Modell beschreibt, wie sich verändernde Zustände der beteiligten Objekte über die Zeit hinweg zu Trajektorien verbunden und als ein Ereignis repräsentiert werden. Das Konzept des mentalen Scannings setzt genau an diesem Punkt an, jedoch auf einer anderen Abstraktionsebene. Es beschreibt, wie ein Sprecher oder Hörer diese Abfolge von Zustandsrepräsentationen mental durchläuft, entweder als sequentielle Abfolge eines Prozesses oder als simultan aktivierte, zusammengefasste Gestalt. Auf diese Weise verbindet mentales Scanning die zeitlich strukturierten Zustandsrepräsentationen aus kognitionswissenschaftlichen Modellen mit der Frage, welche Art von mentaler Ereignisrepräsentation durch unterschiedliche sprachliche Strukturen nahegelegt wird.

Das mentale Scanning von Ereignissen fokussiert sich auf die Zeit, die einen Prozess konzeptuell strukturiert. Sie wird damit als Medium der Konzeptualisierung verstanden und ist unabhängig davon, wie Ereignisse in der wahrnehmbaren Außenwelt in Echtzeit

ablaufen. Konzeptualisierung ist inhärent dynamisch. Sie besteht in der mentalen Verarbeitung, korreliert mit neurologischer Aktivität und vollzieht sich daher in der Zeit. Jede wahrnehmbare oder denkbare Aktivität, selbst solche, die augenblicklich zu geschehen scheinen, z.B. das Fühlen des Einstichs einer Nadel oder das plötzliche Erinnern an einen alten Freund, erstreckt sich über eine gewisse Dauer, wenn sie in einer ausreichend feinen zeitlichen Auflösung betrachtet wird. Weil Konzeptualisierung als kognitiver Prozess immer selbst eine Verarbeitungszeit erfordert, ist die konzeptualisierte Zeit (Zeit als Medium der Konzeptualisierung) in der Regel nicht identisch mit der Konzeptualisierungszeit, gemessen in der objektiven Zeit (Sekunden, Minuten usw.) (vgl. Langacker, 2013, 2019).

Ereignisse entfalten sich ebenfalls über die Zeit. Jeder Zustand eines Teilnehmers tritt in einer bestimmten sequentiellen Reihenfolge auf. Diese Abfolge in der außersprachlichen Welt kann sich in der Reihenfolge widerspiegeln, in der Ereignisse mental konzeptualisiert und sprachlich ausgedrückt werden. Dies wird als *temporale Ikonizität* bezeichnet und stellt einen natürlichen mentalen Zugangsweg dar (vgl. Langacker, 2013).

Allerdings ist die *temporale Ikonizität* nicht zwingend. Ereignisse können auch auf andere Weise konzeptualisiert und sprachlich ausgedrückt werden, als sie tatsächlich stattgefunden haben. Die Konzeptualisierung beschränkt sich also nicht auf eine lineare Reihenfolge, sondern kann gleichzeitig in mehreren strukturellen Dimensionen, auf verschiedenen Organisationsebenen und in unterschiedlichen Zeitskalen erfolgen, wie beim Vorhersagen oder Erinnern. Die verschiedenen kognitiven Verarbeitungsweisen haben zur Folge, dass sprachliche Ausdrücke, die sich auf dieselbe objektive Situation beziehen, semantisch variieren können, da die unterschiedlichen Konzeptualisierungen ihren eigenen zeitlichen Verlauf haben und sich im Laufe der Verarbeitungszeit auf eine bestimmte Art und Weise entfalten (vgl. Langacker, 2013).

Gemäß Langacker (1987) werden für die verschiedenen kognitiven Verarbeitungen zur Strukturierung eines Ereignisses bezüglich der zeitlichen Domäne zwei Modi des

mental Scannings unterschieden: das sequentielle Scanning (*sequential scanning*) und das holistische Scanning (*summary scanning*). Diese Modi des mental Scannings resultieren in unterschiedlichen semantischen und perzeptuellen Interpretationen, die das Verständnis von Ereignisrepräsentationen beeinflussen.

Das sequentielle Scanning bezieht sich auf die kontinuierliche Beobachtung verschiedener Zustände, die durch Handlungen oder Objekte in verschiedenen Zeiträumen hervorgerufen werden (vgl. Langacker, 1987). In dynamischen Ereignissen verändern sich die Zustände der Objekte im Laufe der Zeit. Um eine Reihe verschiedener Zustände zusammen als eine kohärente, sich entwickelnde Gestalt wahrnehmen zu können, müssen Beziehungen zwischen den einzelnen Zuständen hergestellt werden. Dabei dient jeder Zustand als Referenzpunkt, der mit dem nächsten Zustand verglichen wird, um Unterschiede zu identifizieren und Veränderungen wahrzunehmen. Sequentielles Scanning umfasst somit die Transformation eines Zustands in den nächsten oder eine kontinuierliche Abfolge dieser Transformationen. Während des Scanvorgangs liegt der Fokus auf der zeitlichen Beziehung zwischen den Zuständen, die zusammen einen Prozess bilden. Es ist zu beachten, dass beim sequentiellen Scanning die Szene sukzessive und nicht gleichzeitig betrachtet wird (vgl. Langacker, 1987).

Das sequentielle Scanning ist ein typischer kognitiver Verarbeitungsmodus für Verben und verbale Prädikate. Ausgehend von Langackers Konzept (1987, 1991) profilieren Verben bzw. verbale Prädikate Beziehungen, die speziell als sich im Laufe der Zeit entwickelnd betrachtet werden – sei es als abgeschlossenes Ereignis oder als Zustand von unbestimmter Dauer. Hierbei steht die zeitliche Domäne nicht direkt im Fokus, sondern hilft, den Verlauf der profilierten Beziehung über die Zeit hinweg zu organisieren (vgl. Langacker, 2019). Zum Beispiel profiliert das Verb *fallen* in erster Linie die räumlichen Bewegungen. Bei der kognitiven Verarbeitung werden die Positionen des fallenden Objekts entlang der vertikalen Achse über die Zeit hinweg sequentiell verfolgt und gescannt. Erst unter Berücksichtigung des zeitlichen Verlaufs kann zwischen *fallen* und *steigen* unterschieden werden.

Ein dynamisches Ereignis, das in einem verbalen Prädikat erscheint, umfasst eine Abfolge separater, miteinander verbundener Konfigurationen oder Zustände. Diese einzelnen Zustände unterscheiden sich typischerweise voneinander, sodass das Verb eine Veränderung im Laufe der Zeit repräsentieren kann. Damit die durch das Verb ausgedrückten Veränderungen erkannt und der Prozess als zusammenhängend und kohärent wahrgenommen werden kann, ist die Kontinuität von einem Zustand zum nächsten erforderlich. In diesem Fall werden die Zustände sowie die Beziehungen – ähnlich wie beim Betrachten eines Films – seriell nacheinander gescannt, sodass der Betrachter der zeitlichen Abfolge folgen kann (vgl. Langacker, 1987, 1991).

Ein Beispiel hierfür ist das Ereignis im Hauptsatz „*Der Baum fällt*“. Das Verb *fallen* evoziert eine komplexe Szene, indem es den Prozess profiliert, bei dem der Baum eine Transformation von einer aufrechten Position in eine liegende Position durchläuft. Dieser Prozess umfasst mehrere Bewegungszustände des Baums, die sich im Laufe der Zeit verändern, z.B. beginnt der Baum zu kippen, fällt weiter und trifft schließlich auf den Boden. Bei der kognitiven Verarbeitung des Hauptsatzes werden diese verschiedenen Zustände während der Bewegung nacheinander in zeitlicher Abfolge mental gescannt. Währenddessen bewegt sich die Aufmerksamkeit entlang der Abfolge der Zustände. Durch diese Weise wird das Ereignis in einem Hauptsatz als ein kontinuierlicher Prozess konzeptualisiert, der sich über die Zeit erstreckt. Das sequentielle Scanning ermöglicht somit eine dynamische, zeitlich sequentielle Strukturierung des Ereignisses, bei der die zeitliche Abfolge im Mittelpunkt steht. Dies wird in Abbildung 3.10 veranschaulicht.

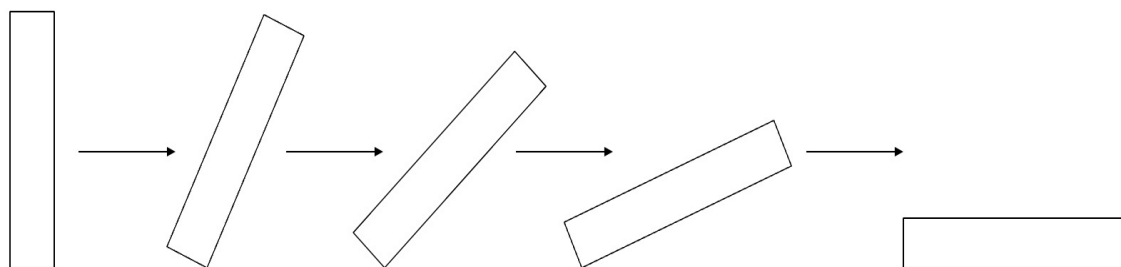


Abbildung 3.10: Schema des sequentiellen Scannings. Die Pfeile indizieren die zeitliche Abfolge. Die fünf Zustände fungieren als Abkürzung für eine kontinuierliche Abfolge solcher Zustände im Prozess.

Auch finite Relativsätze werden sequentiell gescannt, da sie das prozessuale Profil besitzen. Die verschiedenen Aspekte der Ereignisrepräsentationen, die in finiten Subjekt-Relativsätzen auftauchen, werden nacheinander aktiviert, anstatt vollständig in die holistische Perspektive des übergeordneten Kopfnomens eingebunden zu sein (vgl. Langacker, 2013).

Das sequentielle Scanning stellt einen natürlichen mentalen Zugangsweg dar und kann als mentale Simulation der Beobachtung des Ereignisses betrachtet werden, da es Ereignisse in der Reihenfolge verarbeitet, in der sie tatsächlich stattgefunden haben (vgl. Langacker, 2013).

Obwohl sich Ereignisse zeitlich entfalten, können sie kognitiv so betrachtet werden, dass diese zeitliche Domäne nicht im Vordergrund steht. Hierfür steht der andere Modus des mentalen Scannings zur Verfügung, nämlich *das holistische Scanning*. Dieses Scanning beginnt zunächst mit einer kurzen, inkrementellen Aufbauphase, in der die einzelnen Zustandsrepräsentationen eines Ereignisses verarbeitet werden. Im Vergleich zu dem sequentiellen Scanning bleibt jede gescannte Zustandsrepräsentation während des gesamten Scanprozesses aktiv und wird mental übereinandergelegt. Nach dem Scannen des gesamten Ereignisses liegen alle Zustandsrepräsentationen simultan vor und werden miteinander verglichen. Erst durch die Koaktivierung (und ggf. Reaktivierung) aller Zustandsrepräsentationen wird eine kohärente Gestalt gebildet (vgl. Langacker, 1987). Dieser Modus lässt sich mit der Betrachtung eines Mehrfachbelichtungsfotos vergleichen, bei dem alle Details gleichzeitig wahrgenommen werden, was zu einem umfassenden Verständnis des gesamten Bildes führt.

Das holistische Scanning gestattet eine atemporale Konstruktion eines Ereignisses, bei der verschiedene Aspekte bzw. alle Zustände simultan erkannt werden, ohne die zeitliche Abfolge zu betonen, was im Gegensatz zum sequentiellen Scanning steht. Eine typische sprachliche Struktur, bei der das holistische Scanning angewendet wird, ist die Nominalisierung (vgl. Langacker, 1987). Nominalisierungen können auf Ereignisse

Bezug nehmen, wobei das holistische Scanning eine ganzheitliche, statische Perspektive des Ereignisses erzeugt, ohne die zeitliche Domäne zu berücksichtigen. Anhand dieser kognitiven Operation profiliert die Nominalisierung die abstrakte Region, die durch mehrere Zustände etabliert wird, als Gesamtheit.

Ein Beispiel soll dies verdeutlichen: Die Nominalisierung „*das Baumfällen*“ kodiert zwar ein dynamisches Ereignis, doch durch das holistische Scanning wird nicht auf die sukzessive Verarbeitung der verschiedenen Zustände des Fallens (Anfang, Mitte und Ende) in zeitlicher Abfolge fokussiert. Stattdessen werden diese Phasen mental überlagert und als ein Ganzes additiv konzeptualisiert. Sie werden somit nicht als eine Abfolge von Zuständen über die Zeit hinweg betrachtet, sondern als eine einzige komplexe Konfiguration, in der alle Zustände und Facetten simultan koaktiviert und verfügbar sind. Dadurch wird die Gesamtheit des Ereignisses hervorgehoben, wie in Abbildung 3.11 dargestellt.

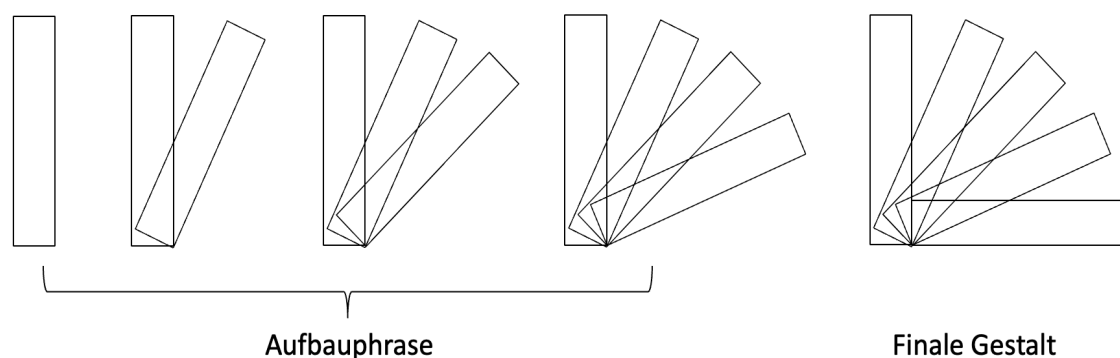


Abbildung 3.11: Schema des holistischen Scannings.

Zusammenfassend lässt sich festhalten, dass bei der kognitiven Verarbeitung eines Ereignisses, das in verschiedenen linguistischen Strukturen mit unterschiedlichen Profilen kodiert wird, zwei Scanning-Modi angewendet werden können. Dies ermöglicht es, unterschiedliche Aspekte eines Ereignisses in den Fokus zu rücken und diese auf vielfältige Weise zu konzeptualisieren und zu verstehen. Bei der Verarbeitung von Haupt- und finiten Relativsätzen wird das sequentielle Scanning angewendet, wobei Ereignisse als dynamische, sich im Laufe der Zeit entfaltende Prozesse betrachtet werden. Für die kognitive Verarbeitung der Nominalisierungen eignet sich hingegen

das holistische Scanning, das additiv und parallel arbeitet. Dabei wird das Ereignis kollektiv als zusammenhängendes Ganzes wahrgenommen, indem alle Zustände simultan aktiviert und zu einer kohärenten Konfiguration komprimiert werden. Dadurch rückt das Gesamtereignis in den Vordergrund.

3.2.3. Specificity

Sprachbenutzer können dasselbe Ereignis durch verschiedene sprachliche Strukturen auf unterschiedlichen Ebenen von *Specificity* und *Schematicity* konstruieren. Beispiele hierfür sind die Nominalisierung „*die Autoreparatur*“, der Relativsatz „*der Mechaniker, der das Auto repariert*“ und der Hauptsatz „*der Mechaniker repariert das Auto*“. Der jeweilige Grad der *Specificity* lässt sich weitergehend analysieren.

Specificity bezeichnet den Grad der „Auflösung“ und „Granularität“ bei der Beschreibung oder Interpretation von Objekten, Ereignissen oder Konzepten, d.h. ob diese fein- oder grobgranular dargestellt werden. Ein hoher Grad an *Specificity* bedeutet eine präzisere und spezifischere Darstellung der Merkmale einer Szene oder eines Objekts, während ein niedriger Grad an *Specificity* zu einer allgemeineren und abstrakteren Beschreibung führt (vgl. Langacker, 1991, 2013). So ist zum Beispiel „*eine blühende gelbe Rose*“ spezifischer als „*Blume*“.

Demgegenüber steht der Begriff *Schematicity*, bei dem Ereignisse auf einer gröberen Ebene betrachtet und dargestellt werden. Dies ermöglicht es einem Sprecher, von Details zu abstrahieren, die in einer bestimmten Szene als irrelevant angesehen werden (vgl. Langacker, 2013, 2017). Beispielsweise ist „*der Behälter mit Flüssigkeit darin*“ schematischer als „*das Glas mit Wasser darin*“.

Specificity und *Schematicity* gelten als Aspekte der Selektions- und Konstruktionsoperationen, die es ermöglichen, Szenen auf verschiedenen Granularitätsstufen zu konstruieren und zu konzeptualisieren (vgl. Langacker, 2013; Mousavi et al., 2019).

Der Grad an *Specificity* und *Schematicity* wird von Talmy (2000) als „*level of particularity*“ bezeichnet. Die Wahl zwischen spezifischen und schematischen Ausdrücken steht in engem Zusammenhang mit der Aufmerksamkeitsverteilung. Spezifischere Formulierungen lenken die Aufmerksamkeit auf detailreiche Aspekte eines Ereignisses und erlauben eine feinere Verarbeitung, während schematischere Formulierungen von Details abstrahieren und die Aufmerksamkeit auf größere Strukturen der Szene bündeln (vgl. Langacker, 1991; Verhagen, 2007).

Langacker (1991) unterscheidet zwischen einfachen Nomen und vollständigen Nominalphrasen sowie zwischen Verbstämmen und finiten Sätzen hinsichtlich der *Specificity* des Referenten. Zentral ist dabei das Konzept der „Verankerung (*ground*)“. Die Verankerung kann durch grammatische Elemente wie Zeit- oder Modusmarkierungen erfolgen, indem sie Informationen über das Wissen oder die Überzeugung des Sprechers in Bezug auf die Sprechzeit oder die Gewissheit des beschriebenen Ereignisses vermitteln. Eine Entität ist verankert (*grounded*), wenn sie von Sprachbenutzern im Sprachkontext eindeutig spezifiziert wird und sich von anderen unterscheidet. Ähnliches gilt für Prozesse, wenn sie zeitlich und modal situiert sind (vgl. Langacker, 1991, 2013).

In diesem Zusammenhang bezeichnet ein einfaches Nomen, wie „*Auto*“, einen übergeordneten Typ, während eine komplexe Nominalphrase mit bestimmtem Artikel wie „*das alte Auto*“, eine verankerte Instanz (*grounded instance*) dieses Typs spezifiziert, die sich von anderen Instanzen unterscheidet und im Wissen des Sprechers bzw. Hörers lokalisiert wird. Hierbei haben grammatische Elemente wie Artikel und Demonstrativa die Funktion, das Nomen in Bezug auf den Sprachkontext zu verankern und zu spezifizieren. In diesem Fall wird das einfache Nomen schematischer betrachtet als die vollständige Nominalphrase.

Analog dazu beschreibt ein Verbstamm wie „*reparieren*“ einen allgemeinen Prozesstyp, der potenziell zu jeder Zeit und mit jedem geeigneten Teilnehmer instanziiert werden kann. Im Gegensatz dazu evoziert ein vollständiger finiter Satz wie „*Der Mechaniker*

reparierte das Auto“ eine verankerte Instanz dieses Prozesstyps, da der Prozess zeitlich spezifiziert und mit konkreten Teilnehmern verbunden ist (vgl. Langacker, 1991). Dies erhöht die Präzision und Auflösung des beschriebenen Ereignisses und lenkt die Aufmerksamkeit stärker auf die spezifischen Informationen, was eine präzisere mentale Repräsentation des Ereignisses im Kopf des Hörers ermöglicht.

Der Grad der *Specificity* von Nominalisierungen hängt mit den zugrunde liegenden sprachlichen Einheiten zusammen. Aus einem Verbstamm wie „*reparieren*“ entsteht durch Nominalisierung ein schematischer Typ (*Reparatur*), der weder Zeit, Ort noch beteiligte Personen spezifiziert. Nominalisierungen wie „*die Autoreparatur*“ oder die Nominalphrasen wie „*die Autoreparatur durch den Mechaniker*“ erhöhen die *Specificity*, indem sie das Patiens und ggf. das Agens explizit machen und den Referenzbereich einengen. Dennoch bleiben diese Phrasen schematischer als finite Sätze, da sie typischerweise keine explizite zeitliche Verankerung enthalten.

Nun kehren wir zu den zu Beginn des Unterkapitels genannten Beispielen zurück, um die Unterschiede an *Specificity* detaillierter herauszuarbeiten. Der Hauptsatz „*Der Mechaniker repariert das Auto*“ charakterisiert eine verankerte Instanz und weist daher eine hohe *Specificity* auf, da er in einen zeitlichen Rahmen verankert ist und die beiden wichtigen Teilnehmer konkret spezifiziert.

Der Relativsatz „*Der Mechaniker, der das Auto repariert*“ zeigt ebenfalls eine hohe *Specificity*. Das Nomen „*Mechaniker*“, das einen allgemeinen Typ bezeichnet, wird durch den bestimmten Artikel „*der*“ zu einer spezifischen Instanz. Das Relativsyntagma „*der das Auto repariert*“ präzisiert das Kopfnomen zusätzlich. Ein finiter Relativsatz enthält ein Verb, das durch Numerus, Tempus und Modus grammatisch markiert ist. Dadurch wird der im Relativsatz kodierte Prozess in einem spezifischen zeitlichen und modalen Kontext verankert (vgl. Langacker, 2013).

Die Nominalisierung „*Die Autoreparatur*“ weist einen höheren Grad an Generalität und *Schematicity* im Vergleich zum Relativ- und Hauptsatz auf. Das Patiens ist zwar explizit,

das Agens bleibt implizit. Die Nominalisierung bietet somit eine schematische Darstellung des Ereignisses, ohne spezifische Informationen zu Akteur, Zeitpunkt oder Kontext. Nach Talmy (2000) rückt die explizite Erwähnung bestimmter Ereignisbestandteile diese in den Vordergrund der Aufmerksamkeit, während nicht genannte Komponenten in den Hintergrund treten. Diesbezüglich lenkt die Nominalisierung „*Die Autoreparatur*“ die Aufmerksamkeit vor allem auf das Konzept der Reparatur sowie auf das explizit erwähnte Patiens, während andere Aspekte des Ereignisses schematisch bleiben.

4. Sprachstrukturen im Chinesischen und Deutschen

Ereignisse können, je nach Kommunikationsziel, in unterschiedlichen sprachlichen Strukturen kodiert werden. Nach den Konzepten der kognitiven Linguistik beeinflussen sich die sprachliche Kodierung und *Construal* eines Ereignisses gegenseitig (vgl. Kapitel 3; Johnson, 1987; Lakoff, 1987; Langacker, 1987, 1991; Talmy, 2000). Diesbezüglich ist anzunehmen, dass den drei untersuchten sprachlichen Strukturen, nämlich einfache aktive transitive Hauptsätze, Subjekt-Relativsätze und deverbale Nominalisierungen, jeweils unterschiedliche Konzeptualisierungen zu Grunde liegen. Diese Unterschiede sind durch unterschiedliche *Construals* bedingt und gehen daher gemäß den oben beschriebenen Theorien mit unterschiedlichen Aufmerksamkeitsfoki einher. Beim Verstehen der drei Strukturen ist daher davon auszugehen, dass sie die Aufmerksamkeit eines Hörers auf unterschiedliche Weise lenken.

Da sich die drei Strukturen abgesehen von ihrem *Construal* auf kognitiver Ebene auch in ihrer Oberflächenform auf sprachlicher Ebene unterscheiden können, insbesondere aus der Perspektive des Sprachvergleichs, ergibt sich die Möglichkeit, *Construal* als kognitiv realen Prozess zu isolieren. Unabhängig von der Realisierung auf der sprachlichen Oberfläche sollten sich beim Vergleich von Hauptsätzen, Subjekt-Relativsätzen und Nominalisierungen zwischen zwei Sprachen – hier Chinesisch und Deutsch – trotz der Unterschiede in der Oberflächenform ähnliche Muster der

Aufmerksamkeitslenkung zeigen.

Chinesisch und Deutsch weisen erhebliche Unterschiede in ihren morphologischen und syntaktischen Systemen auf. Chinesisch ist eine isolierende Sprache und verfügt über keine morphologische Flexion wie Deklination oder Konjugation. Deutsch hingegen zählt zu den flektierenden Sprachen und ist durch eine Vielzahl morphologischer Markierungen gekennzeichnet. Hinsichtlich der Wortstellung weisen beide Sprachen im Hauptsatz grundsätzlich eine Subjekt-Verb-Objekt-Abfolge (SVO) auf. In den Nebensatzstrukturen unterscheiden sie sich jedoch deutlich: Chinesisch hat die kopf-finale Eigenschaft, während Deutsch typischerweise kopf-initiale Strukturen verwendet (vgl. Comrie, 1989; Greenberg, 1963, 1966). Dementsprechend zeigen beide Sprachen unterschiedliche Strategien und Oberflächenstrukturen bei der Bildung der drei untersuchten sprachlichen Strukturen. Besonders die unterschiedliche Wortstellung in Subjekt-Relativsätzen bietet einen wichtigen Ansatzpunkt, um zu untersuchen, ob und inwieweit die syntaktische Struktur die Konzeptualisierung und Dekodierung der Ereignisrepräsentationen beeinflusst, was den Vergleich zwischen Chinesisch und Deutsch besonders aufschlussreich macht.

In diesem Kapitel werden die jeweiligen spezifischen Oberflächenstrukturen und Bildungsstrategien der drei sprachlichen Strukturen im Chinesischen und im Deutschen behandelt. Darüber hinaus wird erörtert, wie die in Kapitel 3 beschriebenen kognitiven Mechanismen und Verarbeitungsprozesse bei diesen spezifischen Strukturen jeweils sprachabhängig zur Anwendung kommen.

4.1. Hauptsatz

Hauptsätze sind sowohl im Chinesischen als auch im Deutschen eine fundamentale Satzstruktur. Ein Hauptsatz wird typischerweise als eigenständiger Satz betrachtet, der grammatisch unabhängig von einem anderen Satz ist (vgl. Bloomfield, 1933). Je nach Natur des Prädikats und Funktion lassen sich Hauptsätze in verschiedene Kategorien einteilen, wie z.B. transitive und intransitive Sätze, Aktiv- und Passivsätze, Deklarativ-

und Imperativsätze, von denen es jeweils mehrere Varianten gibt. Die in der vorliegenden Arbeit betrachteten Hauptsätze beschränken sich auf die einfachen aktiven transitiven Hauptsätze.

In semantischer Hinsicht dient der Hauptsatz vor allem dazu, Aussagen über die Welt zu machen, die wahr oder falsch sein können (vgl. Krifka, 2024; Oppenrieder, 2012). Solche Aussagen können sich auf Ereignisse beziehen. Wie in Kapitel 3.1 erläutert, spiegelt die Struktur eines einfachen aktiven transitiven Hauptsatzes das kanonische Ereignismodell wider. Konkret wird das Agens, das die Handlung initiiert, als Subjekt kodiert und das Patiens, das die Handlung erfährt, als Objekt kodiert. Das Verb spezifiziert die Energieübertragung. Diese Satzstruktur gilt daher als prototypischer Satztyp zur Kodierung eines kanonischen kausativen Ereignisses (vgl. Langacker, 1991).

In dieser Satzstruktur ist das Verb von zentraler Bedeutung. Es fungiert als Kern des Satzes und bestimmt dessen Profil (vgl. Langacker, 1991). Ein Verb an sich profiliert eine Reihe von Beziehungen zwischen inhärenten Teilnehmern, wobei diese Teilnehmer zunächst nur schematisch impliziert sind. Erst durch die Hinzufügung eines Subjektnomens (X) und eines Objektnomens (Y) werden die salienten Teilnehmer, nämlich der Trajektor und die Landmark, explizit spezifiziert. Langacker (1991) veranschaulicht die abstrakte semantische Darstellung dieser Satzstruktur in Abbildung 4.1.

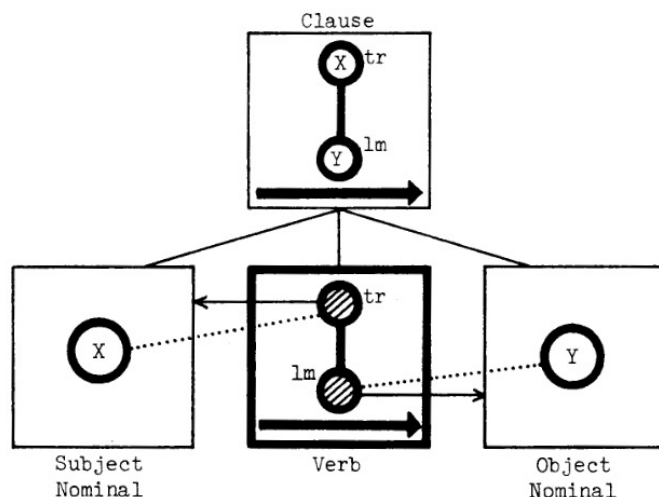


Abbildung 4.1: Abstrakte Darstellung eines finiten Satzes (zit. nach Langacker, 1991, S. 37). Die Korrespondenzen zwischen dem Trajektor und dem Subjekt sowie zwischen der Landmark und dem Objekt sind durch Pfeile dargestellt. Da das Verb den Kern des Satzes bildet, ist das Rechteck, das das Verb repräsentiert, fett hervorgehoben.

Im Folgenden werden die Strukturen des Hauptsatzes im Chinesischen und im Deutschen vorgestellt sowie sprachspezifische Unterschiede in der Satzbildung und deren kognitive Verarbeitung analysiert.

4.1.1. Chinesischer Hauptsatz

Morphologische und syntaktische Struktur

Chinesische Hauptsätze folgen typischerweise der SVO-Wortstellung. Die Grundstruktur lautet (vgl. Greenberg, 1963; Li & Thompson, 1981):

[Subjekt + Verb + Objekt]

Bsp. 4-1 机械师 修理 汽车
 jīxièshī xiūlǐ qìchē
 Mechaniker reparieren Auto
 „der Mechaniker repariert das Auto“

Da das Chinesische eine isolierende Sprache ist, unterliegen Wörter keiner morphologischen Flexion. Substantive werden nicht dekliniert und Verben weder flektiert noch konjugiert. Numerus, Tempus und Modus werden durch das Hinzufügen

von Partikeln oder Markern realisiert. Zudem gibt es keine Kasus- und Genusmarkierungen. Daher spielt die Wortstellung eine zentrale Rolle bei der Bedeutungszuweisung. Insbesondere in einfachen aktiven transitiven Hauptsätzen ohne zusätzliche Marker ist die Reihenfolge von Subjekt und Objekt entscheidend für die Zuweisung der thematischen Rollen. Eine Umstellung dieser Reihenfolge kann in solchen Fällen die Interpretation des Satzes verändern.

Kognitive Verarbeitung

Verben kodieren Prozesse, die konzeptuell von den an ihnen beteiligten Entitäten näher bestimmt sind. In ihrer semantischen Struktur sind bestimmte zentrale Teilnehmerrollen, nämlich Trajektor und Landmark, als logische Argumente angelegt bzw. lizenziert (vgl. Langacker, 1987, 1991). In einem einfachen aktiven transitiven Hauptsatz im Chinesischen werden diese potenziellen Teilnehmer durch Subjekt und Objekt explizit realisiert. Dabei übernimmt der Trajektor des Prozesses die Subjektposition, während die Landmark als Objekt kodiert wird. Diese syntaktische Struktur führt zu einem hohen Grad an *Specificity* in der Ereigniskonzeptualisierung, da durch die Kombination mit konkreten Teilnehmern sowie möglichen Tempus- und Aspektmarkierungen ein spezifisches, situativ verankertes Ereignis ausgedrückt werden kann.

Der Bildungsprozess von Beispiel 4-1 ist in Abbildung 4.2 schematisch dargestellt. Das Verb *xiūlǐ* (reparieren) impliziert zwei Teilnehmer: einen, der die Handlung ausführt, und einen, der von der Handlung betroffen ist. In diesem Beispiel werden sie durch *jīxièshī* (Mechaniker) und *qìchē* (Auto) realisiert. Diese Struktur spiegelt das *Link*-Schema wider, indem das Verb die grammatischen Beziehungen von Subjekt und Objekt zuweist und die beiden miteinander verbindet. Darüber hinaus bildet dieser aktive transitive Hauptsatz die einfache Handlungskette *jīxièshī* (Mechaniker) → *qìchē* (Auto) ab, wobei die gesamte Handlungskette profiliert wird.

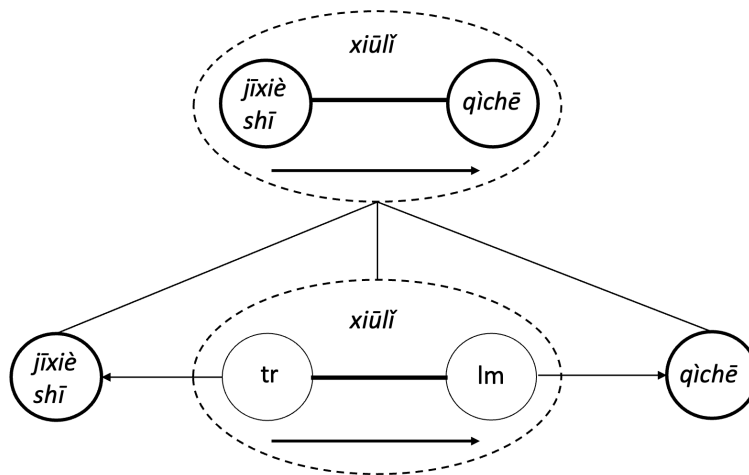


Abbildung 4.2: Abstrakte Darstellung des chinesischen Hauptsatzes „jīxièshī xiūlǐ qìchē“ (der Mechaniker repariert das Auto). Die Linie zwischen dem Trajektor und der Landmark repräsentiert die Beziehung zwischen den Teilnehmern. Die längeren Pfeile in den Ellipsen symbolisieren die zeitliche Dimension des Prozesses. Diese sind fett hervorgehoben, da das Verb und der gesamte Hauptsatz den Prozess über die Zeit profilieren. Die gebildeten abstrakten Regionen sind beim Hauptsatz und Verb latent. Sie sind daher durch gestrichelte Ellipsen gekennzeichnet.

Langacker (1987, 1991) gemäß weisen chinesische Hauptsätze theoretisch ein prozessuales Profil auf und werden durch sequentielles Scanning mental verarbeitet, wobei die verschiedenen Zustände der beteiligten Objekte in zeitlicher Abfolge erfasst werden.

4.1.2. Deutscher Hauptsatz

Morphologische und syntaktische Struktur

Im Deutschen, einer flektierenden Sprache, werden syntaktische Beziehungen und grammatische Funktionen durch Flexionsendungen markiert (vgl. Greenberg, 1963). Das Verb spielt im deutschen Hauptsatz eine entscheidende Rolle. Es weist grammatische Beziehungen und thematische Rollen aller Satzglieder zu und vermittelt durch die Flexion detaillierte grammatische Informationen, die im Chinesischen fehlen. Es zeigt durch verschiedene Flexionsformen das Tempus (z.B. Präsens, Präteritum etc.), den Modus (Indikativ, Konjunktiv, Imperativ) und das Genus Verbi (Aktiv, Passiv) an und gibt durch die Konjugation die Person (1., 2., 3. Person) sowie den Numerus (Singular, Plural) an (vgl. Eisenberg, 2020).

Auch das Deutsche folgt der SVO-Wortstellung, wobei das finite Verb typischerweise an zweiter Stelle steht (vgl. Greenberg, 1963). Die Grundstruktur ist damit formal identisch mit der des chinesischen Hauptsatzes.

[Subjekt + Verb + Objekt]

Bsp. 4-2 Der Mechaniker repariert das Auto.

Die Verbflexion, das Kasus- und Genus-System im Deutschen ermöglichen es, die Stellung von Subjekt und Objekt in gewisser Weise flexibel zu verändern, ohne dass die Satzbedeutung beeinträchtigt wird. Während das Subjekt im Chinesischen oft weggelassen werden kann (vgl. Halliday & McDonald, 2004; Li & Thompson, 1981), ist es im deutschen Hauptsatz in der Regel obligatorisch.

Beispiel 4-2 zeigt einen deutschen transitiven Hauptsatz, in dem das Agens („*der Mechaniker*“) als Subjekt und das Patiens („*das Auto*“) als Objekt realisiert wird. Das Verb *reparieren* kongruiert durch Flexion und Konjugation mit dem Subjekt in Person und Numerus und verankert den beschriebenen Prozess zeitlich, während die Nomina der Kongruenz durch ihre Deklination folgen.

Kognitive Verarbeitung

Hinsichtlich der in Kapitel 3 beschriebenen kognitiven Mechanismen und Verarbeitungsprozesse werden deutsche und chinesische Hauptsätze in ähnlicher Weise verarbeitet. Wie im Chinesischen steht auch hier die gesamte enkodierte Handlungskette im Vordergrund. Deutsche Hauptsätze weisen ebenfalls das prozessuale Profil auf, und die Konzeptualisierung erfolgt über das sequentielle Scanning, bei dem die verschiedenen Zustandsrepräsentationen in zeitlicher Abfolge mental aktiviert werden.

Abbildung 4.3 veranschaulicht die abstrakte Darstellung des deutschen Hauptsatzes anhand des Beispiels 4-2. Diese Darstellung ist analog zur Struktur des chinesischen Hauptsatzes.

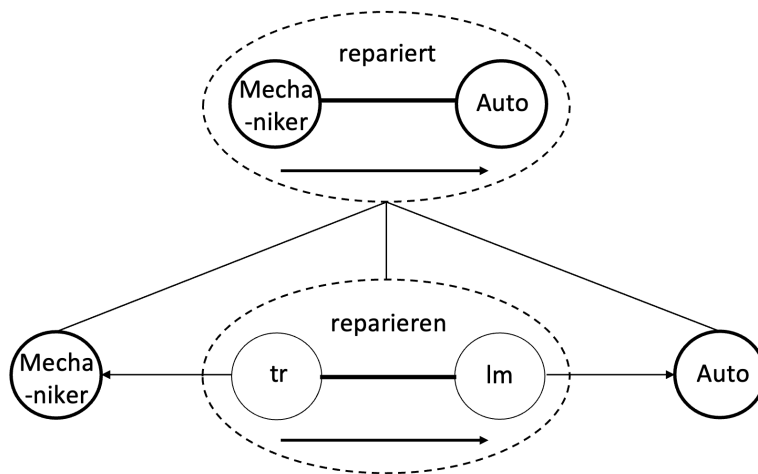


Abbildung 4.3: Abstrakte Darstellung des deutschen Hauptsatzes „Der Mechaniker repariert das Auto“.

4.2. Subjekt-Relativsatz

Relativsätze werden in der Regel nicht als eigenständige Sätze, sondern als untergeordnete Nebensätze betrachtet, die vom Matrixsatz abhängig sind und sich syntaktisch auf einen nominalen Ausdruck im Matrixsatz beziehen (vgl. Lehmann, 1986). Dieser nominale Ausdruck, auch Kopfnominal genannt, verknüpft den Matrixsatz und den Relativsatz sowohl syntaktisch als auch semantisch miteinander.

Semantisch gilt das Relativsyntagma meist als Attribut des Kopfnomens. Es kann zusätzliche Informationen in das Kopfnomen integrieren, um entweder die durch das Kopfnomen bezeichneten Entitäten einzuschränken und dadurch präziser zu identifizieren oder nur ergänzende Informationen zu liefern, ohne die Bedeutung des Kopfnomens wesentlich zu verändern (vgl. Holler, 2013; Lehmann, 1986).

Ein Subjekt-Relativsatz kann aus einem einfachen aktiven transitiven Hauptsatz umgewandelt werden. Wird das Subjekt des Hauptsatzes relativiert und der übrige Satzteil zu Modifikatoren des Kopfnomens degradiert, kann diese strukturelle Veränderung dazu führen, dass sich die Perspektive auf das Ereignis ändert und das ursprüngliche mentale Image des Ereignisses durch ein neues ersetzt wird.

Im Folgenden werden die Strukturen der Subjekt-Relativsätze im Chinesischen und Deutschen vorgestellt, einschließlich ihrer Bildungsstrategien sowie der damit

verbundenen Informationsstruktur.

4.2.1. Chinesischer Subjekt-Relativsatz

Syntaktische Struktur und Bildungsstrategie

Aus typologischer Perspektive besteht ein wesentlicher Unterschied in der Wortstellung von Relativsätzen im Chinesischen und Deutschen. Laut den Korpusanalysen von Dryer (1992, 2003) ist Chinesisch unter den 625 (1992) bzw. 910 (2003) Sprachen eine der wenigen SVO-Sprachen, bei denen das Kopfnomen in Relativsätzen immer rechtsperipher angesiedelt wird, während das Relativsyntagma vorangestellt wird. Das Morphem *DE* funktioniert dabei als Marker und steht immer am Ende des Relativsyntagmas. Die abstrakte Struktur lautet:

[Relativsyntagma + Kopfnomen]

Bsp. 4-3 修理 汽车 的 机械师
 xiūlǐ qìchē de jīxièshī
 reparieren Auto de Mechaniker
 „der Mechaniker, der das Auto repariert.“

Basierend auf sprachübergreifenden Untersuchungen fasst Comrie (1989) vier grundlegende Relativierungsstrategien zusammen, nämlich *non-reduction*, *pronoun-retention*, *relative pronoun* und *gap*. Von diesen wird die *Gap*-Strategie im Chinesischen am häufigsten verwendet, insbesondere bei der Relativierung von Subjekten und direkten Objekten. Die *Pronoun-Retention*-Strategie ist zwar zulässig, wird aber nur bei niedrigeren syntaktischen Funktionen, wie dem indirekten Objekt, verwendet (vgl. Lehmann, 1986; Liu, 2005; Wang, 2011; Wang & Feng, 2012).

Kognitive Verarbeitung

Bei der kognitiven Verarbeitung von Ereignissen werden verschiedene Informationen unterschiedlich gewichtet, sodass bestimmte Objekte oder Teile eines Ereignisses

hervorgehoben werden und mehr Aufmerksamkeit auf sich ziehen, während andere eher in den Hintergrund treten. Die Figur-Grund-Beziehung (*figure-ground alignment*), die von Talmy (1978, 2000) in der kognitiven Linguistik zur Erklärung vieler sprachlicher Phänomene eingeführt wurde und zur Dimension „Prominenz“ des *Construals* gehört, spiegelt diesen Unterschied in der kognitiven Salienz und Aufmerksamkeit wider (vgl. Langacker, 1993; Biên, 2021).

Die Figur-Grund-Beziehung dient der kognitiven Organisation von Informationen und hilft dabei, Vordergrund und Hintergrund zu unterscheiden. Sie beeinflusst die Verteilung der Aufmerksamkeit, indem sie festlegt, welche Elemente eines Ereignisses oder eines kognitiven Kontexts als zentral konstruiert werden.

Die Figur, die auch als „Vordergrund“ bezeichnet wird, ist das zentrale oder dominante Element, auf das sich die Aufmerksamkeit des Betrachters richtet. Sie ist tendenziell auffälliger, beweglicher, besser abgegrenzt und kleiner als der Grund. Der Grund hingegen umfasst die anderen Elemente, die weniger prominent sind und den Kontext bilden. Er erfordert weniger Aufmerksamkeit (vgl. Langacker, 1987, 1991, 2013; Talmy, 2000; Ungerer & Schmid, 1996).

Die grammatische Struktur chinesischer Subjekt-Relativsätze lässt sich in Anlehnung an die Figur-Grund-Unterscheidung analysieren. In der hier vertretenen Analyse fungiert das Kopfnomen typischerweise als semantische und konzeptuelle „Figur“, während das Relativsyntagma den Hintergrund bereitstellt, der zusätzliche Informationen zur näheren Spezifizierung des Kopfnomens liefert (vgl. Langacker, 2013; Yu, 2005).

Da chinesische Relativsätze pränominal sind, erscheint das Kopfnomen linear am Ende der Konstruktion. Die Informationsabfolge verläuft daher vom Grund zur Figur:

Grund (Relativsyntagma) → Figur (Kopfnomen)

Auf diese Weise erhält der Hörer bzw. Leser zunächst kontextuelle bzw. ergänzende

Informationen, die einen Rahmen für das nachfolgende Kopfnomen schaffen. In Beispiel 4-3 wird durch das Relativsyntagma „*xiūlǐ qìchē de*“ (der das Auto repariert) zunächst eine semantische Restriktion eingeführt, bevor das Kopfnomen „*jīxièshī*“ (Mechaniker) erscheint.

Diese Anordnung zwingt den Hörer bzw. Leser, bis zum Ende des Satzes zu warten, um das Kopfnomen zu erhalten, das durch den vorangestellten Relativsatz bereits präzisiert wird. Es kann angenommen werden, dass diese bevorstehenden Informationen im Arbeitsgedächtnis aktiv gehalten und zur Antizipation des Kopfnomens genutzt werden, indem seine semantischen Merkmale vorab verarbeitet werden. Dadurch wird der mögliche Referenzbereich des Kopfnomens bereits vor dessen Auftreten eingeschränkt und eine Erwartungshaltung entwickelt. Daneben ist weniger Rückwärtsverarbeitung erforderlich, um die Beziehung zwischen Modifikator und Nomen zu entschlüsseln. In diesem Sinne kann die Integration von Modifikator und Kopfnomen erleichtert werden (vgl. Gibson, 1998; Karimi, Diaz & Ferreira, 2019).

4.2.2. Deutscher Subjekt-Relativsatz

Syntaktische Struktur und Bildungsstrategie

Je nach Funktion und Struktur lassen sich deutsche Relativsätze in verschiedene Typen und Varianten einteilen. Sie können restriktiv oder nicht-restriktiv, mit oder ohne Relativpronomen sowie wortbezogen oder satzbezogen sein. Aufgrund dieser Vielfalt ist es schwierig, ein einheitliches Muster für alle Relativsätze zu definieren. Dennoch lassen sich gemeinsame Charakter für prototypische Relativsätze im Deutschen benennen.

Holler (2013) identifiziert folgende Hauptmerkmale für prototypische Relativsätze im Deutschen: Sie sind grammatisch abhängig von einem übergeordneten Matrixsatz und können nicht alleine stehen; das Verb steht am Satzende; sie erfüllen eine attributive Funktion, indem sie ein Nomen im Matrixsatz näher beschreiben oder zusätzliche Informationen bereitstellen; sie werden durch Relativpronomina wie „der, die,

das“ oder durch Relativadverbien wie „wo, wann“ eingeleitet.

Deutsche Subjekt-Relativsätze unterscheiden sich in vielerlei Hinsicht von chinesischen. Der Hauptunterschied in der Oberflächenstruktur besteht darin, dass das Kopfnomen in deutschen Relativsätzen vor dem subordinierenden Nebensatz steht, während es im Chinesischen am Ende des Relativsatzes erscheint. Die abstrakte Struktur eines prototypischen deutschen Subjekt-Relativsatzes lässt sich wie folgt darstellen:

[Kopfnomen + Relativsyntagma]

Bsp. 4-4 Der Mechaniker, der das Auto repariert.

Kopfnomen Relativpron. NP Verb

Relativpronomina, die im Chinesischen nicht vorkommen, spielen im Deutschen eine wichtige Rolle, da sie das Kopfnomen und den subordinierenden Nebensatz syntaktisch verbinden. Genus und Numerus des Relativpronomens werden durch das Kopfnomen bestimmt, während der Kasus durch dessen syntaktische Funktion im Relativsatz festgelegt wird. In Beispiel 4-4 bezieht sich das Relativpronomen „*der*“ auf das Kopfnomen „*der Mechaniker*“. „*Das Auto*“ fungiert im Relativsatz als direktes Objekt und wird durch das Verb „*reparieren*“ regiert.

Das Relativpronomen kann bei der kognitiven Verarbeitung als *Cue* dienen, das die Verbindung zwischen Kopfnomen und Relativsatz aktiviert (vgl. Hemforth et al., 2000; Karimi et al., 2019). Die Aktivierung des Relativpronomens kann möglicherweise zur Reaktivierung des Kopfnomens führen, wodurch dessen Salienz erhöht werden kann.

Kognitive Verarbeitung

Die in Abschnitt 4.2.1 beschriebene Figur-Grund-Beziehung findet sich ebenfalls in deutschen Relativsätzen. Auch hier fungiert das Kopfnomen als zentrales Bezugsobjekt, während das Relativsyntagma zusätzliche, eher hintergründige Informationen bereitstellt, die diesen Referenten näher spezifizieren. Der Hauptunterschied zur

chinesischen Struktur liegt in der Reihenfolge dieser Elemente. Während chinesische Relativsätze pränominal sind und den Hintergrund zuerst präsentieren, sind deutsche Relativsätze postnominal organisiert. Die Information verläuft daher von der Figur zum Grund:

Figur (Kopfnomen) → Grund (Relativsyntagma)

Bei dieser Struktur wird die Kerninformation unmittelbar bereitgestellt, da das Kopfnomen am Anfang des Relativsatzes erscheint. Hörer bzw. Leser erkennen ohne Verzögerung, welcher Referent im Mittelpunkt steht. Der Relativsatz liefert im Anschluss zusätzliche Informationen, die den Referenten genauer beschreiben oder den Bereich der möglichen Referenten einschränken. Diese Abfolge ermöglicht eine schnelle Identifikation und Verarbeitung der Kerninformation, gefolgt von der Integration der ergänzenden Beschreibung. In Beispiel 4-4 wird das Kopfnomen „*der Mechaniker*“ am Anfang des Relativsatzes eingeführt, während das Relativsyntagma „*der das Auto repariert*“ den Mechaniker näher spezifiziert und als kontextualisierende Hintergrundinformation dient.

4.3. Nominalisierung

Nominalisierung wird im weiteren Sinne als ein grammatischer Prozess definiert, bei dem ein nicht-nominaler Ausdruck, z.B. ein Verb, ein Adjektiv, ein Satz oder ein Satzteil, in einen nominalen Ausdruck umgewandelt wird (vgl. Comrie & Thompson, 1985; Li & Thompson, 1981). Da die deverbale Nominalisierung einen Transformationsprozess beinhaltet, weist sie gemischte Eigenschaften von Nomen und der ursprünglichen verbalen oder prädikativen Struktur auf (vgl. Alexiadou, 2010; Roy & Soare, 2011).

Das vorliegende Kapitel konzentriert sich auf Nominalisierungen wie „*die Autoreparatur*“, die auf Ereignisse Bezug nehmen. Grimshaw (1990) argumentiert, dass bestimmte Typen von deverbalen Nominalisierungen als *complex event nominals* aufgefasst werden. Sie legen eine ereignishafte Interpretation nahe und weisen eine

interne Ereignisstruktur auf, die der Argumentstruktur eines Verbs ähnelt. Beispiele hierfür sind Rektionskomposita (auch *synthetische Komposita* genannt) im Deutschen sowie Gerundien im Englischen (vgl. Alexiadou, 2010, 2017; Grimshaw, 1990). Die auf Gerundien basierende englische Nominalisierung [V-ing + of-Phrase], z.B. *signing of the contract* (vgl. Langacker, 1991), weist Ähnlichkeiten mit der chinesischen Nominalisierung [NP + DE + VP] auf (vgl. Lu & Pan, 2013; Jiang & Dai, 2004). Dies ermöglicht eine Analyse der chinesischen Nominalisierung in Anlehnung an Langacker (1991) über [V-ing + of-Phrase].

Im Folgenden werden zunächst die chinesische Nominalisierung [NP + DE + VP] und anschließend die deutschen Rektionskomposita beschrieben. Dabei wird jeweils die Struktur erläutert und die kognitive Verarbeitung dieser Nominalisierungen analysiert.

4.3.1. Chinesische Nominalisierung

In Bezug auf die Ereigniskodierung ist im Chinesischen die komplexe Nominalisierung [NP + DE + VP] von besonderer Bedeutung. Diese Nominalisierung kann aus einem Hauptsatz oder Teilsatz umgewandelt werden (vgl. He & Wang, 2007). Dabei kommt ein Ereignis, das ursprünglich in einem Satz kodiert wurde, nun in einem nominalen Ausdruck zum Vorschein. In diesem Fall wird der Marker *DE* zwischen dem Subjekt bzw. Objekt und dem Verb eingefügt. Die abstrakte Grundstruktur der Konstruktion lautet:

$$[\text{NP} + \text{DE} + \text{VP}]^1$$

Bei dieser nominalen Konstruktion wird die VP durch die NP modifiziert und nimmt somit die Kernposition in der Nominalisierung ein (vgl. Chen, 1996; Wang, 2010; Zhang, 1993). Wie in Beispiel 4-5 zu sehen ist, wird der unterstrichene Satz in a) zu

¹ Die gesamte Struktur [NP + DE + VP] ist nominal. In der Fachliteratur besteht jedoch bislang keine einheitliche Auffassung darüber, ob das ursprüngliche Verb innerhalb dieser Struktur selbst als nominalisiert gelten kann. Das Verb übernimmt in dieser Struktur viele nominale Eigenschaften, behält aber auch einige verbale Eigenschaften (vgl. Shen & Wang, 2000; Si, 2004; Zhu, Lu & Ma, 1961). Da diese Frage nicht im Fokus der vorliegenden Arbeit steht, wird sie hier nicht weiter vertieft. Im Einklang mit der gängigen Praxis in der einschlägigen Forschung wird in dieser Struktur „VP“ verwendet, um auf das ursprüngliche Verb zu verweisen.

den in Klammern stehenden Nominalisierungen in b) bzw. c) umgewandelt.

Bsp. 4-5 a) 他 放弃 了 机会, 我 很 惋惜。

tā fàngqì le jīhuì, wǒ hěn wànxī

er verzichten PEF Chance, ich sehr bedauern

„Ich bedauere sehr, dass er auf die Chance verzichtet hat.“

b) 我 很 惋惜 [他 的 放弃]

wǒ hěn wànxī [tā de fàngqì]

ich sehr bedauern [er *de* verzichten]

„Ich bedauere seinen Verzicht sehr.“

c) 我 很 惋惜 [机会 的 放弃]

wǒ hěn wànxī [jīhuì de fàngqì]

Ich sehr bedauern [Chance *de* verzichten]

„Ich bedauere den Verzicht auf die Chance sehr.“

Das Basisverb *fàngqì* (verzichten) weist im Hauptsatz und in den Nominalisierungen denselben Argumenten identische thematische Rollen zu. Das heißt, *tā* (er) gilt in 4-5 a) und b) semantisch als Agens des enkodierten Ereignisses, und *jīhuì* (Chance) in 4-5 a) und c) als Patiens. Unterschiede ergeben sich jedoch in ihrer syntaktischen Funktion. Aus grammatischer Sicht fungiert *tā* (er) nur im Hauptsatz als Subjekt und *jīhuì* (Chance) als Objekt. In der Nominalisierung können Argumente hingegen nicht direkt als Subjekt oder Objekt realisiert werden, sondern nur indirekt über den Marker *DE* eingebunden werden.

Aus den Beispielen wird ersichtlich, dass nach der Nominalisierung grundsätzlich nur ein explizites Argument in der Nominalkonstruktion verbleiben kann (vgl. Li & Thompson, 1981). Wenn das Basisverb im ursprünglichen Prädikat ein einwertiges Verb ist, bleibt das einzige Argument erhalten. Bei zwei- oder mehrwertigen Verben hängt die Wahl der NP vom Verb ab.

Viele Linguisten haben die Regeln und Einschränkungen dieser Struktur in Bezug auf die NP, die VP sowie ihre semantischen Rollen untersucht (z.B. Chen, 1996, 2010; Wang, 2010; Zhan, 1998a, 1998b; Zhang, 1993). Nach einer Korpusanalyse von Zhan (1998a) können die meisten Verben in dieser Struktur nur das Agens als NP wählen. Nicht selten gibt es aber auch viele Verben, die sowohl das Agens (wie in 4-5 b) als auch das Patiens (wie in 4-5 c) als NP zulassen. Daneben kann die NP auch andere thematische Rollen kodieren, wie z.B. Instrument, Lokation und Zeit.

Die Reduktion eines Arguments kann dazu führen, dass *Specificity* und die Individualität des Ausdrucks nach der Nominalisierung abnehmen. So drückt beispielsweise der unterstrichene Satz in 4-5 a) deutlich aus, wer auf was verzichtet hat. Die Nominalisierung in 4-5 b) gibt hingegen nicht explizit an, worauf verzichtet wird, und in 4-5 c) bleibt unklar, wer auf die Chance verzichtet hat. Es könnte sich um „ich“ oder eine andere Person handeln.

Zudem wird in Beispiel 4-5 a) der Aspektmarker *le* dem Verb hinzugefügt, wodurch das Ereignis in einer vergangenen Zeit verankert wird. Im Gegensatz dazu sind Aspektmarker wie *le* (perfektiver Marker), *zhe* (progressiver Marker) und *guò* (perfektiver Marker) in den Nominalisierungen nicht zulässig (vgl. Chen, 1996; Zhang, 1993), sodass das auftauchende Ereignis zeitlich nicht spezifiziert ist und ohne weiteren Kontext unterschiedliche zeitliche Lesarten zulässt.

Daneben dürfen in dieser Struktur weder Modaladverbien noch die meisten Temporaladverbien, Hilfsverben und die Negation *méiyǒu* vor der VP stehen (vgl. Chen, 1996). Diese eingeschränkten Ausdrucksmöglichkeiten führen im Vergleich zur Satzform zu einem Verlust an spezifischer Information.

Kognitive Verarbeitung

Die chinesische Nominalisierung [NP + *DE* + VP] ähnelt der englischen Nominalisierung [V-*ing* + *of*-Phrase] (vgl. Lu & Pan, 2013; Jiang & Dai, 2004). Auf Grundlage von Langackers (1991) Analyse der Struktur [V-*ing* + *of*-Phrase] lässt sich

auch die chinesische Nominalisierung analysieren. Während das Verb im Prädikat direkt mit seinem Objekt verbunden sein kann, ist dies in der Nominalisierung nicht mehr möglich. Stattdessen wird ein ursprüngliches Argument in einer anderen Form realisiert – im Englischen durch die *of*-Phrase und im Chinesischen durch [NP + *DE*]. Dadurch wird ein Teilnehmer des zugrunde liegenden Ereignisses konkretisiert (vgl. Lu & Pan, 2013).

Die Morpheme „*DE*“ im Chinesischen und „*of*“ im Englischen stellen typischerweise eine intrinsische Beziehung zwischen Komponenten dar (vgl. Langacker, 1991), z.B. „*túhuà de yánsè*“ (*the color of the picture*). Diese Konstruktion ist strukturell ähnlich wie die Nominalisierung „*hétóng de qiānshǔ*“ (*signing of the contract*). Der wesentliche Unterschied zwischen diesen Beispielen besteht darin, dass „*qiānshǔ*“ (*signing*) einen verdinglichten Prozess bezeichnet, während „*yánsè*“ (*color*) ein statisches Konzept ist. Das Nomen *hétóng* (*contract*) spezifiziert einen Teilnehmer bzw. ein Argument des ursprünglichen Prozesses. In der Nominalisierungsstruktur profiliert „*DE*“ daher eine statische Beziehung zwischen einem verdinglichten Prozess und einem Teilnehmer des Prozesses.

Dieser verdinglichte Prozess wird auf einer übergeordneten Ebene durch die VP näher spezifiziert. Das ursprüngliche Verb verliert in dieser Struktur seinen prozessualen Charakter. Die im Verb enthaltenen Beziehungen werden nun gemeinsam als abstrakte Einheit konzeptualisiert – als ein „Ding“. Dieser spezifische verdinglichte Prozess entspricht dem durch *DE* assoziierten schematischen verdinglichten Prozess und kann als Trajektor dieser statischen Beziehung betrachtet werden. Der durch [NP + *DE*] spezifizierte Ereignisteilnehmer entspricht einem der vom Verb implizierten Teilnehmer, entweder dem Trajektor oder der Landmark.

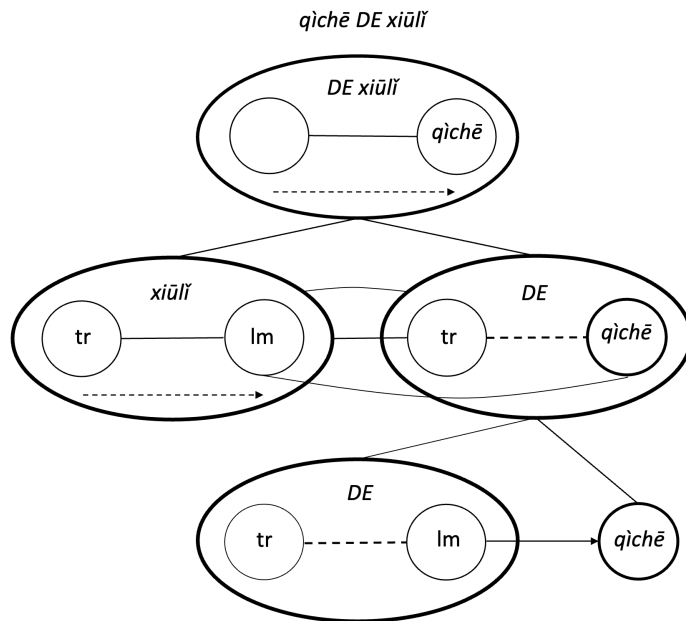


Abbildung 4.4: Abstrakte Darstellung der Nominalisierung „*qìchē de xiūlǐ*“ (die Autoreparatur). Der Marker *DE* bezeichnet kein Prozess, sondern eine statische Beziehung, was durch die gestrichelte Linie zwischen seinem Trajektor und der Landmark gekennzeichnet wird. Der Pfeil bei *xiūlǐ* repräsentiert den zeitlichen Verlauf und ist gestrichelt dargestellt. Das bedeutet, dass das Basisverb gewissermaßen nominalisiert wird und die prozessuale Eigenschaft in den Hintergrund tritt.

Die abstrakte Darstellung der Nominalisierung [NP + *DE* + VP] ist in Abbildung 4.4 am Beispiel „*qìchē de xiūlǐ*“ (die Autoreparatur) zu sehen. Auf der ersten Ebene profiliert der Marker *DE* eine statische Beziehung, deren Landmark durch „*qìchē*“ (Auto) spezifiziert wird. Diese bilden zusammen die Phrase „*qìchē de*“ (Auto *de*). Der Trajektor dieser Beziehung, der ein verdinglichter Prozess ist, bleibt noch schematisch. Auf der zweiten Ebene wird dieser verdinglichte Prozess durch „*xiūlǐ*“ (reparieren) konkretisiert. Da das Basisverb gewissermaßen nominalisiert wird, kann es kein Argument mehr direkt einleiten. Stattdessen wird das Argument in der Nominalisierung durch die [NP + *DE*]-Phrase realisiert, wie „*qìchē de*“ (Auto *de*). Dieses Argument entspricht der Landmark des zugrunde liegenden Prozesses. Zusammen bilden „*qìchē de*“ (Auto *de*) und „*xiūlǐ*“ (reparieren) die vollständige Nominalisierung „*qìchē de xiūlǐ*“ (die Autoreparatur), wobei das Ereignis als Gesamtheit in den Vordergrund rückt.

Nach Talmy (2000) gilt in der sprachlichen Kommunikation, dass das, was explizit genannt wird, salenter ist und kognitiv leichter verarbeitet wird als implizite

Informationen. In der Nominalisierung [NP + DE + VP] wird in der Regel nur ein Argument explizit genannt, wodurch es als salientes Element konstruiert und mit geringem kognitivem Aufwand verarbeitet werden kann. Die impliziten Argumente müssen hingegen aus dem Kontext erschlossen werden, was mehr kognitive Ressourcen erfordert.

Im Chinesischen ist diese Nominalisierung insbesondere in der Schriftsprache verbreitet (vgl. Lu & Pan, 2013; Zhan, 1998b). Sie dient typischerweise dazu, auf vorherige Informationen zurückzugreifen, anstatt neue Informationen einzuführen (vgl. Zhang, 1993). Eine Korpusanalyse von Zhan (1998b) zeigt, dass diese Struktur vorrangig verwendet wird, um auf bereits erwähnte Ereignisse Bezug zu nehmen. Dadurch kann ein zuvor ausführlich beschriebenes Ereignis prägnant und komprimiert in einer kurzen Nominalphrase zusammengefasst werden. Die enkodierte Information wird allerdings gewissermaßen reduziert, was zu einer geringen Spezifität dieser Ausdrucksweise führt.

4.3.2. Deutsche Nominalisierung

Bei den deutschen Nominalisierungen, die auf Ereignisse Bezug nehmen, handelt es sich um deverbale *Rektionskomposita*. In diesen Komposita übernimmt das Erstglied eine syntaktische und semantische Funktion in Bezug auf das Zweitglied. Das Zweitglied ist eine deverbale Nominalisierung und weist die Argumentstruktur auf (vgl. Aptacy, 2015; Olsen, 1986). Die Grundstruktur eines solchen Rektionskompositums lautet:

[N (Erstglied) + N (Zweitglied)]

Bsp. 4-6 Die *Autoreparatur* dauert einen Monat.

In dieser Struktur fungiert das Zweitglied (*Reparatur* in Beispiel 4-6) als Kopf, indem es als „das *Determinatum strukturelle Anknüpfungspunkte für das determinierende Element vorgibt oder anbietet*“ (Eichinger, 2000, S. 128). Es bestimmt die grammatischen Eigenschaften hinsichtlich Wortart und Genus des Kompositums (vgl.

Gaeta, 2010). Außerdem erbt diese Nominalisierung die Argumentstruktur des zugrunde liegenden Verbs und determiniert dadurch, welche Argumentrolle dem Erstglied zugewiesen wird (vgl. Bierwisch, 1990).

Als das Zweitglied des Kompositums können die Nominalisierungen durch verschiedene Methoden gebildet werden, insbesondere durch Derivation und Konversion. Das Deutsche verwendet als flektierende Sprache zumeist morphologische Mittel, insbesondere Suffixe, um Nominalisierungen abzuleiten. Häufig genutzte Suffixe sind wie *-er*, *-ung*, *-ion* und *-ur*, z.B. *reparieren* – *die Reparatur*. Daneben gibt es Nullableitungen, bei denen Verben ohne Affix direkt auf Basis ihres Verbstamms nominalisiert werden können, z.B. *bauen* – *der Bau*. Eine weitere Form ist die Infinitivnominalisierung, die im Prinzip für alle Verben möglich ist, z.B. *lesen* – *das Lesen* (vgl. Scheffler, 2005).

Das Erstglied (*Auto* in Beispiel 4-6) eines Rektionskompositums füllt typischerweise eine Argumentposition in der Argumentstruktur des Zweitglieds aus (vgl. Aptacy, 2015; Haspelmath & Sims 2013; Olsen, 1986). Eichinger (2000) erweitert das Konzept der Rektionskomposita dahingehend, dass das Erstglied neben Argumenten auch adverbiale Angaberelationen einbeziehen kann. Demnach kann das Erstglied vielfältige syntaktische Funktionen in Bezug auf das Zweitglied einnehmen, die normalerweise von Satzgliedern übernommen werden, z.B. Subjekt-, Objekt-, modale und temporale Relationen (vgl. Eichinger, 2000). Das am häufigsten anzutreffende Muster besteht darin, dass das Erstglied das direkte Objekt des zugrundeliegenden Verbs repräsentiert (vgl. Aptacy, 2015; Eichinger, 2000; Gaeta & Zeldes, 2012).

Rektionskomposita stellen eine nützliche und effiziente Strategie dar, um die Argumentstruktur eines Ereignisses in der Nominalisierung zu kodieren (vgl. Aptacy, 2015; Toman, 1983). Dabei weist das Zweitglied dem Erstglied die thematische Rolle zu, während das Erstglied einen Rahmen für die im Zweitglied dargestellte Handlung bereitstellt, wodurch der Bedeutungsumfang der Nominalisierung weiter präzisiert wird.

Deutsche Nominalisierungen können direkt mit ihren Argumenten zu Komposita verbunden werden. Dadurch unterscheidet sich die abstrakte Darstellung der deutschen Rektionskomposita deutlich von derjenigen chinesischer Nominalisierungen. Wie Abbildung 4.5 zeigt, wird das Verb *reparieren* zum Nomen *Reparatur* nominalisiert. Das direkte Objekt *Auto* entspricht dabei der Landmark des Basisverbs. Zusammen bilden beide Komponenten das Rektionskompositum „die Autoreparatur“.

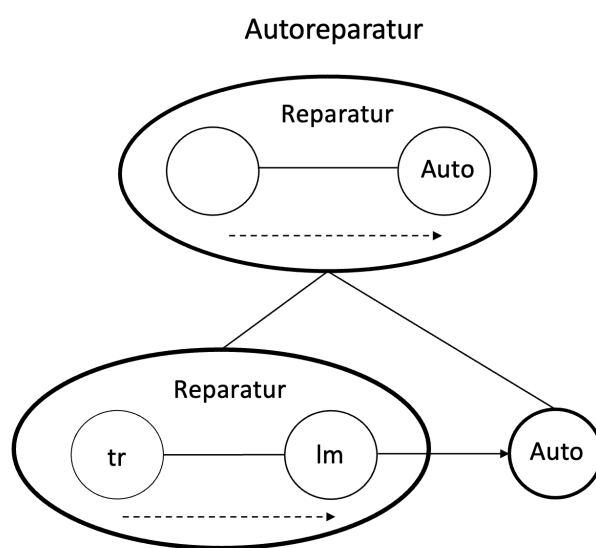


Abbildung 4.5: Abstrakte Darstellung des Rektionskompositums „die Autoreparatur“.

Trotz struktureller Unterschiede weisen deutsche und chinesische Nominalisierungen gewisse Gemeinsamkeiten auf. So weisen auch die deutschen Rektionskomposita ein nominales Profil auf, bei dem infolge der Nominalisierung die zeitliche Dimension der zugrunde liegenden verbalen Prädikation an Bedeutung verliert. Die durch das Kompositum ausgedrückte Sachverhaltsstruktur wird nun als eine holistische Gestalt erfasst. Dies zeigt sich auch grammatisch daran, dass innerhalb dieser Struktur keine Tempusmarkierung möglich ist. Eine eindeutige zeitliche Verankerung des zugrunde liegenden Geschehens ist daher ohne zusätzlichen Kontext nicht möglich.

Darüber hinaus bleibt auch bei deutschen Rektionskomposita nur ein Argument des Basisverbs explizit erhalten, wodurch bestimmte Informationen des ursprünglichen

Prädikats verloren gehen (vgl. Bzdęga, 1981). Dies zeigt sich beispielsweise im Rektionskompositum „*Autoreparatur*“, in dem weder das Agens noch der Zeitpunkt des Geschehens explizit ausgedrückt werden. Daher reduziert auch diese Struktur die Präzision und den Detailgrad der Ereignisdarstellung im Vergleich zu einem entsprechenden finiten Satz.

5. Fragestellung und Forschungsdesign

5.1. Fragestellung

Zur Strukturierung und Interpretation der Außenwelt werden sensorische Informationen vom kognitiven System als diskrete Ereignisse mental repräsentiert und verarbeitet (vgl. Ünal, Ji & Papafragou, 2021; Zacks & Tversky, 2001). Ereignisse werden in diesem Zusammenhang als mentale Repräsentationen betrachtet, mit denen nicht nur passiv auf Inputs reagiert wird, sondern die aktiv generiert, fortlaufend evaluiert und für das Erzeugen von Vorhersagen verwendet werden. Für Ereignisrepräsentationen sind eine Vielzahl von Aspekten von Bedeutung, wie z.B. die Ereignisteilnehmer und ihre Beziehungen, der räumlich-zeitliche Rahmen, die Grenze zwischen Ereignissen und der zeitliche Verlauf des Geschehens einschließlich aller Zustandsveränderungen der Objektrepräsentationen (vgl. Kapitel 2).

Ereignisrepräsentationen können sowohl direkt durch sensorische Wahrnehmung entstehen als auch indirekt durch die Verarbeitung von Sprache vermittelt werden. Die Art und Weise, wie Sprache Ereignisse strukturiert, spiegelt sich in der erzeugten Ereignisrepräsentation wider. Diese Ereignisrepräsentation kann einen Einfluss auf die Konzeptualisierung des enkodierten Ereignisses ausüben. Zahlreiche empirische Studien zeigen, dass Sprache die Wahrnehmung und Aufmerksamkeitslenkung ihrer Sprecher entweder kurzfristig während der Sprachverwendung oder langfristig im Rahmen nicht-sprachlicher kognitiver Prozesse prägen kann (vgl. Gerwien & von Stutterheim, 2018; Levinson, 2003; Papafragou et al., 2008; Slobin, 1996; Whorf, 1956).

Menschen verfügen über die Fähigkeit, auf ein und dasselbe Ereignis auf unterschiedliche Weise sprachlich Bezug zu nehmen. Diese Fähigkeit erfüllt vielfältige Kommunikationszwecke. Demnach spiegelt die Art und Weise, wie wir eine Szene sprachlich kodieren, unsere Konzeptualisierung dieser Szene wider. Unterschiedliche sprachliche Strukturen, die auf verschiedenen kognitiven Modellen und Verarbeitungsprozessen beruhen, rücken unterschiedliche Aspekte eines Ereignisses in den Mittelpunkt und beeinflussen dadurch die relative Salienz der Ereigniskomponenten, zum Beispiel, thematische Rollen.

Vor diesem Hintergrund ergeben sich zwei zentrale Fragestellungen für die vorliegende Arbeit:

- 1) Beeinflussen die drei sprachlichen Strukturen – Hauptsatz, Subjekt-Relativsatz und deverbale Nominalisierung – die Ereigniskonzeptualisierung, indem sie jeweils unterschiedliche Aspekte der Ereignisrepräsentation hervorheben?

Wenn die drei sprachlichen Strukturen auf dasselbe Ereignis Bezug nehmen, sind die jeweiligen thematischen Rollen der Ereignisteilnehmer in allen drei Strukturen identisch. Allerdings haben die verschiedenen Oberflächenstrukturen gemäß den theoretischen Konzepten der kognitiven Linguistik zur Folge, dass bestimmte Komponenten in den Vordergrund treten, während andere im Hintergrund bleiben, was in dieser Arbeit als *Construal* bezeichnet wird (vgl. Kapitel 3). In diesem Zusammenhang wird im empirischen Teil dieser Arbeit untersucht, ob verschiedene *Construals* auf bestimmte Weise mit der visuellen Wahrnehmung interagieren. Diese Frage wird in den Experimenten 1 bis 3 mittels verschiedener empirischer Paradigmen am Beispiel des Chinesischen untersucht.

Des Weiteren bestehen große sprachspezifische Unterschiede in der Art und Weise, wie Ereigniskomponenten segmentiert und in Sätze verpackt werden (vgl. Papafragou et al., 2008). Zahlreiche Studien zeigen, dass Unterschiede in sprachlichen Kategorien und grammatischen Strukturen häufig zu Unterschieden in der Konstruktion von Ereignissen (z.B. Flecken et al., 2015; Gerwien & von Stutterheim, 2018; Papafragou

et al., 2008; Ünal et al., 2021) sowie in anderen kognitiven Domänen wie der zeitlichen (vgl. Boroditsky, 2001) und räumlichen Domäne (vgl. Levinson, 2003; Majid et al., 2004) führen.

Vor diesem Hintergrund richtet sich die zweite Fragestellung der vorliegenden Arbeit auf den Vergleich zwischen Chinesisch und Deutsch:

- 2) Welche Gemeinsamkeiten und Unterschiede zeigen sich im Einfluss sprachlicher Strukturen auf die Ereigniskonzeptualisierung im Chinesischen und im Deutschen?

Chinesisch und Deutsch unterscheiden sich hinsichtlich vieler grammatischer Merkmale deutlich. Besonders wichtig für diese Arbeit ist die Oberflächenstruktur von Relativsätzen. Bei der Verarbeitung sprachlicher Inputs kann die semantische Struktur nicht isoliert von der Wortstellung betrachtet werden. Beim Lesen oder Hören werden Wörter und Phrasen in linearer Abfolge wahrgenommen. Die Wortstellung bestimmt, welche Elemente früher oder später genannt werden. Während das Kopfnomen im Chinesischen am Ende des Relativsatzes steht, befindet es sich im Deutschen am Anfang. Von besonderem Interesse ist daher, wie sich diese oberflächlichen syntaktischen Unterschiede auf die zugrunde liegende Ereignisrepräsentation auswirken. Konkret ergeben sich die Fragen: Orientiert sich die Aufmerksamkeitslenkung an der semantischen Struktur oder an der linearen Abfolge der genannten Ereigniskomponenten? Wird die Aufmerksamkeit früher auf die Elemente gelenkt, die in der sprachlichen Struktur früher genannt werden?

Diese Fragen lassen sich durch den sprachübergreifenden Vergleich zwischen Chinesisch und Deutsch beleuchten. Dieser Vergleich ermöglicht es, den Einfluss der Wortstellung von dem der semantischen Struktur zu trennen. Dabei werden die folgenden Hypothesen aufgestellt:

- Richtet sich die Aufmerksamkeit bei der Verarbeitung chinesischer Relativsätze zunächst auf die im Relativsyntagma ausgedrückten Informationen, während sie sich bei deutschen Relativsätzen zuerst auf das Kopfnomen konzentriert, spricht

dies dafür, dass die Wortstellung die Ereigniskonzeptualisierung beeinflusst.

- Zeigen sich hingegen trotz unterschiedlicher Oberflächenstrukturen in beiden Sprachen ähnliche Effekte, deutet dies darauf hin, dass nicht die Wortstellung, sondern die semantische Struktur der Konstruktionen den entscheidenden Einfluss auf die Ereigniskonzeptualisierung ausübt.

Speziell über den Einfluss der syntaktischen Reihenfolge von Modifikatoren auf die Enkodierung und den späteren Abruf modifizierter Nominalphrasen (NPs) besteht bislang noch kein Konsens. Einige Studien legen nahe, dass prämodifizierte NPs effizienter verarbeitet und besser im Gedächtnis enkodiert werden können als postmodifizierte NPs, da die Repräsentationen im Gedächtnis mit zunehmender Zeit tendenziell an Aktivität verlieren. Dieser Effekt wird als zeitbasierter Verfall (*time-based decay*) bezeichnet (vgl. Gibson, 1998; Lewis et al., 2006). Die Wahrscheinlichkeit des Verfalls von Repräsentationen im Kurzzeitgedächtnis steigt mit dem zeitlichen Abstand zwischen ihrer Kodierung und ihrem Abruf. Bei prämodifizierten NPs tritt das Kopfnomen näher am Abrufpunkt auf, da die Modifikatoren vorangestellt sind. Dies reduziert den zeitlichen Abstand zwischen Kodierung und Abruf des Kopfnomens und minimiert somit den Gedächtnisverfall (vgl. Gibson, 1998). Dagegen erfolgt die semantische Spezifizierung bei postmodifizierten NPs erst nach dem Kopfnomen. Der größere zeitliche Abstand zwischen der Enkodierung des Kopfnomens und seiner späteren Verarbeitung könnte zu einem stärkeren Gedächtnisverfall führen (vgl. Gibson, 1998).

Außerdem können Prämodifikatoren bestimmte Merkmale des Kopfnomens voraktivieren und aktiv im Gedächtnis aufrechterhalten. Dadurch wird die Menge der möglichen nachfolgenden Kopfnomen eingeschränkt und die Vorhersagbarkeit des Kopfnomens erhöht, was wiederum die Integration der Modifikatoren erleichtert (vgl. Altmann & Mirkovic, 2009; Hale, 2001; Karimi et al., 2019; Levy, 2008). Gemäß dem *head-reactivation*-Mechanismus werden vorhersagbare Informationen, z.B. die syntaktische Kategorie des Kopfnomens, voraktiviert und anschließend bei der Verarbeitung jedes Modifikators reaktiviert. Diese wiederholten Aktivierungen tragen

dazu bei, die Repräsentation des Kopfnomens zu stärken (vgl. Lewis & Vasishth, 2005; Lewis et al., 2006; Radvansky & Zacks, 2014).

Im Gegensatz zu dieser Ansicht argumentieren viele Studien, dass postmodifizierte NPs effizienter verarbeitet werden können als prämodifizierte, da sie mit dem „*given-before-new*“-Prinzip übereinstimmen, d.h. bereits bekannte, erwartete oder leicht zugängliche Informationen („*given*“) werden vor neuen Informationen („*new*“) präsentiert. Wenn das Kopfnomen zuerst genannt wird, kann es als kognitiver Anker dienen und die Integration der nachfolgenden Modifikatoren erleichtern. Prämodifikatoren müssen demgegenüber über eine längere Zeitspanne hinweg im Arbeitsgedächtnis aktiv gehalten werden, bis das Kopfnomen erscheint, was zusätzlichen kognitiven Aufwand (*maintenance cost*) erfordert und die Integration verzögern kann (vgl. Benatar & Clifton, 2014; Bock & Irwin, 1980; Gibson, 1998; Halliday, 1967; Karimi et al., 2019; Wagers & Phillips, 2014).

Karimi et al. (2019) stützen diesen Ansatz und liefern empirische Evidenz dafür, dass die lineare Position der Modifikatoren im Englischen einen Einfluss darauf hat, wie die modifizierten NPs im mentalen System kodiert und anschließend aus dem Gedächtnis abgerufen werden. In dem Sprachproduktionsexperiment wurden drei Bedingungen miteinander verglichen: die Baseline ohne Modifikator wie „*The king criticized the prince*“, die prämodifizierte Bedingung wie „*The cruel and merciless king criticized the prince*“ und die postmodifizierte Bedingung wie „*The king who was cruel and merciless criticized the prince*“. Die Versuchspersonen sollten aus den vorgegebenen Sätzen eine Fortsetzung formulieren. Die Ergebnisse zeigten, dass postmodifizierte NPs häufiger durch Pronomen referenziert wurden als prämodifizierte NPs. Dies spricht für eine höhere Zugänglichkeit der postmodifizierten NPs, da Pronomen häufiger für leicht zugängliche Referenten verwendet werden, während weniger zugängliche NPs eher mit vollständigen Nomina wiederholt werden.

Im anschließenden *self-paced-reading* Experiment lasen die Versuchspersonen Texte mit zwei Sätzen. Der erste Satz enthielt entweder eine unmodifizierte, eine

prämodifizierte oder eine postmodifizierte NP. Der zweite Satz enthielt ein Pronomen, das sich auf die Target-NP des ersten Satzes bezog. Es wurde gefunden, dass die Lesezeiten beim vierten Wort nach dem kritischen Pronomen („n+4“) in der postmodifizierten Bedingung kürzer waren als in der prämodifizierten. Zusammengefasst zeigen beide Experimente, dass die Position der Modifikatoren im Englischen die Verarbeitung sprachlicher Reize beeinflusst.

Es ist jedoch zu beachten, dass in der Studie von Karimi et al. (2019) neben der syntaktischen Position der Modifikatoren eine weitere Variable vorliegt, nämlich der Modifikationstyp. Während es sich bei postmodifizierten NPs um Relativsätze handelt, sind alle Prämodifikatoren Adjektive. Basierend auf allgemeiner Spracherfahrung weisen Relativsätze im Vergleich zu Adjektiven unabhängig von ihrer Position eine höhere strukturelle und semantische Komplexität auf. Sie erlauben die Einbindung ganzer Satzteile oder zusätzlicher NPs und können dadurch komplexere Informationen vermitteln. Dieses in der Sprachverwendung erworbene Wissen könnte potenziell die sprachliche Verarbeitung beeinflussen (vgl. Karimi et al., 2019).

In der vorliegenden Arbeit wird nur die Position einer sprachlichen Struktur systematisch variiert. Untersucht werden Subjekt-Relativsätze in prä- und postnominaler Position im Chinesischen und Deutschen. Ein besonderer Schwerpunkt liegt darauf, gezielt zu prüfen, ob allein die lineare Wortstellung einen Einfluss auf die Konzeptualisierung des Ereignisses ausübt.

5.2. Forschungsdesign

5.2.1. Linguistische Stimuli

Da sich die vorliegende Arbeit mit der Ereigniskonzeptualisierung befasst, müssen zunächst die zu enkodierenden Ereignisse manipuliert werden. Untersucht werden kanonische kausative Ereignisse, die sich in allen drei untersuchten Sprachstrukturen sowohl im Chinesischen als auch im Deutschen natürlich ausdrücken lassen. Konkret beinhalten diese Ereignisse jeweils zwei zentrale Teilnehmer, nämlich ein

prototypisches belebtes Agens, das die Handlung absichtlich ausführt, und ein prototypisches unbelebtes Patiens, das von der Handlung betroffen wird. In zwei Stimuli (3,7 % der gesamten Stimuli) sind die Objekte zwar unbelebte Entitäten (*Wunde* und *Haare*), aber Körperteile bzw. Merkmale eines belebten Menschen.

Die analysierten Sprachstrukturen beschränken sich auf einfache aktive transitive Hauptsätze, Subjekt-Relativsätze sowie Nominalisierungen (im Chinesischen: [NP + DE + VP]-Struktur; im Deutschen: Rektionskomposita). Sowohl Subjekt als auch Objekt bestehen ausschließlich aus einfachen Nominalen, d.h. ohne zusätzliche adjektivische oder präpositionale Erweiterungen, um potenzielle Störfaktoren zu minimieren. Im Fall der Relativsatzbedingungen besteht das Subjekt aus einem einfachen Nomen, das lediglich durch den Relativsatz modifiziert wird. Haupt- und Relativsätze werden im Präsens präsentiert. Bei Nominalisierungen wird in der Regel nur ein Argument des Basisverbs explizit kodiert. Da im Deutschen bei Rektionskomposita Objekt-Nominalisierungen am häufigsten vorkommen, wurde in den Experimenten dieses Muster verwendet (vgl. Kapitel 4.3.2; Aptacy, 2015; Eichinger, 2000; Gaeta & Zeldes, 2012). Die chinesischen Nominalisierungen wurden analog strukturiert, um eine sprachliche Konsistenz zwischen den beiden Sprachen sicherzustellen. Beispiel 5-1 zeigt je ein Beispiel der drei untersuchten Sprachstrukturen auf Chinesisch. Um die inhaltliche Vergleichbarkeit zu gewährleisten, wurden die deutschen linguistischen Stimuli durch Übersetzung aus dem Chinesischen erstellt.

Im Chinesischen ist besonders darauf zu achten, dass es eine potenzielle Verwechslung von Subjekt-Relativsätzen mit Hauptsätzen gibt, da sich ihre Anfangsstrukturen ähneln. Die typische Struktur eines Subjekt-Relativsatzes lautet [**Verb** + **Objekt** + DE + Subjekt], während ein Hauptsatz dem Muster [Subjekt + **Verb** + **Objekt**] folgt. Da das Subjekt im Chinesischen häufig ausgelassen wird, besteht die Möglichkeit, dass ein Relativsatz anfangs fälschlicherweise als Hauptsatz mit Subjekt-Drop interpretiert wird. Erst mit dem Auftreten des Markers „DE“ wird eindeutig, dass es sich um einen Relativsatz handelt (vgl. Liu, 2005). Um die potenzielle Verwechslungsgefahr zu

minimieren und die Relativsätze frühzeitig erkennbar zu machen, wird in den Experimenten das Demonstrativpronomen „*nàge*“, das laut Korpusanalysen häufig in Subjekt-Relativsätzen vorkommt (vgl. Tang, 2007; Zhang, 2019), am Anfang jedes Relativsatzes eingefügt. Die Struktur lautet somit [*nàge* + Verb + Objekt + *DE* + Subjekt]. Da die Versuchspersonen das Auftreten von Relativsätzen möglicherweise aufgrund des Demonstrativpronomens antizipieren könnten, wird „*nàge*“ auch in Hauptsätzen hinzugefügt, um die Konsistenz zwischen den Bedingungen sicherzustellen und unerwünschte Vorhersageeffekte zu vermeiden.

Bsp. 5-1

a) 那个 店员 冲泡 咖啡

nàge *diànyuán* *chōngpào* *kāfēi*

DemPron. Barista zubereiten Kaffee

„Die Barista bereitet Kaffee zu“

b) 那个 冲泡 咖啡 的 店员

nàge *chōngpào* *kāfēi* *de* *diànyuán*

DemPron. zubereiten Kaffee *DE* Barista

„Die Barista, die Kaffee zubereitet“

c) 咖啡 的 冲泡

kāfēi *de* *chōngpào*

Kaffee *DE* zubereiten

„Die Kaffeezubereitung“

5.2.2. Visuelle Stimuli

Gemäß den in Kapitel 3 ausführlich dargelegten theoretischen Überlegungen lassen sich die drei untersuchten Sprachstrukturen hinsichtlich ihrer konzeptuellen Fokussierung unterscheiden (vgl. Johnson, 1987; Lakoff, 1987; Langacker, 1987, 1991; Talmy, 2000). Während der einfache transitive Hauptsatz das gesamte kanonische Ereignis widerspiegelt, rückt der Subjekt-Relativsatz spezifisch das Agens in den Fokus.

Bei Nominalisierungen werden die explizit genannten Komponenten hervorgehoben, und zwar das Patiens und die Handlung, während das implizite Agens konzeptuell in den Hintergrund tritt.

Um die sprachlich induzierte Hervorhebung experimentell zu überprüfen, müssen die zentralen Ereigniskomponenten – Agens, Patiens und Handlung – in den visuellen Stimuli möglichst deutlich voneinander getrennt sein. Aus diesem Grund wurden die dargestellten Ereignisse so konzipiert, dass sie typischerweise mithilfe von Instrumenten ausgeführt werden. Die Instrumente fungieren dabei als visuelle Vertreter für die Handlungen.

Zusätzlich zu diesen drei Kernkomponenten wurde für jedes Ereignis ein passender Hintergrund in den visuellen Stimuli generiert, um die gesamte Szene natürlich wirken zu lassen und eine realistischere Wahrnehmungssituation zu erzeugen. Außerdem dienen die peripheren Elemente im Hintergrund in Experiment 1 als *Distractoren*, die dazu beitragen, das Forschungsziel zu verbergen.



Abbildung 5.1: Ein Beispiel der visuellen Stimuli. Es entspricht den linguistischen Stimuli in Bsp. 5-1.

5.2.3. Experimente

Operationalisierung der theoretischen Annahmen

Kapitel 3 hat herausgearbeitet, dass die drei untersuchten sprachlichen Strukturen

dieselbe grundlegende Ereignisstruktur mit denselben Teilnehmerrollen ausdrücken, diese Rollen jedoch unterschiedlich profilieren. Je nach Struktur wird entweder das Agens, die Handlung oder das Patiens konzeptuell in den Vordergrund gerückt. Wenn diese theoretisch postulierten Salienzunterschiede kognitiv wirksam sind, sollten sie sich in unterschiedlichen Verarbeitungsmustern niederschlagen. In den vorliegenden Experimenten wird diese Annahme auf drei Methode empirisch operationalisiert, nämlich Bildauswahl, Reaktionszeitmessung und Eyetracking.

Experiment 1

In Experiment 1 wird in einer *offline Picture-Selection-Task* überprüft, ob Probanden systematisch jene Bildvariante auswählen, in der genau diejenige Ereigniskomponente visuell hervorgehoben ist, die nach den theoretischen Überlegungen durch die jeweilige sprachliche Struktur konzeptuell hervorgehoben wird.

In diesem Experiment wird die visuelle Salienz einzelner Ereigniskomponenten gezielt manipuliert, um eine direkte Zuordnung zwischen der visuellen Hervorhebung im Bild und der konzeptuellen Salienz in der Sprachstruktur herzustellen. Konkret werden für jeden visuellen Stimulus vier Varianten erstellt, bei denen jeweils bestimmte Komponenten farblich markiert werden, während der Rest der Szene in Graustufen mit leichter Unschärfe dargestellt wird. In einer Bildvariante ist beispielsweise nur das Agens farbig, während alle anderen Teile in Graustufen präsentiert werden. Diese Variante sollte gemäß den theoretischen Annahmen als diejenige angesehen werden, die dem Subjekt-Relativsatz am besten entspricht, da das Agens als konzeptuelle Figur in Subjekt-Relativsätzen gilt. Das Design wird in Kapitel 6 ausführlich beschrieben.

Die Logik dieses Experiments besteht darin, dass, wenn unterschiedliche sprachliche Strukturen tatsächlich verschiedene Aspekte eines Ereignisses hervorheben, zu erwarten ist, dass die Versuchspersonen jeweils jene Bildvariante bevorzugen, deren perzeptuelle Hervorhebung mit der durch die sprachliche Struktur induzierten konzeptuellen Hervorhebung übereinstimmt. Gleichzeitig dient Experiment 1 der

Normierung der Bildstimuli für die nachfolgenden Experimente.

Experiment 2

Experiment 2 untersucht mittels Reaktionszeitmessung in einer *Picture-Sentence-Matching-Task*, ob die Übereinstimmung zwischen sprachlich profilierten und visuell hervorgehobenen Ereigniskomponenten die Integration der sprachlich und visuell induzierten Ereignisrepräsentationen erleichtert.

Frühere Studien zeigen, dass sprachliche Verarbeitung und visuelle Wahrnehmung eng miteinander verknüpft sind. Wenn bestimmte perzeptuelle Merkmale eines Objekts (z.B. dessen Form oder Bewegungsrichtung) mit der sprachlich vermittelten Information übereinstimmen, erfolgt die Verarbeitung dieser Merkmale schneller als bei einer Nichtübereinstimmung (vgl. Zwaan, Stanfield & Yaxley, 2002; Zwaan et al., 2004). Zwaan et al. (2004) argumentieren, dass beim Sprachverstehen mentale perzeptuelle Simulationen der beschriebenen Ereignisse erzeugt werden. Beispielsweise aktiviert der Satz „*He saw an eagle in the sky*“ typischerweise eine mentale Repräsentation eines fliegenden Adlers mit ausgebreiteten Flügeln. Wenn das anschließende Bild genau dieser Vorstellung entspricht, wird die sprachlich aufgebaute Repräsentation leichter mit der visuellen Information integriert, was zu kürzeren Reaktionszeiten führt. Weicht das Bild hingegen von der aktivierten Vorstellung ab (z.B. ein Adler mit angelegten Flügeln), ist zusätzliche Verarbeitung zur Anpassung der Repräsentation erforderlich (vgl. Zwaan et al., 2002).

Beim Sprachverstehen werden mentale Modelle der beschriebenen Ereignisse konstruiert (vgl. Radvansky & Zacks, 2014). Dabei sind Menschen in der Lage, die zentrale Ereignisstruktur einschließlich der Rollen wie Agens und Patiens schnell zu extrahieren, selbst bei komplexen visuellen Inputs und kurzen Darbietungszeiten (vgl. Dobel et al., 2007; Griffin & Bock, 2000; Henderson & Ferreira, 2004).

Auf der Grundlage früherer Studien werden in Experiment 2 *Match*- und *Mismatch*-Bedingungen eingesetzt, die sich an den Ergebnissen von Experiment 1 orientieren. Die

konkreten Bedingungen werden in Kapitel 7 detailliert beschrieben. Gemessen wird die Reaktionszeit zwischen dem Bild-Onset und der Entscheidung, ob das Bild zu dem zuvor gehörten linguistischen Stimulus passt.

Die zentrale Annahme ist, dass beim Sprachverstehen je nach Sprachstruktur unterschiedliche Ereignisrepräsentationen mental aktiviert werden. Wenn die farblich und sprachlich hervorgehobenen Komponenten übereinstimmt (*Match*-Bedingung), sollte dies die Integration der sprachlich evozierten Repräsentation und der visuell induzierten Repräsentation erleichtern und zu kürzeren Reaktionszeiten führen. Bei Nichtübereinstimmung ist hingegen mit zusätzlichem Integrationsaufwand und entsprechend längeren Reaktionszeiten zu rechnen (*Mismatch*-Bedingung). Mit diesem Experiment wird geprüft, ob unterschiedliche sprachliche Strukturen tatsächlich spezifische Aspekte eines Ereignisses hervorheben und ob sich diese Unterschiede in der Echtzeitverarbeitung niederschlagen.

Experiment 3

Es bleibt noch offen, wie diese sprachlichen Strukturen die Aufmerksamkeit während der visuellen Wahrnehmung lenken und welche Komponente bei der Verarbeitung der Ereignisrepräsentation priorisiert wird. Diese Fragen werden in Experiment 3 mittels Eyetracking untersucht.

Augenbewegungen sind ein direkter Indikator für die Aufmerksamkeitsverteilung während der sprachlichen Verarbeitung in Echtzeit, da Menschen dazu neigen, auf das zu fixieren, woran sie gerade denken oder was sie gerade verarbeiten. Dieses Phänomen wird als *eye-mind-link* bezeichnet (vgl. Carter & Luke, 2020; Griffin & Bock, 2000; Rayner, 2009). Solche Augenbewegungen werden sowohl durch *bottom-up*-Faktoren, z.B. visuelle Eigenschaften eines Stimulus auf niedriger Ebene, als auch durch *top-down*-Faktoren, z.B. kognitive Prozesse auf hoher Ebene, wie Erwartung, Vorwissen, Gedächtnis oder Sprachgebrauch, gesteuert (vgl. Altmann & Kamide, 1999; Carter & Luke, 2020). Augenbewegungen liefern somit wertvolle Einblicke in die zugrunde

liegenden Ereignisrepräsentationen und die Interaktion zwischen Sprache und visueller Wahrnehmung (vgl. Papafragou et al., 2008).

Eine Fixation bezeichnet eine zeitlich relativ stabile Augenposition, in der visuelle Informationen aufgenommen werden (vgl. Rayner, 2009). Die Position und Dauer von Fixationen spiegeln somit wider, wie Versuchspersonen ihre Aufmerksamkeit innerhalb einer visuellen Szene verteilen. Die Aufmerksamkeitsverteilung hängt stark davon ab, welche Aspekte einer Szene für die Bewältigung einer bestimmten Aufgabe als besonders wichtig oder informativ erachtet werden (vgl. Papafragou et al., 2008).

Insbesondere die erste Fixation auf ein Objekt dient als sensibler Indikator für die Salienz oder Priorität einer Ereigniskomponente. Dies liegt daran, dass sie bereits in den allerersten Millisekunden während der Betrachtung der Szene erfolgt und somit die initiale Aufmerksamkeitszuteilung abbildet und noch nicht von späteren kognitiven Verarbeitungsschritten beeinflusst ist (vgl. Gerwien & Flecken, 2016; Sauppe & Flecken, 2021). All dies ermöglicht es, anhand des Blickverhaltens Rückschlüsse darauf zu ziehen, welche Ereigniskomponenten als zentral wahrgenommen und kognitiv priorisiert werden.

Empirische Studien zeigten, dass Sprache die Salienz bestimmter Ereigniskomponenten bei der visuellen Wahrnehmung modulieren kann (vgl. Divjak et al., 2020; Griffin & Bock, 2000). Sauppe und Flecken (2021) fanden beispielsweise heraus, dass die syntaktische Struktur (aktiv oder passiv) zu systematischen Unterschieden in der ersten Fixation auf eine nachfolgende Szene führte. Demnach richteten die Versuchspersonen nach einer aktiven Beschreibung der *Cue*-Szene ihre erste Fixation bevorzugt auf das Agens der nachfolgenden neuen Target-Szene, während nach einem Passivsatz hingegen das Patiens zuerst fixiert wurde. Diese Ergebnisse weisen auf einen frühen Einfluss sprachlicher Strukturen auf die visuelle Ereigniswahrnehmung hin, selbst bei sehr kurzer Darbietungsdauer (300 ms).

Des Weiteren untersuchten Divjak et al. (2020) ebenfalls mithilfe des Eyetrackings, ob

und wie verschiedene sprachliche Konstruktionen die visuelle Wahrnehmung und Verarbeitung einer Szene beeinflussen. In dieser Studie wurden drei syntaktische Bedingungen behandelt, und zwar Präposition („*The poster is above the bed*“ vs. „*The bed is below the poster*“), Genus Verbi („*The policeman arrested the thief*“ vs. „*The thief was arrested by the policeman*“) und Dativ („*The boy gave the girl a flower*“ vs. „*The boy gave a flower to the girl*“). In jeder Bedingung wurde die erste Variante als typisch und die zweite als atypisch angesehen, da die erste Variante im jeweiligen Kontext als häufiger verwendet oder intuitiv bevorzugt gilt.

Im Experiment hörten die Versuchspersonen eine von zwei Varianten und sahen dieselbe Szene. Gemessen wurden die Reihenfolge der Fixationen auf die definierten *Areas of Interest* (AOIs), die erste Gazedauer und die gesamte Gazedauer pro AOI. Die Reihenfolge der Fixationen auf die AOIs zeigt an, welche Komponenten einer Szene zuerst verarbeitet werden. Die erste Gazedauer dient als Indikator für den frühen Verarbeitungsaufwand, wenn der Blick erstmals auf ein bestimmtes AOI fällt. Die gesamte Gazedauer reflektiert die Integration der Informationen. Die Gazedauer gibt Aufschluss über den kognitiven Aufwand, der durch die Verarbeitung der Informationen entsteht (vgl. Divjak et al., 2020).

Die Ergebnisse zeigten, dass diese drei sprachlichen Bedingungen einen unterschiedlichen Einfluss auf die visuelle Wahrnehmung ausübten. Bei der Präposition-Bedingung wurde das leicht bewegliche Objekt (z.B. „*poster*“) nach einer typischen Variante häufiger zuerst fixiert und länger betrachtet, während seine Salienz nach einer atypischen Variante abnahm. Bei der Genus-Verbi-Bedingung wurde die Handlung (z.B. „*arrested*“) nach einem Aktivsatz überwiegend zuerst fixiert, während dieser Effekt nach einem Passivsatz nicht mehr signifikant war und das Patiens (z.B. „*the thief*“) fast genauso häufig zuerst fixiert wurde. Hinsichtlich der Gazedauer zeigte sich aber kein signifikanter Unterschied zwischen beiden Varianten. Bei der Dativ-Bedingung, bei der Rezipient und Objekt in den beiden Varianten in unterschiedlicher Reihenfolge genannt wurden, wurden jedoch keine signifikanten Effekte beobachtet.

Die Befunde verdeutlichen, dass sprachliche Strukturen die Salienz der Ereigniskomponenten während der visuellen Wahrnehmung der Szenen beeinflussen. Eine Szene wird visuell anders verarbeitet, wenn sie zuvor in einer bestimmten Weise formuliert wurde. Diese Effekte treten jedoch je nach syntaktischer Konstruktion unterschiedlich stark und zu unterschiedlichen Zeitpunkten in der Verarbeitung auf.

Allerdings konzentrierten sich frühere Studien vorwiegend auf den Vergleich zwischen Hauptsätzen und beschränkten sich auf eine einzelne Sprache. Komplexere Vergleiche zwischen unterschiedlichen Konstruktionen, wie eingebettete Nebensätze oder Nominalphrasen, blieben dabei weitgehend unberücksichtigt. Vor diesem Hintergrund zielt Experiment 3 darauf ab, die modulierenden Effekte der drei sprachlichen Strukturen, nämlich Hauptsatz, Subjekt-Relativsatz und Nominalisierung, auf die Salienz spezifischer Ereigniskomponenten mittels Eyetracking zu überprüfen. Dabei werden die Position der Erstfixation, die Fixationslatenz und die gesamte Gazedauer auf jedem AOI gemessen. Dies erlaubt Rückschlüsse darauf, ob und wie die Sprachstrukturen frühe und späte Phasen der visuellen Ereigniswahrnehmung und Ereigniskonzeptualisierung beeinflussen.

Experiment 4

In der vorliegenden Arbeit wird zudem der Frage nachgegangen, ob sich die Auswirkungen der drei sprachlichen Strukturen auf die Wahrnehmung und Konzeptualisierung eines Ereignisses auch in anderen Sprachen beobachten lassen. Zu diesem Zweck wird in Experiment 4 eine direkte Replikation von Experiment 3 mit deutschsprachigen Versuchspersonen durchgeführt. Dabei werden die linguistischen Stimuli ins Deutsche übersetzt, während alle anderen experimentellen Bedingungen konstant gehalten werden.

Ein besonderer Punkt dieses sprachübergreifenden Vergleichs liegt in der Wortstellung von Relativsätzen. Während das Kopfnomen im Chinesischen am Ende des Relativsatzes steht, befindet es sich im Deutschen am Anfang. Dieser Unterschied in

der Wortstellung ermöglicht eine Untersuchung darüber, ob die lineare Anordnung der Satzbestandteile die mentale Ereignisrepräsentation beeinflusst.

Die Studie von Divjak et al. (2020) liefert Hinweise darauf, dass die Reihenfolge der Erwähnung die visuelle Wahrnehmung nicht bei allen sprachlichen Konstruktionen gleichermaßen beeinflusst. In der Präpositionsbedingung spiegelten frühe und späte Fixationsmuster die syntaktische Reihenfolge klar wider. In der Genus-Verbi-Bedingung fand sich ein Effekt auf die Reihenfolge der Fixationen auf AOIs. Aber es wurden keine signifikanten Effekte bei der Gazedauer beobachtet. In der Dativ-Bedingung ergaben sich keine signifikanten Effekte.

Diese unterschiedlichen Effekte können auf den Grad der Perspektiv- bzw. Fokusänderung zurückgeführt werden. Bei der Präposition-Bedingung wird die Figur-Grund-Beziehung der Objekte durch verschiedene Präpositionen wie „*below*“ und „*above*“ verändert. Bei der Genus-Verbi-Bedingung bleiben die thematischen Rollen der Ereignisteilnehmer in beiden Varianten erhalten, aber das *Viewpoint* und die grammatische Funktion ändern sich. Dadurch verschiebt sich der Fokus des Ereignisses. Im Aktivsatz liegt der Fokus auf dem Agens, während im Passivsatz das Patiens hervorgehoben wird. Bei der Dativ-Bedingung gibt es hingegen keine eindeutige Perspektivverschiebung, da beide Varianten traditionell als semantisch gleichwertig betrachtet werden (vgl. Givón, 1984). Obwohl neuere Forschungen auf subtile semantische Unterschiede hinweisen (vgl. Bresnan & Ford, 2010; Goldberg, 2005), sind diese Effekte weniger ausgeprägt als bei den anderen beiden syntaktischen Bedingungen.

Hinsichtlich der Subjektiv-Relativsätze im Chinesischen und Deutschen zeigt sich eine Ähnlichkeit zur Dativ-Bedingung bei Divjak et al. (2020). Ausgehend von den in Kapitel 3 und 4 dargestellten theoretischen Annahmen gibt es hier keine Perspektiv- oder Fokusverschiebung. Die thematischen Rollen und die konzeptuelle Figur (das Kopfnomen) bleiben in beiden Sprachen gleich. Es ändert sich nur die Wortstellung. Der Vergleich zwischen Experiment 3 (Chinesisch) und Experiment 4 (Deutsch)

ermöglicht es, in einem sprachübergreifenden Kontext zu prüfen, ob die Konzeptualisierung von Ereignissen durch die konzeptuelle Salienz der sprachlich hervorgehobenen Komponenten oder durch die lineare Anordnung beeinflusst wird.

6. Experiment 1

6.1. Hypothesen

Das Experiment verfolgt das Ziel, zu untersuchen, ob und wenn ja, welche Komponenten eines Ereignisses durch verschiedene sprachliche Strukturen hervorgehoben werden. Wie in Kapitel 3 auf Grundlage der Konzepte von Johnson (1987), Lakoff (1987), Langacker (1987, 1991, 2013) und Talmy (2000) erläutert, rücken die drei sprachlichen Strukturen unterschiedliche Aspekte eines Ereignisses in den Fokus. Einfache aktive transitive Hauptsätze dienen als der prototypische Satztyp, um ein kanonisches kausatives Ereignis sprachlich zu kodieren (vgl. Langacker, 1991, 2013). Im Subjekt-Relativsatz wird das Agens als Kopfnomen hervorgehoben, das durch das Relativsyntagma weiter spezifiziert wird, wodurch seine konzeptuelle Salienz gegenüber den übrigen Ereigniskomponenten erhöht wird. Nominalisierungen fokussieren demgegenüber primär auf das Patiens und die Handlung, da explizit genannte Ereigniskomponenten generell salienter sind als implizite (vgl. Talmy, 2000).

Um diese theoretischen Überlegungen experimentell zu testen, wurde eine *offline Picture-Selection-Task* durchgeführt, bei der die Salienz einzelner Komponenten eines Ereignisses visuell manipuliert wurde. In den visuellen Stimuli wurden mehrere Bildvarianten erstellt, in denen unterschiedliche Komponenten des Ereignisses farblich hervorgehoben waren (siehe Abbildung 6.1). Die zugrunde liegende Annahme war, dass die Versuchspersonen zu einem gegebenen linguistischen Stimulus dasjenige Bild auswählen würden, dessen visuell betonte Komponente am besten mit dem sprachlich betonten Aspekt korrespondiert.

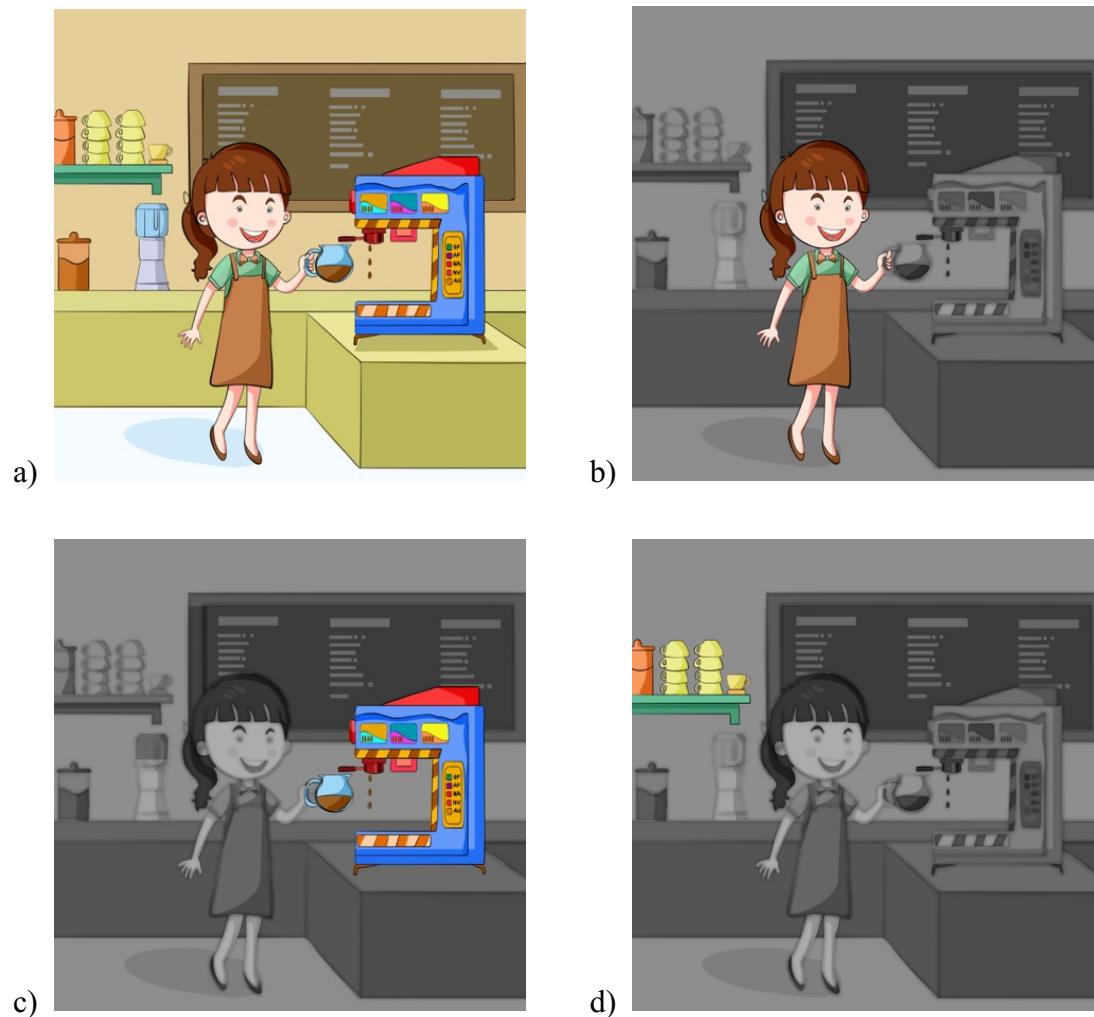


Abbildung 6.1: Ein Beispiel der vier Varianten eines kritischen visuellen Stimulus in Experiment 1.

Daraus ergeben sich die folgenden spezifischen Hypothesen für Experiment 1:

- 1) Bei Hauptsätzen sollte bevorzugt eine Bildvariante ausgewählt werden, in der alle Ereigniskomponenten sowie der Hintergrund, der das *Setting* repräsentiert, gleichermaßen deutlich dargestellt sind, ohne dass ein bestimmter Teilnehmer visuell stark hervorgehoben wird (Abbildung 6.1 a).
- 2) Bei Subjekt-Relativsätzen, in denen das Agens sprachlich besonders salient gemacht wird, sollte häufiger die Bildvariante gewählt werden, in der ausschließlich das Agens farblich hervorgehoben ist (Abbildung 6.1 b).
- 3) Bei Nominalisierungen, die besonders das Patiens und die Handlung betonen, sollte entsprechend häufiger die Bildvariante gewählt werden, in der das Patiens und das Instrument, welches die Handlung symbolisiert, farblich hervorgehoben sind

(Abbildung 6.1 c).

6.2. Methodisches Vorgehen

6.2.1. Versuchspersonen

Insgesamt nahmen 30 Versuchspersonen an dem Experiment teil, davon waren 29 Frauen. Das Alter der Versuchspersonen lag zwischen 18 und 34 Jahren (Durchschnittsalter: 23,2 Jahre). Von ihnen waren 17 Versuchspersonen Studierende am *Jiangnan Art Vocational College*, 9 waren Angestellte des Volksgerichts des Bezirks Qianjiang, China und 4 waren Studierende an der Universität Heidelberg. Alle Versuchspersonen waren chinesische Muttersprachler. Sie hatten alle ein normales Hörvermögen und zeigten keine neurologischen Störungen oder Lesestörungen. Alle Versuchspersonen nahmen freiwillig an dem Experiment teil. Die Versuchspersonen des *Jiangnan Art Vocational College* erhielten jeweils 50 *Yuan* als Belohnung.

6.2.2. Material

In diesem Experiment wurden insgesamt 54 linguistische Stimulus-Sets als kritische Stimuli verwendet. Jedes Set bestand aus drei sprachlichen Strukturen, die sich inhaltlich auf dasselbe kausative Ereignis beziehen, nämlich einem einfachen aktiven transitiven Hauptsatz, einem Subjekt-Relativsatz und einer Nominalisierung. Jeder linguistische Stimulus wird in seiner einfachsten und prototypischsten Oberflächenform verwendet (siehe Beispiel 5-1 in Kapitel 5.2.1). Die vollständigen linguistischen Stimuli sind im Anhang zu finden.

Zudem wurden für die linguistischen Stimuli entsprechende visuelle Stimuli digital erstellt. Jedes Bild stellte ein kausatives Ereignis mit einer menschlichen Figur als Agens dar, die ein Instrument in der Hand hält und damit eine Handlung am Patiens ausführt. In den meisten Fällen handelte es sich beim Patiens um eine unbelebte Entität. Ausnahmen waren zwei Stimuli, bei denen die Patiens Teil einer Person waren (*Wunden* und *Haare*). Dabei wurden diese beiden menschlichen Figuren in ihrer Rückseite aufgezeichnet. Unter der Prämisse, den Bildinhalt natürlich und angemessen zu halten,

sollte sich das Instrument möglichst nicht mit dem Kopf des Agens überlappen, um eine eindeutige visuelle Trennung der Komponenten Agens, Patiens und Instrument zu gewährleisten. Außerdem wurde für jedes Bild ein einfacher, kontextbezogener Hintergrund mit zwei bis drei linguistisch irrelevanten Objekten erstellt (z. B. ein Bürgersteig in einem Stadtviertel).

Für jedes Bild wurden vier visuelle Varianten erstellt, welche sich durch eine unterschiedliche farbliche Hervorhebung der Ereigniskomponenten unterschieden (siehe Abbildung 6.1). In Bildvariante a) wurde die gesamte Szene farbig dargestellt. In Bildvariante b) wurde nur das Agens farbig hervorgehoben, während der Rest des Bildes in Graustufen dargestellt und zusätzlich leicht verwischt wurde. In Bildvariante c) wurden nur das Instrument, das die Handlung repräsentiert, und das Patiens farbig dargestellt, während der übrige Teil des Bildes in Graustufen mit leichter Unschärfe verblieb. Zusätzlich wurde Bildvariante d) erstellt, in der lediglich ein linguistisch irrelevantes Objekt im Hintergrund farbig hervorgehoben wurde, während alle kritischen Ereigniskomponenten grau und leicht unscharf dargestellt waren. Diese Variante diente zur Überprüfung, ob die Versuchspersonen tatsächlich auf die linguistischen Stimuli achteten und den Zusammenhang zwischen Farbe und Fokus erkannten. Zudem trug sie dazu bei, den Zweck des Experiments zu verschleiern.

Im Experiment wurde jeder Sprachstruktur aus demselben linguistischen Stimulus-Set ein identischer visueller Stimulus zugewiesen. Bei jedem Trial wurden alle vier Varianten eines Bildes gleichzeitig in einer Reihe präsentiert. Das Target-Bild variierte je nach sprachlicher Struktur. Konkret war Bildvariante a) das Target-Bild für Hauptsätze, Bildvariante b) für Relativsätze und Bildvariante c) für Nominalisierungen. Jede Bildvariante erschien gleichmäßig häufig an den vier Positionen der Reihe, sodass kein Bias für eine bestimmte Position entsteht. Alle verwendeten visuellen Stimuli befinden sich im Anhang.

Im Experiment wurden zusätzlich 18 Filler eingesetzt, bei denen zwei Objekte in räumlicher Beziehung zueinander beschrieben wurden, wie in Beispiel 6-1. Im

Unterschied zu den kritischen linguistischen Stimuli bestanden alle Filler aus vollständigen Hauptsätzen mit unbelebten Subjekten. Diese Form diente einerseits dazu, die linguistischen Strukturen innerhalb jeder experimentellen Liste auszubalancieren. In jeder Liste bestanden die kritischen Stimuli aus 18 Hauptsätzen, 18 Relativsätzen und 18 Nominalisierungen. Die Relativsätze und Nominalisierungen enthielten dabei jeweils den Marker *DE* und waren Nominalphrasen, während die Hauptsätze vollständige Sätze ohne *DE* waren (siehe Kapitel 4, wo die sprachlichen Strukturen genau beschrieben werden, und Bsp. 5-1). Um die Anzahl der Nominalphrasen und der vollständigen Sätze sowie die Struktur mit und ohne Marker *DE* auszugleichen, wurden die zusätzlichen 18 Filler-Hauptsätze verwendet. Andererseits enthielten die Haupt- und Relativsätze der kritischen Stimuli immer ein belebtes Agens, während in den Nominalisierungen nur unbelebte Entitäten explizit erwähnt wurden. Um dies auszugleichen, wurden die Filler ausschließlich mit unbelebten Subjekten gebildet.

Wie die kritischen Bilder stellten auch die Filler-Bilder Ereignisse mit einer oder zwei menschlichen Figuren dar. Für jedes Filler-Trial wurden auch vier Bildvarianten erstellt, die sich in der farblichen Hervorhebung einzelner Bildkomponenten unterschieden (siehe Abbildung 6.2). Allerdings variierte der farblich hervorgehobene Teil leicht von den kritischen Stimuli. In Bildvariante a) war die gesamte Szene farbig, während in Bildvariante b) nur das unbelebte Subjekt des linguistischen Stimulus farbig markiert war. In Bildvariante c) wurde eine menschliche Figur in der Szene farbig hervorgehoben und in Bildvariante d) ein linguistisch irrelevantes Objekt.

Bsp. 6-1 那面 镜子 挂 在 墙 上
 nàmiàn jìngzǐ guà zài qiáng shàng
 DemPron. Spiegel hängen Aspekt Wand oben
 „der Spiegel hängt an der Wand.“

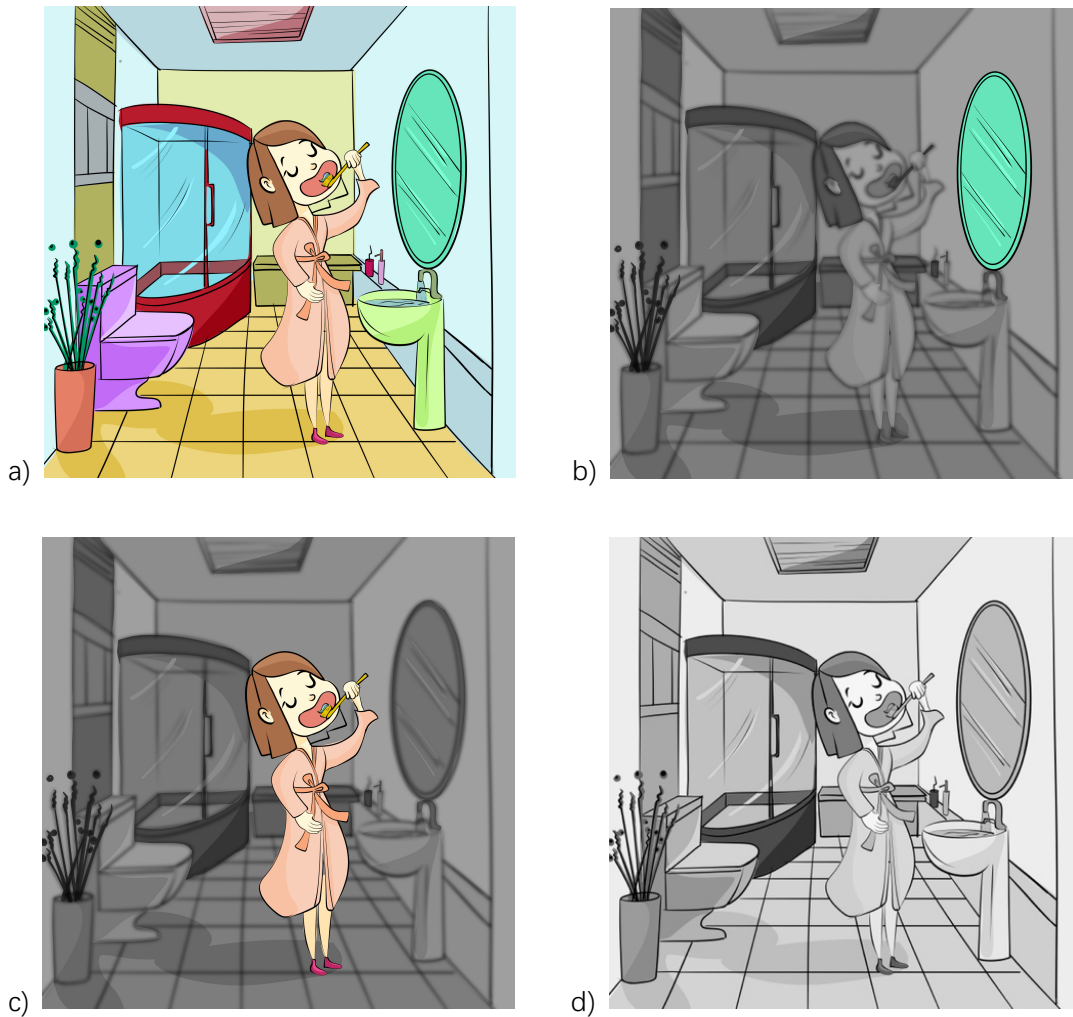


Abbildung 6.2: Ein Beispiel der Bildvarianten der Filler-Bedingung in Experiment 1.

Ange­si­chts der un­ter­schied­li­chen Sprach­struk­tu­ren wur­den drei ex­per­i­men­tel­le Listen er­stellt. Alle lin­gu­is­ti­schen Stimuli wur­den gleich­mä­ßig auf die drei Listen ver­teilt. Jede Liste ent­hielt nur eine sprach­li­che Struktur aus einem kri­ti­schen Stimulus-Set. Die Filler wa­ren in allen drei ex­per­i­men­tel­len Listen iden­tisch. Alle kri­ti­schen Stimuli und Filler wur­den in jeder Liste in ran­do­mi­sier­ter Rei­hen­fol­ge prä­sen­tiert.

Vor Be­ginn des ex­per­i­men­tel­len Blocks wur­de ein Übungs­block an­ge­setzt. Dieser be­stand aus in­ge­sam­ 12 Practice-Trials, wo­bei je­weils drei Stimuli der kri­ti­schen sprach­li­chen Strukturen und drei Filler-Stimuli dar­ge­boten wur­den. Ziel des Übungs­blocks war es, den Ver­suchs­per­so­nen eine Ein­ge­wöhnungs­phase zu er­mög­li­chen. Durch die Übun­gen konn­ten die Ver­suchs­per­so­nen einen klaren Über­blick über den Auf­bau des Ex­per­i­ments er­hal­ten, was ihnen half, sich bes­ser dar­auf

zu konzentrieren. Die Stimuli im Übungsblock waren in allen drei Listen identisch.

6.2.3. Durchführung

Das Experiment wurde mithilfe des Online Fragebogen-Tools *Wenjuanxing* (<https://www.wjx.cn/>) durchgeführt. Jede Versuchsperson nahm einzeln teil und wurde dabei von einem/r Experimentleiter/in begleitet, entweder auf dem Universitätscampus oder im Büro. Um den eigentlichen Zweck des Experiments zu verschleiern, wurden die Versuchspersonen vor Beginn des Experiments darüber informiert, dass das Ziel des Experiments darin bestehe, eine innovative Lehrmethode für Chinesischlernende zu testen. Konkret wurde ihnen vermittelt, dass durch dieses Experiment geprüft werden sollte, ob die visuelle Präsentation den Lernenden hilft, die Bedeutung und den Gebrauch verschiedener sprachlicher Strukturen intuitiv zu erfassen.

Für jede der drei experimentellen Listen wurde ein individueller Link zum Online-Fragebogen erstellt. Jeder Versuchsperson wurde zufällig ein Link zugewiesen, den sie auf ihrem eigenen PC oder Laptop öffnete. Auf der ersten Seite sahen die Versuchspersonen zunächst eine kurze Instruktion auf Chinesisch: *„Das Experiment untersucht, ob Bilder beim Erlernen der Bedeutung und des Gebrauchs verschiedener sprachlicher Strukturen helfen können. Bitte versuchen Sie sich den Verwendungskontext des Satzes vorzustellen und wählen Sie dann das Bild aus, das diesen Verwendungskontext am besten widerspiegelt.“* Anschließend gaben die Versuchspersonen auf der Website ihre demographischen Daten ein, und zwar Geschlecht, Alter und Geburtsort. Danach wurden die Informationen über das Experiment und die konkrete Aufgabe vorgestellt: *„Bei jedem Trial wurde Ihnen ein Satz oder eine Phrase zusammen mit vier Bildern gleichzeitig präsentiert. Bitte lesen Sie zuerst den Satz oder die Phrase sorgfältig und dann wählen Sie das Bild aus, das Ihrer Meinung nach am besten dazu passt“.*

Vor Beginn des experimentellen Blocks bearbeiteten die Versuchspersonen zunächst 12 Übungstrials. Anschließend wurden sie gefragt, ob sie irgendwelche Fragen zur Aufgabe hätten. Etwaige Fragen wurden erklärt und beantwortet. Danach begann der

experimentelle Block. Die Trials wurden einzeln auf dem Bildschirm präsentiert. Die Versuchspersonen wählten eine Option durch Mausklick oder Touchpad aus. Für die Auswahl gab es kein Zeitlimit. Nach der Auswahl konnten die Versuchspersonen auf die Schaltfläche „Zur nächsten Seite“ klicken, um mit dem nächsten Trial fortzufahren. Wenn sie keine Auswahl trafen, wurde dies durch eine Fehlermeldung blockiert. Es war den Versuchspersonen nicht möglich, zum vorherigen Trial zurückzukehren. Sobald alle Trials beantwortet waren, wurde die Umfrage automatisch abgegeben.



Abbildung 6.3: Ein Beispiel-Trial im Fragebogen in Experiment 1. Die deutschen Übersetzungen in den Klammern waren für die Versuchspersonen unsichtbar.

6.2.4. Datenkodierung und -analyse

Das Experiment dauerte durchschnittlich 17,24 Minuten. Vor der Datenanalyse wurden die gesammelten Antworten systematisch organisiert und kodiert. Da das Experiment darauf abzielte, die Auswirkungen unterschiedlicher Sprachstrukturen auf die Hervorhebung spezifischer Ereigniskomponenten zu untersuchen, lag der Schwerpunkt auf den Unterschieden in der Bildauswahl zwischen den drei Sprachstrukturen. Die sprachliche Struktur bildete die unabhängige Variable, und die Bildvariante die abhängige Variable.

Gemäß den Hypothesen wurde erwartet, dass die Bildvarianten jeweils abhängig von der Sprachstruktur bevorzugt gewählt werden: Variante a) sollte für Hauptsätze, Variante b) für Relativsätze und Variante c) für Nominalisierungen gewählt werden. Entsprechend wurden die Varianten als Targets den jeweiligen Sprachstrukturen zugeordnet. Zur statistischen Überprüfung der Hypothesen wurde der *Bayesian Proportion Test* durchgeführt. Die Bayes'sche Inferenz ist mit bestimmten Axiomen der rationalen Inferenz vereinbar und ermöglicht eine einfache Definition und Interpretation aufgrund der Verwendung von Wahrscheinlichkeiten (vgl. O'Hagan & Forster, 2004). Darüber hinaus erlaubt sie eine flexible Modellierung komplexer Datenstrukturen und bezieht sämtliche verfügbare Informationen in die Analyse mit ein (vgl. O'Hagan & Forster, 2004).

Der *Bayesian Proportion Test* wurde separat für die Auswahl der Varianten a), b) und c) durchgeführt. Da Variante d) nur einmal aus den insgesamt 1620 Beobachtungen ausgewählt wurde und nicht als Target einer Sprachstruktur galt, entfiel ein Test für diese Variante. Zur Durchführung des Tests wurde überprüft, ob jede der drei Sprachstrukturen (Hauptsatz, Subjekt-Relativsatz und Nominalisierung) systematisch mit einer bestimmten Bildvariante assoziiert war. Dabei wurde für jede Bildvariante einzeln analysiert, wie häufig diese jeweils in Kombination mit einer bestimmten Sprachstruktur ausgewählt wurde. Die Auswahl derjenigen Variante, die gemäß den Hypothesen zur jeweiligen Sprachstruktur passte, wurde als „Erfolg“ definiert. Der Test schätzt dann die „Erfolgsquoten“, also die relativen Häufigkeiten, mit denen eine Variante in Abhängigkeit von der Sprachstruktur ausgewählt wurde.

Zur Darstellung der Ergebnisse wurden Höchstdichteintervalle (HDIs) zusammen mit den Schätzungen der posterioren Wahrscheinlichkeitsverteilung verwendet. HDIs fassen die glaubwürdigsten Werte der Posteriorverteilungen zusammen. In der vorliegenden Arbeit wurde ein 95%-HDI verwendet, um den Bereich der plausiblen Parameterwerte einzuschränken (vgl. Kruschke, 2011).

Im nächsten Schritt wurde das Konzept der *region of practical equivalence* (ROPE)

angewandt, um die Wahrscheinlichkeit zu berechnen, ob der Unterschied zwischen den sprachlichen Strukturen als gering betrachtet werden kann. ROPE definiert einen Bereich von Parameterwerten, die aus praktischen Gründen als äquivalent zum Nullwert angesehen werden. Wenn der 95%-HDI vollständig außerhalb des ROPE-Bereichs liegt, wird der Nullwert abgelehnt und auf einen signifikanten Unterschied zwischen den Gruppen geschlossen. Liegt der 95%-HDI hingegen vollständig innerhalb des ROPE-Bereichs, wird der Nullwert akzeptiert und es wird davon ausgegangen, dass kein signifikanter Unterschied zwischen den Gruppen besteht (vgl. Kruschke, 2011). Im vorliegenden Experiment wurde die ROPE mit einem Intervall von $[-5, 5]$ Prozentpunkten festgelegt (vgl. Bååth, 2014).

Alle statistischen Analysen wurden mit der Software R, Version 4.2.1, durchgeführt. Für die Datenverarbeitung und -analyse wurden die Pakete „tidyverse“, Version 1.3.2 (Wickham et al., 2019) sowie „BayesianFirstAid“, Version 0.1 (Bååth, 2014) verwendet.

6.3. Ergebnisse

Insgesamt wurden 1620 Beobachtungen gesammelt und in die Auswertung einbezogen, davon jeweils 540 Beobachtungen für die Hauptsatz-, die Relativsatz- und die Nominalisierungsbedingung.

Die Ergebnisse zeigen deutlich, dass es je nach Sprachstruktur große Unterschiede bei der Bildauswahl gibt. Beim Hauptsatz betrug die Proportion der Auswahl der Variante a) 75,19%, während die Proportionen der Varianten b) und c) bei 11,48% bzw. 13,33% lagen. Beim Relativsatz wurde die Variante b) mit 74,07% am häufigsten gewählt, während die Proportionen der Varianten a) und c) bei 22,78% bzw. 3,15% lagen. Bei der Nominalisierung wurde die Variante c) mit einer beeindruckenden Proportion von 86,11% ausgewählt, während die Proportionen für die Varianten a) und b) bei 12,77% bzw. 0,93% lagen (vgl. Abbildung 6.4).

In allen Beobachtungen wurde Bildvariante d) nur einmal ausgewählt. Dies deutet darauf hin, dass die Versuchspersonen den Zusammenhang zwischen visueller und

konzeptueller Hervorhebung klar erkannt haben.

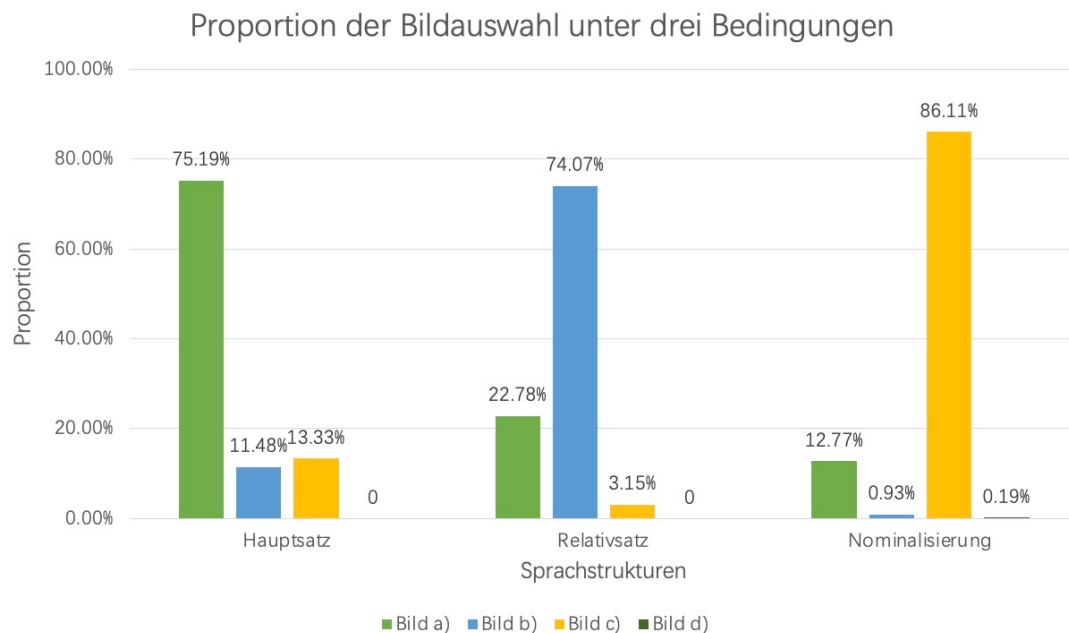


Abbildung 6.4: Proportion der Bildwahl nach sprachlicher Struktur (Experiment 1).

Der *Bayesian Proportion Test* wurde durchgeführt, um die Unterschiede in den Erfolgsquoten zwischen drei Gruppen zu untersuchen. Die Ergebnisse bezüglich Variante a) zeigten, dass diese Variante am häufigsten für den Hauptsatz gewählt wurde (Median posterior $r = 0,75$, 95% HDI = $[0,71, 0,79]$), während die Proportionen für den Relativsatz (Median posterior $r = 0,23$, 95% HDI = $[0,19, 0,26]$) und die Nominalisierung (Median posterior $r = 0,13$, 95% HDI = $[0,10, 0,16]$) vergleichsweise geringer waren. Es gab deutliche und zuverlässige Unterschiede zwischen dem Hauptsatz und dem Relativsatz ($\theta_1 - \theta_2 = 0,52$, 95% HDI = $[0,47, 0,57]$, ROPE = 0%) sowie zwischen dem Hauptsatz und der Nominalisierung ($\theta_1 - \theta_3 = 0,62$, 95% HDI = $[0,58, 0,67]$, ROPE = 0%). Dies spricht dafür, dass diese Unterschiede statistisch zuverlässig sind. Die Wahrscheinlichkeit, den Nullwert (keine Differenz zwischen den Bedingungen) zu verwerfen, betrug jeweils 100%.

Variante b) wurde am häufigsten für den Relativsatz gewählt (Median posterior $r = 0,74$, 95% HDI = $[0,70, 0,78]$), während die Proportionen für den Hauptsatz (Median posterior $r = 0,12$, 95% HDI = $[0,09, 0,14]$) und die Nominalisierung (Median posterior $r = 0,01$, 95% HDI = $[0,004, 0,02]$) geringer waren. Diese Unterschiede waren

jeweils überzeugend belegt. Zwischen dem Relativsatz und dem Hauptsatz ($\theta_2 - \theta_1 = 0,62$, 95% HDI = [0,58, 0,67], ROPE = 0%) sowie zwischen dem Relativsatz und der Nominalisierung ($\theta_2 - \theta_3 = 0,73$, 95% HDI = [0,69, 0,77], ROPE = 0%) wurde der Nullwert jeweils mit einer Wahrscheinlichkeit von 100% abgelehnt.

Die Analyse für Variante c) ergab, dass diese Variante am häufigsten für die Nominalisierung gewählt wurde (Median posterior $r = 0,86$, 95% HDI = [0,83, 0,89]). Die Proportionen für die Hauptsatzbedingung (Median posterior $r = 0,13$, 95% HDI = [0,11, 0,16]) und die Relativsatzbedingung (Median posterior $r = 0,03$, 95% HDI = [0,01, 0,05]) fielen deutlich geringer aus. Die geschätzte paarweise Differenz zwischen Nominalisierung und Hauptsatz betrug $\theta_3 - \theta_1 = 0,73$ (95% HDI = [0,68, 0,76], ROPE = 0%), und die Differenz zwischen Nominalisierung und Relativsatz $\theta_3 - \theta_2 = 0,83$ (95% HDI = [0,79, 0,86], ROPE = 0%). In beiden Fällen lag die posterior berechnete Wahrscheinlichkeit, dass der Unterschied größer als null ist, bei 100%, was auf einen robusten und statistisch bedeutsamen Effekt hinweist.

Sprachstruktur	Variante a [95% HDI]	Variante b [95% HDI]	Variante c [95% HDI]
Hauptsatz	0,75 [0,71, 0,79]	0,12 [0,09, 0,14]	0,13 [0,11, 0,16]
Relativsatz	0,23 [0,19, 0,26]	0,74 [0,70, 0,78]	0,03 [0,01, 0,05]
Nominalisierung	0,13 [0,10, 0,16]	0,01 [0,004, 0,02]	0,86 [0,83, 0,89]

Tabelle 6.1: Ergebnisse der *Bayesian Proportion Tests* (Experiment 1).

6.4. Diskussion

Die Ergebnisse zeigen deutliche und zuverlässige Unterschiede bei der Bildauswahl in Abhängigkeit von den jeweiligen sprachlichen Strukturen. Diese Befunde stützen die theoretische Annahme, dass unterschiedliche sprachliche Strukturen jeweils spezifische Aspekte eines Ereignisses in den Vordergrund rücken.

Gemäß der ersten Hypothese wurde erwartet, dass bei einfachen aktiven transitiven Hauptsätzen die vollfarbige Bildvariante bevorzugt ausgewählt wird. Diese Erwartung beruht auf der theoretischen Annahme, dass ein finiter aktiver transitiver Hauptsatz das

kanonische Ereignismodell abbildet und die gesamte Handlungskette *Agens* → *Patiens* profiliert, wie in Kapitel 3.2.1 ausführlich erläutert. Subjekt und Objekt konkretisieren die zentralen Ereignisteilnehmer, während das Verb deren Interaktion spezifiziert. Sprecher und Leser entsprechen dem externen Betrachter im kanonischen Ereignismodell, der das gesamte Geschehen einschließlich der Teilnehmer und ihrer Interaktion sowie der *Settings* aus seiner Perspektive konstruiert (vgl. Langacker, 1987, 1991).

Die Ergebnisse bestätigen diese Annahme eindeutig. Die vollständig farbige Bildvariante (Variante a) wurde signifikant häufiger gewählt als die anderen Varianten. Dies legt nahe, dass beim *Construal* einfacher aktiver transitiver Hauptsätze die gesamte kodierte Handlungskette in den Vordergrund rückte, ohne dass dabei einzelne Komponenten hervorgehoben wurden.

Die zweite Hypothese wurde ebenfalls bestätigt. Bei Subjekt-Relativsätzen wurde überwiegend Bildvariante b) gewählt, bei der nur das Agens farblich hervorgehoben war. Dieses Ergebnis stimmt mit der theoretischen Annahme überein. Durch die Relativierung verändert sich nicht nur die syntaktische Form, sondern auch die zugrunde liegende semantische Struktur. Zwar werden in Subjekt-Relativsätzen alle zentralen Ereigniskomponenten explizit ausgedrückt, jedoch weisen diese Komponenten eine asymmetrische Salienz auf. Gemäß dem *Center-Periphery*-Schema und der Figur-Grund-Beziehung tritt das Kopfnomen in den Vordergrund, während die Handlung selbst und das *Patiens* im Relativsatz als Modifikatoren in der Peripherie stehen (vgl. Johnson, 1987; Lakoff, 1987; Langacker, 1987, 1991). Dabei erzeugt diese Struktur eine spezifische Perspektive auf das Ereignis, die sich wesentlich von der prototypischen Perspektive aktiver transitiver Hauptsätze unterscheidet. Aus den Ergebnissen lässt sich schließen, dass die Versuchspersonen beim *Construal* eines Subjekt-Relativsatzes tatsächlich primär das Kopfnomen – im vorliegenden Experiment das Agens – fokussierten.

Hinsichtlich der Nominalisierungen wurde angenommen, dass die Handlung und das

Patiens in der mentalen Repräsentation, die durch die Verarbeitung eines Satzes entsteht, salienter sind als das Agens. Diese Hypothese wurde auch durch die erhobenen Daten bestätigt. Die Versuchspersonen wählten überwiegend Bildvariante c), in der Instrument und Patiens farblich hervorgehoben wurden. Dieser Befund lässt sich dadurch erklären, dass Nominalisierungen im Gegensatz zu Haupt- und Relativsätzen nur Patiens und Handlung explizit ausdrücken, während das implizite Agens zahlreiche Interpretationsmöglichkeiten zulässt. Diese explizit spezifizierten Komponenten entsprechen genau den farblich markierten Elementen in Bild c). Hingegen lässt sich das schematisch bleibende Agens möglicherweise auch durch eine andere Figur repräsentieren als die, die im Bild dargestellt war. Daher galt Bildvariante c) als die am besten passende.

Zusammenfassend zeigen die Ergebnisse, dass die untersuchten sprachlichen Strukturen systematisch unterschiedliche Fokussierungen erzeugen und damit klare Präferenzen bei der Bildauswahl hervorrufen. Dabei ist allerdings zu beachten, dass die Versuchspersonen aufgrund der Instruktion metalinguistisches Wissen anwendeten und explizit sowie bewusst darüber nachdachten, welche Bildvariante am besten zu der jeweiligen sprachlichen Struktur passt. Somit liefert Experiment 1 vorläufig empirische Hinweise auf einen Zusammenhang zwischen sprachlicher Struktur und der Salienz spezifischer Ereigniskomponenten, jedoch noch nicht für eine unbewusste Echtzeitverarbeitung. Darüber hinaus dienen die Ergebnisse zur Normierung der Stimuli in den folgenden Experimenten. Sie zeigen, dass die inhaltlichen und visuellen Manipulationen erfolgreich waren. Es bleibt nun zu untersuchen, wie sich die Sprachstruktur auf die Aktivierung von Ereignisrepräsentationen während der Verarbeitung in Echtzeit auswirkt. Diese Fragestellung wurde in den drei folgenden Experimenten behandelt.

7. Experiment 2

7.1. Hypothesen

Die Ergebnisse von Experiment 1 zeigten, dass die Wahl einer bestimmten Bildvariante, bei der unterschiedliche Komponenten eines Ereignisses farblich hervorgehoben sind, maßgeblich von der verwendeten sprachlichen Struktur beeinflusst wird. Diese Befunde bestätigen im ersten Schritt, dass unterschiedliche sprachliche Strukturen jeweils unterschiedliche Aspekte eines Ereignisses hervorheben. Konkret wurde festgestellt, dass ein Subjekt-Relativsatz das Kopfnomen (in dieser Arbeit das Agens) betont, während bei Nominalisierungen die Handlung und das Patiens im Vordergrund stehen. Bei einfachen aktiven transitiven Hauptsätzen wurde hingegen bevorzugt eine Darstellung der gesamten Szene ohne spezifische farbliche Hervorhebung einzelner Komponenten gewählt.

Aufbauend auf diesen Ergebnissen wird in Experiment 2 mithilfe eines *Picture-Sentence-Matching-Tasks* weiter untersucht, inwieweit unterschiedliche sprachliche Strukturen einen Einfluss auf die mentale Ereignisrepräsentation in der Echtzeitverarbeitung haben. Zahlreiche Studien haben gezeigt, dass linguistische Inputs spezifische mentale Repräsentationen aktivieren können, die die visuelle Wahrnehmung oder Antizipation nachfolgender Stimuli beeinflussen können (z.B. Altmann & Kamide, 1999; Kamide et al., 2003; Stanfield & Zwaan, 2001; Zwaan et al., 2002). Insbesondere konnte mittels *Picture-Sentence-Matching-Task* gezeigt werden, dass perzeptuelle Informationen wie Orientierung, Form oder Bewegungsrichtung eines Objekts bei der Sprachverarbeitung mental aktiviert werden und zu schnelleren Reaktionen führen, wenn die Merkmale des nachfolgenden Bildes mit der zuvor sprachlich aktivierten Repräsentation übereinstimmen (*Match*) als bei einer Nichtübereinstimmung (*Mismatch*; vgl. Kapitel 5.2).

Vor diesem Hintergrund widmet sich das vorliegende Experiment der Frage, ob bei der linguistischen Verarbeitung auch konzeptuelle Informationen zur relativen Salienz

einzelner Ereigniskomponenten systematisch aktiviert werden und dadurch die mentalen Ereignisrepräsentationen beeinflussen, was wiederum zu unterschiedlichen *Construals* desselben Ereignisses führen kann.

Ausgehend von den in den Kapiteln 3 und 4 dargestellten theoretischen Annahmen zu den kognitiven Mechanismen der Sprachverarbeitung bezüglich der Aufmerksamkeitslenkung und Informationsselektion (vgl. Johnson, 1987; Lakoff, 1987; Langacker, 1987, 1991; Talmy, 2000) und unter der Annahme, dass diese Prozesse kognitive Realität in der Echtzeitverarbeitung haben, sowie gestützt auf die Ergebnisse aus Experiment 1 und auf frühere Studien mit dem *Picture-Sentence-Matching-Task*, wird angenommen, dass die drei untersuchten sprachlichen Strukturen unterschiedliche mentale Repräsentationen aktivieren und zu verschiedenen Aufmerksamkeitsmustern führen. Diese Unterschiede sollten sich in den Reaktionszeiten der Versuchspersonen zeigen, wenn die farbliche Hervorhebung von Komponenten in den präsentierten Bildern entweder mit den zuvor aktivierten mentalen Repräsentationen übereinstimmt (*Match*) oder davon abweicht (*Mismatch*). Hierbei werden folgende spezifische Hypothesen geprüft:

- 1) Insgesamt reagieren die Versuchspersonen in der *Match*-Bedingung schneller als in der *Mismatch*-Bedingung.
- 2) In der Relativsatz-Bedingung reagieren die Versuchspersonen schneller, wenn das Agens im präsentierten Bild farblich dargestellt ist (*Match*), als wenn das Patiens und die Handlung visuell hervorgehoben sind (*Mismatch*).
- 3) In der Nominalisierungs-Bedingung reagieren die Versuchspersonen schneller, wenn das Patiens und die Handlung im präsentierten Bild farblich markiert sind (*Match*), als wenn nur das Agens farblich markiert ist (*Mismatch*).

Da in der Hauptsatz-Bedingung die gesamte Szene farblich ohne spezifische Hervorhebung dargestellt wird, dient diese Bedingung im vorliegenden Experiment als eine Kontrollbedingung und wird aus der *Match-Mismatch*-Analyse ausgeschlossen.

7.2. Methodisches Vorgehen

7.2.1. Versuchspersonen

Insgesamt nahmen 56 Versuchspersonen an dem Experiment teil, davon 48 Frauen. Das Durchschnittsalter der Versuchspersonen betrug 21,66 Jahre (Bereich zwischen 17 und 32 Jahren). Alle Versuchspersonen waren Studierende des *Jiangnan Art Vocational College, China*, oder der *Westlake University, China*, mit Chinesisch als Muttersprache. Die Versuchspersonen waren normalsichtig oder trugen eine Brille, hatten ein durchschnittliches Hörvermögen und zeigten keine Anzeichen von neurologischen Störungen oder Leseschwierigkeiten. Keine der Versuchspersonen hatte an dem früheren Experiment teilgenommen. Alle Versuchspersonen nahmen freiwillig an dem Experiment teil und erhielten jeweils 50 *Yuan* als Belohnung.

7.2.2. Material

In Experiment 2 wurden die kritischen Stimuli aus Experiment 1 mit leichten Anpassungen übernommen. Alle Stimuli wurden zunächst anhand ihrer Kongruenz zwischen sprachlichen und visuellen Stimuli eingeteilt. Die Kongruenz-Bedingung umfasste Fälle, in denen der sprachliche Stimulus und das anschließende Bild dasselbe Ereignis darstellten, während in der Inkongruenz-Bedingung der linguistische Stimulus und das nachfolgende Bild unterschiedliche Ereignisse zeigten. Alle kritischen Stimuli gehörten zur Kongruenz-Bedingung, die je nach dem Grad der Übereinstimmung nochmals in *Match*- und *Mismatch*-Bedingungen unterteilt wurde. Die *Match*-Bedingung war dadurch charakterisiert, dass der farblich hervorgehobene Teil des Bildes mit dem sprachlich hervorgehobenen Ereignisaspekt übereinstimmte, während in der *Mismatch*-Bedingung hingegen keine solche Übereinstimmung bestand.

Das Experiment folgte einem 2 x 2 faktoriellen Design, bestehend aus den Faktoren Sprachstruktur (Subjekt-Relativsatz vs. Nominalisierung) und Bedingung (*Match* vs. *Mismatch*). Als kritische linguistische Stimuli wurden insgesamt 54 Stimulus-Paare verwendet, wobei jedes Paar aus einem Subjekt-Relativsatz und einer Nominalisierung

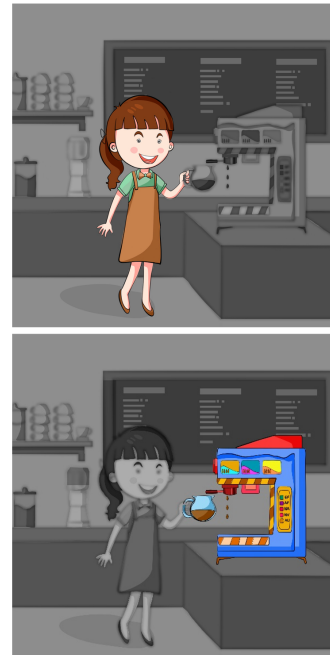
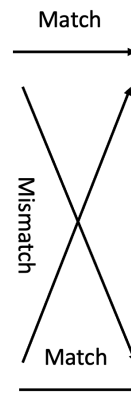
bestand, die Bezug auf dasselbe Ereignis nahmen und identisch mit jenen aus Experiment 1 waren.

Basierend auf den Ergebnissen von Experiment 1 wurde für die Subjekt-Relativsätze die Bildvariante, bei der das Subjekt (Agens) farblich hervorgehoben war, als *Match*-Bedingung definiert. Dieselbe Bildvariante wurde dementsprechend als *Mismatch*-Bedingung für Nominalisierungen verwendet. Im Gegensatz dazu wurde für Nominalisierungen die Bildvariante, bei der das Patiens und Instrument farblich hervorgehoben waren, als *Match*-Bedingung betrachtet, während sie als *Mismatch*-Bedingung für Subjekt-Relativsätze galt. Abbildung 7.1 zeigt ein Beispiel der kritischen Bedingungen.

Die visuellen Stimuli wurden weitgehend aus Experiment 1 übernommen, jedoch wurden kleine Anpassungen vorgenommen, um einen möglichen Positionsbias zu kontrollieren. Frühere Studien zeigen, dass Bilder mit von links nach rechts verlaufenden Handlungstransienten bevorzugt verarbeitet werden und in Satz-Bild-*Matching*-Aufgaben schneller erkannt werden als Bilder mit einer Transientenrichtung von rechts nach links (vgl. Chatterjee, 2001; Dobel et al., 2007). Um diesem Bias entgegenzuwirken, wurden sechs visuelle Stimuli horizontal gespiegelt, sodass Agens und Patiens gleich häufig links oder rechts in den Bildern positioniert waren.

a) Relativsatz:

那个 冲泡 咖啡 的 店员
 nàge chōngpào kāfēi de diànyuán
 DemPron. Zubereiten Kaffee de Barista
 „Die Barista, die Kaffee zubereitet“



b) Nominalisierung:

咖啡 的 冲泡
 kāfēi de chōngpào
 Kaffee de zubereiten
 „Die Kaffeezubereitung“

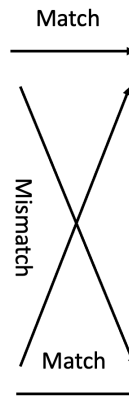
Abbildung 7.1: Ein Beispiel für die kritischen Bedingungen in Experiment 2.

Zur Counterbalance der kritischen Stimuli sowie zur Sicherstellung einer sinnvollen Aufgabenbearbeitung dienten Filler-Stimuli, die sowohl Kongruenz- als auch Inkongruenz-Bedingungen enthielten. Die Inkongruenz-Filler wurden analog zu den kritischen Stimuli gestaltet, indem Relativsätze und Nominalisierungen jeweils mit visuell übereinstimmenden (*Match*) oder nicht übereinstimmenden (*Mismatch*) Bildern kombiniert wurden. Sie unterschieden sich jedoch darin, dass linguistische und visuelle Stimuli jeweils unterschiedliche Szenen zeigten. Diese Inkongruenz-Filler ermöglichten es den Versuchspersonen, die Aufgabe klar mit „*Nein*“ zu beantworten, wodurch die Plausibilität der Aufgabe gewährleistet wurde. Darüber hinaus wurden zwei weitere Filler-Bedingungen aus den kritischen Hauptsätzen von Experiment 1 gebildet. In der Kongruenz-*Match*-Bedingung wurden diese Hauptsätze mit vollfarbig dargestellten, inhaltlich passenden Szenen kombiniert. Demgegenüber wurden in der Inkongruenz-*Match*-Bedingung dieselben Hauptsätze mit vollfarbigen, jedoch inhaltlich irrelevanten Szenen präsentiert. Ein Beispiel der Filler-Bedingungen ist in Abbildung 7.2 zu sehen.

Inkongruenz-Bedingungen:

a) Relativsatz

那个 烹饪 食物 的 主厨
 nàge pēngrèn shíwù de zhǔchú
 DemPron. kochen Speise de Chef
 „Der Chef, der die Speise kocht“



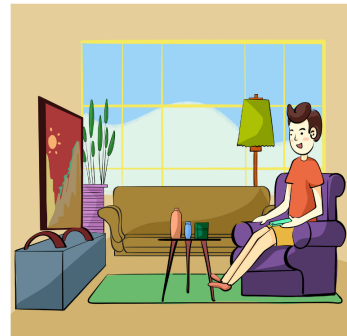
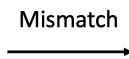
b) Nominalisierung

垃圾 的 回收
 lājī de huíshōu
 Abfall de Verwertung
 „Die Abfallverwertung“



c) Hauptsatz:

那个 厨师 切分 蔬菜
 nàge chùshī qiēfēn shūcài
 DemPron. Koch schneiden Gemüse
 „Der Koch schneidet das Gemüse“



Kongruenz-Bedingung:

d) Hauptsatz:

那个 店员 冲泡 咖啡
 nàge diànyuán chōngpào kāfēi
 DemPron. Barista zubereiten Kaffee
 „Die Barista bereitet Kaffee zu“

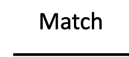


Abbildung 7.2: Beispiele für die Filler-Bedingungen in Experiment 2.

Vor Beginn des experimentellen Blocks gab es einen Übungsblock mit insgesamt 12 Trials, die alle vier experimentellen Bedingungen umfassten. Der Übungsblock diente dazu, die Versuchspersonen mit dem Ablauf des Experiments vertraut zu machen.

Die auditiven Stimuli wurden von einer chinesischen Muttersprachlerin mit konstanter Sprechgeschwindigkeit aufgenommen. Beim Einsprechen wurde darauf geachtet, dass

keine deutlichen Emotionen, Betonungen oder Pausen erkennbar waren, um mögliche Ablenkungen der Aufmerksamkeit der Versuchspersonen zu minimieren.

Für das Experiment wurden insgesamt neun Listen erstellt. Jede Liste enthielt jeweils 12 kritische Subjekt-Relativsätze und 12 Nominalisierungen, von denen die Hälfte jeweils in der *Match*- und die andere Hälfte in der *Mismatch*-Bedingung präsentiert wurde. Jede Struktur aus einem Stimulus-Paar kam nur einmal in einer Liste vor. Zusätzlich enthielt jede Liste 36 Filler-Stimuli, bestehend aus sechs Hauptsätzen in der Kongruenz-*Match*-Bedingung, sechs Hauptsätzen in der Inkongruenz-*Match*-Bedingung sowie jeweils zwölf Relativsätzen und Nominalisierungen, die je zur Hälfte in der Inkongruenz-*Match*- bzw. Inkongruenz-*Mismatch*-Bedingung erschienen. Alle Stimuli wurden in randomisierter Reihenfolge präsentiert.

7.2.3. Durchführung

Das Experiment wurde mithilfe der Online-Experiment-Plattform *PCIBEX* (<https://farm.pcibex.net/>) durchgeführt. Die Versuchspersonen wurden randomisiert auf die Listen verteilt. Sie führten das Experiment eigenständig auf ihrem eigenen Laptop entweder auf dem Universitätscampus oder in einer ruhigen Umgebung durch. Vor Beginn des Experiments wurden die Versuchspersonen gebeten, eine möglichst störungsfreie Umgebung während des Experiments sicherzustellen. Mobiltelefone sollten während der Durchführung stumm geschaltet und keine anderen Webseiten aufgerufen werden.

Vor dem experimentellen Teil wurden demografische Daten der Versuchspersonen erhoben, darunter Geschlecht, Alter, Geburtsort und eine individuelle Experiment-ID. Anschließend wurde den Versuchspersonen die ausführliche Instruktion zur Aufgabe in chinesischer Sprache präsentiert. Die Instruktion lautete wie folgt:

Das Experiment dauert etwa 15 Minuten. Bei jedem Trial hören Sie zunächst ein Audio, gefolgt von einem Bild. Ihre Aufgabe besteht darin, so schnell und präzise wie möglich zu beurteilen, ob das Bild zu dem zuvor gehörten Audio passt. Drücken Sie

dazu die Taste „F“ für „passt“ oder „J“ für „passt nicht“. Bewerten Sie bitte dann auf einer 7-stufigen Skala, wie gut das Bild zu dem Audio passt (1 = „gar nicht“ bis 7 = „sehr gut“).

Die ergänzende Einschätzung auf der 7-stufigen Skala diente zusätzlich als Kontrollmaßnahme, um sicherzustellen, dass die Versuchspersonen sowohl die linguistischen als auch die visuellen Stimuli tatsächlich verarbeiteten. Diese Skalen-Einschätzung wurde jedoch nicht in die statistische Auswertung einbezogen.

Um die Versuchspersonen optimal auf die Aufgabe vorzubereiten, wurde der genaue Ablauf mithilfe von Beispielen visuell dargestellt. Dabei wurde den Versuchspersonen erklärt, dass alle Antworten ausschließlich über die Tastatur einzugeben wären. Die Tasten wurden entsprechend der Skala programmiert: „A“ = 1, „S“ = 2, „D“ = 3, „F“ = 4, „J“ = 5, „K“ = 6 und „L“ = 7. Die Versuchspersonen wurden gebeten, ihre Finger dauerhaft auf folgenden Tasten zu platzieren: Zeigefinger auf „F“ und „J“, Mittelfinger auf „D“ und „K“, Ringfinger auf „S“ und „L“ sowie den linken kleinen Finger auf „A“. Auf diese Weise war es den Versuchspersonen möglich, alle Antworten zu geben, ohne ihre Finger zu bewegen, was eine schnelle und zuverlässige Reaktionsweise ermöglichte. Weiterhin wurden sie angewiesen, das „+“-Zeichen in der Mitte des Bildschirms zu fixieren, sobald es erschien.

Nach der Instruktion durchliefen die Versuchspersonen zunächst einen Übungsblock mit insgesamt 12 Trials. Im Anschluss daran bestand die Möglichkeit, verbleibende Fragen zu klären oder Unklarheiten zu beseitigen. Erst nach dieser Phase begann der eigentliche Experimentalblock.

Zu Beginn jedes Trials wurde ein Fixationskreuz in der Mitte des Bildschirms angezeigt. Nach 500 ms wurde der auditive Stimulus abgespielt. Am Ende des auditiven Inputs wurde ein Bild für 500 ms in der Mitte des Bildschirms präsentiert. Die kurze Darbietungsdauer des Bildes wurde gewählt, da unter Zeitdruck vor allem die salienten Komponenten verarbeitet werden. Die Versuchspersonen sollten schnell

und präzise entscheiden, ob das Bild zu dem linguistischen Stimulus passte, und die entsprechende Taste „F“ oder „J“ drücken. Dabei wurde für alle Kongruenz-Bedingungen eine „Ja“-Antwort (Taste „F“) und für alle Inkongruenz-Bedingungen eine „Nein“-Antwort (Taste „J“) erwartet. Die Reaktionszeit wurde hierbei als Zeitintervall zwischen dem Bild-Onset und dem Tastendruck gemessen.

Nach jedem visuellen Stimulus erschien die Frage auf Chinesisch „*Wie gut passt das gezeigte Bild zu dem Audio?*“ mit einer 7-stufigen Skala auf dem Bildschirm. Die Versuchspersonen konnten einen Wert auswählen. Für diese Aufgabe gab es kein Zeitlimit. Nach der Auswahl eines Skalenwertes bestätigten die Versuchspersonen ihre Wahl durch Drücken der Leertaste. Daraufhin erfolgte automatisch der Übergang zum nächsten Trial. Eine Rückkehr zu vorherigen Trials war nicht möglich. Das Experiment endete automatisch, sobald alle Trials abgeschlossen waren.

Ablauf eines Beispielstimulus

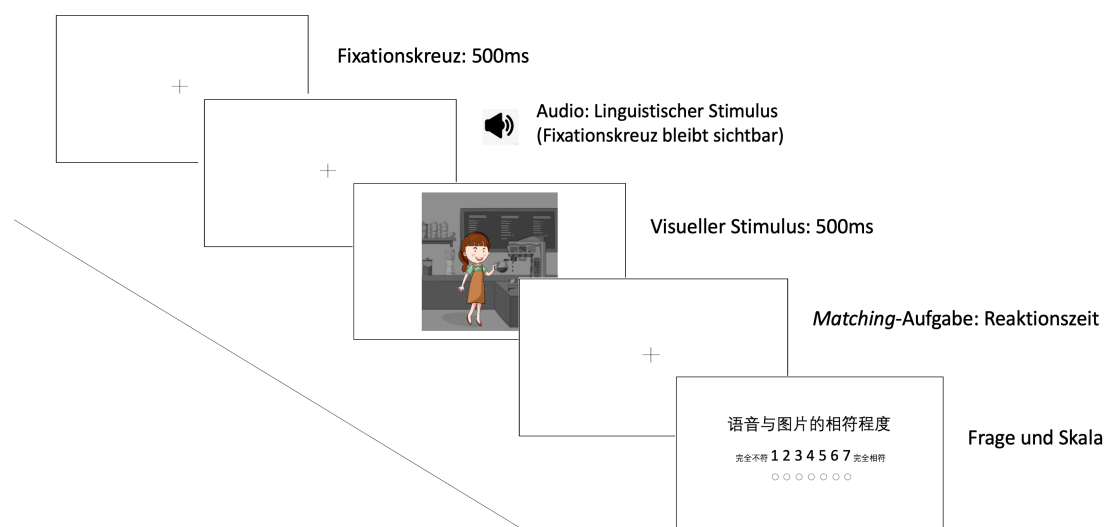


Abbildung 7.3: Beispielablauf eines Trials (Experiment 2).

7.2.4. Datenkodierung und -analyse

Insgesamt wurden 11 Versuchspersonen von der Datenauswertung ausgeschlossen. Bei 7 von ihnen traten technische Probleme mit der Online-Plattform auf. Dabei wurden entweder nicht alle Daten aufgezeichnet oder die Ergebnisse nicht erfolgreich übermittelt. Bei den übrigen 4 Versuchspersonen lag die Fehlerquote jeweils bei über

20% aller Stimuli bzw. der kritischen Stimuli, sodass auch deren Daten aufgrund ihrer geringen Leistung nicht in die Auswertung einbezogen wurden.

Da bei den kritischen Bedingungen die Antwort „Ja“ erwartet wurde, wurden alle falschen Antworten (102 Beobachtungen, 9,4%) aus der Auswertung entfernt. Zudem wurde angenommen, dass extrem langsame oder extrem schnelle Reaktionszeiten möglicherweise nicht das Ergebnis eines gültigen Verarbeitungsprozesses widerspiegeln. Daher wurden die Beobachtungen, die mehr als ± 3 Standardabweichungen von der durchschnittlichen Reaktionszeit aller Versuchspersonen abwichen, von der Auswertung ausgeschlossen (12 Beobachtungen, 1,2%).

Zur statistischen Überprüfung, ob die Reaktionszeiten durch die unterschiedlichen Übereinstimmungsgrade beeinflusst werden, wurde ein Bayes'sches Modell verwendet. Dies ermöglicht eine direkte Bewertung der Hypothese hinsichtlich ihrer Plausibilität und erlaubt die flexible Definition hierarchischer Modelle innerhalb des Bayes'schen Rahmens. Dabei wird eine Überanpassung der Daten vermieden, ohne wertvolle Informationen über die Gruppenvariabilität zu verlieren (vgl. O'Hagan & Forster, 2004).

Vor der statistischen Modellierung wurden die beiden zweistufigen Faktoren *Übereinstimmungsbedingung* (*Match* vs. *Mismatch*) und *Sprachstruktur* (*Relativsatz* vs. *Nominalisierung*) als kategoriale Prädiktoren spezifiziert. Anschließend wurden für beide Faktoren Sum-Kontraste (*contr.sum*) definiert.

Für die Analyse wurde ein Bayes'sches lineares gemischtes Modell mit der *Reaktionszeit* als abhängiger Variable geschätzt. Als *Fixed* Effekte gingen die Haupteffekte von *Übereinstimmungsbedingung* und *Sprachstruktur* sowie deren Interaktion in das Modell ein. Zur Kontrolle von Zufallseffekten wurden für *Versuchspersonen* und *Items* zufällige Intercepts und zufällige Steigungen für alle *Fixed* Effekte spezifiziert, wobei die Kovarianzen zwischen den Zufallseffekten aus Gründen der Modellstabilität auf null gesetzt wurden. Formal entsprach die *Random*-Effekte-

Struktur dem Term $(1 + \textit{Bedingung} + \textit{Sprachstruktur} + \textit{Bedingung} * \textit{Sprachstruktur} || \textit{Versuchsperson}) + (1 + \textit{Bedingung} + \textit{Sprachstruktur} + \textit{Bedingung} * \textit{Sprachstruktur} || \textit{Item})$).

Als Fehlerverteilung wurde eine Student-t-Verteilung gewählt (*family* = *student*), um mögliche Ausreißer in den Reaktionszeiten robuster abzufangen. Die Priorverteilungen wurden explizit wie folgt festgelegt: Für den Interzept wurde ein Normal-Prior mit einem Mittelwert von 1200 ms und einer Standardabweichung von 300 ms angenommen (*normal*(1200, 300)), was einem plausiblen Erwartungsbereich für mittlere Reaktionszeiten in dieser Aufgabe entspricht. Für alle *Fixed-Effects*-Koeffizienten (*b*) wurde ein schwach informativer Normal-Prior mit Mittelwert 0 und Standardabweichung 150 gesetzt (*normal*(0, 150)). Die Standardabweichungen der Zufallseffekte erhielten Student-t-Priors mit 3 Freiheitsgraden, Mittelwert 0 und Skala 150 (*student_t*(3, 0, 150)), die Residualstreuung (*sigma*) einen Student-t-Prior mit 3 Freiheitsgraden, Mittelwert 0 und Skala 200 (*student_t*(3, 0, 200)). Das Modell wurde mit der Funktion *brm()* geschätzt. Es wurden vier Ketten mit jeweils 4000 Iterationen verwendet, wobei die ersten 2000 Iterationen pro Kette als Aufwärmphase (*Warm-up*) verworfen wurden. Zur Reduktion divergierender Übergänge und zur Verbesserung der Konvergenz wurde der Parameter *adapt_delta* auf 0.99 und *max_treedepth* auf 12 gesetzt. Die Ergebnisse wurden in Form der Posterior-Verteilung inklusive 95%-Konfidenzintervall (CIs) angegeben.

Um die Hypothese 1 weiter zu überprüfen, dass die Versuchspersonen bei der *Match*-Bedingung schneller reagieren als bei der *Mismatch*-Bedingung, wurde ein Bayes'scher Hypothesentest mit der Funktion *hypothesis()* durchgeführt. Auf Basis der im Modell verwendeten Kontrastkodierung wurde ein gerichteter Kontrast spezifiziert, der die Differenz der mittleren Reaktionszeiten zwischen *Mismatch*- und *Match*-Bedingung abbildet. Ein positiver Wert dieses Kontrasts entspricht kürzeren Reaktionszeiten in der *Match*- gegenüber der *Mismatch*-Bedingung. Darüber hinaus wurden weitere gerichtete Hypothesentests definiert, um den Effekt der Übereinstimmungsbedingung getrennt innerhalb der beiden Sprachstrukturen (Relativsatz und Nominalisierung) zu

untersuchen. Auf diese Weise lässt sich prüfen, ob der erwartete Vorteil der *Match*- gegenüber der *Mismatch*-Bedingung in beiden Sprachstrukturen im gleichen Ausmaß auftritt oder ob sich strukturspezifische Unterschiede ergeben.

7.3. Ergebnisse

Die Ergebnisse des Bayes'schen Modells zeigten, dass die Versuchspersonen in der *Match*-Bedingung ($M = 1230,60$ ms, $SD = 544,66$) schneller reagierten als in der *Mismatch*-Bedingung ($M = 1292,79$ ms, $SD = 587,14$). Der Haupteffekt der Bedingung ergab einen geschätzten Unterschied von $\beta = 23,78$, 95% CI [0,21, 48,04].

Im Gegensatz dazu lagen die Effekte der Sprachstruktur ($\beta = 19,24$, 95% CI [-7,87, 47,86]) sowie der Interaktion zwischen Bedingung und Sprachstruktur ($\beta = -14,31$, 95% CI [-38,24, 8,57]) mit Konfidenzintervallen, die den Nullwert einschließen, sodass hier keine eindeutigen Schlüsse hinsichtlich signifikanter Unterschiede getroffen werden konnten. Dies legt nahe, dass der Unterschied in den Reaktionszeiten hauptsächlich durch die Bedingung (*Mismatch* vs. *Match*) erklärt wurde, während die Effekte der Sprachstruktur und deren Interaktion mit der Bedingung in diesem Modell weniger robust waren.

Der zusätzliche Bayes'sche Hypothesentest bestätigte den Bedingungseffekt. Es wurde ein geschätzter Effekt von 47,55 ms (Est. Error = 24,32, 90% CI [7,51, 88,14]) zwischen der *Mismatch*- und *Match*-Bedingung festgestellt. Die Evidenzrate war 40,67, und die Posterior-Wahrscheinlichkeit war 0,98. Damit ist es mit hoher posteriorer Wahrscheinlichkeit plausibel, dass die *Match*-Bedingung kürzere Reaktionszeiten aufweist als die *Mismatch*-Bedingung.

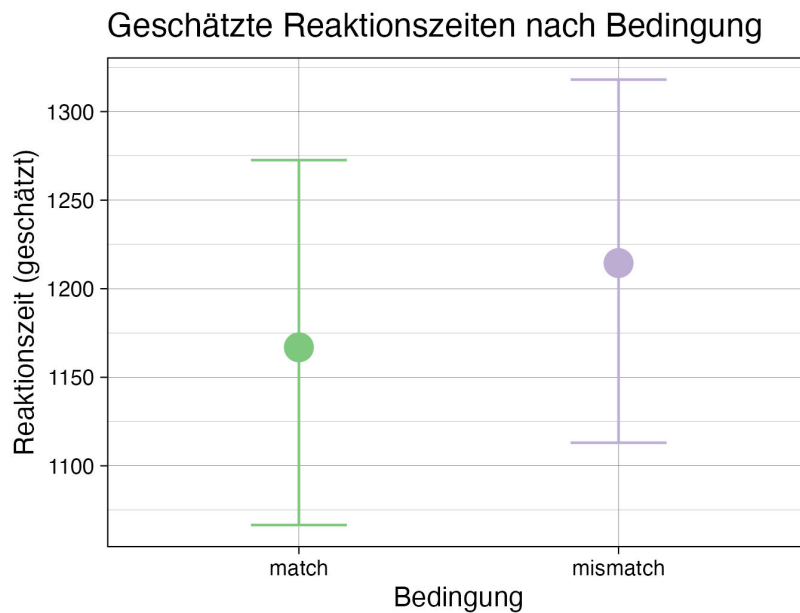


Abbildung 7.4: Geschätzte Reaktionszeiten nach Übereinstimmungsbedingung (Experiment 2).

Zur weiteren Differenzierung der Effekte wurden separate Auswertungen innerhalb der Sprachstrukturen vorgenommen. Für die Struktur „Relativsatz“ gab es geringe Unterschiede in Bezug auf die Reaktionszeiten zwischen der *Match*-Bedingung ($M = 1273,75$ ms, $SD = 602,66$) und der *Mismatch*-Bedingung ($M = 1287,94$ ms, $SD = 602,07$). Der Bayes'sche Vergleich zwischen *Mismatch*- und *Match*-Bedingung ergab einen geschätzten Unterschied von $18,93$ ms (90% CI $[-39,15, 75,48]$, Posterior-Wahrscheinlichkeit = $0,71$). Da das CI-Intervall den Nullwert umfasst, konnte hier kein zuverlässiger Unterschied zwischen den beiden Bedingungen festgestellt werden.

Im Gegensatz dazu zeigte die Analyse der Struktur „Nominalisierung“ einen eindeutigen *Mismatch*-Effekt: Die Reaktionszeit betrug in der *Match*-Bedingung im Mittel $1189,89$ ms ($SD = 481,30$) und in der *Mismatch*-Bedingung $1297,71$ ms ($SD = 572,78$). Der geschätzte Unterschied lag bei $\beta = 76,17$ ms (90% CI $[20,55, 134,22]$, Posterior-Wahrscheinlichkeit = $0,99$).

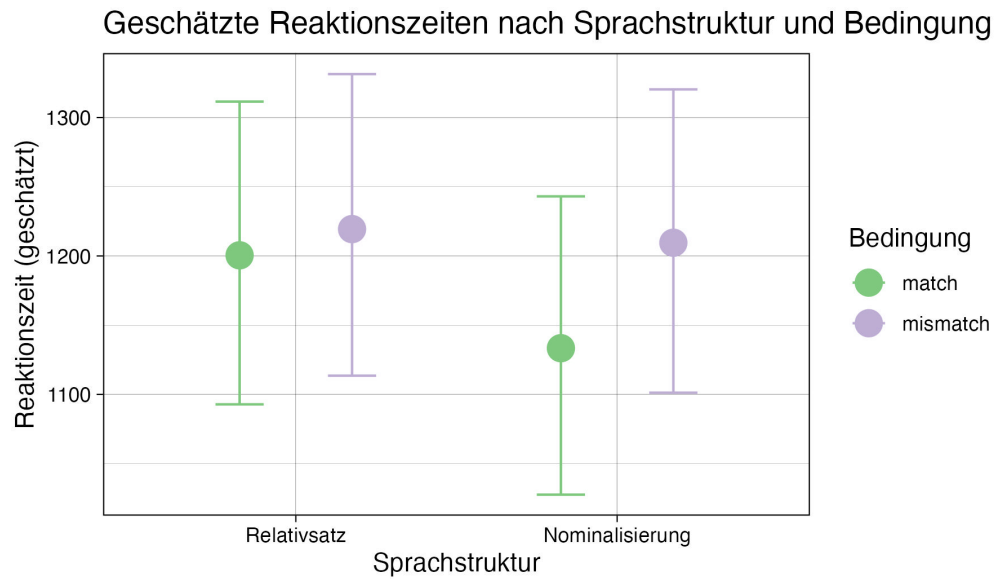


Abbildung 7.5: Geschätzte Reaktionszeiten nach Sprachstruktur und Übereinstimmungsbedingung (Experiment 2).

Sprachstruktur	Match: M (SD)	Mismatch: M (SD)	Geschätzte Differenz: Median [90% CI]
Gesamt	1230,60 (544,66)	1292,79 (587,14)	47,55 [7,51, 88,14]
Relativsatz	1273,75 (602,66)	1287,94 (602,07)	18,93 [-39,15, 75,48]
Nominalisierung	1189,89 (481,30)	1297,71 (572,78)	76,17 [20,55, 134,22]

Tabelle 7.1: Deskriptive Statistiken und Ergebnisse des Bayes'schen Modells zu den Reaktionszeiten (Experiment 2).

7.4. Diskussion

In dem vorliegenden Experiment wurde vor allem der Frage nachgegangen, ob und wie die verschiedenen sprachlichen Strukturen die Ereignisrepräsentation in der Echtzeitverarbeitung beeinflussen. Gemäß den theoretischen Überlegungen, die in Kapitel 3 dargestellt wurden, führen die unterschiedlichen sprachlichen Kodierungen von Ereignissen durch Subjekt-Relativsätze und Nominalisierungen zu jeweils spezifischen *Construals*. Während bei Subjekt-Relativsätzen das Kopfnomen besonders hervorgehoben wird, stehen bei Nominalisierungen die explizit genannten Komponenten im Fokus (vgl. Langacker, 1987, 1991; Talmy, 2000, 2008). In diesem Experiment spiegelten die visuellen Stimuli in den *Match*- und *Mismatch*-Bedingungen diese unterschiedlichen Hervorhebungen mithilfe einer gezielten farblichen Gestaltung

wider.

Frühere Studien haben bereits gezeigt, dass mentale Ereignisrepräsentationen während der Sprachverarbeitung automatisch aktiviert werden und nachfolgende visuelle Verarbeitungen beeinflussen sowie antizipieren können (z.B. Altmann & Kamide, 1999; Kamide et al., 2003; Zwaan et al., 2002). Solche mentalen Ereignisrepräsentationen können eine Vielzahl von Informationen integrieren, einschließlich der semantischen Strukturen sowie der spezifischen Perspektive auf ein Ereignis. Beispielsweise können Aktiv- und Passivsätze in vorherigen Studien zu unterschiedlichen visuellen Präferenzen auf die Ereignisteilnehmer führen. Während nach Aktivsätzen bevorzugt das Agens betrachtet wurde, rückte nach Passivsätzen eher das Patiens in den Fokus (vgl. Sauppe & Flecken, 2021). Diese Befunde belegen, dass sprachliche Strukturen die Perspektive auf ein Ereignis steuern und die Aufmerksamkeit in der nachfolgenden visuellen Szenenverarbeitung lenken können.

Auf Basis der theoretischen und empirischen Grundlagen wurde im vorliegenden Experiment angenommen, dass die mental aktivierten Ereignisrepräsentationen die zugrunde liegende semantische Struktur der Sprachstrukturen widerspiegeln. Daraus sollte folgen, dass sich die Reaktionszeiten beim *Sentence-Picture-Matching* unterscheiden, je nachdem, ob das präsentierte Bild mit der durch die jeweilige Sprachstruktur aktivierten Repräsentation übereinstimmt (*Match*) oder nicht (*Mismatch*).

Die Ergebnisse bestätigten diese Annahme. Insgesamt wurden in der *Match*-Bedingung schnellere Reaktionszeiten beobachtet als in der *Mismatch*-Bedingung. Wurde beispielsweise ein linguistischer Stimulus in Form einer Nominalisierung wie „*Kaffeezubereitung*“ präsentiert, reagierten die Versuchspersonen schneller, wenn im anschließenden Bild nur das Patiens („*Kaffee*“) und das Instrument („*die Kaffeemaschine*“) farblich hervorgehoben waren, während das Agens („*Barista*“) und der Hintergrund in Graustufen erschienen. Waren hingegen im Bild nur das Agens farblich hervorgehoben und der Rest in Graustufen dargestellt, verlängerten sich die

Reaktionszeiten signifikant.

Diese Ergebnisse lassen sich dahingehend interpretieren, dass beim Sprachverstehen die semantischen Rollen und die strukturelle Information eines Ereignisses automatisch aktiviert werden. Die sprachlich induzierte Hervorhebung bestimmter Ereigniskomponenten spiegelt sich in der Aktivierung entsprechender mentaler Ereignisrepräsentationen wider. Beim Hören einer bestimmten sprachlichen Struktur generierten die Versuchspersonen die mentale Repräsentation, in der spezifische Ereigniskomponenten (z.B. das Agens bei Subjekt- Relativsätzen oder die Handlung und das Patiens bei Nominalisierungen) mental besonders salient gemacht wurden. Wenn anschließend ein Bild präsentiert wurde, rief der visuelle Input ebenfalls die entsprechende Repräsentation hervor. Wenn die visuelle Hervorhebung mit der zuvor sprachlich aktivierten Fokussierung übereinstimmte (*Match*-Bedingung), verkürzte sich die Reaktionszeit zur Erkennung des Ereignisses, da die relevanten Repräsentationen bereits aktiviert waren und die Integration erleichtert wurde. Wenn hingegen eine Diskrepanz zwischen der sprachlich aktivierten Repräsentation und der visuellen Hervorhebung bestand (*Mismatch*-Bedingung), erforderte dies einen erhöhten kognitiven Verarbeitungsaufwand, um die Informationen anzupassen. Dieser zusätzliche Integrationsprozess führte zu verlängerten Reaktionszeiten.

Es ist jedoch wichtig anzumerken, dass hiermit nicht behauptet wird, dass die mentale Repräsentation, die bei der Verarbeitung der linguistischen Stimuli entsteht, identisch mit der Repräsentation ist, die bei der Bildbetrachtung generiert wird. Vielmehr wird lediglich angenommen, dass der Grad der Übereinstimmung zwischen diesen beiden Repräsentationen in der *Match*-Bedingung höher ist als in der *Mismatch*-Bedingung.

Des Weiteren wurde der Einfluss der Übereinstimmung (*Match* vs. *Mismatch*) getrennt für Subjekt-Relativsätze und Nominalisierungen analysiert. Dabei zeigte sich ein signifikanter Unterschied in den Reaktionszeiten nur bei den Nominalisierungen, während für die Subjekt-Relativsätze kein signifikanter Effekt festgestellt wurde.

Bei Nominalisierungen wies die *Match*-Bedingung kürzere Reaktionszeiten als die *Mismatch*-Bedingung auf. Dieses Ergebnis lässt sich vor allem dadurch erklären, dass in Nominalisierungen nur das Patiens und die Handlung explizit genannt werden, während das Agens implizit bleibt und daher nur schematisch repräsentiert wird. Aufgrund dieser impliziten Darstellung ist das Agens mental relativ offen und weniger spezifisch festgelegt. Beispielsweise muss das Ereignis „*Kaffeezubereitung*“ nicht zwingend von einer Barista ausgeführt werden, sondern könnte theoretisch ebenso von einer Haushaltshilfe oder einem Landwirt vorgenommen werden.

Beim Hören der linguistischen Stimuli in Form von Nominalisierungen generierten die Versuchspersonen demnach mentale Repräsentationen, die sich in erster Linie aus Ereignisschemawissen speisten. In diesem Fall wurden die explizit spezifizierten Komponenten, nämlich das Patiens und die Handlung, deutlich stärker aktiviert als das nur schematisch repräsentierte Agens. Wenn anschließend ein visueller Stimulus präsentiert wurde, in dem die visuell hervorgehobenen Komponenten mit den sprachlich fokussierten Komponenten übereinstimmten (*Match*-Bedingung), konnten die aus der sprachlichen und der visuellen Verarbeitung entstehenden Repräsentationen leichter integriert werden, sodass die Reaktionszeiten kürzer ausfielen. Hingegen konnte die Diskrepanz zwischen dem linguistischen und visuellen Stimulus (*Mismatch*-Bedingung) zu zusätzlichem kognitivem Verarbeitungsaufwand führen.

Die Abwesenheit eines eindeutigen Effekts bei Subjekt-Relativsätzen könnte auf deren syntaktische Struktur zurückzuführen sein. In chinesischen Relativsätzen steht das Kopfnomen am Ende des Satzes. Die Hörer müssen daher bis zum Ende auf das Kopfnomen warten, um die zuvor bereitgestellten Informationen in das Kopfnomen zu integrieren. Während dieses Prozesses könnte die Information aus dem Relativsatz im Arbeitsgedächtnis aktiv bleiben und wird bei der Verarbeitung sowie Integration jedes Modifikators und schließlich des Kopfnomens reaktiviert (vgl. Karimi et al., 2019; Lewis & Vasishth, 2005; Lewis et al., 2006). Dies kann die Salienz der Komponenten im Relativsyntagma erhöhen.

Im Unterschied zu den Nominalisierungen werden in Subjekt-Relativsätzen alle drei zentralen Komponenten eines kausativen Ereignisses, und zwar Agens, Patiens sowie Handlung, explizit sprachlich konkretisiert. Dadurch werden im Relativsatz alle diese Komponenten mit vergleichbarer *Specificity* mental repräsentiert und aktiviert. In beiden visuellen Bedingungen wurde jedoch nur ein Teil dieser Komponenten farblich hervorgehoben (in der *Match*-Bedingung das Agens und in der *Mismatch*-Bedingung das Patiens und die Handlung). Da alle Komponenten bereits mental verfügbar waren, könnte die visuelle Hervorhebung nur einzelner Elemente die Verarbeitung nicht wesentlich beeinflusst haben.

Darüber hinaus lassen sich diese Befunde durch die informationsstrukturellen Eigenschaften von Relativsätzen erklären. Typischerweise wird das Kopfnomen (hier das Agens) im Sprachgebrauch als bereits gegebene Information betrachtet, während der Relativsatz neue, ergänzende Informationen vermittelt (vgl. Kapitel 5.1). Aufgrund dieser wiederholten sprachlichen Erfahrung könnten im Chinesischen das Patiens und die Handlung im Relativsatz als Informationsträger stark aktiviert werden und mehr kognitive Ressourcen in Anspruch nehmen, wodurch sie Aufmerksamkeit auf sich ziehen.

Es stellte sich die Frage, ob der beobachtete Bedingungseffekt durch eine bewusste Verarbeitungsstrategie der Versuchspersonen erklärt werden könnte. Da die Versuchspersonen wussten, dass nach jedem linguistischen Stimulus ein Bild erscheinen würde, wäre es theoretisch möglich gewesen, dass sie gezielt und bewusst ein mentales Bild vorbereitet hätten. Diese Erklärung erscheint jedoch unwahrscheinlich.

In diesem Experiment stellten die visuellen Stimuli in *Match*- und *Mismatch*-Bedingungen dieselbe Szene dar und unterschieden sich lediglich in der farblichen Markierung bestimmter Ereigniskomponenten. Da die explizite Aufgabe für die Probanden nicht darin bestand, sich auf die Hervorhebung bestimmter Komponenten zu konzentrieren, sondern lediglich die Kongruenz zwischen Sprache und Bild zu

beurteilen, wäre in beiden Bedingungen ein nahezu identisches mentales Bild des betreffenden Ereignisses zu erwarten. Somit könnte eine bewusste strategische Vorbereitung nicht erklären, warum unmittelbar nach der Verarbeitung der Sprachstruktur automatisch bestimmte semantische Rollen oder Ereigniselemente mental hervorgehoben wurden. Zudem wussten die Versuchspersonen im Voraus nicht, welche spezifische Hervorhebung im nachfolgenden Bild tatsächlich gezeigt werden würde, weshalb eine strategische Vorbereitung hinsichtlich der spezifischen Hervorhebung nicht möglich war.

Darüber hinaus wurden Inkongruenz-Bedingungen als Counterbalance-Bedingungen eingesetzt, bei denen die sprachlichen und visuellen Stimuli zwei völlig unterschiedliche Ereignisse darstellten. Unter diesen Bedingungen wäre eine bewusste, strategische Vorab-Erstellung detaillierter mentaler Repräsentationen ineffizient gewesen, da diese überhaupt nicht mit den präsentierten Bildern übereinstimmten. Zusätzlich wurde jedes Bild nur für 500 ms präsentiert, was eine bewusste und gezielte Verarbeitung der sprachlich und visuell betonten Details erheblich erschwerte.

Wenn die Versuchspersonen tatsächlich eine bewusste, strategische Vorgehensweise verfolgt hätten, wäre außerdem davon auszugehen, dass sie diese in der Nachbefragung des Experiments erwähnt hätten. Da in den Interviews jedoch keine entsprechenden Hinweise auf eine bewusste Strategie vorlagen, spricht auch dies deutlich gegen diese Erklärung.

Es lässt sich daher schlussfolgern, dass der beobachtete Bedingungseffekt mit hoher Wahrscheinlichkeit nicht auf eine bewusste Verarbeitungsstrategie zurückzuführen ist. Vielmehr kam er durch die automatische Aktivierung unterschiedlicher konzeptueller Repräsentationen bei der Sprachverarbeitung zustande.

Zusammenfassend lässt sich festhalten, dass unterschiedliche sprachliche Strukturen die mentale Ereignisrepräsentation bei der Echtzeitverarbeitung beeinflussen. Die Versuchspersonen waren demnach nicht nur in der Lage, durch bewusste Überlegungen

zu erkennen, dass verschiedene sprachliche Strukturen spezifische Ereigniskomponenten fokussieren, wie bereits in Experiment 1 gezeigt wurde. Vielmehr spiegelt sich diese Hervorhebung unmittelbar und automatisch in den mentalen Repräsentationen bei der Sprachverarbeitung wider. Besonders deutlich zeigt sich dieser Effekt bei Nominalisierungen, die spezifisch das Patiens und die Handlung in den Fokus rücken und somit ein anderes *Construal* des Ereignisses erzeugen als Relativsätze.

8. Experiment 3

In Experiment 2 wurde mittels der *Sentence-Picture-Matching*-Aufgabe untersucht, wie die unterschiedlichen Sprachstrukturen die Aktivierung mentaler Ereignisrepräsentationen in Echtzeit beeinflussen. Die Ergebnisse zeigten, dass die Versuchspersonen auf die Bilder in der *Match*-Bedingung schneller reagierten als auf jene in der *Mismatch*-Bedingung.

Dennoch bleiben nach Experiment 2 einige Fragen offen. Erstens liefern Reaktionszeiten zwar wertvolle Informationen über den kognitiven Zeitaufwand während der Ereignisverarbeitung, sie bieten jedoch keine detaillierten Informationen darüber, wie genau die Aufmerksamkeit auf einzelne Ereigniskomponenten in Echtzeit verteilt wird und wie die Verarbeitung dieser Komponenten unmittelbar erfolgt. Zweitens wurden in Experiment 2 lediglich zwei Sprachstrukturen, nämlich Subjekt-Relativsätze und Nominalisierungen, als kritische Bedingungen berücksichtigt, während Hauptsätze nicht einbezogen wurden. Da einfache aktive transitive Hauptsätze die prototypische Struktur zur Kodierung eines kausativen Ereignisses sind, ist es sinnvoll zu überprüfen, ob und wie diese Sprachstruktur die Ereigniskonzeptualisierung beeinflusst, insbesondere im Vergleich zu den beiden anderen untersuchten Strukturen. Um diese Lücke zu schließen und ein umfassendes Bild der Einflüsse aller drei Sprachstrukturen auf die Ereignisverarbeitung zu erhalten, wird in Experiment 3 zusätzlich der Hauptsatz als dritte kritische Struktur einbezogen.

Um diese offenen Fragen zu klären und die Prozesse der Aufmerksamkeitslenkung sowie der Aktivierung sprachlich induzierter mentaler Ereignisrepräsentationen genauer und differenzierter zu untersuchen, kommt in Experiment 3 eine Variante des *Visual-World-Paradigmas* zum Einsatz. Basierend auf dem sogenannten *eye-mind-link* spiegeln Augenbewegungen die zugrunde liegenden kognitiven Prozesse wider (vgl. Carter & Luke, 2020; König et al., 2016). Sie dienen daher als Indikator für die mentale Ereignisverarbeitung. Mittels Eyetracking-Messungen lassen sich detaillierte Einblicke in die Verteilung visueller Aufmerksamkeit auf einzelne Ereigniskomponenten in Echtzeit gewinnen. Dadurch wird direkt aufgezeigt, wie unterschiedliche sprachliche Strukturen die mentale Konzeptualisierung eines Ereignisses beeinflussen. Anders als in der klassischen Verwendung des *Visual-World-Paradigmas* werden den Versuchspersonen jedoch auditive Stimuli vor der Präsentation des visuellen Stimulus dargeboten und nicht gleichzeitig. Diese Herangehensweise soll zeigen, inwiefern eine durch sprachlichen Input evozierte Repräsentation die visuelle Informationsaufnahme beeinflusst.

8.1. Hypothesen

Im vorliegenden Experiment wird mittels Eyetracking untersucht, wie die drei sprachlichen Strukturen die Salienz bestimmter Ereigniskomponenten modulieren und wie sich diese Modulation in den Blickbewegungen widerspiegelt. Von besonderem Interesse ist, welche Ereigniskomponenten zuerst fokussiert werden und wie sich die Aufmerksamkeit auf diese Komponenten im Verlauf der Bildbetrachtung verteilt. Ausgehend von Experiment 2, in dem insbesondere bei Nominalisierungen eine verstärkte Hervorhebung von Patiens und Handlung nachgewiesen wurde, sollen diese Befunde nun anhand der Blickbewegungen weiter präzisiert werden.

Wie in den Kapiteln 3 und 4 ausgeführt, variiert die Salienz der Ereigniskomponenten in Abhängigkeit von der verwendeten sprachlichen Struktur. In Subjekt-Relativsätzen steht das Agens als Kopfnomen im Vordergrund, während Handlung und Patiens im Relativsyntagma peripher bleiben. Bei Nominalisierungen werden hingegen Handlung

und Patiens hervorgehoben, während das Agens nur implizit bleibt. In Hauptsätzen wird die gesamte kodierte Handlungskette profiliert, wobei alle zentralen Ereigniskomponenten, nämlich Agens, Handlung und Patiens, vergleichsweise ausgewogen konzeptualisiert werden.

Auf Grundlage dieser Annahmen wird die zentrale Hypothese aufgestellt, dass sprachliche Strukturen die visuelle Wahrnehmung und Ereignisverarbeitung gezielt beeinflussen, indem sie bestimmte Komponenten des betreffenden Ereignisses hervorheben und damit deren Salienz bei der visuellen Verarbeitung erhöhen. Diese sprachlich induzierte Fokussierung sollte sich in spezifischen Blickbewegungsmustern niederschlagen.

Zur Überprüfung dieser Hypothese wurden drei Messgrößen herangezogen: (1) die Position der ersten Fixation, (2) die Latenz der Erstfixation auf jeder AOI als Kontrollmaß für die erste Messgröße, sowie (3) die Gazedauer auf jeder AOI, die die kumulierte visuelle Aufmerksamkeit während der Stimulusdarbietung erfasst.

(1) Hypothesen zur Position der ersten Fixation

Frühere Studien haben gezeigt, dass sprachliche Strukturen die initiale visuelle Aufmerksamkeit gezielt auf bestimmte Ereigniskomponenten lenken und damit die Position der ersten Fixation beeinflussen (vgl. Gerwien & Flecken, 2016; Griffin & Bock, 2000). Menschen neigen dazu, jene Komponenten eines Ereignisses zuerst zu fixieren, die durch die sprachliche Struktur als besonders wichtig hervorgehoben werden (vgl. Divjak et al., 2020; Sauppe & Flecken, 2021). Da die erste Fixation in den frühesten Phasen der visuellen Verarbeitung erfolgt, unterliegt sie selten bewussten Strategien. Daher dient die erste Fixation als frühes Indiz für eine unmittelbare Aufmerksamkeitssteuerung bezüglich der Konzeptualisierung der visuell dargestellten Szene, hier in Abhängigkeit von der sprachlichen Kodierung, und liefert Informationen darüber, welche Ereigniskomponente priorisiert aufgesucht wird.

Auf Grundlage der Ergebnisse aus Experiment 1 und 2, die eine Hervorhebung des

Patiens und der Handlung in Nominalisierungen zeigten, wird die Hypothese aufgestellt, dass bei Nominalisierungen die erste Fixation mit höherer Wahrscheinlichkeit auf Patiens und Handlung fällt als bei Haupt- und Relativsätzen. Demgegenüber wird erwartet, dass bei Haupt- und insbesondere bei Relativsätzen die erste Fixation häufiger auf das Agens gerichtet ist als bei Nominalisierungen. Beim Vergleich von Haupt- und Relativsätzen wird angenommen, dass Relativsätze eine höhere Wahrscheinlichkeit für eine Erstfixation auf das Agens und eine entsprechend geringere Wahrscheinlichkeit für eine Erstfixation auf Patiens und Handlung aufweisen als Hauptsätze.

(2) Hypothesen zur Fixationslatenz

Die Fixationslatenz bezeichnet die Zeitspanne vom Beginn eines visuellen Stimulus bis zur ersten Fixation auf eine definierte AOI. Sie erlaubt differenzierte Rückschlüsse darauf, wie früh eine bestimmte Ereigniskomponente visuell erfasst wird und dient somit als ergänzender Indikator für die initiale Aufmerksamkeitslenkung im Verarbeitungsprozess.

Es wird die Hypothese aufgestellt, dass linguistisch hervorgehobene Komponenten früher fixiert werden. Konkret wird erwartet, dass bei Nominalisierungen die Patiens-/Handlungsregion früher fixiert wird als bei Haupt- und Relativsätzen. Umgekehrt wird erwartet, dass die Latenz der ersten Fixation auf das Agens bei Nominalisierungen länger ausfällt als bei Haupt- und Relativsätzen. Im Vergleich von Haupt- und Relativsätzen wird angenommen, dass das Agens in Relativsätzen aufgrund seiner Prominenz schneller fixiert wird als in Hauptsätzen.

(3) Hypothesen zur Gazedauer

Die gesamte Gazedauer auf einem AOI erfasst die kumulierte visuelle Verweildauer auf diesem Bereich während der Darbietung eines Stimulus. Dieses Maß reflektiert die visuelle und semantische Salienz sowie die Informativität der jeweiligen Ereigniskomponenten (vgl. Holmqvist et al., 2011).

Die Hypothese bezüglich der Gazedauer besagt, dass Ereigniskomponenten, die durch die jeweilige sprachliche Struktur hervorgehoben werden, insgesamt länger visuell fixiert werden, da sie mental als besonders relevant verarbeitet werden. Entsprechend wird erwartet, dass bei Nominalisierungen die Gazedauer für Patiens und Handlung länger ausfällt als bei Haupt- und Relativsätzen. Umgekehrt sollte die Gazedauer für das implizite Agens bei Nominalisierungen kürzer sein als bei Haupt- und Relativsätzen, da das Agens hier nicht als besonders relevant gilt. Im Vergleich zwischen Haupt- und Relativsätzen wird angenommen, dass die Gazedauer auf das Agens bei Relativsätzen länger und auf das Patiens sowie die Handlung kürzer ausfällt als bei Hauptsätzen.

8.2. Methodisches Vorgehen

8.2.1. Versuchspersonen

Insgesamt nahmen 46 Versuchspersonen an dem Experiment teil, davon 27 Frauen. Alle Versuchspersonen waren Studierende der Universität Heidelberg mit Chinesisch als Muttersprache. Das Alter der Versuchspersonen variierte zwischen 19 und 33 Jahren (Durchschnittsalter: 26,3 Jahre). Alle Versuchspersonen berichteten entweder von normaler Sehkraft oder trugen eine Brille bzw. Kontaktlinsen, um ihre Sehfähigkeit auf Normalniveau zu korrigieren. Zudem verfügten sie über ein normales Hörvermögen und wiesen keine neurologischen oder Lesestörungen auf. Keine der Versuchspersonen hatte an den anderen vorliegenden Experimenten teilgenommen. Die Teilnahme war freiwillig, und als kleines Dankeschön erhielten alle Versuchspersonen am Ende des Experiments Süßigkeiten.

8.2.2. Apparatur

Das Eyetracking-Experiment wurde mithilfe des *SR Research EyeLink* Systems durchgeführt, das sowohl die Hardware als auch die Software zur Verfügung stellte. Die Hardware bestand aus einem Host-PC, einem Experimental-PC und dem *EyeLink 1000 Plus* mit der Hochgeschwindigkeitskamera in Form einer Tischhalterung. Auf dem Host-PC lief ein speziell für *Eye-Tracking* entwickeltes Echtzeit-Betriebssystem, das

von der Experimentleiterin bedient wurde. Der Eyetracker, der mit dem Host-PC verbunden war, befand sich unter dem experimentellen Display. Aus der Sicht der Versuchsperson befand sich die Kamera des Eyetrackers etwa auf Höhe der unteren Bildschirmkante. Im Experiment wurde eine 35mm Linse eingesetzt. Der Eyetracker war in der Lage, binokulare Augenbewegungen mit bis zu 2000 Bildern pro Sekunde aufzuzeichnen, was sich als ideal für die Erfassung von bildschirmbasierten Eye-Tracking-Daten erwies.

Um eine stabile Kopfposition der Versuchspersonen während des Experiments zu gewährleisten und damit die Genauigkeit der Blickbewegungsaufzeichnung zu optimieren, wurde eine *SR Research* Kopfstütze verwendet. Dieses Gerät bot eine stabile Kinn- und Stirnstütze und wurde vor dem Eyetracker an der Tischkante montiert. Der Abstand zwischen der Kopfstütze und dem experimentellen Display betrug etwa 92 cm, der Abstand zum Eyetracker ca. 60 cm. Zur Eingabe der Antworten wurde eine Button-Box verwendet.

Die Software *SR Research Experiment Builder* und *EyeLink Data Viewer* wurden auf dem Experiment-PC installiert. Mithilfe des *Experiment Builders* wurde das Experiment programmiert. Die visuellen Stimuli wurden auf einem 24-Zoll-Display mit einer Auflösung von 1920×1080 Pixeln präsentiert. Die Augenbewegungen wurden mit einer Abtastrate von 1000 Hz aufgezeichnet. Nach Abschluss des Experiments wurden die aufgezeichneten Blickdaten mit dem *EyeLink Data Viewer* visualisiert und weiterverarbeitet, wodurch eine detaillierte Analyse der Blickmuster und Augenbewegungen ermöglicht wurde.

8.2.3. Material

Die visuellen Stimuli entsprachen im Wesentlichen denen aus Experiment 1 und 2, wurden jedoch geringfügig angepasst. Der Hauptunterschied bestand darin, dass in diesem Experiment für jeden linguistischen Stimulus nur ein Bild präsentiert wurde, das mithilfe der Software *IrfanView* in Graustufen konvertiert wurde. Diese Anpassung

diente dazu, potentielle Einflüsse der Farbgebung auf das Blickverhalten zu reduzieren. Alle Bilder wurden einheitlich in der Größe von 21.5×21.5 cm bzw. 720×720 Pixel präsentiert (siehe Abbildung 8.1).

Die kritischen linguistischen Stimuli waren identisch mit denen aus Experiment 1 (siehe Beispiel 5-1). In allen drei Sprachstruktur-Bedingungen wurde dasselbe Bild präsentiert. Diese Vorgehensweise ermöglichte es, Unterschiede in den Blickbewegungen in Abhängigkeit von der jeweiligen sprachlichen Struktur zu beobachten und direkt miteinander zu vergleichen. Durch die Verwendung eines identischen Bildes wurden potentielle visuelle Störeffekte minimiert.

Wie bereits erwähnt, wurden die Bilder gezielt so gestaltet, dass das Instrument räumlich möglichst klar vom Oberkörper und Kopf des Agens getrennt war, ohne die Plausibilität der Szene zu beeinträchtigen. Dadurch konnten Fixationen in den Bereichen von Agens und Instrument zuverlässig unterschieden werden. Um eine strategische Verarbeitung zu vermeiden, die durch eine konstante Platzierung der Bilder auf derselben Position des Bildschirms begünstigt werden könnte, wurden die Bilder jeweils zur Hälfte auf der linken und zur Hälfte auf der rechten Seite des Bildschirms präsentiert. Dieses Design sollte zudem verhindern, dass die erste Fixation durch das zentrale Fixationskreuz beeinflusst wird, und die Versuchspersonen dazu anregen, aktiv zu fixieren. Die Bildposition wurde randomisiert, sodass nicht vorhersehbar war, auf welcher Seite das nächste Bild erscheinen würde.

Im vorliegenden Experiment wurden alle Agens durch belebte menschliche Figuren repräsentiert, während nahezu alle Patiens unbelebte Objekte waren. Zahlreiche Studien belegen, dass es eine visuelle Präferenz sowohl für belebte menschliche Figuren gegenüber unbelebten Objekten als auch für Agens gegenüber Patiens gibt (vgl. Dobel et al., 2007; Gerwien & Flecken, 2016; Isasi-Isasmendi et al., 2023; Wilson et al., 2011). Aufgrund dieser Salienzasymmetrie hinsichtlich semantischer Rollen und Belebtheit wurde antizipiert, dass die Versuchspersonen unabhängig von der gehörten Sprachstruktur zunächst bevorzugt das Agens fixieren würden. Um diesen Bias zu

verringern, wurden die Bereiche für Patiens und Instrument stets näher an der Mitte des Bildschirms platziert, während sich das Agens weiter von der Mitte entfernt befand, wenn beide einer Links-Rechts-Anordnung folgten.



Abbildung 8.1: Ein Beispiel für die kritischen visuellen Stimuli in Experiment 3.

Die Filler-Stimuli waren den kritischen Stimuli in ihrer Struktur ähnlich. Sie umfassten ebenfalls Hauptsätze, Relativsätze und Nominalisierungen. Der wesentliche Unterschied bestand darin, dass bei den Filler-Stimuli das gezeigte Bild und der zuvor dargebotene auditive Input zwei unterschiedliche Ereignisse darstellten. Die Versuchspersonen sollten daher bei den Filler-Stimuli die Frage, ob das Bild zum auditiven Input passt, mit „*Nein*“ beantworten.

Dem experimentellen Block ging ein Übungsblock voraus, der aus insgesamt sieben Trials bestand und sowohl kritische Bedingungen als auch Filler-Bedingungen beinhaltete. Ziel dieser Übungsphase war es, die Versuchspersonen mit dem Aufbau und Ablauf des Experiments vertraut zu machen und sicherzustellen, dass sie die Aufgabe klar verstanden hatten, um eine maximale Konzentration während der Datenerhebung zu gewährleisten.

Insgesamt wurden sechs experimentelle Listen erstellt. In allen Listen waren die Übungs- und Filler-Trials identisch. Die kritischen sprachlichen Strukturen wurden

gleichmäßig auf die sechs Listen verteilt, so dass jede Versuchsperson jedes Bild nur einmal zu sehen bekam. Alle Stimuli wurden in randomisierter Reihenfolge präsentiert.

8.2.4. Durchführung

Das Experiment wurde im *Eye-Tracking-Labor* des Instituts für Deutsch als Fremdsprachenphilologie der Universität Heidelberg durchgeführt.

Alle Versuchspersonen betraten das Labor einzeln und führten das Experiment einzeln durch. Während des Experiments blieb eine Experimentleiterin im Labor, hielt jedoch genügend Abstand und saß am Host-PC, um mögliche Interferenzen und Störungen zu minimieren. Die Versuchspersonen nahmen auf einem Stuhl vor dem experimentellen Computer Platz und füllten zunächst eine Tabelle mit Angaben zu Alter, Geschlecht, Geburtsort und dominanter Hand aus. Danach wurden sie gebeten, ihre Stirn und ihr Kinn auf der Stütze abzulegen. Da die Versuchspersonen während des gesamten Experiments möglichst unbeweglich bleiben sollten, hatten sie vor Beginn des Experiments die Möglichkeit, die Kinnstütze, die Sitzhöhe und den Abstand zum Tisch individuell einzustellen, um einen optimalen Komfort zu gewährleisten. Die Höhe der Stirnstütze war fest vorgegeben, um sicherzustellen, dass die Blickrichtung der Versuchspersonen immer zentral auf den Bildschirm gerichtet war und der Abstand zwischen Augen und Bildschirm ungefähr 90 cm betrug.

Anschließend erhielten die Versuchspersonen schriftliche Instruktionen auf dem Bildschirm, in denen ihnen die experimentelle Aufgabe erklärt wurde. Die Aufgabe bestand darin, zunächst einen Audioclip zu hören, danach ein Bild zu betrachten und schließlich mittels einer Button-Box zu entscheiden, ob das Bild zum gehörten Audioclip passte. Wenn ja, sollte der grüne Knopf rechts gedrückt werden; wenn nein, der rote Knopf links.

Nachdem die Versuchspersonen die Aufgabe verstanden hatten, setzten sie Kopfhörer auf und führten eine 9-Punkt-Kalibrierung durch. Dabei fixierten sie zunächst einen kleinen weißen Punkt innerhalb eines schwarzen Punktes in der Mitte des Bildschirms,

der sich anschließend zufällig über den Bildschirm bewegte. Die Versuchspersonen wurden angewiesen, dem weißen Punkt mit ihren Augen zu folgen, ohne dabei den Kopf zu bewegen. Dadurch wurde die Präzision des Eyetrackers gewährleistet. Die Qualität der Kalibrierung wurde durch eine anschließende Validierung überprüft. Die Kalibrierung und Validierung wurden gegebenenfalls mehrfach durchgeführt, bis eine ausreichende Genauigkeit erreicht war. In der Regel genügte eine einzige Kalibrierung. Danach wurde eine Driftkorrektur durchgeführt, bei der die Versuchspersonen den Fixationspunkt in der Mitte des Bildschirms fixierten und die Experimentleiterin die Fixation manuell bestätigte. Die Driftkorrektur wurde vor jedem einzelnen Trial wiederholt, um eine zuverlässige Messung der Augenbewegungen sicherzustellen.

Im Anschluss begann der Übungsblock, der dazu diente, die Versuchspersonen mit der experimentellen Aufgabe vertraut zu machen. Wenn die Versuchspersonen keine Fragen oder Unklarheiten zu den Trials hatten, startete der experimentelle Block, in dem alle Trials in randomisierter Reihenfolge präsentiert wurden. Vor jedem Trial erschien zunächst ein Fixationskreuz in der Mitte des Bildschirms, das von einer für die Versuchspersonen nicht sichtbaren Ellipse (Durchmesser 150 Pixel) umgeben war. Sobald die Fixation für mindestens 500 ms innerhalb dieser Ellipse erfolgte, wurde der Audioclip ausgelöst. Auf diese Weise wurde sichergestellt, dass sich der Blick zu Beginn jedes Trials in der Mitte des Bildschirms befand und kein Positionsbias auftrat.

Während der Präsentation des Audioclips sollten die Versuchspersonen das Fixationskreuz weiterhin fixieren. Nach Abschluss des auditiven Inputs verschwand das Fixationskreuz, und daraufhin wurde das Bild für 1500 ms gezeigt. Danach erschien auf Chinesisch die Frage: „Passt das Bild zu dem auditiven Input?“. Auf diese Weise wurde gewährleistet, dass die Versuchspersonen die Stimuli tatsächlich verarbeiteten. Nach ihrer Entscheidung mittels Buttondruck wurde automatisch der nächste Trial gestartet. Es gab keine Pausen bis zum Ende des Experiments. Das gesamte Experiment dauerte ungefähr 15 Minuten.

Alle Ergebnisse wurden automatisch von der *SR Eyelink 1000+* Software gespeichert.

Ablauf eines Beispielstimulus

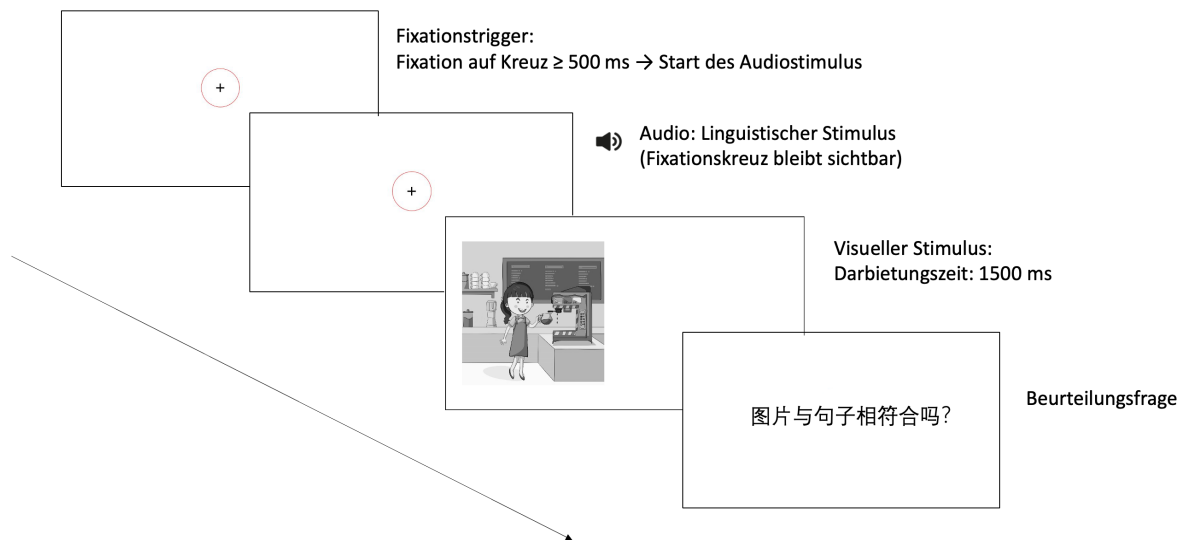


Abbildung 8.2: Beispielablauf eines Trials (Experiment 3).

8.2.5. Datenkodierung

Das vorliegende Experiment geht in erster Linie der Frage nach, ob die Position der Erstfixation, die Fixationslatenz und die Gazedauer durch unterschiedliche sprachliche Strukturen beeinflusst werden. Der Fokus lag dabei insbesondere auf den Ereigniskomponenten Agens, Patiens und Handlung. Dementsprechend wurden mit Hilfe der Software *Data Viewer* für jedes kritische Bild zwei AOIs definiert, die im Wesentlichen den farblich hervorgehobenen Bereichen der Bildvarianten b) und c) aus Experiment 1 entsprachen, jedoch kleinere Anpassungen erhielten.

AOI 1 bezog sich auf die Agens-Region. In allen kritischen Bildern hielt das Agens ein Instrument in der Hand, das die Handlung repräsentierte. Wenn das Agens ein Instrument nur mit einer Hand hielt und die andere Hand relativ weit entfernt vom Patiens war, wurde AOI 1 als die Region definiert, die den Kopf, Körper und die weiter entfernte Hand des Agens umfasste.

AOI 2 bezeichnete die Patiens-/Handlungsregion. Diese AOI umfasste das Instrument, das Patiens sowie die Hände des Agens, sofern diese in direktem Kontakt mit dem Instrument standen. Wenn das Instrument oder das Patiens mit dem Körper des Agens

überlappten, wurde dieser Bereich der AOI 2 zugeordnet. AOI 1 und AOI 2 überlappten sich somit nicht. Ein Beispiel zur Veranschaulichung der AOIs ist in Abbildung 8.3 dargestellt.

Um die Natürlichkeit der Szenen zu gewährleisten, wurde die Größe der dargestellten Komponenten in den visuellen Stimuli nicht kontrolliert. Insgesamt war die Patiens-/Handlungsregion (AOI 2) in der Regel größer als die Agens-Region (AOI 1).

Das Zeitintervall vom Bild-Onset bis zum Bild-Offset wurde als „*interest period*“ definiert und bildete den relevanten Analysezeitraum. Es betrug insgesamt 1500 ms.

Die statistische Auswertung der *Eye-Tracking*-Daten erfolgte mit der Software R (Version 4.2.1) unter Verwendung der Pakete „tidyverse“ (Version 1.3.2; Wickham et al., 2019), „brms“ (Version 2.19.0; Bürkner, 2017), „emmeans“ (Version 1.8.7; Lenth, 2023) und „rstan“ (Version 2.32.6; Stan Development Team, 2025).

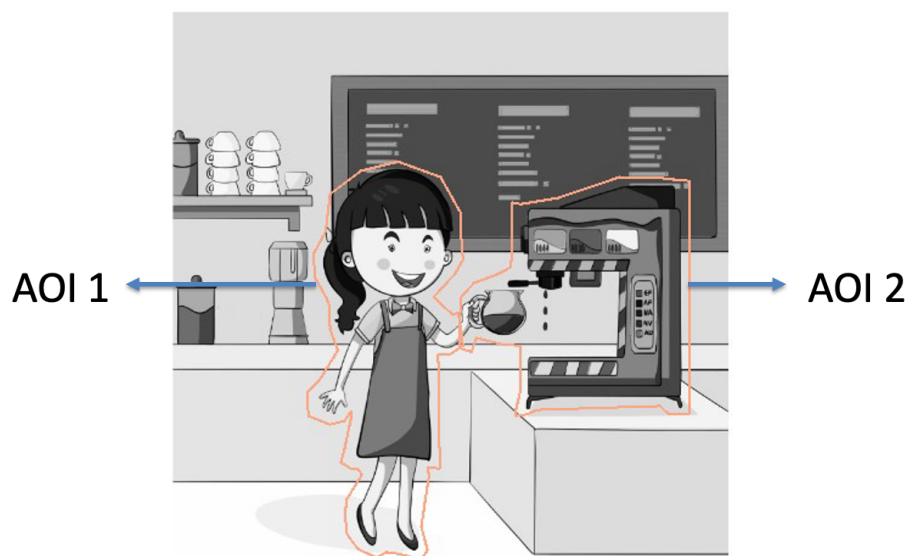


Abbildung 8.3: Ein Beispiel für AOIs der kritischen Bedingung (Experiment 3).

8.2.6. Datenanalyse

Insgesamt wurden 4 Versuchspersonen aus der Datenauswertung ausgeschlossen. Bei einer Versuchsperson trat ein technisches Problem mit dem experimentellen System auf, sodass mehr als 40% der kritischen Daten nicht korrekt aufgezeichnet wurden. Drei

weitere Versuchspersonen mussten aufgrund schlechter Eyetracking-Qualität ausgeschlossen werden. Zudem wurden alle Beobachtungen ausgeschlossen, bei denen eine falsche Antwort auf die Beurteilungsfrage gegeben wurde (3,09%). Darüber hinaus wurden Beobachtungen mit schlechter Datenqualität entfernt (0,79%). Da in der vorliegenden Untersuchung insbesondere die Fixationen auf den beiden kritischen AOIs von Interesse waren, wurden die Fixationen außerhalb dieser AOIs nicht in die statistische Auswertung einbezogen.

Zur statistischen Überprüfung der Forschungsfragen wurden für jede der drei Messgrößen separate Bayes'sche Modelle spezifiziert. Ziel dieser Modellierung war es, die Plausibilität der Hypothesen zu bewerten und die posterioren Verteilungen für die interessierenden Effekte zu analysieren. Dabei wurden folgende Parameter berücksichtigt: (1) das 95% CI zur Quantifizierung der Unsicherheit, (2) *probability of direction* (pd) als Maß für die Richtungssicherheit eines Effekts sowie (3) der Anteil der posterioren Verteilung innerhalb der ROPE, um die praktische Relevanz der Effekte zu beurteilen.

Datenanalyse zur ersten Fixation

Um zu untersuchen, ob sich die Position der ersten Fixation auf den beiden kritischen AOIs zwischen den drei Sprachstrukturen unterscheidet, wurde ein Bayes'sches logistisches gemischtes Modell verwendet. Dieses Modell basierte auf einer Bernoulli-Verteilung mit einer Logit-Link-Funktion und modellierte die Wahrscheinlichkeit, dass die erste Fixation auf AOI 2 im Gegensatz zu AOI 1 landete. Vor der Modellierung wurden zunächst alle ersten Fixationen mit einer Onset-Latenz von weniger als 100 ms (1,56%) ausgeschlossen, da sich diese aufgrund der Fixationsphase vor den visuellen Stimuli noch in dem Fixationskreuz befanden. Anschließend wurde für jede Beobachtung eine binäre abhängige Variable erstellt, die den Wert „1“ annimmt, wenn die erste Fixation auf AOI 2 erfolgte, und ansonsten den Wert „0“.

Die *Sprachstruktur* wurde als unabhängige Variable in das Modell aufgenommen. Sie

wurde mittels der Funktion *factor()* als Faktor und mittels der Funktion *levels()* in drei Stufen definiert, nämlich „Hauptsatz“, „Relativsatz“ und „Nominalisierung“, wobei der Hauptsatz als Referenzkategorie diente. Als abhängige Variable wurde die Wahrscheinlichkeit modelliert, dass die erste Fixation auf AOI 2 (Patien-/Handlungsregion) fiel. Die Daten wurden mithilfe einer Bernoulli-Verteilung mit Logit-Linkfunktion analysiert.

Um interindividuelle Unterschiede zwischen den Versuchspersonen sowie Stimuluspezifische Unterschiede adäquat zu berücksichtigen, wurden sowohl zufällige Intercepts als auch zufällige Slopes für den Faktor Sprachstruktur auf den Ebenen „Versuchsperson“ und „Stimulus“ ins Modell aufgenommen. Dies entspricht der Modellformulierung $(1 + \text{Sprachstruktur} \mid \text{Versuchsperson})$ und $(1 + \text{Sprachstruktur} \mid \text{Stimulus})$. Durch diese Spezifikation wird der Annahme gerecht geworden, dass die Effekte der Sprachstruktur nicht bei allen Versuchspersonen und Stimuli gleich ausfallen, sondern systematisch variieren können. Dadurch wird die Generalisierbarkeit der Ergebnisse über die konkreten Stichproben hinaus verbessert (vgl. Barr et al., 2013).

Mögliche informative Priors für Intercept, *Fixed* Effekte und *Random* Effekte wurden mithilfe der Funktion *get_prior()* ermittelt. Die Priors für Intercept und Koeffizienten wurden je nach erwarteter Größenordnung angepasst, d.h. *Normal(0, 5)* für *b*, *Gamma(1, 1)* für *Residual-σ*. Diese Priors ermöglichen eine ausreichende Flexibilität des Modells, während gleichzeitig der Einfluss extremer Ausreißer reduziert und die Konvergenz der Modellparameter unterstützt wird. Das endgültige Modell wurde mit der Funktion *brm()* geschätzt, wobei vier MCMC-Ketten mit jeweils 5000 Iterationen einschließlich einer Aufwärmphase von 500 Iterationen durchgeführt wurden. Der Parameter *adapt_delta* wurde auf 0,95 gesetzt, um divergierende Übergänge zu minimieren und eine zuverlässige Konvergenz der Ketten zu gewährleisten. Für die Reproduzierbarkeit wurde ein Seed-Wert gesetzt.

Zur weiteren Überprüfung, ob die beobachteten Unterschiede zwischen den linguistischen Bedingungen statistisch signifikant sind, wurden *Post-hoc*-Vergleiche

zwischen den sprachlichen Strukturen mittels *Estimated Marginal Means* (EMMs) und deren paarweisen Vergleiche durchgeführt. Diese Analysen erfolgten mithilfe der Funktionen *emmeans()* und *pairs()*.

Datenanalyse zur Fixationslatenz

Zur Kontrolle der Ergebnisse der Analyse der Erstfixation wurde zusätzlich die Fixationslatenz für beide kritischen AOIs mittels einer Bayes'schen *Mixed-Effects*-Analyse untersucht. Die Fixationslatenz wurde dabei als die Zeitspanne definiert, die vom Onset des visuellen Stimulus bis zur ersten Fixation auf AOI 1 bzw. AOI 2 vergeht.

Zur Analyse der Fixationslatenzen wurden die Daten zunächst getrennt für die beiden kritischen AOIs (AOI 1 und AOI 2) gefiltert. In Anlehnung an die Analyse der Erstfixation wurden nur jene Beobachtungen berücksichtigt, bei denen die Latenz der ersten Fixation auf die jeweilige AOI mindestens 100 ms betrug. Dadurch wurden 0,20% der Datenpunkte für AOI 1 und 1,33% für AOI 2 ausgeschlossen. Zudem wurde eine obere Schranke von 1500 ms für die Fixationslatenz festgelegt.

Für AOI 1 und AOI 2 wurden zwei separate statistische Modelle spezifiziert. In beiden Modellen wurde die Fixationslatenz als lognormalverteilte, abhängige Variable mit einer Identitäts-Link-Funktion modelliert, um der positiven Schiefe der Datenverteilung adäquat Rechnung zu tragen. Als fester Effekt wurde jeweils die Sprachstruktur in das Modell aufgenommen, die mit drei Stufen („Hauptsatz“, „Nominalisierung“, „Relativsatz“) mittels *factor()* und *levels()* spezifiziert wurde. Dabei wurde „Hauptsatz“ als Referenzkategorie definiert. Die Definition der Zufallseffekte entsprach dabei jener im Modell zur Analyse der Erstfixation. Es wurden zufällige Intercepts sowie zufällige Slopes für den Effekt der Sprachstruktur modelliert.

Im Modell für AOI 1 wurde der Intercept mithilfe eines Student-t-Priors mit 3 Freiheitsgraden, einem Mittelwert von 5,9 und einer Standardabweichung von 2,5 spezifiziert (class = „Intercept“). Für das Modell AOI 2 wurde ein Student-t-Prior mit einem Mittelwert von 5,5 verwendet. Für die fixen Effekte der Sprachstruktur (class =

„b“) wurden in beiden Modellen Normalverteilungen mit Mittelwert 0 und Standardabweichung 5 angesetzt. Die Residualvarianz (class = „sigma“) wurde mit einem Student-t-Prior mit 3 Freiheitsgraden, Mittelwert 0 und Standardabweichung 2 definiert, wobei die untere Schranke (lb = 0) sicherstellt, dass nur positive Werte erlaubt sind. Für die Standardabweichungen der zufälligen Effekte (class = „sd“) wurde ebenfalls ein Student-t-Prior mit 3 Freiheitsgraden, Mittelwert 0 und Standardabweichung 2,5 (lb = 0) verwendet, um eine angemessene Variabilität der Zufallseffekte zu gewährleisten.

Das Modell wurde mithilfe der Funktion *brm()* geschätzt. Dabei wurden vier Ketten mit je 5000 Iterationen verwendet, von denen jeweils die ersten 500 Iterationen als Aufwärmphase dienten. Der Parameter „adapt_delta“ wurde auf 0.95 gesetzt, um eine gute Konvergenz der Modelle sicherzustellen. Die Option *sample_prior* wurde auf *TRUE* gesetzt, um die Priorverteilung explizit zu berücksichtigen. Darüber hinaus wurde zur Sicherstellung der Reproduzierbarkeit ein Seed-Wert verwendet.

Zur weiteren detaillierten Untersuchung, ob die beobachteten Unterschiede zwischen den Bedingungen statistisch zuverlässig sind, wurden *Post-hoc*-Vergleiche mittels EMMs durchgeführt. Dabei wurden paarweise Vergleiche zwischen den einzelnen Sprachstrukturen durch *emmeans()* und *pairs()* berechnet.

Datenanalyse zur gesamten Gazedauer

Zur Analyse der Effekte unterschiedlicher Sprachstrukturen auf die Verweildauer wurde für beide AOIs jeweils ein Bayes'sches hierarchisches Modell spezifiziert. Zunächst wurden für jeden AOI nur Datensätze mit einer Gesamtgazedauer von mehr als 50 ms berücksichtigt, um sicherzustellen, dass ausschließlich valide Fixationen in die Analyse eingehen. Werte unter 50 ms gelten als zu kurz, um eine valide visuelle Verarbeitung widerzuspiegeln, und wurden daher ausgeschlossen (0,51% der Datenpunkte für AOI 1 und 0,28% für AOI 2). Im Anschluss erfolgte die separate Analyse von AOI 1 und AOI 2.

Als abhängige Variable wurde jeweils die Gesamtgazedauer auf dem entsprechenden AOI betrachtet. Die unabhängige Variable war die Sprachstruktur, die analog zu den vorherigen Modellen als kategorialer fester Effekt definiert wurde. Die Zufallsstruktur entsprach der des Modells zur Analyse der Erstfixation und umfasste sowohl zufällige Intercepts als auch zufällige Slopes für Versuchspersonen und Stimuli. Aufgrund der annähernd normalverteilten Natur der Gazedauer wurde für beide Modelle eine Gauß'sche Verteilungsfamilie mit Identitäts-Link-Funktion verwendet.

Weiterhin wurden informative Priors definiert, die sich an der beobachteten empirischen Verteilung der Daten orientieren. Konkret wurde für den Intercept ein schwach informativer Student-t-Prior mit 3 Freiheitsgraden gewählt, dessen Mittelwert und Standardabweichung den jeweiligen Mittelwerten und Standardabweichungen der Gazedauer aus den Rohdaten der jeweiligen AOI entsprechen. Für die festen Effekte der Sprachstrukturen wurde ein normalverteilter Prior mit Mittelwert 0 und einer Standardabweichung von 50 ms gesetzt, was Effekte bis zu etwa ± 100 ms zulässt. Für die residuale Standardabweichung (σ) und die zufälligen Effekte (sd) wurden halbnormalverteilte Priors mit Mittelwert 0 und Standardabweichung 100 ms gewählt, wobei explizit eine untere Grenze von 0 ms gesetzt wurde, um sicherzustellen, dass nur positive Werte angenommen werden.

Die Bayes'sche Parameterschätzung wurde mit der Funktion *brm()* über vier Ketten mit jeweils 5000 Iterationen durchgeführt, wobei davon jeweils 500 Iterationen als Warm-Up verwendet wurden. Die Adaptionsrate (*adapt_delta*) wurde auf 0,95 festgelegt. Die Schätzung aller Kontraste wurde mit den Funktionen *emmeans()* und *pairs()* durchgeführt.

8.3. Ergebnisse

8.3.1. Ergebnisse der ersten Fixation

Die Analyse der ersten Fixation zeigte deutliche Unterschiede in der Verteilung der initialen visuellen Aufmerksamkeit auf die beiden kritischen AOIs in Abhängigkeit von

der Sprachstruktur. In der Bedingung „Hauptsatz“ entfielen 41,87% der ersten Fixationen auf AOI 1, während AOI 2 in 58,13% der Fälle zuerst fixiert wurde. Für Relativsätze ergab sich eine vergleichsweise ausgeglichene Verteilung. So wurde AOI 1 in 43,83% der Fälle zuerst fixiert, während AOI 2 in 56,16% der Fälle zuerst fixiert wurde. Hingegen zeigte sich in der Bedingung „Nominalisierung“ ein deutlich asymmetrischeres Fixationsmuster, wobei AOI 2 mit 66,02% der ersten Fixationen deutlich bevorzugt wurde, während AOI 1 lediglich 33,98% der ersten Fixationen erhielt.

Das Bayes'sche logistische gemischte Modell sowie die zusätzlich durchgeführten paarweisen Vergleiche bestätigten einen signifikanten Einfluss der Sprachstruktur auf die Position der ersten Fixation. Konkret war die Wahrscheinlichkeit einer ersten Fixation auf AOI 2 in der Bedingung „Nominalisierung“ signifikant höher als in der Bedingung „Hauptsatz“ (Median $\beta = 0,64$, 95% CI [0,16, 1,17], $p_d = 99,57\%$, ROPE = 0,44%). Ebenso wurde ein deutlicher Unterschied zwischen den Bedingungen „Nominalisierung“ und „Relativsatz“ festgestellt. In der Nominalisierungsbedingung war die Wahrscheinlichkeit, sich zuerst auf AOI 2 zu fixieren, signifikant höher als in der Relativsatzbedingung (Median $\beta = 0,78$, 95% CI [0,27, 1,34], $p_d = 99,84\%$, ROPE = 0%). Diese Ergebnisse verdeutlichen einen ausgeprägten Einfluss der Nominalisierungen auf die frühe visuelle Orientierung hin zu AOI 2 im Vergleich zu Haupt- und Relativsätzen.

Zwischen den Bedingungen „Relativsatz“ und „Hauptsatz“ konnte hingegen kein verlässlicher Unterschied hinsichtlich der ersten Fixation festgestellt werden (Median $\beta = -0,14$, 95% CI [-0,58, 0,32], $p_d = 73,29\%$, ROPE = 51,95%). Somit weisen diese Ergebnisse darauf hin, dass sich Relativ- und Hauptsätze bezüglich ihrer Wirkung auf die Position der ersten Fixation nicht signifikant unterscheiden.

Zusammenfassend bestätigen die Ergebnisse, dass insbesondere Nominalisierungen die Wahrscheinlichkeit, dass AOI 2 zuerst fixiert wurde, robust und signifikant erhöhten, während zwischen Relativ- und Hauptsätzen keine signifikanten Unterschiede

festgestellt werden konnten.

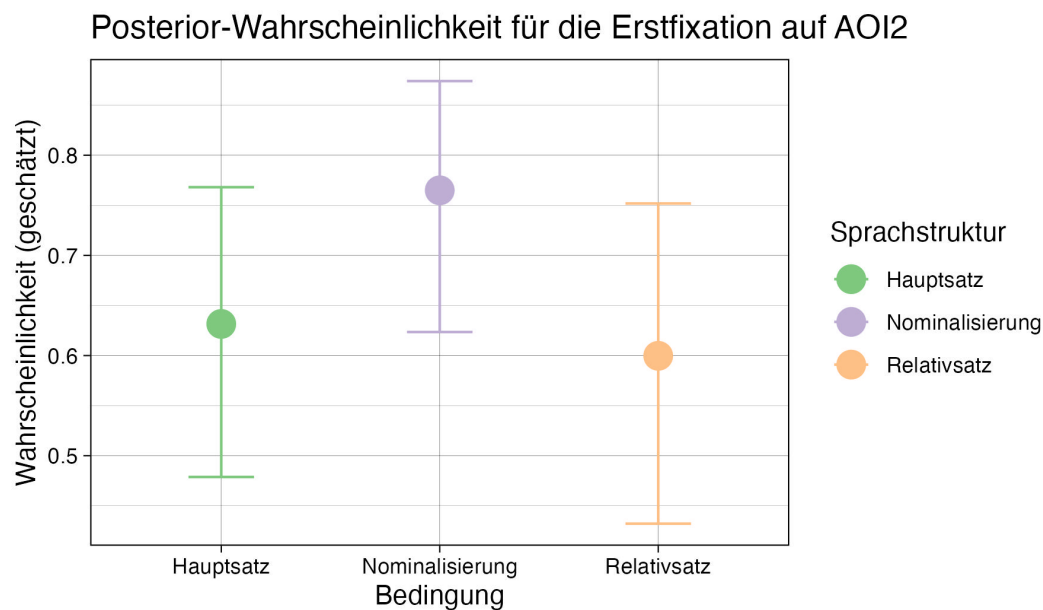


Abbildung 8.4: Posterior-Wahrscheinlichkeit für die Erstfixation auf die Patiens-/Handlungsregion (Experiment 3).

Vergleiche	Median	95% CI	pd	% in ROPE
Hauptsatz - Nominalisierung	-0,64	[-1,17, -0,16]	99,57%	0,44%
Hauptsatz - Relativsatz	0,14	[-0,32, 0,58]	73,29%	51,95%
Nominalisierung - Relativsatz	0,78	[0,27, 1,34]	99,84%	0%

Tabelle 8.1: Ergebnisse der posterioren paarweisen Vergleiche zur Erstfixation (Experiment 3).

8.3.2. Ergebnisse der Fixationslatenz

Die Analyse der Fixationslatenz auf AOI 1 zeigte deutliche Unterschiede zwischen den Sprachstrukturen. Die mediane Latenz auf der logarithmischen Skala betrug im Durchschnitt 5,91 (95% CI [5,78, 6,03]) in der Referenzstufe „Hauptsatz“. Im Vergleich zum Hauptsatz führte die Nominalisierung zu einer signifikant längeren Fixationslatenz auf AOI 1 (Median $\beta = 0,15$, 95% CI [0,05, 0,25], $pd = 99,96\%$, ROPE = 0%), während sich der Relativsatz kaum vom Hauptsatz unterschied (Median $\beta = -0,05$, 95% CI [-0,13, 0,04], $pd = 87,96\%$, ROPE = 9,76%). Die paarweisen Vergleiche bestätigten, dass AOI 1 bei Nominalisierungen signifikant später fixiert wurde als bei Hauptsätzen (Median der Differenz = -0,15, 95% CI [-0,25, -0,06], $pd = 99,96\%$, ROPE = 0%), und

auch später als bei Relativsätzen (Median der Differenz = 0,20, 95% CI [0,11, 0,30], $pd = 99,99\%$, ROPE = 0%). Zwischen Haupt- und Relativsätzen konnte hingegen kein eindeutiger Unterschied festgestellt werden (Median $\beta = 0,05$, 95% CI [-0,03, 0,14], $pd = 87,96\%$, ROPE = 9,76%).

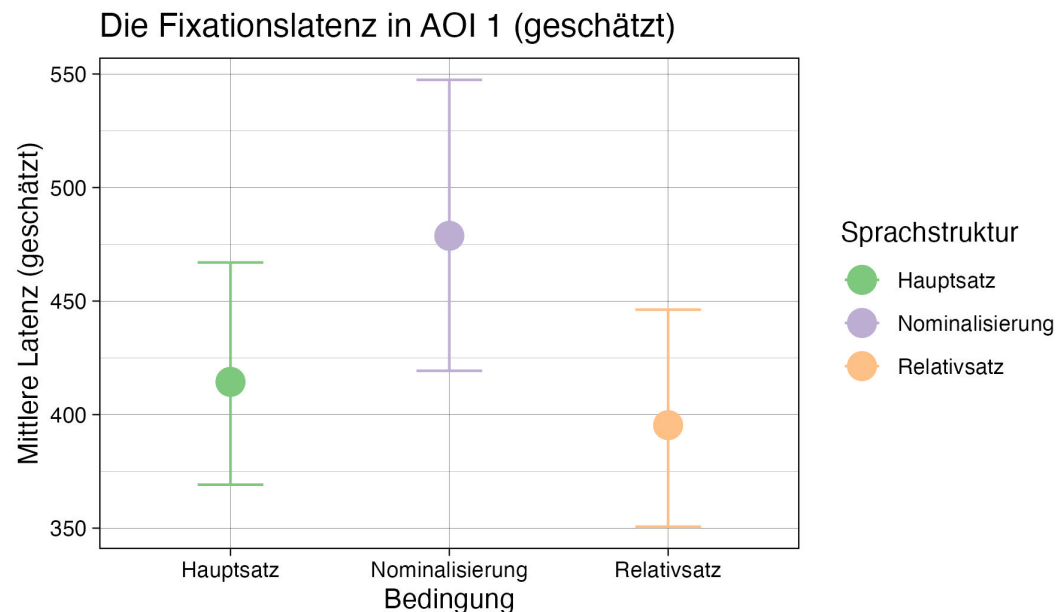


Abbildung 8.5: Geschätzte Fixationslatenz in die Agens-Region (Experiment 3).

In Bezug auf AOI 2 zeigte sich ein anderes Muster. Der Median der Fixationslatenz betrug hier für Hauptsätze 5,82 (95% CI [5,67, 5,97]). Bei Nominalisierungen wurden AOI 2 signifikant früher fixiert als bei Hauptsätzen (Median $\beta = -0,20$, 95% CI [-0,30, -0,10], $pd = 100\%$, ROPE = 0%). Im Gegensatz dazu zeigten Relativsätze keinen klaren Unterschied zu Hauptsätzen (Median $\beta = -0,03$, 95% CI [-0,14, 0,09], $pd = 69,13\%$, ROPE = 13,22%). Die *Post-hoc*-Vergleiche bestätigten ebenfalls signifikante Unterschiede zwischen Hauptsatz und Nominalisierung (Median der Differenz = 0,20, 95% CI [0,10, 0,30], $pd = 100\%$, ROPE = 0%) sowie zwischen Nominalisierung und Relativsatz (Median der Differenz = -0,17, 95% CI [-0,29, -0,06], $pd = 99,71\%$, ROPE = 0%). Der Unterschied zwischen Hauptsatz und Relativsatz war wiederum nicht eindeutig (Median der Differenz = 0,03, 95% CI [-0,09, 0,14], $pd = 69,13\%$, ROPE = 13,22%).

Zusammenfassend zeigen die Ergebnisse, dass Nominalisierungen im Vergleich zu

Haupt- und Relativsätzen mit einer verlängerten Latenz zur ersten Fixation auf die Agens-Region und mit einer verkürzten Latenz zur ersten Fixation auf die Patiens/Handlung-Region assoziiert sind. Dies deutet darauf hin, dass Nominalisierungen zu unterschiedlichen Mustern der visuellen Aufmerksamkeit führen als die beiden anderen untersuchten Sprachstrukturen.

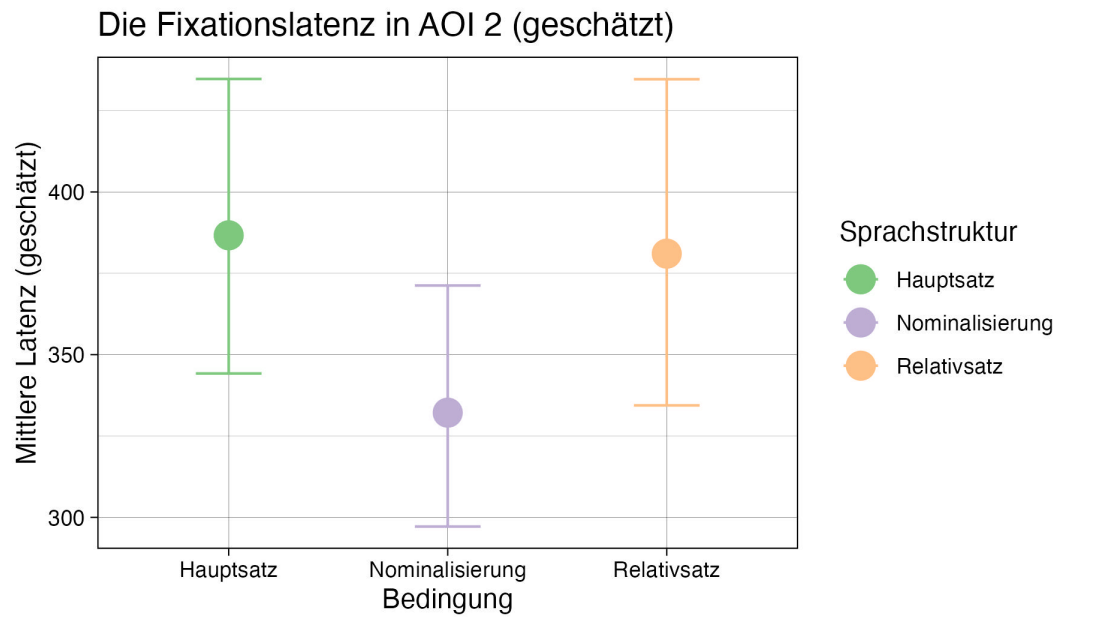


Abbildung 8.6: Geschätzte Fixationslatenz in die Patiens-/Handlungsregion (Experiment 3).

Vergleiche	Median	95% CI	pd	% in ROPE
AOI 1				
Hauptsatz - Nominalisierung	-0,15	[-0,25, -0,06]	99,96%	0%
Hauptsatz - Relativsatz	0,05	[-0,03, 0,14]	87,96%	9,76%
Nominalisierung - Relativsatz	0,20	[0,11, 0,30]	99,99%	0%
AOI 2				
Hauptsatz - Nominalisierung	0,20	[0,10, 0,30]	100%	0%
Hauptsatz - Relativsatz	0,03	[-0,09, 0,14]	69,13%	13,22%
Nominalisierung - Relativsatz	-0,17	[-0,29, -0,06]	99,71%	0%

Tabelle 8.2: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz (Experiment 3).

8.3.3. Ergebnisse der Gazedauer

Um die Auswirkungen unterschiedlicher Sprachstrukturen auf die gesamte Gazedauer an den kritischen AOIs zu untersuchen, wurden separate Bayes'sche hierarchische Modelle für AOI 1 und AOI 2 berechnet. Die Ergebnisse für beide AOIs sind nachfolgend dargestellt.

Die Auswertung des posterioren Modells für AOI 1 ergab eine geschätzte mediane Gazedauer im Hauptsatz von ca. 477 ms (95% CI [424,60, 528,67]). Der direkte Vergleich der Bedingungen mittels paarweiser Vergleiche ergab einen deutlichen Unterschied zwischen den Bedingungen „Hauptsatz“ und „Nominalisierung“ (Median Differenz = 123,27, 95% CI [81,17, 162,56], $pd = 100\%$, ROPE = 0%). Demnach war die Gazedauer an AOI 1 bei Nominalisierungen signifikant kürzer als bei Hauptsätzen. Ebenso zeigte sich ein eindeutiger Unterschied zwischen der Nominalisierung und dem Relativsatz (Median Differenz = -143,32, 95% CI [-188,39, -95,92], $pd = 100\%$, ROPE = 0%), wobei die Nominalisierung hier ebenfalls signifikant kürzere Verweildauern aufwies. Zwischen den Bedingungen „Hauptsatz“ und „Relativsatz“ ergab sich kein überzeugender Unterschied (Median Differenz = -20,25, 95% CI [-56,52, 17,57], $pd = 85,7\%$, ROPE = 0,01%).

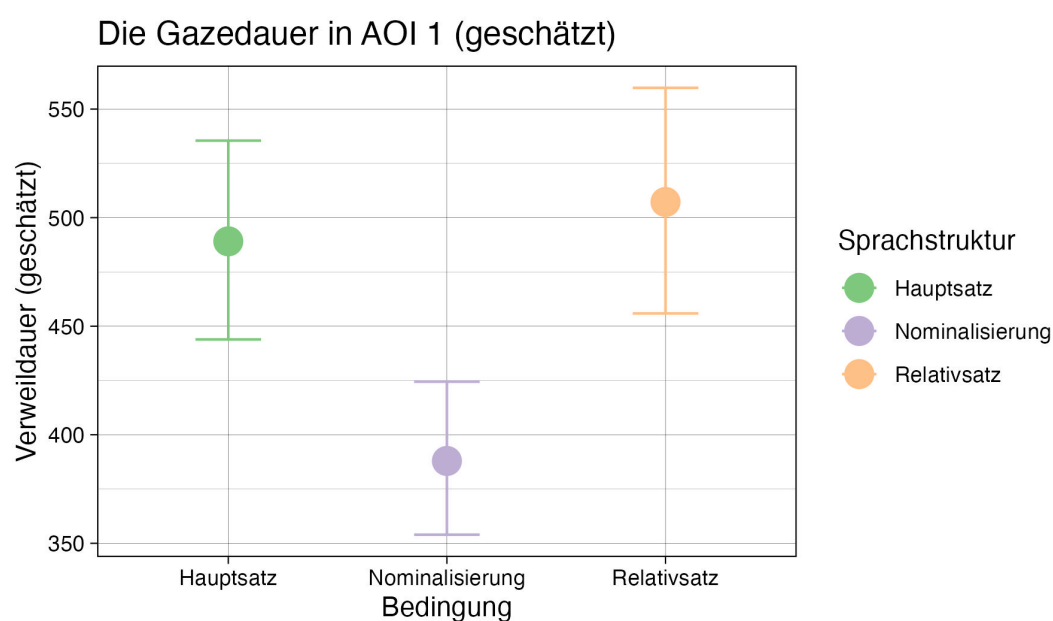


Abbildung 8.7: Geschätzte gesamte Gazedauer in der Agens-Region (Experiment 3).

Für AOI 2 ergaben sich vergleichbare Ergebnisse, jedoch in entgegengesetzter Richtung. Die mediane Gazedauer im Hauptsatz betrug hier ca. 631,81 ms (95% CI [588,09, 675,63]). Ein signifikanter Unterschied zeigte sich zwischen den Bedingungen „Hauptsatz“ und „Nominalisierung“ (Median Differenz = -106,65, 95% CI [-144,33, -67,84], $pd = 100\%$, ROPE = 0%), wobei bei Nominalisierungen längere gesamte Gazedauern an AOI 2 beobachtet wurden als bei Hauptsätzen. Auch zwischen Nominalisierung und Relativsatz zeigte sich ein deutlicher Unterschied zugunsten längerer Gazedauern bei der Nominalisierung (Median Differenz = 112,61, 95% CI [68,55, 155,20], $pd = 100\%$, ROPE = 0%). Zwischen den Bedingungen „Hauptsatz“ und „Relativsatz“ zeigte sich hingegen kein überzeugender Unterschied (Median Differenz = 5,95, 95% CI [-29,31, 41,50], $pd = 63,19\%$, ROPE = 0,02%).

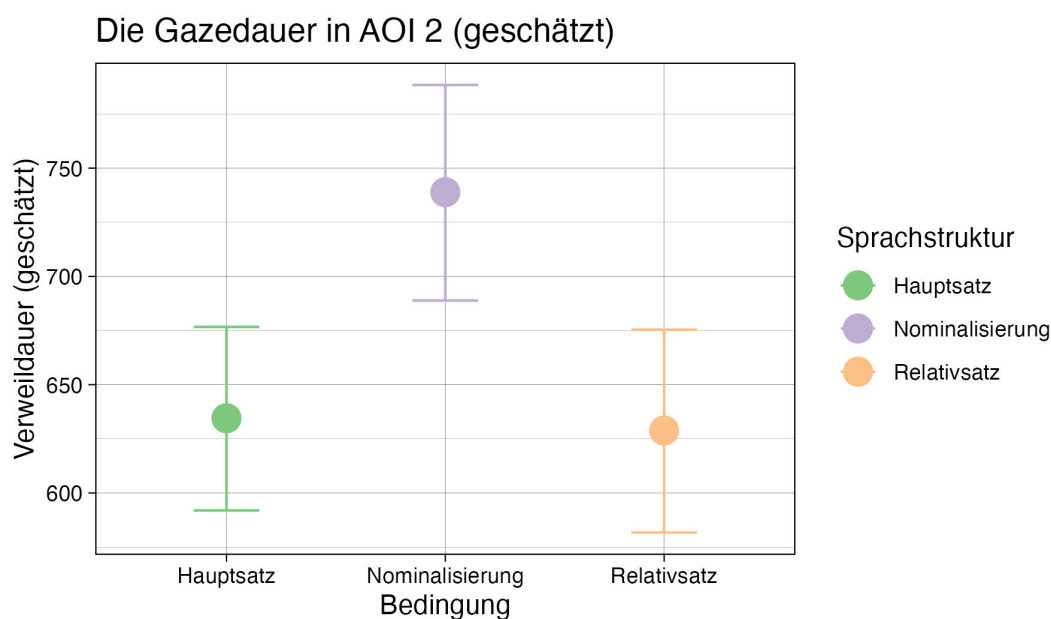


Abbildung 8.8: Geschätzte gesamte Gazedauer in der Patiens-/Handlungsregion (Experiment 3).

Vergleiche	Median	95% CI	pd	% in ROPE
AOI 1				
Hauptsatz - Nominalisierung	123,27	[81,17, 162,56]	100%	0%
Hauptsatz - Relativsatz	-20,25	[-56,52, 17,57]	85,70%	0,01%
Nominalisierung - Relativsatz	-143,32	[-188,39, -95,92]	100%	0%
AOI 2				

Hauptsatz - Nominalisierung	-106,65	[-144,33, -67,84]	100%	0%
Hauptsatz - Relativsatz	5,95	[-29,31, 41,50]	63,19%	0,02%
Nominalisierung - Relativsatz	112,61	[68,55, 155,20]	100%	0%

Tabelle 8.3: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer (Experiment 3).

8.4. Diskussion

Im vorliegenden Experiment wurde mithilfe von Eyetracking untersucht, wie sich unterschiedliche linguistische Strukturen auf die Verteilung der visuellen Aufmerksamkeit bei der Sprach- und Ereignisverarbeitung auswirken und welche Ereigniskomponenten dabei priorisiert werden. Gemäß der sogenannten „*eye-mind*“-Hypothese stehen Augenbewegungen in einem direkten, messbaren Zusammenhang mit zugrunde liegenden kognitiven Prozessen (vgl. Carter & Luke, 2020; Divjak et al., 2020; Rayner, 2009).

Durch die Verarbeitung sprachlicher Strukturen werden sich dynamisch entfaltende Repräsentationen eines Ereignisses generiert, wobei mentale Modelle über das enkodierte Ereignis aufgebaut werden. Augenbewegungen, die nach Abschluss der Sprachverarbeitung auf einen visuellen Stimulus gemessen werden, spiegeln wider, inwieweit die visuell wahrgenommene Szene mit der mental konstruierten Ereignisstruktur übereinstimmt. Je stärker ein Objekt oder eine Region im Bild mit der aus sprachlich vermittelten Informationen aufgebauten mentalen Repräsentation übereinstimmt, desto wahrscheinlicher wird sie visuell fixiert (z.B. Altmann & Kamide, 2007; Kang, Joergensen & Altmann, 2020). Dies verdeutlicht, dass Blickbewegungen maßgeblich von der sprachlich gesteuerten Konzeptualisierung eines Ereignisses in *top-down*-Weise beeinflusst werden. Demnach kann die visuelle Aufmerksamkeitsverteilung reflektieren, welche Elemente der Szene mental besonders salient sind.

Frühere Eyetracking-Studien zur Untersuchung des *Construals* von Ereignissen zeigen, dass unterschiedliche sprachliche Konstruktionen aufgrund variierender syntaktischer und semantischer Perspektiven die Salienz einzelner Ereigniskomponenten

beeinflussen und sich dies entsprechend in unterschiedlichen Blickbewegungsmustern niederschlägt (z.B. Divjak et al., 2020; Sauppe & Flecken, 2021). Diese Studien konzentrierten sich jedoch vor allem auf die Perspektivverschiebung in englischen Hauptsätzen, beispielsweise im Vergleich von Aktiv- und Passivsätzen. Andere Strukturen und Sprachen wurden bislang kaum berücksichtigt.

Die vorliegende Arbeit erweitert diesen Ansatz, indem sie den Einfluss komplexerer Sprachstrukturen auf *Construal* untersucht. Insbesondere weisen Relativsätze und Nominalisierungen wesentliche Unterschiede zu Hauptsätzen auf, da sie als Nominalphrasen in einen übergeordneten sprachlichen Kontext eingebettet werden können. Ausgehend von den im theoretischen Teil erläuterten kognitiven Prozessen wird angenommen, dass diese beiden Konstruktionen unterschiedliche Perspektiven auf ein Referenz-Ereignis abbilden und jeweils spezifische, komplexe Repräsentationen hervorrufen, die in ein übergeordnetes Ereignismodell integriert werden können. Auf diese Weise beeinflussen sie das *Construal* eines Referenz-Ereignisses.

Vor diesem theoretischen und empirischen Hintergrund wurde im vorliegenden Experiment die Hypothese geprüft, dass die drei untersuchten sprachlichen Strukturen einen Einfluss auf die visuelle Aufmerksamkeitsverteilung ausüben, messbar an drei Blickbewegungsparametern: Position der ersten Fixation, Fixationslatenz und Gazedauer. Es wurde spezifisch erwartet, dass Nominalisierungen aufgrund der sprachlich induzierten Hervorhebung von Patiens und Handlung dazu führen, dass diese Komponenten mit höherer Wahrscheinlichkeit zuerst, früher und länger fixiert werden als bei Haupt- und Subjekt-Relativsätzen, während das Agens in Nominalisierungen seltener zuerst fixiert wird und eine kürzere Gazedauer aufweist. Für Subjekt-Relativsätze wurde angenommen, dass das Agens mit höherer Wahrscheinlichkeit zuerst, früher und länger fixiert wird als bei Hauptsätzen.

In diesem Experiment wurde die Darbietungszeit der visuellen Stimuli auf 1500 ms beschränkt, wodurch ein Zeitdruck entstand. Die zugrunde liegende Logik bestand

darin, dass die Versuchspersonen aufgrund der begrenzten Darbietungszeit schnell entscheiden mussten, wohin sie zuerst blicken sollten. Diese Entscheidung kann in der Regel sowohl auf *bottom-up*-Faktoren wie visuelle Merkmale als auch auf *top-down*-Faktoren wie Vorwissen, Absichten oder spezifische Aufgabenanforderungen basieren (vgl. Isasi-Isasmendi et al., 2023). Von besonderem Interesse war, welche Informationen die Versuchspersonen innerhalb dieser begrenzten Zeit aufnehmen konnten und welche Ereigniskomponenten dabei spontan priorisiert wurden.

Die Ergebnisse bestätigten die Hypothesen zu den Nominalisierungen. Die Analyse der Erstfixation zeigte, dass die unterschiedlichen sprachlichen Strukturen bereits in einer sehr frühen Phase während der Präsentation der Szenen einen Einfluss auf die Aufmerksamkeitverteilung ausübten. Bei Nominalisierungen war die Wahrscheinlichkeit einer ersten Fixation auf die Patiens-/Handlungsregion signifikant höher als bei Haupt- und Relativsätzen. Umgekehrt entfielen bei Haupt- und Relativsätzen deutlich mehr Erstfixationen auf die Agens-Region als bei Nominalisierungen. Diese Muster wurden durch die Fixationslatenzen weiter untermauert. In der Nominalisierungsbedingung erfolgten Fixationen auf die Patiens-/Handlungsregion früher und auf die Agens-Region später als in den Bedingungen mit Haupt- oder Relativsätzen.

Die Ergebnisse zur Gazedauer legen nahe, dass die linguistischen Bedingungen einen signifikanten Einfluss auf die visuelle Aufmerksamkeitsverteilung in den kritischen AOIs während des gesamten Prozesses hatten. Die Gazedauer war bei Nominalisierungen auf der Patiens-/Handlungsregion signifikant länger und auf der Agens-Region kürzer als bei Haupt- und Relativsätzen.

Diese Befunde lassen sich darauf zurückführen, dass sich die konzeptuelle Struktur der Ereignisrepräsentation in Abhängigkeit von der linguistisch evozierten Salienz bestimmter Ereigniskomponenten unterscheidet. Wie im theoretischen Teil erläutert, besteht eine linguistische Bedeutung nicht nur in dem Inhalt einer Äußerung, sondern auch in *Construal*, d.h. der spezifischen Art und Weise, wie dieser Inhalt

konzeptualisiert wird (vgl. Langacker, 2013, 2019). Obwohl alle drei untersuchten Strukturen dasselbe kausative Ereignis kodierten, unterschieden sie sich darin, welcher Aspekt des Ereignisses linguistisch besonders hervorgehoben wurde. Während Subjekt-Relativsätze das Subjekt bzw. Agens in den Mittelpunkt rücken, profilieren Nominalisierungen eher das explizit erwähnte Patiens und die Handlung. Bei Hauptsätzen wurden theoretisch die drei zentralen Ereigniskomponenten relativ ausgeglichen im Vergleich zu den anderen beiden Strukturen (vgl. Kapitel 3 & 4).

Diese unterschiedliche linguistische Fokussierung führt dazu, dass beim Hören der linguistischen Inputs bestimmte Ereigniskomponenten je nach Sprachstruktur als konzeptuelle Figur hervorgehoben und entsprechend mental besonders salient repräsentiert werden. Diese aktivierten mentalen Repräsentationen bleiben auch nach dem sprachlichen Input im Arbeitsgedächtnis erhalten. In der anschließenden visuellen Wahrnehmung des Bildes gelang es den Versuchspersonen trotz der hohen Komplexität der Szene, die zentrale Ereignisstruktur bereits innerhalb von Sekundenbruchteilen zu erfassen (vgl. Dobel et al., 2007; Griffin & Bock, 2000). Es wurde angenommen, dass diese visuellen Repräsentationen bei allen drei Sprachstrukturbedingungen sehr ähnlich waren, da bei den drei Bedingungen das identische Bild gezeigt wurde. Die beobachteten Unterschiede in den Augenbewegungen lassen sich daher auf die vorab durch Sprache aufgebauten Ereignisrepräsentationen zurückführen.

Die Versuchspersonen greifen in diesem Fall auf die zuvor aktivierten linguistischen Repräsentationen zurück, die sich *top-down* auf die visuelle Wahrnehmung auswirken. Sie spiegeln die unterschiedlichen Profilierungen der sprachlichen Strukturen wider und steuern damit automatisch, welche Ereigniskomponente als besonders aufmerksamkeitsregend wahrgenommen wurde und zuerst fixiert werden sollte. Augenbewegungen dienen somit nicht der Aufnahme visueller Information aus dem Bild, sondern reflektieren die durch die Sprachverarbeitung etablierte Profilierung der Ereignisrepräsentation.

Die Tatsache, dass die erste Fixation innerhalb weniger Millisekunden erfolgte,

verdeutlicht, dass diese visuelle Präferenz für die linguistisch hervorgehobenen Ereigniskomponenten nicht durch bewusste Überlegungen oder Strategien zustande kam. Vielmehr handelte es sich hierbei um eine spontane, unbewusste kognitive Verarbeitung. Die beobachteten Effekte der sprachlichen Strukturen auf die erste Fixation bestätigen somit die Hypothese, dass unterschiedliche linguistische Hervorhebungen bereits die frühe visuelle Aufmerksamkeit und Ereigniswahrnehmung steuern. Diese Ergebnisse unterstützen damit die früheren Untersuchungen, die zeigten, dass Menschen jene Elemente zuerst fixieren, die für sie von besonderer Relevanz erscheinen (z.B. Divjak et al., 2020; Sauppe & Flecken, 2021).

Frühere Studien weisen darauf hin, dass die Gazedauer ein Indikator für den kognitiven Aufwand bei der Informationsverarbeitung ist (vgl. Divjak et al., 2020). Sie korreliert eng mit der Komplexität, Informativität und Salienz einer Region (vgl. Holmqvist et al., 2011). Blickbewegungen verteilen sich nicht zufällig auf bestimmte Bereiche eines Bildes, sondern werden durch deren inhaltliche Relevanz gesteuert. Je relevanter ein Bereich für die Interpretation einer Szene oder für eine aktuelle Aufgabe ist, desto häufiger und intensiver wird dieser fixiert. Eine längere Gazedauer kann daher auf einen erhöhten kognitiven Aufwand hinweisen, z.B. weil die Informationen in diesem Bereich besonders bedeutsam sind. Die im Experiment beobachtete verlängerte Gazedauer auf die Patiens-/Handlungsregion bei Nominalisierungen spricht dafür, dass diese Region im Kontext dieser sprachlichen Struktur als besonders salient konzeptualisiert wurde.

Diese Befunde stützen die theoretische Annahme, dass Nominalisierungen die Ereigniskomponenten Patiens und Handlung stärker in den Fokus rücken als bei Haupt- und Subjekt-Relativsätzen. Bei Nominalisierungen werden Patiens und Handlung explizit benannt, während das Agens nur schematisch impliziert wird. Es bleibt offen, ob und welches konkrete Agens die Handlung ausführt. Daher liefern die spezifizierte Handlung und das Patiens konkretere und salientere Informationen im Vergleich zum impliziten Agens und gewinnen an Salienz (vgl. Talmy, 2000). In Haupt- und Subjekt-Relativsätzen hingegen werden alle zentralen Ereigniskomponenten explizit kodiert,

sodass deren *Specificity* auf einem vergleichbaren Niveau liegt. Bei Haupt- und Subjekt-Relativsätzen erscheinen folglich Patiens und Instrument nicht salienter als Agens im Vergleich zu Nominalisierungen.

Im Gegensatz zu den Vergleichen zwischen Nominalisierungen einerseits und Haupt- sowie Relativsätzen andererseits, bei denen hinsichtlich aller drei verwendeten Blickbewegungsmaße signifikante Unterschiede nachgewiesen wurden, konnten zwischen Haupt- und Relativsätzen in keiner der drei Messgrößen verlässliche Unterschiede festgestellt werden.

Wie im theoretischen Teil erläutert wurde, rückt das Kopfnomen in Subjekt-Relativsätzen in den Fokus, während die übrigen Satzteile zum Relativsyntagma herabgestuft und in den Hintergrund verschoben werden. Bei Hauptsätzen steht hingegen die gesamte Handlungskette bzw. stehen alle drei zentralen Ereigniskomponenten im Vordergrund gegenüber den hintergründigen *Settings* (vgl. Kapitel 3; Langacker, 1991). Daher wäre zu erwarten gewesen, dass sich Relativsätze und Hauptsätze in ihren Aufmerksamkeitsmustern unterscheiden. Im Experiment zeigten sich jedoch vergleichbare Blickbewegungsmuster bei beiden Bedingungen, was auf eine geringere Differenzierung in der visuellen Verarbeitung hinweist.

Eine mögliche Erklärung liegt darin, dass Haupt- und Subjekt-Relativsätze viele Eigenschaften auf kognitiver Ebene teilen. Beide Strukturen enthalten dasselbe Verb und dieselben Argumente, werden nach Langacker (1987, 1991) mental sequentiell gescannt und profilieren einen dynamischen Prozess. Außerdem sind drei zentrale Ereigniskomponenten in beiden Strukturen explizit ausgedrückt. Diese starken Gemeinsamkeiten in der Ereignisstruktur und ihrer Konzeptualisierung können dazu führen, dass die Unterschiede in der Profilierung in der konkreten experimentellen Umsetzung nicht hinreichend stark hervortraten, um signifikante Unterschiede in den Blickbewegungen zu erzeugen.

Hinzu kommen mögliche Erklärungen zum experimentellen Design. In diesem

Experiment wurden die visuellen Stimuli nur für 1500 ms präsentiert. Diese kurze Darbietungszeit erlaubt typischerweise nur zwei bis vier Fixationen. Zudem erschienen die Bilder entweder links oder rechts auf dem Bildschirm, sodass die erste Aufnahme der Szene über das periphere Sehen und nicht über eine unmittelbare zentrale Fixation erfolgte. Diese Kombination aus kurzer Darbietungszeit und seitlicher Präsentation dürfte die Blickbewegungen zusätzlich beeinflusst haben, sodass feinere Unterschiede zwischen Haupt- und Relativsatzbedingungen möglicherweise überdeckt wurden.

Darüber hinaus lassen sich diese Ergebnisse wahrscheinlich auf mehrdimensionale Asymmetrien zwischen den beiden Ereignisteilnehmern hinsichtlich der Belebtheit, der thematischen Rolle und der grammatischen Funktion zurückführen. Im vorliegenden Experiment wurde das Agens immer durch belebte, menschliche Figuren repräsentiert, während das Patiens sowie die Instrumente überwiegend unbelebte Objekte waren. Zahlreiche Studien zeigen, dass unser visuelles Aufmerksamkeitssystem seine Ressourcen nicht gleichmäßig auf alle Elemente einer Szene verteilt (vgl. Simons & Levin, 1997). Belebte, potenziell handlungsfähige Entitäten, insbesondere menschliche Figuren, werden als besonders salient wahrgenommen und ziehen daher mehr Aufmerksamkeit auf sich als unbelebte Objekte (z.B. Altman et al., 2016; New, Cosmides & Tooby, 2007).

Zudem genießt das Agens gemäß der thematischen Hierarchie und zahlreichen empirischen Studien aufgrund prototypischer Eigenschaften wie Handlungsfähigkeit, Belebtheit und Intentionalität sowie seiner grammatischen Prominenz grundsätzlich eine höhere kognitive Präferenz als das Patiens (vgl. Bock & Warren, 1985; Dowty, 1991; Dobel et al., 2007; Fillmore, 1968; Isasi-Isasmendi et al., 2023; Haspelmath, 2011). Beispielsweise beobachteten Isasi-Isasmendi et al. (2023) in einem Eyetracking-Experiment, dass die erste Fixation häufiger auf das Agens gerichtet wurde. Ebenso stellten Cohn und Paczynski (2013) fest, dass das Agens unabhängig von seiner linearen Reihenfolge in einer linguistischen Konstruktion länger fixiert wurde als Patiens.

Des Weiteren kodiert das Subjekt in den Hauptsätzen des vorliegenden Experiments

den Trajektor eines Ereignisses, der der zentrale und prominenteste Teilnehmer des Ereignisses ist. Im Vergleich dazu kodiert das Objekt die Landmark, die der zweitwichtigste Teilnehmer ist (vgl. Langacker, 1987, 1991). Das Subjekt gilt also grundsätzlich als salienter als das Objekt (vgl. Bock & Warren, 1985).

Aufgrund der vorliegenden Asymmetrien zwischen den beiden Ereignisteilnehmern ist davon auszugehen, dass das Agens bei der Verarbeitung von Haupt- und Relativsätzen tatsächlich eine vergleichbare Salienz aufweist, was zu vergleichbaren Mustern der Aufmerksamkeitsverteilung führt.

Die ähnlichen Muster bei Haupt- und Relativsätzen schließen nicht aus, dass kognitive Unterschiede zwischen beiden Strukturen bestehen. Sie legen lediglich nahe, dass das vorliegende Design möglicherweise nicht hinreichend sensitiv war, um die subtilen Unterschiede zwischen diesen Repräsentationen sichtbar zu machen.

Ein möglicher Einwand gegenüber den hier berichteten Befunden betrifft die Frage, ob die beobachteten Unterschiede zwischen Hauptsätzen und Nominalisierungen sowie zwischen Subjekt-Relativsätzen und Nominalisierungen möglicherweise auf sogenannte *Primacy*- oder *Recency*-Effekte zurückzuführen sind. Diese beiden Effekte beschreiben gegensätzliche kognitive Bias. Der *Primacy*-Effekt bezeichnet das Phänomen, dass die am Anfang einer Sequenz präsentierten Informationen stärker gewichtet und besser im Gedächtnis verankert werden als nachfolgende Informationen (vgl. Murdock, 1962; Glanzer & Cunitz, 1966). Demgegenüber besagt der *Recency*-Effekt, dass zuletzt wahrgenommene Informationen aufgrund ihrer unmittelbaren Verfügbarkeit im Kurzzeitgedächtnis besonders gut erinnert werden, da sie zeitlich näher am Abrufpunkt liegen (vgl. Tzeng, 1973). Beide Effekte könnten theoretisch einen Einfluss darauf nehmen, welche Ereigniskomponente zuerst fixiert oder als zentral wahrgenommen wird.

Allerdings werden diese Erklärungen für das vorliegende Experiment als höchst unwahrscheinlich eingeschätzt. Zwar steht das Agens im Hauptsatz am Satzanfang, was

gemäß dem *Primacy*-Effekt zu erhöhter kognitiver Prominenz führen könnte, während es im Subjekt-Relativsatz am Ende des Satzes erscheint, was eine Verstärkung durch den *Recency*-Effekt nahelegen würde. Dennoch zeigten sich zwischen diesen beiden Bedingungen keine signifikanten Unterschiede in den Blickbewegungen, was gegen eine Erklärung durch positionsbasierte *Primacy*- oder *Recency*-Effekte spricht. Darüber hinaus wiesen die Effekte zwischen Hauptsätzen und Nominalisierungen sowie zwischen Relativsätzen und Nominalisierungen ähnliche Muster auf und verliefen jeweils in die gleiche Richtung. Dies deutet ebenfalls darauf hin, dass weder *Primacy*- noch *Recency*-Effekte die beobachteten Aufmerksamkeitsverteilungen hinreichend erklären können. Die Blickbewegungen wurden daher nicht von der linearen Position der Ereigniskomponenten in der Satzstruktur beeinflusst, sondern von der linguistisch evozierten Salienz und konzeptuellen Fokussierungen der Komponenten.

Zusammenfassend zeigen die Ergebnisse, dass unterschiedliche sprachliche Strukturen die Aufmerksamkeit sowohl in sehr frühen als auch in gesamten Phasen der visuellen Wahrnehmung beeinflussen. Die sprachinduzierten Effekte erwiesen sich über alle drei Blickbewegungsmaße hinweg als konsistent und äußerten sich in Verschiebungen der Aufmerksamkeitsverteilung über die AOIs. Diese Blickbewegungsmuster können so interpretiert werden, dass sie die durch die jeweiligen Strukturen etablierten Ereignisrepräsentationen widerspiegeln. Sprachliche Strukturen modulieren die konzeptuelle Salienz einzelner Ereigniskomponenten und steuern dadurch die visuelle Verarbeitung *top-down*. Darüber hinaus replizierte das vorliegende Experiment die in Experiment 2 beobachtete Präferenz für die Hervorgehobenheit von Patiens und Handlung bei Nominalisierungen. Die hier erhobenen Fixationsdaten liefern zusätzlich direkte und detaillierte Hinweise auf die konzeptuell gesteuerte Aufmerksamkeitslenkung während der visuellen Ereignisverarbeitung.

9. Experiment 4

Die zuvor durchgeführten drei Experimente ergaben mithilfe verschiedener

experimenteller Methoden konsistente Hinweise darauf, dass die drei untersuchten sprachlichen Strukturen im Chinesischen die Konzeptualisierung von Ereignissen modulieren, indem sie unterschiedliche Ereigniskomponenten hervorheben. Experiment 1 belegte mittels einer *offline Picture-Selection*-Aufgabe, dass die jeweils konzeptuell hervorgehobenen Komponenten der Sprachstrukturen den theoretischen Annahmen entsprachen. Experiment 2 zeigte anhand von Reaktionszeitmessungen, dass die Verarbeitung beschleunigt wurde, wenn die linguistisch hervorgehobene Ereigniskomponente mit der visuell betonten Komponente eines präsentierten Bildes übereinstimmte.

Experiment 3 verdeutlichte mittels Eyetracking, dass die sprachlich evozierte Hervorhebung bestimmter Ereigniskomponenten die visuelle Aufmerksamkeit lenkt. Besonders die Nominalisierungsbedingung wich deutlich von den Bedingungen mit Haupt- und Subjekt-Relativsätzen ab. In dieser Bedingung entfiel ein größerer Anteil der frühen und kumulierten Fixationen auf die Patiens-/Handlungsregion, während das Agens relativ an visueller Gewichtung verlor.

Eine noch offene Frage betrifft jedoch die Generalisierbarkeit dieser Ergebnisse: Lassen sich vergleichbare Effekte der untersuchten Sprachstrukturen – Hauptsatz, Subjekt-Relativsatz und Nominalisierung – auch in anderen Sprachen nachweisen, oder bestehen sprachspezifische Unterschiede hinsichtlich der Aufmerksamkeitsverteilung bei der Ereigniskonzeptualisierung? Frühere Studien deuten darauf hin, dass typologische Unterschiede zwischen Sprachen kurzfristige Einflüsse während der Sprachproduktion (z.B. Gennari et al., 2002; Papafragou et al., 2008) oder langfristige Einflüsse auf nicht-linguistische kognitive Prozesse wie das Gedächtnis (z.B. Flecken, Athanasopoulos et al., 2015; Gerwien & von Stutterheim, 2018) haben können. Allerdings konzentrieren sich solche Studien meist auf die Konzeptualisierung von Ereignissen während der Vorbereitung des Sprechens. Wenig untersucht ist bislang hingegen die Frage, ob und wie sprachliche Unterschiede die durch Sprachstrukturen erzeugte Ereignisrepräsentation beeinflussen und dadurch die visuelle Aufmerksamkeit auf nachfolgende Stimuli lenken. Dieser Frage wird im vorliegenden Experiment

empirisch nachgegangen.

Vor diesem Hintergrund legt das vorliegende Experiment besonderen Wert auf den Einfluss der sprachlichen Unterschiede zwischen Deutsch und Chinesisch auf die Konzeptualisierung kanonischer kausativer Ereignisse. Experiment 4 verfolgt das Ziel, die Effekte dieser drei sprachlichen Strukturen auf die visuelle Aufmerksamkeitsverteilung im Deutschen zu analysieren und somit einen sprachübergreifenden Vergleich mit dem Chinesischen zu ermöglichen.

9.1. Hypothesen

Ausgehend von Experiment 3 stellt sich die Frage, ob die dort beobachteten Blickbewegungsmuster bei den drei untersuchten Sprachstrukturen auch im Deutschen auftreten oder ob sich aufgrund sprachspezifischer Eigenschaften abweichende Muster ergeben.

Wie in Kapitel 4 ausführlich dargelegt, unterscheiden sich Deutsch und Chinesisch in der morphologischen und syntaktischen Realisierung von Hauptsätzen, Subjekt-Relativsätzen und Nominalisierungen, während ihre semantische Struktur theoretisch jeweils vergleichbar bleibt. Besonders Relativsätze unterscheiden sich in der linearen Abfolge der konzeptuellen Komponenten. Daraus ergibt sich die Frage, ob diese Unterschiede in der Oberflächenstruktur die Profilierung und visuelle Verarbeitung der Ereigniskomponenten modifizieren oder ob die Ereigniskonzeptualisierung in beiden Sprachen weitgehend ähnlich ausfällt. Basierend auf diesen Überlegungen wurden folgende Hypothesen zu den drei Sprachstrukturen formuliert:

Deutsche und chinesische einfache transitive Hauptsätze besitzen dieselbe grundlegende SVO-Wortstellung. Auch die Reihenfolge der thematischen Rollen *Agens* - *Patiens* ist identisch. Aufgrund dieser strukturellen Gemeinsamkeiten auf der Oberfläche und in der semantischen Struktur wird angenommen, dass Hauptsätze in beiden Sprachen eine vergleichbare visuelle Aufmerksamkeitsverteilung hervorrufen.

Nominalisierungen unterscheiden sich zwar hinsichtlich ihrer sprachlichen Realisierung zwischen Deutsch und Chinesisch – während im Chinesischen die Struktur [NP + *DE* + VP] verwendet wird, liegen im Deutschen Rektionskomposita vor –, jedoch zeigen beide Strukturen eine vergleichbare Reihenfolge der Komponenten, wobei zunächst das Patiens und gefolgt von der Handlung kommt. In beiden Sprachen darf nur ein Argument explizit genannt werden, wodurch das Agens implizit bleibt. Daher wird angenommen, dass Nominalisierungen in beiden Sprachen ähnliche Effekte auf die visuelle Verarbeitung von Ereignissen zeigen.

Ein besonderes Augenmerk liegt auf Subjekt-Relativsätzen, da sich chinesische und deutsche Relativsätze in ihrer linearen Abfolge der Komponenten unterscheiden. Im Deutschen steht das Kopfnomen in Relativsätzen an erster Stelle. Chinesische Relativsätze haben hingegen die kopf-finale Eigenschaft, d.h., das Kopfnomen befindet sich hier am Ende der Struktur. Trotz dieser Unterschiede weisen beide Sprachen eine vergleichbare semantische Struktur auf. Das Kopfnomen (im Experiment stets das Agens) wird hervorgehoben, während die Handlung und das Patiens im Relativsyntagma den Hintergrund bilden. Wenn die lineare Wortstellung für die kognitive Verarbeitung eine Rolle spielt, wären unterschiedliche Blickbewegungsmuster in beiden Sprachen zu erwarten. Wenn jedoch nicht die Wortstellung, sondern die semantische Struktur, insbesondere die Figur-Grund-Beziehung, die Ereignisrepräsentation beeinflusst, sollten vergleichbare Effekte in beiden Sprachen auftreten.

9.2. Methodisches Vorgehen

9.2.1. Versuchspersonen

An dem Experiment nahmen insgesamt 42 Versuchspersonen teil, darunter 28 Frauen. Ihr Alter lag zwischen 20 und 36 Jahren ($M = 24,9$ Jahre). Alle Versuchspersonen waren deutsche Muttersprachler und Studierende der Universität Heidelberg. Sie verfügten über normale oder mittels Brille bzw. Kontaktlinsen korrigierte Sehkraft sowie ein

normales Hörvermögen. Keine der Versuchspersonen berichtete über neurologische Beeinträchtigungen oder Lesestörungen. Die Teilnahme erfolgte freiwillig. Als Dankeschön erhielten alle Versuchspersonen nach Abschluss des Experiments Kekse.

9.2.2. Apparatur

Die technische Apparatur war identisch mit der in Experiment 3 beschriebenen.

9.2.3. Material

Die in diesem Experiment verwendeten Materialien, bestehend aus den kritischen Stimuli sowie den Übungs- und Filler-Stimuli, entsprachen im Wesentlichen denen aus Experiment 3. Sie wurden jedoch sprachlich an das Deutsche angepasst. Die visuellen Stimuli blieben identisch zu denen in Experiment 3. Die linguistischen Stimuli wurden inhaltlich unverändert aus dem Chinesischen ins Deutsche übersetzt. Dabei wurden alle experimentellen Bedingungen konstant gehalten. Ein Beispiel für ein kritisches deutsches linguistisches Stimulus-Set ist das folgende:

- Bsp. 9-1 Hauptsatz: Die Barista bereitet Kaffee zu.
 Relativsatz: Die Barista, die Kaffee zubereitet.
 Nominalisierung: die Kaffeezubereitung

Die auditiven Stimuli wurden mithilfe einer Sprachausgabe-App erzeugt. Dabei kam eine deutsche weibliche Stimme mit konstanter Sprechgeschwindigkeit zum Einsatz. Es wurde besonders darauf geachtet, dass weder prosodische Auffälligkeiten wie Betonungen oder emotionale Färbungen noch unnatürliche Pausen auftraten, um potenzielle Ablenkungen bei der Verarbeitung zu vermeiden.

9.2.4. Durchführung

Die Durchführung dieses Experiments entsprach in allen wesentlichen Aspekten derjenigen von Experiment 3.

Alle Versuchspersonen wurden einzeln im *Eye-Tracking-Labor* des Instituts für Deutsch als Fremdsprachenphilologie der Universität Heidelberg getestet. Sie gaben zunächst demografische Angaben zu Geschlecht, Alter und Händigkeit an und nahmen anschließend vor dem Experiment-PC Platz. Die Kopfstütze wurde individuell auf die jeweilige Körpergröße eingestellt, um stabile Messbedingungen zu gewährleisten.

Anschließend erhielten die Versuchspersonen die Instruktion und Beispiele zur Aufgabe in deutscher Sprache. Die Experimentleiterin erklärte zusätzlich mündlich die Aufgabe und den Ablauf eines Durchgangs sowie etwaige Rückfragen. Nach erfolgreicher Durchführung der Kalibrierung, Validierung und Drift-Korrektur begann der experimentelle Block. Alle Beurteilungsfragen wurden den Versuchspersonen in deutscher Sprache präsentiert.

9.2.5. Datenkodierung

Die Definition der AOIs in den kritischen Stimuli entsprach vollständig derjenigen aus Experiment 3.

9.2.6. Datenanalyse

Zunächst wurden die Daten für die statistische Auswertung aufbereitet. Dabei wurden 7,41% der Beobachtungen aufgrund falscher Antworten bei den Beurteilungsfragen ausgeschlossen. Weitere 1,85% wurden wegen der schlechten Datenqualität entfernt.

Die statistischen Modelle und die *Post-hoc*-Vergleiche zur Analyse der ersten Fixation, der Fixationslatenz und der gesamten Gazedauer entsprachen für die deutschen Daten denjenigen aus Experiment 3, um eine methodische Konsistenz zwischen beiden Sprachen zu gewährleisten und einen direkten Vergleich der Ergebnisse zu ermöglichen.

Da das zentrale Ziel dieses Experiments darin bestand, zu untersuchen, ob die drei Sprachstrukturen im Deutschen und Chinesischen vergleichbare Effekte auf die Ereigniskonzeptualisierung aufweisen, war insbesondere der Interaktionseffekt zwischen Sprachstruktur und Sprache von zentraler Bedeutung. Für diesen

sprachübergreifenden Vergleich wurden zusätzliche separate statistische Modelle für die drei Messgrößen erstellt, in denen die Datensätze beider Sprachgruppen zusammengeführt und mit der Variable „*Sprache*“ kodiert wurden.

Die beiden unabhängigen Variablen „*Sprachstruktur*“ und „*Sprache*“ wurden mithilfe der Funktion *factor()* kodiert. Für die Variable „*Sprachstruktur*“ wurden drei Stufen, nämlich „*Hauptsatz*“, „*Relativsatz*“ und „*Nominalisierung*“, mittels der Funktion *levels()* definiert, wobei „*Hauptsatz*“ als Referenzstufe festgelegt wurde. Für die Variable „*Sprache*“ wurden mittels *levels()* zwei Stufen definiert, wobei die Stufe „*Deutsch*“ als Referenz gesetzt wurde. Der Effekt der Stufe „*Chinesisch*“ wurde somit als Differenz zur deutschen Bedingung modelliert.

Für die statistische Auswertung des sprachübergreifenden Vergleichs in Bezug auf die drei Messgrößen wurden grundsätzlich die gleichen Bayes'schen Modelle wie für die Einzelsprachen verwendet. Jedoch wurden für den sprachübergreifenden Vergleich die unabhängigen Variablen „*Sprachstruktur*“ und „*Sprache*“ sowie deren Interaktion (*Sprachstruktur* $\times$ *Sprache*) in die Modelle aufgenommen, während in den Modellen für die Einzelsprachen nur die *Sprachstruktur* als unabhängige Variable berücksichtigt wurde.

Da der Schwerpunkt des sprachübergreifenden Vergleichs darauf liegt, ob sich das Muster der Effekte der Sprachstrukturen zwischen den Sprachen unterscheidet, wurden zusätzlich Bayes'sche „Differenz-der-Differenzen“-Analysen durchgeführt, um den Interaktionseffekt *Sprachstruktur* $\times$ *Sprache* genauer auszuwerten. Konkret wurden drei Kontraste berechnet, die jeweils die Unterschiede zwischen zwei Sprachstrukturen innerhalb der beiden Sprachen miteinander vergleichen: *Hauptsatz* vs. *Nominalisierung* (Deutsch vs. Chinesisch), *Hauptsatz* vs. *Relativsatz* (Deutsch vs. Chinesisch) sowie *Nominalisierung* vs. *Relativsatz* (Deutsch vs. Chinesisch). Diese Kontraste verdeutlichen, ob sich die Abstände zwischen den Bedingungen (Sprachstrukturen) in den beiden Sprachen signifikant unterscheiden. Die relevanten posterioren Verteilungen wurden mithilfe der Funktionen *emmeans()*, *pairs()* und *contrast()*

berechnet.

Zusätzlich wurde der Haupteffekt der Sprache mithilfe paarweiser *Post-hoc*-Vergleiche analysiert, wofür die Funktionen *emmeans()* und *pairs()* verwendet wurden.

9.3. Ergebnisse

9.3.1. Ergebnisse zu den deutschen Daten

Ergebnisse zur ersten Fixation

Die Ergebnisse der Bayes'schen Analyse zur Position der ersten Fixation zeigten einen deutlichen Einfluss der Sprachstruktur auf die initiale visuelle Aufmerksamkeitslenkung. Der Effekt der Nominalisierungsbedingung im Vergleich zur Hauptsatzbedingung war signifikant positiv (Median $\beta = 0,60$, 95% CI [0,04, 1,17], $pd = 98,61\%$), was darauf hindeutete, dass Nominalisierungen die Wahrscheinlichkeit, die Patiens-/Handlungsregion (AOI 2) zuerst zu fixieren, deutlich erhöhten. Hingegen zeigte sich für die Relativsatzbedingung gegenüber der Hauptsatzbedingung kein eindeutiger Effekt, da das Konfidenzintervall den Nullpunkt enthält (Median $\beta = -0,47$, 95% CI [-1,07, 0,07], $pd = 95,77\%$), was lediglich auf einen tendenziellen Rückgang hinwies.

Die nachfolgenden paarweisen Vergleiche bestätigten diese Muster. Die Wahrscheinlichkeit, AOI 2 zuerst zu fixieren, war in der Nominalisierungsbedingung signifikant höher als in der Hauptsatzbedingung (Median $\beta = -0,60$, 95% CI [-1,19, -0,06], $pd = 98,61\%$, ROPE = 3,83%) sowie signifikant höher als in der Relativsatzbedingung (Median $\beta = 1,08$, 95% CI [0,49, 1,75], $pd = 99,99\%$, ROPE = 0%). Demgegenüber zeigte sich zwischen Hauptsatz und Relativsatz kein signifikanter Unterschied (Median $\beta = 0,47$, 95% CI [-0,07, 1,07], $pd = 95,77\%$, ROPE = 12,44%). Diese Ergebnisse legen nahe, dass insbesondere Nominalisierungen die initiale visuelle Aufmerksamkeit deutlich auf die AOI 2 lenken.

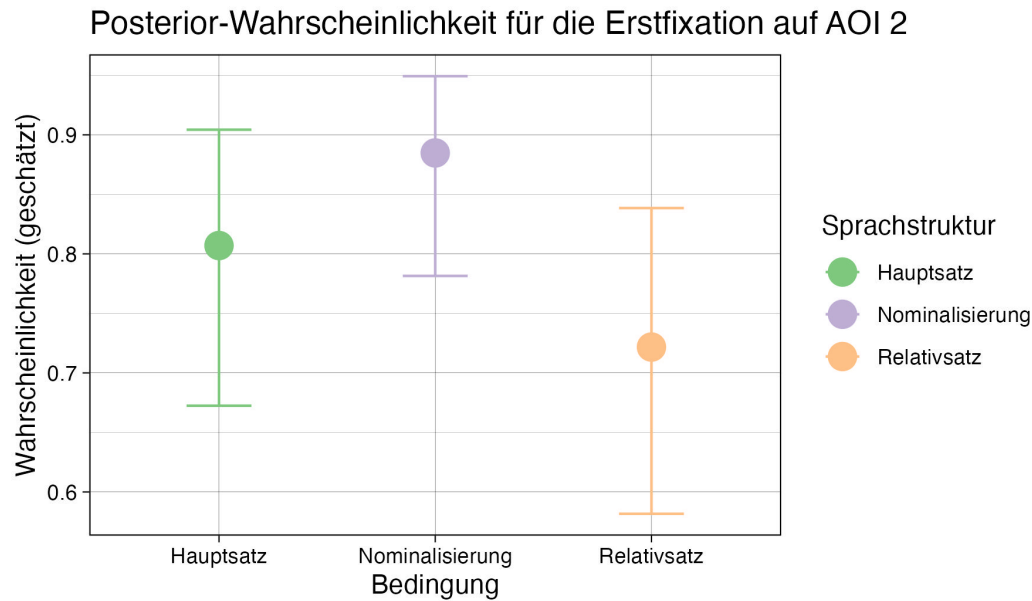


Abbildung 9.1: Posterior-Wahrscheinlichkeit für die Erstfixation auf die Patiens-/Handlungsregion (Experiment 4).

Vergleiche	Median	95% CI	pd	% in ROPE
Hauptsatz - Nominalisierung	-0,60	[-1,19, -0,06]	98,61%	3,83%
Hauptsatz - Relativsatz	0,47	[-0,07, 1,07]	95,77%	12,44%
Nominalisierung - Relativsatz	1,08	[0,49, 1,75]	99,99%	0%

Tabelle 9.1: Ergebnisse der posterioren paarweisen Vergleiche zur Erstfixation (Experiment 4).

Ergebnisse zur Fixationslatenz

Die Analyse der Fixationslatenz auf die Agens-Region (AOI 1) ergab deutliche Unterschiede in Abhängigkeit von der Sprachstruktur. Der Median des Interzepts für die Referenzbedingung „Hauptsatz“ betrug 5,92 (95% CI [5,81, 6,03], $pd = 100\%$). Im Vergleich dazu zeigte sich für die Bedingung „Nominalisierung“ eine signifikant höhere Latenz (Median $\beta = 0,24$, 95% CI [0,16, 0,33], $pd = 100\%$), was darauf hinwies, dass die erste Fixation auf die Agens-Region bei Nominalisierungen später erfolgte als bei Hauptsätzen. Für die Bedingung „Relativsatz“ ergab sich hingegen kein signifikanter Unterschied zur Referenzbedingung (Median $\beta = -0,01$, 95% CI [-0,09, 0,07], $pd = 61,69\%$).

Die *Post-hoc*-Vergleiche bestätigten diesen Befund. Die Bedingung

„Hauptsatz“ unterschied sich deutlich von der Bedingung „Nominalisierung“ (Median $\beta = -0,24$, 95% CI $[-0,33, -0,16]$, $pd = 100\%$, ROPE = 0%), während der Vergleich zwischen den Bedingungen „Hauptsatz“ und „Relativsatz“ keinen zuverlässigen Unterschied ergab (Median $\beta = 0,01$, 95% CI $[-0,07, 0,09]$, $pd = 61,69\%$, ROPE = 20,21%). Zwischen den Bedingungen „Nominalisierung“ und „Relativsatz“ wurde wiederum ein deutlicher Unterschied festgestellt (Median $\beta = 0,26$, 95% CI $[0,17, 0,35]$, $pd = 100\%$, ROPE = 0%). Dies bedeutet, dass die Erstfixation auf die Agens-Region bei Nominalisierungen signifikant später einsetzte als bei Relativsätzen und Hauptsätzen.

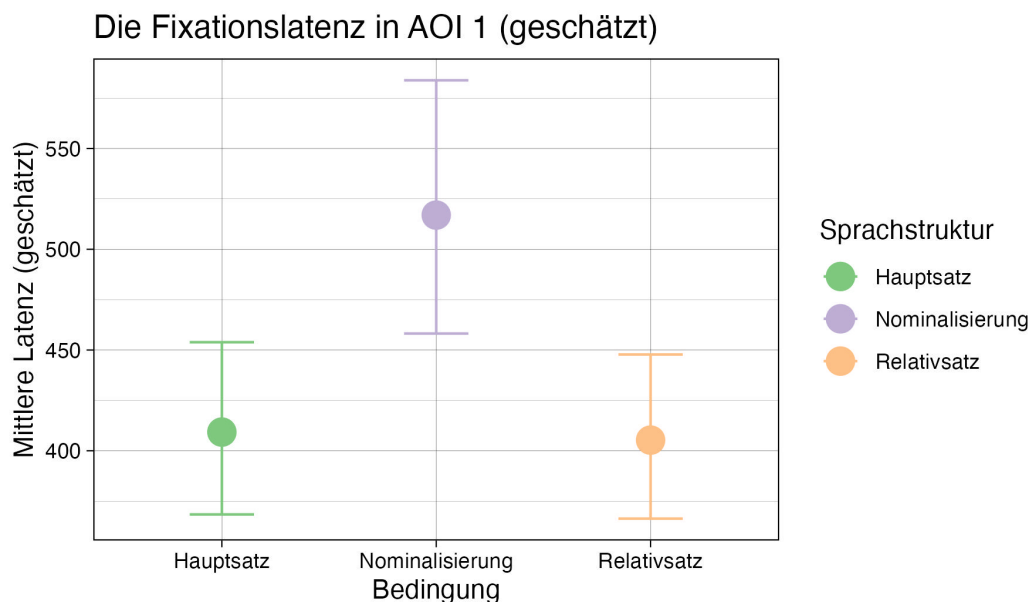


Abbildung 9.2: Geschätzte Fixationslatenz in die Agens-Region (Experiment 4).

Auch bei der Fixationslatenz in AOI 2 wurden Unterschiede zwischen den Bedingungen beobachtet. Der Median des Interzepts in der Referenzbedingung „Hauptsatz“ lag bei 5,65 (95% CI $[5,54, 5,76]$). Im Vergleich dazu wies die Bedingung „Nominalisierung“ eine tendenziell kürzere Fixationslatenz auf (Median $\beta = -0,07$, 95% CI $[-0,15, 0,00]$, $pd = 97,49\%$), ein Ergebnis, das nahe an der Grenze zur Signifikanz liegt. Für die Bedingung „Relativsatz“ ergab sich hingegen kein verlässlicher Unterschied im Vergleich zum Hauptsatz (Median $\beta = 0,05$, 95% CI $[-0,04, 0,14]$, $pd = 85,55\%$).

Die ergänzenden paarweisen *Post-hoc*-Vergleiche lieferten hier weitere Klarheit. Der Vergleich zwischen „Hauptsatz“ und „Nominalisierung“ zeigte eine kürzere Latenz in AOI 2 signifikant zugunsten der Nominalisierungsbedingung (Median $\beta = 0,07$, 95% CI [0,00, 0,15], $pd = 97,49\%$, ROPE = 2,26%). Der Vergleich zwischen „Hauptsatz“ und „Relativsatz“ blieb ohne signifikanten Unterschied (Median $\beta = -0,05$, 95% CI [-0,14, 0,04], $pd = 85,55\%$, ROPE = 9,98%). Hingegen wurde zwischen „Nominalisierung“ und „Relativsatz“ ein deutlicher Effekt festgestellt, mit einer kürzeren Latenz auf AOI 2 in der Nominalisierungsbedingung (Median $\beta = -0,12$, 95% CI [-0,21, -0,03], $pd = 99,57\%$, ROPE = 0%).

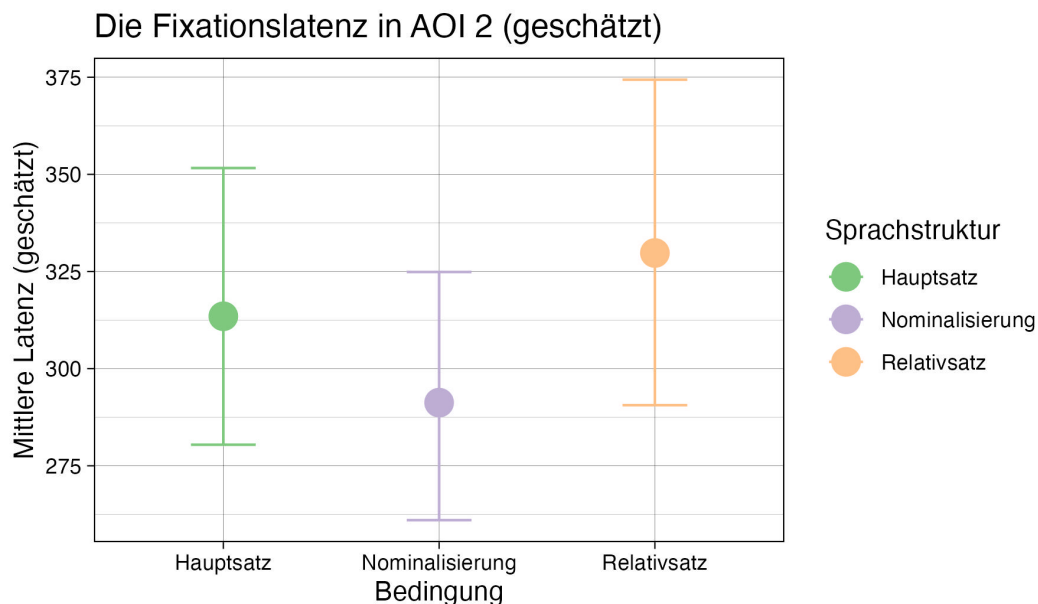


Abbildung 9.3: Geschätzte Fixationslatenz in die Patiens-/Handlungsregion (Experiment 4).

Vergleiche	Median	95% CI	pd	% in ROPE
AOI 1				
Hauptsatz - Nominalisierung	-0,24	[-0,33, -0,16]	100%	0%
Hauptsatz - Relativsatz	0,01	[-0,07, 0,09]	61,69%	20,21%
Nominalisierung - Relativsatz	0,26	[0,17, 0,35]	100%	0%
AOI 2				
Hauptsatz - Nominalisierung	0,07	[0,00, 0,15]	97,49%	2,26%
Hauptsatz - Relativsatz	-0,05	[-0,14, 0,04]	85,55%	9,98%

Nominalisierung - Relativsatz	-0,12	[-0,21, -0,03]	99,57%	0%
-------------------------------	-------	----------------	--------	----

Tabelle 9.2: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz (Experiment 4).

Ergebnisse zur Gazedauer

Die statistische Analyse zur gesamten Gazedauer ergab signifikante Effekte der Sprachstrukturen auf beide kritischen AOIs. Für AOI 1, nämlich die Agens-Region, lag der Median der Gazedauer der Hauptsatzbedingung bei 480,82 ms (95% CI [429,43, 531,36]). Im Vergleich zur Hauptsatzbedingung führten die Nominalisierungen zu einer signifikanten Verringerung der Gazedauer auf AOI 1 (Median β = -112,18 ms, 95% CI [-151,33, -71,76], pd = 100%). Dagegen zeigte sich bei Relativsätzen im Vergleich zu Hauptsätzen lediglich eine schwache Tendenz zu einer erhöhten Gazedauer auf AOI 1 (Median β = 24,25 ms, 95% CI [-8,32, 57,33], pd = 92,77%), die statistisch jedoch nicht signifikant ausfiel.

Die anschließenden paarweisen Vergleiche bestätigen diese Beobachtung. Der Unterschied zwischen Hauptsatz und Nominalisierung war hoch signifikant (Median = 112,18 ms, 95% CI [71,59, 151,18], pd = 100%, ROPE = 0%), ebenso der Unterschied zwischen Nominalisierung und Relativsatz (Median = -136,25 ms, 95% CI [-177,47, -93,57], pd = 100%, ROPE = 0%). Der Unterschied zwischen Hauptsatz und Relativsatz war hingegen nicht eindeutig und deutet auf keinen robusten Unterschied hin (Median = -24,25 ms, 95% CI [-56,81, 8,91], pd = 92,77%, ROPE = 0,00585%).

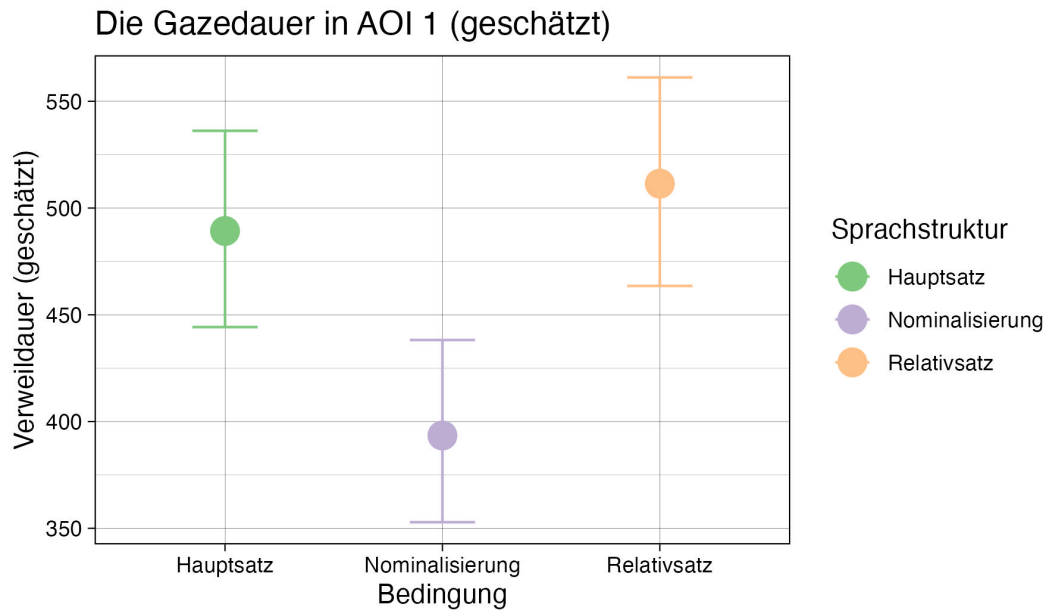


Abbildung 9.4: Geschätzte gesamte Gazedauer in der Agens-Region (Experiment 4).

Auf AOI 2 lag der Median der Gazedauer der Hauptsatzbedingung bei 624,97 ms (95% CI [580,28, 669,05]). Im Vergleich zu AOI 1 ergab sich hier ein gegenteiliges Muster. Die Nominalisierungsbedingung führte zu einer signifikanten Erhöhung der gesamten Gazedauer im Vergleich zur Hauptsatzbedingung (Median $\beta = 92,01$ ms, 95% CI [54,37, 130,54], $pd = 100\%$, ROPE = 0%). Relativsätze zeigten hingegen im Vergleich zu Hauptsätzen eine tendenziell kürzere Gazedauer auf AOI 2 (Median $\beta = -29,23$ ms, 95% CI [-62,94, 2,42], $pd = 96,09\%$, ROPE = 0,00585%), wobei dieser Unterschied nur schwach ausgeprägt war und statistisch nicht signifikant.

Die paarweisen Vergleiche bestätigten wiederum die Stärke des Effekts zwischen Hauptsatz und Nominalisierung (Median $\beta = -92,01$ ms, 95% CI [-129,00, -52,36], $pd = 100\%$) sowie zwischen Nominalisierung und Relativsatz (Median $\beta = 120,98$ ms, 95% CI [79,63, 160,32], $pd = 100\%$). Der Vergleich zwischen Hauptsatz und Relativsatz deutete hingegen auf keinen eindeutigen Unterschied hin (Median $\beta = 29,23$ ms, 95% CI [-3,10, 62,41], $pd = 96,09\%$).

Zusammenfassend zeigten die Ergebnisse, dass Nominalisierungen zu einer deutlichen Verschiebung der visuellen Aufmerksamkeit von AOI 1 zu AOI 2 führten, während Haupt- und Relativsätze eine ähnliche Verteilung der Gazedauer auf die beiden AOIs

aufwiesen.

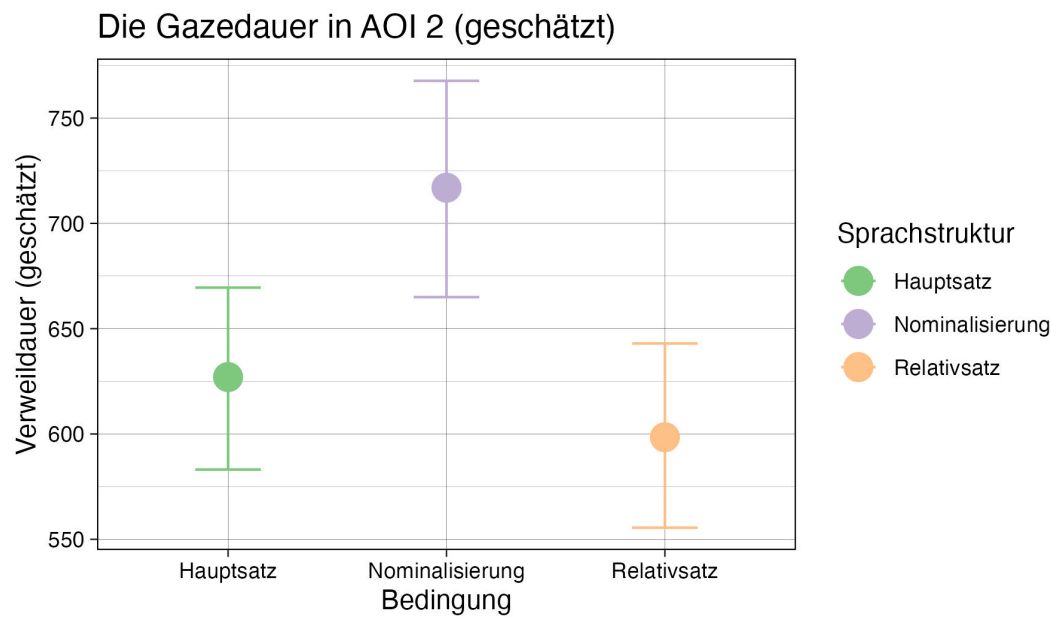


Abbildung 9.5: Geschätzte gesamte Gazedauer in der Patiens-/Handlungsregion (Experiment 4).

Vergleiche	Median	95% CI	pd	% in ROPE
AOI 1				
Hauptsatz - Nominalisierung	112,18	[71,59, 151,18]	100%	0%
Hauptsatz - Relativsatz	-24,25	[-56,81, 8,91]	92,77%	0,00585%
Nominalisierung - Relativsatz	-136,25	[-177,47, -93,57]	100%	0%
AOI 2				
Hauptsatz - Nominalisierung	-92,01	[-129,00, -52,36]	100%	0%
Hauptsatz - Relativsatz	29,23	[-3,10, 62,41]	96,09%	0,00585%
Nominalisierung - Relativsatz	120,98	[79,63, 160,32]	100%	0%

Tabelle 9.3: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer (Experiment 4).

9.3.2. Ergebnisse zum Sprachvergleich

Ergebnisse zur ersten Fixation

Die Ergebnisse zeigten zunächst signifikante Haupteffekte der Sprachstruktur sowie der Sprache. Der Median des Intercepts betrug 1,30 (95% CI [0,63, 1,94], $pd = 99,99\%$), was darauf hindeutete, dass in der Referenzbedingung (Hauptsatz, Deutsch) eine klare

Präferenz dafür besteht, AOI 2 zuerst zu fixieren.

Bezüglich des Haupteffekts der Sprache wurde festgestellt, dass chinesische Probanden AOI 2 signifikant seltener zuerst fixierten als deutsche Probanden (Median $\beta = -0,69$, 95% CI [-1,22, -0,14], $p_d = 99,28\%$, ROPE = 0,62%). Dieser Unterschied deutet darauf hin, dass bei gleicher visueller Szene und inhaltlich äquivalentem sprachlichen Stimulus die Muttersprache der Probanden einen Einfluss auf die initiale visuelle Aufmerksamkeit hat.

Die kritische Interaktion zwischen Sprachstruktur und Sprache erwies sich als nicht signifikant (*Nominalisierung* $\times$ *Sprache*: Median $\beta = -0,03$, 95% CI [-0,65, 0,56], $p_d = 54,19\%$; *Relativsatz* $\times$ *Sprache*: Median $\beta = 0,14$, 95% CI [-0,46, 0,73], $p_d = 67,92\%$). Dies spricht dafür, dass sich die Effekte der Sprachstrukturen auf die Erstfixation zwischen den beiden Sprachen nicht systematisch unterscheiden.

In Bezug auf den Haupteffekt der Sprache zeigten sich bei den *Post-hoc*-Vergleichen zwischen Chinesisch und Deutsch signifikante Unterschiede in den Bedingungen „Hauptsatz“ (Median $\beta = -0,69$, 95% CI [-1,24, -0,16], $p_d = 99,28\%$, ROPE = 0,62%) und „Nominalisierung“ (Median $\beta = -0,72$, 95% CI [-1,31, -0,15], $p_d = 99,23\%$, ROPE = 1,00%). Dies bedeutet, dass in diesen beiden Bedingungen die chinesischen Probanden AOI 2 seltener zuerst fixierten als die deutschen Probanden. Der Vergleich für den Relativsatz (Median $\beta = -0,55$, 95% CI [-1,15, 0,04], $p_d = 96,57\%$, ROPE = 8,89%) zeigte zwar eine ähnliche Tendenz, die aber wegen des überlappenden Nullpunkts im 95% CI nicht als signifikant bewertet werden konnte.

Die Ergebnisse der paarweisen Interaktionsanalyse zwischen Sprachstruktur und Sprache zeigten, dass sich kein Hinweis auf eine bedeutsame Interaktion zwischen Sprache und Sprachstruktur für den Vergleich zwischen Hauptsatz und Nominalisierung ergab (Median $\beta = -0,03$, 95% CI = [-0,64, 0,57], $p_d = 54,19\%$, ROPE = 46,95%). Auch zwischen Hauptsatz und Relativsatz in beiden Sprachen (Median $\beta = 0,14$, 95% CI = [-0,45, 0,73], $p_d = 67,92\%$, ROPE = 41,89%) ergab sich keine

überzeugende Evidenz für eine signifikante Interaktion. Die Unterschiede zwischen Nominalisierungen und Relativsätzen waren ebenso in beiden Sprachen vergleichbar (Median $\beta = 0,17$, 95% CI = [-0,44, 0,79], pd = 70,79%, ROPE = 40,13%). Diese Ergebnisse weisen darauf hin, dass die Effekte der sprachlichen Strukturen in Bezug auf die erste Fixation in beiden Sprachen ähnlich waren.

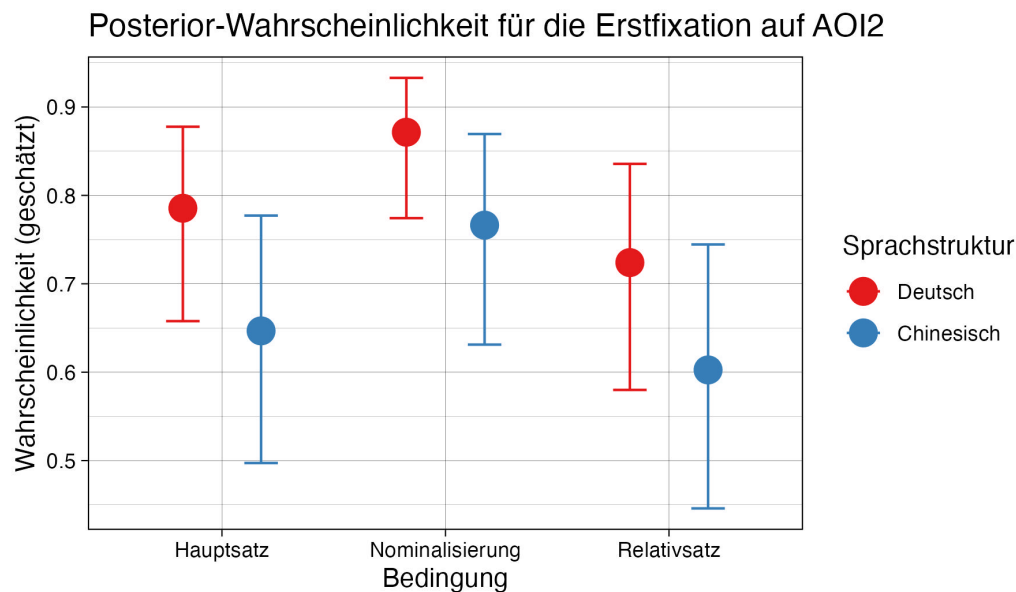


Abbildung 9.6: Posterior-Wahrscheinlichkeit für die Erstfixation auf die Patiens-/Handlungsregion im Sprachvergleich.

Vergleiche	Median	95% CI	pd	% in ROPE
Haupteffekt der Sprache (Deutsch - Chinesisch)				
Hauptsatz	0,69	[0,16, 1,24]	99,28%	0,62%
Nominalisierung	0,72	[0,15, 1,31]	99,23%	1,00%
Relativsatz	0,55	[-0,04, 1,15]	96,57%	8,89%
Interaktionseffekt (Sprachstruktur × Sprache; Deutsch - Chinesisch)				
Hauptsatz - Nominalisierung	-0,03	[-0,64, 0,57]	54,19%	46,95%
Hauptsatz - Relativsatz	0,14	[-0,45, 0,73]	67,92%	41,89%
Nominalisierung - Relativsatz	0,17	[-0,44, 0,79]	70,79%	40,13%

Tabelle 9.4: Ergebnisse der posterioren paarweisen Vergleiche zur ersten Fixation auf der Patiens-/Handlungsregion im Sprachvergleich.

Ergebnisse zur Fixationslatenz

Für AOI 1 ergab das Bayes'sche Modell einen deutlichen Haupteffekt der Sprachstruktur. Nominalisierungen wiesen eine signifikant längere Fixationslatenz auf als Hauptsätze (Median $\beta = 0,25$, 95% CI [0,16, 0,34], $pd = 100\%$, ROPE = 0%), während zwischen Relativ- und Hauptsätzen kein zuverlässiger Unterschied bestand (Median $\beta = -0,009$, 95% CI [-0,09, 0,07], $pd = 58,72\%$, ROPE = 20,40%). Ein Haupteffekt der Sprache konnte hingegen nicht nachgewiesen werden (Median $\beta = 0,007$, 95% CI [-0,11, 0,13], $pd = 54,27\%$, ROPE = 12,90%).

Ebenso waren die Interaktionseffekte zwischen Sprache und Sprachstruktur statistisch nicht signifikant. Für die Interaktion *Nominalisierung* $\times$ *Chinesisch* ergab sich ein Medianunterschied von -0,10 (95% CI [-0,21, 0,01], $pd = 95,92\%$, ROPE = 3,26%). Die Interaktion *Relativsatz* $\times$ *Chinesisch* zeigte ebenfalls keine signifikante Abweichung (Median $\beta = -0,04$, 95% CI [-0,14, 0,07], $pd = 75,92\%$, ROPE = 11,77%).

Diese Ergebnisse wurden durch die *post-hoc* paarweisen Vergleiche bestätigt. Der Sprachvergleich zwischen Deutsch und Chinesisch zeigte keine eindeutigen Unterschiede bei Hauptsätzen (Median $\beta = -0,007$, 95% CI [-0,13, 0,11], $pd = 54,27\%$, ROPE = 12,90%) und Relativsätzen (Median $\beta = 0,03$, 95% CI [-0,09, 0,15], $pd = 69,74\%$, ROPE = 12,00%). Bei den Nominalisierungen gab es zwar eine leichte Tendenz zu einer längeren Fixationslatenz im Deutschen im Vergleich zum Chinesischen (Median $\beta = 0,09$, 95% CI [-0,04, 0,22], $pd = 91,96\%$, ROPE = 4,74%), jedoch wurde auch hier keine statistische Signifikanz erreicht.

Die „Differenz-der-Differenzen“-Analysen belegten, dass der Unterschied zwischen Hauptsatz und Nominalisierung (Median $\beta = -0,10$, 95% CI [-0,21, 0,01], $pd = 95,92\%$, ROPE = 3,26%), zwischen Hauptsatz und Relativsatz (Median $\beta = -0,04$, 95% CI [-0,15, 0,07], $pd = 75,92\%$, ROPE = 11,77%) und zwischen Nominalisierung und Relativsatz (Median $\beta = 0,06$, 95% CI [-0,05, 0,17], $pd = 85,07\%$, ROPE = 8,61%) konnten keine signifikanten Unterschiede zwischen beiden Sprachen festgestellt

werden.

Diese Ergebnisse sprechen für eine robuste Verzögerung der Erstfixation auf die Agens-Region bei Nominalisierungen im Vergleich zu Haupt- und Relativsätzen, unabhängig von der Sprache. Weder für den Haupteffekt der Sprache noch für Interaktionseffekte zwischen Sprachstruktur und Sprache konnten signifikante Unterschiede festgestellt werden.

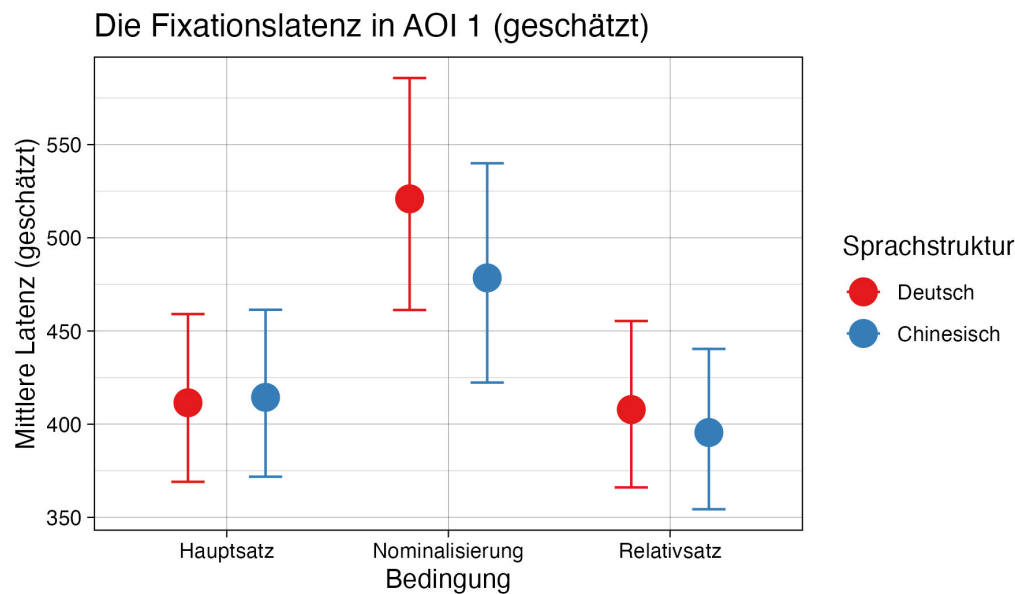


Abbildung 9.7: Geschätzte Fixationslatenz in die Agens-Region im Sprachvergleich.

Vergleiche	Median	95% CI	pd	% in ROPE
Haupteffekt der Sprache (Deutsch - Chinesisch)				
Hauptsatz	-0,007	[-0,13, 0,11]	54,27%	12,90%
Nominalisierung	0,09	[-0,04, 0,22]	91,96%	4,74%
Relativsatz	0,03	[-0,09, 0,15]	69,74%	12,00%
Interaktionseffekt (Sprachstruktur × Sprache; Deutsch - Chinesisch)				
Hauptsatz - Nominalisierung	-0,10	[-0,21, 0,01]	95,92%	3,26%
Hauptsatz - Relativsatz	-0,04	[-0,15, 0,07]	75,92%	11,77%
Nominalisierung - Relativsatz	0,06	[-0,05, 0,17]	85,07%	8,61%

Tabelle 9.5: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz in der Agens-Region im Sprachvergleich.

Für AOI 2 zeigte sich ein eindeutiger Haupteffekt der Sprache, wobei die Fixationslatenz im Chinesischen durchgehend länger war als im Deutschen (Median $\beta = 0,24$, 95% CI [0,13, 0,35], $pd = 100\%$). Hinsichtlich der Sprachstrukturen zeigte sich, dass Nominalisierungen tendenziell zu einer kürzeren Latenz der ersten Fixation führten als Hauptsätze (Median $\beta = -0,09$, 95% CI [-0,19, 0,00], $pd = 96,88\%$), während Relativsätze eine leicht längere Latenz aufwiesen als Hauptsätze, wobei dieser Unterschied nicht eindeutig war (Median $\beta = 0,06$, 95% CI [-0,05, 0,17], $pd = 85,79\%$).

Die Interaktion zwischen Sprache und Sprachstruktur zeigte keine klaren Evidenzen für einen substanziellen Unterschied. Für die Interaktion *Nominalisierung* $\times$ *Chinesisch* (Median $\beta = -0,10$, 95% CI [-0,23, 0,04], $pd = 92,74\%$) sowie *Relativsatz* $\times$ *Chinesisch* (Median $\beta = -0,08$, 95% CI [-0,23, 0,06], $pd = 86,56\%$) fanden sich keine robusten Hinweise auf einen deutlichen Effekt.

Zur detaillierten Untersuchung des Spracheffekts wurden paarweise *Post-hoc*-Vergleiche innerhalb der einzelnen Sprachstrukturen durchgeführt. Für Hauptsätze zeigte sich ein deutlicher Unterschied. Chinesisch hatte eine längere Fixationslatenz als Deutsch (Median $\beta = -0,24$, 95% CI [-0,35, -0,13], $pd = 100\%$, ROPE = 0%). Bei Nominalisierungen ergab sich ebenfalls ein klarer Unterschied zwischen den Sprachen, mit einer längeren Latenz für Chinesisch (Median $\beta = -0,14$, 95% CI [-0,25, -0,03], $pd = 99,34\%$, ROPE = 0%). Auch bei Relativsätzen fand sich eine signifikant längere Fixationslatenz für Chinesisch (Median $\beta = -0,16$, 95% CI [-0,30, -0,02], $pd = 98,63\%$, ROPE = 0%).

Zusätzlich wurden die Differenzen der Unterschiede berechnet, um zu prüfen, ob sich die Effekte der Sprachstrukturen zwischen den Sprachen unterscheiden. Diese Analyse zeigte jedoch keine Evidenzen für substanzielle Unterschiede zwischen den Strukturen zwischen Deutsch und Chinesisch. Konkret war der Unterschied zwischen Hauptsatz und Nominalisierung zwischen Deutsch und Chinesisch als nicht signifikant betrachtet

(Median $\beta = -0,10$, 95% CI $[-0,23, 0,03]$, $pd = 92,74\%$, ROPE = 4,17%). Ebenso wurden keine signifikanten Unterschiede zwischen Hauptsatz und Relativsatz (Median $\beta = -0,08$, 95% CI $[-0,23, 0,07]$, $pd = 86,56\%$, ROPE = 5,93%) sowie zwischen Nominalisierung und Relativsatz (Median $\beta = 0,02$, 95% CI $[-0,13, 0,16]$, $pd = 60,14\%$, ROPE = 11,16%) zwischen Deutsch und Chinesisch festgestellt.

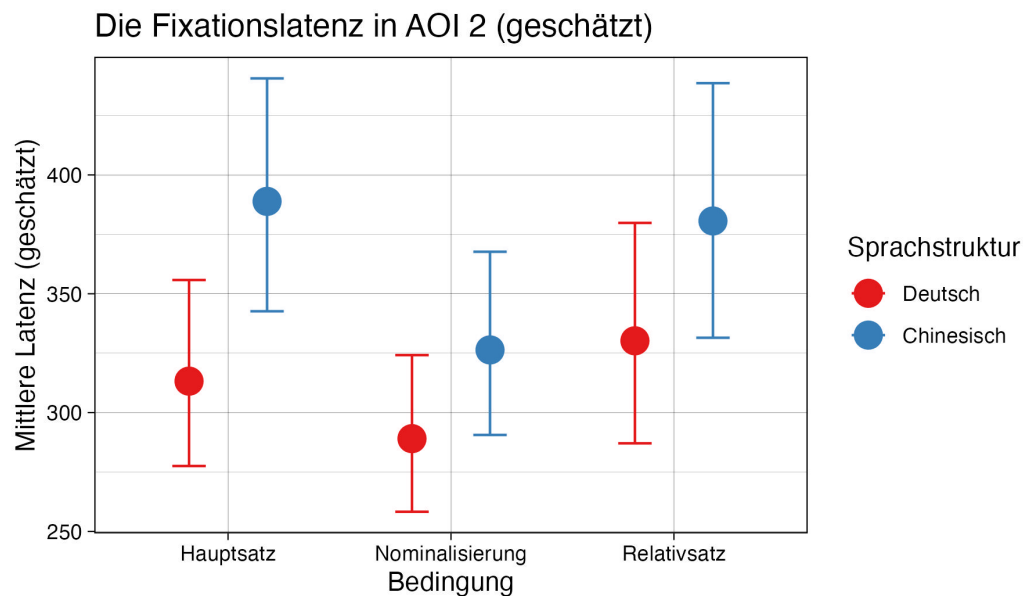


Abbildung 9.8: Geschätzte Fixationslatenz in die Patiens-/Handlungsregion im Sprachvergleich.

Vergleiche	Median	95% CI	pd	% in ROPE
Haupteffekt der Sprache (Deutsch - Chinesisch)				
Hauptsatz	-0,24	$[-0,35, -0,13]$	100%	0%
Nominalisierung	-0,14	$[-0,25, -0,03]$	99,34%	0%
Relativsatz	-0,16	$[-0,30, -0,02]$	98,63%	0%
Interaktionseffekt (Sprachstruktur $\times$ Sprache; Deutsch - Chinesisch)				
Hauptsatz - Nominalisierung	-0,10	$[-0,23, 0,03]$	92,74%	4,17%
Hauptsatz - Relativsatz	-0,08	$[-0,23, 0,07]$	86,56%	5,93%
Nominalisierung - Relativsatz	0,02	$[-0,13, 0,16]$	60,14%	11,16%

Tabelle 9.6: Ergebnisse der posterioren paarweisen Vergleiche zur Fixationslatenz in der Patiens-/Handlungsregion im Sprachvergleich.

Ergebnisse zur Gazedauer

Die Ergebnisse des Bayes'schen Modells zur gesamten Gazedauer auf AOI 1 zeigten, dass sich die beiden Sprachen nicht systematisch darin unterschieden, wie lange die Probanden in diesem Bereich verweilten. Der Haupteffekt der Sprache war sehr klein und statistisch nicht bedeutsam (Median $\beta = 6,12$ ms, 95% CI [-39,87, 53,89], $pd = 60,02\%$, ROPE = 0,04%). Die Interaktionseffekte (*Sprachstruktur* $\times$ *Sprache*) zeigten, dass sich weder für Nominalisierung (*Nominalisierung* $\times$ *Chinesisch*: $\beta = -26,11$ ms, 95 % CI [-71,65, 19,81], $pd = 86,70\%$, ROPE = 0,02%) noch für Relativsatz (*Relativsatz* $\times$ *Chinesisch*: $\beta = -1,17$ ms, 95% CI [-43,67, 40,24], $pd = 52,28\%$, ROPE = 0,03%) ein signifikanter Interaktionseffekt mit der Sprache nachweisen ließ.

Die *Post-hoc*-Vergleiche in Bezug auf den Haupteffekt der Sprache bestätigten, dass es in keiner der drei Sprachstrukturen signifikante Unterschiede zwischen Deutsch und Chinesisch bezüglich der Gazedauer auf AOI 1 gab (Hauptsatz: Median $\beta = -6,12$ ms, 95% CI [-53,19, 40,74], $pd = 60,02\%$, ROPE = 0,04%; Nominalisierung: Median $\beta = 20,01$ ms, 95% CI [-31,12, 71,14], $pd = 78,19\%$, ROPE = 0,006%; Relativsatz: Median $\beta = -5,15$ ms, 95% CI [-58,03, 48,75], $pd = 57,51\%$, ROPE = 0,03%).

Um zu prüfen, ob sich das Muster der Unterschiede zwischen den Sprachstrukturen in den beiden Sprachen unterscheidet, wurden zusätzlich die „Differenz-der-Differenzen“-Analysen durchgeführt. Insgesamt zeigten die Analysen, dass es keine signifikante Interaktion zwischen Sprache und Sprachstruktur gab. Im einzelnen paarweisen Vergleich zeigten sich folgende Ergebnisse: Zwischen Hauptsatz und Nominalisierung (Deutsch vs. Chinesisch) betrug der Median -26,11 ms (95% CI [-71,76, 19,70], $pd = 86,70\%$, ROPE = 0,02%). Zwischen Hauptsatz und Relativsatz (Deutsch vs. Chinesisch) betrug der Median -1,17 ms (95% CI [-42,92, 41,02], $pd = 52,28\%$, ROPE = 0,03%). Zwischen Nominalisierung und Relativsatz (Deutsch vs. Chinesisch) war der Median 25,26 ms (95% CI [-27,14, 76,47], $pd = 82,56\%$, ROPE = 0,02%).

Zusammenfassend ließen sich weder bei den Haupteffekten der Sprache noch bei den Interaktionen mit der Sprachstruktur relevante Unterschiede in der Gazedauer auf AOI 1 zwischen deutschen und chinesischen Probanden feststellen.

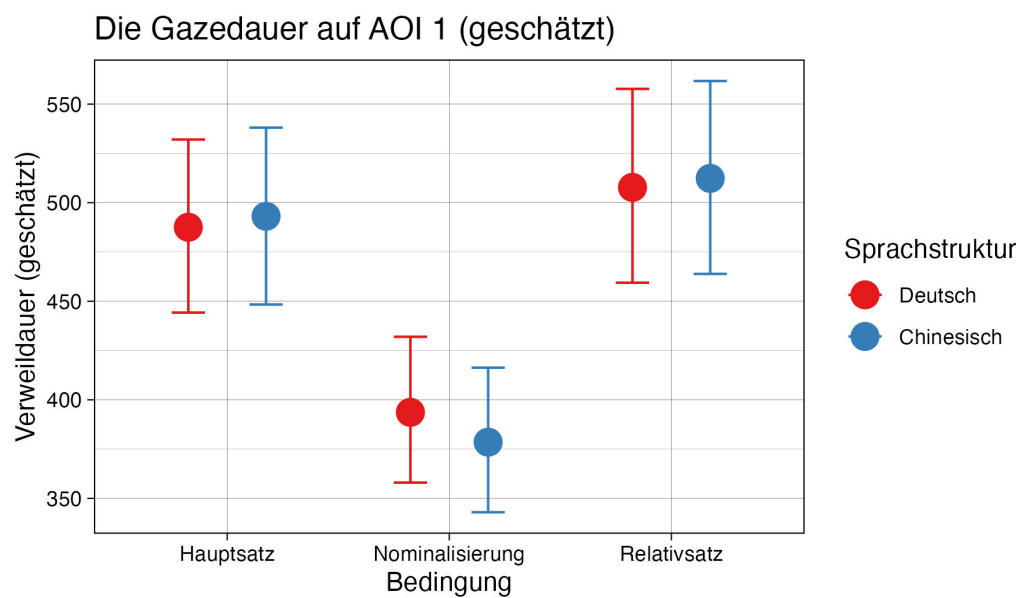


Abbildung 9.9: Geschätzte gesamte Gazedauer in der Agens-Region im Sprachvergleich.

Vergleiche	Median	95% CI	pd	% in ROPE
Haupteffekt der Sprache (Deutsch - Chinesisch)				
Hauptsatz	-6,12	[-53,19, 40,74]	60,02%	0,04%
Nominalisierung	20,01	[-31,12, 71,14]	78,19%	0,006%
Relativsatz	-5,15	[-58,03, 48,75]	57,51%	0,03%
Interaktionseffekt (Sprachstruktur × Sprache; Deutsch - Chinesisch)				
Hauptsatz - Nominalisierung	-26,11	[-71,76, 19,70]	86,70%	0,02%
Hauptsatz - Relativsatz	-1,17	[-42,92, 41,02]	52,28%	0,03%
Nominalisierung - Relativsatz	25,26	[-27,14, 76,47]	82,56%	0,02%

Tabelle 9.7: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer in der Agens-Region im Sprachvergleich.

Hinsichtlich der Gazedauer auf AOI 2, die Patiens-/Handlungsregion, zeigten die Ergebnisse deutliche Unterschiede zwischen den Sprachstrukturen. Die Gazedauer in

AOI 2 war bei Nominalisierungen im Vergleich zu Hauptsätzen signifikant länger (Median $\beta = 93,39$ ms, 95% CI [56,65, 131,20], $pd = 100\%$). Zwischen Relativsätzen und Hauptsätzen zeigte sich eine leicht kürzere Gazedauer auf AOI 2 bei Relativsätzen (Median $\beta = -25,91$ ms, 95% CI [-57,96, 3,85], $pd = 95,02\%$). Dieser Effekt war jedoch nicht signifikant.

Hinsichtlich des Haupteffekts der Sprache wurde kein signifikanter Effekt festgestellt (Median $\beta = 4,91$ ms, 95% CI [-36,00, 42,42], $pd = 59,61\%$, ROPE = 0,03%). Auch die *Post-hoc*-Vergleiche zwischen deutschen und chinesischen Probanden innerhalb jeder Sprachstruktur bestätigten, dass es keine bedeutsamen sprachlichen Unterschiede in Bezug auf die Gazedauer auf AOI 2 gab. In der Hauptsatzbedingung betrug die Differenz -4,91 ms (95% CI [-43,71, 34,83], $pd = 59,61\%$, ROPE = 0,03%), bei Nominalisierungen -27,58 ms (95% CI [-72,05, 15,88], $pd = 89,21\%$, ROPE = 0,01%) und bei Relativsätzen -25,37 ms (95% CI [-69,82, 19,86], $pd = 86,24\%$, ROPE = 0,03%). Keine dieser Differenzen war statistisch überzeugend.

Weiterhin wurden die Interaktionseffekte zwischen Sprache und Sprachstruktur paarweise geprüft, um festzustellen, ob sich die Effekte der Sprachstrukturen zwischen Deutsch und Chinesisch unterscheiden. Auch hierbei zeigten sich keine signifikanten Interaktionseffekte. Die Differenz der Unterschiede zwischen Hauptsatz und Nominalisierung zwischen Deutsch und Chinesisch betrug 22,83 ms (95% CI [-19,66, 66,61], $pd = 85,01\%$, ROPE = 0,03%), diejenige zwischen Hauptsatz und Relativsatz betrug 19,98 ms (95% CI [-21,14, 62,04], $pd = 83,06\%$, ROPE = 0,03%), und für den Vergleich zwischen Nominalisierung und Relativsatz betrug sie -2,58 ms (95% CI [-51,44, 45,46], $pd = 54,38\%$, ROPE = 0,04%). Keiner dieser Werte wies auf eine bedeutsame Interaktion hin.

Es ließ sich somit festhalten, dass die Sprache weder direkt noch in Interaktion mit den untersuchten Sprachstrukturen die Gazedauer auf beiden kritischen AOIs signifikant beeinflusste. Die beobachteten Unterschiede der Sprachstrukturen hinsichtlich ihrer Wirkung auf die Gazedauer gelten somit konsistent für beide Sprachgruppen.

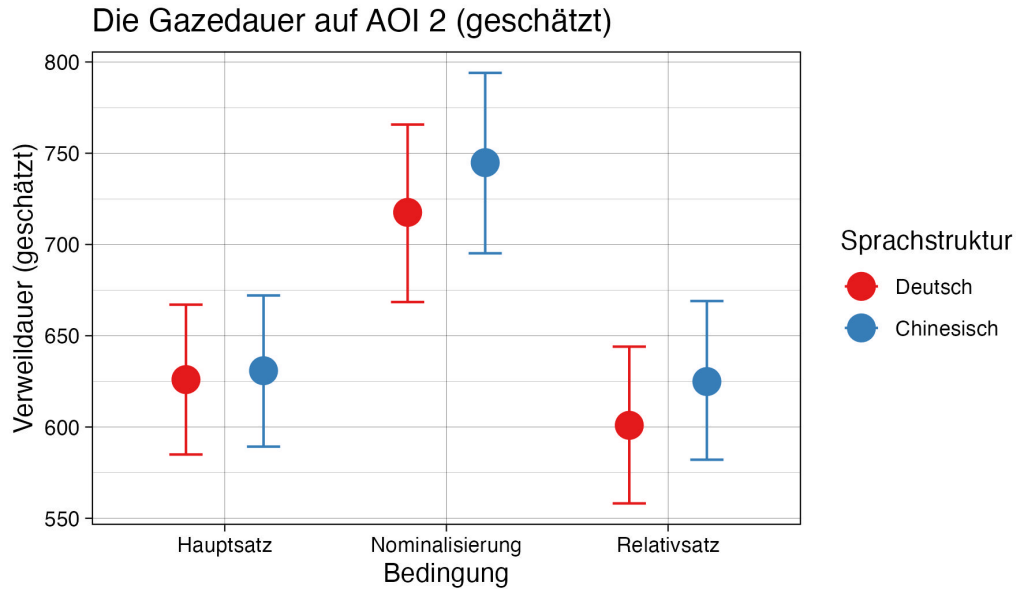


Abbildung 9.10: Geschätzte gesamte Gazedauer in der Patiens-/Handlungsregion im Sprachvergleich.

Vergleiche	Median	95% CI	pd	% in ROPE
Haupteffekt der Sprache (Deutsch - Chinesisch)				
Hauptsatz	-4,91	[-43,71, 34,83]	59,61%	0,03%
Nominalisierung	-27,58	[-72,05, 15,88]	89,21%	0,01%
Relativsatz	-25,37	[-69,82, 19,86]	86,24%	0,03%
Interaktionseffekt (Sprachstruktur × Sprache; Deutsch - Chinesisch)				
Hauptsatz - Nominalisierung	22,83	[-19,66, 66,61]	85,01%	0,03%
Hauptsatz - Relativsatz	19,98	[-21,14, 62,04]	83,06%	0,03%
Nominalisierung - Relativsatz	-2,58	[-51,44, 45,46]	54,38%	0,04%

Tabelle 9.8: Ergebnisse der posterioren paarweisen Vergleiche zur Gazedauer in der Patiens-/Handlungsregion im Sprachvergleich.

9.4. Diskussion

Das vorliegende Experiment folgte derselben methodischen Vorgehensweise wie Experiment 3 und hat zum Ziel, zu prüfen, ob die in Experiment 3 beobachteten Effekte sprachlicher Strukturen auch im Deutschen in vergleichbarer oder abweichender Form

auftreten.

Frühere sprachvergleichende Studien zeigen, dass Unterschiede in grammatischen Kategorien und Kodierungsmustern mit Unterschieden in der Ereignisverarbeitung einhergehen können. So konnten Flecken, Athanasopoulos et al. (2015) beispielsweise mittels ERP-Messungen zeigen, dass deutsche Muttersprachler bei der Wahrnehmung von Bewegungsereignissen stärker auf den Endpunkt und weniger auf die Trajektorie achteten, während dieser Unterschied bei englischen Muttersprachlern nicht beobachtet wurde. Papafragou et al. (2008) zeigen, dass englische Sprecher *Manner* und griechische Sprecher *Path* der Bewegung bevorzugt ins Blickfeld rücken (vgl. Kapitel 2).

Allerdings beschäftigten sich frühere Studien überwiegend mit Bewegungsereignissen und der Schwerpunkt lag zumeist auf der Aufmerksamkeitsverteilung bei der Sprachproduktion. Die vorliegende Arbeit geht über diesen Rahmen hinaus, indem sie komplexere linguistische Unterschiede zwischen drei sprachlichen Strukturen in Deutsch und Chinesisch bei der Konzeptualisierung kausativer Ereignisse nach dem Sprachverstehen vergleichend untersucht.

Wie bereits in Kapitel 4 diskutiert, weisen Deutsch und Chinesisch in vielerlei Hinsicht typologische Unterschiede auf, wodurch die konkrete morphologische und syntaktische Realisierung sprachlicher Strukturen in beiden Sprachen deutlich variieren kann. Dennoch bleibt die zugrunde liegende semantische Struktur der jeweiligen Sprachstruktur theoretisch sprachübergreifend stabil. Daher bietet die vorliegende sprachvergleichende Untersuchung wertvolle Einblicke in die Frage, ob die sprachliche Oberfläche die Profilierung auf konzeptueller Ebene moduliert.

Ein prototypischer deutscher einfacher aktiver transitiver Hauptsatz folgt der kanonischen SVO-Wortstellung, wobei das Subjekt typischerweise das Agens und das Objekt das Patiens ausdrückt, z.B. „*Der Mechaniker (Agens) repariert (Verb) das Auto (Patiens)*.“ Auch im Chinesischen entspricht der einfache transitive Hauptsatz dieser

grundlegenden Wortstellung sowie der Zuweisung der thematischen Rollen. Somit besitzen beide Sprachen hinsichtlich des Hauptsatzes eine ähnliche Oberflächen- und semantische Struktur.

Hauptsatz			
	Subjekt (Agens)	Verb (Handlung)	Objekt (Patiens)
Deutsch	der Mechaniker	repariert	das Auto
Chinesisch	jīxièshī	xiūlǐ	qìchē
	Mechaniker	reparieren	Auto
	„Der Mechaniker repariert das Auto“		

Tabelle 9.9: Vergleich der linearen und semantischen Struktur von Hauptsätzen im Deutschen und Chinesischen.

Ähnliches gilt für Nominalisierungen. Die deutschen Nominalisierungen, die ein kausatives Ereignis kodieren können, bilden Rektionskomposita. Sie folgen der abstrakten Konstruktion $[N \text{ (Patiens)} + N_V \text{ (Handlung)}]$ ², z.B. „die Autoreparatur“ (Patiens: Auto; Handlung: reparieren). Im Chinesischen erfolgen die Nominalisierungen über die Struktur $[N \text{ (Patiens)} + DE + V \text{ (Handlung)}]$, z.B. „汽车的修理“ (qìchē de xiūlǐ - „Auto DE Reparatur“). Diese Strukturen in beiden Sprachen weisen eine vergleichbare lineare Abfolge der konzeptuellen Komponenten auf, in der das Patiens der Handlung vorangestellt ist. In beiden Fällen bleibt zudem das Agens implizit.

Nominalisierung			
	Patiens		Handlung
Deutsch	Auto		Reparatur
Chinesisch	qìchē	de	xiūlǐ
	Auto	DE	Reparatur
	„die Autoreparatur“		

Tabelle 9.10: Vergleich der linearen und semantischen Struktur von Nominalisierungen im Deutschen

² N_V bezieht sich auf die Nominalisierung, die aus einem Verb nominalisiert.

und Chinesischen.

Besonders erwähnenswert sind Subjekt-Relativsätze. In dieser Struktur werden zwar alle zentralen Ereigniskomponenten, nämlich Agens, Patiens und Handlung, explizit benannt, jedoch in umgekehrter Reihenfolge in den beiden Sprachen. Im Deutschen steht das Kopfnomen (Agens) in der Regel vor dem Relativsyntagma (Patiens + Handlung). Im Chinesischen steht hingegen das Relativsyntagma vor dem Kopfnomen. Trotz dieser Unterschiede liegen Subjekt-Relativsätze beider Sprachen gemeinsam dem *Center-Peripher-Schema* zugrunde (vgl. Johnson, 1987; Lakoff, 1987). Das Kopfnomen bildet als semantischer und konzeptueller Kern die Figur, während das Relativsyntagma das Kopfnomen modifiziert und die Hintergrundinformationen bereitstellt. Tabelle 9.11 zeigt die Oberflächen- und die semantische Struktur des Subjekt-Relativsatzes in beiden Sprachen.

Subjekt-Relativsatz					
Deutsch	Agens (Kopfnom.)		Patiens		Handlung
	<i>Figur</i>		<i>Grund</i>		
	der Mechaniker,	der	das Auto		repariert
Chinesisch	Handlung		Patiens		Agens (Kopfnom.)
	<i>Grund</i>				<i>Figur</i>
	xiūlǐ		qìchē	de	jīxièshī
	reparieren		Auto	DE	Mechaniker
	„Der Mechaniker, der das Auto repariert“				

Tabelle 9.11: Vergleich der linearen und semantischen Struktur von Relativsätzen im Deutschen und Chinesischen.

Ausgehend von den grammatischen Eigenschaften beider Sprachen wurde in Kapitel 9.1 angenommen, dass bei Hauptsätzen und Nominalisierungen vergleichbare Effekte der sprachlichen Struktur auf die Ereigniskonzeptualisierung in beiden Sprachen zu erwarten sind, da sowohl die lineare Abfolge der Ereigniskomponenten als auch die semantische Struktur in Deutsch und Chinesisch übereinstimmen. Bei Subjekt-Relativsätzen wären, je nachdem, ob die Wortstellung selbst oder die durch die Struktur

induzierte konzeptuelle Profilierung ausschlaggebend ist, entweder unterschiedliche oder aber vergleichbare Aufmerksamkeitsmuster zu erwarten.

Die Ergebnisse des deutschen Experiments zeigen insgesamt ähnliche Muster hinsichtlich der ersten Fixation, der Fixationslatenz und der Gazedauer wie die des chinesischen Experiments. Konkret wurde bei diesem Experiment ebenfalls beobachtet, dass das Patiens und die Handlung bei Nominalisierungen eher zuerst, früher und insgesamt länger fixiert wurden als bei Haupt- und Relativsätzen. Demgegenüber wurde das Agens bei Nominalisierungen weniger wahrscheinlich zuerst, später und insgesamt kürzer fixiert als bei den anderen beiden Sprachstrukturen. Zwischen Haupt- und Relativsätzen waren ähnliche Blickbewegungsmuster festzustellen. Diese Befunde zeigen, dass sich die Effekte der drei untersuchten sprachlichen Strukturen auch im Deutschen nachweisen lassen. Die linguistisch evozierte Salienz steuert die visuelle Aufmerksamkeit und führt zu unterschiedlichen Blickbewegungsmustern.

Die statistische Analyse des sprachübergreifenden Vergleichs ergab zudem keine signifikanten Unterschiede zwischen deutschen und chinesischen Muttersprachlern bezüglich des Einflusses der Sprachstruktur auf die Verteilung der visuellen Aufmerksamkeit. Bemerkenswert ist, dass für Relativsätze in beiden Sprachgruppen eine vergleichbare Aufmerksamkeitsverteilung in Hinblick auf den Interaktionseffekt von Sprache und Sprachstruktur festgestellt wurde, obwohl diese Struktur in beiden Sprachen eine umgekehrt lineare syntaktische Reihenfolge aufweist. Bereits die ersten Fixationen innerhalb der ersten Millisekunden, die kaum durch bewusste Strategien beeinflusst sind, zeigten ähnliche Muster der Effekte von Sprachstrukturen in beiden Sprachen.

Der in Kapitel 8.4 rekonstruierte kognitive Prozess, nämlich vom Aufbau sprachlich induzierter Ereignisrepräsentationen über die Integration mit der visuellen Szene bis hin zur Planung von Blickbewegungen, bietet auch einen Rahmen für die Interpretation der Ergebnisse von Experiment 4. Unmittelbar nach dem Hören der linguistischen Stimuli konstruierten die Probanden mentale Repräsentationen des kodierten

Ereignisses. Beim anschließenden Betrachten der visuellen Szenen erfassten sie die Kernstruktur äußerst schnell. Dabei erfolgten die Planung und Ausführung der Blickbewegungen gemäß der durch die Sprachstruktur voraktivierten konzeptuellen Repräsentation, in der spezifische Ereigniskomponenten im Fokus standen, sowie der aktuellen visuellen Repräsentation. Da die visuellen Inputs in den drei linguistischen Bedingungen identisch waren, wirkten sich besonders diese sprachlich induzierten Fokusverschiebungen darauf aus, welche Bildelemente bevorzugt fixiert wurden.

Zudem waren die visuellen Stimuli in beiden Experimenten identisch, während sich die linguistischen Inputs lediglich in der linearen Reihenfolge unterschieden und der durch die jeweilige Sprachstruktur vorgegebene konzeptuelle Fokus in beiden Sprachen gleich blieb. Dennoch ergaben beide Experimente vergleichbare Effekte. Insbesondere bei Subjekt-Relativsätzen führte die unterschiedliche lineare Reihenfolge der Erwähnung nicht zu unterschiedlichen Blickbewegungsmustern. Dies liefert Evidenz dafür, dass beim *Construal* eines Ereignisses die Verarbeitung von Ereigniskomponenten nicht auf ihrer Wortstellung beruht, sondern auf ihrer konzeptuellen Salienz, die durch die Sprachstruktur induziert ist. Entsprechend wird bei der anschließenden visuellen Verarbeitung der Szene nicht einfach das zuerst Wahrgenommene zuerst verarbeitet. Die Blickbewegungen spiegeln diese sprachlich evozierte Profilierung wider. Daraus kann gefolgert werden, dass vergleichbare Konstruktionen in verschiedenen Sprachen, hier Chinesisch und Deutsch, zu ähnlichen Konzeptualisierungen und Aufmerksamkeitsmustern führen, obwohl sie sich in der Oberflächenstruktur unterscheiden.

Diese Befunde stehen im Einklang mit der Studie von Divjak et al. (2020) (vgl. Kapitel 5) und bestätigen deren Annahme, dass die Perspektiv- bzw. Fokusänderung das *Construal* eines Ereignisses beeinflusst, und zwar auch bei komplexeren sprachlichen Strukturen. Zudem erweitern die Ergebnisse diese Annahme auf eine sprachübergreifende Ebene, indem gezeigt wird, dass die durch die sprachliche Struktur induzierte Fokusänderung unabhängig von der Sprache konsistente Effekte auf die Aufmerksamkeitslenkung ausübt.

Die sprachübergreifenden Ergebnisse beantworten die zweite zentrale Fragestellung dieser Arbeit und zeigen, dass die Effekte der drei untersuchten Sprachstrukturen auf die Ereigniskonzeptualisierung in beiden Sprachen vergleichbar sind. Diese Experimente fokussieren auf die durch das Sprachverstehen entstehende Ereignisrepräsentation, wobei keine explizite Sprachproduktion erforderlich ist. Die Frage, ob die Wortstellung von Subjekt-Relativsätzen in beiden Sprachen während der Sprachproduktion einen Einfluss auf die Blickbewegungen ausübt, bleibt offen und sollte in zukünftigen Experimenten weiter untersucht werden.

Obwohl kein Interaktionseffekt zwischen Sprachstruktur und Sprache festgestellt wurde, ergab sich dennoch ein interessanter Spracheffekt bezüglich der ersten Fixation auf die Patiens-/Handlungsregion. Unabhängig von der Sprachstruktur richteten die deutschen Probanden ihre erste Fixation häufiger auf die Patiens-/Handlungsregion als die chinesischen Probanden, während entsprechend weniger Erstfixationen auf die Agens-Region entfielen. Dieser Effekt war bei Hauptsätzen und Nominalisierungen signifikant, bei Relativsätzen zeigte sich eine starke Tendenz, die jedoch nicht statistisch signifikant war.

Diese Blickverteilung ist wahrscheinlich auf die erhöhte Salienz der Handlung bzw. des Patiens im Deutschen zurückzuführen und nicht auf eine verminderte Aufmerksamkeit für das Agens. Diese Vermutung wird durch die Ergebnisse der Fixationslatenz auf beiden AOIs gestützt. Die Fixationslatenz auf die Patiens-/Handlungsregion war bei den deutschen Probanden signifikant kürzer als bei den chinesischen Probanden. Für das Agens ergaben sich dagegen keine signifikanten Unterschiede zwischen den Sprachgruppen. Dies deutet darauf hin, dass das Patiens und die Handlung im Deutschen früher und effizienter als salient erkannt und kognitiv priorisiert werden.

Dieser Spracheffekt lässt sich wahrscheinlich nicht auf die oberflächlichen sprachlichen Strukturen selbst zurückführen. Er trat unabhängig davon auf, ob die Sprachstruktur in beiden Sprachen die gleiche lineare Wortstellung (wie bei Hauptsätzen) oder eine unterschiedliche Wortstellung (wie bei Subjekt-Relativsätzen) hat. Ebenso trat er

unabhängig davon auf, ob das Agens explizit erwähnt wird (wie in Haupt- und Relativsätzen) oder implizit kodiert ist (wie in Nominalisierungen). Auch eine Erklärung über die Salienz der Ereigniskomponenten erscheint unwahrscheinlich, da die drei Sprachstrukturen unterschiedliche Komponenten hervorheben und die semantische Struktur jeweiliger Sprachstruktur in beiden Sprachen konstant bleibt.

Ein möglicher Erklärungsansatz liegt in den typologischen und grammatischen Unterschieden zwischen Deutsch und Chinesisch. Wie in Kapitel 4 erläutert, ist Deutsch eine flektierende Sprache, während Chinesisch zu den isolierenden Sprachen zählt (vgl. Greenberg, 1963). Im Deutschen tragen Verben morphologische Markierungen, die grammatische Informationen wie Person, Numerus, Tempus und Modus enkodieren. Demgegenüber verfügt das Chinesische nicht über ein Flexionssystem. Die genannten Informationen werden nicht durch morphologische Veränderungen am Verb selbst, sondern durch andere sprachliche Mittel wie Partikeln vermittelt. Dadurch liefern Verben im Deutschen mehr grammatische Informationen zur Satzstruktur als im Chinesischen.

Aufgrund dieser typologischen Unterschiede ergibt sich auch eine höhere morphologische Komplexität im Deutschen. Je nach syntaktischem Kontext verändert sich die morphologische Form der Verben. So werden Verben in Hauptsätzen wie „*Der Mechaniker **repariert** das Auto*“ und in Relativsätzen wie „*Der Mechaniker, der das Auto **repariert***“ konjugiert. In Nominalisierungen wie „*die Autoreparatur*“ erfolgt ein eindeutiger Wortartenwechsel durch Derivation. Im Gegensatz dazu bleibt im Chinesischen die morphologische Form des Verbs in allen Strukturen unverändert, z.B. im Hauptsatz „那个工人修理汽车“, im Relativsatz „那个修理汽车的工人“ und in der Nominalisierung „汽车的修理“.

Aufgrund der zusätzlichen morphologischen und grammatischen Informationen, die deutsche Verben enkodieren, könnte angenommen werden, dass die Verarbeitung dieser Informationen komplexer ist. Dies könnte dazu führen, dass bei der Dekodierung und Integration der verbbezogenen Informationen im Deutschen ein höherer kognitiver

Verarbeitungsaufwand erforderlich ist, da mehrere morphosyntaktische Merkmale extrahiert und interpretiert werden müssen. Dadurch entsteht möglicherweise eine stärkere konzeptuelle Salienz der Handlungskomponente im Deutschen im Vergleich zum Chinesischen.

Darüber hinaus bildet das Verb im Deutschen einen unverzichtbaren Bestandteil eines vollständigen Satzes und gilt als dessen Kern. Während ein vollständiges Prädikat im Deutschen zwingend ein Verb enthalten muss, können chinesische Sätze auch nicht-verbale Prädikate aufweisen. Beispielsweise kann eine Nominal- oder Adjektivphrase in prädikativer Funktion verwendet werden (vgl. Li & Thompson, 1981; Wong & Ginzburg, 2013; Zhu, 1982). Eine Korpusanalyse von Mo und Shan (1985) ergab, dass Nomen mit einer Häufigkeit von 0,18% und Adjektive mit 26,2% in prädikativer Funktion auftreten können. Wegen dieser grammatischen Unterschiede lässt sich vermuten, dass Verben im Deutschen erfahrungsbasiert eine bedeutendere Rolle für die Interpretation sprachlicher Äußerungen einnehmen als im Chinesischen. Diese sprachspezifische Erfahrung könnte sich *top-down* auf die Aufmerksamkeit auswirken, sodass im Deutschen der Handlungskomponente eines Ereignisses während der Verarbeitung eine höhere Salienz zukommt als im Chinesischen. Infolgedessen fixierten die deutschen Probanden mit höherer Wahrscheinlichkeit zuerst die Handlungsregion als die chinesischen Probanden.

Dass grammatische Eigenschaften hinsichtlich Verben die visuelle Aufmerksamkeit beeinflussen können, wurde auch von Flecken, Gerwien et al. (2015) nachgewiesen. In ihrer Studie beschrieben deutsch- und englischsprachige Versuchspersonen Ereignisse in kurzen Videoclips. Englisch, eine Aspektsprache, markiert den progressiven Aspekt grammatisch am Verb, z.B. „*A man is folding a paper airplane*“. Diese morphologische Aspektmarkierung erlaubt es, sowohl die Finitheit als auch die Spezifität eines Ereignisses direkt über das Verb auszudrücken. Im Gegensatz dazu verfügt das Deutsche als Nicht-Aspekt-Sprache über keine vergleichbare grammatische Kodierung. Verben tragen hier keine Informationen über Aspekt und Spezifität eines Ereignisses, z.B. „*Ein Mann mit einer Glatze faltet einen Papierflieger*“. Diese

grammatischen Unterschiede spiegelten sich in abweichenden Aufmerksamkeitsmustern während der präverbalen Sprachplanung wider. Beide Sprachgruppen richteten zunächst ihre Blicke auf das Agens. In einem späteren Zeitfenster von etwa 600 bis 300 ms vor Sprechbeginn zeigten englischsprachige Probanden jedoch eine erhöhte Fixation auf die Handlungsregion, während deutschsprachige Probanden das Agens weiterhin stärker fokussierten.

Eine weitere mögliche Erklärung für die beobachteten Unterschiede könnte in methodisch-technischen Faktoren liegen. Zwar wurden beide Experimente weitgehend unter vergleichbaren Bedingungen durchgeführt, einschließlich identischer Apparatur, paralleler Materialsätze, identischer Durchläufe sowie einheitlicher Datenaufbereitung, dennoch gab es einige Unterschiede, die das Ergebnis potenziell beeinflusst haben könnten. Zum einen unterscheiden sich die auditiven Stimuli. Während die chinesischen auditiven Stimuli von einer chinesischen Muttersprachlerin aufgenommen wurden, wurden die deutschen mithilfe einer synthetischen Sprach-App generiert. Trotz sorgfältiger Kontrolle der Sprechgeschwindigkeit sowie der Intonation seitens der Sprecherin und der App könnten qualitative Unterschiede in der Natürlichkeit und prosodischen Struktur der Sprachstimuli bestehen. Diese Unterschiede könnten sich möglicherweise auf das Sprachverstehen und die dadurch aktivierten mentalen Repräsentationen auswirken.

Zum anderen wurden in der Datenaufbereitung alle Trials ausgeschlossen, in denen die Beurteilungsfrage falsch beantwortet wurde oder die Aufzeichnung aufgrund technischer Artefakte – etwa durch Blinzeln, Kopfbewegungen oder Augenbewegungsstörungen – von unzureichender Qualität war. Danach wurden insgesamt mehr chinesische als deutsche Daten in die statistische Analyse einbezogen. Diese ungleiche Datenbasis könnte sich auf die Repräsentativität der Blickbewegungsmuster ausgewirkt haben.

Ein weiterer Aspekt betrifft die visuelle Gestaltung der Stimuli. Alle kritischen Bildmaterialien wurden von einer Chinesin gezeichnet und von einer chinesischen

Versuchsleiterin aufgenommen. Es ist nicht auszuschließen, dass bestimmte dargestellte Szenen, Instrumente oder Objekte den deutschen Probanden kulturell möglicherweise weniger vertraut waren als den chinesischen. Da die Probanden explizit beurteilen sollten, ob die visuelle Szene zum zuvor gehörten Audio passt, könnte diese geringere kulturelle Vertrautheit bei den deutschen Probanden zu einer verstärkten Aufmerksamkeitszuwendung insbesondere auf die Patiens-/Handlungsregion geführt haben. In dieser Region waren typischerweise mehrere visuell prüfbare Objekte enthalten, während die Agens-Region meist nur eine einzelne Person zeigte.

Da dieser Spracheffekt nicht im Mittelpunkt der vorliegenden Arbeit steht, wird an dieser Stelle nicht weiter über die Erklärungen diskutiert. Bemerkenswert ist jedoch, dass dieser Spracheffekt nur in der Hauptsatz- und Nominalisierungsbedingung statistisch signifikant war. Bei Relativsätzen war zwar ein deutlicher Trend in dieselbe Richtung zu beobachten, dieser erreichte jedoch keine statistische Signifikanz. Diese Befunde stützen die theoretische Annahme, dass im Relativsatz die im Relativsyntagma enthaltenen Informationen über Handlung und Patiens eher im Hintergrund stehen, während das Kopfnomen (das Agens) die zentrale Salienz aufweist. Diese Hervorhebung des Agens kann die visuelle Aufmerksamkeit auf Handlung und Patiens potenziell abschwächen.

Des Weiteren liefern diese unterschiedlichen Spracheffekte in Haupt- und Relativsätzen einen Hinweis darauf, dass die Hervorhebung von Agens und die gleichzeitige Unterdrückung von Patiens und Handlung in Relativsätzen stärker ausgeprägt sind als in Hauptsätzen. Die in Relativsätzen beobachtete Abschwächung der allgemeinen Tendenz deutscher Probanden, zunächst die Patiens-/Handlungsregion zu fixieren, lässt sich auf eine erhöhte konzeptuelle Salienz des Agens zurückführen. Diese erhöhte Salienz wirkt hemmend auf die visuelle Zuwendung zur Patiens-/Handlungsregion und könnte erklären, warum der ansonsten bestehende generelle Spracheffekt in dieser Bedingung nicht statistisch signifikant wurde.

10. Allgemeine Diskussion

10.1. Zusammenfassung

Die vorliegende Arbeit widmet sich der zentralen Frage, inwiefern unterschiedliche sprachliche Strukturen das *Construal* eines Ereignisses – d. h. dessen mentale Konzeptualisierung – beeinflussen. Ausgehend von Langacker (1987, 1991, 2013, 2019) wird angenommen, dass sich die Bedeutung sprachlicher Ausdrücke nicht allein aus dem konzeptuellen Inhalt ergibt, sondern auch maßgeblich durch die Art und Weise bestimmt wird, wie dieser Inhalt mental konstruiert wird. Die Wahl und Strukturierung des Inhalts sind dabei selbst Teil des *Construal*-Prozesses, der durch verschiedene Dimensionen wie Prominenz, Perspektive und Dynamik gesteuert wird.

Theoretisch wird davon ausgegangen, dass die drei untersuchten sprachlichen Strukturen auf dasselbe Referenz-Ereignis Bezug nehmen können, dabei jedoch unterschiedliche Perspektiven darauf vermitteln. Sie beruhen auf jeweils spezifischen kognitiven Mechanismen, u. a. in Bezug auf *Image Schema*, *Profiling*, *mentales Scanning* und den Grad an *Specificity* (vgl. Kapitel 3; Lakoff, 1987; Langacker, 1987, 1991, 2013, 2019). Dies erlaubt es, bestimmte Komponenten desselben Ereignisses in den Vordergrund zu rücken und konzeptuell besonders salient zu machen, während andere Aspekte defokussiert werden (vgl. Langacker, 1987, 1991). Der einfache aktive transitive Hauptsatz dient als prototypische Struktur zur Kodierung eines kanonischen kausativen Ereignisses, wobei Agens, Handlung und Patiens als eine gesamte Handlungskette profiliert sind. Subjekt-Relativsätze und Nominalisierungen sind demgegenüber in übergeordnete sprachliche Kontexte eingebettet und können dabei ein im Relativsatz bzw. in der Nominalisierung kodiertes Referenz-Ereignis mit einem übergeordneten Hauptereignis verbinden. Subjekt-Relativsätze verschieben den Fokus auf das Kopfnomen (im vorliegenden Fall das Agens), während die im Relativsyntagma kodierten Komponenten, und zwar Patiens und Handlung, in der Peripherie bleiben. Nominalisierungen profilieren hingegen Handlung und Patiens, während das Agens nur schematisch impliziert bleibt (vgl. Kapitel 3).

Vor diesem Hintergrund wurde in drei Experimenten mit chinesischen Muttersprachlern empirisch geprüft, ob Hauptsätze, Subjekt-Relativsätze und Nominalisierungen tatsächlich zu unterschiedlichen Ereignisrepräsentationen führen, indem sie spezifische Ereigniskomponenten unterschiedlich hervorheben.

Ein viertes Experiment replizierte das Design von Experiment 3 mit deutschsprachigen Versuchspersonen. Besonders im Fokus standen hierbei die Subjekt-Relativsätze, die in beiden Sprachen zwar eine vergleichbare semantische Struktur aufweisen, sich jedoch in der linearen Oberflächenstruktur unterscheiden. Dieser sprachübergreifende Vergleich erlaubt Rückschlüsse darauf, ob die visuelle Ereignisverarbeitung von der durch die Sprachstruktur induzierten konzeptuellen Profilierung oder von der linearen Wortstellung gesteuert wird.

Experiment 1 überprüfte die theoretischen Annahmen zunächst mithilfe einer *offline Picture-Selection*-Aufgabe. Das Ziel bestand darin, die linguistisch evozierte Hervorhebung einzelner Ereigniskomponenten durch die Auswahl entsprechender visueller Darstellungen sichtbar zu machen. Zu diesem Zweck wurden zu jedem kritischen Bild vier Varianten erstellt, in denen jeweils unterschiedliche Komponenten des dargestellten Ereignisses farblich hervorgehoben waren. Dadurch wurde eine direkte Assoziation zwischen sprachlicher und visueller Hervorhebung hergestellt.

Die Ergebnisse zeigten ein deutliches und konsistentes Muster. Bei einfachen aktiven transitiven Hauptsätzen wählten die Probanden überwiegend die Bildvariante, in der keine spezifische Komponente visuell hervorgehoben war. Dies deutet darauf hin, dass Hauptsätze das Ereignis enkodieren, ohne eine bestimmte Komponente besonders hervorzuheben. Bei Subjekt-Relativsätzen bevorzugten die Probanden hingegen die Variante, in der das Agens bzw. Subjekt farblich markiert war. Dies bestätigt die Annahme, dass diese Struktur linguistisch eine stärkere Hervorhebung des Kopfnomens bewirkt. Im Gegensatz dazu entschieden sich die Probanden bei Nominalisierungen am häufigsten für die Bildvariante, in der das Patiens und die Handlung visuell hervorgehoben wurden. Dieses Ergebnis stützt die theoretische Annahme, dass

Nominalisierungen insbesondere das Patiens und die Handlung betonen.

Experiment 2 knüpfte an die Ergebnisse von Experiment 1 an und überprüfte anhand einer *Picture-Sentence-Matching*-Aufgabe, ob die durch verschiedene sprachliche Strukturen evozierte Salienz einzelner Ereigniskomponenten die Verarbeitung von Ereignisrepräsentationen in Echtzeit beeinflusst. Die Hypothese lautete, dass die Probanden kürzere Reaktionszeiten zeigen sollten, wenn die linguistisch induzierten Ereignisrepräsentationen mit den aus visueller Informationsaufnahme konstruierten Repräsentationen übereinstimmen (*Match*-Bedingung), als wenn sie nicht übereinstimmen (*Mismatch*-Bedingung).

Die Ergebnisse belegen einen signifikanten Einfluss der linguistischen Struktur auf das *Construal*, indem die Versuchspersonen in der *Match*-Bedingung deutlich schneller reagierten, insbesondere bei Nominalisierungen. Dies lässt sich dadurch erklären, dass unterschiedliche sprachliche Strukturen verschiedene Ereignisrepräsentationen aktivieren, die anschließend mit den visuellen Szenen integriert werden. In der *Match*-Bedingung erleichtert die Übereinstimmung zwischen linguistisch hervorgehobener und visuell salienter Komponente die Integration, wodurch der kognitive Aufwand reduziert und damit die Reaktionszeit verkürzt wird. In der *Mismatch*-Bedingung ist dagegen eine stärkere mentale Anpassung erforderlich, um die Diskrepanz zwischen sprachlicher und visueller Hervorhebung auszugleichen. Dies erforderte mehr kognitive Ressourcen und führte zu einer längeren Reaktionszeit.

Diese beiden Experimente lieferten empirische Hinweise darauf, dass unterschiedliche sprachliche Strukturen *Construal* eines Ereignisses sowohl *offline* im Sinne von Präferenzen für bestimmte visuell saliente Komponenten als auch in Echtzeitverarbeitung im Sinne einer schnelleren Verarbeitung entsprechender visuell hervorgehobener Informationen maßgeblich modulierten. Aufbauend auf diesen Befunden widmete sich Experiment 3 der genaueren Analyse der visuellen Aufmerksamkeitsverteilung während der Verarbeitung von Ereignisrepräsentationen. Mithilfe des Eyetrackings wurde detailliert untersucht, welche Komponenten einer

dargestellten Szene von den Versuchspersonen bevorzugt fixiert wurden, wie früh diese Fixationen erfolgten und wie lange die Aufmerksamkeit auf den jeweiligen Regionen verweilte. Das Ziel bestand darin, zu überprüfen, inwiefern sich die in Experiment 1 und 2 nachgewiesenen sprachlich induzierten Hervorhebungen auch in den Blickbewegungen widerspiegeln.

Die Ergebnisse zeigten konsistente Effekte über alle drei analysierten Blickparameter hinweg. Insbesondere zeigte sich bei Nominalisierungen eine signifikante visuelle Präferenz für die Patiens-/Handlungsregion, indem sie mit höherer Wahrscheinlichkeit zuerst, früher und insgesamt länger fixiert wurde. Im Vergleich dazu lenkten Haupt- und Relativsätze die Aufmerksamkeit deutlich weniger auf diese Region.

Diese Befunde lassen sich durch die erhöhte Salienz bestimmter Ereigniskomponenten erklären, die durch die jeweilige sprachliche Struktur auf konzeptueller Ebene induziert wurden. Bei der Verarbeitung linguistischer Inputs wurden entsprechende mentale Ereignisrepräsentationen aktiviert. Dabei rücken die jeweiligen hervorgehobenen Ereigniskomponenten in den Fokus. Diese mentale Repräsentation bleibt im Arbeitsgedächtnis aktiv. Bei der anschließenden visuellen Wahrnehmung gelang es den Versuchspersonen, die zentrale Ereignisstruktur bereits innerhalb von Sekundenbruchteilen zu erfassen (vgl. Dobel et al., 2007; Griffin & Bock, 2000). Die zuvor aktivierten sprachlichen Repräsentationen, die die unterschiedlichen Profilierungen widerspiegeln, beeinflussten in der *top-down*-Weise, welche Bildregionen bevorzugt beachtet wurden. Die Daten belegen, dass unterschiedliche sprachliche Hervorhebungen die visuelle Aufmerksamkeitsverteilung beeinflussen. Insbesondere erfolgte dieser gesamte kognitive Prozess innerhalb weniger Millisekunden. Zudem unterstützen diese Ergebnisse die theoretische Annahme, dass bei Nominalisierungen das Patiens und die Handlung stärker betont werden als bei Haupt- und Subjekt-Relativsätzen.

Allerdings konnten in diesem Experiment keine signifikanten Unterschiede zwischen Haupt- und Relativsätzen festgestellt werden. Wahrscheinlich waren die

sprachstrukturell bedingten Unterschiede in der Profilierung zu subtil, um sich unter den konkreten Versuchsbedingungen deutlich in den Blickbewegungen niederzuschlagen. Zum einen teilen beide Strukturen viele zentrale Eigenschaften. Sie enthalten dasselbe Verb mit derselben Argumentstruktur, profilieren einen dynamischen Prozess und nennen alle drei Ereignisteilnehmer explizit. Zum anderen lagen Asymmetrien zwischen den Ereignisteilnehmern vor. Das Agens wurde in allen Stimuli durch belebte menschliche Figuren repräsentiert, während das Patiens überwiegend durch unbelebte Objekte dargestellt wurde. Außerdem gilt das Agens gemäß thematischer Hierarchie und zahlreicher empirischer Befunde als die prominenteste thematische Rolle und wird typischerweise als Subjekt kodiert (vgl. Dowty, 1991; Fillmore, 1968; Levin & Rappaport Hovav, 2005). Diese Merkmale begünstigen generell eine erhöhte Salienz des Agens sowohl in Hauptsätzen als auch in Relativsätzen, und können dazu geführt haben, dass die Unterschiede in der Profilierung zwischen beiden Strukturen in diesem Experiment nicht sichtbar wurden.

Bemerkenswert ist hierbei jedoch, dass das Agens in Hauptsatz- und Relativsatzstrukturen im Chinesischen an unterschiedlichen Positionen der linguistischen Oberfläche erscheint, aber dennoch vergleichbare Muster der Aufmerksamkeitsverteilung hervorruft. Dieser Befund verdeutlicht, dass die durch Sprachstrukturen evozierte konzeptuelle Salienz einzelner Ereigniskomponenten die visuelle Aufmerksamkeitsverteilung beeinflusst, nicht jedoch deren lineare syntaktische Position. Dadurch lassen sich positionsabhängige *Primacy*- oder *Recency*-Effekte als alternative Erklärungen weitgehend ausschließen.

Um weiter zu klären, ob diese beobachteten Blickbewegungsmuster auch dann auftreten, wenn die gleiche semantische Struktur aufgrund sprachspezifischer syntaktischer Strukturen in unterschiedlicher linearer Reihenfolge realisiert wird, wurde in Experiment 4 ein analoges Eyetracking-Design mit deutschen linguistischen Stimuli eingesetzt.

Die Ergebnisse von Experiment 4 zeigten, dass die Blickbewegungsmuster der

untersuchten Sprachstrukturen auch im Deutschen weitgehend mit denen im Chinesischen übereinstimmten. Nominalisierungen lenkten die Aufmerksamkeit insbesondere auf das Patiens und die Handlung. Im Vergleich dazu richteten sich die Blicke bei Haupt- und Relativsätzen weniger auf diese beiden Komponenten. Zwischen Haupt- und Relativsätzen zeigten sich wiederum keine signifikanten Unterschiede.

Diese Befunde liefern Evidenz dafür, dass unterschiedliche sprachliche Strukturen die konzeptuelle Strukturierung von Ereignissen systematisch beeinflussen, was sich konsistent in der Verteilung visueller Aufmerksamkeit widerspiegelt. Die spezifischen *Construals*, die mit bestimmten sprachlichen Strukturen einhergehen, sind bei vergleichbaren Strukturen in beiden Sprachen relativ stabil und unabhängig von typologischen Unterschieden in der linearen Oberflächenstruktur.

Zusätzlich wurde ein interessanter Spracheffekt beobachtet: Die deutschen Probanden tendierten dazu, die Patiens-/Handlungsregion früher und mit höherer Wahrscheinlichkeit zuerst zu fixieren als die chinesischen Probanden. Dies weist auf eine höhere konzeptuelle Zugänglichkeit dieser Ereigniskomponenten für die deutschen Probanden hin. Trotz dieser allgemeinen Tendenz war der Spracheffekt hinsichtlich der Erstfixation nur bei Hauptsätzen und Nominalisierungen statistisch signifikant, nicht aber bei Subjekt-Relativsätzen. Dieses Muster liefert einen empirischen Hinweis darauf, dass sich Haupt- und Relativsätze in ihrer semantischen Struktur unterscheiden. Subjekt-Relativsätze erhöhen die Salienz des Agens derart so stark, dass sie die generelle Tendenz deutscher Probanden zur bevorzugten Fokussierung auf Patiens und Handlung überlagern.

Insgesamt bietet die vorliegende Arbeit empirische Evidenz dafür, dass Sprache aktiv steuert, welche Aspekte eines Ereignisses konzeptuell hervorgehoben und entsprechend mental repräsentiert werden. Unterschiedliche sprachliche Strukturen rufen verschiedene mentale Repräsentationen desselben Ereignisses hervor und beeinflussen somit die Konzeptualisierung des Ereignisses maßgeblich. Diese Effekte konnten mithilfe verschiedener experimenteller Methoden sowie in zwei unterschiedlichen

Sprachen konsistent nachgewiesen werden. Besonders stabil zeigte sich dabei die Präferenz für Patiens und Handlung bei Nominalisierungen, sowohl im Chinesischen als auch im Deutschen und über unterschiedliche Aufgaben hinweg. Der sprachübergreifende Vergleich belegt darüber hinaus, dass die Aufmerksamkeitssteuerung nicht durch die lineare Wortstellung einer Sprache bestimmt wird. Vielmehr spielt die konzeptuelle Salienz einzelner Ereigniskomponenten eine dominante Rolle in der sprachgesteuerten kognitiven Ereignisverarbeitung.

10.2. Reflexion

Die vorliegende Dissertation konnte empirisch belegen, dass ein enger Zusammenhang zwischen Sprache und Ereigniskognition besteht. Unterschiedliche sprachliche Strukturen, nämlich Hauptsätze, Subjekt-Relativsätze und Nominalisierungen, beeinflussen die mentale Konzeptualisierung von Ereignissen, indem sie jeweils bestimmte Ereigniskomponenten in den Fokus rücken. Die beobachteten Effekte traten konsistent über verschiedene experimentelle Paradigmen hinweg sowie sprachübergreifend im Chinesischen und Deutschen auf. Zugleich ergeben sich aus der Durchführung, Konzeption und Auswertung der Experimente mehrere Anhaltspunkte für weiterführende Forschungsperspektiven und methodische Optimierungen.

Ein methodisches Desiderat betrifft die Asymmetrien zwischen den Ereignisteilnehmern in den verwendeten Stimuli. Theoretisch wurde angenommen, dass bei Hauptsätzen beide Ereignisteilnehmer, nämlich Agens und Patiens, in etwa gleichgewichtig berücksichtigt werden im Vergleich zu den anderen beiden Sprachstrukturen. In Experiment 3 und 4 zeigten sich jedoch vergleichbare Blickbewegungsmuster bei Haupt- und Relativsätzen, was darauf hindeutet, dass das Agens auch in Hauptsätzen eine erhöhte Salienz aufwies – vergleichbar mit Relativsätzen. Eine mögliche Erklärung hierfür liegt in der Belebtheit des Agens, das in den verwendeten Szenen stets durch menschliche Figuren repräsentiert wurde, während das Patiens meist unbelebte Objekte darstellte. Für zukünftige Studien wäre

es daher sinnvoll, die Belebtheit beider Ereignisteilnehmer zu manipulieren, indem sowohl das Agens als auch das Patiens entweder als belebte Figuren oder unbelebte Objekte dargestellt werden, um die Unterschiede in der Aufmerksamkeitslenkung zwischen Haupt- und Relativsätzen präziser erfassen zu können.

Ein weiteres methodisches Problem betrifft die räumliche Anordnung der beiden AOIs in Experiment 3 und 4. Dabei war das Agens weiter vom Fixationskreuz entfernt platziert, während sich die Patiens-/Handlungsregion näher zur Bildschirmmitte befand. Diese asymmetrische Anordnung könnte zu einem Positionseffekt geführt haben. Tatsächlich zeigte sich in beiden Experimenten ein Blickbewegungsbias zugunsten der Patiens-/Handlungsregion, d.h. unabhängig von der sprachlichen Struktur und Sprachgruppe wurde diese Region insgesamt früher, mit höherer Wahrscheinlichkeit zuerst und über längere Zeit fixiert als die Agens-Region. Da die Blickbewegungen typischerweise vom Fixationskreuz in der Bildschirmmitte ausgehen, ist es naheliegend, dass die näher gelegene Patiens-/Handlungsregion schneller und häufiger fixiert wurde. Eine sinnvolle methodische Verbesserung für zukünftige Studien bestünde daher darin, die Positionen von Agens und Patiens systematisch zu counterbalancieren, um positionsbedingte Verzerrungen zu vermeiden und die Effekte der sprachlichen Struktur auf die visuelle Aufmerksamkeitsverteilung zuverlässiger untersuchen zu können.

Darüber hinaus sollten die visuellen Stimuli in zukünftigen Studien stärker kontrolliert werden, da visuelle Merkmale auf niedriger Ebene einen Einfluss auf Reaktionszeiten und Blickbewegungen ausüben können. Um die natürliche Plausibilität der dargestellten Szenen zu wahren, wurde in den vorliegenden Experimenten die Größe von Objekten und Personen innerhalb der Bilder nicht normiert. Dies führte dazu, dass die Patiens-/Handlungsregion in vielen Fällen größer ausfiel als die Agens-Region. Zudem setzte sich die Patiens-/Handlungsregion immer aus zwei visuellen Elementen zusammen, nämlich dem Instrument und dem Patiens, wohingegen die Agens-Region nur eine einzelne menschliche Figur abbildete. Diese Unterschiede in visueller Größe und struktureller Komplexität könnten potenziell das Verhalten der Versuchspersonen beeinflusst haben, da größere oder komplexere Bildbereiche einen erhöhten kognitiven

Verarbeitungsaufwand erfordern. Um solchen Verzerrungen vorzubeugen und die Interpretation linguistisch induzierter Effekte zu erleichtern, erscheint es für künftige Studien unabdingbar, die visuellen Eigenschaften der AOIs wie Größe, Anzahl der Elemente und visuelle Komplexität systematisch zu kontrollieren. Dabei sollte jedoch gleichzeitig darauf geachtet werden, dass diese Kontrollen nicht zu einer unnatürlichen Darstellung führen. Nur wenn diese visuellen Faktoren über die AOIs hinweg vergleichbar sind, lassen sich beobachtete Unterschiede eindeutig auf sprachliche Strukturen zurückführen.

Ein weiterer kritischer Aspekt betrifft die Generalisierbarkeit der Ergebnisse auf beide Sprachen. Wie Boroditsky (2001) betont, besteht bei sprachvergleichenden Studien grundsätzlich das methodische Problem, dass selbst bei formal identischen Aufgaben nicht gewährleistet werden kann, dass Stimuli und Instruktionen in verschiedenen Sprachen exakt dieselbe Bedeutung transportieren. Bereits geringfügige Unterschiede in der Formulierung von Instruktionen und Stimuli können zwischen den Sprachen unterschiedliche Interpretationen hervorrufen, was möglicherweise dazu führt, dass sich Unterschiede im Verhalten der Versuchspersonen aus abweichenden semantischen Interpretationen ergeben und nicht aus tatsächlichen Unterschieden in der kognitiven Verarbeitung. Allerdings betrifft das Problem grundsätzlich alle sprachübergreifenden Studien, da eine vollständige äquivalente Bedeutungsübertragung zwischen Sprachen kaum garantiert werden kann. Um dieses Problem zu minimieren, wurde in Experiment 4 besonderer Wert auf präzise und idiomatisch natürliche Übersetzungen sämtlicher Instruktionen und linguistischer Stimuli vom Chinesischen ins Deutsche gelegt. Die Übersetzungen wurden von zwei deutschen Muttersprachlern geprüft, darunter eine Person mit umfassenden Chinesischkenntnissen, um sowohl inhaltliche Äquivalenz als auch sprachliche Natürlichkeit im Deutschen sicherzustellen.

10.3. Ausblick

In der vorliegenden Dissertation wurde untersucht, wie unterschiedliche sprachliche Strukturen die konzeptuelle Salienz spezifischer Ereigniskomponenten in der

Echtzeitverarbeitung beeinflussen. Dabei standen prototypische kausative Ereignisse mit zwei klar definierten thematischen Rollen, und zwar dem Agens und Patiens, im Mittelpunkt. Ein kausatives Ereignis wie „*Der Junge zerbrach das Glas*“ beinhaltet die Handlungskette, bei der das Agens (*der Junge*) als Quelle der Energieübertragung und das Patiens (*das Glas*) als Empfänger fungiert. Dies entspricht der prototypischen und kanonischen Form der Ereigniskodierung (vgl. Langacker, 1991).

Allerdings gibt es für dasselbe Ereignis zahlreiche sprachliche Kodierungsmöglichkeiten, die vom gewählten Umfang der Prädikation (*scope of predication*) abhängen, d. h. davon, welche Entitäten überhaupt als relevante Ereignisteilnehmer sprachlich kodiert werden. Je nach Umfang der Prädikation ergeben sich unterschiedliche Profilierungsoptionen (vgl. Langacker, 1991). So kann das zuvor genannte Ereignis beispielsweise durch Einbezug des Instruments als „*Der Junge zerbrach das Glas mit dem Hammer*“ ausgedrückt werden, wobei die Handlungskette mit drei Teilnehmern gebildet wird, nämlich *Junge* → *Hammer* → *Glas*. Ebenso könnte es durch „*Der Hammer zerbrach das Glas*“ dargestellt werden. Damit wird die Interaktion zwischen dem *Hammer* und *Glas* profiliert. Das Ereignis könnte auch als „*Das Glas wird zerbrochen*“ beschrieben werden, wobei theoretisch nur die Zustandsänderung beim Patiens profiliert wird (vgl. Langacker, 1991).

Solche sprachlichen Varianten unterscheiden sich nicht nur in ihrer syntaktischen Struktur, sondern auch hinsichtlich der konzeptuellen Perspektive. Sie ermöglichen unterschiedliche Profilierungen und beeinflussen somit potenziell deren kognitive Salienz im Prozess der Ereignisverarbeitung. Zukünftige Studien könnten daher untersuchen, wie sich weitere Kodierungsvarianten mit unterschiedlichem Prädikationsumfang bzw. Profilierung auf die Konzeptualisierung und Aufmerksamkeitssteuerung beim Sprachverstehen auswirken. Insbesondere wäre es von Interesse, inwiefern der Einbezug anderer thematischer Rollen wie Instrument, die bislang nicht berücksichtigt wurden, die linguistisch induzierte Salienz einzelner Ereigniskomponenten beeinflusst.

Ein weiterer offener Forschungsaspekt betrifft die Frage, inwieweit sich die sprachlichen Strukturen auf die Konzeptualisierung komplexerer Ereignisse auswirken, die mehr als zwei Teilnehmer mit komplexeren thematischen Rollen umfassen, wie beispielsweise Transferereignisse. Bei *Construal* solcher Ereignisse wären neben der linearen Abfolge der Teilnehmer auch deren komplexerer Fokus- und Topikstatus von besonderem Interesse, da diese Faktoren maßgeblich zur Perspektivierung des Geschehens beitragen.

Während bei kausativen Ereignissen die Aufmerksamkeitslenkung zwischen Agens und Patiens erfolgt, bieten Transferereignisse wie „*Der Junge gibt dem Mädchen den Ball*“ zusätzliche semantische und syntaktische Variationsmöglichkeiten. Solche Ereignisse umfassen mindestens drei Teilnehmer mit distinkten thematischen Rollen, und zwar Agens (Geber), Rezipient (Empfänger) und Thema (das transferierte Objekt). Es wäre beispielsweise interessant zu untersuchen, ob eine ditransitive Konstruktion („*Der Junge gibt dem Mädchen den Ball*“) andere Aufmerksamkeitsmuster hervorruft als eine komplexe Nominalisierung („*Die Übergabe des Balls an das Mädchen*“) oder eine Relativsatzkonstruktion („*Der Junge, der dem Mädchen den Ball gibt*“). Ebenso könnten Vergleiche hinsichtlich der Profilierung weiterer Ereignisteilnehmer vorgenommen werden, z.B. zwischen einer Rezipient-zentrierten Relativsatzstruktur wie „*Das Mädchen, dem der Junge den Ball gibt*“ und einer Phrase wie „*Die Übergabe des Balls an das Mädchen*“. Solche Konstruktionen erlauben die Abbildung vielfältigerer Fokus- und Topikrelationen. Es wäre sinnvoll zu untersuchen, wie diese linguistischen Kodierungen komplexerer Ereignisse, die jeweils spezifische semantische Salienz und unterschiedliche Wortstellung besitzen, die Ereignisrepräsentation beeinflussen.

Des Weiteren konzentriert sich die vorliegende Arbeit auf den Einfluss der sprachlichen Strukturen auf die Ereignisrepräsentation, die nach dem Sprachverstehen mental konstruiert wird. Da sich *Construal* und sprachliche Kodierung gegenseitig beeinflussen, wäre es in künftigen Studien sinnvoll, zu untersuchen, wie sich unterschiedliche *Construals* eines Ereignisses auf die Wahl der Ausdrucksformen

auswirken. Wie bereits dargestellt, müssen Sprecher bei der Sprachproduktion entscheiden, welche Inhalte der gesamten Situation sprachlich ausgedrückt werden sollen. Daneben müssen sie auch entscheiden, mit welchem Grad an *Specificity* dies erfolgt (vgl. Langacker, 1991). Dabei könnte gezielt geprüft werden, ob und wie sich die verschiedenen Konzeptualisierungen hinsichtlich der Auswahl, Hervorhebung und Spezifizierung der Ereignisteilnehmer in der Wahl spezifischer Sprachstrukturen niederschlagen und wie diese mit der semantischen bzw. visuellen Salienz interagieren. Solche Untersuchungen könnten vertiefende Einsichten in die kognitiven Prozesse bieten, die der Entscheidung zwischen verschiedenen linguistischen Ausdrucksmöglichkeiten zugrunde liegen, und somit das Verständnis der Wechselwirkung zwischen *Construal* und sprachlicher Kodierung erweitern.

Darüber hinaus könnten weitere Faktoren, wie z.B. die Belebtheit der Ereignisteilnehmer oder grammatische Markierungen wie Tempus in die Analyse einbezogen werden, um deren Einfluss auf die Ereigniskonzeptualisierung umfassender zu verstehen. Ferner könnte der sprachvergleichende Ansatz auf weitere Sprachen ausgeweitet werden, um zu prüfen, ob die beobachteten Effekte der spezifischen sprachlichen Strukturen im Chinesischen und Deutschen auch in anderen Sprachen auftreten.

Appendix I : Kritische Stimuli in den chinesischen Experimenten

Item01	Hauptsatz	nage banggong lengcang shiwu (那个帮工冷藏食物)
	Glossen	DemPron. Aushilfe kühlen Lebensmittel
	Übersetzung	Die Aushilfe kühlt die Lebensmittel.
	Relativsatz	nage lengcang shiwu de banggong (那个冷藏食物的帮工)
	Glossen	DemPron. kühlen Lebensmittel <i>DE</i> Aushilfe
	Übersetzung	Die Aushilfe, die die Lebensmittel kühlt.
	Nominalisierung	shiwu de lengcang (食物的冷藏)
	Glossen	Lebensmittel <i>DE</i> Kühlung
	Übersetzung	Die Lebensmittelkühlung
Item02	Hauptsatz	nage zhuli fangying huandengpian (那个助理放映幻灯片)
	Glossen	DemPron. Assistent präsentieren Dia
	Übersetzung	Der Assistent präsentiert die Dias.
	Relativsatz	nage fangying huandengpian de zhuli (那个放映幻灯片的助理)
	Glossen	DemPron. präsentieren Dias <i>DE</i> Assistent
	Übersetzung	Der Assistent, der die Dias präsentiert.
	Nominalisierung	huandengpian de fangying (幻灯片的放映)
	Glossen	Dia <i>DE</i> Präsentation
	Übersetzung	Die Diapäsentation
Item03	Hauptsatz	nage chushi qiefen shucai (那个厨师切分蔬菜)
	Glossen	DemPron. Koch schneiden Gemüse
	Übersetzung	Der Koch schneidet das Gemüse.
	Relativsatz	nage qiefen shucai de chushi (那个切分蔬菜的厨师)
	Glossen	DemPron. schneiden Gemüse <i>DE</i> Koch

Übersetzung	Der Koch, der das Gemüse schneidet.
Nominalisierung	shucaí de qiēfēn (蔬菜的切分)
Glossen	Gemüse <i>DE</i> Schneiden
Übersetzung	Das Gemüseschneiden
Item04 Hauptsatz	nàge yìshùjiā huìzhì yóuhuà (那个艺术家绘制油画)
Glossen	DemPron. Künstler malen Bild
Übersetzung	Der Künstler malt das Bild.
Relativsatz	nàge huìzhì yóuhuà de yìshùjiā (那个绘制油画的艺术家)
Glossen	DemPron. malen Bild <i>DE</i> Künstler
Übersetzung	Der Künstler, der das Bild malt.
Nominalisierung	yóuhuà de huìzhì (油画的绘制)
Glossen	Bild <i>DE</i> Malen
Übersetzung	Das Bildmalen
Item05 Hauptsatz	nàge bānyúngōng kùnbāng bāoguó (那个搬运工捆绑包裹)
Glossen	DemPron. Träger schnüren Paket
Übersetzung	Der Träger schnürt das Paket.
Relativsatz	nàge kùnbāng bāoguó de bānyúngōng (那个捆绑包裹的搬运工)
Glossen	DemPron. schnüren Paket <i>DE</i> Träger
Übersetzung	Der Träger, der das Paket schnürt.
Nominalisierung	bāoguó de kùnbāng (包裹的捆绑)
Glossen	Paket <i>DE</i> Schnürung
Übersetzung	Die Paketeschnürung
Item06 Hauptsatz	nàge gāodiǎnshī hōngbēi dāngāo (那个糕点师烘焙蛋糕)
Glossen	DemPron. Konditor backen Kuchen
Übersetzung	Der Konditor backt den Kuchen.
Relativsatz	nàge gōngbēi dāngāo de gāodiǎnshī (那个烘焙蛋糕的糕点师)
Glossen	DemPron. backen Kuchen <i>DE</i> Konditor

Übersetzung	Der Konditor, der den Kuchen backt.
Nominalisierung	dangao de hongbei (蛋糕的烘焙)
Glossen	Kuchen <i>DE</i> Backen
Übersetzung	Das Kuchenbacken
Item07 Hauptsatz	nage gongchengshi zhangliang guandao (那个工程师丈量管道)
Glossen	DemPron. Ingenieur abmessen Rohr
Übersetzung	Der Ingenieur misst das Rohr ab.
Relativsatz	nage zhangliang guandao de gongchengshi (那个丈量管道的工程师)
Glossen	DemPron. abmessen Rohr <i>DE</i> Ingenieur
Übersetzung	Der Ingenieur, der das Rohr abmisst.
Nominalisierung	guandao de zhangliang (管道的丈量)
Glossen	Rohr <i>DE</i> Abmessung
Übersetzung	Die Rohrabmessung
Item08 Hauptsatz	nage kexuejia guance liuxing (那个科学家观测流星)
Glossen	DemPron. Wissenschaftler beobachten Stern
Übersetzung	Der Wissenschaftler beobachtet Sterne.
Relativsatz	nage guance liuxing de kexuejia (那个观测流星的科学家)
Glossen	DemPron. beobachten Stern <i>DE</i> Wissenschaftler
Übersetzung	Der Wissenschaftler, der Sterne beobachtet.
Nominalisierung	liuxing de guance (流星的观测)
Glossen	Stern <i>DE</i> Beobachten
Übersetzung	Das Sternenbeobachten
Item09 Hauptsatz	nage luren tianxie wenjuan (那个路人填写问卷)
Glossen	DemPron. Fußgänger ausfüllen Fragebogen
Übersetzung	Der Fußgänger füllt den Fragebogen aus.
Relativsatz	nage tianxie wenjuan de luren (那个填写问卷的路人)

	Glossen	DemPron. ausfüllen Fragebogen <i>DE</i> Fußgänger
	Übersetzung	Der Fußgänger, der den Fragebogen ausfüllt.
	Nominalisierung	wenjuan de tianxie (问卷的填写)
	Glossen	Fragebogen <i>DE</i> Ausfüllung
	Übersetzung	Die Fragebogenausfüllung
Item10	Hauptsatz	nage famugong kanfa mucai (那个伐木工砍伐木材)
	Glossen	DemPron. Waldarbeiter fällen Baum
	Übersetzung	Der Waldarbeiter fällt den Baum.
	Relativsatz	nage kanfa mucai de famugong (那个砍伐木材的伐木工)
	Glossen	DemPron. fällen Baum <i>DE</i> Waldarbeiter
	Übersetzung	Der Waldarbeiter, der den Baum fällt.
	Nominalisierung	mucai de kanfa (木材的砍伐)
	Glossen	Baum <i>DE</i> Fällung
	Übersetzung	Die Baumfällung
Item11	Hauptsatz	nage daozei wajue caibao (那个盗贼挖掘财宝)
	Glossen	DemPron. Räuber graben Schatz
	Übersetzung	Der Räuber gräbt nach Schätzen.
	Relativsatz	nage wajue caibao de daozei (那个挖掘财宝的盗贼)
	Glossen	DemPron. graben Schatz <i>DE</i> Räuber
	Übersetzung	Der Räuber, der nach Schätzen gräbt.
	Nominalisierung	caibao de wajue (财宝的挖掘)
	Glossen	Schätzen <i>DE</i> Grabung
	Übersetzung	Die Schatzgrabung
Item12	Hauptsatz	nage huaxuejia chengliang shiji (那个化学家称量试剂)
	Glossen	DemPron. Chemiker bemessen Flüssigkeit
	Übersetzung	Der Chemiker bemisst die Flüssigkeit.
	Relativsatz	nage chengliang shiji <i>DE</i> huaxuejia (那个称量试剂的化学家)
	Glossen	DemPron. bemessen Flüssigkeit <i>DE</i> Chemiker
	Übersetzung	Der Chemiker, der die Flüssigkeit bemisst.

Nominalisierung	shiji de chengliang (试剂的称量)
Glossen	Flüssigkeit <i>DE</i> Bemessen
Übersetzung	Das Flüssigkeitsbemessen
Item13 Hauptsatz	nage ayi fengbu yifu (那个阿姨缝补衣服)
Glossen	DemPron. Tante nähen Kleid
Übersetzung	Die Tante näht das Kleid.
Relativsatz	nage fengbu yifu de ayi (那个缝补衣服的阿姨)
Glossen	DemPron. nähen Kleid <i>DE</i> Tante
Übersetzung	Die Tante, die das Kleid näht.
Nominalisierung	yifu de fengbu (衣服的缝补)
Glossen	Kleid <i>DE</i> Nähen
Übersetzung	Das Kleidernähen
Item14 Hauptsatz	nage fensi zhantie haibao (那个粉丝粘贴海报)
Glossen	DemPron. Fan aufkleben Plakat
Übersetzung	Der Fan klebt das Plakat auf.
Relativsatz	nage zhantie haibao de fensi (那个粘贴海报的粉丝)
Glossen	DemPron. aufkleben Plakat <i>DE</i> Fan
Übersetzung	Der Fan, der das Plakat aufklebt.
Nominalisierung	haibao de zhantie (海报的粘贴)
Glossen	Plakat <i>DE</i> Aufkleben
Übersetzung	Das Plakataufkleben
Item15 Hauptsatz	nage cunmin guangai nongtian (那个村民灌溉农田)
Glossen	DemPron. Landwirt bewässern Feld
Übersetzung	Der Landwirt bewässert die Felder.
Relativsatz	nage guangai nongtian de cunmin (那个灌溉农田的村民)
Glossen	DemPron. bewässern Feld <i>DE</i> Landwirt
Übersetzung	Der Landwirt, der die Felder bewässert.
Nominalisierung	nongtian de guangai (农田的灌溉)
Glossen	Feld <i>DE</i> Bewässern
Übersetzung	Das Feldbewässern

Item16 Hauptsatz	nage nüren liangshai yifu (那个女人晾晒衣服)
Glossen	DemPron. Frau aufhängen Kleidung
Übersetzung	Die Frau hängt die Kleidung auf.
Relativsatz	nage liangshai yifu de nüren (那个晾晒衣服的女人)
Glossen	DemPron. aufhängen Kleidung <i>DE</i> Frau
Übersetzung	Die Frau, die die Kleidung aufhängt.
Nominalisierung	yifu de liangshai (衣服的晾晒)
Glossen	Kleidung <i>DE</i> Aufhängen
Übersetzung	Das Kleideraufhängen
Item17 Hauptsatz	nage yanjiuyuan guancha xibao (那个研究员观察细胞)
Glossen	DemPron. Forscher beobachten Zelle
Übersetzung	Der Forscher beobachtet die Zellen.
Relativsatz	nage guancha xibao de yanjiuyuan (那个观察细胞的研究员)
Glossen	DemPron. beobachten Zelle <i>DE</i> Forscher
Übersetzung	Der Forscher, der die Zellen beobachtet.
Nominalisierung	xibao de guancha (细胞的观察)
Glossen	Zelle <i>DE</i> Beobachten
Übersetzung	Die Zellenbeobachtung
Item18 Hauptsatz	nage zhuanjia jiating wenwu (那个专家鉴定文物)
Glossen	DemPron. Experte identifizieren Relikt
Übersetzung	Der Experte identifiziert die Relikte.
Relativsatz	nage jiating wenwu de zhuanjia (那个鉴定文物的专家)
Glossen	DemPron. identifizieren Relikt <i>DE</i> Experte
Übersetzung	Der Experte, der die Relikte identifiziert.
Nominalisierung	wenwu de jiating (文物的鉴定)
Glossen	Relikt <i>DE</i> Identifizierung
Übersetzung	Die Reliktidentifizierung
Item19 Hauptsatz	nage shouhuoyuan chenlie shangpin (那个售货员陈列商)

	品)	
Glossen	DemPron. Verkäufer ausstellen Ware	
Übersetzung	Die Verkäuferin stellt die Waren aus.	
Relativsatz	nage chenlie shangpin de shouhuoyuan (那个陈列商品的售货员)	
Glossen	DemPron. ausstellen Ware <i>DE</i> Verkäufer	
Übersetzung	Die Verkäuferin, die die Waren ausstellt.	
Nominalisierung	shangpin de chenlie (商品的陈列)	
Glossen	Ware <i>DE</i> Ausstellung	
Übersetzung	Die Warenausstellung	
Item20 Hauptsatz	nage jiazhengyuan cashi canzhuo (那个家政员擦拭餐桌)	
Glossen	DemPron. Haushaltsgehilfe abwischen Tisch	
Übersetzung	Die Haushaltsgehilfin wischt den Tisch ab.	
Relativsatz	nage cashi canzhuo de jiazhengyuan (那个擦拭餐桌的家政员)	
Glossen	DemPron. abwischen Tisch <i>DE</i> Haushaltsgehilfe	
Übersetzung	Die Haushaltsgehilfin, die den Tisch abwischt.	
Nominalisierung	canzhuo de cashi (餐桌的擦拭)	
Glossen	Tisch <i>DE</i> Abwischen	
Übersetzung	Das Tischabwischen	
Item21 Hauptsatz	nage zhongdiangong qingjie diban (那个钟点工清洁地板)	
Glossen	DemPron. Putzfrau reinigen Boden	
Übersetzung	Die Putzfrau reinigt den Boden.	
Relativsatz	nage qingjie diban de zhongdiangong (那个清洁地板的钟点工)	
Glossen	DemPron. reinigen Boden <i>DE</i> Putzfrau	
Übersetzung	Die Putzfrau, die den Boden reinigt.	
Nominalisierung	diban de qingjie (地板的清洁)	
Glossen	Boden <i>DE</i> Reinigung	

	Übersetzung	Die Bodenreinigung
Item22	Hauptsatz	nage tongshi jiare fancai (那个同事加热饭菜)
	Glossen	DemPron. Kollege aufwärmen Speise
	Übersetzung	Der Kollege wärmt die Speise auf.
	Relativsatz	nage jiare fancai de tongshi (那个加热饭菜的同事)
	Glossen	DemPron. aufwärmen Speise <i>DE</i> Kollege
	Übersetzung	Der Kollege, der die Speise aufwärmt.
	Nominalisierung	fancai de jiare (饭菜的加热)
	Glossen	Speise <i>DE</i> Aufwärmen
	Übersetzung	Das Speiseaufwärmen
Item23	Hauptsatz	nage youchai toudi xinjian (那个邮差投递信件)
	Glossen	DemPron. Postbote zustellen Brief
	Übersetzung	Der Postbote stellt den Brief zu.
	Relativsatz	nage toudi xinjian de youchai (那个投递信件的邮差)
	Glossen	DemPron. zustellen Brief <i>DE</i> Postbote
	Übersetzung	Der Postbote, der den Brief zustellt.
	Nominalisierung	xinjian de toudi (信件的投递)
	Glossen	Brief <i>DE</i> Zustellung
	Übersetzung	Die Briefzustellung
Item24	Hauptsatz	nage jixieshi genghuan luntai (那个机械师更换轮胎)
	Glossen	DemPron. Mechaniker wechseln Reifen
	Übersetzung	Der Mechaniker wechselt den Reifen.
	Relativsatz	nage genghuan luntai de jixieshi (那个更换轮胎的机械师)
	Glossen	DemPron. wechseln Reifen <i>DE</i> Mechaniker
	Übersetzung	Der Mechaniker, der den Reifen wechselt.
	Nominalisierung	luntai de genghuan (轮胎的更换)
	Glossen	Reifen <i>DE</i> Wechsel
	Übersetzung	Der Reifenwechsel

Item25	Hauptsatz	nage hushi baoza shangkou (那个护士包扎伤口)
	Glossen	DemPron. Krankenschwester verbinden Wunde
	Übersetzung	Die Krankenschwester verbindet die Wunde.
	Relativsatz	nage baoza shangkou de hushi (那个包扎伤口的护士)
	Glossen	DemPron. verbinden Wunde <i>DE</i> Krankenschwester
	Übersetzung	Die Krankenschwester, die die Wunde verbindet.
	Nominalisierung	shangkou de baoza (伤口的包扎)
	Glossen	Wunde <i>DE</i> Verbindung
	Übersetzung	Die Wunde Verbindung
Item26	Hauptsatz	nage fuwuyuan dabao waimai (那个服务员打包外卖)
	Glossen	DemPron. Kellner verpacken Essen
	Übersetzung	Der Kellner verpackt das Essen.
	Relativsatz	nage dabao waimai de fuwuyuan (那个打包外卖的服务员)
	Glossen	DemPron. verpacken Essen <i>DE</i> Kellner
	Übersetzung	Der Kellner, der das Essen verpackt.
	Nominalisierung	waimai de dabao (外卖的打包)
	Glossen	Essen <i>DE</i> Verpacken
	Übersetzung	Das Essenverpacken
Item27	Hauptsatz	nage xuetu chongxi qiche (那个学徒冲洗汽车)
	Glossen	DemPron. Lehrling waschen Auto
	Übersetzung	Der Lehrling wäscht das Auto.
	Relativsatz	nage chongxi qiche de xuetu (那个冲洗汽车的学徒)
	Glossen	DemPron. waschen Auto <i>DE</i> Lehrling
	Übersetzung	Der Lehrling, der das Auto wäscht.
	Nominalisierung	qiche de chongxi (汽车的冲洗)
	Glossen	Auto <i>DE</i> Wäsche
	Übersetzung	Die Autowäsche
Item28	Hauptsatz	nage xiaofan laozhi xianbing (那个小贩烙制馅饼)
	Glossen	DemPron. Straßenverkäufer braten Pfannkuchen

Übersetzung	Die Straßenverkäuferin brät die Pfannkuchen.
Relativsatz	nage laozhi xianbing de xiaofan (那个烙制馅饼的小贩)
Glossen	DemPron. braten Pfannkuchen <i>DE</i> Straßenverkäufer
Übersetzung	Die Straßenverkäuferin, die die Pfannkuchen brät.
Nominalisierung	xianbing de laozhi (馅饼的烙制)
Glossen	Pfannkuchen <i>DE</i> Braten
Übersetzung	Das Pfannkuchenbraten
Item29 Hauptsatz	nage shibing hongzha chengbao (那个士兵轰炸城堡)
Glossen	DemPron. Soldat beschießen Schloss
Übersetzung	Der Soldat beschießt das Schloss.
Relativsatz	nage hongzha chengbao de shibing (那个轰炸城堡的士兵)
Glossen	DemPron. beschießen Schloss <i>DE</i> Soldat
Übersetzung	Der Soldat, der das Schloss beschießt.
Nominalisierung	chengbao de hongzha (城堡的轰炸)
Glossen	Schloss <i>DE</i> Beschießung
Übersetzung	Die Schlossbeschießung
Item30 Hauptsatz	nage zhuangzhu niangzao putaojiu (那个庄主酿造葡萄酒)
Glossen	DemPron. Winzer herstellen Wein
Übersetzung	Der Winzer stellt Wein her.
Relativsatz	nage niangzao putaojiu de zhuangzhu (那个酿造葡萄酒的庄主)
Glossen	DemPron. herstellen Wein <i>DE</i> Winzer
Übersetzung	Der Winzer, der Wein herstellt.
Nominalisierung	putaojiu de niangzao (葡萄酒的酿造)
Glossen	Wein <i>DE</i> Herstellung
Übersetzung	Die Weinherstellung
Item31 Hauptsatz	nage yundongyuan huaxing pihuating (那个运动员划行)

	皮划艇)	
Glossen	DemPron. Sportler fahren Kanu	
Übersetzung	Der Sportler fährt das Kanu.	
Relativsatz	nage huaxing pihuating de yundongyuan (那个划行皮划艇的运动员)	
Glossen	DemPron. fahren Kanu <i>DE</i> Sportler	
Übersetzung	Der Sportler, der das Kanu fährt.	
Nominalisierung	pihuating de huaxing (皮划艇的划行)	
Glossen	Kanu <i>DE</i> Fahren	
Übersetzung	Das Kanufahren	
Item32 Hauptsatz	nage yumin dalao shuicao (那个渔民打捞水草)	
Glossen	DemPron. Fischer entfernen Wasserpflanze	
Übersetzung	Der Fischer entfernt die Wasserpflanzen.	
Relativsatz	nage dalao shuicao de yumin (那个打捞水草的渔民)	
Glossen	DemPron. entfernen Wasserpflanze <i>DE</i> Fischer	
Übersetzung	Der Fischer, der die Wasserpflanzen entfernt.	
Nominalisierung	shuicao de dalao (水草的打捞)	
Glossen	Wasserpflanze <i>DE</i> Entfernung	
Übersetzung	Die Wasserpflanzenentfernung	
Item33 Hauptsatz	nage nongfu shouge liangshi (那个农夫收割粮食)	
Glossen	DemPron. Bauer ernten Getreide	
Übersetzung	Der Bauer erntet das Getreide.	
Relativsatz	nage shouge liangshi de nongfu (那个收割粮食的农夫)	
Glossen	DemPron. ernten Getreide <i>DE</i> Bauer	
Übersetzung	Der Bauer, der das Getreide erntet.	
Nominalisierung	liangshi de shouge (粮食的收割)	
Glossen	Getreide <i>DE</i> Ernte	
Übersetzung	Die Getreideernte	
Item34 Hauptsatz	nage nühai shouji shuiguo (那个女孩收集水果)	
Glossen	DemPron. Mädchen sammeln Frucht	

Übersetzung	Das Mädchen sammelt die Früchte.
Relativsatz	nage shouji shuiguo de nühai (那个收集水果的女孩)
Glossen	DemPron. sammeln Frucht <i>DE</i> Mädchen
Übersetzung	Das Mädchen, das die Früchte sammelt.
Nominalisierung	shuiguo de shouji (水果的收集)
Glossen	Frucht <i>DE</i> Sammlung
Übersetzung	Die Fruchtsammlung
Item35 Hauptsatz	nage huajiang jiaoguan huacao (那个花匠浇灌花草)
Glossen	DemPron. Florist gießen Blume
Übersetzung	Die Floristin gießt die Blumen.
Relativsatz	nage jiaoguan huacao de huajiang (那个浇灌花草的花匠)
Glossen	DemPron. gießen Blume <i>DE</i> Florist
Übersetzung	Die Floristin, die die Blumen gießt.
Nominalisierung	huacao de jiaoguan (花草的浇灌)
Glossen	Blume <i>DE</i> Gießen
Übersetzung	Das Blumengießen
Item36 Hauptsatz	nage mentong yunsong xingli (那个门童运送行李)
Glossen	DemPron. Portier liefern Gepäck
Übersetzung	Der Portier liefert das Gepäck.
Relativsatz	nage yunsong xingli de mentong (那个运送行李的门童)
Glossen	DemPron. liefern Gepäck <i>DE</i> Portier
Übersetzung	Der Portier, der das Gepäck liefert.
Nominalisierung	xingli de yunsong (行李的运送)
Glossen	Gepäck <i>DE</i> Lieferung
Übersetzung	Die Gepäcklieferung
Item37 Hauptsatz	nage guke zhifu zhangdan (那个顾客支付账单)
Glossen	DemPron. Kunde bezahlen Rechnung
Übersetzung	Die Kundin bezahlt die Rechnung.
Relativsatz	nage zhifu zhangdan de guke (那个支付账单的顾客)
Glossen	DemPron. bezahlen Rechnung <i>DE</i> Kunde

Übersetzung	Die Kundin, die die Rechnung bezahlt.
Nominalisierung	zhangdan de zhifu (账单的支付)
Glossen	Rechnung <i>DE</i> Zahlung
Übersetzung	Die Rechnungszahlung
Item38 Hauptsatz	nage dashi diaoke shixiang (那个大师雕刻石像)
Glossen	DemPron. Bildhauer meißeln Statue
Übersetzung	Der Bildhauer meißelt die Statue.
Relativsatz	nage diaoke shixiang de dashi (那个雕刻石像的大师)
Glossen	DemPron. meißeln Statue <i>DE</i> Bildhauer
Übersetzung	Der Bildhauer, der die Statue meißelt.
Nominalisierung	shixiang de diaoke (石像的雕刻)
Glossen	Statue <i>DE</i> Meißeln
Übersetzung	Die Statuenmeißeln
Item39 Hauptsatz	nage huanweigong qingsao luoye (那个环卫工清扫落叶)
Glossen	DemPron. Reinigungskraft fegen Laub
Übersetzung	Die Reinigungskraft fegt das Laub.
Relativsatz	nage qingsao luoye de huanweigong (那个清扫落叶的环卫工)
Glossen	DemPron. fegen Laub <i>DE</i> Reinigungskraft
Übersetzung	Die Reinigungskraft, die das Laub fegt.
Nominalisierung	luoye de qingsao (落叶的清扫)
Glossen	Laub <i>DE</i> Fegen
Übersetzung	Das Laubfegen
Item40 Hauptsatz	nage xiaofangyuan pujiu huozai (那个消防员扑救火灾)
Glossen	DemPron. Feuerwehrmann löschen Brand
Übersetzung	Der Feuerwehrmann löscht den Brand.
Relativsatz	nage pujiu huozai de xiaofangyuan (那个扑救火灾的消防员)
Glossen	DemPron. löschen Brand <i>DE</i> Feuerwehrmann
Übersetzung	Der Feuerwehrmann, der den Brand löscht.

Nominalisierung	huozai de pujiu (火灾的扑救)
Glossen	Brand <i>DE</i> Löschen
Übersetzung	Das Brandlöschen
Item41 Hauptsatz	nage yuanding xiujian shuzhi (那个园丁修剪树枝)
Glossen	DemPron. Gärtner schneiden Zweig
Übersetzung	Der Gärtner schneidet Zweige.
Relativsatz	nage xiujian shuzhi de yuanding (那个修剪树枝的园丁)
Glossen	DemPron. schneiden Zweig <i>DE</i> Gärtner
Übersetzung	Der Gärtner, der Zweige schneidet.
Nominalisierung	shuzhi de xiujian (树枝的修剪)
Glossen	Zweig <i>DE</i> Schneiden
Übersetzung	Das Zweigeschneiden
Item42 Hauptsatz	nage DJ bofang yinyue (那个 DJ 播放音乐)
Glossen	DemPron. DJ ablegen Musik
Übersetzung	Der DJ legt die Musik auf.
Relativsatz	nage bofang yinyue de DJ (那个播放音乐的 DJ)
Glossen	DemPron. ablegen Musik <i>DE</i> DJ
Übersetzung	Der DJ, der die Musik auflegt.
Nominalisierung	yinyue de bofang (音乐的播放)
Glossen	Musik <i>DE</i> Auflegen
Übersetzung	Das Musikauflegen
Item43 Hauptsatz	nage gongren hanjie jinshu (那个工人焊接金属)
Glossen	DemPron. Arbeiter schweißen Metall
Übersetzung	Der Arbeiter schweißt das Metall.
Relativsatz	nage hanjie jinshu de gongren (那个焊接金属的工人)
Glossen	DemPron. schweißen Metall <i>DE</i> Arbeiter
Übersetzung	Der Arbeiter, der das Metall schweißt.
Nominalisierung	jinshu de hanjie (金属的焊接)
Glossen	Metall <i>DE</i> Schweißen
Übersetzung	Das Metallschweißen

Item44 Hauptsatz	nage youqigong fenshua qiangbi (那个油漆工粉刷墙壁)
Glossen	DemPron. Maler anstreichen Wand
Übersetzung	Der Maler streicht die Wand an.
Relativsatz	nage fenshua qiangbi de youqigong (那个粉刷墙壁的油漆工)
Glossen	DemPron. anstreichen Wand <i>DE</i> Maler
Übersetzung	Der Maler, der die Wand anstreicht.
Nominalisierung	qiangbi de fenshua (墙壁的粉刷)
Glossen	Wand <i>DE</i> Anstreichen
Übersetzung	Der Wandanstrich
Item45 Hauptsatz	nage yongren yuntang chenshan (那个佣人熨烫衬衫)
Glossen	DemPron. Hausangestellte bügeln Hemd
Übersetzung	Die Hausangestellte bügelt das Hemd.
Relativsatz	nage yuntang chenshan de yongren (那个熨烫衬衫的佣人)
Glossen	DemPron. bügeln Hemd <i>DE</i> Hausangestellte
Übersetzung	Die Hausangestellte, die das Hemd bügelt.
Nominalisierung	chenshan de yuntang (衬衫的熨烫)
Glossen	Hemd <i>DE</i> Bügeln
Übersetzung	Das Hemdenbügeln
Item46 Hauptsatz	nage nüpu dianran lazhu (那个女仆点燃蜡烛)
Glossen	DemPron. Dienstmädchen anzünden Kerze
Übersetzung	Das Dienstmädchen zündet die Kerze an.
Relativsatz	nage dianran lazhu de nüpu (那个点燃蜡烛的女仆)
Glossen	DemPron. anzünden Kerze <i>DE</i> Dienstmädchen
Übersetzung	Das Dienstmädchen, das die Kerze anzündet.
Nominalisierung	lazhu de dianran (蜡烛的点燃)
Glossen	Kerze <i>DE</i> Anzündung
Übersetzung	Die Kerzenanzündung

Item47 Hauptsatz	nage mishu dayin ziliao (那个秘书打印资料)
Glossen	DemPron. Sekretär ausdrucken Dokument
Übersetzung	Die Sekretärin druckt die Dokumente aus.
Relativsatz	nage dayin ziliao de mishu (那个打印资料的秘书)
Glossen	DemPron. ausdrucken Dokument <i>DE</i> Sekretär
Übersetzung	Die Sekretärin, die die Dokumente ausdruckt.
Nominalisierung	ziliao de dayin (资料的打印)
Glossen	Dokument <i>DE</i> Ausdruck
Übersetzung	Der Dokumentenausdruck
Item48 Hauptsatz	nage zhuchu pengren shiwu (那个主厨烹饪食物)
Glossen	DemPron. Chef kochen Speise
Übersetzung	Der Chef kocht die Speise.
Relativsatz	nage pengren shiwu de zhuchu (那个烹饪食物的主厨)
Glossen	DemPron. kochen Speise <i>DE</i> Chef
Übersetzung	Der Chef, der die Speise kocht.
Nominalisierung	shiwu de pengren (食物的烹饪)
Glossen	Speise <i>DE</i> Kochen
Übersetzung	Das Speisekochen
Item49 Hauptsatz	nage zhufu bianzhi maoyi (那个主妇编织毛衣)
Glossen	DemPron. Hausfrau stricken Pullover
Übersetzung	Die Hausfrau strickt den Pullover.
Relativsatz	nage bianzhi maoyi de zhufu (那个编织毛衣的主妇)
Glossen	DemPron. stricken Pullover <i>DE</i> Hausfrau
Übersetzung	Die Hausfrau, die den Pullover strickt.
Nominalisierung	maoyi de bianzhi (毛衣的编织)
Glossen	Pullover <i>DE</i> Stricken
Übersetzung	Das Pulloverstricken
Item50 Hauptsatz	nage baomu qingxi yiwu (那个保姆清洗衣物)
Glossen	DemPron. Haushaltshilfe waschen Wäsche
Übersetzung	Die Haushaltshilfe wäscht die Wäsche.

Relativsatz	nage qingxi yiwu de baomu (那个清洗衣物的保姆)
Glossen	DemPron. waschen Wäsche <i>DE</i> Haushaltshilfe
Übersetzung	Die Haushaltshilfe, die die Wäsche wäscht.
Nominalisierung	yiwu de qingxi (衣物的清洗)
Glossen	Wäsche <i>DE</i> Waschen
Übersetzung	Das Wäschewaschen
Item51 Hauptsatz	nage dianyuan chongpao kafei (那个店员冲泡咖啡)
Glossen	DemPron. Barista zubereiten Kaffee
Übersetzung	Die Barista bereitet Kaffee zu.
Relativsatz	nage chongpao kafei de dianyuan (那个冲泡咖啡的店员)
Glossen	DemPron. zubereiten Kaffee <i>DE</i> Barista
Übersetzung	Die Barista, die Kaffee zubereitet.
Nominalisierung	kafei de chongpao (咖啡的冲泡)
Glossen	Kaffee <i>DE</i> Zubereitung
Übersetzung	Die Kaffeezubereitung
Item52 Hauptsatz	nage zaoxingshi shouna shipin (那个造型师收纳饰品)
Glossen	DemPron. Stylist aufräumen Schmuck
Übersetzung	Die Stylistin räumt Schmuck auf.
Relativsatz	nage shouna shipin de zaoxingshi (那个收纳饰品的造型师)
Glossen	DemPron. aufräumen Schmuck <i>DE</i> Stylist
Übersetzung	Die Stylistin, die Schmuck aufräumt.
Nominalisierung	shipin de shouna (饰品的收纳)
Glossen	Schmuck <i>DE</i> Aufräumen
Übersetzung	Das Schmuckaufräumen
Item53 Hauptsatz	nage lifashi shuli faxing (那个理发师梳理发型)
Glossen	DemPron. Friseur stylen Haar
Übersetzung	Der Friseur stylt die Haare.
Relativsatz	nage shuli faxing de lifashi (那个梳理发型的理发师)
Glossen	DemPron. stylen Haar <i>DE</i> Friseur

	Übersetzung	Der Friseur, der die Haare stylt.
	Nominalisierung	faxing de shuli (发型的梳理)
	Glossen	Haar <i>DE</i> Styling
	Übersetzung	Das Haarstyling
Item54	Hauptsatz	nage xuanshou sheji feidie (那个选手射击飞碟)
	Glossen	DemPron. Spieler schießen Tontaube
	Übersetzung	Der Spieler schießt auf Tontauben.
	Relativsatz	nage sheji feidie de xuanshou (那个射击飞碟的选手)
	Glossen	DemPron. schießen Tontaube <i>DE</i> Spieler
	Übersetzung	Der Spieler, der auf Tontauben schießt.
	Nominalisierung	feidie de sheji (飞碟的射击)
	Glossen	Tontaube <i>DE</i> Schießen
	Übersetzung	Das Tontaubenschießen

Appendix II : Kritische Stimuli im deutschen Experiment

Item01	Hauptsatz	Die Aushilfe kühlt die Lebensmittel.
	Relativsatz	Die Aushilfe, die die Lebensmittel kühlt.
	Nominalisierung	Die Lebensmittelskühlung
Item02	Hauptsatz	Der Assistent präsentiert die Dias.
	Relativsatz	Der Assistent, der die Dias präsentiert.
	Nominalisierung	Die Diapräsentation
Item03	Hauptsatz	Der Koch schneidet das Gemüse.
	Relativsatz	Der Koch, der das Gemüse schneidet.
	Nominalisierung	Das Gemüseschneiden
Item04	Hauptsatz	Der Künstler malt das Bild.
	Relativsatz	Der Künstler, der das Bild malt.
	Nominalisierung	Das Bildmalen
Item05	Hauptsatz	Der Träger schnürt das Paket.
	Relativsatz	Der Träger, der das Paket schnürt.
	Nominalisierung	Die Paketeschnürung
Item06	Hauptsatz	Der Konditor backt den Kuchen.
	Relativsatz	Der Konditor, der den Kuchen backt.
	Nominalisierung	Das Kuchenbacken
Item07	Hauptsatz	Der Ingenieur misst das Rohr ab.
	Relativsatz	Der Ingenieur, der das Rohr abmisst.
	Nominalisierung	Die Rohrabmessung
Item08	Hauptsatz	Der Wissenschaftler beobachtet Sterne.
	Relativsatz	Der Wissenschaftler, der Sterne beobachtet.
	Nominalisierung	Das Sternenbeobachten
Item09	Hauptsatz	Der Fußgänger füllt den Fragebogen aus.

	Relativsatz	Der Fußgänger, der den Fragebogen ausfüllt.
	Nominalisierung	Die Fragebogenausfüllung
Item10	Hauptsatz	Der Waldarbeiter fällt den Baum.
	Relativsatz	Der Waldarbeiter, der den Baum fällt.
	Nominalisierung	Die Baumfällung
Item11	Hauptsatz	Der Räuber gräbt nach Schätzen.
	Relativsatz	Der Räuber, der nach Schätzen gräbt.
	Nominalisierung	Die Schatzgrabung
Item12	Hauptsatz	Der Chemiker bemisst die Flüssigkeit.
	Relativsatz	Der Chemiker, der die Flüssigkeit bemisst.
	Nominalisierung	Das Flüssigkeitsbemessen
Item13	Hauptsatz	Die Oma näht das Kleid.
	Relativsatz	Die Oma, die das Kleid näht.
	Nominalisierung	Das Kleidernähen
Item14	Hauptsatz	Der Fan klebt das Plakat auf.
	Relativsatz	Der Fan, der das Plakat aufklebt.
	Nominalisierung	Das Plakataufkleben
Item15	Hauptsatz	Der Landwirt bewässert die Felder.
	Relativsatz	Der Landwirt, der die Felder bewässert.
	Nominalisierung	Das Feldbewässern
Item16	Hauptsatz	Die Nachbarin hängt die Kleidung auf.
	Relativsatz	Die Nachbarin, die die Kleidung aufhängt.
	Nominalisierung	Das Kleideraufhängen
Item17	Hauptsatz	Der Forscher beobachtet die Zellen.
	Relativsatz	Der Forscher, der die Zellen beobachtet.
	Nominalisierung	Die Zellenbeobachtung
Item18	Hauptsatz	Der Experte identifiziert die Relikte.
	Relativsatz	Der Experte, der die Relikte identifiziert.
	Nominalisierung	Die Reliktidentifizierung
Item19	Hauptsatz	Die Verkäuferin stellt die Waren aus.

	Relativsatz	Die Verkäuferin, die die Waren ausstellt.
	Nominalisierung	Die Warenausstellung
Item20	Hauptsatz	Die Haushaltsgehilfin wischt den Tisch ab.
	Relativsatz	Die Haushaltsgehilfin, die den Tisch abwischt.
	Nominalisierung	Das Tischabwischen
Item21	Hauptsatz	Die Putzfrau reinigt den Boden.
	Relativsatz	Die Putzfrau, die den Boden reinigt.
	Nominalisierung	Die Bodenreinigung
Item22	Hauptsatz	Der Kollege wärmt die Speise auf.
	Relativsatz	Der Kollege, der die Speise aufwärmt.
	Nominalisierung	Das Speiseaufwärmen
Item23	Hauptsatz	Der Postbote stellt den Brief zu.
	Relativsatz	Der Postbote, der den Brief zustellt.
	Nominalisierung	Die Briefzustellung
Item24	Hauptsatz	Der Mechaniker wechselt den Reifen.
	Relativsatz	Der Mechaniker, der den Reifen wechselt.
	Nominalisierung	Der Reifenwechsel
Item25	Hauptsatz	Die Krankenschwester verbindet die Wunde.
	Relativsatz	Die Krankenschwester, die die Wunde verbindet.
	Nominalisierung	Die Wundeverbindung
Item26	Hauptsatz	Der Kellner verpackt das Essen.
	Relativsatz	Der Kellner, der das Essen verpackt.
	Nominalisierung	Das Essenverpacken
Item27	Hauptsatz	Der Lehrling wäscht das Auto.
	Relativsatz	Der Lehrling, der das Auto wäscht.
	Nominalisierung	Die Autowäsche
Item28	Hauptsatz	Die Straßenverkäuferin brät die Pfannkuchen.
	Relativsatz	Die Straßenverkäuferin, die die Pfannkuchen brät.
	Nominalisierung	Das Pfannkuchenbraten
Item29	Hauptsatz	Der Soldat beschießt das Schloss.

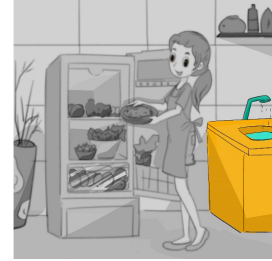
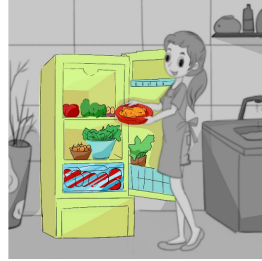
	Relativsatz	Der Soldat, der das Schloss beschießt.
	Nominalisierung	Die Schlossbeschießung
Item30	Hauptsatz	Der Winzer stellt Wein her.
	Relativsatz	Der Winzer, der Wein herstellt.
	Nominalisierung	Die Weinherstellung
Item31	Hauptsatz	Der Sportler fährt das Kanu.
	Relativsatz	Der Sportler, der das Kanu fährt.
	Nominalisierung	Das Kanufahren
Item32	Hauptsatz	Der Fischer entfernt die Wasserpflanzen.
	Relativsatz	Der Fischer, der die Wasserpflanzen entfernt.
	Nominalisierung	Die Wasserpflanzenentfernung
Item33	Hauptsatz	Der Bauer erntet das Getreide.
	Relativsatz	Der Bauer, der das Getreide erntet.
	Nominalisierung	Die Getreideernte
Item34	Hauptsatz	Das Mädchen sammelt die Früchte.
	Relativsatz	Das Mädchen, das die Früchte sammelt.
	Nominalisierung	Die Fruchtsammlung
Item35	Hauptsatz	Die Floristin gießt die Blumen.
	Relativsatz	Die Floristin, die die Blumen gießt.
	Nominalisierung	Das Blumengießen
Item36	Hauptsatz	Der Portier liefert das Gepäck.
	Relativsatz	Der Portier, der das Gepäck liefert.
	Nominalisierung	Die Gepäcklieferung
Item37	Hauptsatz	Die Kundin bezahlt die Rechnung.
	Relativsatz	Die Kundin, die die Rechnung bezahlt.
	Nominalisierung	Die Rechnungszahlung
Item38	Hauptsatz	Der Bildhauer meißelt die Statue.
	Relativsatz	Der Bildhauer, der die Statue meißelt.
	Nominalisierung	Die Statuenmeißeln
Item39	Hauptsatz	Die Reinigungskraft fegt das Laub.

	Relativsatz	Die Reinigungskraft, die das Laub fegt.
	Nominalisierung	Das Laubfegen
Item40	Hauptsatz	Der Feuerwehrmann löscht den Brand.
	Relativsatz	Der Feuerwehrmann, der den Brand löscht.
	Nominalisierung	Das Brandlöschen
Item41	Hauptsatz	Der Gärtner schneidet Zweige.
	Relativsatz	Der Gärtner, der Zweige schneidet.
	Nominalisierung	Das Zweigeschneiden
Item42	Hauptsatz	Der DJ legt die Musik auf.
	Relativsatz	Der DJ, der die Musik auflegt.
	Nominalisierung	Das Musikauflegen
Item43	Hauptsatz	Der Arbeiter schweißt das Metall.
	Relativsatz	Der Arbeiter, der das Metall schweißt.
	Nominalisierung	Das Metallschweißen
Item44	Hauptsatz	Der Maler streicht die Wand an.
	Relativsatz	Der Maler, der die Wand anstreicht.
	Nominalisierung	Der Wandanstrich
Item45	Hauptsatz	Die Hausangestellte bügelt das Hemd.
	Relativsatz	Die Hausangestellte, die das Hemd bügelt.
	Nominalisierung	Das Hemdenbügeln
Item46	Hauptsatz	Das Dienstmädchen zündet die Kerze an.
	Relativsatz	Das Dienstmädchen, das die Kerze anzündet.
	Nominalisierung	Die Kerzenanzündung
Item47	Hauptsatz	Die Sekretärin druckt die Dokumente aus.
	Relativsatz	Die Sekretärin, die die Dokumente ausdruckt.
	Nominalisierung	Der Dokumentenausdruck
Item48	Hauptsatz	Der Chef kocht die Speise.
	Relativsatz	Der Chef, der die Speise kocht.
	Nominalisierung	Das Speisekochen
Item49	Hauptsatz	Die Hausfrau strickt den Pullover.

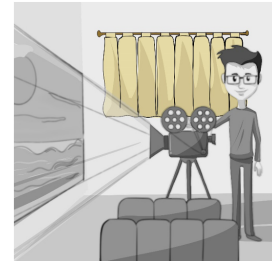
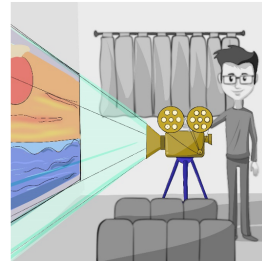
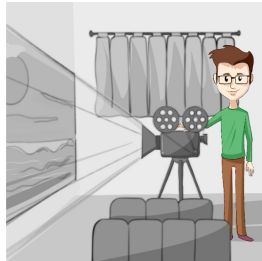
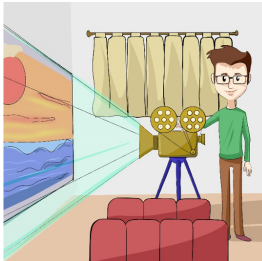
	Relativsatz	Die Hausfrau, die den Pullover strickt.
	Nominalisierung	Das Pulloverstricken
Item50	Hauptsatz	Die Haushaltshilfe wäscht die Wäsche.
	Relativsatz	Die Haushaltshilfe, die die Wäsche wäscht.
	Nominalisierung	Das Wäschewaschen
Item51	Hauptsatz	Die Barista bereitet Kaffee zu.
	Relativsatz	Die Barista, die Kaffee zubereitet.
	Nominalisierung	Die Kaffeezubereitung
Item52	Hauptsatz	Die Stylistin räumt Schmuck auf.
	Relativsatz	Die Stylistin, die Schmuck aufräumt.
	Nominalisierung	Das Schmuckaufräumen
Item53	Hauptsatz	Der Friseur stylt die Haare.
	Relativsatz	Der Friseur, der die Haare stylt.
	Nominalisierung	Das Haarstyling
Item54	Hauptsatz	Der Spieler schießt auf Tontauben.
	Relativsatz	Der Spieler, der auf Tontauben schießt.
	Nominalisierung	Das Tontaubenschießen

Appendix III : Visuelle Stimuli

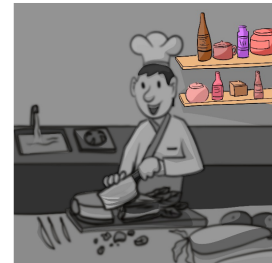
Item01:



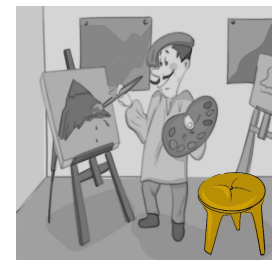
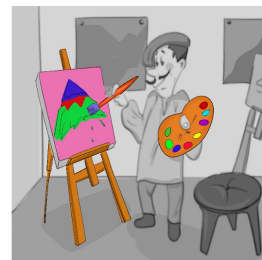
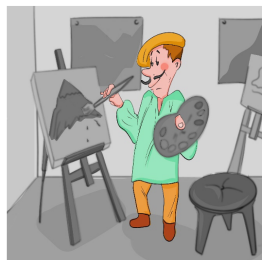
Item02:



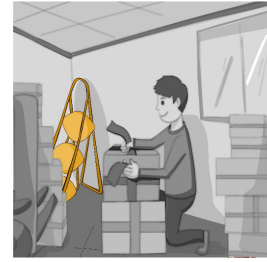
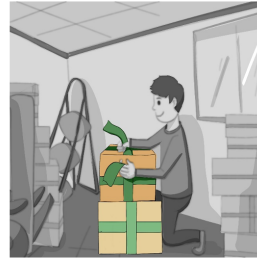
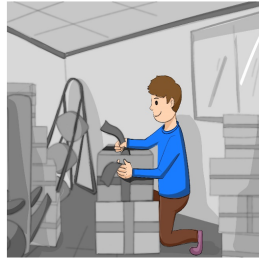
Item03:



Item04:



Item05:



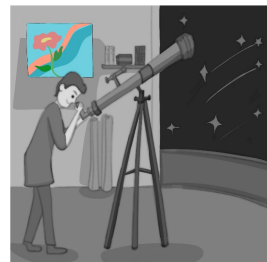
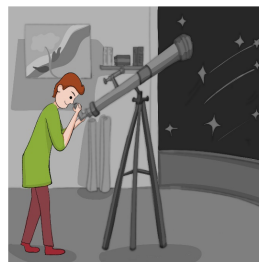
Item06:



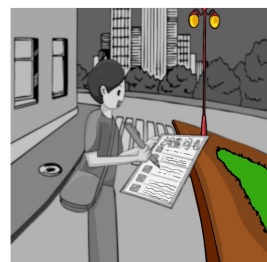
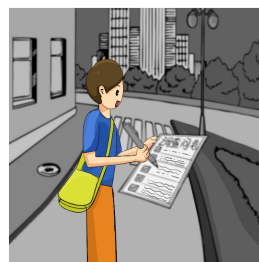
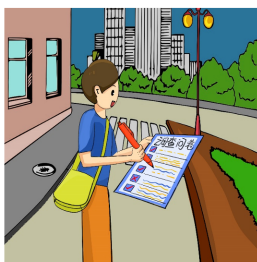
Item07:



Item08:



Item09:



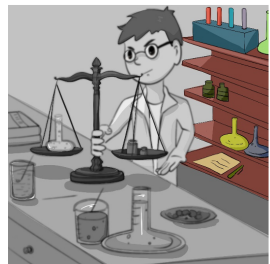
Item10:



Item11:



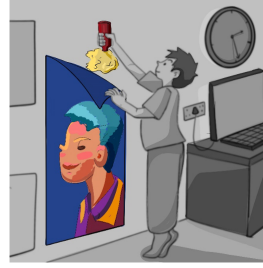
Item12:



Item13:



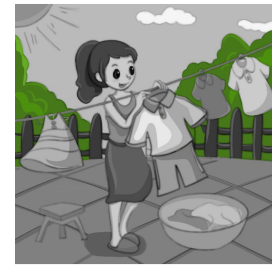
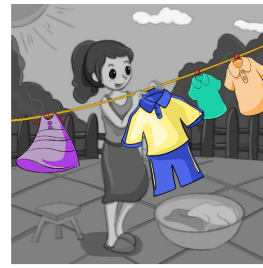
Item14:



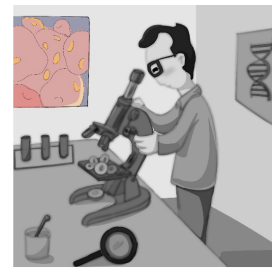
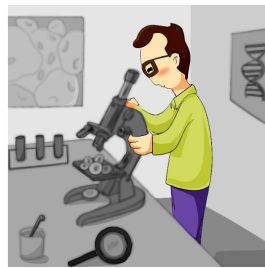
Item15:



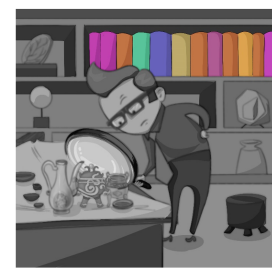
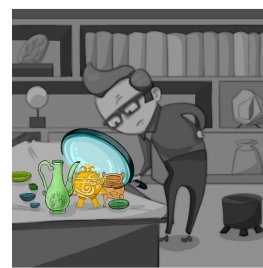
Item16:



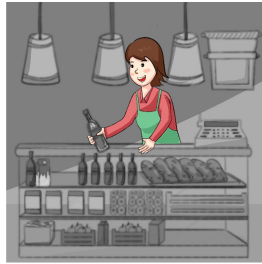
Item17:



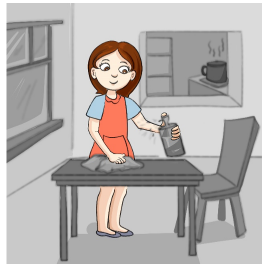
Item18:



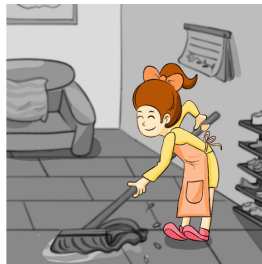
Item19:



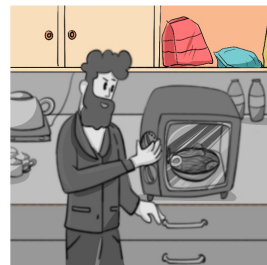
Item20:



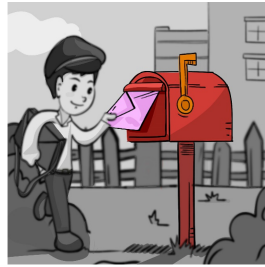
Item21:



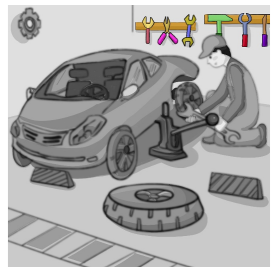
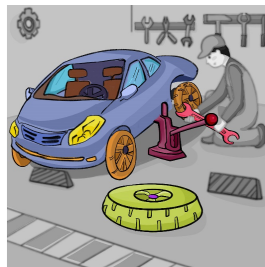
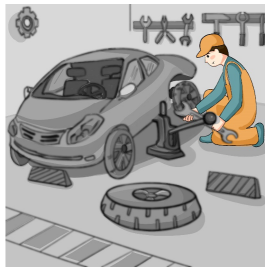
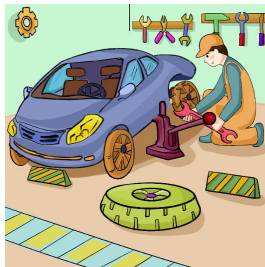
Item22:



Item23:



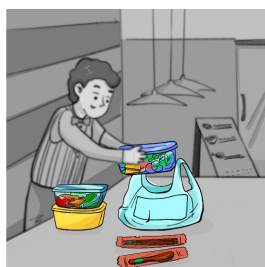
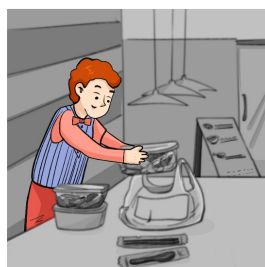
Item24:



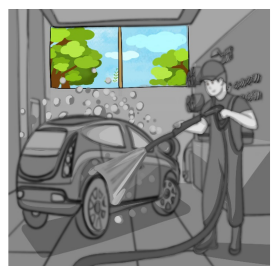
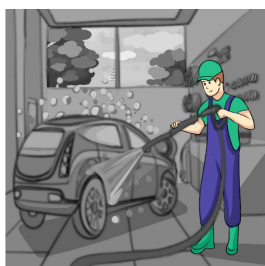
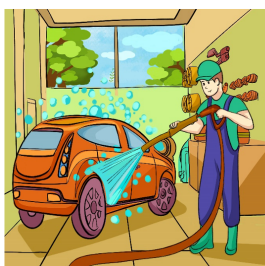
Item25:



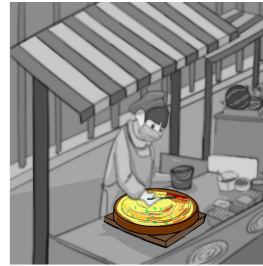
Item26:



Item27:



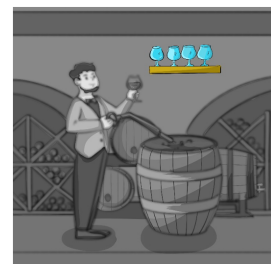
Item28:



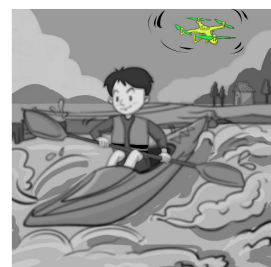
Item29:



Item30:



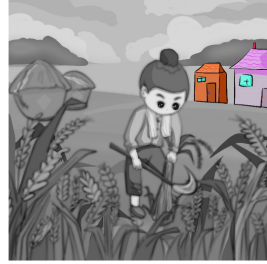
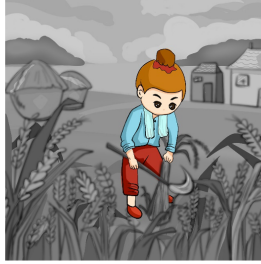
Item31:



Item32:



Item33:



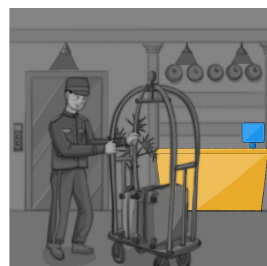
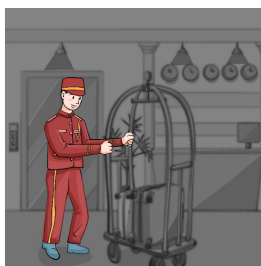
Item34:



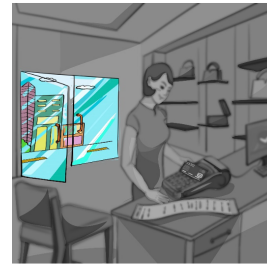
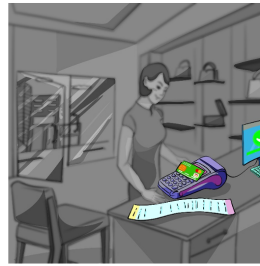
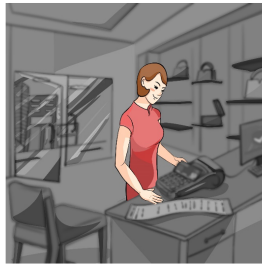
Item35:



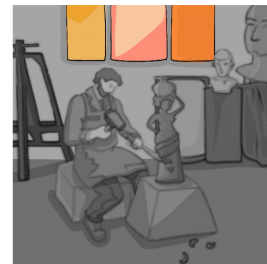
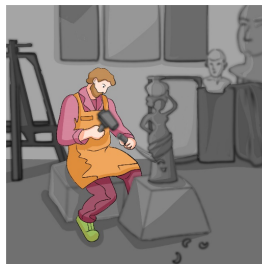
Item36:



Item37:



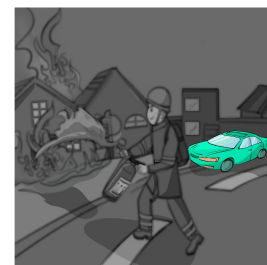
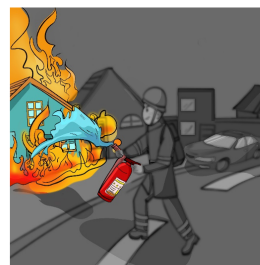
Item38:



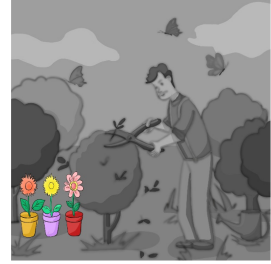
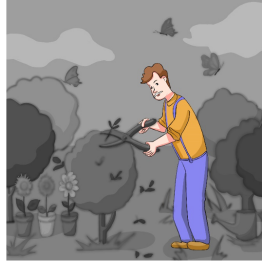
Item39:



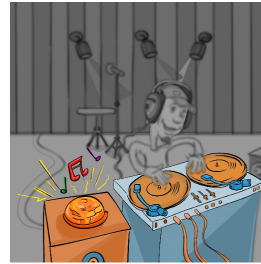
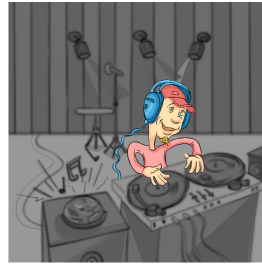
Item40:



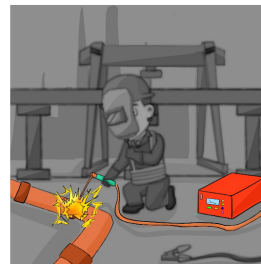
Item41:



Item42:



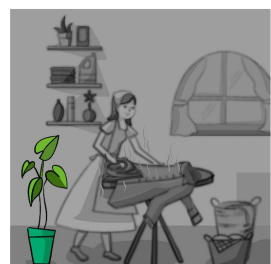
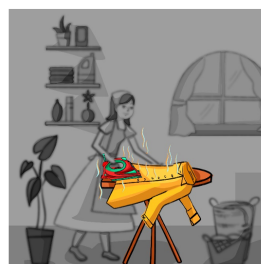
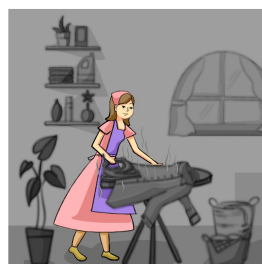
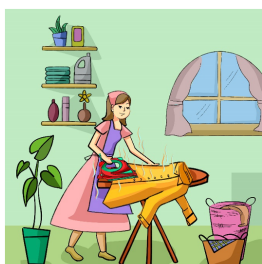
Item43:



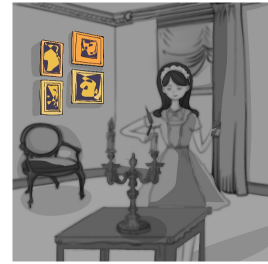
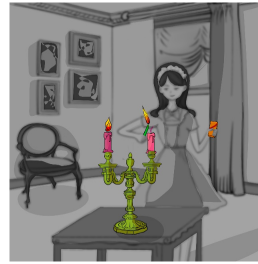
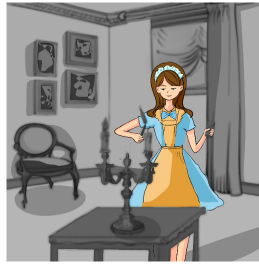
Item44:



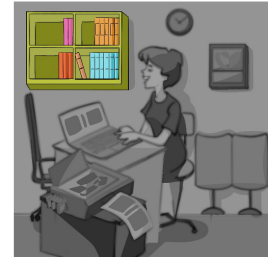
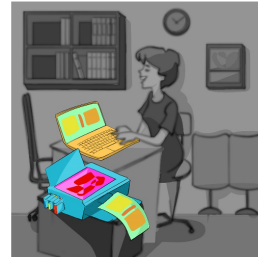
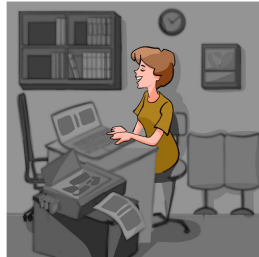
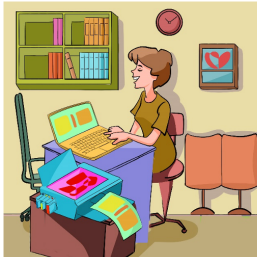
Item45:



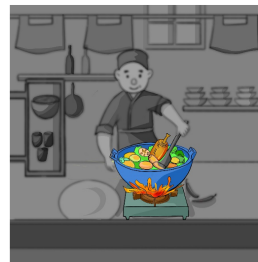
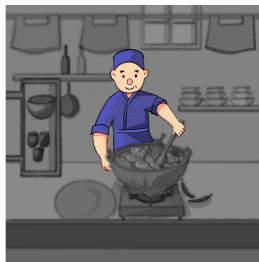
Item46:



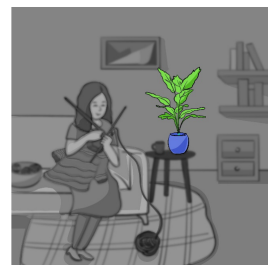
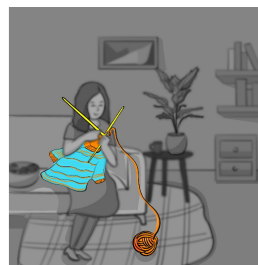
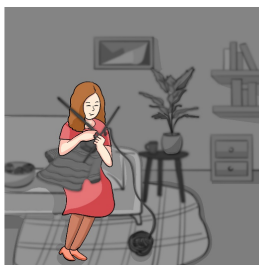
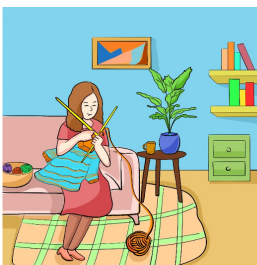
Item47:



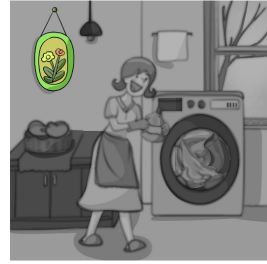
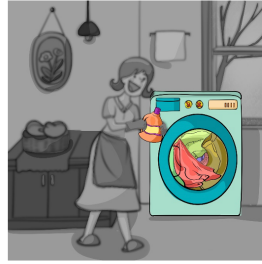
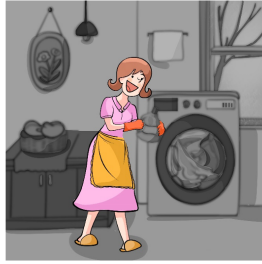
Item48:



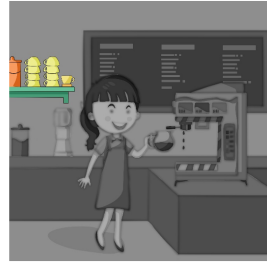
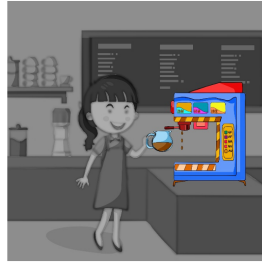
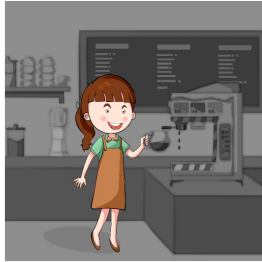
Item49:



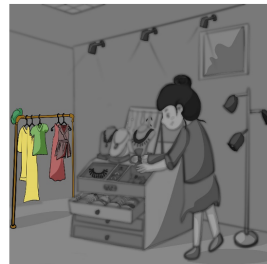
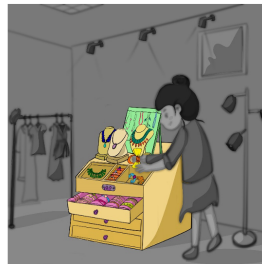
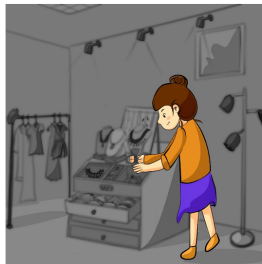
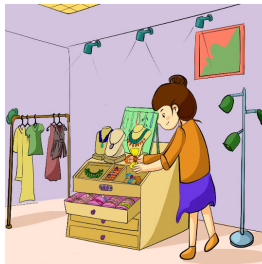
Item50:



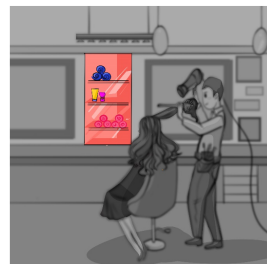
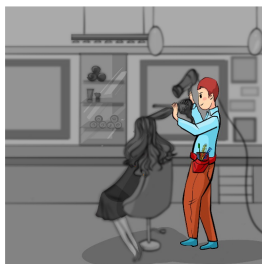
Item51:



Item52:



Item53:



Item54:



Literatur

- Alexiadou, A. (2010). Nominalizations: A probe into the architecture of grammar part I: The nominalization puzzle. *Language and Linguistics Compass*, 4(7), 496-511.
- Alexiadou, A. (2017). On the complex relationship between deverbal compounds and argument supporting nominals. In M. Bloch-Trojnar & A. Malicka-Kleparska (Eds.), *Aspect and valency in nominals* (pp. 53-82). Berlin: Walter de Gruyter.
- Altmann, G., & Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review*, 126(6), 817-840.
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- Altmann, G. T., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action* (pp. 347-386). New York and Hove: Psychology Press.
- Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57(4), 502-518.
- Altmann, G. T., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33(4), 583-609.
- Altman, M. N., Khislavsky, A. L., Coverdale, M. E., & Gilger, J. W. (2016). Adaptive attention: How preference for animacy impacts change detection. *Evolution and Human Behavior*, 37(4), 303-314.
- Aptacy, J. (2015). Rektionskomposita und Argumentvererbung—ein deutsch-polnischer Vergleich. *Zeitschrift des Verbandes Polnischer Germanisten/Czasopismo*

Stowarzyszenia Germanistów Polskich (ZVPG), 4(1), 1-16.

- Bååth, R. (2014). Bayesian First Aid: A Package that Implements Bayesian Alternatives to the Classical *.test Functions in R. In the proceedings of UseR! 2014 - *the International R User Conference*.
- Baker, M. C. (1997). Thematic roles and syntactic structure. In L. Haegeman (Ed.), *Elements of grammar: Handbook in generative syntax* (pp. 73-137). Dordrecht: Springer Netherlands.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278.
- Benatar, A., & Clifton Jr, C. (2014). Newness, givenness and discourse updating: Evidence from eye movements. *Journal of Memory and Language*, 71(1), 1-16.
- Biên, D. H. (2021). Construal and its representative forms in cognitive linguistics. *Dalat University Journal of Science*, 11(3), 3-27.
- Bierwisch, M. (1990). Event nominalizations: Proposals and problems. *Acta Linguistica Hungarica*, 40(1/2), 19-84.
- Bloomfield, L. (1933). *Language*. New York: Henry Holt.
- Bock, J. K., & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior*, 19(4), 467-484.
- Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47-67.
- Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, 43(1), 1-22.

- Bresnan, J., & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, 86(1), 168-213.
- Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28.
- Bzdęga, A. Z. (1981). Grundsätzliches zur Nominalisierung im Deutschen und Polnischen. *Studia Germanica Posnaniensia*, 11, 3-20.
- Cai, Z. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language*, 66(4), 833-849.
- Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, 155, 49-62.
- Chatterjee, A. (2001). Language and space: Some interactions. *Trends in Cognitive Sciences*, 5(2), 55-61.
- Chen, Q. (1996). “N 的 V”短语的句法分析 (the syntactic analysis of “N de V”-phrase). *Henan Daxue Xuebao*, (4), 64-67.
- Chen, Q. (2010). “N 的 V”短语的语义结构 (The semantic structure of “N de V”). *Journal of Xinyang Normal University*, 30(1), 106-110.
- Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, 67(3), 73-97.
- Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology*. University of Chicago Press.
- Comrie, B., & Thompson, S. A. (1985). Lexical nominalization. *Language Typology and Syntactic Description*, 3, 349-398.

- De Vries, M. (2001). Patterns of relative clauses. *Linguistics in the Netherlands*, 18(1), 231-243.
- Dietrich, R. (2002). *Psycholinguistik* (Vol. 342). Stuttgart: Verlag JB Metzler.
- Divjak, D., Milin, P., & Medimorec, S. (2020). Construal in language: A visual-world approach to the effects of linguistic alternations on event perception and conception. *Cognitive Linguistics*, 31(1), 37-72.
- Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, 125(2), 129-143.
- Dowty, D. R. (1989). On the semantic content of the notion of ‘thematic role’. In G. Chierchia, B. Partee & R. Turner (Eds.), *Properties, types and meaning: Volume II: Semantic issues* (pp. 69-129). Dordrecht: Springer Netherlands.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3), 547-619.
- Dryer, M. S. (1992). The Greenbergian word order correlations. *Language*, 68(1), 81-138.
- Dryer, M. S. (2003). Word order in Sino-Tibetan languages from a typological and geographical perspective. In G. Thurgood & R. J. LaPolla (Eds.), *The Sino-Tibetan Languages* (pp. 43-55). London: Routledge.
- Eichinger, L. M. (2000). *Deutsche Wortbildung: eine Einführung*. Gunter Narr Verlag.
- Eisenberg, P. (2020). *Grundriss der deutschen Grammatik: Band 2: Der Satz* (5. Auflage). Berlin: Springer-Verlag.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7(2), 195-225.

- Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, 33(4), 547-582.
- Fillmore, C. J. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1-88). New York: Holt, Rinehart and Winston.
- Flecken, M., Athanasopoulos, P., Kuipers, J. R., & Thierry, G. (2015). On the road to somewhere: Brain potentials reflect language effects on motion event perception. *Cognition*, 141, 41-51.
- Flecken, M., Gerwien, J., Carroll, M., & von Stutterheim, C. (2015). Analyzing gaze allocation during language planning: a cross-linguistic study on dynamic events. *Language and Cognition*, 7(1), 138-166.
- Gaeta, L. (2010). Synthetic compounds: With special reference to German. In S. Scalise & I. Vogel (Eds.), *Cross-disciplinary issues in compounding* (pp. 219-235). John Benjamins.
- Gaeta, L., & Zeldes, A. (2012). Deutsche Komposita zwischen Syntax und Morphologie: Ein korpusbasierter Ansatz. In L. Gaeta & B. Schlücker (Eds.), *Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte* (pp. 197-217). Walter de Gruyter.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, 83(1), 49-79.
- Gerwien, J., & Flecken, M. (2016). First things first? Top-down influences on event apprehension. In A. Papafragou, D. Grodner, D. Mirman, & J.C. Trueswell (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 38, pp. 2633-2638).
- Gerwien, J., Marberg, I., & Nicolaisen, K. (2024). What are events? In J. Gerwien, I. Marberg & K. Nicolaisen (Hrsg.), *Die kognitive Perspektive: Wie Menschen über*

- die Welt sprechen. Festschrift zu Ehren von Christiane von Stutterheim* (pp. 225-255). Heidelberg: Heidelberg University Publishing.
- Gerwien, J., & von Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition*, 180, 225-237.
- Gibbs Jr, R. W., & Colston, H. L. (1995). The cognitive psychological reality of image schemas and their transformations. *Cognitive Linguistics*, 6(4), 347-378.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1-76.
- Givón, T. (1984). *Syntax: A functional-typological introduction*. Amsterdam: John Benjamins.
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 5(4), 351-360.
- Goldberg, A. E. (2005). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of Language*, 2, 73-113.
- Greenberg, J. H. (1966). *Language universals: with special reference to feature hierarchies* (Bd. 59). The Hague; Paris: Mouton.
- Griffin, Z. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, 82(1), B1-B14.
- Griffin, Z. M., & Bock, J. K. (2000). What the eyes say about speaking? *Psychological Science*, 11, 274-279.
- Grimshaw, J. (1990). *Argument structure*. Cambridge, MA: MIT Press.

- Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition*, 175, 36-52.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Second meeting of the north american chapter of the association for computational linguistics* (pp. 1-8). Association for Computational Linguistics.
- Halliday, M. A. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, 3(2), 199-244.
- Halliday, M. A. K., & McDonald, E. (2004). Metafunctional profile of the grammar of Chinese. In A. Caffarel, J. R. Martin, & C. Matthiessen (Eds.), *Language typology: A functional perspective* (pp. 305-396). Amsterdam and Philadelphia: John Benjamins.
- Hanson, C., & Hirst, W. (1989). On the representation of events: A study of orientation, recall, and recognition. *Journal of Experimental Psychology: General*, 118(2), 136-147.
- Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition*, 34(6), 1221-1235.
- Haspelmath, M. (2011). On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology*, 15(3), 535-567.
- Haspelmath, M., & Sims, A. (2013). *Understanding morphology*. Routledge.
- He, Y., & Wang, L. (2007). 论汉语中的名物化结构 (On nominalization in Chinese). *Chinese Language Learning*, (1), 13-24.
- Hemforth, B., Konieczny, L., Seelig, H., & Walter, M. (2000). Case matching and relative clause attachment. *Journal of Psycholinguistic Research*, 29, 81-88.

- Henderson, J., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. Henderson & F. Ferreira (Eds.), *The Interface of Language, Vision, and Action* (pp. 1-58). New York: Psychology Press.
- Hindy, N. C., Altmann, G. T., Kalenik, E., & Thompson-Schill, S. L. (2012). The effect of object state-changes on event processing: Do objects compete with themselves? *Journal of Neuroscience*, 32(17), 5795-5803.
- Holler, A. (2013). d-und w-Relativsätze. In J. Meibauer, M. Steinbach & H. Altmann (Eds.), *Satztypen im Deutschen* (pp. 266-300). Berlin: Walter de Gruyter.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- Isasi-Isasmendi, A., Andrews, C., Flecken, M., Laka, I., Daum, M. M., Meyer, M., Bickel, B., & Sauppe, S. (2023). The agent preference in visual event apprehension. *Open Mind: Discoveries in Cognitive Science*, 7, 240-282.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.
- Jiang, Y. & Dai, T. (2004). 英汉名词化对比研究 (Contrastive study of nominalization between English and Chinese). *Journal of Western Chongqing University (social sciences edition)*, 3(4), 42-46.
- Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
- Johnson, M. (1989). Image-schematic basis of meaning. *Recherches semiotiques/Semiotic Inquiry*, 9, 109-118.

- Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133-156.
- Kang, X., Joergensen, G. H., & Altmann, G. T. (2020). The activation of object-state representations during online language comprehension. *Acta Psychologica*, 210, 1-14.
- Karimi, H., Diaz, M., & Ferreira, F. (2019). “A cruel king” is not the same as “a king who is cruel”: Modifier position affects how words are encoded and retrieved from memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(11), 2010-2035.
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2), 205-225.
- König, P., Wilming, N., Kietzmann, T. C., Ossandón, J. P., Onat, S., Ehinger, B. V., Gameiro, R. R., & Kaspar, K. (2016). Eye movements as a window to cognitive processes. *Journal of Eye Movement Research*, 9(5), 1-16.
- Krifka, M. (2024). Structure and interpretation of declarative sentences. *Journal of Pragmatics*, 226, 51-63.
- Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, 6(3), 299-312.
- Kuperberg, G. R. (2021). Tea with milk? A hierarchical generative framework of sequential event comprehension. *Topics in Cognitive Science*, 13(1), 256-298.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial

- language. *Cognition*, 96(1), 1-33.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume I: Theoretical prerequisites*. Stanford University Press.
- Langacker, R. W. (1991). *Foundations of cognitive grammar: Volume II: Descriptive application*. Stanford University Press.
- Langacker, R. W. (1993). Universals of construal. *Berkeley Linguistics Society*, 19, 447-463.
- Langacker, R. W. (2013). *Essentials of cognitive grammar*. Oxford University Press.
- Langacker, R. (2017). *Ten lectures on the elaboration of Cognitive Grammar* (Vol. 18). Brill.
- Langacker, R. W. (2019). Construal. In E. Dąbrowska & D. Divjak (Eds.), *Cognitive Linguistics – Foundations of Language* (pp. 140-166). Walter de Gruyter.
- Lehmann, C. (1986). On the typology of relative clauses. *Linguistics*, 24, 663-680.
- Lehmann, C. (1992). Relativization. *International Encyclopedia of Linguistics*, 3, 333-335.
- Lenth, R. (2023). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.8.7, <https://rvlenth.github.io/emmeans/>.
- Levelt, W. (1989). *Speaking*. Cambridge, MA: MIT Press.
- Levin, B., & Rappaport Hovav, M. (2005). *Argument realization* (Vol. 10). Cambridge: Cambridge University Press.
- Levinson, S. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge University Press.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126-

- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3), 375-419.
- Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10(10), 447-454.
- Li, C., & Thompson, S. (1981). *A functional reference grammar of Mandarin Chinese*. Berkeley, CA: University of California Press.
- Li, L. Z., & Yang, T. (2020). A cognitive study of nominalization in Mandarin. *Journal of Literature and Art Studies*, 10(9), 835-847.
- Liu, D. (2005). 语法调查与研究中的从属小句问题 (Subordinate clauses as an issue in language investigation and description). *Contemporary Linguistics*, (3), 193-212.
- Lu, S. & Pan, H. (2013). 从英汉比较看汉语名物化结构 (Chinese gerundive structures from a perspective of Chinese-English comparative study). *Foreign Language Teaching and Research (bimonthly)*, 45(5), 643-656.
- Majid, A., Bowerman, M., Kita, S., Haun, D., & Levinson, S. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8(3), 108-114.
- McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, 33, 1174-1184.
- Mo, P., & Shan, Q. (1985). 三大类实词句法功能的统计分析 (A Statistical Analysis of the Syntactic Functions of Three Major Categories of Content Words), *Nanjing Shida Xuebao*, 3, 55-63.
- Mousavi, S., Golfam, A., Agha Golzadeh, F., & KZ Kambuziya, A. (2019). Explaining

- the construction of euphemistic meaning via schematization in Persian: a cognitive approach. *The International Journal of Humanities*, 26(4), 30-43.
- Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64(5), 482-488.
- New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*, 104(42), 16598-16603.
- Newton, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, 28(1), 28-38.
- Newton, D., Engquist, G. A., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology*, 35(12), 847-862.
- O'Hagan, A., & Forster, J. (2004). *The advanced theory of statistics, Vol. 2B: Bayesian inference*. London, UK: Arnold.
- Olsen, S. (1986). *Wortbildung im Deutschen: Eine Einführung in die Theorie der Wortstruktur*. Stuttgart: Kröner.
- Oppenrieder, W. (2012). Deklarativsätze. In J. Meibauer, M. Steinbach & H. Altmann (Eds.), *Satztypen im Deutschen* (pp. 20-50). Berlin: Walter de Gruyter.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155-184.
- Pütz, M. (2007). Cognitive linguistics and applied linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 1139-1159). New York: Oxford University Press.
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R->

project.org/.

- Radden, G., & Dirven, R. (2007). *Cognitive English grammar*. John Benjamins Publishing.
- Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science*, 21(4), 269-272.
- Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(6), 608-620.
- Radvansky, G. A., & Zacks, J. M. (2014). *Event cognition*. Oxford University Press.
- Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457-1506.
- Rickheit, G., Sichelschmidt, L., & Strohner, H. (2002). *Psycholinguistik. Die Wissenschaft vom sprachlichen Verhalten und Erleben*. Tübingen: Stauffenburg.
- Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, 26(6), 1850-1869.
- Roy, I., & Soare, E. (2011). Nominalizations: new insights and theoretical implications. *Recherches Linguistiques de Vincennes*, (40), 7-23.
- Sauppe, S., & Flecken, M. (2021). Speaking for seeing: Sentence structure guides visual event apprehension. *Cognition*, 206, 1-7.
- Scheffler, T. (2005). Nominalization in German. *Unpublished Manuscript, University of Pennsylvania*.
- Shen, J. & Wang, D. (2000). “N 的 V”和“参照体 – 目标”构式 (“N de V” as a reference point construction). *Shijie Hanyu Jiaoxue*, (4), 25-32.

- Si, F. (2004). 中心语理论和汉语的 DeP (Head theory and DeP in Chinese). *Contemporary Linguistics*, (1), 26-34.
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1(7), 261-267.
- Slobin, D. (1996). From thought and language to thinking for speaking. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). Cambridge, UK: Cambridge University Press.
- Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, 18(5), 449-455.
- Sridhar, S. N. (1988). *Cognition and sentence production: A cross-linguistic study*. New York: Springer-Verlag.
- Stan Development Team (2025). "RStan: the R interface to Stan." R package version 2.32.6, <https://mc-stan.org/>.
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153-156.
- Talmy, L. (1978). Figure and Ground in complex sentences. In J. Greenberg, C. Ferguson, & J. Moravcsik (Eds.), *Universals of human language* (Vol. 4, pp. 625-649). Stanford University Press.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description. Grammatical categories and the lexicon* (Vol. 3, pp. 57-149). Cambridge: Cambridge University Press.
- Talmy, L. (2000). *Toward a cognitive semantics, Vol.1: Concept structuring systems*. Cambridge, MA: MIT Press.

- Talmy, L. (2008). Aspects of attention in language. In P. Robinson, & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 27-38). Routledge.
- Tang, Z. (2007). 关系化对象与关系从句的位置 – 基于真实语料和类型分析 (Correlations between demonstrative constituents and the positions of relative clauses in spontaneous Conversation: A typological perspective). *Contemporary Linguistics*, 9(2), 139-150.
- Thierry, G. (2016). Neurolinguistic relativity: How language flexes human perception and cognition. *Language Learning*, 66(3), 690-713.
- Thierry, G., Giraud, A., & Price, C. (2003). Hemispheric dissociation in access to the human semantic system. *Neuron*, 38, 499-506.
- Tınaz, B., & Ünal, E. (2024). Event segmentation in language and cognition. In L. K. Samuelson, S. L. Frank, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Meeting of the Cognitive Science Society (CogSci 2024)* (pp. 184-191).
- Toman, J. (1983). *Wortsyntax. Eine Diskussion ausgewählter Probleme deutscher Wortbildung*. Tübingen: Niemeyer.
- Tuggy, D. (2010). Schematicity. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 82-116). New York: Oxford University Press.
- Tzeng, O. J. (1973). Positive recency effect in a delayed free recall. *Journal of Verbal Learning and Verbal Behavior*, 12(4), 436-439.
- Ungerer, F., & Schmid, H. (1996). *An introduction to cognitive linguistics*. London: Longman.
- Ünal, E., Ji, Y., & Papafragou, A. (2021). From event representation to linguistic

- meaning. *Topics in Cognitive Science*, 13(1), 224-242.
- Ünal, E., Richards, C., Trueswell, J. C., & Papafragou, A. (2021). Representing agents, patients, goals and instruments in causative events: A cross-linguistic investigation of early language and cognition. *Developmental Science*, 24(6), 1-13.
- Verhagen, A. (2007). Construal and perspectivization. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 48-81). New York: Oxford University Press.
- Wagers, M. W., & Phillips, C. (2014). Going the distance: Memory and control processes in active dependency construction. *Quarterly Journal of Experimental Psychology*, 67(7), 1274-1304.
- Wang, C. (2010). *The study of “N de V”-construction*. (Master’s thesis, Shanghai International Studies University).
- Wang, D. (2010). “N 的 V”构式与谓语动词的选择 (The “N de V” Construction and the Selection of Predicate Verbs). *Haiwai Huawen Jiaoyu*, (2), 46-51.
- Wang, Y. (2011). 关系化和汉语关系从句 (Relativization and Chinese relative clauses). *Journal of Yancheng Teachers University (Humanities & Social Sciences)*, 31(3), 74-77.
- Wang, Y., & Feng, L. (2012). 汉语语义角色的关系化及关系化难度等级序列分析 (Relativization of Chinese semantic roles and difficulty hierarchy of relativization). *Journal of Yunnan Normal University (Teaching & Studying Chinese as a Foreign Language Edition)*, 10(5), 6-14.
- Whorf, B. L. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V.,

- Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. doi:10.21105/joss.01686.
- Wilson, F., Papafragou, A., Bungler, A., & Trueswell, J. (2011). Rapid extraction of event participants in caused motion events. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 33(33), 1206-1211.
- Wilson, F., Ünal, E., Trueswell, J., & Papafragou, A. (2014). Homologies between language and event cognition: Evidence from event role prominence. In *39th Annual Boston University Conference on Language Development, Boston, MA*.
- Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3), 253-265.
- Wong, K. C., & Ginzburg, J. (2013). Investigating non-sentential utterances in a spoken Chinese corpus. In *PACLING 2013*, Tokyo, Japan.
- Wu, Y. J., & Thierry, G. (2012). How reading in a second language protects your heart. *Journal of Neuroscience*, 32(19), 6485-6489.
- Yu, L. (2005). 基于图形-背景理论的英汉关系分句认知对比分析 (A cognitive comparison to English and Chinese relative clause – from the point of figure-ground theory). *Journal of Yuxi Teachers College*, 21(7), 56-59.
- Zacks, J. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, 28, 979-1008.
- Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6), 651-655.
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: a mind-brain perspective. *Psychological Bulletin*, 133(2), 273-

- Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, 138(2), 307-327.
- Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127(1), 3-21.
- Zhan, W. (1998a). 关于“NP+的+VP”偏正结构 (Attributive noun phrase “NP de VP” in Chinese). *Chinese Language Learning*, 2, 24-28.
- Zhan, W. (1998b). “NP + 的 + VP”偏正结构在组句谋篇中的特点 (Characteristics of the “NP + de + VP” Modifier-Head Structure in Sentence Construction and Discourse Organization). *Yuwen Yanjiu*, (1), 16-23.
- Zhang, B. (1993). “N 的 V” 结构的构成 (The composition of the “N de V” structure). *Studies of the Chinese Language*, (4), 252-259.
- Zhang, Q. (2019). 句法可及和语篇可及 – 汉语关系从句中指数量标记分布考察 (Syntax accessibility and discourse accessibility: The distribution of demonstrative quantitative classifier markers in Chinese relative clauses). *Journal of Foreign Languages*, 42(5), 62-71.
- Zhu, D. (1982). 语法讲义 (A Course in Grammar). Beijing: Shangwu yinshuguan.
- Zhu, D., Lu, J., & Ma, Z. (1961). 关于动词形容词“名物化”问题 (The topic regarding the nominalization of verbs and adjectives). *Journal of Peking University*, (4), 51-64.
- Ziegler, J., & Snedeker, J. (2018). How broad are thematic roles? Evidence from structural priming. *Cognition*, 179, 221-240.
- Zifonun, G. (2001). *Grammatik des Deutschen im europäischen Vergleich: der Relativsatz*. Mannheim: Institut für deutsche Sprache.

- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6(5), 292-297.
- Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). *Moving words: Dynamic representations in language comprehension*. *Cognitive Science*, 28(4), 611-619.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168-171.