

*Verarbeitungsprozesse in der Satzproduktion:
Die Interaktion zwischen visueller, semantischer und
kontextueller Salienz*

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

der Neuphilologischen Fakultät der Ruprecht-Karls-Universität Heidelberg

vorgelegt von

Xiaogang Wu (吴晓钢)

aus Jiangxi, VR China

Erstgutachterin: Prof. Dr. Christiane von Stutterheim

Zweitgutachterin: PD Dr. Sandra Pappert

Heidelberg, im Dezember 2018

Danksagung

Ohne die Unterstützung zahlreicher Menschen wäre diese Arbeit nicht möglich gewesen. Diesen möchte ich an dieser Stelle danken. Mein besonderer Dank gilt ...

Prof. Dr. Christiane von Stutterheim, die mir durch ihre Betreuung die Dissertation ermöglichte und mich mit ihrer kompetenten wissenschaftlichen Beratung, ihrer wertvollen Unterstützung und ihrem steten Glauben an dieses Forschungsprojekt in allen Phasen der Promotion begleitete. Vielen herzlichen Dank für die sehr gute Betreuung und das mir entgegengebrachte Vertrauen.

PD Dr. Sandra Pappert für die wertvolle Unterstützung bei der Vorbereitung des AmLaP-Talks und für die Bereitschaft, die vorliegende Arbeit zu lesen.

Dr. Johannes Gerwien, der mein Interesse für Psycholinguistik während meines Masterstudiums geweckt hat und mir stets mit seinem Rat und seiner wissenschaftlichen Erfahrung zur Seite gestanden ist. Unsere produktiven Diskussionen und seine wertschätzende Unterstützung trugen maßgeblich zum Gelingen der Arbeit bei.

China Scholarship Council (CSC). Ohne die finanzielle Unterstützung durch das CSC-Stipendium hätte diese Arbeit nicht zustande kommen können.

Allen Probanden, die an der Studie teilgenommen und das Gelingen dieser Arbeit ermöglicht haben.

Allen meinen Freunden in Heidelberg, die mir beim Verfassen der Arbeit geholfen haben, insbesondere Takara Baumbach, Qili Wang, Rike Herweg, Romina Ott, Abbassia Bouhaous, Ouaffa Qaddioui, Philip Kehl, Andreas Barmin und Hedwig Schilling.

Meinen Eltern, Gong Chunhong 龚春红 und Wu Jianjun 吴建军, für ihr Vertrauen in mich.

Meiner Frau Zhu Heng 朱珩, dass sie diese lange Zeit mit mir zusammen durchgestanden hat. Ohne diesen Rückhalt hätte ich es nicht geschafft.

Zusammenfassung

In der vorliegenden Dissertation geht es um Verarbeitungsprozesse innerhalb der Satzproduktion. Die im Rahmen der Arbeit durchgeführte Studie basiert auf dem Satzproduktionsmodell von Bock und Levelt (1994). Im Gegensatz zu früheren Studien wird das Satzproduktionssystem jedoch nicht nur in pragmatisch neutralen, sondern auch in pragmatisch variierenden Kommunikationssituationen untersucht. Ziel ist es, mit Hilfe der Eye-Tracking-Technologie die Art der Informationsverarbeitung während des Sprechens festzustellen.

In insgesamt drei Experimenten werden visuell dargebotene Ereignisse von deutschsprachigen Probanden beschrieben, wobei die visuelle, semantische und kontextuelle Salienz der Ereignisteilnehmer als Einflussfaktoren systematisch manipuliert werden. Analysen der Sprech- und Blickbewegungsdaten führen zu der Aussage, dass die Satzproduktion im Deutschen generell ein strukturell-inkrementell verlaufender Prozess ist, bei dem nicht nur der semantischen Ereignisstruktur, sondern auch der Informationsstruktur eine zentrale Bedeutung zukommt. Mit diesem Befund gelingt es, die Inkrementalitätstheorie auf die pragmatische Verarbeitungsebene zu erweitern und die vorhandenen Satzproduktionsmodelle, welche dem Kontext bisher nur eine vergleichsweise geringe Beachtung geschenkt haben, zu bereichern. Des Weiteren wird eine Verbindung zwischen der Informationsstruktur und der semantischen Struktur auf der Mikroplanungsebene des Satzproduktionssystems festgestellt. Dieser Befund erlaubt somit neue Erkenntnisse über die Arbeitsweise der Mikroplanung. Darüber hinaus deuten Ergebnisse der Studie darauf hin, dass kontextuelle Informationen bei der Satzproduktion im Deutschen vorrangig verarbeitet werden. Um eine Übersicht über die Verarbeitungshierarchie verschiedener Faktoren bei der Satzplanung zu geben, wird im Rahmen der Dissertation eine Prioritätenliste entwickelt.

Inhaltsverzeichnis

1	EINLEITUNG	1
2	MODELLIERUNG DER SATZPRODUKTION	6
2.1	KONZEPTUALISIERUNG	8
2.2	GRAMMATISCHE ENKODIERUNG	15
2.2.1	<i>Lexikalische Selektion</i>	15
2.2.2	<i>Funktionszuweisung</i>	17
2.2.3	<i>Konstituenten-Zusammenstellung</i>	23
2.2.4	<i>Flexion</i>	29
2.3	PHONOLOGISCH-PHONETISCHE ENKODIERUNG UND ARTIKULATION	31
2.4	MAPPINGS	33
2.5	FAZIT	39
3	INKREMENTALITÄT	40
3.1	<i>DAS STARTING POINTS PROBLEM</i>	42
3.2	LINEARE INKREMENTALITÄT UND STRUKTURELLE INKREMENTALITÄT	44
3.2.1	<i>Implikationen der linearen und der strukturellen Inkrementalität</i>	44
3.2.2	<i>Befunde zur linearen Inkrementalität</i>	48
3.2.3	<i>Befunde zur strukturellen Inkrementalität</i>	53
3.3	VARIABILITÄT DER INKREMENTELLEN VERARBEITUNG	58
3.4	FAZIT	63
4	SALIENZ UND SATZPRODUKTION	65
4.1	SALIENZ	66
4.2	VISUELLE SALIENZ UND SATZPRODUKTION	69
4.3	SEMANTISCHE SALIENZ UND SATZPRODUKTION	72
4.4	KONTEXTUELLE SALIENZ UND SATZPRODUKTION	76
4.4.1	<i>Gegebenheit und Satzproduktion</i>	77
4.4.2	<i>Topikalität und Satzproduktion</i>	81
4.5	FAZIT	88
5	UNTERSUCHUNGSFRAGEN UND VERSUCHSDESIGN	90
5.1	UNTERSUCHUNGSFRAGEN	90
5.2	ENTWICKLUNG EINES VERSUCHSDESIGNS	95
5.3	MESSGRÖßEN	104
5.3.1	<i>Linguistische Messgrößen</i>	104

5.3.2	<i>Visuelle Messgrößen</i>	106
6	EXPERIMENT 1	110
6.1	HYPOTHESEN	110
6.1.1	<i>Hypothesen über die Arbeitsweise des Satzproduktionssystems: lineare vs. strukturelle Inkrementalität</i>	110
6.1.2	<i>Hypothesen über die Prioritätenliste der mitwirkenden Faktoren: visuelle vs. semantische Salienz</i>	115
6.2	METHODISCHES VORGEHEN	117
6.2.1	<i>Versuchspersonen</i>	117
6.2.2	<i>Technische Apparatur</i>	117
6.2.3	<i>Material</i>	118
6.2.4	<i>Durchführung</i>	119
6.2.5	<i>Datenkodierung</i>	123
6.2.6	<i>Datenbearbeitung und -analyse</i>	124
6.3	ERGEBNISSE	126
6.4	DISKUSSION.....	130
7	EXPERIMENT 2	146
7.1	HYPOTHESEN	147
7.1.1	<i>Hypothesen über die Rolle der kontextuellen Salienz bzw. der Topikalität bei der Satzproduktion</i>	147
7.1.2	<i>Hypothesen über den Abbildungsprozess des Topiks auf die Syntax</i>	148
7.1.3	<i>Hypothesen über die Prioritätenliste der mitwirkenden Faktoren</i>	152
7.2	METHODISCHES VORGEHEN	153
7.2.1	<i>Versuchspersonen</i>	153
7.2.2	<i>Technische Apparatur</i>	153
7.2.3	<i>Material</i>	154
7.2.4	<i>Durchführung</i>	154
7.2.5	<i>Datenkodierung</i>	155
7.2.6	<i>Datenbearbeitung und -analyse</i>	155
7.3	ERGEBNISSE	159
7.4	DISKUSSION.....	165
8	EXPERIMENT 3	184
8.1	HYPOTHESEN	185
8.1.1	<i>Hypothesen über den Abbildungsprozess des Topiks auf die Semantik</i>	185
8.1.2	<i>Hypothesen über den Effekt der kontextuellen Salienz bzw. der Topikalität auf die Satzplanung</i>	186
8.2	METHODISCHES VORGEHEN	186

8.2.1	<i>Versuchspersonen</i>	186
8.2.2	<i>Technische Apparatur</i>	187
8.2.3	<i>Material</i>	187
8.2.4	<i>Durchführung</i>	187
8.2.5	<i>Datenkodierung</i>	187
8.2.6	<i>Datenbearbeitung und -analyse</i>	187
8.3	ERGEBNISSE	188
8.4	DISKUSSION.....	193
9	ALLGEMEINE DISKUSSION	200
9.1	ZUSAMMENFASSUNG DER ERGEBNISSE.....	200
9.2	VERSUCH EINER EIGENEN KONZEPTION.....	211
9.3	AUSBLICKE.....	216
9.4	ALLGEMEINE SCHLUSSFOLGERUNG	218
10	LITERATURVERZEICHNIS	223

Abbildungsverzeichnis

ABBILDUNG 2.1: SATZPRODUKTIONSMODELL ADAPTIERT VON LEVELT 1989 UND BOCK UND LEVELT 1994.....	7
ABBILDUNG 3.1: INKREMENTELLE VERARBEITUNG DER SATZPRODUKTION (ADAPTIERT VON KEMPEN UND HOENKAMP 1987).....	41
ABBILDUNG 3.2: BEISPIELSTIMULUS AUS TOMLIN 1997.....	50
ABBILDUNG 3.3: ZEITLICHER VERLAUF EINES DURCHGANGS NACH GLEITMANN ET AL. 2007.....	51
ABBILDUNG 3.4: BEISPIELSTIMULUS AUS GRIFFIN UND BOCK 2000.....	56
ABBILDUNG 4.1: BEISPIELSTIMULI AUS DOBEL ET AL. 2007; KOHÄRENTES EREIGNIS (LINKS) UND INKOHÄRENTES EREIGNIS (RECHTS).....	74
ABBILDUNG 4.2: BEISPIELSTIMULUS AUS PRAT-SALA UND BRANIGAN 2000.....	80
ABBILDUNG 5.1: BEISPIELSTIMULUS AUS DER VORLIEGENDEN STUDIE.....	97
ABBILDUNG 5.2: SCHEMATISCHE ILLUSTRATION DER STIMULI AUS THEEUWES ET AL. 1998. GRAUE KREISPERIPHERIEN SIND DURCH GESTRICHELTE LINIEN ANGEZEIGT, UND ROTE KREISPERIPHERIEN SIND DURCH DURCHGEZOGENE LINIEN ANGEZEIGT; ZITIERT IN MCCARLEY & KRAMER 2007.....	100
ABBILDUNG 5.3: DREI PHASEN DER SAKKADENGENERIERUNG.....	106
ABBILDUNG 5.4: INTERSAKKADISCHES INTERVALL (IM OBEREN TEIL) UND VERKÜRZTES INTERSAKKADISCHES INTERVALL (IM UNTEREN TEIL).....	109
ABBILDUNG 6.1 BEISPIELABLAUF DER KRITISCHEN BEDINGUNG UND DER FILLER-BEDINGUNG (EXPERIMENT 1).....	122
ABBILDUNG 6.2 PROPORTION DER ERSTEN FIXATIONEN, DIE AUF VISUELL MARKIERTE REFERENTEN GERICHTET WURDEN, IM VERHÄLTNISS ZUR SEMANTISCHEN ROLLE DER REFERENTEN.....	128
ABBILDUNG 6.3: PROPORTION DER ZWEITEN SAKKADEN, DIE AUF EINEN ANDEREN REFERENTEN ALS DEN ZUERST FIXIERTEN GERICHTET WURDEN, IM VERHÄLTNISS ZUR POSITION DER ERSTEN FIXATIONEN.....	129
ABBILDUNG 6.4: PROPORTION PRODUZIERTER KANONISCHER AKTIVSÄTZE IM VERHÄLTNISS ZUR POSITION DER ERSTEN FIXATIONEN.....	130
ABBILDUNG 6.5: SIMULTANE PROGRAMMIERUNGEN DER ERSTEN UND ZWEITEN SAKKADE.....	141
ABBILDUNG 7.1: VIER HYPOTHESEN ÜBER DEN ABBILDUNGSPROZESS DES TOPIKS AUF DIE WORTSTELLUNG.....	149
ABBILDUNG 7.2: BEISPIELABLAUF DER KRITISCHEN BEDINGUNG UND DER FILLER-BEDINGUNG (EXPERIMENT 2).....	156
ABBILDUNG 7.3: PROPORTION DER DURCH DEN VISUELLEN CUE GETRIEBENEN ERSTEN FIXATIONEN IM VERHÄLTNISS ZUM DISKURSTATUS DES PATIENS (DAS PATIENS WURDE IMMER DURCH DEN CUE MARKIERT).....	160

ABBILDUNG 7.4 PROPORTION DER DURCH DAS TOPIK GETRIEBENEN ERSTEN FIXATIONEN IM VERHÄLTNIS ZUR VISUELLEN SALIENZ DER TOPIKALISCHEN REFERENTEN	161
ABBILDUNG 7.5 PROPORTION DER ZWEITEN SAKKADEN, DIE AUF DAS AGENS GESTEUERT WURDEN, IM VERHÄLTNIS ZUM DISKURSTATUS DES AGENS.....	162
ABBILDUNG 7.6: PROPORTION DER GEBILDETEN AKTIVSÄTZE IM VERHÄLTNIS ZUM DISKURSTATUS DES AGENS.....	163
ABBILDUNG 7.7 SPEECH-ONSET-LATENZ DER SELEKTIERTEN SÄTZE AUS EXPERIMENT 1 UND 2.....	164
ABBILDUNG 8.1 PUNKTDIAGRAMM DER SPEECH-ONSET-LATENZ DER TOPIK-INITIALEN PASSIVSÄTZE (LINKS) UND DER TOPIK-INITIALEN AKTIVSÄTZE (RECHTS) (IN MILLISEKUNDE).....	190
ABBILDUNG 8.2 PROPORTION DER DURCH DEN VISUELLEN CUE GETRIEBENEN ERSTEN FIXATIONEN IM VERHÄLTNIS ZUM DISKURSTATUS DES AGENS (DAS AGENS WURDE IMMER DURCH DEN CUE MARKIERT).....	190
ABBILDUNG 8.3 PROPORTION DER ZWEITEN SAKKADEN, DIE AUF DAS AGENS GESTEUERT WURDEN, IM VERHÄLTNIS ZUM DISKURSTATUS DES AGENS	191
ABBILDUNG 8.4 PROPORTION DER GEBILDETEN AKTIVSÄTZE IM VERHÄLTNIS ZUM DISKURSTATUS DES AGENS.....	192
ABBILDUNG 9.1 DER ABBILDUNGSPROZESS DES TOPIKS INNERHALB DER SATZPLANUNG.....	209
ABBILDUNG 9.2 SPOTLIGHT-MODELL ZUR BESCHREIBUNG DER SELEKTIONSWEISE VON STARTING POINTS.....	211

Tabellenverzeichnis

TABELLE 3.1: ZUSAMMENFASSUNG DER KERNAUSSAGEN DER LINEAREN UND DER STRUKTURELLEN INKREMENTALITÄTSHYPOTHESE	47
TABELLE 3.2: ERGEBNISSE AUS KUCHINSKY & BOCK 2010.....	61
TABELLE 6.1: FEHLER (EXPERIMENT 1).....	126
TABELLE 6.2: LOGISTISCHE REGRESSIONSANALYSE AUF DIE ABHÄNGIGKEIT DES VISUELLEN CUEINGEFFEKTS VON DER SEMANTISCHEN ROLLE DER VISUELL MARKIERTEN REFERENTEN.....	128
TABELLE 6.3: LOGISTISCHE REGRESSIONSANALYSE AUF DIE ABHÄNGIGKEIT DER ZIELSELEKTION DER ZWEITEN SAKKADEN VON DER POSITION DER ERSTEN FIXATIONEN	129
TABELLE 6.4: LOGISTISCHE REGRESSIONSANALYSE AUF DIE ABHÄNGIGKEIT DER SYNTAKTISCHEN SELEKTION VON DER POSITION DER ERSTEN FIXATIONEN	130
TABELLE 7.1: FEHLER (EXPERIMENT 2).....	159
TABELLE 7.2: LOGISTISCHE REGRESSIONSANALYSE AUF DIE ABHÄNGIGKEIT DES VISUELLEN CUEINGEFFEKTS VOM DISKURSTATUS DES PATIENS	160
TABELLE 7.3: LOGISTISCHE REGRESSIONSANALYSE AUF DIE ABHÄNGIGKEIT DES TOPIKALITÄTS-EFFEKTS VON DER VISUELLEN SALIENZ DER TOPIKALISCHEN REFERENTEN.....	161
TABELLE 7.4: LOGISTISCHE REGRESSIONSANALYSE AUF DIE ABHÄNGIGKEIT DER SELEKTION DER ZWEITEN SAKKADISCHEN ZIELPOSITION VOM DISKURSTATUS DES AGENS	162
TABELLE 7.5: LOGISTISCHE REGRESSIONSANALYSE AUF DIE ABHÄNGIGKEIT DER SELEKTION SYNTAKTISCHER STRUKTUREN VOM DISKURSTATUS DES AGENS	163
TABELLE 8.1: FEHLER (EXPERIMENT 3).....	189
TABELLE 8.2 LOGISTISCHE REGRESSIONSANALYSE MIT GEMISCHTEN EFFEKTEN AUF DIE ABHÄNGIGKEIT DER VISUELLEN SALIENZ VOM DISKURSTATUS DES AGENS.....	191
TABELLE 8.3 LOGISTISCHE REGRESSIONSANALYSE MIT GEMISCHTEN EFFEKTEN AUF DIE ABHÄNGIGKEIT DER SELEKTION DER ZWEITEN SAKKADISCHEN ZIELPOSITION VOM DISKURSTATUS DES AGENS.....	191
TABELLE 8.4 LOGISTISCHE REGRESSIONSANALYSE MIT GEMISCHTEN EFFEKTEN AUF DIE ABHÄNGIGKEIT DER SELEKTION SYNTAKTISCHER STRUKTUREN VOM DISKURSTATUS DES AGENS.....	192

Abkürzungen

Ps. grammatische Person

Sg. Singular

ms Millisekunde

NP Nominal-Phrase

SVO Subjekt-Verb-Objekt

SOV Subjekt-Objekt-Verb

VSO Verb-Subjekt-Objekt

VOS Verb-Objekt-Subjekt

OSV Objekt-Subjekt-Verb

VOS Verb-Objekt-Subjekt

DO-Konstruktion *double-object*-Konstruktion

PO-Konstruktion *prepositional-object*-Konstruktion

1 Einleitung

Sprechen scheint für uns Menschen ein einfacher Vorgang zu sein – in einer Sekunde können wir mühelos durchschnittlich drei bis fünf Wörter produzieren. Es ist ein hochgradig automatisierter Prozess, dessen Ablauf uns in der Regel kaum bewusst ist. Sieht man aber genauer hin, wird deutlich, dass beim Sprechakt eine Menge an Teilprozessen miteinander verwoben sind. Einen Gedanken zu formulieren, erfordert die Koordination mehrerer kognitiver und motorischer Enkodierungen. Von der Intention ausgehend, etwas formulieren zu wollen, planen Sprecher zuerst, was sie sagen möchten. Dabei müssen sie entscheiden, welche Perspektive sie für ihre Äußerung wählen. Es geht vor allem darum, was im vorausgegangenen Gespräch bereits gesagt wurde. Dieses gemeinsame Wissen teilen Gesprächspartner. Es muss dann entschieden werden, welche Information hervorgehoben werden soll. Dann werden Wörter ausgewählt, die in eine syntaktische Struktur gesetzt und in einer grammatischen Abfolge angeordnet werden. Anschließend enkodieren Sprecher die phonologische und phonetische Information und artikulieren schließlich die zu produzierenden Laute.

Eines der im Rahmen der Sprachproduktion wichtigsten und deshalb auch am intensivsten untersuchten Forschungsgebiete ist die Satzproduktion. Wilhelm M. Wundt (1904), Begründer der Psychologie als eigenständige Wissenschaft, definiert einen Satz wie folgt:

Psychologisch betrachtet ist demnach der Satz beides zugleich, ein simultanes und ein sukzessives Ganzes: ein simultanes, weil er in jedem Moment seiner Bildung in seinem ganzen Umfang im Bewußtsein ist, wenn auch einzelne Nebenelemente gelegentlich aus diesem verschieden mögen; ein sukzessives, weil sich das Ganze von Moment zu Moment in seinem Bewusstseinszustand verändert, indem nacheinander bestimmte Vorstellungen in den Blickpunkt treten und andere dunkler werden. (S. 241)

Diese Definition beschreibt demnach nicht nur den Satz selbst, sondern auch den Sprachproduktionsprozess. Das lässt sich gut in folgendem Beispiel zeigen: Stellen

wir uns die Szene vor, in der ein Sprecher ein Ereignis beschreibt, indem eine Frau einen Mann mit einer Pistole erschießt. Wie lässt sich dies nun versprachlichen bzw. wie verläuft nun der Satzproduktionsprozess? Wundts These nach wird die Szene zum einen als Gesamtheit mehrerer Darstellungskomponenten mental verarbeitet und zum anderen werden die einzelnen Teilnehmer (Frau, Pistole, Mann) nach und nach versprachlicht. Die Satzproduktion ist daher ein Prozess, wie ein Gedanke mit der zu formulierenden Szene als Ganzes einerseits und simultan vorhandenen konzeptuellen Segmenten andererseits in eine lineare Sequenz aus lexikalisch, syntaktisch und lautlich enkodierten Konzepten transformiert wird.

Innerhalb der Psycholinguistik wird die überwiegende Anzahl der Untersuchungen zur Satzproduktion in pragmatisch neutralen Situationen durchgeführt, in denen Komponenten eines Satzes einen vergleichbaren Diskursstatus aufweisen. Jedoch werden Sätze im Alltagsleben kaum isoliert gebildet, sondern müssen immer in einen bestimmten Kontext eingeordnet werden. Der Kontext, der theoretisch als wichtiger Faktor einen Einfluss auf die Satzproduktion hat, wurde allerdings empirisch wenig betrachtet.

Der Fokus der vorliegenden Arbeit liegt auf Verarbeitungsprozessen der Satzproduktion im Deutschen und deren Verhältnisse zu drei Einflussfaktoren, nämlich der visuellen, semantischen und insbesondere kontextuellen Salienz. So wird die Satzplanung in dieser Arbeit im Gegensatz zu vorangegangenen Studien nicht nur in pragmatisch neutralen Situationen, in denen die zu versprachlichenden Referenten einen gleichen Diskursstatus aufweisen, sondern auch in kontextuellen Situationen, in denen der Diskursstatus der Referenten unterschiedlich ist, untersucht. Untersucht wird hierbei die Arbeitsweise des deutschen Satzproduktionssystems in diesen beiden Situationen, sowie die Repräsentation und die Art der Verarbeitung der kontextuellen Information bzw. des Topiks im Planungsprozess.

Die vorliegende Arbeit gliedert sich in neun Kapitel. Zu Beginn stelle ich in Kapitel 2 die Modellierung der Satzproduktion dar. Die Satzproduktion, so die Annahme,

besteht aus vier Teilsystemen – der Konzeptualisierung, der grammatischen Enkodierung, der phonologisch-phonetischen Enkodierung sowie der Artikulation (vgl. Bock & Levelt, 1994). Man kann sich nun fragen, welche Informationen auf der jeweiligen Ebene verarbeitet werden und in welchem Verhältnis diese Teilsysteme zueinanderstehen. Diese Frage steht im Vordergrund des Kapitels. Dann wird basierend auf dem Satzproduktionsmodell in diesem Kapitel ebenfalls der Informationsfluss innerhalb der Satzplanung betrachtet. Die Frage, die es in diesem Zusammenhang zu beantworten gilt, lautet, wie Informationen von der konzeptuellen auf der positionellen Enkodierungsebene, die für die Erstellung der Wortstellung zuständig ist, abgebildet werden.

In Kapitel 3 wird die Arbeitsweise des Satzplanungssystems thematisiert. Es wird basierend auf der These zur inkrementellen Verarbeitung zwischen zwei Inkrementalitätshypothesen – der linearen und der strukturellen – unterschieden. Den Ausgangspunkt bildet die Frage, wie das erste lexikalische Konzept, also der sog. *starting point* der Satzplanung, selektiert wird (das sog. *starting point problem*). Der linearen Inkrementalitätshypothese zufolge liegt dem Selektionsprozess eine Bottom-Up-Verarbeitung der Eigenschaften einzelner Referenten (wie etwa des visuellen oder konzeptuellen Merkmals) zugrunde. Im Gegensatz dazu erfordert die Selektion von *starting points* der strukturellen Inkrementalitätshypothese nach den Aufbau der semantisch-konzeptuellen Struktur des Gesamtereignisses. Von diesen beiden polarisierenden Hypothesen ausgehend, wurde ein weiterer und dritter Ansatz aufgestellt, der besagt, dass je nach konkreter Satzproduktionssituation beide inkrementelle Arbeitsweisen möglich sind. Da die Satzproduktion im Deutschen – soweit dem Autor dieser Arbeit bekannt – in vorangegangenen Studien im Rahmen der Inkrementalitätshypothese wenig untersucht wurde, ist es interessant zu hinterfragen, welche Arbeitsweise sie verfolgt.

Drei bei der Satzproduktion wirkende Einflussfaktoren, nämlich die semantische, visuelle und kontextuelle Salienz, werden anschließend in Kapitel 4 thematisiert. Bei der visuellen Salienz handelt es sich um die durch die Augen greifbare Auffälligkeit,

die durch eine Menge distinktiver visueller Eigenschaften erzeugt wird, während die semantische und kontextuelle Salienz jeweils auf der semantischen Struktur und Informationsstruktur basieren und inhalts- oder diskursindizierte Implikationen aufweisen. Dieses Kapitel steht in einem engen Zusammenhang mit Kapitel 3, da die Hervorhebung des Effekts der visuellen und der semantischen Salienz jeweils der linearen und der strukturellen Inkrementalitätshypothese entspricht. Obwohl der Effekt der kontextuellen Salienz in vorangegangenen Studien mehrfach bestätigt wurde, ist bis dato allerdings noch nicht klar, wie kontextuelle Informationen, hier vor allem das Topik, während der Satzplanung repräsentiert werden und in welcher Art und Weise sie sich auswirken. Des Weiteren wurde die kontextuelle Salienz, die den Diskursstatus von Referenten beschreibt und somit die Informationsstruktur widerspiegelt, bisher nicht in der Untersuchung der strukturellen Inkrementalität mit einbezogen.

In Kapitel 5 werden die zugrundeliegenden Fragestellungen, die Entwicklung des Versuchsdesigns und die Auswahl der analytischen Methoden ausführlich beschrieben.

In dieser Studie wurden insgesamt drei *eye-tracking*-Experimente durchgeführt. Jedem Experiment ist ein eigenes Kapitel gewidmet, in dem nach den einführenden Fragestellungen und den Hypothesen das methodische Vorgehen und die Ergebnisse dargestellt werden, bevor sich die Diskussion anschließt. Das erste Experiment, das die Arbeitsweise des deutschen Satzproduktionssystems in pragmatisch neutralen Situationen untersucht, wird in Kapitel 6 vorgestellt. Hierfür werden die visuelle und die semantische Salienz als Einflussfaktoren der Satzplanung eingeführt und systematisch manipuliert. Es soll herausgefunden werden, welcher Faktor eine dominante Rolle bei der Selektion von *starting points* während des Planungsvorgangs spielt.

Im zweiten Experiment, das in Kapitel 7 dargestellt wird, kommt als weiterer Faktor zu den Faktoren aus Experiment 1 die kontextuelle Salienz hinzu. Die Wirksamkeit,

die Repräsentation und die Art der Verarbeitung der kontextuellen Information (und insbesondere des Topiks) werden untersucht. Mithilfe des Vorwissens des Informationsflusses, welches in Kapitel 2 vorgestellt wurde, wird dem Abbildungsprozess der kontextuellen Information innerhalb der grammatischen Enkodierung nachgegangen. Aus diesen Ergebnissen sollte die Arbeitsweise des Satzplanungssystems in kontextuellen Situationen abgeleitet werden. Darüber hinaus werden die Effekte der kontextuellen Salienz mit denen der visuellen und semantischen Salienz verglichen. Ziel des Vergleichs ist es, eine Prioritätenliste, welche die Verarbeitungshierarchie verschiedener Faktoren bei der Satzplanung beschreiben kann, zu entwickeln.

In Kapitel 8, in dem das dritte Experiment vorgestellt wird, steht die Untersuchung zu den Mechanismen der Mikroplanung, bei der sowohl die semantische Struktur als auch die Informationsstruktur generiert werden, im Vordergrund. Dabei geht es darum, aufzuzeigen, ob die kontextuelle Information wie beispielsweise das Topik, zunächst auf die semantisch strukturelle Repräsentation abgebildet wird, bevor sie in die grammatische Enkodierungsebene überführt wird.

In Kapitel 9 schließt die Arbeit mit einer allgemeinen Diskussion der Ergebnisse ab. Zusammenfassend sollen noch einmal die im Zuge der Arbeit gewonnenen Erkenntnisse über die Arbeitsweise des Satzproduktionssystems und die Rolle der kontextuellen Salienz in dem Planungsprozess sowie das Verhältnis der untersuchten drei Faktoren dargestellt werden. Im Anschluss wird dann versucht, ein eigenes Modell zur Skizzierung der Selektionsmechanismen von *starting points* aufzustellen. Abschließend wird ein Ausblick auf zukünftige Untersuchungen geboten.

2 Modellierung der Satzproduktion

Bevor wir auf das Satzproduktionsmodell eingehen, betrachten wir zuerst die folgenden zwei mit dem Substantiv *Bank* gebauten Sätze:

(1a) Der Student eröffnete ein Konto bei der Bank.

(1b) Der Student setzt sich auf eine Bank.

Das Wort *Bank* ist ein einfaches Wort im Deutschen und bezeichnet zum einen etwas, woauf man sich setzen kann, und zum anderen ein Unternehmen, das Geld- und Kreditgeschäfte betreibt. Trotz der Mehrdeutigkeit dieses Wortes wird es in beiden Sätzen richtig interpretiert und nicht mehrdeutig aufgefasst. Dahinter steht die Annahme, dass Wörter im mentalen Lexikon nicht als einzelne unabhängige Einheiten, sondern mit anderen in semantischer Beziehung stehenden Konzepten in einem Netzwerk abgespeichert werden. So steht das Wort *Bank* einerseits als Finanzunternehmen in Zusammenhang mit Konzepten wie *KONTO*, *GELD* oder *ÜBERWEISUNG*; andererseits wird das Substantiv *Bank* als Sitzgelegenheit gemeinsam mit statischen Handlungen wie z. B. *SITZEN* oder *LIEGEN* gespeichert. Dank der konzeptuellen Zusammenhänge wird sichergestellt, dass die Bedeutung des Wortes *Bank* in Satz (1a) nicht mit der in Satz (1b) verwechselt wird. Die beiden Beispiele deuten darauf hin, dass sich hinter der Oberflächenstruktur und den Bedeutungen einer sprachlichen Formulierung immer eine konzeptuelle Struktur versteckt. Der Prozess, bei dem begriffliche und strukturelle Repräsentationen versprachlicht werden, wird als Sprachproduktionsprozess bezeichnet.

Die Modellierung der Sprach- bzw. Satzproduktion, auf die in diesem Kapitel eingegangen wird, befasst sich vor allem mit der Umwandlung konzeptueller Repräsentationen in mündliche oder schriftsprachliche Texte (Levelt, 1989, 1993, 1994, 1995; vgl. auch Bock, 1982; Dell, 1986; Dietrich, 2002; Ferreira & Slevc, 2007; Garrett, 1975, 1988; Vigliocco & Hartsuiker, 2002). Ich werde mich im Folgenden auf die mündliche Sprachproduktion beschränken. Als theoretische

Grundlage zur Darstellung der Sprechabläufe werden das Sprachproduktionsmodell von Levelt (1989) und das daraus weiterentwickelte Satzproduktionsmodell von Bock und Levelt (1994) herangezogen. In dem Modell von Levelt werden drei Komponenten oder Stufen des Sprachproduktionsprozesses unterschieden: die Konzeptualisierungs-, die Formulierungs- und die Artikulationsstufe. In dem Konzeptualisierer wird die Botschaft inhaltlich geplant, welche dann als Input an den Formulator übergeben und dort sowohl grammatisch als auch phonologisch ausformuliert wird. Der Output der Formulierung ist der Artikulationsplan, mithilfe dessen ein physikalisch wahrnehmbarer Sprechakt produziert wird. Basierend auf dem allgemeinen Modell von Levelt wurde ein weiteres Modell, das die Satzproduktion spezifiziert, von Bock und Levelt (1994) entwickelt. Der Schwerpunkt dieses Modells wird von beiden Psycholinguisten auf die grammatische Enkodierung der Botschaft gelegt, die sich in zwei Verarbeitungsebenen, die funktionale und die positionelle Enkodierung, gliedert. Das Verhältnis zwischen den beiden Modellen wird durch Abbildung 2.1 veranschaulicht.

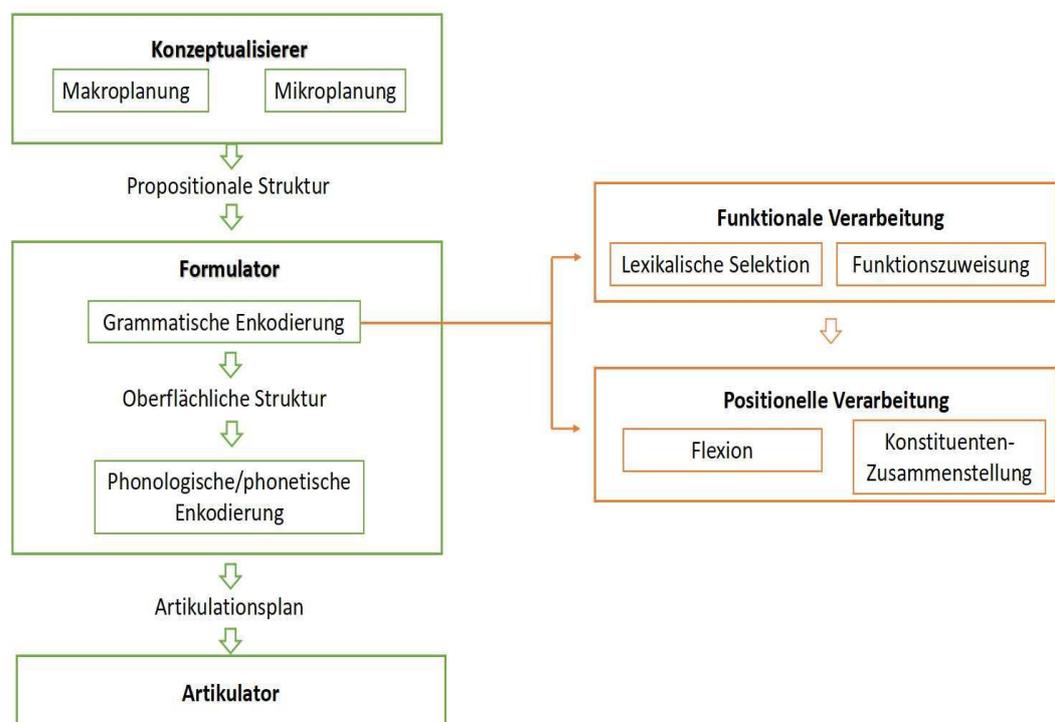


Abbildung 2.1: Satzproduktionsmodell adaptiert von Levelt 1989 und Bock und Levelt 1994

Es ist zu beachten, dass, obwohl der Satzproduktionsprozess schematisch durch aufeinander folgende Module illustriert wird, die Verarbeitung in den jeweiligen Modulen jedoch nicht strikt nacheinander, sondern auch parallel abläuft (vgl. De Smedt, 1990; Dietrich, 1999; Kempen & Hoenkamp, 1982, 1987; Levelt, 1989, 1994; Roelof, 1998; Wheeldon, Meyer & Smith, 2003). Wie zahlreiche Experimente gezeigt haben, planen Sprecher nicht zuerst die gesamte Satzeinheit auf der konzeptuellen Ebene, bevor dann die nachfolgende grammatische Verarbeitung beginnt. Sobald eine Verarbeitungseinheit, ein Inkrement, auf einer höheren Verarbeitungsebene generiert worden ist, wird es unmittelbar an die anschließende Ebene weitergereicht. Während auf einer Verarbeitungsebene noch an einem Inkrement gearbeitet wird, kann bereits auf einer höheren Ebene an einem neuen Inkrement gearbeitet werden. Diese Eigenschaft des Satzproduktionsprozesses wird als Inkrementalität bezeichnet, auf die ich in Kapitel 3 genauer eingehen werde. Im Folgenden werden zuerst die jeweiligen Verarbeitungsebenen der Satzproduktion veranschaulicht. Besonderes Gewicht wird dabei auf die Konzeptualisierung und die grammatische Enkodierung gelegt.

2.1 Konzeptualisierung

Bevor nun nachfolgend auf die Konzeptualisierungsebene eingegangen wird, ist festzuhalten, dass, obwohl das Satzproduktionsmodell von Bock und Levelt (1994) eine aussagekräftige Erklärung für die Satzproduktion liefert, es sich schwerpunktmäßig mit der grammatischen Enkodierung beschäftigt. Die Konzeptualisierungsebene, auf der propositionale Strukturen als Input der linguistischen Verarbeitung ausgeführt werden, wurde in diesem Modell allerdings nur wenig beleuchtet. Da das Sprachproduktionsmodell von Levelt (1989) eine globale Erklärung für alle Komplexitätsstufen der sprachlichen bzw. mündlichen Rede bietet, die sich von der Wortproduktion bis zur Satz- und Diskursproduktion erstrecken, liefert es als übergeordnetes Modell eine gute theoretische Ergänzung für das Modell von Bock und Levelt (1994). So wird die Konzeptualisierung in diesem

Teil in dem theoretischen Rahmen von Levelts Modell vorgestellt.

Der Konzeptualisierer ist im Wesentlichen „Nervenzentrum“ der Sprachproduktion, in dem alles, vom Inhalt bis hin zur Art des Sprechens, geplant und organisiert wird. Sprecher stellen sich in dieser Phase hauptsächlich folgende Fragen: Warum spreche ich? Was spreche ich? Wie spreche ich? – Von diesen Fragen ausgehend wird nun der Prozess der Konzeptualisierung eingehender betrachtet.

Menschen sprechen immer aus einem bestimmten Grund. Dieser Grund wird hier als *kommunikative Intention* bezeichnet. Wenn ein Tourist in der Stadt nach dem Weg zum Universitätsplatz fragt, liegt die Intention seitens des Gefragten darin, diese Frage hinreichend zu beantworten und dem Touristen somit helfen zu können. In Gefahrensituationen beispielsweise ist die kommunikative Absicht eines laut geäußerten *Halt!* eine Warnung über eine mögliche Gefahr zu geben. Äußerungsintentionen können aktiv und passiv getriggert werden und sind unendlich vielfältig.

Ist eine Äußerungsintention gegeben, müssen Sprecher in einem weiteren Schritt eine konzeptuelle Äußerungsbasis erzeugen, die eine gedankliche Struktur für die linguistische Verarbeitung vorgibt. Nach Levelt (1989) gelingt die Erstellung solcher Äußerungsbasis aufgrund zweier Verarbeitungsprozesse: der Makroplanung und der Mikroplanung. Die Definitionen dieser beiden Begriffe werden anhand eines Zitats verdeutlicht:

Macroplanning involves the elaboration of some communicative goal into a series of subgoals, and the retrieval of the information to be expressed in order to realize each of these subgoals. Microplanning assigns the right propositional shape to each of these „chunks“ of information, as well as the informational perspective (the particular topic and focus) that will guide the addressee's allocation of attention.
(Levelt 1989, S. 11)

Während des Makroplanungsprozesses wird zuerst an der selektierten kommunikativen Intention gearbeitet. Dabei handelt es sich vor allem um eine eventuell notwendige Aufspaltung der Intentionen in Teilziele. Basierend darauf

werden die zu verbalisierenden Inhalte aus der Wissensbasis festgelegt. Altmann und Trafton (2002) und Levelt (1999b) zufolge sind Informationen, die während dieses Prozesses ausgewählt werden, *lexikalische Konzepte*. Unter dem lexikalischen Konzept werden abstrakte konzeptuelle Repräsentationen verstanden, die einen Zusammenhang mit den im mentalen Lexikon der Zielsprache gespeicherten Worteinträgen aufweisen. Zur Veranschaulichung ein Beispiel: *Die Stute jagt den Hengst*. Um diesen Satz zu generieren, müssen Sprecher die lexikalischen Konzepte *STUTE*, *HENGST* und *JAGEN* mental aufrufen. Dies ist allerdings nicht immer einfach. Ein zentrales Problem bei der Aktivierung lexikalischer Konzepte, welches von Levelt und Kollegen als *verbalisation problem* bezeichnet wurde, ist, dass es kein einfaches eins-zu-eins-Mapping der zu verbalisierenden Information auf die präverbale Message gibt (Levelt, Roelofs & Meyer 1999, S. 3). Ein Beispiel für das Problem sind Formulierungen wie beispielsweise *Stute* versus *weibliches Zebra*: Während zur Verbalisierung von *Stute* ein einziges lexikalisches Konzept, welches die konzeptuellen Merkmale *+WEIBLICH* und *+PEFERD* trägt, zur Verfügung steht, gibt es im Deutschen kein lexikalisches Konzept, in dem sich die Merkmale *weiblich* und *Zebra* vereinigen.

Da in der Regel ein lexikalisches Konzept mit mehreren lexikalischen Einträgen in Verknüpfung steht, müssen Sprecher nach dessen Generierung weiter entscheiden, mit welchem Eintrag das Konzept zu realisieren ist. So kann zum Beispiel der Sprecher das Konzept *STUTE* auch als *Tier* oder *Pferd* verbalisieren. Welcher lexikalische Eintrag zum Schluss zur lexikalischen Enkodierung des selektierten Konzeptes ausgewählt wird, ist nach Levelt (1999b) von den Situationen oder Kontexten, in denen sich die Satzproduktion vollzieht, abhängig. Darüber hinaus befindet sich ein lexikalisches Konzept immer in einem semantisch zusammenhängenden Netzwerk, welches von Roelofs (1992) als *konzeptuelles Stratum* (engl.: *conceptual stratum*) bezeichnet wird. In diesem Stratum ist jedes lexikalische Konzept z. B. *JAGEN* durch einen konzeptuellen Knoten repräsentiert. Dazu wird es auch mit konzeptuellen Anbindungen „etikettiert“, die konzeptuelle

Relationen ausführen (z. B. *JAGEN* und *VERFOLGEN*). Die Verbindung zwischen semantisch verwandten lexikalischen Konzepten führt dazu, dass, sobald ein Konzept aktiviert wurde, andere Konzepte, die dem gleichen Netzwerk zugeordnet sind, zum Teil mit aktiviert werden.

Im Anschluss an die Makroplanung wird die generierte Information in der Mikroplanung zu einer vorsprachlichen Bedeutungsstruktur weiterverarbeitet. Hauptaufgabe auf dieser Ebene ist es, eine propositionale Form für die Botschaft zu erstellen. Hierzu wird zuerst eine semantische Struktur aufgebaut, durch die die verbalisierenden Referenten mit bestimmten semantischen Funktionen (z. B. Agens, Patiens, Instrument) belegt werden. Zu beachten ist, dass noch Uneinigkeit darüber herrscht, ob die semantische Struktur – wie Levelt (1989) behauptet – mit der konzeptuellen Struktur gleichgesetzt wird und daher der konzeptuellen Ebene zugeordnet werden kann. Diesbezüglich bestehen mittlerweile verschiedene Annahmen. Relevant hierbei sind die konzeptuelle Semantiktheorie von Jackendoff (1983) und die Zwei-Stufen-Semantik von Bierwisch (1983) und Bierwisch und Lang (1987). Im Rahmen der konzeptuellen Semantiktheorie wird von Jackendoff angenommen, dass die semantische Struktur mit der konzeptuellen gleichbedeutend ist. Sprache ist kein eigenständiges System, sondern ein Subsystem des konzeptuellen Wissens. Dementsprechend ist die semantische Struktur nach Jackendoff eine Variante konzeptueller Struktur, die zum Zweck der Versprachlichung aufgebaut wird. Entgegengesetzt geht die Zwei-Stufen-Semantik nach Bierwisch und Lang davon aus, dass kognitive und sprachliche Leistungen des Sprechers zwei verschiedene Module der menschlichen Intelligenz sind und daher nicht undifferenziert betrachtet werden dürfen. Die semantische Struktur eines Ereignisses, bei Jackendoff als Subsystem der konzeptuellen Struktur aufgefasst, wird ähnlich wie die syntaktische Struktur von Bierwisch und Lang als ein Teil der Grammatik angenommen. Die Verknüpfung der konzeptuellen und der semantischen Struktur bei der Sprachplanung gelingt durch ein sog. Mappingverfahren, bei dem die konzeptuelle Einzelsachverhaltsvorstellung auf die sprachliche Bedeutung

abgebildet wird (vgl. auch Herweg & Maienborn, 1992). Da ich in der vorliegenden Studie dem Sprachproduktionsmodell von Levelt (1989) folge, dessen Grundgedanke bezüglich der Zuordnung der semantischen Struktur im Einklang mit der konzeptuellen Semantiktheorie von Jackendoff steht, betrachte ich die semantische Konstruktion als einen Bestandteil der konzeptuellen Enkodierung.

Neben dem Aufbau der semantischen Struktur müssen Sprecher in dieser Phase noch entscheiden, in welcher Art und Weise bzw. aus welcher Perspektive Informationen sprachlich dargestellt werden. Diese Entscheidungen fasst Levelt unter dem Begriff *Perspektivierung* (engl.: *Perspectivation*) zusammen, die von v. Stutterheim und Klein (2002, S. 62) auch als *L-Perspektivierung* (engl.: *L-Perspectivation*; *L* für *language*) bezeichnet werden.¹ Dieser Prozess zeichnet sich vor allem durch strukturelle, lexikalische und kontextuelle Selektionen aus, die zwar die Formen von Äußerungen, jedoch nicht ihren Gehalt verändern.² Die Perspektivierung ist derart wichtig, dass sie von Clark (1997) als „the very core of all conceptual preparation for speech“ angesehen wird (zitiert aus Levelt 1999a, S.92).

Angesichts der Bedeutsamkeit der Perspektivierung stellt sich die Frage, weshalb eine Perspektive bevorzugt bzw. vor einer anderen präferiert wird. Die Auswahl einer Perspektive kann etwa auf die Zugänglichkeit von Referenten zurückzuführen sein. So wird der Satz *Der Hund jagt den Ball* mit höherer Wahrscheinlichkeit ausgewählt, wenn die Entität *Hund* kurz zuvor im Gespräch erwähnt wurde, da durch die Vorerwähnung die Zugänglichkeit eines bestimmten Referenten erhöht wird. Der kognitive Aufwand wird somit auch reduziert, wenn vorerwähnte Referenten früher verarbeitet werden. Natürlich ist die Vorerwähnung nicht der einzige Faktor für eine

¹ Es ist zu beachten, dass die Gültigkeiten der Perspektivierung nach v. Stutterheim und Klein und nach Levelt unterschiedlich sind. V. Stutterheim und Klein zufolge ist die Perspektivierung für alle getroffenen Entscheidungen einschließlich lexikalischer, syntaktischer und kontextueller Entscheidungen zuständig, während sie nach Levelt sich lediglich auf die Auswahl spezifischer pragmatischer Relationen begrenzt.

² Eine Ausnahme hier sind Ereignisse, die sich mit perspektivischen Prädikatspaaren wie *kaufen* und *verkaufen* beschreiben lassen. So kann man eine geschäftliche Transaktion entweder mit *kaufen* (aus der Perspektive des Käufers) oder mit *verkaufen* (aus der Perspektive des Verkäufers) beschreiben. Dabei werden unterschiedliche Gehalte vermittelt.

frühe Positionierung eines Referenten. Darüber hinaus geht die Selektion einer Perspektive mit der Festlegung der Topik-Kommentar-Struktur einher. Der Topik-Begriff bezeichnet das, worüber etwas ausgesagt wird (vgl. Hocket, 1958). Durch die Festlegung des Topiks wird auch der Bezugsrahmen, in den Sprecher ihre Äußerung setzen, festgelegt. Der Kommentar wiederum bezeichnet das, was über das Topik ausgesagt wird. Ein Beispiel hierfür ist: *Der Hund schläft neben der Katze* und *Die Katze schläft neben dem Hund*. Obwohl die beiden Sätze die gleichen Wahrheitsbedingungen haben, unterscheiden sie sich dahingehend, dass der Sachverhalt aus der Perspektive verschiedener Referenten dargestellt ist. So würde man die Frage *Wo ist der Hund?* primär mit dem ersten Satz beantworten. Die Perspektive kann in dem Gespräch von der Frage festgelegt und durch die Topik-Kommentar-Struktur ausgedrückt werden.

Zum besseren Verständnis über die Festlegung einer Proposition der zu verbalisierenden Äußerung wird die Situation der bereits oben angeführten Wegbeschreibung nochmals herangezogen. Wenn ein Tourist einen Fußgänger Folgendes fragt: *Mit welchem Bus kann ich zum Universitätsplatz fahren?*, werden während der darauffolgenden beginnenden Antwort *Zum Universitätsplatz können Sie mit dem Bus mit der Nummer 32 fahren* im Prinzip vier lexikalische Konzepte aktiviert: – nämlich UNIVERSITÄTSPLATZ, SIE, BUS MIT DER NUMMER 32 und FAHREN. Dies passiert aufgrund der Sprechintention des Gefragten, nämlich einen Weg zum Uniplatz zu beschreiben. Mithilfe solcher Information wird eine semantisch-konzeptuelle Struktur aufgebaut, durch welche die involvierten Bedeutungen miteinander verknüpft werden und gemeinsam ein verständliches Ereignis bilden. Die semantische Strukturierung ist wichtig, da sie „in irgendeiner Weise Zugang zum Sprachsystem“ hat (Härtil 1999, S. 189). Ein Sachverhalt kann in unterschiedlichen syntaktischen bzw. grammatischen Konstellationen realisiert werden, wenn die daran beteiligten Entitäten unterschiedlichen thematischen Rollen zugeordnet werden (z. B. *Andreas fürchtet Moritz* und *Moritz ängstigt Andreas*). In Bezug auf das oben angeführte Beispiel fungiert das lexikalische Konzept SIE in

dem Satz als Agens, der eine Handlung initiiert; das Konzept *BUS MIT DER NUMMER 32* ist das Instrument, mit dem das Agens den Universitätsplatz erreichen kann, und das Konzept *UNIVERSITÄTSPLATZ* ist das Ziel. Die Handlung, die durch das Konzept *FAHREN* dargestellt wird, verbindet dann die thematischen Rollen Agens, Ziel und Instrument miteinander. Nachdem die zu verbalisierenden Informationen festgelegt wurden, muss entschieden werden, aus welcher Perspektive heraus die Antwort geäußert werden soll. Da durch die Frage das Ziel, nämlich *UNIVERSITÄTSPLATZ* in den Fokus des Gesprächsinteresses gestellt wurde und somit mit höherer Wahrscheinlichkeit in der Antwort als Topik dienen wird, wird dieses Konzept demensprechend auch eher als Ausgangspunkt der Antwort angenommen. Selbstverständlich kann der Gefragte auch mit dem Satz *Sie können mit dem Bus 32 zum Universitätsplatz fahren* antworten, in dem das Agens *SIE* als Topik hervortreten kann und alle anderen Informationen als ergänzende Topik-Beschreibungen dienen. Dazu wird für ein Gespräch mit einem Unbekannten wahrscheinlich eine höfliche Formulierungsvariante ausgewählt, die u.a. ein Modalverb enthält (hier *können*). Durch das Modalverb wird auch die temporale Information (*Indikativ Präsens*) dargestellt.

Mit der präverbalen Botschaft als Output der Konzeptualisierung bzw. der Makro- und Mikroplanung wird laut Levelt eine propositionale Struktur gebildet, der ein Bündel semantischer, relationaler und kontextueller Bedeutungen zugeordnet wird (vgl. auch Ferreira & Slevc, 2009; Garrett, 1975; King, 1996; Klein & v. Stutterheim, 1987). Der *innere Kern* der propositionalen Struktur besteht aus der Referenz auf Ereignisse, Vorgänge und Zustände (Prädikat) sowie beteiligte Figuren, während die relationale und kontextuelle Information sowie Angaben zur Modalität, Zeit usw. den *äußeren Kern* ausmachen (vgl. Klein & v. Stutterheim, 1987). Diese Struktur wird weiter auf der grammatischen Enkodierungsebene verarbeitet, auf die ich nun im anschließenden Abschnitt eingehen werde.

2.2 Grammatische Enkodierung

Die grammatische Enkodierung ist ein Prozess, indem die auf der konzeptuellen Ebene generierte präverbale Information lexikalisch und grammatisch erfasst und gereiht wird. Das Ziel dieses Prozesses besteht darin, lexikalische und morphosyntaktische Informationen der von Sprechern ausgewählten Konzepte für die lautliche Verarbeitung bereitzustellen. Er wird von Bock und Levelt (1994) in zwei Unterprozesse eingeteilt: In einem ersten Schritt werden syntaktische Funktionen einzelner Elemente festgelegt (*funktionaler Verarbeitungsprozess*) und in einem zweiten Schritt die syntaktische lineare Sequenz generiert (*positioneller Verarbeitungsprozess*). Durch die funktionale Verarbeitung werden passende Wörter für die zu verbalisierenden Konzepte ausgewählt (*lexikalische Selektion*), die dann weiterhin mit syntaktischen Funktionen (z. B. Subjekt und Objekt) belegt werden (*Funktionszuweisung*). In der darauffolgenden positionellen Verarbeitung werden die Wörter dann flektiert (*Flexion*) und in einer linearen Sequenz angeordnet (*Konstituenten-Zusammenstellung*). Da die vorliegende Arbeit schwerpunktmäßig die Planung der Satzstruktur thematisiert, setze ich folglich den Akzent auf die Funktionszuweisung und die Konstituenten-Zusammenstellung.

2.2.1 Lexikalische Selektion

Wie bereits in Abschnitt 2.1 ausgeführt, besteht die präverbale Botschaft aus verschiedenen Verarbeitungseinheiten, die bestimmte semantische, relationale und kontextuelle Merkmale aufweisen. Ausgehend von der semantischen Information wird eine semantische Suche nach lexikalischen Einheiten, welche den entsprechenden Eigenschaften des lexikalischen Konzepts genügen, im mentalen Lexikon vorgenommen (Zimmermann, 1997). Diese erfolgt unter anderen in zwei Phasen, die sich jeweils mit der Auswahl von *Lemmata* und *Lexemen* beschäftigen (vgl. Levelt, 1999a; Dell, Chang & Griffin, 1999; Mizzo & Caramazza, 1997; Vigliocco, Antonini & Garrett, 1997). Unter dem Lemma versteht man Wortwissen, welches die semantischen und syntaxkategorialen Bedeutungen eines Wortes

aufweist, während das Lexem die Wortform bezeichnet. Die lexikalische Selektion, die sich in dieser Phase vollzieht, beschränkt sich auf die Auswahl von semantischen und syntaktischen Informationen und wird auch als *lemma selection* bezeichnet (vgl. Ferreira & Slevc, 2009).

Studien haben gezeigt, dass lexikalische Einheiten kaum isoliert im mentalen Lexikon gespeichert werden, sondern in den meisten Fällen in einem semantischen oder kategorialen Zusammenhang mit anderen Einheiten stehen (Levelt, 1989; Roelofs, 1992, 2011). Beim Wortzugriff werden auch andere lexikalische Konzepte und Wörter, die Bedeutungsüberschneidungen mit dem Zielwort aufweisen, teilweise mit aktiviert (z. B. *Hund* und *Katze*; vgl. La Heij, 2005; Griffin & Ferreira, 2006). Diese Eigenschaft der lexikalischen Selektion wird beispielsweise in einer Studie von Glaser und Düngelhoff (1984) nachgewiesen, bei der anhand eines Wort-Bild-Interferenz-Paradigmas das Verhältnis zwischen der Bildererkennung und der Wort-Kategorisierung erforscht wurde. In der Studie wurden Probanden visuelle Stimuli von Objekten gezeigt und sie dann dazu aufgefordert, die Objekte so schnell wie möglich zu benennen. Sogenannte Störwörter, die als Distraktoren fungierten und in verschiedenen semantischen Beziehungen zu den visuell dargestellten Sachverhalten standen, wurden unter die Stimuli gemischt. Diesen Beziehungen entsprechend wurden insgesamt vier Bedingungen aufgestellt, wovon drei für die vorliegende Arbeit relevant sind: (1) Die kategorial kongruente Bedingung, in der der Distraktor zu derselben Kategorie der präsentierten Figur gehörte (z. B. *church* und *house*); (2) die inkongruente Bedingung, in der der Distraktor zu einer anderen Kategorie gehört (z. B. *car* und *house*); (3) die neutrale Bedingung, in der anstelle bedeutungstragender Wörter nur bedeutungslose Symbole als Distraktor verwendet wurden (z. B. *xxxxxxx* und *house*). Eine Gruppe der Versuchspersonen wurde instruiert, die dargestellten Objekte ohne Rücksicht auf den Distraktor so schnell wie möglich zu benennen. Es zeigte sich, dass die Benennzeit von der semantischen Beziehung zwischen dem Distraktor und dem präsentierten Objekt abhängig ist. Gehört der Distraktor wie in der ersten Bedingung, derselben Kategorie wie die

dargestellte Figur an, wurde die Benennung deutlich verlängert gegenüber den anderen zwei Bedingungen. Diese Effekte wurden von den Autoren als Hinweis darauf gesehen, dass andere mit den Targets semantisch verwandte Informationen bei der Selektion eines Worteintrags mit aktiviert werden. Angesichts dieser Eigenschaft ist die lexikalische Selektion daher ein wettbewerblicher Ablauf, bei dem der Abruf eines bestimmten Wortes vom Verhältnis der Lemma-Aktivierung und der Aktivierung der gesamten in Beziehung stehenden Lemmata abhängt. Die Wörter, deren Aktivierungsniveau am höchsten ist, werden zum Schluss als Targets ausgewählt (vgl. Dell, 1986; Rapp & Goldrick, 2000).

Lexikalische Einträge sind nicht nur ein Bündel semantischer Merkmale, sondern auch mit spezifischen syntaktischen Eigenschaften wie etwa grammatischen Kategorien wie dem Genus versehen. Zum Beispiel ist das Wort *Bleistift* bezüglich seiner Semantik ein Schreibgerät und bezüglich seiner grammatischen Eigenschaft ein Maskulinum. Nach der *lexicalist theory* werden grammatische Eigenschaften, mit deren Abruf eine Voraussetzung für weitere Funktionszuweisungen und morphologische Anpassungen geschaffen wird, zusammen mit den semantischen und phonologischen Informationen eines Wortes beim Satzbau automatisch mit abgerufen (vgl. z. B. Levelt, 1989; Kempen & Hoenkamp, 1987; Pickering & Branigan, 1998).

2.2.2 Funktionszuweisung

Wie bereits in Abschnitt 2.1 ausgeführt, werden auf der konzeptuellen Ebene nicht nur einzelne lexikalische Konzepte, sondern auch strukturelle Informationen generiert, durch die die Konzepte zur Verknüpfung gebracht werden. Die strukturelle Information lässt sich auf verschiedenen Ebenen unterschiedlich definieren. Auf der semantischen Ebene bezieht sie sich auf die semantische Struktur, in der Konzepte nach deren semantischen Relationen mit den anderen in einem bestimmten Ereignis verschiedene semantische Rollen erhalten. Zum Beispiel werden den Figuren, die eine Tätigkeit oder eine Handlung initiieren, die Agen-Rolle zugewiesen und die,

die von einem Vorgang betroffen sind, als Patiens bezeichnet. Auf der kontextuellen Ebene wird die strukturelle Information als Informationsstruktur aufgefasst, die die Art und Weise der Informationsdarstellung spezifiziert und das augenblickliche Interesse und den Informationsstand des Adressaten darstellt (z. B. Topikalität oder Gegebenheit; vgl. Haliday, 1967). In der Satzproduktion müssen auch strukturelle Informationen verbalisiert werden, wofür die Funktionszuweisung und die Konstituenten-Zusammenstellung zuständig sind. Im nachstehenden Abschnitt wird zuerst die Zuweisung syntaktischer Funktionen beleuchtet.

Die Satzproduktion verläuft inkrementell, d.h. sobald die Konzepte auf der konzeptuellen Ebene aufgerufen worden sind, werden sie unmittelbar zur weiteren Enkodierung an die funktionale Verarbeitungsebene übergeben. Dementsprechend vollzieht sich die syntaktische Funktionszuweisung auch sukzessiv, d.h. syntaktische Funktionen werden nicht gleichzeitig, sondern nacheinander zugeteilt. In diesem Zusammenhang stellt sich die Frage, ob es eine bestimmte Reihenfolge gibt, nach der syntaktische Funktionen zugewiesen werden, oder ob der Zuweisungsprozess unstrukturiert geschieht.

Eine mögliche Antwort findet man in der *Zugänglichkeitshierarchie* (engl.: *accessibility hierarchy*) von Keenan und Comrie (1977). Die Studie handelt von der typologischen Forschung zum grammatischen Verhalten von Relativsatzkonstruktionen, präsentiert aber auch Anregungen zur Lösung solcher Fragen bezüglich der sprachlichen Verarbeitung. Der Fokus der Studie liegt auf der Identifikation jeglicher Art von Nominalphrasen in einer gegebenen Sprache, deren Bestimmung durch Relativsätze zugänglich ist. Durch den Vergleich von 50 Sprachen wurde herausgefunden, dass nicht alle Nominalphrasen in demselben Ausmaß durch einen Relativsatz determiniert werden können. Stattdessen besteht hierbei eine hierarchische Abstufung, die die Zugänglichkeit der sechs untersuchten syntaktischen Funktionen zur Determination darstellt (Die Funktionen vor dem Symbol „>“ stehen in einem höheren Rang als die hinteren):

Subjekt > Direktes Objekt > Indirektes Objekt > Obliques Objekt > Possessor > Vergleichsobjekt (Keenan & Comrie 1977, S. 66)

Es zeigt sich, dass insgesamt das Subjekt am häufigsten näher durch einen Relativsatz bestimmt werden kann als alle Objektarten. Da die vorliegende Arbeit die Satzproduktion thematisiert, stellt sich im Hinblick darauf die Frage, ob sich die oben erwähnte Hierarchie mit dem syntaktischen Funktionszuweisungsprozess verknüpfen lässt, um eine primäre Zuweisungsreihenfolge beschreiben zu können. In der Tat haben einige psycholinguistische Experimente gezeigt, dass das Subjekt in der Regel früher „zugeteilt“ wird als das Objekt. Ein Beispiel dafür ist eine semantische Priming-Studie von Bock (1986), in der festgestellt wurde, dass englische Sprecher beim Beschreiben eines transitiven Ereignisses dazu tendieren, die Figuren, deren semantische Eigenschaften durch das vorige Zeigen eines anderen Wortes voraktiviert wurden, eher als Subjekt zu realisieren als als Objekt. Ein vergleichbares Ergebnis war auch in einem weiteren Experiment von Bock (1987a) zu beobachten, bei dem die Wörter, deren phonologische Eigenschaften zuvor aktiviert und daher früher selektiert und verarbeitet wurden, häufiger als Subjekt anstatt als Objekt aufgetreten sind. Insgesamt stehen die genannten psycholinguistischen Befunde im Einklang mit der Zugänglichkeitshierarchie von Keenan und Comrie (1977). Demnach bleibt an dieser Stelle festzuhalten, dass der Subjektstatus bei der syntaktischen Funktionszuweisung eine zentrale Rolle spielt und während dieses Prozesses zumeist früher zugewiesen wird als der Objektstatus. Das wirft nun die Frage auf, wie diese syntaktische Funktionsvergabe überhaupt gelingt?

Bock und Levelt (1994) zufolge werden semantische Rollen, sobald auf der konzeptuellen Ebene hervorgerufen, an die syntaktische Strukturierung weitergereicht und syntaktischen Funktionen zugeordnet, was eine Verknüpfung zwischen der semantischen und syntaktischen Struktur darstellt. In welcher Art und Weise semantische Rollen auf die funktionale Verarbeitungsebene abgebildet werden, wird nach Fillmore (1968, 1977), Jackendoff (1990) und auch Primus (1996) von der

thematischen Hierarchie, in der thematische Rollen gemäß einer Salienzskala (mit verschiedenen Dimensionen wie etwa *Aktivität, Handlungsfähigkeit, Involviertheit*; vgl. Zifonun, Hoffmann & Strecker, 1997) angeordnet werden, bestimmt. Rollen, die semantisch prominenter sind und somit einen höheren Status in der Hierarchie haben (z. B. Agens), werden grammatisch früher enkodiert als diejenigen, die weniger prominent sind (z. B. Instrument). Eine nähere Auseinandersetzung mit der semantischen Hierarchie findet in Abschnitt 4.1 statt. Da die semantischen Rollen, ähnlich wie die syntaktischen Funktionen, gemäß deren Prominenz nach einer primären Reihenfolge abgerufen und auf die syntaktische Enkodierungsebene abgebildet werden, wird deswegen oft eine Korrespondenz zwischen bestimmten semantischen und syntaktischen Funktionen vermutet. Beispielsweise findet man in unmarkierten Sätzen das Agens sehr häufig in Subjektposition und das Patiens wird zumeist als Objekt realisiert. Allerdings ist auch zu beachten, dass trotz der Primärstruktur semantische Funktionen auch andere syntaktische Funktionen als die bevorzugte besetzen können. Ein Beispiel dafür ist die Passivkonstruktion. In dieser Konstruktion kann das Agens entweder eine *propositional-object*-Position besetzen oder in der syntaktischen Struktur gar nicht auftauchen. Dementsprechend kann die Subjektposition auch von einer anderen semantischen Rolle wie etwa dem Patiens okkupiert werden.

Neben der Abbildung der semantischen Struktur auf die syntaktische Ebene scheint die Verbselektion auch eine Rolle bei der Funktionszuweisung zu spielen. Der folgende Satz *Thomas gibt Jana einen Schlüssel* kann problemlos mit dem Satz *Jana kriegt einen Schlüssel von Thomas* beschrieben werden. Obwohl die beiden Sätze ein identisches Ereignis beschreiben, unterscheiden sich die syntaktischen Funktionen der Ereignisteilnehmer aufgrund der verschiedenen Verben voneinander. Ein anderes Beispiel ist das Wortpaar *give* und *donate* im Englischen. Das Verb *give* erlaubt die Funktionszuweisung des Empfängers sowohl als indirektes Objekt als auch als Präpositionalobjekt, während in einem Satz mit dem Verb *donate* der Empfänger lediglich als Präpositionalobjekt realisiert werden kann. Es ist jedoch zu beachten,

dass noch keine Einigkeit darüber herrscht, ob das Verb vor oder erst nach der Funktionszuweisung aktiviert wird. Frühere Studien haben einen einzelsprachbezogenen Unterschied festgestellt. Die Studie von Hwang und Kaiser (2014) hat beispielsweise zeigen können, dass das Verb im Englischen vor Sprechbeginn abgerufen wird und daher bei der Bestimmung der syntaktischen Struktur von Bedeutung ist. Es wurden keine vergleichbaren Verb-Effekte in der koreanischen Satzproduktion gefunden. Momma, Slevc und Phillips (2015) konnten auf der japanischen syntaktischen Struktur Verb-Effekte feststellen, dies jedoch nur bei der Produktion von Sätzen mit einer OSV-Struktur. Da der zeitliche Verlauf der Verb-Selektion kein Schwerpunkt dieser Arbeit ist, wird darauf nicht weiter eingegangen.

Darüber hinaus können kontextuelle Informationen ebenso relevant für die Zuweisung syntaktischer Funktionen sein, was beispielsweise oftmals durch Fragestellungen gelingen kann. Wenn ein bestimmtes Konzept aufgrund einer zuvor gestellten Frage in den Fokus des Gesprächsinteresses rückt (z. B. *Was passiert mit der Frau?*), wird diesem Konzept (*die Frau*) als Topik der Antwort mit höherer Wahrscheinlichkeit der Subjektstatus zugeteilt (vgl. Bock, 1977; Carroll, 1958; Johnson-Laird, 1968; Prentice, 1966; Tannenbaum & Williams, 1968). Außerdem haben psycholinguistische Experimente und Korpusstudien Evidenzen dafür geliefert, dass es einen Zusammenhang zwischen der Gegebenheit und bestimmten syntaktischen Funktionen gibt (z. B. Clark & Haviland, 1977; Arnold, Wasow, Losongco & Ginstrom, 2000). Ein Beispiel dafür ist die Korrespondenz zwischen vorerwähnten Referenten und Subjekten im Englischen. Gegebene Informationen sind aufgrund der Vorerwähnung im Vergleich zu neuen Informationen leichter zugänglich und können daher früher auf der Funktionszuweisungsebene auftauchen. Infolgedessen werden diese vorerwähnten Referenten eher mit der syntaktisch höherrangigeren Funktion, nämlich dem Subjektstatus, belegt.

Neben semantischen und kontextuellen Informationen, die jeweils durch semantische Extraktion und Perspektivierung an Konzepte verteilt werden, spielen die inhärenten

Eigenschaften der Konzepte, die nicht abhängig von der Satzproduktionsumwelt sind, auch eine Rolle bei der grammatischen Enkodierung. Die am häufigsten untersuchten inhärenten Eigenschaften sind konzeptuelle und wahrnehmende Eigenschaften, deren Einflüsse als konzeptuelle und perzeptuelle Zugänglichkeit aufgefasst werden. Unter der konzeptuellen Zugänglichkeit versteht man, wie leicht ein Referent bzw. ein Wort aus dem Langzeitgedächtnis abgerufen werden kann. Relevant hierfür sind Eigenschaften wie Belebtheit, Konkretheit und Vorstellbarkeit. Eine ganze Reihe experimenteller Untersuchungen hat gezeigt, dass konzeptuell zugänglichere Figuren häufiger als Subjekt auftreten als konzeptuell weniger zugängliche Figuren (z. B. Bock, Loebell & Morey, 1992; Tanaka, Branigan, McLean & Pickering, 2011; Branigan, Pickering & Tanaka, 2008; van Nice & Dietrich, 2003; Kempen & Harbusch, 2004). Zur Interpretation dieses Phänomens behaupten Bock und Warren (1985), dass zugänglichere Konzepte eine höhere Vorhersagbarkeit haben bzw. besitzen, was wiederum einen schnelleren Zugriff ermöglicht. Mit der Vorhersagbarkeit ist die Menge konzeptueller Relationen gemeint, durch die ein Konzept mit anderen Konzepten in Verbindung gebracht werden kann (Keil, 1979). Belebte Figuren wie *ein Mann* oder *eine Frau* sind einfacher aufzurufen, da sie in massenhaften konzeptuellen Relationen (z. B. *schlafen, essen, fernsehen, baden, fahren, rauchen*) auftreten können. Unbelebte Figuren hingegen wie zum Beispiel *eine Brille* können nur mit wenigen konzeptuellen Verbindungen (z. B. *tragen, putzen* oder *abnehmen*) verlinkt werden (und sind somit beschränkt zugänglich).³ Da konzeptuell zugänglichere Figuren aufgrund ihrer höheren Vorhersagbarkeit schneller aufgerufen und dementsprechend früher in die Funktionszuweisung überführt werden können, werden sie daher mit höherer Wahrscheinlichkeit als Subjekt realisiert, da das Subjekt wegen seiner höheren grammatischen Prominenz früher zugewiesen wird als andere Funktionen. Eine ähnliche Tendenz findet man auch bei visuell salienten Referenten, die wegen

³ Es ist zu beachten, dass die Wortfrequenz auch einen Einfluss auf die konzeptuelle und lexikalische Zugänglichkeit haben kann.

ihrer Auffälligkeit in Farben, Größen, Orientierungen usw. schneller wahrgenommen und konzeptuell verarbeitet werden. Eine nähere Betrachtung über die Funktion der visuellen Zugänglichkeit bei der Satzproduktion findet man in Abschnitt 4.2.

Manche Psycholinguisten plädieren dafür, dass die genannten Faktoren nicht die Funktionszuweisung, sondern die Positionierung von Konstituenten beeinflussen. Da diesbezüglich die meisten vorhandenen Ergebnisse aus Studien zur englischen Sprache stammen, in der das Subjekt in der Regel die initiale Satzposition belegt, können unter Umständen falsche Schlussfolgerungen gezogen werden: Anstelle eines Effekts auf die Funktionszuweisung hat die konzeptuelle oder visuelle Zugänglichkeit wahrscheinlich einen direkten Effekt auf die Linearisierung (vgl. Cai, Pickering & Branigan, 2012; Shin & Christianson, 2009). Dieses Bedenken eröffnet eine neue Möglichkeit zur Beschreibung der Verarbeitungsmechanismen der grammatischen Enkodierung – statt über die funktionale Verarbeitungsebene werden konzeptuelle Repräsentationen direkt auf die positionelle Verarbeitungsebene abgebildet, was für eine direkte Verbindung zwischen der Konzeptualisierungsebene und der Linearisierungsebene spricht. So ergeben sich widersprüchliche Annahmen zur Funktionsweise der grammatischen Enkodierung, die im anschließenden Abschnitt und Abschnitt 2.4 behandelt werden.

Zusammenfassend lässt sich sagen, dass im Prozess der syntaktischen Funktionszuweisung Selektionen passender syntaktischer Funktionen stattfinden. Dieser Prozess läuft inkrementell ab, wobei syntaktische Funktionen nicht gleichzeitig, sondern nacheinander und vermutlich in einer bestimmten Reihenfolge zugewiesen werden. Welche Figur zuerst syntaktisch enkodiert wird, ist von deren semantischen, kontextuellen und konzeptuellen sowie visuellen Merkmalen abhängig. Gemeinsam mit der lexikalischen Selektion bildet die syntaktische Funktionszuweisung die funktionale Verarbeitungsebene.

2.2.3 Konstituenten-Zusammenstellung

Ein Satz ist im Wesentlichen eine lineare Sequenz aus lexikalisch, grammatisch und

phonologisch enkodierten Konzepten bestehend. Nachdem Konzepte und konzeptuelle Strukturen in der funktionalen Verarbeitungsphase lexikalisch und grammatisch erfasst wurden, müssen sie nun in eine Reihenfolge gebracht werden (*Linearisierung*). Der Linearisierungsprozess, durch welchen Konstituenten-Strukturen oder Wortstellungen hervorgebracht werden, wird Konstituenten-Zusammenstellung genannt und generell von Levelt (1989) als „[...] deciding what to say first, what to say next, and so on“ zusammengefasst (S. 138). Demzufolge besteht die zentrale Aufgabe dieser Verarbeitungsstufe darin, Einzelsachverhaltsrepräsentationen zusammen mit den zugeteilten semantischen, syntaktischen und eventuell kontextuellen Merkmalen in lineare Anordnungen zu transformieren.

In Bezug auf die Konstituenten-Zusammenstellung thematisieren Studien in den letzten Jahren überwiegend den Ursprung von Wortstellungen. Untersucht wurden hier vor allem, ob die Erarbeitung der Wortstellung direkt unter der Kontrolle der konzeptuellen Verarbeitung steht oder durch die Funktionszuweisung modifiziert wird.

Die wichtigste und meist verwendete Methode zur Beantwortung der Frage ist das strukturelle (bzw. syntaktische) Priming-Paradigma. Dieses Paradigma geht von einem zentralen Phänomen der Sprachproduktion aus, und zwar der Tendenz seitens der Sprecher, die in vorigen Äußerungen verwendeten syntaktischen Strukturen in aktuellen Kommunikationssituationen zu repetieren. In einer typischen strukturellen Priming-Aufgabe werden Versuchspersonen instruiert, einen gehörten Prime-Satz zu wiederholen und direkt danach eine visuell dargebotene Szene zu beschreiben. Da ein Satz auf verschiedenen Ebenen unterschiedliche Strukturen, z. B. thematische und kontextuelle Strukturen, aufweist, kann aufgrund der Beobachtung, welche Struktur des Prime-Satzes im Target-Satz wiederholt wird, nachvollzogen werden, auf welcher Ebene das Priming aufgetreten ist. Diese Ebene wird dann psycholinguistisch als der Ursprung der Wortstellung identifiziert.

Es ist davon auszugehen, dass es mehrere Antworten auf die Frage bezüglich des Ursprungs der Wortstellung gibt, denn es liegen unterschiedliche empirische Befunde dazu vor, welche Strukturen bei dem Linearisierungsprozess befolgt werden. Generell lassen sich diese in zwei Gruppen aufteilen: konzeptuelle und grammatische Strukturen. In Bezug auf die konzeptuelle Struktur lässt sich weiter zwischen der semantischen und der kontextuellen Struktur unterscheiden (vgl. Abschnitt 2.1). Wenn die Wortstellung der konzeptuellen Ebene entstammt, sollen eine geteilte semantische Struktur oder Informationsstruktur zwischen dem Prime- und dem Target-Satz gefunden werden. Wenn die Linearisierung andererseits jedoch in keiner direkten Verbindung mit der Konzeptualisierung steht und stattdessen den Aufbau der syntaktischen Struktur als Bedingung verlangt, sollten Priming-Effekte nur auf der linguistischen Enkodierungsebene beobachtet werden können.

Die erste Annahme, dass die Konstituenten-Abfolge der konzeptuellen Verarbeitungsebene entstammt, wird durch verschiedene Studien gestützt (z. B. Chang, Bock & Goldberg, 2003; Hare & Goldberg, 1999; Pappert & Pechmann, 2014; Salamoura & Williams, 2007). Einen Nachweis liefern Hare und Goldberg (1999), die zeigen konnten, dass die Wortstellung in einem direkten Zusammenhang mit der thematischen Struktur steht. In ihrer Studie hatten die Versuchspersonen die Aufgabe, visuell dargebotene Bilder mit entweder der *prepositional-object*-Konstruktion (PO) oder der *double-object*-Konstruktion (DO) zu beschreiben, während ihnen zuvor ein PO-Satz, ein DO-Satz oder ein besonderer *provide-with*-Satz als Prime-Sätze präsentiert wurden. Beispiele für die Prime-Sätze sind wie folgt:

- (2a) His editor offered the hot story to Bob. – *PO*-Satz
- (2b) His editor promised Bob the hot story. – *DO*-Satz
- (2c) The officers provided the soldiers with guns. – *provide-with*-Satz

Im ersten Satz mit der *prepositional-object*-Konstruktion (PO) ist die Reihenfolge

der Verbargumente folgendermaßen: Subjekt, direktes Objekt, Präpositionalobjekt; auf der konzeptuellen Ebene lautet die Reihenfolge der thematischen Rollen so: Agens, Thema, Rezipient. Der zweite Satz mit der *double-object*-Konstruktion (DO) weist auf der syntaktischen Ebene folgende Reihenfolge auf: Subjekt, indirektes Objekt, direktes Objekt; Die semantische Reihenfolge lautet hier Agens, Rezipient, Thema. Von zentralem Interesse hierbei ist die *provide-with*-Konstruktion wie in (2c), die eine einzigartige Priming-Bedingung für die Studie darbot: Einerseits teilt diese Konstruktion auf der syntaktischen Ebene die *prepositional-object*-Konstruktion (PO) mit dem ersten Satz (2a), andererseits steht ihre thematische Abfolge in Übereinstimmung mit dem zweiten Satz (2b). Entscheidend ist, welche Struktur von den Versuchspersonen nach der Wiederholung eines *provide-with*-Satzes zur Beschreibung der visuellen Stimuli repliziert wird. Die Ergebnisse zeigten, dass sich die Bildbeschreibungen an der thematischen Reihenfolge des *provide-with*-Satzes orientierten, wobei deutlich mehr DO-Sätze produziert wurden als PO-Sätze. Hare und Goldberg schließen aus den Ergebnissen, dass im Gegensatz zum Priming auf der syntaktischen Ebene das Priming auf der semantischen Ebene relevanter für die Konstituenten-Zusammenstellung ist, was wiederum grundsätzlich mit der Hypothese, dass der Linearisierungsprozess die semantische Struktur befolgt, übereinstimmt.

Nachdem Priming-Effekte auf der semantischen Strukturierungsebene festgestellt wurden, lassen sich auch experimentelle Evidenzen dafür finden, dass die Konstituenten-Zusammenstellung unmittelbar von der Perspektivierung, die sich auf der konzeptuellen Ebene vollzieht, beeinflusst wird. In dem Priming-Experiment von Fleischer, Pickering und McLean (2012) mussten polnisch-englische Bilinguale und Konfidente der Forscher abwechselnd Bilder beschreiben. Nach jeder Beschreibung mussten die Gesprächsteilnehmer das beschriebene Bild aus einer Bildgruppe identifizieren, was den Autoren zufolge den Eindruck der durch den Prime-Satz vermittelten Strukturen im Arbeitsgedächtnis verstärken kann. Die Konfidenten bildeten in der kritischen Bedingung polnische Aktiv- und Passivsätze, Aktivsätze

mit OVS-Struktur und in der Kontrollbedingung koordinierte NPs, während die bilingualen Versuchspersonen die Bilder immer auf Englisch beschreiben mussten. Die Ergebnisse zeigten, dass deutlich mehr englische Passivsätze gebildet wurden, nachdem die Versuchspersonen einen polnischen OVS-Satz (z. B. *Baletnicę przygniata sportowiec*; deutsch: „Die Ballettänzerin quetscht der Sportler“) im Vergleich zur Kontrollbedingung (z. B. *baletnica i sportowiec*; deutsch: „Die Ballettänzerin und der Sportler“) gehört hatten. Von Bedeutung ist, dass zwischen den Proportionen der gebildeten Passivsätze nach dem Priming der polnischen Passiv- und OVS-Sätze keine auffälligen Unterschiede auftraten. Da die syntaktischen Strukturen des OVS-Prime-Satzes und des gebildeten englischen Passivsatzes verschieden sind, nehmen die Autoren an, dass die Präferenz für die Bildung eines Passivsatzes nach einem OVS-Prime-Satz nicht auf die syntaktischen Relationen, sondern auf die geteilte Formulierungsperspektive der Prime- und Target-Sätze zurückzuführen ist: In den beiden Satztypen steht das Patiens wegen der Besetzung der Satzanfangsposition im Fokus. Diese Erklärung bietet somit eine neue Ursprungsquelle für die Wortstellung an, nämlich die Informationsstruktur. Allerdings ist diese Interpretation nicht unproblematisch. Ein Kritikpunkt liegt darin, dass sich die semantischen Strukturen des OVS-Prime-Satzes und des produzierten Passivsatzes nicht voneinander unterscheiden ließen. D.h., wenn man die Interpretation lediglich auf die semantische Ebene begrenzt und behauptet, dass die Präferenz für die Passivkonstruktion unter der OVS-Prime-Bedingung aus deren geteilten thematischen Reihenfolgen resultiert, können die Ergebnisse ebenfalls gut erklärt werden. Diese Frage wird in Abschnitt 5.2 näher betrachtet.

Konträr zu der Hypothese, dass die Generierung der Wortstellung auf die konzeptuelle Ebene zurückgeführt wird, existieren auch zahlreiche Arbeiten zur Untermauerung der Effekte der syntaktischen Struktur auf die Konstituenten-Zusammenstellung (z. B. Bock & Loebell, 1990; Ferreira, 2003; Messenger, Branigan & McLean, 2011). Beispielsweise haben Bock und Loebell (1990) in ihrer strukturellen Priming-Studie herausgefunden, dass Veränderungen der

semantisch-konzeptuellen Struktur keine Veränderung der Wortstellung verursachen und die Wortstellung des Target-Satzes sich an der syntaktischen Struktur des Prime-Satzes orientiert. In ihrem Experiment wurden Versuchspersonen dazu aufgefordert, bildhaft dargebotene Ereignisse so schnell wie möglich zu beschreiben, nachdem sie zuvor auditiv präsentierte Prime-Sätze wiederholt hatten. Die Prime-Sätze beinhalteten entweder eine präpositionale Lokativphrase (z. B. *The wealthy widow drove her Mercedes to the church*) oder eine präpositionale Benefaktivphrase (z. B. *The wealthy widow gave her Mercedes to the church*). Die Funktionen der präpositionalen Phrase *to the church* wurde durch die Nutzung verschiedener Verben (*give, drive*) verändert. Wenn die Abfolge der Konstituenten unmittelbar durch die Abbildung der thematischen Abfolge erzeugt wird, sollten keine Priming-Effekte zwischen den beiden Prime-Sätzen auftreten, da sich deren semantische Strukturen unterscheiden (*church* in der Lokativphrase als ZIEL und in der Benefaktivphrase als REZIPIENT). Diese Hypothese wurde jedoch nicht bestätigt. Die Autoren haben festgestellt, dass sich die Wortstellung der produzierten Target-Sätze nach der Reihenfolge der syntaktischen Funktionen in den Prime-Sätzen richtet. Die Versuchspersonen berücksichtigten die Unterschiede in der semantischen Struktur nicht und produzierten dennoch unter beiden Bedingungen gleich viele Sätze mit präpositionalen Benefaktivphrasen, was darauf hinweist, dass die Priming-Effekte auf der syntaktischen Ebene stattfinden. Ein weiteres Experiment konnte zeigen, dass sich Passivsätze (z. B. *The construction worker was hit by the bulldozer*) durch Sätze mit einer lokativen *by*-Phrase (z. B. *The construction worker was digging by the bulldozer*) primen lassen. Aus den Ergebnissen ihrer Studie schlussfolgerten die Autoren, dass lediglich die syntaktische Struktur für die Erarbeitung der Wortstellung relevant ist, was wiederum in Übereinstimmung mit der Hypothese steht, die den modifizierenden Effekt der funktionalen Verarbeitung hervorhebt.⁴ Diese Schlussfolgerung ist

⁴ Die Schlussfolgerung, dass das Auftreten der strukturellen Priming-Effekte lediglich geteilte syntaktische Funktionen zwischen dem Prime- und Target-Satz verlangt und die konzeptuelle Ebene keine Rolle bei der Linearisierung spielt, gilt wahrscheinlich nicht für die L2-Produktion. In der Studie von Gerwien und Flecken (2015) wurde herausgefunden, dass bei der L2-Produktion trotz

möglicherweise zu weitreichend. Denn das Fehlen von Priming-Effekten auf der semantischen Ebene ist wohl nur der Tatsache geschuldet, dass die semantischen Rollen zwischen dem Prime- und dem Target-Satz nicht übereinstimmen. Die Studie überprüfte zudem nicht, ob semantische Priming-Effekte auftreten, wenn der Prime- und der Target-Satz die gleichen semantischen Strukturen aufweisen. Aus diesem Grund lässt sich anhand der Daten dieser Studie nicht nachvollziehen, ob syntaktische Funktionen der einzige Ausgangspunkt der Wortstellung sind. Trotzdem ist nicht von der Hand zu weisen, dass die Studie eine aussagekräftige Evidenz für den Einfluss der syntaktischen Funktionen auf die Wortstellung liefert.

Ausgehend von den bislang angeführten Studien werden zwei unterschiedliche und auch widersprüchliche Annahmen in Bezug auf den Ursprung der Konstituenten-Struktur bzw. der Wortstellung expliziert. Einerseits konnte empirisch nachgewiesen werden, dass die Generierung der Wortstellung von der semantischen und kontextuellen Struktur beeinflusst wird, was den Nachweis dafür liefert, dass die konzeptuelle Struktur die Konstituenten-Zusammenstellung beeinflusst; andererseits belegen einige Befunde Priming-Effekte auf der syntaktischen Ebene. Der erste Befund fordert das Satzproduktionsmodell von Bock und Levelt (1994) und Garrett (1980) heraus und zeigt, dass die funktionale und die positionelle Enkodierungsstufe nicht zwei nacheinander gereihete Stufen sind, wobei die Erstellung der Wortstellung immer die syntaktische Funktionszuweisung voraussetzt. Stattdessen sind sie möglicherweise zwei eigenständige Systeme, die beide direkt unter der Kontrolle der konzeptuellen Ebene stehen. Um auf diese Frage näher einzugehen, werden in Abschnitt 2.4 die konzeptuelle und grammatische Verarbeitungsebene gegenübergestellt und mit experimentellen Befunden erklärt.

2.2.4 Flexion

Die Wortstellung ist nicht das einzige Instrument zur Darstellung syntaktischer

einer geteilten syntaktischen Struktur zwischen dem Prime- und Target-Satz keine Priming-Effekte beobachtet wurden, wenn sich die konzeptuellen Strukturen der beiden Sätze voneinander unterscheiden.

Relationen. Um die grammatischen Eigenschaften der jeweiligen Konstituente weiter zu spezifizieren, werden dazu in manchen Sprachen Funktionswörter (z. B. Artikel, Hilfsverben, Konjunktionen usw.) und Flexionsendungen benutzt. Da Funktionswörter und Flexionsendungen vor allem grammatische Funktionen erfüllen und keine lexikalische Bedeutung haben, wird angenommen, dass deren Abruf über einen anderen Kanal als den der inhaltlichen Wörter erfolgt (Bock & Levelt, 1994).

Als Beispiel soll die Erzeugung der s-Pluralform (z. B. *Autos*) dienen. Um diese Pluralform formulieren zu können, müssen Sprecher zuerst eine nominale Konstruktion aufbauen, bei der ein Affix /s/ zum Wortstamm (*Auto*) hinzugefügt wird. Wie erfolgt dieser Prozess? Es wird angenommen, dass bereits auf der konzeptuellen Ebene die Information der Konstituenten-Anzahl in einer Form von *notional numbers* verarbeitet wird (vgl. Bock & Eberhard, 1993; Bock & Ferreira, 2014; Humphreys & Bock, 2005). Diese stehen in einer engen Verbindung mit der semantischen Information des zu formulierenden Konzeptes. Die *notional numbers* werden zusammen mit den semantischen Informationen in die lexikalische Ebene überführt, auf der Sprecher das Pluralaffix auswählen, das der semantischen Bedeutung des aufgerufenen Wortes entspricht. So wird etwa das Pluralaffix /s/ aufgerufen, wenn das Wort *Auto* ausgewählt wird, wohingegen bei der Wortselektion von *Wagen* keine speziellen Affixe benötigt werden. Analog zur Pluralform steht die Enkodierung der Zeitform eines Verbs ebenfalls unter einer direkten Kontrolle der konzeptuellen Ebene. So wird die temporale Information zuerst auf der konzeptuellen Ebene verarbeitet und dann nach der Verbselektion ins Verb eingebaut.

Im Unterschied zur Verb-Zeitform ist die Bildung der Subjekt-Prädikat-Kongruenz von der Funktionszuweisung abhängig. Da in manchen Sprachen wie zum Beispiel im Deutschen und Englischen die Prädikate in Person und Numerus mit dem Satzsubjekt kongruieren müssen, werden die Verbformen jedoch erst nach der Subjektzuordnung bestimmt. Wie etwa bei dem Beispielsatz *Das Handy fällt auf den Boden*. Hier entspricht das Verb *fallen* (in 3. Ps. Sg.) dem Subjekt (*das Handy*). Eine

komplexere Situation findet man bei der Bildung einer Wortstruktur, die zeitgleich unterschiedliche Informationen aufweist. Ein Beispiel hierfür ist die Dativobjekt-Phrase *den Schweinen*, wobei *Schwein* für das Wort, /e/ für den Plural, /n/ für den Dativ und der Artikel *den* sowohl für den Plural, aber auch für den Dativ stehen. Obwohl die Flexionsendungen /e/ und /n/ aneinanderreihend in dasselbe Lemma eingebaut sind, werden sie trotz allem auf unterschiedlicher Ebene und zu unterschiedlichem Zeitpunkt aufgerufen: Während die Pluralinformation auf der konzeptuellen Ebene geplant und bei der Selektion des Lemmas ausgewählt wird, ist das Dativaffix erst nach der Funktionszuweisung zugänglich. Das grammatische Geschlecht von *Schwein*, das zur lexikalischen Eigenschaften dieses Nomens gehört, wird zusammen mit der Lemma-Information aktiviert, während die Kasus-Beugung der Funktionszuordnung folgt.

2.3 Phonologisch-phonetische Enkodierung und Artikulation

Am Ende der grammatischen Enkodierung liegt schließlich eine syntaktische Oberflächenstruktur vor, die in der nachfolgenden Verarbeitungsebene mit phonologischen und phonetischen Informationen „aufgefüllt“ wird. Die phonologisch-phonetische Enkodierung bezeichnet den Prozess, der die syntaktische Oberflächenstruktur sozusagen „aussprechbar“, also zum Aussprechen geeignet, macht. Bei der phonologischen Enkodierung geht es darum, eine abstrakte phonologische Repräsentation, die morphologische, segmentale und metrische Informationen enthält, zu ermitteln. Sie dient dann als Input für die phonetische Enkodierung, bei der der phonetische Plan erzeugt wird, durch welchen die motorischen Programme der Artikulation spezifiziert werden.

In Bezug auf die phonologische Enkodierung müssen Sprecher im ersten Schritt die morphologische Struktur, die aus Morphemen besteht, abrufen. Morpheme sind die kleinsten lexikalischen und bedeutungstragenden Einheiten; zum Beispiel besteht das Lemma *Handynummer* aus zwei Morphemen - nämlich *Handy* und *Nummer*.

Nach dem Aufbau der morphologischen Struktur werden in einem zweiten Schritt dann die phonemischen Wortsegmente abgerufen. Die Eigenschaft, dass Wortformen nicht als Ganze zur Aussprache gebracht, sondern von individuellen phonemischen Segmenten zusammengesetzt werden, zeigt sich meist in Segmentfehlern wie etwa in *guten Gorgen*. Wenn Wortformen als Ganzes abgerufen werden, sollen solche Fehler aber gar nicht auftreten. Die Tatsache, dass schätzungsweise 60 bis 90% aller sublexikalischen Fehler segmentelle Informationen betreffen, gilt als Evidenz für den Status von Segmenten als Verarbeitungseinheiten (vgl. Meyer, 2003). In einem weiteren Schritt werden dann die metrischen Formen erarbeitet. Hierbei handelt es sich vor allem um die Festlegung der Rahmenstruktur einzelner Wörter, also die Silbenanzahl und die Betonung des Wortes. Zudem wird die Reihenfolge der phonemischen Segmente festgesetzt. Beispielsweise hat das Substantiv *Mode* als metrische Struktur einen Trochäus, d.h., eine akzentuierte lange und eine kurze Silbe. Dazu besteht die segmentale Komposition vom Wort *Mode* aus den Phonemen /m/, /o:/, /d/, /ə/. Steht die segmentale und metrische Form eines Wortes schließlich bereit, wird es sodann im sprachlichen Kontext eingebettet. Während dieses Ablaufs kann die silbische Struktur des lexikalischen Wortes verändert werden (*un-der-stand-it* als *un-der-stan-dit*; Beispiel aus Levelt, Roelofs & Meyer 1999, S. 20). Als Ergebnis dieser Stufe wird ein phonologisches Wort generiert, das eine abstrakte Repräsentation einer metrischen Segmentalstruktur darstellt.

Im anschließenden phonetischen Enkodierungsprozess wird der phonetische Plan, den Levelt (1992) auch als *internal speech* bezeichnet, für einzelne Wörter erstellt. Diese bestehen aus einem phonetischen Plan für Silben, die Anweisungen für Artikulationsmuster liefern. Sie werden aus einem mentalen Silbarium (engl.: *syllabary*) abgerufen, in dem für die hochfrequenten Silben motorische Programme gespeichert sind.

Am Ende des Sprachproduktionsprozesses wird dann der phonetische Plan in die Artikulation eingespeist, wo die Sequenz der artikulatorischen Bewegungen in ein motorisches Programm umgesetzt wird. Dabei werden die relevanten Regionen des

neuronalen Systems aktiviert, um die Ausführung des motorischen Programms zu wahrnehmbaren Schallakten realisieren und kontrollieren zu können. Dieser Prozess ist hoch automatisiert und benötigt keine zusätzliche Aufmerksamkeit eines geübten Sprechers.

2.4 Mappings

Dieser Abschnitt greift nun thematisch noch einmal die schon zuvor behandelte, offene Frage aus dem Abschnitt 2.2.3 zum Ursprung der Wortstellung auf. Die Rede ist von der Verbindung der konzeptuellen, funktionalen und positionellen Verarbeitungsebene. Das Modell von Bock und Levelt (1994) sieht vor, dass die syntaktische Enkodierung der präverbalen Message auf zwei nacheinander folgenden Ebenen vorgenommen wird: der funktionalen und der positionellen Ebene. Durch die Funktionszuweisung werden syntaktische Funktionen ausgewählt und zugeteilt, welche dann in der folgenden Zusammenstellungsphase in eine Reihenfolge gebracht werden. Da dieses Modell die Wichtigkeit der Verteilung syntaktischer Rollen für die weitere Linearisierung hervorhebt und annimmt, dass die Konstituenten-Anordnung die Festlegung syntaktischer Relationen voraussetzt, wird es dementsprechend als zweistufiges Modell oder auch *grammatical function account* bezeichnet (vgl. Bock & Warren, 1985; Bock, 1987b; Bock, Loebell, & Morey, 1992). Im Gegensatz dazu fordern andere Forscherinnen und Forscher, dass Konzepten nicht unbedingt syntaktische Funktionen zugewiesen werden müssen, bevor sie an die positionelle Verarbeitungsebene weitergereicht werden (vgl. z. B. De Smedt, 1990; Kempen & Hoenkamp, 1987). Sobald ein Konzept aufgerufen wurde, kann es direkt in der Wortstellung angeordnet werden. In diesem Sinn ist die Wortstellung ein sog. Doppelgänger der Reihung der Konzept-Aufrufung, steht aber in keinem direkten Zusammenhang mit der Zuweisung syntaktischer Funktionen. Dieses Modell wird als einstufiges Modell oder *word order account* bezeichnet, bei dem die Funktionszuweisung nicht notwendigerweise vor der Konstituenten-Zusammenstellung vorgenommen werden muss. Da konzeptuell

zugänglichere Elemente schneller aus dem Gedächtnis abgerufen und dementsprechend früher in die grammatische Enkodierung übertragen werden können, bietet die Untersuchung über das Verhältnis zwischen der konzeptuellen Zugänglichkeit und der grammatischen Enkodierung eine Chance zur Unterscheidung zwischen dem *word order account* und dem *grammatical function account*.

Es existieren zahlreiche Studien, die das zweistufige Modell untermauern. Ein Beispiel dafür ist die Studie von Bock und Warren (1985). In der Studie wurden Versuchspersonen zwei verschiedene Satztypen, ein transitiver Satz und ein koordinierter Objektsatz, gezeigt und sie wurden dazu aufgefordert, diese nach einer weiteren Aufgabe, die zur Ablenkung diente, wieder aufzurufen und zu wiederholen. Beispiele für die Satztypen sind folgende:

(3a) *The shock was administered by the doctor*

(3b) *The hiker fought time and winter.*

Die Konkretheit der Wörter wurde in beiden Sätzen manipuliert. In (3a) ist *doctor* konkreter als *shock* und in (3b) ist *winter* konkreter als *time*. Es wurde gezeigt, dass Satz (3a) eher als *The doctor administered the shock* erinnert wurde. Dies ist laut den Autoren auf die höhere Konkretheit von *doctor* zurückzuführen. Allerdings tauchen keine Konkretheitseffekte bei der Wiedergabe von Sätzen mit koordinierten Objekten wie (3b) auf: Als die Probanden sich an den zweiten Satz erinnerten, wurde *winter* trotz dessen höherer Konkretheit nicht häufiger vor *time* platziert. Die Autoren gehen davon aus, dass die konzeptuelle Zugänglichkeit die Zusammenstellung von Konstituenten nicht beeinflusst. Dies rührt daher, dass sich die syntaktischen Funktionen von *time* und *winter* nicht voneinander unterscheiden und es bei der Wiedergabe dieses Satzes nur um die Positionierung der beiden Konstituenten geht. Aufgrund des Auftretens der Konkretheitseffekte auf der syntaktischen Funktionszuweisungsebene schlussfolgern die Autoren, dass die konzeptuelle Zugänglichkeit nur für die funktionale Verarbeitung relevant ist, was

für das zweistufige Modell spricht.

Diese Schlussfolgerung wurde von McDonald, Bock und Kelly (1993) getestet, indem die Effekte der Belebtheit auf die syntaktische Strukturierung überprüft wurden. Ähnlich wie bei Bock und Warren (1985) mussten Versuchspersonen transitive Sätze und Sätze mit koordinierten Objekten hören und anschließend wiedergeben. Alle zu wiederholenden Sätze beinhalteten eine belebte und eine unbelebte Entität. Die Ergebnisse zeigten, dass bei der Wiedergabe eines transitiven Satzes belebte Entitäten deutlich häufiger als Subjekt realisiert wurden (z. B. wurde der Satz *The music soothes the child* häufiger als *The child was soothed by the music* erinnert), während sich keine Effekte der Belebtheit auf die Reihenfolgebildung der syntaktisch gleichbedeutenden NPs feststellen ließen (beispielsweise wurde der Satz *The telephone and the dog were both making noise* selten als *The dog and the telephone were both making noise* wiedergegeben). Die Ergebnisse werden als Hinweis darauf interpretiert, dass die Auswirkung der Belebtheit die grammatische Funktionszuweisung von Entitäten voraussetzt.

Zusammenfassend unterstützen die beiden oben erwähnten Studien eine strikt serielle Verarbeitung von Konzepten innerhalb der grammatischen Enkodierung, welche zwei gereifte Verarbeitungsstufen, die funktionale und die positionelle Verarbeitung, umfasst. Eine direkte Abbildung der konzeptuell zugänglichen Konzepte auf die positionelle Verarbeitungsebene wurde nicht festgestellt.

Obwohl die oben genannten Studien aussagekräftige Evidenzen zur Untermauerung des zweistufigen Modells geliefert haben, ist zu beachten, dass solche Evidenzen vorwiegend aus Studien über die englische Satzproduktion stammen, bei der das Subjekt in den meisten Fällen mit der ersten Nominalphrase des Satzes identisch ist. Ob das zweistufige Modell daher als universal für andere Sprachen gelten kann, ist fraglich. Tatsächlich haben andere Sprachen, die eine flexible Wortstellung besitzen, eine entgegengesetzte Richtung gezeigt: Anstatt durch die Ebene der funktionalen Verarbeitung kann die präverbale Message ohne Rücksicht auf grammatische

Funktionen unmittelbar auf die Wortstellung abgebildet werden (Deutsch: Kempen & Harbusch, 2004; Griechisch: Branigan & Feleki, 1999; Odawa: Christianson & Ferreira, 2005).

Beispielsweise konnten Branigan und Feleki (1999) zeigen, dass die konzeptuelle Zugänglichkeit die griechische Wortstellung unabhängig von der Funktionszuweisung beeinflusst. Im Vergleich zum Englischen ist die griechische Wortstellung sehr flexibel, denn das Objekt kann durch das Scrambling sowohl in einem Aktiv- als auch in einem Passivsatz am Satzanfang erscheinen, was wiederum zu einer OVS-Konstruktion führt. In der Studie haben die Autoren herausgefunden, dass bei der Satz wiederholung griechische Sprecher dazu tendierten, belebte Entitäten vor unbelebten Entitäten zu platzieren, ohne die syntaktische Funktion der vorangestellten Entitäten zu ändern. Wenn das Subjekt (das Agens) beispielsweise in einem kanonischen Aktivsatz unbelebt und das Objekt (das Patiens) belebt war, wurden eindeutig mehr OVS-Aktivsätze als Passivsätze mit SVO-Konstruktion als umgekehrt gebildet. Ein Beispiel: Der Satz *Sta dimokratika politevmata, to sindagm_[NOM] sevete ton Politi_[AKK]* (deutsch: „Im demokratischen Regime, das Gesetz respektiert den Bürger“) wurde eher als *Sta dimokratika politevmata, ton Politi_[AKK] sevete to sindagm_[NOM]* (deutsch: „Im demokratischen Regime, den Bürger respektiert das Gesetz“) wiedergegeben. Die Ergebnisse untermauern das einstufige Modell, da sie zeigen, dass anstatt einer Veränderung syntaktischer Funktionen der belebten Figuren und somit der Bildung eines Passivsatzes die Effekte der Belebtheit durch das Scrambling zum Tragen kommen. Ein Einfluss der konzeptuellen Zugänglichkeit auf die syntaktische Funktionszuweisung wurde in der genannten Studie aber nicht beobachtet. So kommen die Autoren zu der Schlussfolgerung, dass die konzeptuelle Information im Griechischen unmittelbar auf die positionelle Verarbeitungsebene abgebildet wird. Das gleiche Phänomen wurde in anderen Sprachen mit freier Wortstellung, wie Japanisch (Tanaka, 2006), Katalanisch (Prat-Sala, Schillcock & Sorace, 2000) und Russisch (Myachykov & Tomlin, 2008) festgestellt.

Eine dritte Möglichkeit wurde durch eine Studie über den Zusammenhang zwischen der konzeptuellen Zugänglichkeit und der syntaktischen Strukturierung bei der japanischen Satzproduktion geliefert. In der Studie von Tanaka, Branigan, McLean und Pickering (2011) wurde herausgefunden, dass die Belebtheit der Figuren sowohl die Positionierung als auch die Funktionszuweisung der Figuren beeinflusst. Im Vergleich mit dem Englischen ist Japanisch eine Kasussprache, in der syntaktische Funktionen wenig mit der Wortstellung zu tun haben, sondern vor allem durch Kasus-Marker dargestellt werden, was eine sehr flexible Wortstellung ermöglicht. So kann die SOV-Struktur, sowohl aktiv als passiv, ähnlich wie im Griechischen durch Scrambling, und zwar die Objekt-Voranstellung, zu einer OSV-Struktur verändert werden. Nachfolgend ein paar Sätze aus der Studie als Beispiele für die japanische Wortstellung:

(4a) *Fisher*_[NOM] *Boot*_[AKK] *befördern*_[ACT] – SOV-Aktiv

(4b) *Boot*_[AKK] *Fisher*_[NOM] *befördern*_[ACT] – OSV-Aktiv

Der Fischer befördert das Boot.

(4c) *Boot*_[NOM] *Fischer*_[AKK] *befördern*_[PAS] – SOV-Passiv

(4d) *Fischer*_[AKK] *Boot*_[NOM] *befördern*_[PAS] - OSV-Passiv

Das Boot wurde von dem Fischer befördert.

Die Grundbedeutungen der oben genannten vier Sätze sind identisch. Die Sätze (4a) und (4b) sind Aktivsätze, bei denen die belebte Entität *Fischer* als Subjekt und die unbelebte Entität *Boot* als Objekt dient. Die Sätze (4c) und (4d) sind Passivsätze, bei denen die unbelebte Entität *Boot* als Subjekt und die belebte Figur *Fischer* als Objekt auftritt. Die Sätze (4b) und (4d) sind Scrambled-Sätze, bei denen das Objekt am Satzanfang steht. In der Studie von Tanaka et al. (2011) wurde ein Haupteffekt der Belebtheit auf die Positionierung der Satzglieder gefunden. So wurde ein OSV-Aktivsatz wie (4b), in dem das belebte Subjekt die zweite Position des Satzes einnimmt, häufiger als ein SOV-Aktivsatz wie (4a) wiedergegeben, in dem das

belebte Subjekt am Satzanfang steht. Dabei wurden die syntaktischen Funktionen der Figuren aber nicht geändert. Die Voranstellung der belebten Entität, ohne deren syntaktische Funktion zu verändern, deutet darauf hin, dass die konzeptuelle Zugänglichkeit die Positionierung der Figuren beeinflusst, ohne die syntaktische Funktionszuweisung beeinflussen zu müssen. Interessant ist jedoch, dass japanische Sprecher dazu tendieren, belebte Figuren als Subjekt wiederzugeben, selbst dann, wenn es an zweiter Satzposition erscheint. Beispielsweise wurden SOV-Passivsätze wie in (4d), in dem die belebte Figur wie *Fischer* als Objekt in der Satzmitte steht, häufig als OSV-Aktivsätze wie in (4b) erinnert, in dem die belebte Figur als Subjekt realisiert ist. Insgesamt hat die Studie gezeigt, dass sowohl die Funktionszuweisung als auch die Wortstellung separat von der konzeptuellen Zugänglichkeit beeinflusst werden können, was auf ein paralleles „Mapping“ der konzeptuellen Information sowohl auf der Funktionszuweisungsebene als auch der Linearisierungsebene hindeutet.

Als Fazit dieser Überlegungen lässt sich festhalten, dass drei Möglichkeiten bezüglich des Mappings der konzeptuellen Information auf die Wortstellung aufgezeigt wurden. Das zweistufige Modell, welches auch als *grammatical function account* bezeichnet wird, geht von der Rolle der Funktionszuweisung bei der Linearisierung aus. Konzeptuelle Informationen müssen zuerst syntaktischen Funktionen zugeordnet werden, bevor sie weiter in die Wortstellung überführt werden können. Im Gegensatz dazu verlangt das einstufige Modell (*word order account*), dass konzeptuelle Informationen unmittelbar auf die Wortstellung abgebildet werden, ohne die Rolle der syntaktischen Funktionen zu berücksichtigen. Eine dritte Möglichkeit, die ich hier vorläufig als parallelstufiges Modell bezeichne, nimmt an, dass die konzeptuelle Zugänglichkeit sowohl die syntaktische Funktionszuweisung als auch die Linearisierung gleichermaßen beeinflussen kann. Insgesamt lässt sich auf der Grundlage der bis zu diesem Punkt beschriebenen Studien nicht eindeutig entscheiden, in welcher Art und Weise die Abbildung der konzeptuellen Information auf die Wortstellung stattfindet. Die Tatsache, dass alle

drei Möglichkeiten von empirischen Studien mit verschiedenen Sprachen bestätigt wurden, lässt den Rückschluss zu, dass die Satzproduktion je nach Zielsprache sprachspezifische Unterschiede aufweisen kann und hinsichtlich der Verallgemeinerung eines bestimmten Modells Vorsicht geboten ist.

2.5 Fazit

In diesem Kapitel wurden die mentalen Verarbeitungen der Satzplanung auf der Grundlage des Satzproduktionsmodells von Bock und Levelt (1994) vorgestellt. Bock und Levelt gliedern mentale Prozesse zur Satzproduktion in vier Teilsysteme: die Konzeptualisierung, die grammatische Enkodierung, die phonologisch-phonetische Enkodierung sowie die Artikulation. Jedes Teilsystem fungiert als autonomer Experte, das einen charakteristischen Input in einen charakteristischen Output umwandelt. Alle Ergebnisse, die in dieser Studie gewonnen wurden, werden im Rahmen dieses Modells diskutiert.

Basierend auf dem Modell von Bock und Levelt (1994) wird dazu der Abbildungsprozess der konzeptuellen Information von der Konzeptualisierungsebene auf die Syntax auseinandergesetzt. Diesbezüglich wurden grundsätzlich zwei Annahmen eingeführt: Der *grammatical function account* nimmt an, dass der Beginn des Linearisierungsprozesses die Festlegung syntaktischer Relationen voraussetzt. Im Gegensatz dazu besagt der *word order account*, dass sich die Erstellung der Wortstellung über die syntaktische Funktionszuweisungsebene hinwegsetzt. Die Auseinandersetzung in diesem Bezug ist wichtig, da sie einen theoretischen Rahmen für ein wichtiges Thema dieser Arbeit, und zwar die Abbildung kontextueller Informationen innerhalb der Satzplanung, liefert.

3 Inkrementalität

Wie bereits im zweiten Kapitel dargestellt wurde, gliedert sich der Satzproduktionsprozess in vier Verarbeitungsebenen. Die zwei Ebenen, die im Fokus der vorliegenden Arbeit stehen, sind die Konzeptualisierungs- und die grammatische Enkodierungsebene. Auf der Konzeptualisierungsebene werden nach Levelt (1989) individuelle lexikalische Konzepte identifiziert, die zusammen ein Ereignis ausmachen. Überdies werden auch strukturelle Repräsentationen generiert, anhand derer ersichtlich ist, wie die identifizierten lexikalischen Konzepte zueinander in Verbindung stehen. Auf der grammatischen Enkodierungsebene werden dann die auf der vorangehenden Ebene aufgerufenen Konzepte als Worteinträge realisiert. Dazu werden auch syntaktische Strukturen und Wortstellungen aufgebaut, denen strukturelle Informationen zugeschrieben werden. Eine Schlüsselfrage, die bislang noch nicht geklärt ist, betrifft das *Timing* der genannten Verarbeitungsprozesse. Dabei stellt sich vor allem die Frage, in welcher Art und Weise die Teilsysteme des Satzproduktionsmodells Informationen verarbeiten.

Zur Beantwortung dieser Frage wurde von Kempen und Hoenkamp (1982, 1987) der Begriff der inkrementellen Verarbeitungsweise vorgeschlagen, deren Grundannahme darin besteht, dass die einzelnen Verarbeitungsstufen der Satzproduktion in einer sowohl seriellen als auch parallelen Weise ablaufen (vgl. Abbildung 3.1; Bock & Levelt, 1994; De Smedt, 1990; Dietrich, 1999; Levelt, 1989, 1994; Roelof, 1998; Wheeldon et al., 2003).

Die serielle Verarbeitungsweise beschreibt den Prozess, dass ein Satz nicht ganzheitlich, sondern stückweise auf den genannten Ebenen generiert wird. Innerhalb des Satzproduktionssystems verfolgt die Verarbeitung eines bestimmten Segments eine bestimmte Reihenfolge. Ein Segment wird immer zunächst konzeptualisiert, bevor es dann grammatisch und lautlich enkodiert und schließlich artikuliert wird. Die serielle Weise gilt auch für die Verarbeitung der Äußerungsfragmente auf jeder einzelnen Verarbeitungsebene. Die Satzkonstituenten

durchlaufen das Satzproduktionssystem in der Abfolge, in der sie in einem Satz angeordnet sind. Die Verarbeitung einer Konstituente wird somit erst dann beginnen, nachdem zuvor die der vorherigen auf der gleichen Enkodierungsebene abgeschlossen wurde.

Daneben laufen die Enkodierungen auf allen Verarbeitungsebenen parallel ab. Sprecher müssen nicht abwarten bis alle Informationen auf einer vorgeordneten Ebene ausgegeben werden, um die Verarbeitung auf einer nachgeordneten Ebene anzufangen. Sobald ein Verarbeitungssegment bzw. ein Inkrement in einer Phase generiert wurde, wird sie unmittelbar an die nachgeordnete Ebene weitergereicht. Nachdem eine Einheit linguistisch enkodiert und dem Artikulator zur Verfügung gestellt wurde, beginnen Sprecher dann mit dem Sprechen.

Als Beispiel soll an dieser Stelle die Informationsverarbeitung auf der konzeptuellen und der grammatischen Enkodierungsebene herangezogen werden. Sobald die Konzeptualisierung einen Teil der präverbale Botschaft erzeugt hat, beginnt die grammatische Enkodierungsebene zum einen damit, die dazu passenden Lemma-Informationen auszuwählen, und zum anderen einen Teil der Oberflächenstruktur aufzubauen. Währenddessen werden weitere Einheiten auf der konzeptuellen Ebene erstellt, um die präverbale Botschaft zu vervollständigen. Die verschiedenen Ebenen während des Satzproduktionsprozesses arbeiten demnach zeitgleich an ein und derselben Äußerung, jedoch an unterschiedlichen Inkrementen.

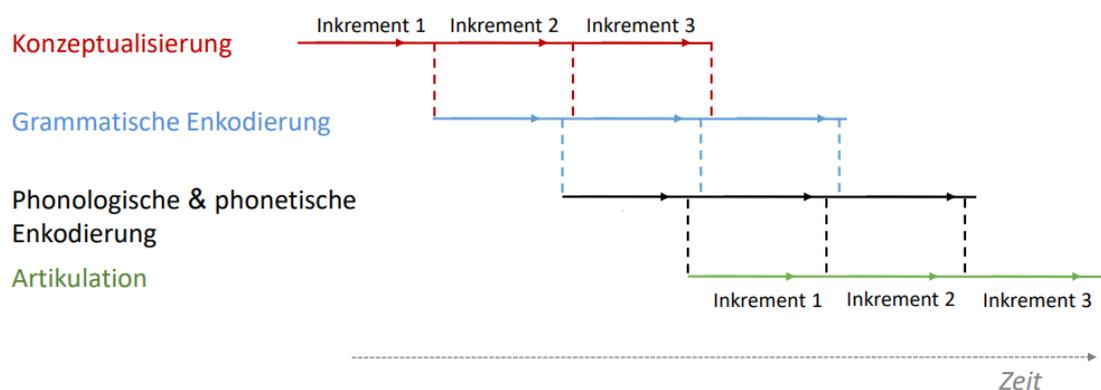


Abbildung 3.1: Inkrementelle Verarbeitung der Satzproduktion (adaptiert von Kempen und Hoenkamp 1987)

Aus der kognitiven Sicht wird die Inkrementalität als eine sich anpassende Charakteristik des Sprechens angesehen: Aufgrund der Kapazitätsbeschränkung des Arbeitsgedächtnisses erlaubt die inkrementelle Verarbeitungsweise eine Unterteilung komplexerer Formulierungen in kleinere und schrittweise erzeugte Elemente. Demzufolge brauchen Sprecher zu einem bestimmten Zeitpunkt nur ein kleines Stück der zu verbalisierenden Information zu enkodieren, was für sie kognitiv weniger aufwendig ist (vgl. Levelt, 1989; Ferreira & Swets, 2002; Wagner, Jescheniak & Schriefers, 2010). Somit wird auch sichergestellt, dass Sprecher zwischen den Sätzen keine großen Pausen machen müssen, was wiederum dem kommunikativen Redefluss dienlich ist.

Dieses Kapitel thematisiert die inkrementelle Verarbeitungsweise der Satzplanung. Im Folgenden wird nun zuerst das bekannte *starting points problem* erläutert, welches die Selektion des als Erstes zu verbalisierenden lexikalischen Konzeptes betrifft (vgl. Abschnitt 3.1). Ausgehend davon werden dann zwei stark polarisierende Annahmen zur Lösung dieses Problems präsentiert: die Annahme von Gleitman, January, Nappa und Trueswell (2007) hinsichtlich einer linear verlaufenden Inkrementalität und die Annahme von Griffin und Bock (2000), welche eine strukturell verlaufende Inkrementalität annehmen. Der Hauptunterschied zwischen diesen zwei Annahmen besteht darin, ob die Selektion von *starting points* auf Eigenschaften einzelner Ereignisteilnehmer oder auf die strukturelle Repräsentation des Gesamtereignisses zurückzuführen ist (Abschnitt 3.2). Des Weiteren wird noch eine integrierende Perspektive vorgestellt, die als einen Mittelweg der beiden Annahmen angesehen wird und behauptet, dass die Arbeitsweise des Satzplanungssystems von der Interaktion zwischen der Kodierbarkeit der Ereignisstruktur und der Ereignisteilnehmer beeinflusst wird (Abschnitt 3.3).

3.1 Das Starting Points Problem

Da Sätze im Allgemeinen nicht in einem Arbeitsschritt als Ganzes, sondern

stückweise geplant werden, müssen Sprecher vor Beginn der Satzplanung zuerst entscheiden, mit welchem Stück der Botschaft mit lexikalischem Konzept sie beginnen. Das lexikalische Konzept, das zuerst zu verarbeiten ist, bezeichnet man als sog. *starting point*, und das Entscheidungsproblem über die Selektion dieses Konzeptes wurde von MacWhinney (1977) als *starting points problem* beschrieben, das zu den wichtigsten Fragen in der Satzproduktion zählt.

Die Festlegung solcher *starting points* ist deshalb so wichtig, weil mit verschiedenen Arten von *starting points* auch verschiedene Oberflächenformen hervorgerufen werden können. Nachfolgend zwei englische Beispiele, die einen solchen möglichen Fall zeigen:

(5a) *The arrow hit the target.*

(5b) *The target was hit by the arrow.*

Die semantischen Funktionen der beiden Ereignisteilnehmer in (5a) und (5b) sind identisch (*the arrow* als Thema und *the target* als Ziel), zudem können sich diese zwei Sätze auch auf das gleiche Ereignis beziehen – was sie jedoch unterscheidet, ist, dass sie mit verschiedenen *starting points* beginnen. In Satz (5a) ist der *starting point* das Thema, mit dessen Selektion die Aktivstruktur hervorgerufen wurde. In Satz (5b) hingegen ist der *starting point* das Ziel, was wiederum zu einem Passivsatz führt.

Angesichts der großen Bedeutung von *starting points* für die syntaktische Strukturierung ergibt sich nun die Frage, wie diese festgelegt werden. Bereits Anfang des 20. Jahrhunderts haben einige Wissenschaftler der Psychologie und Sprachwissenschaft versucht, diese Frage zu beantworten. Beispielsweise war die Annahme von Wundt (1900), dass die Auswahl von *starting points* auf einer Gesamtvorstellung eines Ereignisses beruht, dem ein Bündel von Gegenständen, Eigenschaften und Beziehungen, die das Ereignis ausmachen, zugeordnet wird. Im Gegensatz dazu legte Jespersen (1924) sein Augenmerk auf die Zugänglichkeit

einzelner Entitäten, anstatt auf die Gesamtvorstellung eines Ereignisses. Beispielsweise wird von ihm angenommen, dass der Aufbau eines Passivsatzes auf die höhere inhärente Zugänglichkeit des Patiens im Verhältnis zum Agens zurückzuführen ist. Solche Annahmen bzw. Hypothesen haben sich in den letzten Jahrzehnten weiterentwickelt und münden nun hauptsächlich in zwei Verlaufsinterpretationen: die lineare und die strukturelle Inkrementalitätshypothese, deren Unterschiede in erster Linie darin bestehen, welche Information eine entscheidende Rolle bei der Selektion von *starting points* spielt. Im folgenden Abschnitt sollen die beiden Hypothesen mithilfe experimenteller Befunde diskutiert werden.

3.2 Lineare Inkrementalität und strukturelle Inkrementalität

3.2.1 Implikationen der linearen und der strukturellen Inkrementalität

Die Hypothese der linearen Inkrementalität besagt, dass inhärente Eigenschaften einzelner Ereignispartizipanten ausschlaggebend für die Bestimmung von *starting points* sind (vgl. Bock et al., 1992; Bock & Warren, 1985; Gleitman et al., 2007; Ibbotson, Lieven & Tomasello, 2013). Unter der inhärenten Eigenschaft wird hier vor allem die konzeptuelle und die visuelle Zugänglichkeit verstanden. Entitäten, deren diesbezüglichen Zugänglichkeiten höher sind, werden mit größerer Wahrscheinlichkeit als *starting points* ausgewählt und enkodiert im Gegensatz zu denen, die niedrigere Zugänglichkeiten aufweisen - so die Annahme. Sobald eine Entität als *starting point* festgelegt wurde, wird sie umgehend an die grammatische Enkodierung weitergereicht. Die Organisation sämtlicher weiterer Elemente der Botschaft wird dann auf diese *starting-point*-Entität bezogen, koordiniert. Da die konzeptuelle und die visuelle Eigenschaft inhärent in Stimuli verankert sind, wird die darauf basierende Auswahl von *starting points* als Bottom-Up-Prozess aufgefasst.

Im Gegensatz dazu sieht die strukturelle Inkrementalität die Wichtigkeit in der

strukturellen Repräsentation, die von einer holistischen Interpretation des Gesamtereignisses ausgeht (vgl. Griffin & Bock, 2000; Lee, Brown-Schmidt & Watson, 2013). Durch den Aufbau einer Gesamtvorstellung bzw. einer abstrakten strukturellen Repräsentation werden den zu verbalisierenden Entitäten, je nachdem wie deren Beziehungen zueinander sind, verschiedene Rollen zugewiesen (z. B. semantische Rollen wie Agens oder Patiens). Es wird angenommen, dass die Selektion von *starting points* nur anhand dieser strukturellen Informationen vorgenommen werden kann. Diese Annahme erläutern Bock, Irwin und Davidson (2004) wie folgt:

[...] people begin with what it is they have something to say about, and they tend to avoid initiating utterances until they are fairly sure what the beginning of that something is. This implies that the rudiments of an utterance frame are present when an utterance is initiated. In short, speakers don't start what they can't finish. (S. 270)

Demnach sind *starting points* kein spontan ausgewähltes Produkt aufgrund der konzeptuellen oder visuellen Eigenschaften einzelner Ereignisteilnehmer, sondern innerhalb der Gesamtvorstellung eines Ereignisses gefasst und deren Rollen nach in einem Ereignis sortiert und ausgewählt. Wenn Sprecher keinen ganzheitlichen Überblick über das Ereignis haben, sind sie nicht in der Lage, wichtige Informationen zu versprachlichen und andere irrelevante Informationen trotz ihrer hohen konzeptuellen oder visuellen Auffälligkeit (z. B. Belebtheit oder Luminanz) außen vor zu lassen – so die Annahme. Da der Aufbau der strukturellen Repräsentation einen synthetischen Vorgang erfordert, durch den unterschiedliche Entitäten durch die Vorstellung verknüpft werden, wird die strukturell inkrementelle Verarbeitung demgemäß als Top-Down-Prozess aufgefasst.

Ausgehend von den Grundideen der beiden Annahmen werden verschiedene Hypothesen über den Umfang der Vorausplanung auf der Konzeptualisierungsebene aufgestellt. Unter der Vorausplanung versteht man die konzeptuelle Verarbeitung, die sich vor Beginn der grammatischen Enkodierung vollzieht. Unter der Annahme der linearen Inkrementalität, dass sich also die Auswahl von *starting points* nach der

Eigenschaft einzelner Konzepte richtet, ist die Größe der Vorausplanung infolgedessen so klein wie ein individuelles Konzept. Das bedeutet, dass die Selektion lediglich die Aktivierung eines Konzeptes voraussetzt. Sobald eine Figur beispielsweise aufgrund ihrer höheren visuellen Zugänglichkeit als Erstes wahrgenommen und identifiziert wurde, wird sie direkt in die grammatische Enkodierungsebene überführt und dient als Ausgangspunkt weiterer Verarbeitungen. Die strukturelle Inkrementalitätshypothese hingegen nimmt an, dass die Selektion von *starting points* ein holistisches Verständnis des Gesamtereignisses verlangt. Zudem wird davon ausgegangen, dass die Verarbeitungsabfolge einzelner Konzepte nicht von deren visuellen oder konzeptuellen Zugänglichkeiten, sondern von ihren Rollen in dem zu beschreibenden Ereignis bestimmt wird. In diesem Sinn kann der Umfang der Vorausplanung so groß wie die strukturelle Repräsentation eines Gesamtereignisses sein, die grundsätzlich viel größer ist als ein lexikalisches Konzept.

Unterschiedliche Verarbeitungsweisen bringen auch unterschiedliche Hypothesen über deren Zeitverlauf mit sich, und der zentrale Punkt diesbezüglich betrifft die Frage, wann der Aufbau der strukturellen Repräsentation relativ zum Beginn der grammatischen Enkodierung abgeschlossen ist. Beide Inkrementalitätstypen messen der strukturellen Information in Bezug auf das Auswählen von *starting points* ein unterschiedliches Maß von Wichtigkeit bei. Dadurch kann der Zeitverlauf während des Aufbaus der strukturellen Repräsentation diesbezüglich auch als unterschiedlich angenommen werden. In Bezug auf die strukturelle Inkrementalität wird zusätzlich die Annahme vertreten, dass die strukturelle Repräsentation eines Gesamtereignisses (z. B. Referenten sind als Agens und Patiens spezifiziert) vor Beginn der grammatischen Enkodierung aufgebaut wird. Diese Annahme baut auf der Auffassung auf, dass allen Verarbeitungen einschließlich der Auswahl von *starting points* diese Repräsentation zugrunde liegt. Da das Erfassen der Ereignisstruktur eine Notwendigkeit für die syntaktische Enkodierung ist, muss eine derartige Struktur vor der grammatischen Enkodierung aufgebaut werden. Die lineare Inkrementalität

hingegen nimmt an, dass die strukturelle Repräsentation nicht unbedingt in der Vorausplanungsphase aufgebaut werden muss. Da die grammatische Enkodierung mit der Aktivierung eines einzigen Konzepts bereits begonnen werden kann, sobald dieses Konzept als zugänglicher als andere wahrgenommen wurde, ist es möglich, dass die Enkodierung struktureller Information erst nach Beginn der grammatischen Enkodierung abgeschlossen wird.

Zusammenfassend lässt sich sagen, dass der Unterschied zwischen der linearen und der strukturellen Inkrementalitätshypothese in der Frage besteht, auf welche Information die Selektion von *starting points* zurückzuführen ist. Diese Frage kann durch zwei konkrete Fragen spezifiziert werden:

- Nach dem inhaltlichen Aspekt: In welchem Maß wird das Sprechen in der Vorausplanungsphase, also vor Beginn der grammatischen Enkodierung, geplant?
- Nach dem zeitlichen Aspekt: Wann wird relativ zum Beginn der grammatischen Enkodierung der Aufbau der strukturellen Repräsentation abgeschlossen?

Nachfolgend sind die Unterschiede der beiden Hypothesen in einer Tabelle zusammengefasst (vgl. Tabelle 3.1):

	Lineare Inkrementalität	Strukturelle Inkrementalität
Entscheidende Faktoren für die Selektion von <i>starting points</i>	Eigenschaften einzelner Konzepte	Die Gesamtvorstellung bzw. strukturelle Repräsentation des Ereignisses
Ausmaß der Vorausplanung	Ein individuelles Konzept	Strukturelle Repräsentation
Zeitverlauf des Aufbaus der strukturellen Repräsentation	Nach Beginn der grammatischen Enkodierung	Vor Beginn der grammatischen Enkodierung

Tabelle 3.1: Zusammenfassung der Kernaussagen der linearen und der strukturellen Inkrementalitätshypothese

Ausgehend von den oben aufgeführten Fragen, werden in den nächsten Abschnitten die lineare und die strukturelle Inkrementalität sowie eine neue, integrierende Perspektive bezüglich des Verhältnisses zwischen diesen beiden Hypothesen anhand experimenteller Befunde näher diskutiert.

3.2.2 Befunde zur linearen Inkrementalität

Brown-Schmidt und Konopka (2008) erforschten die Größe der Verarbeitungseinheit in einem Phrasenbildungsexperiment. Die Forscherinnen gingen von dem typologischen Unterschied zwischen dem Englischen und dem Spanischen bezüglich des Aufbaus modifizierter Nominalphrasen aus: Während skalare Adjektive im Englischen vor modifizierten Nomina erscheinen (z. B. *the small butterfly*), treten sie im Spanischen postnominal auf (z. B. *la mariposa pequeña*). Spanisch-englischen Bilingualen wurden zuerst zwei gleich aussehende, jedoch unterschiedlich große Objekte (z. B. ein großer und ein kleiner Schmetterling) präsentiert und dann wurden sie aufgefordert, ein bestimmtes Objekt von den beiden entweder auf Englisch oder Spanisch zu benennen. Da die beiden Objekte visuell gleich aussahen, mussten die Probanden beim Benennen Modifikatoren (z. B. *small/large* oder *pequeño/gran*) verwenden, damit das Target-Objekt von dem Nichttarget-Objekt unterschieden werden kann. Die Augen- bzw. Blickbewegungen der Probanden wurden zeitgleich aufgezeichnet.

Die Autorinnen fanden heraus, dass die Fixationen hinsichtlich der Größen-Kontraste (also den Nichttarget-Schmetterling mit der jeweilig anderen Größe) einen sprachbezogenen Unterschied aufwiesen: Während der englischen Phrasenproduktion traten deutlich früher Fixationen hinsichtlich des Größenkontrastes auf als in der spanischen Versprachlichung. Da Fixationen auf einer Entität die Verarbeitung der von der Entität tragenden Information aufzeigen können, lässt sich aus den verschiedenen Blickbewegungsmustern im Spanischen und im Englischen schließen, dass die Größeninformation in den beiden Sprachen zu unterschiedlichen Zeitpunkten verarbeitet und dem modifizierten Target-Objekt

hinzugefügt wurde. Ebenfalls für unsere Zwecke interessant zu beobachten war, dass, wenn englische Sprecher den Größenkontrast zu einem relativ späteren Zeitpunkt fixierten, dies zu einer Störung des Redeflusses beitrug.

Die Ergebnisse deuten darauf hin, dass sich der Umfang der Vorausplanung lediglich auf ein individuelles lexikalisches Konzept begrenzt. Würden die Größeninformation und das Target-Objekt zusammen als eine Gesamtheit vorausgeplant, was zum Aufbau einer rudimentären Struktur der modifizierten Nominalphrase vor Beginn der Formulierungsphase führen kann, sollten keine unterschiedlichen Blickbewegungsmuster bei der Phrasenbildung in den beiden Sprachen erkennbar sein. Außerdem sollte das irrtümliche Platzieren der Fixation den Redefluss nicht beeinflussen. Aufgrund der Tatsache, dass sich ganz unterschiedliche Blickbewegungsparameter ergaben, wenn die Versuchspersonen modifizierte Nominalphrasen auf Englisch und Spanisch produzierten, schlussfolgerten die Autorinnen, dass die zu verbalisierende Information Konzept für Konzept verarbeitet wurde. Demzufolge kann die Verarbeitungseinheit auf der konzeptuellen Ebene so klein wie ein individuelles lexikalisches Konzept sein. Diese Schlussfolgerung stimmt im Wesentlichen mit der linearen Inkrementalitätshypothese überein, nach der die Versprachlichung von einem einzigen lexikalischen Konzept ausgehen kann.

Tomlin (1995) lieferte einen weiteren positiven Beleg für diese Hypothese. In seiner Studie sahen sich englischsprachige Versuchspersonen Videos an, in denen sich zwei Fische zuerst einander annäherten und letzten Endes ein Fisch vom anderen gefressen wurde. Ein Fisch wurde in den Videos immer mit einem endogenen Cue, einem Pfeil, markiert (Beispiel siehe Abbildung 3.2). Die Versuchspersonen sollten diesen allerdings ignorieren und sich auf die Videos und deren Beschreibung konzentrieren. Das Ergebnis war, dass eben dieser markierte Fisch, unabhängig von seiner semantischen Funktion (Agens vs. Patiens), bei der Beschreibung als Subjekt realisiert bzw. am Satzanfang positioniert wurde. Es wurden deutlich mehr Aktivsätze produziert, wenn sich der Cue auf das Agens (*white fish*) bezog, wohingegen die Passivkonstruktion bevorzugt wurde, wenn der Cue das Patiens

(*gray fish*) traf. Die hohe Antizipation der visuell prominenten Figur für die *starting points* wurde deswegen von dem Autor als Hinweis auf die lineare Inkrementalität interpretiert: Der Konzeptualisierung und der syntaktischen Enkodierung liegen die visuellen Eigenschaften einzelner an einem Ereignis beteiligten Entitäten zugrunde.

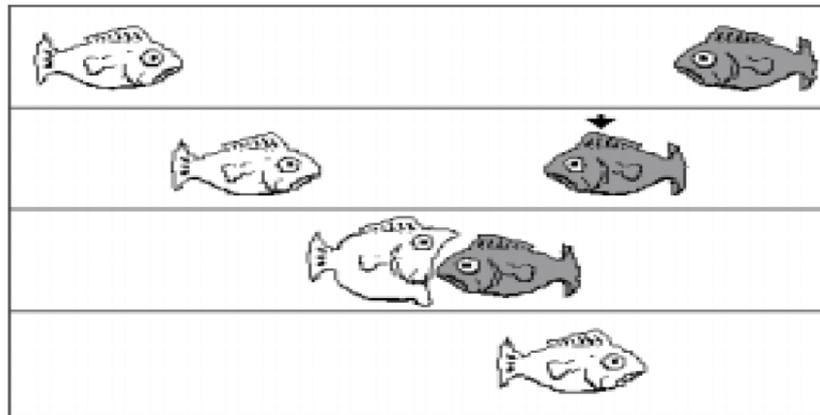


Abbildung 3.2: Beispielstimulus aus Tomlin 1997

In ihrem Ereignisbeschreibungsexperiment haben Gleitman et al. (2007) mithilfe des Eye-Trackings herausgefunden, dass die visuelle Aufmerksamkeit eine ausschlaggebende Rolle für das Auswählen von *starting points* spielt. Im Gegensatz zu Tomlins Versuchsdesign wurde in dieser Studie zur Manipulation der visuellen Aufmerksamkeit ein exogener visueller Cue benutzt, bei dem es sich um „saliente, hervorstechende Reize, wie beispielsweise ein kurzes Aufblitzen [handelt], die Aufmerksamkeit automatisch auf sich und damit auf den Ort lenken, an dem sie erschienen sind“ (Weidner & Fink 2007, S. 303). Bereits einige Untersuchungen haben gezeigt, dass durch das Nutzen eines solchen Cues die anschließende Aufmerksamkeit bzw. die Augenbewegungen manipuliert werden kann (z. B. Henderson & Macquistan, 1993; Vecera, Flevaris & Filapek, 2004). Dieser Manipulationseffekt spiegelt sich vor allem in der Platzierung der ersten Fixation wider: Im Gegensatz zu anderen Positionen wurden die, die kurz zuvor von einem visuellen Cue angezeigt wurden, früher von Versuchspersonen wahrgenommen und daher zuerst fixiert. Aufgrund dieses Befundes wurde in der Studie von Gleitman et al. (2007) immer ein exogener Cue auf einen Ereignisteilnehmer gerichtet. Ziel war,

herauszufinden, ob dieser markierte Ereignisteilnehmer, der so in der Wahrnehmungsphase visuell auffälliger war und daher mit größerer Wahrscheinlichkeit zuerst fixiert wurde, bei der späteren Ereignisbeschreibung eher als *starting point* selektiert würde.

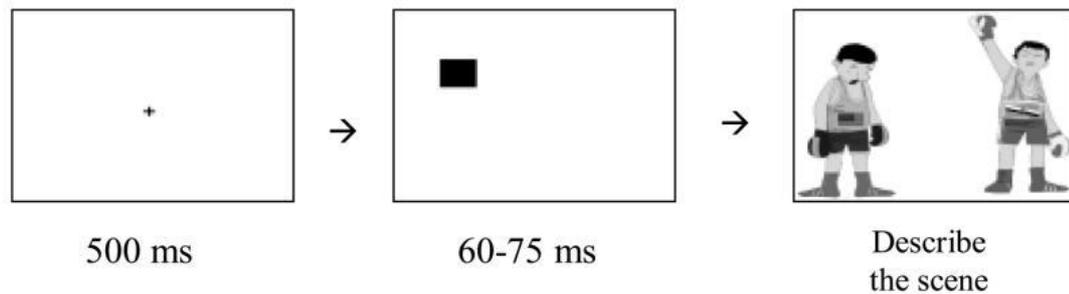


Abbildung 3.3: Zeitlicher Verlauf eines Durchgangs nach Gleitmann et al. 2007

In dieser Studie wurde ein kleines schwarzes Viereck als exogener Cue ausgewählt. Das Experiment wurde wie folgt aufgebaut: Vor jedem Durchgang war ein Fixationskreuz in der Mitte des Bildschirms zu sehen. Danach wurde der visuelle Cue auf einer Figur für ca. 60 - 75 ms positioniert. Nach der Präsentation des Cues erschien dann das jeweilige Target-Bild, welches immer zwei Figuren auswies, deren Belebtheitsgrad gleich war. Die Aufgabe der Probanden war, das Target-Bild so schnell wie möglich in ihrer Muttersprache, dem Englischen, zu beschreiben (Beispiellauf siehe Abbildung 3.3). Die produzierten Sätze und Blickbewegungen wurden von Beginn der Bilddarstellung bis zum Ende der Beschreibung rechnergestützt erfasst.

Die Stimuli ließen sich in vier Gruppen einteilen, die jeweils unterschiedliche linguistisch-strukturelle Relationen aufwiesen: (1) Aktiv-Passiv-Paare (z. B. *A girl is spraying a boy* oder *A boy is being sprayed by a girl*), (2) perspektivische Prädikate (z. B. *A dog is chasing a man* oder *A man is running away from a dog*), (3) symmetrische Prädikate (z. B. *A policeman shakes hands with a construction worker* oder *A construction worker shakes hands with a policeman*), (4) koordinierte NPs (z. B. *The cat and the dog* oder *The dog and the cat*).

Es wurde zuerst überprüft, ob der visuelle Cue in allen Bedingungen die frühzeitigen Augenbewegungen manipulieren kann. Ähnlich wie in vielen anderen Studien (z. B. Henderson & Macquistan, 1993; Vecera, Flevaris & Filapek, 2004) wurde ein Einfluss des visuellen Cues auf das Verorten der Erstfixation beobachtet: Die Sprecher tendierten dazu, zuerst diejenigen Figuren zu fixieren, die der Cue markiert.

Da die visuelle Wahrnehmung bzw. die perzeptuelle Verarbeitung in der konzeptuellen und grammatischen Enkodierungsebene münden muss, wurde in einem weiteren Schritt überprüft, ob die Figur, die aufgrund ihrer höheren visuellen Auffälligkeit als Erstes fixiert wurde, auch als Erstes konzeptualisiert und grammatisch enkodiert wird. Diese Frage kann anhand der produzierten Sätze und der aufgezeichneten Blickbewegungsdaten beantwortet werden. In Bezug auf die produzierten Sätze wurde herausgefunden, dass im Allgemeinen eine Korrespondenz zwischen dem Referenten, der zuerst fixiert wurde, und dem Referenten, der zuerst lexikalisch enkodiert wurde, besteht. In der ersten Bedingung war das Verhältnis der gebildeten Passivsätze deutlich höher, wenn die Erstfixation das Patiens im Gegensatz zum Agens traf. Ein weiteres Ergebnis war, dass die erste Fixation nicht nur die Oberflächenform von gebildeten Sätzen, sondern auch die Perspektive, aus der ein Ereignis betrachtet wird, beeinflussen kann. Ein Beispiel dafür sieht man in der zweiten Bedingung mit dem JAGEN/FLIEHEN-Ereignis. Wenn der Jagende als Erstes fixiert wurde, tendierten die Versuchspersonen dazu, diese Figur am Satzanfang zu platzieren und das Verb *jagen* zur Beschreibung auszuwählen. Im Gegensatz dazu wurde der Gejagte häufiger am Satzanfang platziert, wenn dessen Figur als Erstes fixiert wurde, was wiederum zur Aktivierung des Verbs *fliehen* bei der Beschreibung führte. Ähnlich verhielt es sich mit den symmetrischen Prädikaten (3. Bedingung) und den koordinierten NPs (4. Bedingung). Die Figuren, denen zuerst die visuelle Aufmerksamkeit gewidmet wurde, wurden häufiger als Erstes enkodiert und eher verbalisiert als diejenigen, die später fixiert wurden. Weitere Verarbeitungen betrafen weiterhin die Festlegung dieser Figur. Der Befund, dass die

zuerst fixierte Figur meistens als der erste Referent eines Satzes oder einer Phrase realisiert wurde, steht im Wesentlichen in Übereinstimmung mit der Idee der linearen Inkrementalitätstheorie, nach der die Satzproduktion konzeptgesteuert ist.

Die Übereinstimmung zwischen der ersten Fixation und dem ersten Wort wurde auch in den Augenbewegungsmustern beobachtet: Innerhalb 200 ms nach der Präsentation des Target-Bildes richtet sich das Augenmerk durch den visuellen Cue auf die spezifizierte Entität. Die Fixationen auf dieser Entität hielten meistens bis kurz vor dem Sprechanfang an, was auf einen anschließenden lexikalischen Zugriff der fixierten Entität hinweist. Diese Befunden galten für alle Bedingungen, in denen unterschiedliche linguistisch-strukturelle Relationen vorherrschten.

Aufgrund dieser Beobachtungen schlussfolgerten die Autoren, dass die Satzplanung visuell gesteuert ist und linear inkrementell abläuft. Sobald eine Entität aufgrund ihrer höheren visuellen Auffälligkeit zuerst fixiert wurde, wird sie umgehend als angemessener *starting point* des zu bildenden Satzes selektiert und konzeptuell und linguistisch verarbeitet. Der grammatischen Enkodierung einschließlich der Erstellung einer syntaktischen Struktur und einer Wortstellung liegt dann dieses selektierte lexikalische Konzept zugrunde. In diesem Prozess ist der Aufbau der semantisch-konzeptuellen Struktur eines Gesamtereignisses, an der parallel zur grammatischen Enkodierung gearbeitet wird, nicht erforderlich.

3.2.3 Befunde zur strukturellen Inkrementalität

Eine andere Studie, die sich ebenfalls mit der inkrementellen Verarbeitungsweise befasst, ist die von Hwang und Kaiser (2015). Zwar hat diese Studie keine direkte Evidenz für die strukturelle Inkrementalitätshypothese geliefert, dennoch soll sie an dieser Stelle aus zwei Gründen zitiert werden: Obwohl die Studie zum einen eine strukturell inkrementelle Verarbeitung nicht beweisen kann, hat sie mit dem gleichen Versuchsparadigma (*visual-cueing-Paradigma*) wie Gleitman et al. (2007) einen einzelsprachbezogenen Unterschied zwischen der englischen und der koreanischen Satzproduktion feststellen können. Dieser Befund fordert die universale Gültigkeit

der linearen Inkrementalität heraus. Zum anderen konnte die Studie zeigen, dass die Mechanismen der Satzplanung wahrscheinlich auch von den grammatischen typologischen Eigenschaften der Zielsprache beeinflusst werden.

Der Versuchsablauf glich dem der Gleitman et al. Studie. In dem Experiment sahen die koreanischsprachigen Versuchspersonen zunächst ein Fixationskreuz auf dem Bildschirm. Nach der Ausblendung des Kreuzes wurde ein visueller Cue für 80 ms gezeigt. Nach dem Erscheinen des Cues wurde ein aus zwei belebten Figuren bestehendes transitives Ereignis präsentiert. Die Versuchspersonen hatten die Aufgabe, die präsentierten Ereignisse so schnell wie möglich in ihrer Muttersprache, und zwar dem Koreanischen, zu beschreiben. In einer Hälfte der Trials trat der Cue auf der Region des Agens auf, während in der anderen Hälfte die Region des Patiens vorab von dem exogenen Cue angezeigt wurde. Die Augenbewegungen der Versuchspersonen wurden im Laufe der Versprachlichung aufgezeichnet.

Es zeigte sich, dass der visuelle Cue das Platzieren der ersten Fixation manipulieren konnte: Die Sprecher richteten ihre ersten Sakkaden eher auf diejenige Region, die zuvor vom visuellen Cue markiert wurde. Diese Ergebnisse bestätigen den Befund früherer Studien (z. B. Gleitman et al., 2007; Myachykov, Garrod & Scheepers, 2010; Myachykov & Tomlin, 2008). Im Gegensatz zu den Befunden der Satzproduktion im Englischen ist es interessant anzumerken, dass keine Effekte der frühen visuellen Aufmerksamkeitssteuerung auf die Selektion von *starting points* während der Satzplanung im Koreanischen beobachtet wurden. Das Verhältnis der gebildeten Passivsätze und OSV-Aktivsätze, in denen das Patiens satzinitial auftritt, erhöhte sich nicht, wenn sich Erstfixationen auf das Patiens bezogen. Dieser Befund widerspricht daher der Vorstellung der linearen Inkrementalitätsthese, nach welcher die Selektion von *starting points* von Bottom-Up-Faktoren wie etwa der visuellen Zugänglichkeit beeinflusst wird. Vergleichbare Ergebnisse findet man auch bei der Satzproduktion im Finnischen (Myachykov et al., 2010).

Als Grund für das Fehlen des visuellen Effekts auf die Auswahl von *starting points*

gaben die Autorinnen die im Koreanischen zur Verfügung stehende Kasusmarkierung an. Verglichen mit dem Englischen ist das Koreanische eine Sprache, in der grammatische Relationen überwiegend durch Kasus anstatt Wortstellung markiert werden. Mithilfe dieses grammatischen Mittels ist eine Satzposition (z. B. Satzanfang) nicht unbedingt mit einer bestimmten grammatischen Funktion (z. B. Subjekt) verbunden. So kann beispielsweise die satzinitiale Position im Koreanischen bei der Beschreibung eines transitiven Ereignisses entweder von der Subjekt-Funktion, was zur Bildung eines SOV-Satzes führt, oder von der Objekt-Funktion, was eine OSV-Struktur herbeiführt, belegt werden. Sollten Sprecher Informationen linear inkrementell verarbeiten und *starting points* aufgrund der visuellen Eigenschaft der Handlungsbeteiligten auswählen, müssten sie nach der Selektion noch weiter entscheiden, welche syntaktische Funktion dem *starting point* des zu bildenden Satzes zuzuweisen ist. Im Prozess der Funktionszuweisung wird eine Default-Regel, nach der die Subjekt-Funktion im Gegensatz zur Objekt-Funktion mit größerer Wahrscheinlichkeit früher zugeteilt wird, verfolgt wird (vgl. die Diskussion über die Zugänglichkeitshierarchie in Abschnitt 2.2.2). Dennoch ist die Möglichkeit, dass dem ersten Referenten die Objekt-Funktion zugewiesen wird, nicht völlig auszuschließen. Diese beiden syntaktischen Funktionen können zeitgleich Kandidaten für die grammatische Funktion des satzinitial aufzutretenden Wortes sein. Dies kann wiederum dazu führen, dass eine Konkurrenz bezüglich dieser syntaktischen Funktionszuweisung hervorgerufen wird, was die Satzplanung allerdings verlangsamen würde. Diese Arbeitsweise spricht somit gegen die lineare Inkrementalität.

Dennoch muss man berücksichtigen, dass, obwohl die hier zitierte Studie die universale Gültigkeit der linearen Inkrementalität in Frage stellt, sie keine klare Erklärung dafür liefert, in welcher Verbindung die koreanische Satzplanung mit der strukturellen Inkrementalität steht. Mit anderen Worten lassen sich aus diesen Ergebnissen keine direkten Evidenzen dafür finden, dass die semantische Ereignisstruktur während der Vorausplanung eines koreanischen Satzes konstruiert

wurde.

Stärkere und direktere Evidenz für die strukturell inkrementelle Verarbeitungsweise liefern Griffin und Bock (2000). In ihrer Studie wurde festgestellt, dass im Kontext der Ereignisbeschreibung Sprecher in der Lage sind, in einigen Bruchteilen der ersten Sekunde nach der Darbietung eines Ereignisses eine semantisch strukturelle Repräsentation dieses Ereignisses aufzubauen. Diese Erkenntnis gelang den Forscherinnen durch den Vergleich der Augenbewegungsparameter während einer Identifizierungs- und einer Beschreibungsaufgabe.

Jedes Trial in dem Experiment begann mit einem Fixationskreuz in der Mitte des Bildschirms. Danach wurde ein transitives Ereignis gezeigt (z. B. *A woman is shooting a man*; vgl. Abbildung 3.4). Mit der Darbietung des Stimulus begann auch das Aufzeichnen der Augenbewegungen. Die Versuchspersonen wurden in vier Gruppen unterteilt, die je eine der vier Aufgabenbedingungen erledigen mussten. Zwei Aufgaben, die für diese Diskussion am relevantesten sind, ist zum einen die *patient-detection*-Aufgabe. Hier mussten die Probanden das Patiens im Ereignis durch ein Knopfdrücken identifizieren. Zum anderen mussten die Probanden das präsentierte Ereignis so schnell wie möglich nach der Präsentation beschreiben (*extemporaneous-speech*-Aufgabe).

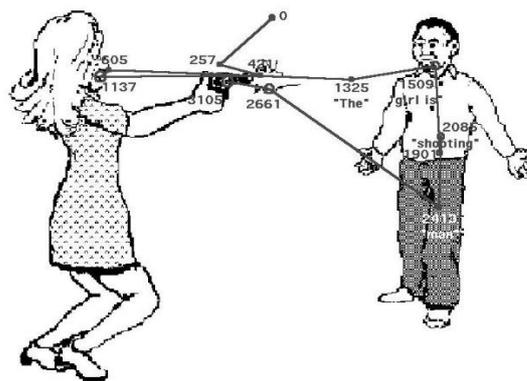


Abbildung 3.4: Beispielstimulus aus Griffin und Bock 2000

Die *patient-detection*-Gruppe diente hier im Wesentlichen nur als Vergleichsgruppe. Um das Patiens nach dem Erscheinen des Stimulus so schnell wie möglich zu

identifizieren, mussten die Versuchspersonen die semantische Struktur des Ereignisses zügig aufbauen. Dieser Konstruktionsprozess sollte sich in den aufgezeichneten Augenbewegungsdaten widerspiegeln. Wenn bei einer mündlichen Ereignisbeschreibung die semantische Struktur des jeweiligen Ereignisses ebenfalls sofort nach dessen Präsentation aufgebaut wird, müsste demnach ein vergleichbarer Blickbewegungsparameter wie bei der Patiens-Identifizierung beobachtet werden.

Die Ergebnisse der *patient-detection*-Aufgabe zeigten, dass die visuelle Aufmerksamkeit in der frühen Präsentationsphase gleichermaßen auf das Agens und das Patiens gerichtet ist. Eine Präferenz für das Patiens trat gegen 288 ms nach der Bildpräsentation auf und erreichte ca. 456 ms nach Stimulus-Onset eine statistische Signifikanz. Griffin und Bock interpretierten die visuelle Präferenz für das Patiens als Hinweis dafür, dass dieses zu diesem Zeitpunkt bereits von den Versuchspersonen erkannt wurde. Da das Erkennen den semantischen Strukturaufbau voraussetzt, wurde die Phase vor der Divergenz, in der gleichmäßige Fixationen auf das Agens und das Patiens gerichtet waren, deshalb als semantische Strukturierungsphase angesehen.

Wenn der mündlichen Satzproduktion der Aufbau der semantischen Struktur zugrunde liegt, dann ist während der Beschreibungsaufgabe nach Stimulus-Onset ein mit der Identifizierungsphase vergleichbares Augenbewegungsmuster zu beobachten. Diese Hypothese wurde bestätigt: In den ersten 316 ms richteten die Versuchspersonen ihren Blick gleichermaßen auf das Agens und das Patiens. Es konnte erst 336 ms nach Stimulus-Onset eine Präferenz über die Verteilung der Blickfrequenzen für eine visuelle Entität, die in der Beschreibung als Subjekt enkodiert wurde, ermittelt werden. Die Ergebnisse sprechen deswegen für die oben präsentierte Hypothese, da die Sprecher in der initialen Sprechphase versuchten, ein grobes Verständnis des Ereignisses und somit ein semantisches Gerüst für die Satzplanung zu haben. Ebenso verhielten sich die Sprecher auch in der nonverbalen Aufgabe. Allen weiteren Verarbeitungsprozessen, einschließlich der Selektion von *starting points*, liegt dann dieses semantische Gerüst zugrunde.

Zusammenfassend gesagt legen die oben erwähnten Studien die Schlussfolgerung nahe, dass der Satzproduktionsprozess strukturell inkrementell verläuft. In der Vorausplanungsphase wird die semantisch strukturelle Repräsentation, die die weiteren Verarbeitungen dirigiert, konstruiert. Die Selektion von *starting points* wird nicht von Eigenschaften einzelner Ereignisteilnehmer beeinflusst, sondern hängt von deren semantischen Rollen in der Gesamtvorstellung eines Ereignisses ab. Die grammatische Enkodierung beginnt erst dann, nachdem Sprecher auf die strukturelle Repräsentation zugegriffen haben.

3.3 Variabilität der inkrementellen Verarbeitung

Von den anscheinend widersprüchlichen Befunden bezüglich der englischen Satzproduktion ausgehend, die von Gleitman et al. (2007) und Griffin und Bock (2000) geliefert wurden, wurde in jüngerer Zeit eine weitere Betrachtungsmöglichkeit bezüglich des Verhältnisses zwischen der linearen und der strukturellen Inkrementalität vorgeschlagen. Ältere wissenschaftliche Ansichten betrachten beide Hypothesen als stark polarisierend. Eine neuere Annahme ist nun, dass sie sich nicht gegenseitig ausschließen, sondern im Gegenteil gemeinsam eine kontinuierliche Erklärung für die Arbeitsweise des Satzproduktionssystems sein können. In Anbetracht der Lage, dass bei der Satzproduktion verschiedene Faktoren (z. B. visuelle Zugänglichkeit und Ereignisstruktur) mitwirken und deren Interaktionen von einer Situation zur nächsten verschieden sein können, sollen Sprecher in der Lage sein, eine passende Strategie zur Informationsorganisation je nach konkreter Satzbildungssituation zu entwickeln.

Im Rahmen eines Bild-Wort-Interferenzparadigmas (engl.: *picture word interference paradigm*) setzten sich Wagner, Jescheniak und Schriefers (2010) kritisch mit dem Zusammenhang zwischen dem Ausmaß der Vorausplanung und der Komplexität der zu verbalisierenden Information auseinander. Deutsche Muttersprachler wurden dazu aufgefordert, Bilder, die zwei nebeneinanderstehende Figuren zeigten, mit dem

Satzmuster *Das [linke Objekt] ist neben dem [rechten Objekt]* (z. B. *Der Frosch ist neben dem Becher*) zu beschreiben. Zeitgleich hörten sie ein Störwort als Distraktor, das es zu ignorieren galt. In kritischen Trials war der Distraktor semantisch entweder mit der Bezeichnung des linken oder des rechten Objekts verwandt (z. B. *Kröte* für *Frosch* oder *Tasse* für *Becher*). In den Kontrolltrials stand der Distraktor in keinem namentlichen Zusammenhang mit den beiden Objekten (z. B. *Socke* für *Frosch*). Außerdem wurde der auditive Distraktor zu unterschiedlichen Zeitpunkten, also vor, synchron mit oder nach der Darbietung des Bildes präsentiert. Die Forscher wollten somit herausfinden, ob und in welchem Ausmaß die Vorausplanung des zu bildenden Satzes von den Distraktoren beeinflusst wird.

Vorangegangene Studien haben gezeigt, dass die Speech-Onset-Latenz, also das zeitliche Intervall vom Stimulus-Onset bis zum Aussprechen des ersten Lautes, von Störwörtern deutlich verlangsamt wird, wenn sie semantisch mit den Zielwörtern korrelieren (vgl. Klein, 1964; La Heij, Boelens & Kuipers, 2010; Neumann, 1980). Dieses linguistische Phänomen wird als semantischer Interferenzeffekt bezeichnet: Die Präsentation von Distraktoren führte zu Koaktivierungen semantischer Konkurrenten des Zielwortes, was eine Störung beim Wortzugriff aus dem mentalen Lexikon auslösen kann. Anhand der Größe der Interferenzeffekte in Bezug auf die Speech-Onset-Latenz ist es den Wissenschaftlern möglich, Rückschlüsse auf das Ausmaß der Vorausplanung zu ziehen.

Es wurde gezeigt, dass die Speech-Onset-Latenz in der kritischen Bedingung deutlich langsamer war als in der Kontrollbedingung. Beide Referenten, sowohl der zuerst als auch der zweit erwähnte, wurden von dem semantischen Distraktor beeinflusst. Dies spricht dafür, dass vor Sprechbeginn die semantischen Merkmale der beiden Referenten bereits aktiviert und dem Sprecher zugänglich waren. Wagner und Kollegen erachteten die Ergebnisse daher als Beleg für ein größeres Vorausplanungsausmaß, das zwei Entitäten umfasst.

In einem weiteren Experiment mussten die Versuchspersonen die Farbinformation

der dargestellten Objekte der Planung zuschreiben und in Sätze einbauen (z. B. *Der blaue Frosch ist neben dem blauen Becher*). Mit dem Hinzufügen der Farbinformation erhöhte sich bei der Vorausplanung der Verarbeitungsaufwand der Probanden. Es zeigte sich, dass die Interferenzeffekte für das Objekt in der zweiten Position in dieser experimentellen Situation verschwunden sind. Dies ist ein Zeichen dafür, dass die semantische Bedeutung des zweiten Objekts aufgrund der Planung der Farbinformation des ersten Objekts nicht mehr in die Vorausplanung miteingeschlossen war. In Verbindung mit den Ergebnissen des ersten Experiments schlussfolgerten die Autoren, dass das Ausmaß der Verarbeitungseinheit nicht feststehend ist, sondern von der Komplexität der zu vermittelnden Information beeinflusst wird.

Anhand der festgestellten Flexibilität bei der Vorausplanung kann man berechtigterweise vermuten, dass die lineare oder die strukturelle inkrementelle Arbeitsweise wohl nicht die einzige Arbeitsweise des Satzplanungssystems ist. Kuchinsky und Bock (2010) stellten die These auf, dass Sprecher wahrscheinlich in der Lage sind, die Verarbeitungsweise der konkreten Satzbildungssituation anzupassen.

Mithilfe des *visual-cueing*-Paradigmas, das auch in der Studie von Gleitman et al. (2007) und Hwang und Kaiser (2015) verwendet wurde, gingen Kuchinsky und Bock (2010) in ihrem Experiment der Frage nach, ob die Auswahl einer bestimmten inkrementellen Arbeitsweise (lineare vs. strukturelle Inkrementalität) von der Interaktion zwischen der Kodierbarkeit der Ereignisstruktur und einzelner an dem Ereignis beteiligten Entitäten abhängt. Die Autorinnen definierten die Kodierbarkeit einer Entität als Anzahl möglicher alternativer Bezeichnungen, die zur Beschreibung dieser Entität zunutze gemacht werden können (vgl. auch Wingfield, 1967, 1968; Griffin, 2001). Je mehr alternative Bezeichnungen ein Charakter hat, desto aufwändiger ist die Enkodierung dieses Charakters. Im Gegensatz dazu steht die Kodierbarkeit einer Ereignisstruktur in einem umgekehrten Verhältnis zur Verbanzahl, die Sprechern zur Ereignisbeschreibung zur Verfügung stehen:

Strukturelle Repräsentationen werden schneller aufgebaut, wenn sie nur mit einer kleinen Menge von Verben beschrieben werden können. Die zwei Ereignistypen („einfache“ vs. „schwere“ Ereignisse) und die zwei Charaktertypen („einfache“ vs. „schwere“ Charaktere) ergaben vier kritische Bedingungen: einfache Ereignisse mit einfachen Charakteren, einfache Ereignisse mit schweren Charakteren, schwere Ereignisse mit einfachen Charakteren und schwere Ereignisse mit schweren Charakteren.

In einem Experiment dieser Studie wurden englischsprachigen Versuchspersonen Ereignisse mit zwei Charakteren (beides belebt oder unbelebt) präsentiert und sie gebeten, das Ereignis zu beschreiben. Kurz vor Stimulus-Onset erschien ein exogener Cue auf der Position eines darzustellenden Charakters. Analog zu früheren Studien wurde hierbei herausgefunden, dass das Platzieren der ersten Fixation von der Position des visuellen Cues beeinflusst wurde. Allerdings wurde der Effekt des Cues nicht unbedingt bis zur Konzeptualisierung und lexikalischen Verarbeitung aufrechterhalten: Es zeigte sich, dass die Frage, ob die als Erstes fixierte Figur auch zuerst linguistisch verarbeitet und artikuliert wurde oder nicht, in einem engen Zusammenhang mit der Kodierbarkeit der Ereignisstruktur und der Konzepte der zu formulierenden Charaktere stand. In Tabelle 3.2 wurden die Effekte des visuellen Cues in den vier Bedingungen zusammengefasst. Die Tabelle zeigt die proportionale Differenz zwischen dem Prozentsatz der Sätze, die mit einem gecueteten Charakter begonnen haben, und dem der Sätze, die durch den gleichen Charakter initiiert wurden, welcher allerdings nicht vorab vom visuellen Cue spezifiziert war.

	Einfacher Charakter	Schwerer Charakter
Einfaches Ereignis	10%	1%
Schweres Ereignis	27%	-2%

Tabelle 3.2: Ergebnisse aus Kuchinsky & Bock 2010

Wenn ein Ereignis insgesamt einfach zu verstehen ist, tendierten die Probanden dazu, ohne Rücksicht auf die Position der ersten Fixation den Referenten, der in der Subjekt-Funktion stand, als den *starting point* der Verbalisierung auszuwählen. Wenn das Target-Ereignis schwer zu enkodieren war, neigten die Versuchspersonen dazu, Sätze mit dem durch den Cue markierten Charakter zu beginnen. Außerdem ließ sich auch eine Abhängigkeit des visuellen Effekts von der Kodierbarkeit des Charakters feststellen. Generell wurden Charaktere, die durch den visuellen Cue markiert wurden, häufiger als *starting points* selektiert, wenn sie einfach zu enkodieren waren.

Aus den Ergebnissen schlussfolgerten die Autoren, dass die Satzplanung weder rein reizinduziert (linear inkrementell) noch rein strukturinduziert (strukturell inkrementell) verläuft. Die Verarbeitungsstrategien, die in der Satzproduktion von Sprechern eingesetzt wurden, sind von einer Situation zur nächsten verschieden, und welche verfolgt wird, steht in Relation mit der Kodierbarkeit des zu verbalisierenden Ereignisses und der Ereignisteilnehmer (vgl. auch Konopka & Meyer, 2014).

Insgesamt deuteten die oben zitierten Studien auf eine Variabilität der Verarbeitungsweise bei der Satzplanung hin. Diese Charakteristik der Satzplanung steht grundsätzlich im Einklang mit dem sog. *minimal-load principle* im Diskurs, für welches Levelt (1989) plädiert: „When there are more alternatives, the principle says, do the simplest thing first“ (S. 144; vgl. auch Ferreira & Henderson, 1998; Konopka & Meyer, 2014). Demnach werden zuerst relativ einfache Informationen verarbeitet, damit sie sich nicht im Arbeitsgedächtnis anhäufen und zu einer hohen kognitiven Belastung führen. Die frühzeitige Enkodierung einfacher Ereignisteilnehmer in einem schwer enkodierbaren Ereignis erlaubt zum einen eine unverzügliche Verarbeitung weiterer Entitäten, und zum anderen liefert sie mit der Festlegung der zuerst zu verbalisierenden Entität eine Bottom-Up-Unterweisung für den Aufbau der Ereignisstruktur. Im Gegensatz dazu ermöglicht eine frühe Verarbeitung einfacher Strukturen eine schnelle Zusammensetzung aller relevanten Konzepte und somit einen beschleunigten Satzbau unter der Direktion struktureller

Informationen.

Zusammengenommen legen die oben genannten Studien nahe, dass bei der Satzplanung dem Sprecher verschiedene Verarbeitungsstrategien zur Verfügung stehen. Weder die lineare noch die strukturelle Inkrementalitätshypothese liefern eine ausreichende Erklärung für die beobachteten Verarbeitungsmechanismen der Satzplanung. Welche Vorgehensweise als Verarbeitungsstrategie ausgewählt wird, ist von der Kodierbarkeit der Ereignisstruktur und einzelner lexikalischer Konzepte abhängig.

3.4 Fazit

Ausgehend von dem von MacWhinney (1977) formulierten *starting points problem* wurden in dem vorliegenden Kapitel drei Annahmen zur Beschreibung des Satzproduktionsprozesses eingeführt. Die These der linearen Inkrementalität nimmt an, dass der Selektion von *starting points* Eigenschaften einzelner Ereignisteilnehmer zugrunde liegen. Entitäten, die aufgrund ihrer hohen Zugänglichkeit früher konzeptualisiert und identifiziert werden, werden dementsprechend auch früher syntaktisch enkodiert und linear angeordnet. Da nach dieser These der Verarbeitungsvorgang auf Eigenschaften einzelner Handlungsbeteiligten aufbaut, wird die lineare Inkrementalität dementsprechend auch als Bottom-Up-Inkrementalität bezeichnet. Im Gegensatz dazu besagt die These der strukturellen Inkrementalität, dass die Konstruktion bzw. das Erfassen einer strukturellen Repräsentation (hier vor allem der semantischen Struktur) für die Vorausplanung einer Szenenbeschreibung notwendig ist. Die Selektion von *starting points* erfolgt nicht aufgrund der visuellen oder konzeptuellen Merkmale von Ereignisteilnehmern, sondern wird von ihren semantischen Rollen beeinflusst, die den jeweiligen Teilnehmern zugeordnet werden (Top-Down-Hypothese). Sowohl die lineare als auch die strukturelle Inkrementalitätshypothese wurden durch frühere experimentelle Studien untermauert (z. B. Tomlin, 1997; Gleitman et al., 2007;

Myavhov, Garrod & Sheepers, 2010; Griffin & Bock, 2000; Lee, Brown-Schmidt & Watson, 2013). Es wurde gezeigt, dass die Arbeitsweise des Satzplanungssystems wohl in einem engen Zusammenhang mit den typologischen Merkmalen der Zielsprache steht. Was allerdings noch nicht völlig geklärt ist, dass selbst für dieselbe Sprache (z. B. Englisch) zwei entgegengesetzte Schlussfolgerungen gezogen wurden (vgl. Gleitman et al, 2007; Griffin & Bock, 2000). Ausgehend von dieser Beobachtung wurde eine dritte Hypothese aufgestellt, die behauptet, dass die Satzplanung in einer Sprache keinem festgelegten Ablaufplan folgt. Welche Strategie gewählt wird, ist abhängig von der Verfügbarkeit der strukturellen Repräsentation des Gesamtereignisses und der an dem Ereignis beteiligten individuellen Entitäten.

Da die Planungs- bzw. Verarbeitungsweise der Satzproduktion sprachspezifische Unterschiede aufweist, ist es erforderlich, mehrere Sprachen neben dem Englischen zu untersuchen. In der vorliegenden Studie steht die Vorgehensweise der deutschen Satzplanung im Fokus. Der Grund dafür liegt darin, dass im Gegensatz zum Englischen Deutsch eine Kasusprache ist, in der syntaktische Funktionen wenig mit der Wortstellung zu tun haben, sondern viel mehr durch eine Kasusmarkierung sichtbar werden. Ausgehend von dieser syntaktischen Eigenschaft des Deutschen ist deswegen anzunehmen, dass das Satzproduktionssystem im Deutschen wohl einer unterschiedlichen Arbeitsweise folgt als das im Englischen. Dazu ist auch zu bemerken, dass in der früheren Diskussion über die Inkrementalität der Einfluss der kontextuellen Repräsentation, die ebenfalls der strukturellen Repräsentation auf der konzeptuellen Enkodierungsebene zugeordnet wird, wenig berücksichtigt wurde. So wird sich mit dieser Frage ebenfalls in der vorliegenden Arbeit auseinandergesetzt.

4 Salienz und Satzproduktion

Die Satzproduktion ist ein hochgradig komplexer kognitiver Prozess, bei dem verschiedene kognitive Prozesse, darunter die Wahrnehmung, Aufmerksamkeit, Wissensbasis und das Arbeitsgedächtnis zusammenwirken. Da sich der Satzproduktionsvorgang inkrementell aufbaut, stehen Sprecher zu Beginn der Planung vor der Aufgabe, den ersten Referenten bzw. das erste Konzept zu selektieren, mit dem sie das Sprechen beginnen wollen. Das lexikalische Konzept und somit der Worteintrag, der zur Verbalisierung als Erstes selektiert wird, bezeichnet man als den sog. *starting point* einer Äußerung.

Einige Psycholinguisten nehmen an, dass die Auswahl von *starting points* von unserer Aufmerksamkeit geleitet wird, die bei der Sprachplanung grundsätzlich als ein mentales Sieb dient (vgl. z. B. Myachykov, Thompson, Garrod & Scheepers, 2012; Zarcone, van Schijndel, Vogels & Demberg, 2016). Referenten, die für die visuelle, konzeptuelle oder lexikalische Verarbeitung am zugänglichsten bzw. am relevantesten sind, werden in der Regel früher herausgefiltert und als *starting points* selektiert als diejenigen, die sich in der Umgebung befinden und wenig zugänglich oder wichtig sind. Eine experimentelle Unterstützung für diese Annahme liefert eine Studie von Tomlin (1995), die in Abschnitt 3.2.2 vorgestellt wurde. Ein zentrales Ergebnis der Studie war, dass bei der Beschreibung einer FRESSEN-Szene mit zwei Fischen immer ein Pfeil auf einen Fisch gerichtet war. Dieser wurde dann öfter als *starting point* der erzeugten Sätze ausgewählt als der andere Fisch. Laut Tomlin (1995) und Myachykov, Garrod und Scheepers (2009) ist der Befund auf die fokussierte Aufmerksamkeit zurückzuführen: Da mit einem Pfeil längere Zeit auf einen der beiden Fische gezeigt wurde, richteten die Sprecher ihre visuelle Aufmerksamkeit mehr auf diesen Fisch. Dies führte dann wiederum zu einer früheren konzeptuellen und linguistischen Verarbeitung dieser Entität.

Angesichts der großen Bedeutung der Aufmerksamkeit für die Selektion von *starting points* stellt sich nun die Frage, von welchen Faktoren die Lenkung der

Aufmerksamkeit im Kontext von Ereignisbeschreibungen beeinflusst wird. Im Folgenden wird nun zum Forschungsstand zunächst ein Überblick über zwei mitwirkende Faktoren - die visuelle und die linguistische Salienz - gegeben. Ausgehend von der Unterscheidung werden in Abschnitt 4.2 – 4.4 die Interaktionen zwischen den beiden Salienz-Formen und der Satzproduktion genauer mithilfe von experimentellen Befunden diskutiert. Bereits in Kapitel 2 wurde das Verhältnis zwischen verschiedenen Faktoren und der Satzplanung kurz erwähnt. Im folgenden Kapitel werden die Funktionen solcher Faktoren sowohl terminologisch als auch empirisch näher beschrieben. Der Schwerpunkt der Diskussion soll auf der Frage liegen, in welcher Art und Weise die Selektion von *starting points* und die syntaktische Strukturierung von distinktiven Salienz-Formen beeinflusst werden.

4.1 Salienz

Salienz ist ein viel verwendeter Begriff, der in unterschiedlichen theoretischen Kontexten unterschiedliche Bedeutungsinhalte und deshalb auch verschiedene Implikationen hat. Generell bezieht sich die Salienz auf die Auffälligkeit bzw. Zugänglichkeit, die die Aufmerksamkeit steuert und dadurch das sowohl perzeptive als auch mentale Durchforsten verkürzen kann. Diese Arbeit konzentriert sich, wie oben erwähnt, auf die folgenden zwei Salienz-Dimensionen, die visuelle und die linguistische Salienz.

Bei der visuellen Salienz handelt es sich um die durch die optische Sinneswahrnehmung greifbare Auffälligkeit, die durch ein Bündel distinktiver visueller Merkmale (z. B. Farben, Größen, Positionen, Orientierungen und Distanzen) erzeugt wird (vgl. Treisman & Gelade, 1980; Itti & Koch, 2000). Entitäten, die auffälligere visuelle Merkmale aufweisen und sich somit von anderen abheben, haben dementsprechend eine höhere Wahrnehmungswahrscheinlichkeit, das heißt, ihnen wird visuell mehr Aufmerksamkeit gewidmet. Da die visuell wahrgenommenen Eigenschaften inhärent in Stimuli verankert sind, wird die

Aufmerksamkeitslenkung durch die visuelle Salienz als Bottom-Up-Prozess verstanden. Bei diesem Prozess ist das Blickverhalten des Betrachters reizinduziert (vgl. Wolfe, 1992, 1998).

Im linguistischen Kontext wird die Salienz als „die kognitive Auffälligkeit eines sprachlichen Merkmals [verstanden], in dem Sinne, dass ein sprachliches Element aus einem Kontext hervorgehoben wird und dadurch dem Sprachbewusstsein leichter und schneller zugänglich ist als nicht-saliente Varianten“ (Lenz 2010, S. 94). Im Unterschied zur visuellen Salienz weist die linguistische Salienz überwiegend inhalts- oder diskursinduzierte Implikationen auf, die von semantischen und kontextuellen Merkmalen abhängig sind. Dementsprechend lässt sie sich in zwei weitere Elemente - die semantische und kontextuelle Salienz - aufteilen, die beide eine hohe Relevanz für die vorliegende Arbeit haben.

Die semantische Salienz fußt auf der semantischen Struktur eines zu beschreibenden Ereignisses und fällt im Wesentlichen mit der thematischen Hierarchie zusammen. Als eine Default-Regel liefert die thematische Hierarchie Hinweise darauf, in welcher Reihenfolge die semantischen Rollen auf die grammatische Ebene abgebildet und dann syntaktisch enkodiert werden. Typischerweise wird sie wie folgt in einer Salienzskala repräsentiert (nach Jackendoff, 1990; Bresnan & Kanerva, 1989; Primus, 1996):

Agens > Wahrnehmungsträger > Instrument > Thema/Patiens >
Lokation/Ausgangspunkt/Endpunkt

In dieser Skala nimmt die Salienz der thematischen Rollen von links nach rechts ab. Zwei thematische Funktionen, die am häufigsten untersucht und auch in dieser Arbeit betrachtet werden, sind das Agens und das Patiens. Das Agens bezeichnet Figuren, die eine Tätigkeit oder Handlung initiieren, während dem Patiens Entitäten zugeschrieben werden, die von einem Vorgang oder Zustand betroffen sind. Insgesamt wird das Agens in einem transitiven Ereignis als mental auffälliger als das Patiens angesehen, da es mehr Aktivität und Handlungsfähigkeit aufweist bzw.

besitzt. Diese Tendenz kann sich z. B. in den Abbildungen semantischer Funktionen auf die syntaktische Ebene widerspiegeln: Für die meisten Sprachen gilt, dass in einem transitiven Ereignis die Entität mit der Agens-Rolle häufiger als Subjekt und die mit der Patiens-Rolle häufiger als Objekt realisiert wird (vgl. Fillmore, 1968; Fischer, 2013). Da das Subjekt - wie oben in Abschnitt 2.2.2 erwähnt - im Vergleich zu anderen syntaktischen Funktionen generell früher zugewiesen wird, kann man aus dieser Tendenz rückschließen, dass das Agens im Gegensatz zum Patiens dementsprechend zu einem früheren Zeitpunkt identifiziert und konzeptualisiert wird.

In Bezug auf die kontextuelle Salienz geht man davon aus, dass die zu verbalisierende Information diskurs- und situationsbedingt häufig kontextuelle Eigenschaften wie etwa Gegebenheit (oder Neuheit), Kontrast oder Topikalität aufweist. Ähnlich wie semantische Rollen sind kontextuelle Merkmale hinsichtlich deren Prominenz nicht gleichwertig stark. Ob ein Referent als kontextuell salient wahrgenommen wird oder nicht, steht in einem engen Zusammenhang mit dem Wissensstand der am Gespräch Beteiligten und der kommunikativen Absicht. Beispielsweise sind Informationen, die in einer vorigen Äußerung erwähnt wurden, in der Regel salienter als diejenigen, die neu in den Diskurs eingeführt werden. Der Grund dafür liegt darin, dass gegebene Informationen aufgrund ihrer durch die Vorerwähnung erhöhten Zugänglichkeit aktuell im Arbeitsgedächtnis verankert sind. Ein anderes Beispiel der kontextuellen Salienz ist die Topik-Kommentar-Struktur: Laut Lambrecht (1994) ist das Topik ein Referent, der „of current interest“ ist (S. 163), und ist daher mental auffälliger als die anderen Informationen, die im anschließenden Kommentar erwähnt werden.

Da die semantische und kontextuelle Salienz einer Entität nicht inhärent sind, sondern sich über eine Interpretation der zu beschreibenden Szene und des gesamten Kontextes aufbauen, ist die diesen beiden Formen zugrundeliegende Verarbeitung ein Top-Down-Prozess, welcher der visuell bedingten Bottom-Up-Verarbeitung gegenübersteht (Claus, 2011). Im Folgenden wird auf die Interaktionen zwischen der

Satzproduktion und den eben beschriebenen drei Saliens-Typen näher eingegangen.

4.2 Visuelle Saliens und Satzproduktion

Sowohl in der visuell-kognitiven als auch psycholinguistischen Forschung ist die visuelle Saliens ein äußerst intensiv erforschter Bereich. Wie bereits im vorigen Abschnitt erwähnt, ergibt sich diese aus physikalisch messbaren Eigenschaften wie zum Beispiel Farben, Größen, Positionen oder Orientierungen (vgl. Rosenholtz, 1999). In Bezug auf das Verhältnis zwischen der visuellen Saliens und der Satzproduktion haben viele Studien gezeigt, dass die visuelle Saliens die Selektion von *starting points* und die syntaktische Strukturierung in vielerlei Hinsicht beeinflusst. Zum Beispiel wurde in einigen Studien herausgefunden, dass bei der Ereignisbeschreibung die Entität, die aufgrund ihrer hohen visuellen Saliens zuerst fixiert wird, mit größerer Wahrscheinlichkeit als der *starting point* des zu produzierenden Satzes selektiert wird (z. B. Gleitman et al., 2007; Myachykov & Tomlin, 2008). Zudem wurde herausgefunden, dass es im Allgemeinen einen Zusammenhang zwischen Blick- und Wortabfolgen gibt: Bei der Beschreibung eines Ereignisses können Sprecher mit ihrem Blick die Abfolge der auf der sprachlichen Oberfläche erscheinenden Referenten antizipieren (z. B. Griffin & Bock, 2000; Norcliffe, Konopka, Brown & Levinson, 2015; Sauppe, Norcliffe, Konopka, van Valin & Levinson, 2013).

Die Rolle der visuellen Wahrnehmung wurde in den vorhandenen Satzproduktionsmodellen bisher kaum berücksichtigt. Daher ist es notwendig, diesen Verarbeitungsprozess, der im Kontext einer linguistischen Bildbeschreibung den Input für weitere Verarbeitungen liefert, dem Satzproduktionsprozess zuzuordnen. Dazu werden wir uns an dem einflussreichen und allgemein anerkannten modularen Modell, welches vor allem zur Beschreibung der Enkodierung eines visuell dargestellten Ereignisses entwickelt wurde (z. B. Marr, 1982; Ullmann, 1996; Henderson & Hollingworth, 1999), orientieren. Laut diesem Modell wird die visuelle

Wahrnehmung nicht als ein eigenständiger Vorgang, sondern als die Vorphase der Konzeptualisierung betrachtet. Demnach besteht die Konzeptualisierung nicht nur – wie Bock und Levelt (1994) annehmen - aus einem Top-Down-Verarbeitungsprozess, in dem der Inhalt und die Art und Weise der Formulierung festgelegt werden, sondern auch aus einem Bottom-Up-Verarbeitungsprozess, durch welchen visuelle Informationen als Input anschließender Top-Down-Verarbeitungen aufgenommen werden. Diese beiden Prozesse laufen stufenweise ab: Die perzeptuelle Verarbeitung bildet eine niedrige Stufe der Konzeptualisierung, bei der es sich vor allem um „extraction of physical properties such as depth, color, and texture from an image as well as the generation of representations of surfaces and edges“ handelt, während die konzeptuelle Verarbeitung, vor allem die Extraktion der Ereignisstruktur und die Identifizierung von Referenten, einer höheren Konzeptualisierungsstufe angehören, die sich auf „the active acquisition of information, short-term memory for visual information, and the identification of objects and scenes“ bezieht (Henderson & Hollingworth 1999, S. 244). Basierend auf der niedrigen Stufe der Konzeptualisierung, nämlich der Wahrnehmung, wird die konzeptuelle Verarbeitung der wahrgenommenen Entitäten und Sachverhalte vorgenommen. In dieser Arbeit verfolgen wir diese Zuordnung.

Bei der sprachlichen Beschreibung eines visuell dargebotenen Ereignisses stehen Sprecher meist unter einem gewissen Zeitdruck. Um eine flüssige Sprachproduktion zu sichern, müssen Sprecher das wahrgenommene Ereignis mit minimaler Zeitverzögerung konzeptuell verarbeiten. Deswegen müssen Sprecher laut mancher Psycholinguisten alle Informationen benutzen, die ihnen umgehend nach Stimulus-Onset zur Verfügung stehen (vgl. z. B. Gleitman et al., 2007). Diese Informationen können sich vor allem auf die visuellen Informationen beziehen, da sie im Vergleich zu semantischen und kontextuellen inhärent sind und ohne weitere Interpretationen, die auf dem Inhalt oder dem Kontext basieren, sofort nach der Präsentation der Stimuli entnommen werden können. Deshalb wird von manchen Psycholinguisten angenommen, dass bei der Satzproduktion visuellen Eigenschaften

– insbesondere bei der Bestimmung von *starting points* – eine zentrale Rolle zukommt. Sobald Figuren in der Konzeptualisierungsstufe aufgrund ihrer höheren visuellen Salienz früher wahrgenommen und identifiziert wurden, werden sie unmittelbar in die folgende grammatische Enkodierungsebene übertragen und linguistisch verarbeitet. Dieser Prozess kann hochgradig automatisiert sein (vgl. z. B. Gleitman et al., 2007).

Die Annahme, dass die visuelle Salienz ausschlaggebend für die Selektion von *starting points* ist, stimmt grundsätzlich mit der linearen Inkrementalitätshypothese überein. Daher liefert das Ergebnis der Studie von Gleitman et al. (2007), die bereits in Abschnitt 3.2.2 als Evidenz für die lineare Inkrementalitätshypothese vorgestellt wurde, ebenfalls Evidenz für die zentrale Rolle der visuellen Salienz bei der Satzplanung. Wenn bei einer Szenen-Beschreibung ein Referent durch einen zuvor präsentierten exogenen Cue früh die Aufmerksamkeit auf sich gelenkt hat, dann wird dieser im Gegensatz zu anderen Kandidaten mit einer höheren Wahrscheinlichkeit als Subjekt realisiert und satzinitial angeordnet als andere Kandidaten. Die Aufmerksamkeit manifestiert sich in der Platzierung der ersten Fixation. Neben der Untersuchung der Satzplanung im Englischen von Gleitman et al. haben Myakykov und Tomlin (2008) in ihrer Studie zur Untersuchung der russischen Satzproduktion mit dem gleichen Versuchsparadigma, eine zentrale Rolle der visuellen Salienz bei der Auswahl von *starting points*, festgestellt. Um den Zusammenhang zwischen der visuellen Salienz und der ersten Fixation aufzuklären, haben die Autoren die Funktion der visuellen Aufmerksamkeit hervorgehoben. Ihnen zufolge tendieren Sprecher meist dazu, ihre Aufmerksamkeit bei einer Szenen-Beschreibung auf diejenigen Ereignisteilnehmer zu richten, deren visuelle Salienz am größten ist. Die visuelle Salienz eines Referenten wurde dadurch manipuliert, dass die Autoren einen bestimmten Referenten in der Position, die vorab durch einen visuellen Cue markiert wurde und deswegen bereits eine große Menge visueller Aufmerksamkeit anzog, präsentierten. Die Akkumulation von Aufmerksamkeit kann weiterhin dazu führen, dass die visuell fokussierte Figur im Vergleich zu den anderen, die visuell weniger

prominent sind und deshalb wenig oder kaum betrachtet werden, schneller identifiziert und grammatisch enkodiert wird (vgl. auch Ferreira & Dell, 2000; Ferreira, 2003; Arnold et al., 2000). Aufgrund dieser Analyse nehmen die Autoren daher an, dass die visuelle Salienz der Ereignisteilnehmer eine wesentliche Rolle bei der Festlegung von *starting points* spielt. Zieht beispielsweise die Figur *Katze* bei der Ereignisbeschreibung *Ein Hund jagt eine Katze* reizinduziert die frühe Aufmerksamkeit auf sich, wird sie dadurch schneller erkannt und lexikalisch sowie syntaktisch funktional früher enkodiert - so die Annahme. Der frühere Zugriff auf die Lemma-Information und die syntaktische Funktion von *Katze* führt weiter dazu, dass sie auch frühzeitiger syntaktisch enkodiert und linearisiert wird. So sind hier in diesem Fall mehr Passivsätze zu erwarten, als in einer Satzproduktionssituation, wo sich die visuelle Salienz von *Hund* und *Katze* nicht voneinander unterscheiden.

Aufgrund der vorgebrachten Diskussion lässt sich schlussfolgern, dass die Satzproduktion einschließlich der syntaktischen Strukturierung und der Wortselektion durch visuelle Salienz manipuliert werden kann. Entitäten mit höherer visueller Salienz werden früher als visuell weniger saliente Entitäten wahrgenommen und sind der Konzeptualisierung dementsprechend zugänglicher. Dies erlaubt eine frühere Identifizierung und Selektion als *starting points* der zu produzierenden Äußerung. Sobald *starting points* festgelegt wurden, werden sie ohne Verzögerung in die grammatische Enkodierungsebene überführt, auf der syntaktische Strukturen und Wortstellungen erzeugt werden. Dieser Prozess kann automatisch ablaufen und setzt keine Interpretation des Gesamtereignisses, sowohl semantisch als auch kontextuell, voraus (Wolfe, 1992, 1998). Aufgrund dieser Analyse nehmen manche Psycholinguisten an, dass der visuellen Salienz eine zentrale Rolle bei der Satzproduktion zukommt.

4.3 Semantische Salienz und Satzproduktion

In einem Ereignis sind deren Teilnehmer nicht als einzelne unabhängige Elemente zu

betrachten, sondern als Partizipanten, die durch zugewiesene semantische Rollen miteinander interagieren. Somit wird zu einem Ereignis eine thematisch strukturelle Repräsentation aufgebaut. Aufgrund von verschiedenen semantischen Rollen in einem Ereignis variiert die Komplexität der semantischen Repräsentation von einer Szene zur anderen. Trotz der verschiedenen Komplexitätsstufen wird angenommen, dass die in einem Ereignis involvierten thematischen Rollen nicht unstrukturiert aufgerufen werden. Stattdessen tendieren Sprecher dazu, Entitäten mit höherer semantischer Salienz früher zu aktivieren als diejenigen, die semantisch weniger auffällig sind. Diese tendenzielle Neigung spiegelt sich insbesondere bei der Beschreibung von transitiven Ereignissen wider, bei der zum Aufbau einer semantischen Struktur die Agens- und Patiensrolle zugewiesen werden müssen.

Eine Reihe von Korpusstudien hat gezeigt, dass Sprecher dazu geneigt sind, transitive Ereignisse in einer Agens-Patiens-Reihenfolge (engl.: *agent-patient-ordering*) zu beschreiben. Bei Dryer (2013) wurde die Anordnung von Subjekt, Objekt und Verb von insgesamt 1377 Sprachen untersucht. Die deklarativen Sätze aller untersuchten Sprachen wurden insgesamt sechs Wortstellungsstrukturen (SOV, SVO, VSO, VOS, OVS und OSV) zugeordnet. Es kam heraus, dass 1188 (86%) Sprachen eine dominante Wortstellung haben; darunter wird in 565 (48%) Sprachen überwiegend nach der kanonischen SOV-Struktur (z. B. Japanisch) und 488 (41%) nach der kanonischen SVO-Struktur (z. B. Mandarin) gereiht. Da in einer kanonischen Satzstruktur die Subjekt- und Objekt-Funktion jeweils mit dem Agens und dem Patiens belegt sind, trifft der Autor anhand der Daten auch eine Aussage über die primäre Reihenfolge der semantischen Rollen in den untersuchten Sprachen: Die Referenz auf das Agens und Patiens verfolgt primär die Agens-Patiens-Reihenfolge, wobei das Agens meist als Subjekt realisiert wird und am Satzanfang erscheint. Ähnliche Ergebnisse wurden ebenfalls von Greenberg (1963) und Tomlin (1986) geliefert, welche jedoch auf einem kleineren Datenbestand beruhen.

Weitere Studien, die den Zeitverlauf des Aufbaus der semantischen Struktur zum

Forschungsgegenstand haben, liefern zudem Evidenzen dafür, dass das Agens einen Vorrang bei der konzeptuellen Verarbeitung hat (z. B. Dobel, Gumnior, Bölte & Zwisserlood, 2007; Segalowitz, 1982; Webb, Knott & MacAskill, 2010). Eine Beispielstudie dafür ist die von Dobel et al. (2007), in der festgestellt wurde, dass das Agens im Gegensatz zum Patiens im Allgemeinen früher identifiziert wird. In diesem Experiment wurde der Verarbeitungsprozess zwischen der visuellen Reizaufnahme und der Verbalisierung zu verschiedenen Zeitpunkten unterbrochen (100, 200, 250 und 300 ms nach Stimulus-Onset). Außerdem wurde die Ereigniskohärenz der im Experiment verwendeten Stimuli, die entweder ein GEBEN- oder ein SCHIEßEN-Ereignis darstellten, systematisch manipuliert: Bei kohärenten Stimuli standen sich das Agens und das Patiens gegenüber, während sie bei inkohärenten Stimuli Rücken an Rücken standen (Beispiel siehe Abbildung 4.1). Die Aufgabe der Versuchsteilnehmer bestand darin, direkt nach der Präsentation der Stimuli (1) die Kohärenz der präsentierten Szene zu entscheiden und (2) dargestellte Agentia, Patientia, Objekte und Handlungen zu benennen. Die unterschiedliche Länge der zeitlichen Darstellung erlaubt einen Einblick darin, inwiefern die Bildinformation bis zu einem der erwähnten Zeitpunkte bereits mental repräsentiert und spezifiziert wurde. Diese Studie ist für die vorliegende Arbeit insofern von Relevanz, als sie zumindest darüber Auskünfte gibt, wie viel Zeit für den Aufbau einer Repräsentation, in der die Ereignisstruktur spezifiziert wird, und für die Identifizierung der semantischen Rollen nach Stimulus-Onset beansprucht wird.

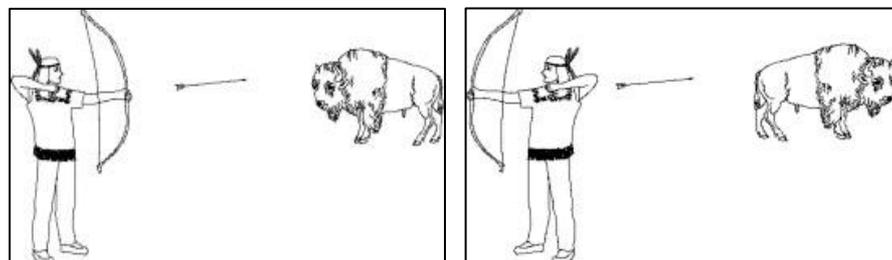


Abbildung 4.1: Beispielstimuli aus Dobel et al. 2007; kohärentes Ereignis (links) und inkohärentes Ereignis (rechts)

Hinsichtlich der Szenenkohärenz war es den Versuchsteilnehmern möglich, in ca. 80% der Trials die Kohärenz der Stimuli richtig zu beurteilen. Es wurden keine signifikanten Unterschiede zwischen den Variationen der Präsentationsdauer identifiziert. In Bezug auf die Benennung der Ereignisteilnehmer ergaben sich allerdings signifikante Unterschiede zwischen den unterschiedlichen experimentellen Bedingungen, die die Länge der zeitlichen Darstellung des Stimulus betrafen, ebenso zwischen den unterschiedlichen Ereignisteilnehmern. In der experimentellen Bedingung mit einer Präsentationsdauer von 100 ms wurden in 30% der Trials das Agens, in 28% das Patiens, in 4% das Objekt und in 19% die Handlung richtig benannt. In einer Bedingung mit einer Präsentationsdauer von 200 ms lag die Benennungsrichtigkeit des Agens bereits bei ca. 74 %, welche mit 14% deutlich höher als die des Patiens war. Das Objekt und die Handlung wurden jeweils in 12% und in 34% der Trials richtig benannt. Bei einer längeren Präsentationsdauer war eine weitere Erhöhung des Wertes weder für das Agens noch für das Patiens zu beobachten.

Der Befund, dass in allen experimentellen Bedingungen in ca. 80% der Trials die richtige Entscheidung über die Kohärenz getroffen wurde, ist ein Zeichen dafür, dass Sprecher in der Lage sind, innerhalb kürzester Zeit (100 ms nach Stimulus-Onset) eine grobe Vorstellung über die Gesamtszene zu bekommen. Allerdings kann man die zu diesem Zeitpunkt erworbene Gesamtvorstellung noch nicht mit der semantischen Struktur gleichsetzen, da die Richtigkeit des Benennens der Ereignisteilnehmer zu diesem Zeitpunkt noch sehr niedrig war. Eine große Performanzsteigerung bei der Identifizierung der Ereignisteilnehmer war zwischen der experimentellen Bedingung mit 100 ms und 200 ms Präsentationsdauer zu erkennen, was darauf hinweist, dass in diesem Zeitraum die grobe Vorstellung über das Ereignis in eine relativ eindeutige semantische Repräsentation umgesetzt wird. Zudem ist an dieser Stelle außerdem relevant, dass ab einer Präsentationsdauer von 200 ms eine beträchtliche Divergenz zwischen der Proportion der richtig identifizierten Agentia und Patientia besteht. Dabei wurde das Agens deutlich

häufiger korrekt erkannt als alle anderen semantischen Rollen, was dafür spricht, dass unter allen semantischen Rollen das Agens von Sprechern und Beobachtern bevorzugt und folglich auch früher identifiziert wird.⁵

Abschließend betrachtet haben die eben erwähnten Studien gezeigt, dass Sprecher in der Lage sind, in Szenenbeschreibungen ein Ereignis innerhalb kürzester Zeit holistisch zu interpretieren. Die Extraktion der semantischen Relationen bzw. die Zuweisung von semantischen Rollen vollzieht sich nicht unstrukturiert, sondern nach einer Reihenfolge, die von der Salienz einzelner semantischer Rollen beeinflusst wird. Dabei weist das Agens ein durchgängiges Verarbeitungsprivileg sowohl bei der konzeptuellen als auch bei der grammatischen Enkodierung auf. In Bezug auf die konzeptuelle Verarbeitung wird das Agens im Vergleich zu anderen semantischen Rollen in der Regel am schnellsten und frühesten identifiziert und hinsichtlich der grammatischen Enkodierung eher als Subjekt realisiert bzw. satzinitial angeordnet.

4.4 Kontextuelle Salienz und Satzproduktion

Bei der Satzplanung müssen Sprecher nicht nur die semantische und syntaktische Struktur eines Satzes aufbauen, sondern auch die zu vermittelnden Informationen in einer bestimmten Art und Weise „arrangieren“ bzw. reihen oder anordnen, damit sie den kontextuellen Bedürfnissen und Anforderungen in optimaler Weise entsprechen. Dieser Prozess wird von Chafe (1976) als das Verpacken von Information (engl.: *packaging of information*) bezeichnet und bezieht sich vor allem auf die Perspektivierung der Satzplanung - wie beispielsweise folgende Sätze: *Die Katze jagt einen Hund* oder *Einen Hund jagt die Katze*. Hinsichtlich der semantischen und syntaktischen Funktion sind die beiden Sätze identisch: Das Agens mit verknüpften Subjektstatus ist die *Katze* und dem *Hund* kommt die Patiensrolle zu, die syntaktisch durch das Objekt realisiert wurde. Darf man jedoch die beiden Beispielsätze

⁵ Das Ergebnis, dass das Agens im Gegensatz zum Patiens häufiger korrekt erkannt wurde, kann nicht auf die visuelle Salienz der Agens- und Patiens-Referenten zurückzuführen sein, da die Größe und andere visuelle Eigenschaften der beiden Referenten in dieser Studie systematisch manipuliert wurden.

miteinander gleichsetzen? Dieser Frage kann durch die Betrachtung des Kontextes nachgegangen werden. Wenn man nun einen Kontext mit der Frage *Was macht die Katze da?* bildet, dann würde der zweite Satz als weniger passende Antwort betrachtet werden. Das Beispiel zeigt, dass neben den semantischen und syntaktischen Eigenschaften einzelner Ereignisteilnehmer die Selektion einer passenden Satzstruktur ebenfalls von Kontexten beeinflusst wird.

Ein häufig verwendetes Mittel, um kontextuelle Eigenschaften abzubilden, ist die Nutzung von definiten und indefiniten Artikeln. Während definite Artikel gegebene (bzw. bereits vorerwähnte) Referenten wiederaufnehmen, werden unbekannte Referenten meist durch Nominalphrasen mit indefiniten Artikeln eingeführt. So lässt sich aus den obigen Beispielen ableiten, dass der Referent *Katze* mit dem definiten Artikel *die* wahrscheinlich in vorigen Äußerungen vorgegeben wurde, während der Referent *Hund* mit dem indefiniten Artikel *ein* als neue Information in die aktuelle Äußerung eingeführt wurde. Darüber hinaus kann der Referent *Hund* in dem zweiten Satz auch als Satztopik betrachtet werden, dem ein Prädikat (*jagt die Katze*) als Kommentar zugeschrieben wurde, während im ersten Satz das Satztopik *Katze* ist und die Entität *Hund* dem Kommentar zugeordnet wurde. Die Zweiteilung zwischen neuer und gegebener Information und zwischen Topik und Kommentar spiegelt den Informationsstatus bzw. die kontextuelle Salienz und führt zu unterschiedlichen Informationsstrukturen. Im Folgenden werden die wichtigsten Ergebnisse der Effekte, die jeweils von der Gegebenheit und der Topikalität ausgehen, im theoretischen Rahmen der kontextuellen Salienz vorgestellt. Dabei wird besonders die Wirkungsweise dieser beiden Faktoren während des Satzproduktionsprozesses beleuchtet.

4.4.1 Gegebenheit und Satzproduktion

Diskursreferenten können entweder als neue Informationen in den Diskurs eingeführt werden oder aufgrund von Vorerwähnungen, thematischen Beziehungen oder dem Weltwissen erschließbar sein und daher als gegeben bzw. bekannt

betrachtet werden (Musan, 2010). Linguistisch gesehen kann die Aufteilung zwischen gegebener und neuer Information etwa durch Mittel wie die Akzentuierung, Pronominalisierung, die Wahl zwischen indefinitem und definitem Artikel oder die Wortfolge erfolgen. In Bezug auf die Wortfolge stehen Elemente, die im Rahmen einer Aussage als *gegeben* angesehen werden, gewöhnlich am Anfang der Aussage, während neue oder unbekannte Elemente üblicherweise am Ende der Aussage stehen (Halliday, 1967a; Gundel, 1988).

Im psycholinguistischen Bereich existieren auch zahlreiche Arbeiten, die die präferierte Abfolge *gegebene Information vor neuer Information* untermauern (z. B. Arnold et al., 2000; Bock, 1977; Bock & Irwin, 1980; Ferreira & Yoshita, 2003; Prat-Sala & Branigan, 2000). Beispielsweise hat die Studie von Bock und Irwin (1980) nahegelegt, dass die Gegebenheit einzelner Ereignisteilnehmer im Zusammenhang mit der syntaktischen Funktionszuweisung bzw. Zusammenstellung von Konstituenten steht. Jeder Durchgang dieses Experimentes gliederte sich in eine Lern- und eine Testphase. In der Lernphase hörten die Versuchspersonen eine randomisierte Liste von Fragen, durch die verschiedene Ereignisse dargestellt werden und ein Ereignisteilnehmer als gegebene Information eingeführt wird. Ein Beispiel dafür ist der Satz (6a), in dem ein Ereignisteilnehmer (*a horse*) als gegebener Referent in den Diskurs eingeführt wurde. Dazu hörten die Versuchspersonen ebenfalls eine randomisierte Liste von Antworten, mit der die gestellten Fragen beantwortet werden konnten (6b und 6c).

(6a) *A rancher had a horse who kept running away. What did the rancher do?*

(6b) *The rancher sold the horse to the cowboy.*

(6c) *The rancher sold the cowboy the horse.*

(6d) *The rancher sold the stallion to the cowboy*

In der anschließenden Testphase wurden die Fragen noch einmal gestellt und die Probanden aufgefordert, die Fragen mit den Antworten, die sie sich in der Lernphase

bereits eingeprägt hatten, nun schriftlich zu beantworten. Die Autoren stellten fest, dass die Probanden dazu tendierten, Antworten mit syntaktischen Strukturen zu produzieren, in denen gegebene (bzw. bereits vorerwähnte) Wörter früher auftraten. Da durch die Frage (6a) ein Referent (*horse*) in den Diskurs eingeführt wurde, benutzten die Probanden mit höherer Wahrscheinlichkeit eine syntaktische Struktur, die eine Voranstellung dieses Referenten zulässt, nämlich Satz (6b). Der Befund, dass die im Kontext früher vorgegebenen Informationen eher vor neuen Informationen verzeichnet wurden, liefert einen Beweis für den Einfluss der Gegebenheit auf die Linearisierung lexikalischer Konstituenten. Dennoch stellt sich die Frage, woher dieser Einfluss rührt.

Es wurde von den beiden Autoren angenommen, dass das sog. *given-new ordering* in einem engen Zusammenhang mit der lexikalischen Zugänglichkeit steht. Da die gegebene Entität in der vorangehenden Frage mit einem bestimmten Worteintrag zum Ausdruck gebracht wurde, ist dieser Worteintrag zugänglich im Arbeitsgedächtnis behalten. Wenn Sprecher zu einem späteren Zeitpunkt im Gespräch denselben Referenten wieder aufnehmen wollen, wird dieser Worteintrag wegen der erhöhten Zugänglichkeit, die durch die Vorerwähnung ausgelöst wurde, schneller aus dem Arbeitsgedächtnis abgerufen und in einer früheren Position des zu produzierenden Satzes platziert. Aber resultiert der Einfluss der Gegebenheit lediglich aus der erhöhten lexikalischen Zugänglichkeit oder wird die konzeptuelle Zugänglichkeit des erneut referierten Konzepts auch durch die Vorerwähnung gesteigert?

Um eben diese Frage zu beantworten wurde ein ähnliches Experiment durchgeführt, jedoch mit dem einzigen Unterschied, dass auf die gegebene Entität in der Frage- und Antwortliste mit zwei verschiedenen Wörtern Bezug genommen wurde (z. B. *horse* in der Frage (6a) und *stallion* in der Antwort (6d)). Durch die Manipulation der lexikalischen Einstimmigkeit war es den Autoren möglich, die lexikalische und konzeptuelle Zugänglichkeit voneinander zu unterscheiden. Es zeigte sich, dass trotz der Nicht-Übereinstimmung der lexikalischen Einträge in Frage und in Antwort der

Effekt der Gegebenheit aufrechterhalten wurde, was von den Autoren als Hinweis dafür interpretiert wurde, dass die konzeptuelle Zugänglichkeit allein die Effekte der Gegebenheit auslösen kann. Zusammengefasst kann demnach gesagt werden, dass das sogenannte *given-new ordering* die bevorzugte Verarbeitungsreihenfolge der Satzplanung ist (vgl. auch Arnold et al., 2000; Bock, 1977; Bock & Irwin, 1980; Prat-Sala & Branigan, 2000; Ferreira & Yoshita, 2003). Aus der psycholinguistischen Sicht kann dieses Prinzip auf die durch die Vorerwähnung erhöhte lexikalische und konzeptuelle Zugänglichkeit zurückführen sein.

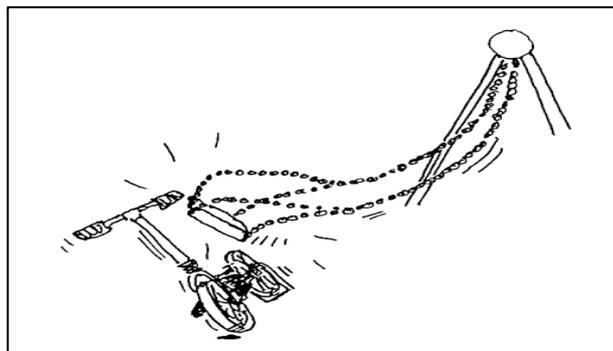


Abbildung 4.2: Beispielstimulus aus Prat-Sala und Branigan 2000.

Prat-Sala und Branigan (2000) zeigten in ihrer Studie, dass, auch wenn alle zu verbalisierenden Entitäten gegeben sind, die Selektion von syntaktischen Strukturen bzw. Wortstellungen von den Gegebenheitsgraden einzelner Entitäten beeinflusst wird. Das Experiment wurde wie folgt aufgebaut: Am Anfang jedes Versuchsdurchgangs wurde englischen und spanischen Muttersprachlern ein transitives Ereignis mit zwei unbelebten Entitäten präsentiert (z. B. *A swing hitting a scooter*; Beispiel siehe Abbildung 4.2); parallel dazu hörten sie einen sog. Einführungssatz, in dem das Agens und das Patiens in unterschiedlichem Maße beschrieben und daher unterschiedliche kontextuelle Gewichtungen bzw. Gegebenheitsgrade aufgezeigt wurden:

(7a) *There was this old rusty swing standing in a playground near a scooter, swaying and creaking in the wind. What happened?* – Agens mit höherem Gegebenheitsgrad

(7b) *There was this old red scooter standing in a playground near a swing, with rusty wheels and scratched paint. What happened?* – Patiens mit höherem Gegebenheitsgrad

Nach dem Abspielen des Einführungssatzes mussten die Versuchspersonen die Frage *What happened?* anhand des dargestellten Bildes beantworten. Es wurde beobachtet, ob der Gegebenheitsgrad einen Einfluss auf die syntaktische Strukturierung hat. Die Ergebnisse legen nahe, dass die selektierten syntaktischen Strukturen bzw. Wortstellungen nach der relativen Gegebenheit des Agens und des Patiens variierten: Wenn das Patiens gegebener bzw. kontextuell salienter war als das Agens, steigerte sich die Proportion der nicht-kanonischen Sätze (Passivsätze für beide Sprachen und Scrambled-Aktivsätze spezifisch fürs Spanische). Im Gegensatz dazu wurden deutlich mehr kanonische Sätze produziert, wenn das Agens hinsichtlich der kontextuellen Salienz dem Patiens überlegen war.

Resümierend lässt sich feststellen, dass die Abfolge *gegebene vor neuer Information* ein wichtiges Prinzip der Satzproduktion bzw. der syntaktischen Strukturierung ist. Dies könnte damit zu tun haben, dass gegebene Informationen im Gedächtnis voraktiviert und der konzeptuellen und lexikalischen Verarbeitung entsprechend schneller zugänglich sind.

4.4.2 Topikalität und Satzproduktion

Ein anderer Faktor, der ebenfalls einen Einfluss auf die syntaktische Strukturierung hat, ist die Topikalität. Der Topiktyp, der in der vorliegenden Arbeit behandelt wird, ist das Satztopik. Laut Hocket (1958) ist das Satztopik derjenige Gegenstand eines Satzes, über den eine Aussage gemacht wird (“what the sentence is about”; S. 201). Lambrecht (1994) geht von der kognitiven Ansicht aus und definiert das Topik auf der satzbezogenen Ebene wie folgt:

A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e. as expressing information which is relevant to and which increases the addressee’s knowledge of this referent.

(S. 131)

In Anlehnung dazu legt Krifka (2007) beim Definieren von Topiks besonderes Gewicht auf die zugrundeliegenden kognitiven Verarbeitungsprozesse. Hierbei wird die Rolle vom sog. Common Ground hervorgehoben: „The topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG [common ground] content“ (S. 41), wobei CG (Common Ground) das Wissen beinhaltet, das von Hörer und Sprecher geteilt wird. Dieses Wissen wird im Verlauf der Kommunikation immer wieder modifiziert. Dennoch gelingt die Modifizierung bzw. die Hinzufügung neuer Information zum Common Ground nicht ungeordnet, sondern geht immer mit einer bestimmten Entität, auf die sie sich bezieht, zusammen. Diese hier bezogene Entität ist das Topik.

Ähnlich wie semantische Rollen muss das Topik ebenfalls auf die grammatische Enkodierungsebene abgebildet werden. In Bezug auf das Verhältnis zwischen der Topikalität und der Wortstellung gibt es eine gewisse sprachspezifische Tendenz: Das Topik nimmt meist die Satzanfangsposition ein (vgl. Choi, 1996; Frey, 2004; Rosengren, 1993). Aus dem Gesichtspunkt der Informationsorganisation, die sich während der Satzproduktion vollzieht, ist dies nachvollziehbar, da es offenbar leichter ist, einen Satz zu bauen, wenn man bereits weiß, worauf sich der Satz bezieht. In Bezug auf das Verhältnis zwischen der Topikalität und der syntaktischen Funktion korreliert das Topik sehr oft – jedoch nicht immer – mit dem Subjekt des Satzes: Das Topik ist häufig das Subjekt eines Satzes. Diese Korrelation ist wohl auf die Bevorzugung des Subjekts und des Topiks für den Satzanfang zurückzuführen: Einerseits tendiert das Subjekt, das aufgrund höherer grammatischer Prominenz früher zugewiesen wird, dazu, die satzeinleitende Position zu besetzen; andererseits weist das Topik als das, worüber etwas ausgesagt wird, auch eine starke Tendenz dazu auf, am Anfang eines Satzes zu erscheinen. Allerdings ist zu beachten, dass, obwohl Topiks im Default-Fall als Subjekt realisiert werden, diese Verbindung in einigen Situationen unterbrochen werden kann. Beispiele hier sind das freie Topik wie *Was Jonas trifft, den hat Moritz gerufen* oder die Linksversetzung wie *Den*

Jonas, den hat Moritz gerufen, auf die in dieser Arbeit jedoch nicht weiter eingegangen wird.

Eine Frage, auf die sich die vorliegende Arbeit konzentriert, ist, in welcher Art und Weise die Topikalität die Satzproduktion beeinflusst. Da Topiks im Diskurs bereits vorher eingeführt wurden, nehmen Gesprächsteilnehmer sie daher als bereits “gegeben” wahr. Deswegen ist anzunehmen, dass ihre Effekte zum Teil mit den Effekten der Gegebenheit gleichzusetzen sind. Die zentrale Frage, die daraus resultiert, ist nun, ob die Topikalität im Gegensatz zur Gegebenheit noch einen zusätzlichen Effekt auf die Satzplanung hat. So ein Effekt unterscheidet sich nämlich von der erhöhten konzeptuellen und lexikalischen Zugänglichkeit aufgrund von Vorerwähnung.

Diesen Fragen gingen Cowles und Ferreira (2012) nach. Im Fokus ihrer Studie steht der Vergleich zwischen dem Effekt der Gegebenheit und der Topikalität auf die Wortanordnung. Hierzu wurden zwei Bedingungen, eine Gegebenheits- und eine Topikalitätsbedingung, durch den Einsatz verschiedener Einführungssätze und somit den Aufbau verschiedener Kontexte erstellt. In der Gegebenheitsbedingung hörten die Versuchspersonen zuerst einen Einführungssatz wie *The passenger realized something when he saw the traffic*. In der Topikalitätsbedingung hörten sie einen Satz wie *The passenger realized something about the traffic* als Einführung. Durch die beiden Sätze wurde für die Probanden ein Common Ground geschaffen, in dem *traffic* nicht nur gegeben ist, sondern auch mithilfe der *about*-Phrase in der Topikalitätsbedingung als Topik-Kandidat in den Fokus der kommenden Erzählung rücken. Nach der Präsentation eines Einführungssatzes wird das Wort am Satzende (hier *traffic*) in den beiden Bedingungen zusammen mit einem Verb (z. B. *annoyed*) und einem anderen Nomen (wie z. B. *driver*) dargestellt. Die Versuchspersonen wurden aufgefordert, mithilfe der gegebenen Wörter die von dem Einführungssatz initiierte Geschichte weiter zu erzählen. Es wurde gemessen, wie häufig die in Einführungssätzen aufgezeigten Wörter (z. B. *traffic*) satzinitial angeordnet wurden.

Die Ergebnisse zeigten, dass in den beiden genannten Bedingungen die durch die vorangegangenen Einführungssätze eingeführten Wörter überzufällig häufig am Satzanfang positioniert und als erste Referenten versprachlicht wurden. Es ist wichtig anzumerken, dass diese Tendenz in der Topikalitätsbedingung stärker war, in der die durch die Einführungssätze eingeführten Referenten nicht nur als gegebene Informationen, sondern auch als Topiks weiterer Erzählungen eingestuft wurden. Die daraus resultierende Schlussfolgerung der Autoren war, dass, zum einen, die Effekte der Gegebenheit und die der Topikalität teilweise überschneidend sind, und dass zum anderen, die Topikalität noch einen zusätzlichen Einfluss auf die syntaktische Strukturierung hat.

Die Forscher gehen daher von der Annahme aus, dass dem Topikalitäts-Effekt - neben dem darin eingebauten Gegebenheitseffekt, der auf die erhöhte lexikalische und konzeptuelle Zugänglichkeit des topikalischen Referenten zurückzuführen ist - auch ein anderer Mechanismus zugrundeliegen muss. Diesen Mechanismus nennen die Autoren *+topic feature*. Allerdings wurden in der Studie keine eindeutigen Erkenntnisse über die Funktionsweise des sogenannten *+topic feature* geliefert.

Zur Erklärung des Einflusses, der vom Topikstatus ausgelöst wurde, schlugen Gleitman et al. (1996; 2007) eine aufmerksamkeitsbasierte Hypothese vor, nach der Elemente mit Topikstatus durch die Lenkung der Aufmerksamkeit zum Zeitpunkt der Rede in den Vordergrund gestellt werden. Diese Erklärung rührt vor allem von Beobachtungen über das Verhältnis zwischen der visuellen Aufmerksamkeit und der Figur-Grund-Wahrnehmung (engl.: *figure ground perception*). Die Figur-Grund-Wahrnehmung bezeichnet die menschliche Fähigkeit, ein bestimmtes visuelles Element aus seinem Hintergrund- und Nebenreizen herauszulösen. Dabei wird die Aufmerksamkeit ständig auf die wichtigen Reize gerichtet. Unauffällige Reize werden übersehen oder weniger deutlich wahrgenommen. „Die ausgewählten Reize bilden die Figur innerhalb unseres Wahrnehmungsfeldes, die anderen - unwichtigen Reize - bilden den nur ungenau wahrgenommenen Hintergrund“ (Zimmer 2005, S. 69). Ein Beispiel für die Figur-Grund-Wahrnehmung

ist, dass Menschen in der Lage sind, ein Objekt aus einer großen Objektmenge zu entdecken, selbst wenn sich viele Objekte in unmittelbarer Nähe befinden.

Zwar steht die im visuellen Kontext stattfindende Figur-Grund-Unterscheidung nicht in einer direkten Verbindung mit der Sprachproduktion, jedoch wurde sie von manchen Forschern in die Linguistik eingeführt, um damit sprachliche Phänomene erklären zu können (Talmy, 1975, 1978, 2000 Kap. 1; Thiering, 2011). Beispielsweise wurde dieses Begriffspaar von Talmy (1975) als eine kognitiv-semantische Kategorie zur Beschreibung der räumlichen Lokalisation genutzt. In diesem Zusammenhang ist die Figur aus der semantischen Sicht „a moving or conceptually movable point whose path or site is conceived as a variable the particular value of which is the salient issue“, demgegenüber ist der Grund „a reference-point, having a stationary setting within a reference-frame, with respect to which the FIGURE’s path or site receives characterization“ (S. 419). Vereinfacht gesagt ist die Figur meist die zu lokalisierende Entität (Referent) und der Grund der Bezugssachverhalt (Relatum). Weitere Beispiele für die Figur-Grund-Unterscheidung im linguistischen Kontext sind Sätze wie folgt:

(8a) *Die Sparkasse* [Figur] *ist neben der Deutschen Bank* [Grund].

(8b) *Die Deutsche Bank* [Figur] *ist neben der Sparkasse* [Grund].

Die beiden Beispielsätze sind wahrheitswertfunktional äquivalent. Dabei handelt es sich nur um die inversen Formen einer symmetrischen Relation. Laut Talmy liegt in diesen Sätzen in der Tat eine Asymmetrie vor. Dies zeigt sich beispielsweise darin, dass nur (8a) als passende Antwort zur Beantwortung von Fragen wie *Wo ist die Sparkasse* infrage kommt. (8b) wird hierbei nicht als angemessen angesehen, da die beiden semantischen Rollen (Figur vs. Grund) den falschen Entitäten zugewiesen wurden.

Den Umstand, dass der Unterschied zwischen Figur und Grund als semantische Rollenunterschiede gekennzeichnet wird, halten Dowty (1995) und Dürscheid (2012)

jedoch als nicht angemessen. Sie gehen eher von dem kontextuellen Gesichtspunkt aus und nehmen an, dass die Figur-Grund-Unterscheidung informationsstrukturell charakterisiert und zur Beschreibung der kommunikativen bzw. informationsstrukturellen Gewichtung verwendet werden soll. So vertritt Dürscheid (2012) die Auffassung, dass die Topik-Kommentar-Strukturierung, die auf der Satzebene vorgenommen wird, im Wesentlichen eine Instanz der Figur-Grund-Anordnung ist. Satz (8a) wird als Antwort zur Frage *Wo ist die Sparkasse?* aufgrund der irrtümlichen Zuteilung des Topikstatus und nicht wegen der falschen semantischen Rollenzuteilung von *Sparkasse* als nicht geeignet angesehen.

Die Ansicht von Dowty und Dürscheid über die Figur-Grund-Strukturierung teile ich und sehe sie eher als kontextuelle Strukturierung, da es im Kern um die Perspektivierung geht. Dazu gehört – wie von Dürscheid angenommen – auch die Topik-Kommentar-Strukturierung. Unter der Annahme, dass sich die Figur-Grund-Unterscheidung im visuellen Kontext durch die Lenkung der Aufmerksamkeit erklären lässt, ist zu vermuten, dass die Aufteilung zwischen Topiks und Kommentaren im linguistischen Kontext einer ähnlichen Funktionsweise folgen kann (vgl. auch Wagner, 2016). Auf diese Hypothese wird in Experiment 2 der vorliegenden Studie näher eingegangen.

In einem weiteren Experiment der Studie von Cowles und Ferreira (2012) versuchten die Forscher zu identifizieren, auf welcher grammatischen Enkodierungsebene (der funktionalen vs. der positionellen Enkodierungsebene) sich die Topikalität auswirkt. Hierfür wurde das Experiment wie folgt aufgebaut: In jedem Durchgang wurden Versuchspersonen zuerst drei Wörter (z. B. *passenger*, *traffic* und *driver*) präsentiert, die sie sich einprägen sollten. Nach der Präsentation der Wörter mussten die Versuchspersonen dann eine zur Ablenkung dienende Aufgabe erledigen, bevor sie einen Einführungssatz hörten. Die abgespielten Sätze waren genau dieselben, die in dem vorigen Experiment benutzt wurden. Dementsprechend wurden ebenso eine Gegebenheits- und eine Topikalitätsbedingung gestellt. Als Aufgabe sollten die Versuchspersonen nach dem Abspielen der Sätze nach dem Bildschirmhinweis

entweder die präsentierten Wörter (kritische Aufgabe) oder die gehörten Sätze wiederholen (Filler-Aufgabe). Von zentralem Interesse hierbei war, zu beobachten, ob ähnlich wie im vorigen Experiment in der Topikalitätsbedingung auch ein stärkerer Einfluss der Topikalität auf die Wortanordnung beim Erledigen der kritischen Aufgabe, bei der keine syntaktische Funktionszuweisung gebraucht wurde, eintritt.

Es zeigte sich, dass sich die Häufigkeit, die vorgegebenen Wörter zuerst wiederzugeben, in beiden Bedingungen nicht unterschied. In der Topikalitätsbedingung wurde das Wort, das sowohl gegeben als auch mit dem Topikstatus versehen war, im Gegensatz zur Gegebenheitsbedingung, unter der dasselbe Wort lediglich gegeben war, nicht öfter als Erstes erwähnt. Das bedeutet, dass kein zusätzlicher Topikalitäts-Einfluss auf das Wiedergeben präsentierter Wörter beobachtet wurde. Da der Unterschied zwischen der Aufgabe dieses Experimentes und der des vorigen nur darin besteht, dass bei der Wiederholung dargestellter Wörter keine syntaktischen Funktionszuweisungen benötigt werden, spricht das Fehlen des additiven topikalischen Einflusses daher dafür, dass die Auswirkung der Topikalität nicht auf der positionellen, sondern auf der funktionalen Verarbeitungsebene zum Tragen kommt (vgl. auch Wagner, 2016).

Resümierend lässt sich der Topikalitäts-Effekt in zwei Kategorien einteilen: Zum einen überschneidet er sich mit dem Effekt der Gegebenheit und kann daher auf die konzeptuelle und lexikalische Zugänglichkeit zurückgeführt werden; zum anderen übt das Topik auch einen speziellen Einfluss aus, dessen Mechanismus bislang allerdings noch vage ist. Eine Hypothese dafür ist, dass die Auswirkung der Topikalität mit der Beeinflussung der Aufmerksamkeit korreliert (vgl. Gleitman, 1996, 2007). Des Weiteren wurde auch herausgefunden, dass die Topikalität keinen direkten Einfluss auf die positionelle Verarbeitungsebene hat und sich auf der syntaktischen Funktionszuweisungsebene auswirkt. Diese Aussage wird ebenfalls in der vorliegenden Arbeit überprüft.

4.5 Fazit

Im Zusammenhang mit der Frage nach der Inkrementalität in Kapitel 3 wurden in diesem Kapitel drei Salienz-Typen vorgestellt, die sowohl die konzeptuelle als auch die grammatische Verarbeitung beeinflussen können – die visuelle, die semantische und die kontextuelle Salienz. Viele Studien haben gezeigt, dass alle drei Salienz-Formen einen Einfluss auf die Zugänglichkeit einzelner Konzepte haben, die sich als ausschlaggebend für die Selektion von *starting points* erwiesen hat. Informationen, die höhere Zugänglichkeiten aufweisen, werden schneller aus dem Langzeitgedächtnis abgerufen und früher an die linguistische Verarbeitung weitergereicht als diejenigen, die mental weniger zugänglich sind. Diese Verarbeitungsweise ist für die Satzplanung von großer Bedeutung, da Sprecher dadurch reduzierter kognitiver Belastung ausgesetzt sind und zugleich mehr Zeit für die Verarbeitung der weniger zugänglichen Entitäten gewinnen können.

In Bezug auf das Verhältnis zwischen der visuellen Salienz und der Satzproduktion wird angenommen, dass Entitäten, die auffälligere visuelle Eigenschaften aufweisen, früher wahrgenommen und identifiziert werden, was dann zur Selektion derselben Entitäten als erstes Verarbeitungsinkrement führen kann. In Bezug auf die semantische Salienz wird davon ausgegangen, dass die Selektion von *starting points* bzw. die syntaktische Strukturierung auf der semantischen bzw. konzeptuellen Struktur des zu beschreibenden Ereignisses basiert. Entitäten, die eine stärkere semantische Salienz aufweisen (z. B. das Agens), werden mit höherer Wahrscheinlichkeit als *starting points* ausgewählt als die mit schwächerer Salienz (z. B. das Patiens). Diese Aussage basiert vor allem auf zwei Beobachtungen: Erstens, während der semantischen Strukturierung eines Ereignisses wird das Agens in der Regel früher aufgerufen als andere thematische Rollen; zweitens, viele Studien haben nahegelegt, dass das Agens einen Verarbeitungsvorteil hat und im Gegensatz zum Patiens häufiger als der erste Referent verbalisiert wird. Eine ähnliche Annahme kann von der kontextuellen Salienz vertreten werden, die sich aus informationsstrukturellen Eigenschaften ergibt. Hierbei sind zwei Elemente von

großer Bedeutung, die Gegebenheit und die Topikalität. Es wurde herausgefunden, dass Referenten, die in der vorigen Äußerung erwähnt wurden und somit gegeben sind, eher vor neuen Informationen platziert werden. Dies ist vor allem auf die Steigerung der konzeptuellen und lexikalischen Zugänglichkeit zurückzuführen. Hinsichtlich der Topikalität wird nicht nur ein mit der Gegebenheit überschneidender, sondern auch ein topikalisch spezifischer Effekt beobachtet. Allerdings ist bis hierhin noch offengeblieben, wie das Topik kognitiv repräsentiert wird und in welcher Art und Weise es sich in dem Satzproduktionsprozess auswirkt.

5 Untersuchungsfragen und Versuchsdesign

5.1 Untersuchungsfragen

Die vorliegende Arbeit geht in erster Linie von dem berühmten *starting points problem* aus (MacWhinney, 1977). Dabei handelt es sich vor allem darum, wie *starting points* der Verarbeitung während des Satzproduktionsprozesses selektiert werden und in welchem Maß das Sprechen in der konzeptuellen Vorausplanungsphase, also vor Beginn der grammatischen Enkodierung, geplant wird.

Zu dieser Frage wurden in vorangehend dargestellten Studien zwei polarisierende Annahmen aufgestellt – die lineare und die strukturelle Inkrementalitätshypothese. In Bezug auf die lineare Inkrementalitätshypothese wird eine auf einer reizbedingten Funktionsweise gründender Arbeitsweise des Satzproduktionssystems angenommen, der einen geringen Umfang der Vorausplanung auf der konzeptuellen Verarbeitungsebene hypothetisiert. Demnach werden Entitäten, deren konzeptuelle oder visuelle Salienz am höchsten ist, früher wahrgenommen und deswegen eher als *starting points* ausgewählt. Die frühere Identifizierung dieser Entitäten führt wiederum zur früheren syntaktischen Enkodierung dieser Entitäten (Gleitman et al., 2007). Nach diesem Ansatz ist es für Sprecher daher nicht nötig, vor der Selektion von *starting points* zuerst eine holistische Interpretation des Gesamtereignisses aufzubauen. Sobald ein Referent aufgrund seiner höheren Zugänglichkeit als das erste Konzept der Verarbeitung ausgewählt wurde, beginnt sofort dessen grammatische Enkodierung. Das bedeutet, dass sich der Umfang der Vorausplanung lediglich auf ein individuelles Konzept beschränken kann. Im Gegensatz dazu geht die strukturelle Inkrementalitätshypothese davon aus, dass das Auswählen von *starting points* nicht auf Eigenschaften einzelner Konzepte (wie etwa die konzeptuelle oder die visuelle Salienz) zurückzuführen ist, sondern einem Vorgang unterliegt, durch welchen eine Gesamtvorstellung über das zu beschreibende

Ereignis generiert wird. Die Grundidee besteht darin, dass Sprecher zuerst ein Ereignis grob als Ganzes verstanden haben müssen, bevor sie dann entscheiden können, mit welchem Konzept sie die Formulierung beginnen. Demnach ist die Vorausplanung mit dem Aufbau einer strukturellen Ereignisrepräsentation viel größer als ein einziges Konzept. Neben diesen beiden polarisierenden Annahmen gibt es noch eine weitere dritte Hypothese, nach der sich die lineare und die strukturelle Inkrementalitätshypothese gegenseitig nicht ausschließen, sondern im Gegenteil gemeinsam eine Erklärung für die Arbeitsweise des Satzplanungssystems sein können.

Von diesen Hypothesen ausgehend, beschäftige ich mich in der vorliegenden Studie zuerst mit der Frage, ob die Satzproduktion im Deutschen in einer pragmatisch neutralen Situation linear oder strukturell inkrementell verläuft. Dafür soll vor allem beobachtet werden, welcher Faktor eine dominante Rolle beim Auswählen von *starting points* spielt (visuelle oder semantische Salienz) und in welchem Ausmaß ein Satz in der Vorausplanungsphase geplant wird (ein einziges Konzept oder eine strukturelle Repräsentation). Auf diese Fragen wird in Experiment 1 eingegangen.

Die zweite Frage der vorliegenden Arbeit betrifft das Verhältnis zwischen der Satzproduktion und dem situativen bzw. kommunikativen Kontext. Die Untersuchung der kontextuellen Salienz im Kontext der Satzplanung ist zur Bereicherung der bereits vorhandenen Satzproduktionsmodelle wichtig, da die bisherigen Modelle der Satzproduktion der Rolle der kontextuellen Information nur eine vergleichsweise geringe Beachtung geschenkt haben. Außerdem ist sie für die Untersuchung der Inkrementalitätstheorie auch von großer Bedeutung. Bei der Überprüfung der strukturellen Inkrementalitätshypothese wurde in früheren Studien nur der gemessene semantische Effekt als zuverlässiger Indikator angesehen. Dennoch beschränkt sich die strukturelle Repräsentation nicht nur auf die semantische Ebene. Denn neben der semantischen Struktur bauen Sprecher während der Mikroplanung auch Informationsstrukturen auf. Die Einbeziehung der kontextuellen Salienz erlaubt es, Theorien zur inkrementellen Verarbeitung zu

vervollständigen bzw. sie zu ergänzen und somit einen neuen Maßstab zur Unterscheidung der linearen und strukturellen Inkrementalität zu setzen.

Hierbei steht in erster Linie die Rolle der kontextuellen Salienz während der Satzplanung im Mittelpunkt. Obwohl Psycholinguisten einen Einfluss der kontextuellen Salienz auf die Satzproduktion als selbstverständlich ansehen, wurde dieser Einfluss in früheren Studien nur begrenzt experimentell belegt. Es ist bis dato noch wenig darüber bekannt, wie sich kontextuelle Informationen im Satzplanungsprozess auswirken (vgl. Arnold, Kaiser, Kahn & Kim, 2013). Auch das bekannte Satzproduktionsmodell von Bock und Levelt (1994) beleuchtet diese Frage nicht näher.

Im Rahmen der kontextuellen Salienzhypothese werden zwischen zwei Prinzipien unterschieden: Zum einen die Gegebenheit, die den Informationsstand des Adressanten betrifft und zum anderen die Topikalität, die auf dem augenblicklichen Gesprächsinteresse gründet (vgl. Halliday, 1967a). Der Effekt der Gegebenheit zeigt sich vor allem in dem sog. *given-new ordering*, nach der gegebene Informationen generell häufiger neuen Informationen vorangestellt werden. Eine Evidenz für dieses Prinzip liefert die Studie von Prat-Sala und Branigan (2000), in der eine Abhängigkeit der Wortabfolge auf Gegebenheitsgrade der an einem Ereignis beteiligten Figuren festgestellt wurde: Figuren, deren Gegebenheitsgrade als relativ hoch wahrgenommen wurden, traten häufiger vor denjenigen auf, deren Grade weniger hoch eingestuft wurden. Ein Grund dafür könnte nach Bock und Irwin (1980) die erhöhte lexikalische und konzeptuelle Zugänglichkeit der gegebenen Information sein.

Im Gegensatz zur Gegebenheit beinhaltet das Konzept der Topikalität komplexere Verarbeitungsmechanismen. Topiks sind Gesprächsteilnehmern in der Regel bereits bekannt, denn Sprecher tendieren dazu, über Sachverhalte zu sprechen, die in der vorangehenden Äußerung oder in ihren Vorstellungen bereits bestehen. Diese Eigenschaft erlaubt es Sprechern daher, den topikalischen Effekt zum Teil mit dem

Effekt der Gegebenheit gleichzusetzen, welcher auf die erhöhte konzeptuelle und lexikalische Zugänglichkeit angewiesen ist (Bottom-Up-Effekt; Cowles & Ferreira, 2012). Andererseits ist aber anzumerken, dass die Topikalität auch in einem engen Zusammenhang mit den augenblicklichen Gesprächsinteressen steht, welche wenig mit der Gegebenheit und mehr mit der Selektion verschiedener Erzählperspektiven zu tun haben. Daher kann es durchaus passieren, dass ein Topik-Ausdruck trotz niedrigem Gegebenheitsgrad an den Satzanfang platziert wird – auch wenn eine andere Entität einen höheren Gegebenheitsgrad aufweist, jedoch nicht im Fokus der momentanen Rede steht. Diesem einzigartigen Effekt der Topikalität (Top-Down-Effekt) wird in meiner Studie viel Aufmerksamkeit gewidmet.

Die Studie von Cowles und Ferreira (2012) ist eine der wenigen, die sich mit dem Verhältnis zwischen der Topikalität und der Satzproduktion auseinandergesetzt haben (vgl. Abschnitt 4.4.2). Diese Studie fand aus, dass Sprecher dazu neigen, Referenten mit Topikstatus am Satzanfang zu platzieren. Um diesen Befund zu interpretieren schlagen die Autoren das Konzept des sog. „+topic feature“ vor. Jedoch geben sie auch zu, dass ihre Arbeit keine eindeutige Erklärung für die Substanz und die Funktionsweise des „+topic feature“ liefern könne. Daraus ergibt sich eine zentrale Frage, die in der vorliegenden Untersuchung beantwortet werden soll: Wie werden Topiks während der Satzproduktion repräsentiert und wie wirkt sich die Topikalität auf die konzeptuelle und die syntaktische Verarbeitung aus.

Um diese Fragen beantworten zu können, wird zuerst der Einfluss der Topikalität auf die syntaktisch strukturelle Selektion getestet. Damit soll herausgefunden werden, ob der Effekt, den Cowles und Ferreira (2012) in ihrer Studie im Englischen finden konnten, auch für das Deutsche zutrifft. Es wird ebenso untersucht, welchen Einfluss die Bedeutung der Topikalität bei der Planung der frühen Augenbewegungen hat. Denn dadurch könnte eine eindeutige Erklärung für die Repräsentation und Verarbeitungsmechanismen der Topikalität während der Sprachplanung gefunden werden.

Eine andere Frage, der ebenfalls in dieser Studie nachgegangen wird, betrifft das Mapping des Topiks auf die Syntax. Alle relevanten Strukturen und Referenten, egal auf welcher Satzproduktionsebene sie erzeugt werden und welche Inhalte sie umfassen, werden in einer linearen Wortsequenz angeordnet. Diese Regel gilt auch für die Verarbeitung kontextueller Informationen. Das wirft daher die Frage auf, in welcher Vorgehensweise Informationsstruktur, die ähnlich wie thematische Struktur ebenfalls auf der Mikroplanungsebene geplant wird, in die Wortstellung eingefügt und gereiht werden. Diese Forschungsfrage kann auch in zwei weitere Fragen unterteilt werden: (1) Müssen kontextuelle Informationen wie das Topik zuerst semantisch enkodiert werden, was auf eine Interaktion zwischen der kontextuell und der semantisch strukturellen Information innerhalb des Mikroplanungssystems hinweisen würde? (2) Müssen kontextuelle Informationen auf die funktionale Enkodierungsebene abgebildet werden, bevor sie zum Schluss in die Wortstellung eingefügt werden?

Hinsichtlich der Relation zwischen der kontextuellen Salienz und der Satzplanung im Deutschen sollen nun in Experiment 2 und 3 folgende wichtige Fragen geklärt werden:

1. Spielt die kontextuelle Salienz bzw. die Topikalität bei der Satzplanung im Deutschen eine Rolle?
2. Wie werden Topiks während der Satzplanung repräsentiert und wie wirkt sich die Topikalität auf die Satzplanung aus?
3. In welcher Art und Weise werden kontextuelle Informationen bzw. das Topik auf die Wortstellung abgebildet?

Des Weiteren beleuchtet die vorliegende Studie das Zusammenspiel der visuellen, semantischen und kontextuellen Salienz, die im Satzproduktionsprozess zusammenwirken. In früheren Studien wurden solche Faktoren nur isoliert untersucht, d.h., um den Einfluss eines bestimmten Faktors zu ermitteln, werden alle

anderen Faktoren bewusst kontrolliert, damit sie nicht als störende Faktoren hervortreten. Allerdings kann das Sprechen dadurch nur sehr schlecht simuliert werden, da die Satzproduktionsumwelt im Alltagsleben nicht so kontrolliert ist wie in einer experimentellen Situation. Alle Faktoren, die die Satzplanung beeinflussen, können in der realen Sprachproduktionsumwelt miteinander interagieren, was wiederum zum Schluss eine sehr komplexe Situation erzeugen kann. Im Idealfall können die maßgeblichen Faktoren zusammentreffen (z. B. ist das Agens sowohl visuell als auch kontextuell am salientesten), wodurch die Satzplanung sehr vereinfacht wird. Allerdings ist dies nicht immer der Fall. In den meisten Situationen korrelieren die verschiedenen Salienz-Typen auch mit unterschiedlichen Referenten. Doch wie lösen Sprecher das Problem, trotz einer komplexen Satzbildungssituation einen fließenden Sprechvorgang sicherstellen zu können? Genauer formuliert: Verfolgen Sprecher in so einer Situation zur Verarbeitung einzelner Referenten eine bestimmte Reihenfolge oder verläuft die Enkodierung unstrukturiert?

Eine Möglichkeit ist, dass die Verarbeitungsreihenfolge der Referenten wohl in einem Zusammenhang mit der Stärke des Einflusses der Eigenschaften, die von den Referenten mitgebracht werden, steht. Obwohl alle Faktoren einen Einfluss auf die Satzplanung haben, variiert die Einflussstärke. Die Faktoren, die für die Satzplanung relevanter sind, könnten während der Verarbeitung im Gegensatz zu denjenigen, die weniger wichtig sind, früher berücksichtigt werden. Somit wird eine sog. Prioritätenliste der wirkenden Faktoren erzeugt, die auf den Einflussstärken der jeweiligen Faktoren beruht und als Leitfaden zur Lösung komplexerer Situationen dienen kann. Da in der vorliegenden Studie die Einflüsse der visuellen, semantischen und kontextuellen Salienz überprüft werden, ist dabei auch festzulegen, an welcher Position sich diese Faktoren in der Prioritätenliste befinden.

5.2 Entwicklung eines Versuchsdesigns

Das Anliegen dieser Studie ist, zum einen zu untersuchen, wie und in welcher Art

und Weise die syntaktische Enkodierung im Deutschen gelingt. Dabei stehen zwei unterschiedliche Ansätze im Mittelpunkt: die lineare und die strukturelle Inkrementalitätshypothese. In Bezug auf die lineare Inkrementalität wird bei der Satzproduktion die Rolle der Eigenschaft von Ereignisteilnehmern hervorgehoben. Die Eigenschaft bezieht sich in der vorliegenden Dissertation vor allem auf die visuelle Salienz. Im Vergleich dazu geht die strukturelle Inkrementalitätshypothese von der semantischen Salienz als dem dominanten Faktor für das Auswählen des angemessenen *starting point* aus. Von diesen zwei verschiedenen Annahmen ausgehend, ist es daher notwendig, die semantische und visuelle Salienz der Stimuli zu manipulieren und deren Effekte auf die Satzplanung im Deutschen zu messen.

Zum anderen hat diese Studie zum Ziel, die Rolle der kontextuellen Salienz bzw. der Topikalität bei der Satzplanung zu untersuchen. Damit dies gelingen kann, ist es wichtig, einen gewissen Kontext für Versuchspersonen zu schaffen, damit die Satzproduktion immer in einen bestimmten Kontext einzuordnen ist. Dabei sollten die zu beschreibenden Figuren einen unterschiedlichen Diskursstatus aufweisen. Nachfolgend wird ausführlich darauf eingegangen, wie die oben erwähnten Faktoren – die semantische, visuelle und kontextuelle Salienz – in dieser Arbeit manipuliert werden.

Manipulation der semantischen Salienz

Viele Studien haben gezeigt, dass Figuren eines Ereignisses, welche unterschiedliche semantische Funktionen aufweisen, bei ihrer Enkodierung auch an ungleichgewichtiger Präferenz gewinnen (vgl. die thematische Hierarchie). Diejenigen Figuren, deren semantische Salienz hoch ist, werden früher verarbeitet als diejenigen, die eine semantisch geringere Auffälligkeit aufweisen und zu einem späteren Zeitpunkt enkodiert werden (vgl. Segalowitz, 1982; Dobel et al., 2007). Zwei semantische Rollen, die in vorangegangenen Studien intensiv untersucht wurden, sind das Agens und das Patiens. Viele Korpus- und Eye-Tracking-Studien haben gezeigt, dass im Gegensatz zum Patiens das Agens bei der Verarbeitung eines

transitiven Ereignisses früher identifiziert und lexikalisch enkodiert wird (z. B. Dryer, 2013; Greenberg, 1963; Tomlin, 1986; Dobel et al., 2007; Webb et al., 2010). Das deutet auf eine höhere semantische Salienz dieser thematischen Rolle hin.

Anhand dieses Ungleichgewichts werden in dieser Studie transitive Ereignisse verwendet, in denen die Ereignisteilnehmer jeweils als Agens und Patiens auftreten und daher unterschiedliche semantische Salienz aufweisen. Ein Beispiel für die Stimuli sieht man in Abbildung 5.1, in der ein *bespritzen*-Ereignis gezeigt wird, bei dem die *Maus* als Agens semantisch salienter als das Patiens, die *Schildkröte*, ist.

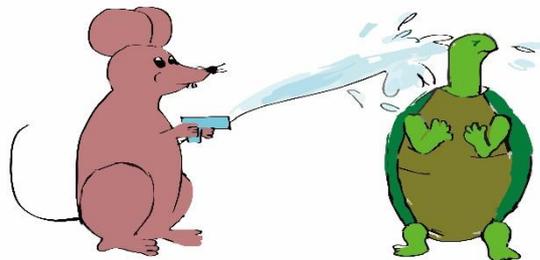


Abbildung 5.1: Beispielstimulus aus der vorliegenden Studie

Des Weiteren ist zu erwähnen, dass Figuren eines Ereignisses nicht nur semantische Salienz, sondern auch inhärente Eigenschaften besitzen, die ebenfalls einen Einfluss auf die Verarbeitungsreihenfolge haben können. Zwei hervorstechende Merkmale davon sind die Belebtheit (z. B. Branigan et al., 2008; Bock & Warren, 1985; Bock et al., 1992; Prat-Sala, 1997; Van Nice & Dietrich, 2003) und die visuelle Salienz (z. B. Gleitman et al., 2007; Myachykov et al., 2008). Auf die visuelle Salienz wird im nächsten Abschnitt näher eingegangen, da sie einer der Faktoren ist, die in der vorliegenden Untersuchung berücksichtigt werden. In Bezug auf die Belebtheit hat eine Menge von Studien gezeigt, dass belebte Figuren mit größerer Wahrscheinlichkeit als Subjekt realisiert bzw. satzinitial arrangiert werden. Da der Effekt der Belebtheit kein Forschungsgegenstand der vorliegenden Studie ist, wird sie bewusst kontrolliert, d.h., die Belebtheiten aller Figuren sind kongruent. Dadurch

beeinträchtigt sie dann nicht als unerwartete Variable die Funktionen der im Fokus der vorliegenden Arbeit stehenden Faktoren.

Manipulation der visuellen Salienz

Das Ziel der Manipulation der visuellen Salienz ist, dass ein Referent im Vergleich zu einem anderen visuell salienter ist und ihm deshalb mehr anfängliche visuelle Aufmerksamkeit (erste Fixationen) nach Stimulus-Onset gewidmet wird. Um dieses Ziel zu erreichen, wird in dieser Studie, im Gegensatz zu Tomlin (1995, 1997) und analog zu Gleitman et al. (2007), ein exogener visueller Cue verwendet.

Im psycholinguistischen Bereich stehen Forschern zur Manipulation der visuellen Salienz vor allem zwei Methoden zur Verfügung, bei denen entweder endogene (z. B. Tomlin, 1995) oder exogene visuelle Cues eingesetzt werden (z. B. Gleitman et al., 2007; Hwang & Kaiser, 2015; Myachokov et al., 2010). Der Grund, warum in der vorliegenden Arbeit ein exogener Cue als passendes Instrument zur Manipulation ausgewählt wurde, ist die Art und Weise, in der er die visuelle Aufmerksamkeit steuert. Während die endogene Aufmerksamkeitsmanipulation von der willentlichen Kontrolle ausgelöst wird, läuft die exogene Aufmerksamkeitssteuerung reflexartig, unbewusst und automatisch ab.

Um die Mechanismen des endogenen visuellen Cues beleuchten zu können, werfen wir einen Blick auf das Experiment von Tomlin (1995), das bereits in Abschnitt 3.2.2 vorgestellt wurde. In diesem Experiment sahen die Versuchspersonen ein Video mit zwei sich einander annähernden Fischen. Ein Fisch wurde zum Schluss von dem anderen gefressen. Während des ganzen Videoverlaufs wurde ein Fisch von einem endogenen Cue (einem Pfeil) angezeigt. Es zeigte sich, dass dieser Fisch bei der Beschreibung früher erwähnt wurde. Zur Erklärung der beobachteten Korrespondenz zwischen dem zuerst erwähnten Referenten und dem vom endogenen Cue markierten Fisch nahm Tomlin an, dass der Pfeil die visuelle Salienz des angezeigten Fisches so erhöht hat, dass er die Aufmerksamkeit des Sprechers auf sich gezogen hat. Das hat wiederum dazu geführt, dass eben dieser Fisch als *starting point* der

Verbalisierung ausgewählt wurde.

Dennoch ist das Experiment auf verschiedenartige Kritik gestoßen, die vor allem die Methodologie betraf (z. B. Bock, 2004): Obwohl sich die Versuchspersonen vor dem Experiment nicht über die Funktion des endogenen Cues bzw. dessen Verhältnis mit der Selektion syntaktischer Strukturen informierten, erlaubte die vermehrte Aufmerksamkeit neben der Erhöhung der visuellen Salienz auch eine beträchtliche Steigerung der konzeptuellen Zugänglichkeit der markierten Figur. Zudem konnte die dauerhafte Präsentation eines Pfeils auf einen Fisch auch perspektivische Informationen über das Ereignis vermittelt haben. Das heißt, dass der markierte Referent durch den Pfeil den Topikstatus erhält, was wiederum zur Voranstellung dieses Referenten führen kann. In diesem Zusammenhang übt der endogene visuelle Cue im Gegensatz zu anderen visuellen Manipulationen (z. B. Änderung der Größe und Farbe), die auf einer Bottom-Up-Funktionsweise beruhen, eher eine Top-Down-Auswirkung aus.

Von diesen beiden Kritikpunkten ausgehend, kann man sagen, dass die Darbietung eines endogenen visuellen Cues zusammen mit dem Stimulus zwar die visuelle Aufmerksamkeit steuert, sie dennoch die Möglichkeit, dass eine parallele Steigerung der konzeptuellen und kontextuellen Salienz der durch den endogenen Cue angezeigten Entität während dieses Prozesses ebenfalls stattfindet, nicht vollständig ausschließen kann. Demnach ist die Schlussfolgerung, dass die Tendenz, die visuell markierte Figur an den Satzanfang zu positionieren, nur auf die erhöhte visuelle Salienz zurückzuführen ist, nicht ausreichend gerechtfertigt. Aus diesem Grund wird der endogene Cue nicht als geeignete Methode zur Manipulation der visuellen Salienz angesehen. Stattdessen wird in dieser Studie ein exogener visueller Cue verwendet, durch welchen die Manipulation rein reizgesteuert und reflexartig erfolgt. Eine Evidenz für diese Funktionsweise des exogenen visuellen Cues wurde durch die Studie von Theeuwes, Kramer, Hahn und Irwin (1998) geliefert.

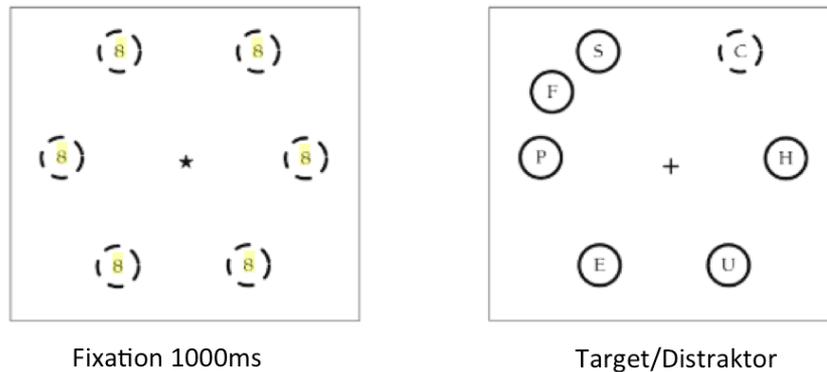


Abbildung 5.2: Schematische Illustration der Stimuli aus Theeuwes et al. 1998. Graue Kreisperipherien sind durch gestrichelte Linien angezeigt, und rote Kreisperipherien sind durch durchgezogene Linien angezeigt; zitiert in McCarley & Kramer 2007

Bei Theeuwes et al. (1998) mussten die Versuchspersonen zuerst einen Punkt in der Mitte des Bildschirms fixieren, der von sechs grauen Kreisperipherien umschlossen war. Jede Kreisperipherie enthielt eine Figur-8 als Premasks. Kurz nach dem Erscheinen der Stimuli änderte sich die Farbe aller Kreisperipherien außer einer von grau zu rot und alle Figur-8-Premasks verwandelten sich in verschiedene Buchstaben. Die Aufgabe der Versuchspersonen bestand darin, den einzig übrig gebliebenen grauen Kreis zu fixieren und den Buchstaben in der Mitte dieses Kreises zu identifizieren. In einigen Durchgängen wurde während der Farbveränderung kurz ein zusätzlicher Kreis, der keinen Zusammenhang mit dem eigentlichen Ziel des Experimentes hatte, präsentiert (vgl. Abbildung 5.2). Es zeigte sich, dass die Probanden ihre Augen stets auf den zusätzlichen Kreis gerichtet hatten, bevor sie den eigentlichen Target-Kreis fixierten und den sich darin befindlichen Buchstaben identifizierten. Nach eigener Aussage der Probanden hat keiner von ihnen wahrgenommen, dass ihre Blicke von dem unerwartet erscheinenden Kreis angezogen wurden. Einige Probanden wussten nicht einmal, dass der Kreis existierte, obwohl ihre Augen bereits auf ihn gerichtet waren.

Aus diesen Ergebnissen schlussfolgerten die Autoren, dass das plötzliche Erscheinen eines Reizes den Blick automatisch an sich zieht - unabhängig davon, ob die Versuchspersonen ein anderes Ziel fixieren wollten oder nicht. Das Experiment zeigt auch, dass diese Manipulation in einer reflexartigen und automatischen Art und

Weise verläuft und deshalb rein reizgesteuert ist.

Basierend auf den genannten Befunden und Analysen wird in der vorliegenden Arbeit ähnlich wie Gleitman et al. (2007) ein exogener visueller Cue verwendet, um die visuelle Salienz eines bestimmten Charakters zu erhöhen und somit die frühe visuelle Aufmerksamkeit zu manipulieren.

Manipulation der kontextuellen Salienz

Wie in Abschnitt 5.1 vorgestellt, besteht der Effekt der Topikalität einerseits aus dem Gegebenheitseffekt, der auf der Vorerwähnung des Topik-Referenten in der vorangegangenen Äußerung begründet, und andererseits aus einem anderen Effekt, der auf das Interesse des momentanen Gesprächs angewiesen ist. Dieser spezifische Effekt der Topikalität, welcher wenig mit dem konzeptuell und lexikalisch bedingten Gegebenheitseffekt zu tun hat, sondern überwiegend von Motivationen oder Intentionen von Gesprächsteilnehmern gesteuert ist, soll in dieser Studie getestet werden.

Eine Schwierigkeit der Manipulation der kontextuellen Salienz besteht darin, den gesprächsinteressen-bedingten Effekt von dem der Gegebenheit zu trennen. Wenn Sprecher die Frage *Was passiert mit der Schildkröte?* mit dem Satz *Die Schildkröte wird von einer Maus bespritzt* beantworten, können die Gründe für die Voranstellung der Figur *Schildkröte* sowohl in dem Topikstatus dieser Figur als auch in deren Vorerwähnung liegen. Damit ergibt sich die Frage, wie der Effekt der Topikalität, der auf dem augenblicklichen Gesprächsinteresse gründet, von dem der Vorerwähnung getrennt werden kann.

Eine Möglichkeit, die das oben genannte Problem lösen kann, liefert das strukturelle Priming-Paradigma, welches auch in der Studie von Fleischer et al. (2012) verwendet wurde. Wenn polnischsprachige Versuchspersonen vor der Beschreibung eines transitiven Ereignisses einen Prime-Satz gehört hatten, in dem das Patiens am Satzanfang stand und daher „hervorgehoben“ wurde, produzierten sie bei der

Beschreibung deutlich mehr Passiv- und OVS-Aktivsätze. In diesen Sätzen platzierten sie das Patiens satzinitial, wodurch dieses weiterhin im Aufmerksamkeitsfokus blieb. Es sollte in dieser Studie kein Effekt der Gegebenheit gefunden werden, da die von den Versuchspersonen beschriebenen Ereignisse aus Entitäten gebildet wurden, die in den Primesätzen nicht erwähnt wurden. Somit können die Autoren annehmen, dass die durch das strukturelle Priming erworbenen Priming-Effekte allesamt auf den topikalisch-spezifischen Effekt der Patiens-Funktion angewiesen sind.

Obwohl durch das strukturelle Priming-Paradigma eine erfolgreiche Trennung zwischen den beiden Effekttypen gelungen ist, dürfen dabei zwei Probleme nicht außer Acht gelassen werden. Das erste Problem betrifft den Ursprung der kontextuellen Salienz. Die kontextuelle Salienz ist im Wesentlichen ein Resultat des Zusammenspiels von Informationsstand und Aufmerksamkeitsfokus eines Gesprächs. Allerdings entsteht die Informationsstruktur der gebildeten Sätze in der zitierten Studie nicht aus vorangegangenen Äußerungen oder momentanen Gesprächsinteressen, sondern aus einem Prime-Satz, der in Wirklichkeit sowohl bedeutungsmäßig als auch kontextuell nicht mit dem gebildeten Satz zu tun hat. Deswegen ist die Schlussfolgerung, dass die kontextuelle Salienz durch Prime-Sätze erzeugt werden kann, höchst bedenklich. Zum anderen ist fragwürdig, ob der gefundene Priming-Effekt wirklich auf die kontextuelle Salienz zurückzuführen ist. Da die Prime-Sätze (mit Patiens am Satzanfang) und die gebildeten Sätze (entweder Passivsätze oder OVS-Aktivsätze) eine identische thematische Abfolge haben, ist nicht auszuschließen, dass der Priming-Effekt von der geteilten semantischen Struktur anstatt der geteilten Informationsstruktur, ausgelöst wurde. Auf Grundlage der beiden Kritikpunkte wird das strukturelle Priming-Paradigma nicht als angemessene Methode für die vorliegende Untersuchung angesehen.

Zur Einrichtung eines Kontextes wird in dieser Studie von dem begrifflichen Kern der kontextuellen Salienz ausgegangen, also einem Resultat des Zusammenspiels des Informationsstands und des augenblicklichen Gesprächsinteresses. Es wird zuerst ein

Common Ground für die Versuchspersonen geschaffen, in dem die Gegebenheitsgrade der an dem Ziel-Ereignis beteiligten zwei Figuren gleich sind (*Ermittlung der Gegebenheit*). Darauf beruhend wird eine von den beiden Figuren durch eine bestimmte Frage spezifiziert, damit diese im Gegensatz zur anderen neben dem gleichen Gegebenheitsstatus noch einen Topikstatus erhält (*Ermittlung des Topikstatus*).

Zur Ermittlung der Gegebenheit wird vor der Präsentation eines Target-Ereignisses eine sog. Preview-Phase eingesetzt, in der die an diesem Target-Ereignis beteiligten Figuren im Voraus präsentiert werden. So sehen die Versuchspersonen beispielsweise zunächst eine *Maus* und eine *Schildkröte*, bevor das Target-Ereignis *Eine Maus bespritzt eine Schildkröte* präsentiert wird. Während der Darstellung der Figuren hören die Probanden noch einen Einführungssatz wie etwa *Auf diesem Bild siehst du eine Maus und eine Schildkröte*, durch welchen die Ereignisteilnehmer auch lexikalisch als gegebene Informationen festgelegt werden. Die visuelle Vorschau und der auditiv präsentierte Einführungssatz sollen gemeinsam einen Common Ground für die darauffolgende von den Versuchspersonen zu produzierende Äußerung schaffen, in dem die beiden Ereignisteilnehmer das gleiche Gegebenheitsniveau haben.

Nach der Einführung der Figuren und vor der Präsentation des Target-Ereignisses wird den Probanden eine Frage gestellt wie z. B. *Auf dem nächsten Bild* (bezieht sich auf das Bild mit dem Target-Ereignis) *siehst du die Maus wieder, sag mir bitte, was mit der Schildkröte passiert?*. Dadurch wird dann eine bestimmte Figur (hier *Schildkröte*) als Topik der nächsten Äußerung selektiert. Des Weiteren ist zu erwähnen, dass das Nicht-Topik in der gestellten Frage noch einmal erwähnt wird, damit es genauso oft angesprochen wurde und den gleichen Gegebenheitsgrad aufweisen kann (Beispiel s. Abbildung 6.1).

Insgesamt soll durch eine visuelle Vorschau der involvierten Figuren, durch eine sprachliche Einführung der Figurenbezeichnungen sowie durch eine Target-Frage ein

Kontext aufgebaut werden. Die Satzplanung findet innerhalb dieses Kontexts statt.

5.3 Messgrößen

Diese Studie ist eine experimentelle Eyetracking Studie, die den Satzproduktionsprozess und insbesondere den dazugehörigen Konzeptualisierungsprozess im Deutschen zum Forschungsgegenstand hat. Um hier relevante Ergebnisse zu erzielen, muss der Konzeptualisierungsprozess durch Werte, die gemessen werden können, explizit dargestellt werden. Hierfür bedienen wir uns linguistischer und visueller Messgrößen als Indikatoren für diesen Verarbeitungsprozess. Im Folgenden werden die beiden Messgrößen-Typen charakterisiert.

5.3.1 Linguistische Messgrößen

Syntaktische Strukturen der produzierten Sätze

Eine wichtige Aufgabe der Konzeptualisierung ist, *starting points* der Verbalisierung auszuwählen. Zur Beschreibung eines transitiven Ereignisses stehen deutschen Sprechern mindestens drei syntaktische Strukturen zur Verfügung: kanonischer Aktivsatz, Scrambled-Aktivsatz und Passivsatz. Die Spezifizierung der syntaktischen Struktur kann die Selektion von *starting points* reflektieren: Figuren, die als *starting points* der sprachlichen Enkodierung ausgewählt werden, werden dementsprechend zuerst verarbeitet und in einer frühen bzw. prominenten Position, in der Regel am Satzanfang, platziert (vgl. Ferreira & Dell, 2000). Daher erlaubt die Beobachtung der konstruierten syntaktischen Struktur, die selektierten *starting points* festzustellen.

Da die vorliegende Arbeit zum Ziel hat, den Einfluss der visuellen, semantischen und kontextuellen Salienz auf die Selektion von *starting points* festzustellen, gestattet ein Vergleich der Vorhersagbarkeit dieser Faktoren auf den zuerst erwähnten Referenten somit zu evaluieren, welcher Faktor eine dominante Rolle in diesem Prozess spielt.

Als Nebenaufgabe interessieren wir uns ebenfalls dafür, in welcher Art und Weise die kontextuelle Information wie beispielsweise das Topik auf die Syntax abgebildet wird. Diese Frage kann ebenfalls mithilfe der syntaktischen Strukturen der produzierten Sätze beantwortet werden. Dabei steht vor allem die Frage im Vordergrund, ob das Patiens als Topik bei der Satzproduktion Scrambled-Aktivsätze hervorrufen kann, in denen es als Objekt die satzinitiale Position einnimmt, oder ob ihm bei der syntaktischen Strukturierung durchaus die Subjekt-Funktion zugewiesen wird. Auf diese Frage wird in Abschnitt 7.1.2 näher eingegangen.

Speech-Onset-Latenz

Unter der Speech-Onset-Latenz ist das Zeitfenster zwischen dem Stimulus-Onset und dem artikulatorischen Onset des ersten Äußerungssegments zu verstehen. Diese zeitliche Messgröße kann Auskunft über den kompletten kognitiven Prozessablauf von der visuellen Wahrnehmung bis zur Artikulation geben. Der Grund dafür, dass diese Messung durchzuführen und in der Datenanalyse zu berücksichtigen ist, liegt darin, dass es sich durch den Vergleich zwischen der Speech-Onset-Latenz beim Produzieren in einer pragmatisch neutralen Situation (Experiment 1), in der die Diskursstatus der Referenten nicht voneinander abweichen, und in einer Situation (Experiment 2/3), in der ein Referent mit dem Topik-Status kontextuell salienter als der andere hervortritt, feststellen lässt, in welchem Zeitverlauf die Topik-Information während des Satzproduktionsvorgangs prozessiert wird. Es sei daran erinnert, dass die Zuweisung des Topikstatus in dieser Studie durchaus in der Preview-Phase, also vor der Präsentation des Target-Bildes, stattfindet. Wenn das Topik direkt nach seiner Einführung in der Preview-Phase als *starting point* ausgewählt und ausgesprochen wird, ohne zuerst das Target-Ereignis näher betrachten zu müssen, sollte eine kürzere Sprech-Onset-Latenz zu erwarten sein als in der Situation, in der kein Topikstatus vorab zugewiesen wird und die Sprecher den *starting point* dann erst nach der Präsentation des Bildes selektieren können. Nähere Diskussion darüber findet man in Abschnitt 7.1.2.

5.3.2 Visuelle Messgrößen

In Bezug auf die visuellen Messgrößen sind hierbei vor allem verschiedene aufgezeichnete sakkadische Parameter im Zuge der Satzplanung gemeint.

Mit Sakkaden werden schnelle ballistische Augenbewegungen beschrieben, durch welche ein Objekt mit der Fovea erfasst und ein Fokuswechsel vollgezogen wird. Die Sakkadengenerierung verläuft in 3 Phasen (vgl. Abbildung 5.3): In der ersten Phase werden visuelle Informationen, die durch die Fovea aufgenommen werden, verarbeitet, um das Ziel der Sakkade festzustellen (vgl. Becker, 1991; Becker & Jürgens, 1979; Engbert, Longtin & Kliegl, 2002; Nuthmann, Smith, Engbert & Henderson, 2010). Nach der Verarbeitung erreicht die Sakkadenplanung eine sogenannte *dead-time*-Phase, welche auf sensorische Transduktion und Verspätungen des motorischen Pfades zurückgeführt wird. Die *dead-time*-Phase dauert ca. 70 - 80 ms (vgl. Ludwig, Mildinhall & Gilchrist, 2007; Hooge & Erkelens, 1996). Während dieses Zeitraums wird die Zielposition der Sakkade nicht von der visuellen Information beeinflusst (vgl. Caspi, Beutter & Eckstein 2004). Nach der neuronalen Bereitstellung wird die Sakkade dann in der dritten Phase ausgeführt. Im Folgenden werden einige sakkadische Parameter charakterisiert, welche für Analyse der Augenbewegungsdaten in der vorliegenden Arbeit relevant sind.

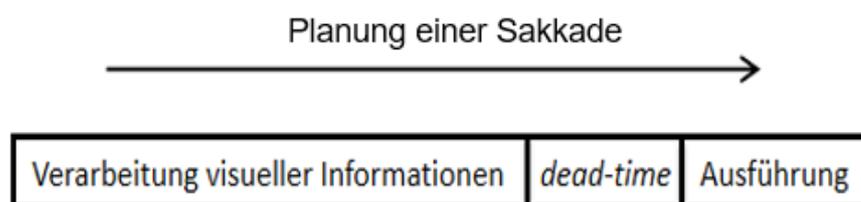


Abbildung 5.3: Drei Phasen der Sakkadengenerierung

Zielposition der ersten und zweiten Sakkade

Menschen "blicken" nicht beliebig. Bei der Planung einer Sakkade muss das kognitive System in erster Linie berechnen, auf welche Regionen die Augen gerichtet werden. Die Selektion einer sakkadischen Zielposition gelingt somit nicht

wahllos, sondern ihr liegt ein bestimmter kognitiver Verarbeitungsvorgang zugrunde, der es erlaubt, nur relevante Informationen zu fokussieren und den Rest unbeachtet zu lassen (vgl. Gilchrist, 2011). Daher können wir durch die Beobachtung der selektierten Zielposition einer Sakkade darauf rückschließen, welchem Referenten vor der Ausführung dieser Sakkade Aufmerksamkeit gewidmet und dieser somit verarbeitet wurde.

In diesem Experiment interessieren wir uns in erster Linie für die Zielposition der ersten Sakkade, da durch den Vergleich zwischen der ersten Zielposition und der Position des exogenen Cues abgeleitet werden kann, ob der visuelle Cue einen Einfluss auf die frühe Aufmerksamkeit während der Satzplanung hat. Des Weiteren wird die Zielposition der zweiten Sakkade überprüft, und zwar aus zwei Gründen: Einerseits fällt - ähnlich wie die erste Sakkade - die Planung und Ausführung der zweiten Sakkade in der Regel ebenfalls in die von Griffin und Bock (2000) und Bock et al. (2004) angenommene Zeitspanne der konzeptuellen Vorausplanung (0 – 400 ms). Andererseits ist die zweite Sakkade - anders als die erste Sakkade, welche überwiegend reizbedingt in einem Bottom-Up-Prozess geplant wird - in der Tat die erste willkürlich ausgeführte Sakkade, deren Planung nicht von dem vorab dargebotenen visuellen Cue beeinflusst wird. Ihr liegt stattdessen ein Top-Down-Planungsvorgang zugrunde. Deswegen lassen sich aus der selektierten Zielposition der zweiten Sakkade Rückschlüsse darauf ziehen, welcher Faktor den Top-Down-Planungsvorgang der Information auf der Konzeptualisierungsebene dirigiert.

Intersakkadisches Intervall

Das intersakkadische Intervall bezeichnet die Zeit, die zwischen dem Ausführungsabschluss einer Sakkade und dem Ausführungsbeginn einer folgenden Sakkade vergeht (vgl. oberer Teil der Abbildung 5.4). Studien zur visuellen Verarbeitung haben gezeigt, dass durch das intersakkadische Intervall der Umfang und der Zeitverlauf der konzeptuellen Verarbeitung reflektiert werden (z. B.

Carpenter & Williams, 1995; Gilchrist, 2011;). Ein Beispiel zur Erklärung dieses Zusammenhangs liefern Studien über Antisakkade (z. B. Hallett, 1978; Hallett & Adams, 1980; Everling & Fischer, 1998).

In einer klassischen Antisakkadenstudie werden Versuchspersonen dazu aufgefordert, zunächst einen Fixationspunkt anzusehen und beim Erscheinen eines zweiten Punktes, welcher zufällig links oder rechts zu dem Fixationspunkt gezeigt wird, nicht zu diesem hin, sondern in dessen entgegengesetzte Richtung zu blicken. Es zeigte sich, dass die Versuchspersonen oft entgegen ihrem Willen zuerst zum neu erschienenen Reiz schauen. Allerdings schaffen sie es auch, die Augenbewegungen sehr schnell umzukehren, damit der Fehler korrigiert werden kann. Was hier relevant ist, ist, dass das intersakkadische Intervall zwischen der fehlerhaften und der korrektiven Sakkade (ca. 10-150 ms) deutlich geringer als ein normales intersakkadisches Intervall ist. Da die Planung einer Sakkade in der Regel ca. 200 – 300 ms in Anspruch nimmt (vgl. Fischer & Weber, 1993; Griffin & Bock, 2000; Saslow, 1967), wird dieses kurze intersakkadische Intervall von Psychologen als Hinweis auf eine parallele Programmierung der irrtümlichen und der korrektiven Sakkade interpretiert (z. B. Becker & Jürgens, 1979; Walker & McSorley, 2006). Das heißt, dass während der Planung und Ausführung einer vorangehenden Sakkade die Planung einer nachkommenden Sakkade bereits angefangen hat und daher die intersakkadische Latenz verkürzt wird (vgl. unterer Teil der Abbildung 5.4).

Von dieser Interpretation ausgehend, wird in dieser Studie die intersakkadische Latenz der ersten und der zweiten Sakkade ausgewertet. Der Grund für diese Auswertung liegt darin, dass anhand des Wertes der intersakkadischen Latenz wir möglicherweise den Zeitverlauf der semantischen Strukturierung berechnen können. Es wird zuerst beobachtet, ob die zweite Sakkade statt visuell gesteuert, von der semantischen Salienz beeinflusst wird. Sollte dies der Fall sein, dann bedeutet das Ergebnis, dass die semantische Struktur bereits vor der Ausführung der zweiten Sakkade aufgebaut ist. Wenn dazu noch ein kurzes intersakkadisches Intervall

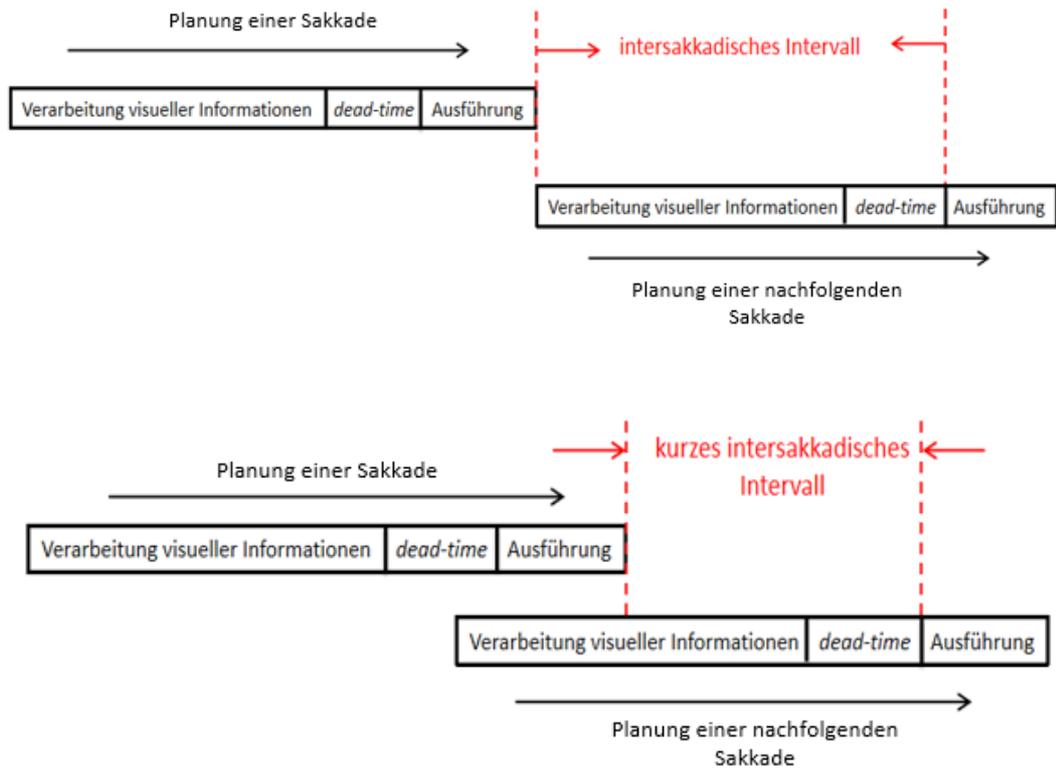


Abbildung 5.4: Intersakkadisches Intervall (im oberen Teil) und verkürztes intersakkadisches Intervall (im unteren Teil)

beobachtet werden kann, welches auf eine parallele Programmierung der ersten und zweiten Sakkade hinweist, dann können wir die Zeitspanne der semantischen Strukturierung weiter verengen, denn dies weist darauf hin, dass der semantische Strukturierungsprozess vor dem Abschluss der ersten Sakkade bereits abgeschlossen ist.

6 Experiment 1

Das nachfolgende Experiment beschäftigt sich mit verschiedenen Perspektiven des Satzproduktionsprozesses im Deutschen in pragmatisch neutralen Situationen. Hierbei handelt es sich um zwei divergierende Arbeitsweisen des Satzproduktionssystems, die entweder der strukturellen Repräsentation eines Gesamtereignisses wie etwa der semantischen Struktur oder der Eigenschaft der Ereignisteilnehmer wie etwa der visuellen Salienz die entscheidende Rolle für die Selektion von *starting points* zuschreiben. Des Weiteren ist von Interesse, in welchem Verhältnis die bei der Satzproduktion mitwirkenden Faktoren zueinanderstehen. Die Faktoren im Fokus dieses Experiments sind: Die visuelle Salienz, die durch einen exogenen, aufmerksamkeitslenkenden Cue erhöht wird, und die semantische Salienz, die durch die Zuweisung unterschiedlicher semantischer Rollen hervorgerufen wird. Die generellen Fragen, die dieses Experiment leiten, sind: (1) Ist die Satzproduktion im Deutschen in einer pragmatisch neutralen Situation ein linear oder ein strukturell inkrementell verlaufender Prozess? (2) Welcher Faktor spielt eine dominante Rolle bei der Konzeptualisierung und der grammatischen Enkodierung im Rahmen der Satzproduktion im Deutschen?

6.1 Hypothesen

6.1.1 Hypothesen über die Arbeitsweise des Satzproduktionssystems: lineare vs. strukturelle Inkrementalität

Blickbewegungsdaten

Eine der zentralen Fragen, der in der vorliegenden Arbeit nachgegangen wird, ist, ob die Satzplanung im Deutschen linear oder strukturell inkrementell abläuft. Das dritte Kapitel dieser Arbeit widmete sich den Evidenzen, sowohl für die lineare als auch für die strukturelle Inkrementalitätshypothese, die vorwiegend in Eye-Tracking-Studien gefunden wurden. Zwei Indikatoren wurden zur Festlegung der Funktionsweise der inkrementellen Verarbeitung benutzt: einerseits die frühen

Blickverteilungen (ca. 300 – 400 ms nach Stimulus-Onset) und das Verhältnis zwischen der visuell bedingten ersten Fixation und den darauffolgenden Fixationen andererseits.

Der erste Indikator wurde beispielsweise von Griffin und Bock (2000) verwendet. Griffin und Bock kommen in ihrer Studie zu dem Ergebnis, dass Fixationen anfänglich gleichmäßig auf das Agens und das Patiens verteilt waren, bevor eine visuelle Präferenz für eine bestimmte thematische Rolle (z. B. das Patiens) auftrat. Diese Präferenz für eine thematische Rolle setzt den Aufbau der semantischen bzw. konzeptuellen Struktur voraus. Daher fassen die Autoren die gleichmäßige Verteilung der visuellen Aufmerksamkeit, die ca. 300 – 400 ms nach Stimulus-Onset andauert, als Phase der semantischen Strukturierung auf. Vergleichbare Befunde zum Blickmuster finden sich ebenfalls in der Studie von Bock et al. (2004). Ehe die Fixationen auf Entitäten mit zwei verschiedenen thematischen Rollen auseinandergehen, was den Beginn der linguistischen Verarbeitung kennzeichnet, gab es ein Zeitfenster, in dem keine Präferenz für eine der beiden Rollen festgestellt werden konnte.

Der zweite Indikator wird zumeist in einer Situation verwendet, in der die Ereignisteilnehmer des zu beschreibenden Ereignisses unterschiedlich visuell salient sind. Beispielsweise belegen Gleitman und Kollegen (2007), dass, nachdem ein Ereignisteilnehmer mit höherer visueller Salienz an erster Stelle fixiert wurde, die darauffolgenden Fixationen weiterhin auf demselben Ereignisteilnehmer verharrten, bis dieser als der satzeinleitende Referent versprachlicht wurde. Da die Verbalisierung einer bestimmten Figur intensive Aufmerksamkeit in Anspruch nimmt, weist die Aufrechterhaltung der Fixationen auf dem visuell salienteren Referenten nach den Autoren darauf hin, dass sich die Versuchspersonen bis kurz vor Sprechbeginn lediglich auf die Verarbeitung der zuerst fixierten Figur konzentrierten. Alle anderen Verarbeitungen, dazu gehören auch der Aufbau struktureller Repräsentationen und die Identifizierung anderer Ereignisteilnehmer, vollziehen sich zu einem späteren Zeitpunkt.

Da das experimentelle Design der vorliegenden Studie im Großen und Ganzen dem von Gleitman et al. (2007) folgt, wird der zweite Indikator, also die Beziehung zwischen der visuell bedingten ersten Fixation und den darauffolgenden Fixationen, zur Unterscheidung der linearen und der strukturellen Inkrementalität verwendet. Allerdings interessieren uns an dieser Stelle nicht für die gesamten Augenbewegungen während des Sprechens, sondern nur die in der konzeptuellen Enkodierungsphase stattfindende Augenbewegungen (ca. die erste und zweite Sakkade bzw. Fixation). Dass die Blickbewegungen nach diesem Zeitfenster nicht mehr berücksichtigt werden, liegt darin begründet, dass diese vor allem die Reihenfolge der auf der sprachlichen Oberfläche erscheinenden Wörter widerspiegeln und nicht den Konzeptualisierungsprozess an sich (vgl. Griffin & Bock, 2000; Norcliff et al., 2015; Sauppe et al., 2013). Die vorliegende Arbeit konzentriert sich jedoch nur auf den Prozess der konzeptuellen Vorausplanung.

Das wirft nun die Frage auf, wie es möglich ist, lediglich aus der ersten und zweiten Sakkade bzw. Fixation auf die Mechanismen der Konzeptualisierung zu schließen. Im Folgenden beziehen wir uns kurz auf die Implikationen, die sich aus den beiden Sakkaden ziehen lassen, bevor wir auf die Hypothesen eingehen.

Für einige Psycholinguisten ist es fraglich, ob sich die Konzeptualisierung, die nach Stimulus-Onset sehr schnell stattfinden kann, mithilfe der Augenbewegungsmessung untersuchen lässt. Und dies, obwohl die linguistische Verarbeitung bzw. die entstehende Reihenfolge der auf der linguistischen Oberfläche erscheinenden Worteinträge bei einer Ereignisbeschreibung durch die Reihenfolge der Fixationen sichtbar gemacht werden kann.

Diese Bedenken sind auf den ersten Blick gerechtfertigt, da eine Reihe experimenteller Untersuchungen bereits gezeigt hat, dass Menschen in der Lage sind, ein Ereignis gleich nach dessen Präsentation grob zu verstehen. Dieser Prozess kann sehr schnell abgeschlossen werden, bevor sich die Augen überhaupt bewegen (vgl. Biderman, 1981; Dobel et al., 2007; Potter, 1975, 1976). Beispielsweise wurden den

Versuchspersonen in der Studie von Dobel et al. (2007), die in Abschnitt 4.3 erwähnt wurde, Bild-Stimuli mit einem SCHIEßEN- oder einem GEBEN-Ereignis präsentiert und sie in einer Bedingung gebeten, das Agens, das Patiens und das Objekt nach einer Darstellungslatenz von entweder 100, 200, 250 oder 300 ms zu benennen. Ab einer Präsentationsdauer von 200 ms nach Stimulus-Onset steigerte sich die Menge der korrekt identifizierten Agens-Benennungen. Das weist darauf hin, dass semantische bzw. konzeptuelle Strukturen selbst nach einer sehr kurzen Präsentationsdauer konstruiert werden können. Dieser Befund gleicht dem von Griffin und Bock (2000), die in ihrer Studie zu folgendem Schluss kommen: Die Extraktion der sog. *event gist* ist innerhalb 400 ms nach Stimulus-Onset abgeschlossen. Wenn wir diese Zeit mit Augenbewegungen in Verbindung bringen, so reicht die Zeitspanne von ca. 200 - 400 ms nach Stimulus-Onset lediglich für Planungen von einer bis zwei Sakkaden. Dies scheint aber für die Identifizierung des zeitlichen Verlaufs und des Inhalts der Konzeptualisierung nicht ausreichend zu sein. Man kann sich nun fragen, ob die Augenbewegungen genügend Auskunft über die sich auf der konzeptuellen Ebene vollziehenden Prozesse geben können.

Auch wenn in dem Zeitfenster der konzeptuellen Vorausplanung lediglich zwei Sakkaden geplant werden können, sollten deren Implikationen für die Identifizierung des Inhalts und des Zeitverlaufs der Konzeptualisierung nicht unterschätzt werden, weil den Sakkaden Verarbeitungsmechanismen unterschiedlicher Art zugrunde liegen.

In Bezug auf die erste Sakkade haben frühere Studien gezeigt, dass sie in den meisten Fällen reizbedingt auf die Entität, deren Position durch einen Cue markiert ist, gerichtet wurde, was auf eine Bottom-Up-Planungsweise hindeutet (z. B. Gleitman et al., 2007; Hwang & Kaiser, 2015; Myachkov & Tomlin, 2008). Da dieses Experiment-Design denen der früheren Studien gleicht, ist daher ein vergleichbares Ergebnis zu erwarten. Das soll heißen, dass die erste Fixation ungeachtet der semantischen Funktionen der Ereignisteilnehmer durchgehend auf denjenigen Referenten platziert werden sollte, der durch den Cue markiert ist.

Im Gegensatz zur ersten Sakkade ist die zweite Sakkade in der Tat die erste selbstbestimmte und willentlich ausgeführte Sakkade. Deswegen sollten uns diese Daten erlauben, diejenige Information, die deutsche Sprecher in der konzeptuellen Vorausplanungsphase top-down verarbeitet haben, abzuleiten. Entscheidend hierbei ist zu ermitteln, in welchem Verhältnis die zweite Sakkade zur ersten steht. Wenn der Planung der zweiten Sakkade parallel zur ersten die visuellen Eigenschaften der Ereignisteilnehmer zugrunde liegen, sollte danach eine zweite Fixation beobachtet werden, die auf der visuell durch den Cue angezeigten Entität bleibt. Wenn der semantischen Salienz eine zentrale Rolle bei der Planung der zweiten Sakkade zukommt, sollte die zweite Sakkade ohne Rücksicht auf die Position der ersten Fixation immer auf die semantisch salientere Figur, also das Agens, gerichtet sein.

Die erste Hypothese steht grundsätzlich in Übereinstimmung mit dem Konzept der linearen Inkrementalität, nach der die visuelle Salienz bei der Auswahl von *starting points* eine zentrale Rolle spielt. Die Figuren, die aufgrund ihrer höheren visuellen Salienz – unabhängig davon, welche semantische Rolle sie haben - als Erstes fixiert werden, werden weiterhin fixiert, bis kurz vor dem Sprechanfang, was auf einen anschließenden lexikalischen Zugriff hinweisen kann. Dagegen spricht die zweite Hypothese, nämlich, dass die Planung der zweiten (also der ersten willentlich ausgeführten) Sakkade von den semantischen Funktionen der Ereignisteilnehmer modifiziert wird, für die strukturelle Inkrementalitätsthese, nach der die semantische Struktur, sobald sie aktiviert wurde, die Funktion der visuellen Salienz übernimmt und weitere konzeptuelle Verarbeitungen dirigiert. So wird die zweite Sakkade bei der Beschreibung eines transitiven Ereignisses ungeachtet der visuellen Salienz immer auf das Agens gerichtet.

Produzierte Sätze

In Bezug auf die produzierten Sätze gilt es noch zu ermitteln, ob der zuerst fixierte Referent mit dem satzeinleitenden korrespondiert. Nach der Grundidee der linearen Inkrementalität spielt die visuelle Salienz bei der Auswahl von *starting points* eine

entscheidende Rolle. Diejenigen Figuren, die visuell am salientesten sind, werden an erster Stelle fixiert, konzeptualisiert und linguistisch enkodiert. Gleitman et al. (2007) kommen zu dem Schluss, dass die visuell markierte und daher zuerst fixierte Figur eher an den Satzanfang gestellt und als das erste Argument kodiert wurde. Ein vergleichbares Ergebnis wurde allerdings bei der Satzproduktion im Koreanischen und Finnischen nicht beobachtet. In der Studie von Hwang und Kaiser (2015) und von Myachokov et al. (2010) fanden die Autoren heraus, dass Fixationen auf dem Patiens die Proportion der gebildeten Passivsätze und Scrambled-Aktivsätze nicht signifikant erhöhten. Die überdurchschnittlich häufige Auswahl des Agens als satzinitialer Referent wird von den Autoren als Beweis dafür angesehen, dass die Verbalisierung trotz der anfänglichen reizinduzierten Verarbeitung überwiegend von der semantisch-konzeptuellen Struktur bestimmt wird.

In dem vorliegenden Experiment wird die gleiche Auswertung durchgeführt, um zu entscheiden, welcher inkrementellen Verarbeitungsweise die Satzproduktion im Deutschen folgt. Wenn die lineare Inkrementalität dominant wäre, dann sollte analog zu Gleitman et al. (2007) eine Korrespondenz zwischen dem zuerst fixierten und dem satzinitialen Referenten beobachtet werden. Wenn jedoch die strukturelle Inkrementalität die dominante Verarbeitungsstrategie in der deutschen Satzproduktion ist, sollte das Agens aufgrund seiner höheren semantischen Salienz ohne Rücksicht auf den anfänglichen Fixationsort durchgehend als *starting point* der Verbalisierung ausgewählt, satzinitial kodiert und als Erstes ausgesprochen werden.

6.1.2 Hypothesen über die Prioritätenliste der mitwirkenden Faktoren: visuelle vs. semantische Salienz

Die Satzproduktion ist ein hochgradig komplexer Prozess, bei dem unterschiedliche Faktoren zusammenwirken. Obwohl alle Faktoren zur Selektion von *starting points* gewissermaßen beitragen, kann ihre Bedeutsamkeit allerdings verschieden sein. So stellt sich die Frage, ob es eine Prioritätenliste für die beiden Faktoren gibt, die in diesem Experiment untersucht werden: die visuelle und die semantische Salienz. Diese Frage soll anhand der Blickbewegungsdaten und der produzierten Sätze

beantwortet werden.

Blickbewegungsdaten

Anhand der die Satzproduktion begleitenden Blickbewegungsdaten soll eine Prioritätenliste der mitwirkenden Faktoren auf der Konzeptualisierungsebene ausgearbeitet werden. Dafür sollen die ersten zwei ausgeführten Sakkaden, die in das Zeitfenster der konzeptuellen Vorausplanung fallen, berücksichtigt werden. Mehrere Studien, in denen das *visual-cuing*-Paradigma benutzt wurde, haben gezeigt, dass die erste Sakkade in meisten Fällen erfolgreich vom vorab präsentierten visuellen Cue manipuliert und daher reflexartig auf die dadurch markierte Figur ausgeführt wurde. Da in diesem Experiment ein analoges Design gewählt wurde, liegt die Vermutung nahe, dass die erste Sakkade im vorliegenden Experiment ebenfalls vom visuellen Cue angetrieben wird. Besonderes Augenmerk liegt daher auf der zweiten Sakkade, um die Wirkungsstärke der visuellen und semantischen Salienz auf die Konzeptualisierung zu identifizieren. Es wird vermutet, dass in die Verarbeitungen der zweiten Sakkade neben der visuellen Salienz auch andere Faktoren wie etwa die semantische Information einfließen. So können wir aus der Beobachtung, ob der zweite Fixationsort eher mit der semantisch oder visuell salienteren Figur korrespondiert, schließen, welche Eigenschaft einen stärkeren Einfluss auf die konzeptuelle Enkodierung hervorrufen kann.

Produzierte Sätze

Die Prioritätenliste der visuellen und der semantischen Salienz während der Satzplanung kann sich auch in den produzierten Sätzen widerspiegeln. Hierfür werden die beiden Faktoren in ihrer Wirksamkeit auf die Selektion des satzinitialen Referenten miteinander verglichen. Der Faktor, der einen größeren Einfluss auf die Selektion hat, sollte eine vorrangige Position auf der Prioritätenliste einnehmen. Im Gegensatz dazu sollte der Faktor, der eine schwächere Wirkung aufzeigt, von Sprechern als weniger vorrangig angesehen werden.

6.2 Methodisches Vorgehen

6.2.1 Versuchspersonen

Insgesamt nahmen 16 Versuchspersonen an dem Experiment teil (davon 12 Frauen). Alle Teilnehmer waren Studierende an der Universität Heidelberg im Alter zwischen 19 und 28 Jahren (durchschnittliches Alter: 22,6) und wuchsen mit Deutsch als Muttersprache auf. Ihre Sehstärke war normal oder mithilfe von Kontaktlinsen oder Brillen auf das Normalniveau korrigiert. Die Teilnahme erfolgte freiwillig, und alle Versuchspersonen erhielten 7 Euro als Aufwandsentschädigung für ihre Teilnahme.

6.2.2 Technische Apparatur

Zur Erstellung der Stimuli in Form von Strichzeichnungen wurde ein digitales Graphiktablett der chinesischen Firma *HUION* in Verbindung mit dem Graphikprogramm *GIMP* verwendet, welches über Windows 10 lief. Das Graphiktablett besteht aus einem USB-Schreibpad und einem Batteriestift. Nach dem Zeichnen der Umrise von Figuren und Ereignissen wurden die Strichzeichnungen im *GIMP* mit Farben ausgefüllt, eventuell geschnitten oder zusammengesetzt und zum Schluss in das JPEG-Format exportiert.

Für das Aufzeichnen der Blickbewegungen wurde ein *RED 500 Eye-Tracking-System* der deutschen Firma *SMI (SensoMotoric Instruments)* verwendet. Das System enthält zwei Teile: die Hardware und die Software. Zur Hardware gehören ein Host-Monitor, ein Eyetracker und eine Gesichtskamera. Der Host-Monitor ist ein 19-Inch TFT-Flachbildschirm (*DELL*) mit einer Auflösung von 1024×968 Pixeln, auf dem die visuellen Stimuli präsentiert wurden. Der Host-Monitor war mit einem Präsentationsrechner verbunden, auf dem die Software *SMI ExperimentCenter* installiert wurde, welche für die Präsentation der Stimuli zuständig war. An dem unteren Rand des Bildschirms war der Eyetracker befestigt, der die Augenbewegungen aufzeichnete. Des Weiteren war die Gesichtskamera, die die Einstellung der Position der Versuchspersonen zu dem Bildschirm ermöglichte,

an dem oberen Rand des Monitors befestigt. Sowohl der Eyetracker als auch die Gesichtskamera waren mit einem Recordingrechner verbunden, mit dem die aufgezeichneten Blickbewegungsdaten gespeichert wurden. Zudem erlaubte es der Recordingrechner auch, den Ablauf des Experimentes bzw. den Status der Augenbewegungen der Versuchspersonen zu überwachen.

Das für die Aufzeichnung der Audio-Daten eingesetzte Gerät war ein *SONY* Ansteckmikrofon, das in die Gesichtskamera eingebaut und mit einer in dem Präsentationsrechner integrierten Soundkarte verbunden war.

6.2.3 Material

Die kritischen Stimuli bestanden aus 28 farbigen Strichzeichnungspaaren, welche vom Autor der vorliegenden Arbeit selbst erstellt wurden (Beispiel im oberen Teil der Abbildung 6.1). Jedes Paar bestand aus einem Einführungsbild, auf dem zwei Figuren (z. B. eine *Maus* und eine *Schildkröte*) nebeneinander präsentiert wurden, und einem Target-Bild, das die beiden Figuren in einem transitiven Ereignis zeigte. Dabei wurde eine der beiden Figuren aus dem Einführungsbild als Initiator der Handlung dargestellt (z. B. *Eine Maus bespritzt eine Schildkröte*). Das Target-Bild erschien immer nach dem Einführungsbild. Zu jedem Paar wurden eine Einführungsaussage und eine Target-Frage auditiv präsentiert. Durch die Einführungsaussage wurden die Bezeichnungen der beiden im Einführungsbild präsentierten Figuren vorgestellt (z. B. *Auf diesem Bild siehst du eine Maus und eine Schildkröte*). Die anschließende Target-Frage richtete sich nach dem Inhalt des Target-Ereignisses und lautete wie folgt: *Sag mir bitte, was auf dem zweiten Bild passiert?* Alle Ereignisteilnehmer waren belebte Entitäten. Es gab immer zwei spiegelverkehrte Versionen von jedem Target-Bild – in einer Version befand sich das Agens immer auf der linken Seite, in der weiteren Version auf der rechten Bildseite.

Als Filler dienten 36 andersfarbige Strichzeichnungspaare (bestehend auch aus einem Einführungsbild und einem Target-Bild), welche ebenfalls vom Autor dieser Arbeit selbst erstellt wurden (Beispiel im unteren Teil der Abbildung 6.1). Das

Einführungsbild war dasselbe wie in den kritischen Stimuli, in denen zwei belebte Figuren einzeln präsentiert wurden (z. B. eine *Maus* und eine *Schildkröte*). In dem anschließenden Target-Bild wurden die gleichen Figuren noch einmal nebeneinander gezeigt. In einer Hälfte aller Filler-Durchgänge waren die Target-Bilder identisch mit den Einführungsbildern, während in der anderen Hälfte die Figuren in ihrer Position ausgetauscht wurden. Die Einführungsaussage war dieselbe wie in den Target-Stimuli (z. B. *Auf diesem Bild siehst du eine Maus und eine Schildkröte*). Hinsichtlich der Target-Frage wurde allerdings anstatt nach dem Inhalt des Target-Ereignisses nach der Positionsänderung der beiden Figuren gefragt: *Haben sich ihre Positionen auf dem zweiten Bild geändert?*

Für eine der Testphase vorangehende Übungsphase wurden 3 andere Trials vorbereitet (2 in der kritischen Bedingung und 1 in der Filler-Bedingung).

Neben den bildlichen und sprachlichen Materialien wurde noch ein exogener Cue verwendet, der – wie frühere Studien gezeigt haben – die frühe visuelle Aufmerksamkeit bzw. die erste Fixation von Versuchspersonen manipulieren kann. Der Cue erschien immer kurz vor der Präsentation eines Target-Bildes. In einer Hälfte der kritischen Trials erschien der Cue in der Region des Agens und in der anderen Hälfte in der des Patiens.

Angesichts der Position des Agens (entweder auf der linken oder rechten Seite) und der des visuellen Cues (entweder auf der Agens- oder Patiens-Position) wurden insgesamt 4 experimentelle Listen erstellt. Innerhalb jeder Liste gab es mindestens einen Filler-Trial zwischen zwei kritischen Trials.

6.2.4 Durchführung

Die Datenaufzeichnung des ersten Experimentes und aller weiteren Experimente wurde in dem Eye-Tracking-Labor des Instituts für Deutsch als Fremdsprachenphilologie der Universität Heidelberg durchgeführt. Die Versuchspersonen wurden einzeln aufgenommen. Zu Beginn der Untersuchung füllten die Versuchspersonen einen Fragebogen aus. Dieser erfasste demographische

Daten wie Name, Alter, Geschlecht und den Bildungsstand. Im Anschluss daran wurde sich nach der Sehfähigkeit der Versuchspersonen erkundigt. Nachfolgend wurde den Teilnehmern der gesamte Ablauf der Datenaufnahme mithilfe eines Beispiels (ähnlich wie Abbildung 6.1) auf Deutsch erklärt. Es wurde sichergestellt, dass die Versuchspersonen vor Beginn des Experimentes alle Instruktionen verstanden hatten. Anschließend wurden die Versuchspersonen gebeten, sich an einen Arbeitsplatz zu begeben, an dem das Experiment durchgeführt wurde. Die Versuchspersonen saßen vor einem 19-inch Flachbildschirm mit einer Auflösung von 1024×968 Pixeln. Die Position der Versuchspersonen wurde so eingestellt, dass der Abstand zwischen den Augen und dem Bildschirm ca. 25 Inch betrug und die Augen auf der Höhe des Fixationskreuzes auf dem Bildschirm waren. Nach der Einstellung wurden die Versuchspersonen instruiert, den Blick während des Experimentes stets auf den Bildschirm zu richten und mögliche Kopf- oder Körperbewegungen zu vermeiden. Die Kalibrierung der Blickmessung wurde vor der tatsächlichen Aufnahme durchgeführt. Dazu erschien auf dem Bildschirm ein sich bewegendes roter Punkt. Die Versuchspersonen wurden dazu aufgefordert, den Punkt mit den Augen zu fixieren und seine Bewegungen auf dem Bildschirm zu verfolgen, wodurch die Augenbewegungsmessung durch den Eyetracker durchgeführt wurde. Daraufhin wurden die Ergebnisse validiert. Nach der Kalibrierung und Validierung machten sich die Versuchspersonen in einer Übungsphase mit dem Experiment vertraut. Es wurden 3 Trials zur Übung durchgeführt, welche nicht in die Datenauswertung einbezogen wurden. Nach der Übungsphase wurden die experimentellen Trials in pseudo-randomisierter Reihenfolge abgespielt. Die Größe der Stimuli lag bei ca. 15.4 × 11 Inch. Bei einem Abstand von 25 Inch vom Bildschirm entfernt betrug die Größe der Stimuli ca. einen 34° × 24° Sehwinkel. Die Blickmessungen erfolgten durch den Eyetracker mit einer Sampling-Rate von 500Hz, und die Audio-Daten wurden von einem Ansteckmikrofon aufgezeichnet. Nach der Aufnahme erhielten die Versuchspersonen ihre Aufwandsentschädigung in Höhe von 7 Euro.

Ablauf eines Trials

Die Abbildung 6.1 zeigt ein Displaybeispiel der kritischen Bedingung und Filler-Bedingung. Jeder Durchgang der kritischen Trials begann mit der Präsentation eines Einführungsbildes. Während der Präsentation des Einführungsbildes hörten die Versuchspersonen zeitgleich eine Einführungsaussage und eine Target-Frage, die sich je nach dem Target-Ereignis richtete. Die Frage lautete *Sag mir bitte, was auf dem zweiten Bild passiert?*. Die Versuchspersonen sollten das Bild gut betrachten. Außerdem wurde vorher betont, dass sie den gestellten Fragen aufmerksam zuhören und ihre Antworten gemäß den Fragen planen sollten. Anschließend erschien ein Fixationskreuz in der Mitte des Bildschirms. Die Versuchspersonen wurden aufgefordert, das Fixationskreuz zu fixieren. Nach 500 ms wurde ein leerer Bildschirm gezeigt. Weitere 200 ms später blitzte der exogene Cue mit einer Größe von $1^\circ \times 1^\circ$ Sehwinkel für 80 ms auf. In einer Hälfte der kritischen Trials erschien der Cue in der Region, wo später das Agens des Target-Ereignisses auftrat, während in der anderen Hälfte der Cue in der Position des kommenden Patiens gezeigt wurde. Nach dem Darbietungsende des Cues erschien das Target-Bild. Die Versuchspersonen wurden gebeten, die zuvor gestellte Frage anhand des Target-Bildes in einem kompletten Satz so schnell wie möglich zu beantworten.

Der Ablauf der Filler-Trials gestaltete sich im Prinzip genauso wie in der kritischen Bedingung. Der Unterschied bestand nur darin, dass die in der Preview-Phase gestellte Target-Frage *Haben sich ihre Positionen auf dem zweiten Bild geändert?* lautete. Des Weiteren wurde in der Filler-Bedingung kein visueller Cue nach dem Blank dargestellt. Stattdessen wurde die Darstellungsdauer des Blank um 80 ms verlängert.

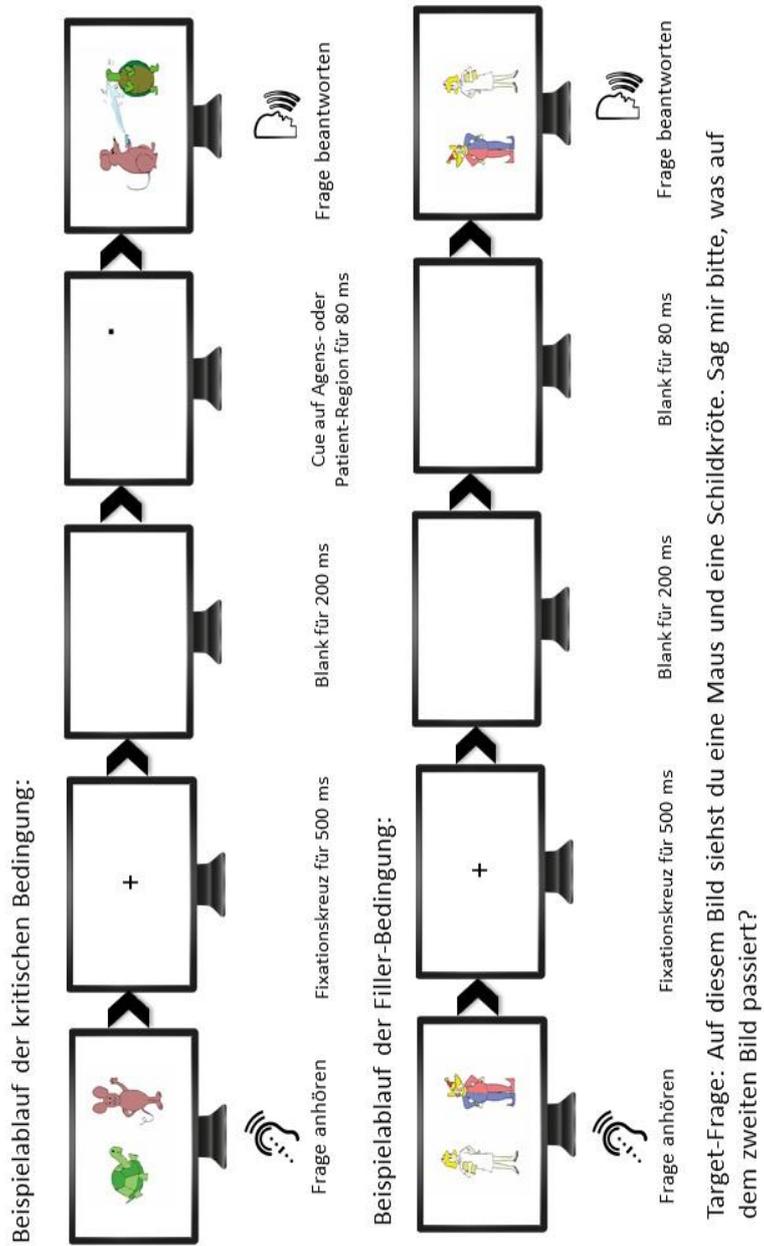


Abbildung 6.1 Beispielablauf der kritischen Bedingung und der Filler-Bedingung (Experiment 1)

6.2.5 Datenkodierung

Vor der statistischen Analyse mussten die aufgenommenen Daten in unterschiedlicher Hinsicht aufbereitet werden. Die für Experiment 1 relevanten Daten betreffen vor allem die syntaktischen Strukturen der produzierten Sätze, die Speech-Onset-Latenzen sowie die Blickbewegungen. Nachfolgend soll vorgestellt werden, wie die Daten organisiert und enkodiert werden.

Syntaktische Strukturen der produzierten Sätze

Die produzierten Antworten wurden von dem Autor der vorliegenden Arbeit transkribiert, anonymisiert und nach den Probanden-Nummern und Stimuli-Nummern in *Microsoft Excel* kodiert. Besonderes Gewicht der Analyse wurde auf die Selektion syntaktischer Strukturen (kanonischer Aktivsatz, Scrambled-Aktivsatz und Passivsatz) gelegt. Folgende Sätze mit folgenden Eigenschaften wurden von der Analyse ausgeschlossen (vgl. Abschnitt 6.3):

- Sätze mit Vertauschung des Satzanfangs und somit der Satzstruktur: *Die Schildkröte, uh, die Maus bespritzt die Schildkröte.*
- Sätze mit koordinierten NPs: *Die Katze und das Eichhörnchen laufen.*
- Sätze mit intransitiven Verben, in denen sich das Patiens nicht identifizieren ließ: *Die Katze läuft dem Eichhörnchen hinterher.*
- Sätze mit nur einem Ereignisteilnehmer: *Der Schwimmer streckt die Hand aus.*
- Antworten mit mehreren Sätzen: *Das Kind schläft und die Krankenschwester behandelt das.*

Speech-Onset-Latenz

Neben den syntaktischen Strukturen der produzierten Sätze wurde auch die Speech-Onset-Latenz analysiert. Alle erhobenen Daten einschließlich der Audio- und Blickbewegungsdaten wurden in die Software *SMI BeGaze* eingefügt. Mit der

Thumbnail Control von *SMI BeGaze* können für jedes Bild die Tonspuren der einzelnen Antworten, synchron beginnend mit dem Anfangszeitpunkt des Bildes, extrahiert werden. Da nicht alle Sätze mit einem inhaltlichen Wort angefangen wurden, wurde die Speech-Onset-Latenz anhand der visualisierten Tonspuren von Hand bestimmt. Damit wurde vermieden, dass Verzögerungslaute (z. B. *um*, *uh* oder *mhh*) als Sprechanfang eingestuft wurden.

Blickbewegungsdaten

Die aufgenommenen Blickbewegungsdaten können direkt mithilfe der Software *SMI BeGaze* ausgewertet werden. Für jedes kritische Target-Bild sind unter allen Blickbewegungsdaten zwei Indikatoren von großer Bedeutung - die sakkadische Zielposition und die Latenz zwischen der ersten und der zweiten Sakkade. In Bezug auf die sakkadische Zielposition wurde ausgewertet, wie häufig die erste und zweite Sakkade auf die durch den Cue angezeigte Figur oder die Figur mit dem Agens-Status gerichtet wurden. In Bezug auf die Latenz zwischen der ersten und der zweiten Sakkade wird das Zeitintervall zwischen dem Abschluss der ersten und dem Beginn der zweiten sakkadischen Bewegung, also die Dauer der ersten Fixation, errechnet.

6.2.6 Datenbearbeitung und -analyse

In einem ersten Schritt der Datenanalyse sollte der visuelle Cueingeffekt überprüft, der sich vor allem in der ersten Fixationsposition widerspiegelt. Zur Auswertung des Verhältnisses zwischen dem visuellen Cue und der Position der ersten Fixation wurde ein zweiseitiger Einstichproben-*t*-Test sowohl aus Sicht der Teilnehmer als auch der Stimuli durchgeführt. Ähnlich wie bei Gleitman et al. (2007) und Hwang und Kaiser (2015) wurde die Zufälligkeit von 50% als die vergleichende Basislinie ausgewählt, da ohne die Präsentation eines visuellen Cues zwei ereignisbildende Figuren mit der gleichen Chance zuerst fixiert werden können.

Basierend auf die Ergebnisse bezüglich des visuellen Cueingeffekts wird das Verhältnis zwischen der visuellen Salienz und der syntaktischen Struktur analysiert.

Die Auswertung erfolgt mit einer sog. logistischen Regressionsanalyse mit gemischten Effekten, in dem *die Position der vom visuellen Cue gesteuerten ersten Fixation (Agens- vs. Patiens-Region)* als feste unabhängige Variable und die *Stimuli* und *Versuchspersonen* als zufällige Faktoren fungierten. Die abhängige Variable ist die *syntaktische Struktur*. Die Art der Datenanalyse ermöglicht es, Zusammenhänge zwischen kategorialen Variablen zu analysieren, wobei sowohl *feste* als auch *zufällige* unabhängige Variable in Betracht gezogen werden (vgl. Baayen, 2008). Unter „festen“ unabhängigen Variablen werden diejenigen verstanden, deren Auswirkungen von zentralem Interesse eines Experimentes sind, hier in erster Linie die Position des visuellen Cues (Agens- vs. Patiens-Region). Eine zufällige unabhängige Variable wird stichprobenartig aus seiner Gesamtmenge ausgewählt und belegt. Das gilt insbesondere für die Auswahl von Versuchspersonen und Stimulus-Materialien. Da wir mit dem vorliegenden Experiment eine Aussage über alle deutschen Muttersprachler treffen wollen, die uns natürlich nicht alle zur Verfügung stehen, unterliegt die Selektion von Versuchspersonen dem Zufall und deren Ausprägungen gelten dementsprechend als zufällige Effekte. Hinsichtlich der Stimulus-Materialien wurde ebenfalls ein zufällig ausgewähltes Set von Trials aus einer unendlichen Gesamtheit selektiert. Mittels der logistischen Regressionsanalyse mit gemischten Effekten war es möglich, die Auswirkungen der festen unabhängigen Variablen von den zufälligen zu trennen.

Neben den syntaktischen Strukturen der produzierten Sätze bedient sich die Zielposition der zweiten Sakkade ebenfalls als Indikator für die Selektion von *starting points*. Für die Evaluierung des Verhältnisses zwischen der visuellen oder semantischen Salienz und der Zielposition der zweiten Sakkade wurde eine logistische Regressionsanalyse mit gemischten Effekten durchgeführt, in der *die Position der ersten Fixation (Agens vs. Patiens)* als feste unabhängige Variable eingeführt wird. Dazu werden die *Stimuli* und die *Probanden* als zufällige unabhängige Variablen mit einbezogen. Die abhängige Variable ist *das Auftreten eines Richtungswechsels bei der zweiten sakkadischen Ausführung*. Wenn die

semantische Struktur zum Zeitpunkt der zweiten sakkadischen Planung nicht aufgebaut worden wäre und die visuelle Salienz weitergehend eine zentrale Rolle bei der Vorausplanung spielt, dann dürfte kein Richtungswechsel passieren. Das bedeutet, dass die zweite Fixation weiterhin auf der zuerst fixierten Figur bleiben sollte. Wenn die semantische Struktur vor der zweiten Sakkade extrahiert würde, dann sollte eine Agens-Präferenz bei der Ausführung der zweiten Sakkade beobachtet werden.

Alle statistischen Auswertungen wurden in *R* (Version 1.0.143) vorgenommen. Als Signifikanzniveau wurde $p < .05$ gewählt.

6.3 Ergebnisse

Fehler

Insgesamt wurden 448 Sätze von den Probanden produziert. Davon wurden ca. 11% der erhobenen Sätze von der statistischen Analyse ausgeschlossen. Tabelle 6.1 zeigt genau die Entstehung der Fehler.

Fehlerkategorie	Anzahl
Technische Messfehler	14 (3.1%)
Fehler Performanz der Sätze insgesamt	36 (8.0%)
Fehler Performanz der Sätze detailliert	36 (100%)
Sätze mit Vertauschung	5 (14%)
Sätze mit intransitiven Verben	22 (61%)
Sätze mit nur einem Ereignisteilnehmer	1 (3%)
Sätze mit koordinierten NPs	3 (8%)
Antworten mit mehreren Sätzen	5 (14%)

Tabelle 6.1: Fehler (Experiment 1)

Visuelle/semantische Salienz und erste Fixation

Es konnte ein signifikanter Zusammenhang zwischen der Position des visuellen Cues

und dem ersten Fixationsort festgestellt werden: Unter allen Durchgängen wurden 69% der ersten Sakkaden auf den mit dem Cue markierten Charakter gerichtet. Diese Menge ähnelt den Daten der von Gleitman et al. durchgeführten Studie, bei der die Versuchspersonen in 70% aller Durchgänge die visuell markierte Figur als Erstes fixierten. Ein zweiseitiger Einstichproben-*t*-Test (separat für Probanden und Stimuli) wurde durchgeführt. Dies zeigt deutlich, dass die Proportion der ersten Fixation auf dem durch den Cue spezifizierten Charakter zuverlässig höher war, als durch den Zufall zu erwarten war (50%), $t_1(15) = 7.4722$, $p_1 < .01$, $t_2(27) = 8.7097$, $p_2 < .01$. Das lässt die Schlussfolgerung zu, dass die frühe Aufmerksamkeit der Versuchspersonen vom visuellen Cue manipuliert wurde und die Planung der ersten Sakkade bzw. Fixation vor allem in einem Bottom-Up-Prozess verläuft.

Eine Voraussetzung für die Wahl von 50% als Basislinie für den Vergleich ist die, dass die semantische Struktur vor der Planung der ersten Sakkade nicht extrahiert wurde. Sollten die semantischen Funktionen (hier Agens und Patiens) den Ereignisteilnehmern vor der Planung der ersten Sakkade zugewiesen werden, werden diese nicht mit der gleichen Chance fixiert. Die Begründung ist, dass das Agens mit einer höheren semantischen Salienz bei der Informationsverarbeitung privilegiert ist und im Gegensatz zum Patiens häufiger als Erstes fixiert werden kann. Um diese Bedenken über die Rationalität der selektierten Basislinie zu verwerfen, ist es notwendig, eine zusätzliche Analyse über die potenziellen Effekte der semantischen Funktion auf die erste Fixation durchzuführen. Mittels einer logistischen Regressionsanalyse mit gemischten Effekten wurde der Zusammenhang zwischen *der Position des Cues (Agens- vs. Patiens-Region)* und *dem Manipulationseffekt des visuellen Cues* analysiert. Wenn der Cue auf der Agens-Region präsentiert wurde, landeten 72.8% der ersten Fixation auf dem Agens. Wurde hingegen die Patiens-Region vorab vom visuellen Cue angezeigt, richteten die Versuchspersonen 65.2% ihrer ersten Sakkaden auf das Patiens (vgl. Abbildung 6.2). Die Ergebnisse der Statistik zeigten, dass die Position des Cues keinen signifikanten Einfluss auf den visuellen Cueingeffekt hat (vgl. Tabelle 6.2). Mit diesen Daten wurde die Gültigkeit

von 50% als Basislinie für den oben erwähnten *t*-Test bestätigt. Das bedeutet, dass bei der Planung der ersten Fixation potenzielle Auswirkungen der semantischen Struktur ausgeschlossen werden können.

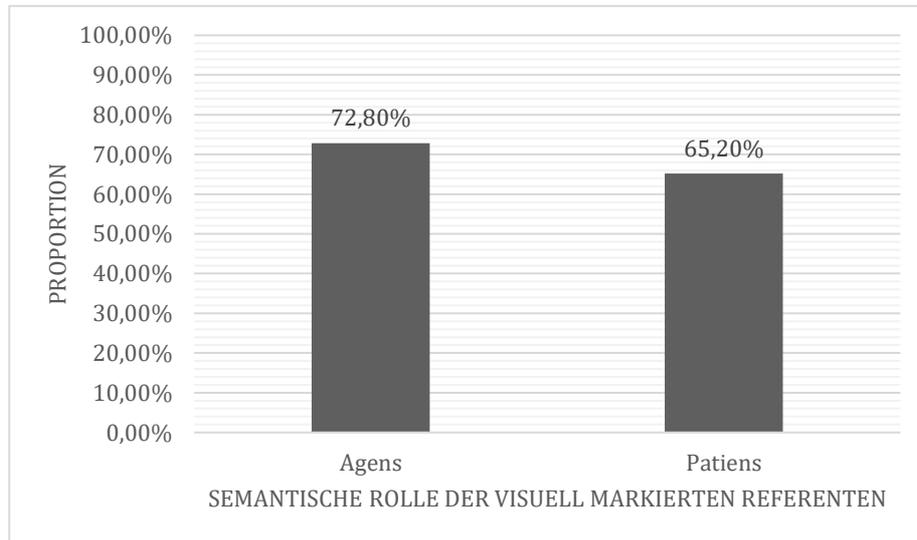


Abbildung 6.2 Proportion der ersten Fixationen, die auf visuell markierte Referenten gerichtet wurden, im Verhältnis zur semantischen Rolle der Referenten

Effekt	Est.	Std. Error	z-value
Intercept	0.8105	0.1232	6.579
Position des Cues	-0.1978	0.1071	-1.856

Tabelle 6.2: Logistische Regressionsanalyse auf die Abhängigkeit des visuellen Cueingeffekts von der semantischen Rolle der visuell markierten Referenten

Visuelle/semantische Salienz und Zielposition der zweiten Sakkade

Die Tatsache, dass bei der Planung der ersten Sakkade kein Effekt der semantischen Salienz beobachtet wurde, kann als Beweis dafür angesehen werden, dass die semantische Struktur eines Ereignisses während der Planung der ersten Sakkade noch nicht abgerufen wurde.

Wie sieht nun die zweite Sakkade aus, die ebenfalls als Indikator für die konzeptuelle Vorausplanung gilt? Um diese Frage zu beantworten, wurde der Zusammenhang zwischen der ersten und zweiten sakkadischen Zielposition analysiert. Ein Haupteffekt der ersten Fixationsposition (Agens- vs. Patiens-Region) auf die

Zielselektion der zweiten Sakkade wurde durch eine logistische Regressionsanalyse festgestellt (vgl. Tabelle 6.3). Wenn die erste Fixation auf dem Agens landete, richteten sich 16.3% der zweiten Sakkaden auf das Patiens. Wenn jedoch zuerst auf das Patiens geschaut wurde, wurden 69.4% der zweiten Sakkaden auf das Agens gerichtet (vgl. Abbildung 6.3). Unabhängig von der Position der ersten Fixation wurde in 79% der Trials das Agens als das Ziel der zweiten Sakkade selektiert. Aus diesem Ergebnis ergibt sich die Tatsache, dass die Planung der zweiten Sakkade von der semantischen Rolle der Ereignisteilnehmer beeinflusst wurde. Der Mittelwert des Zeitintervalls zwischen der ersten und der zweiten Sakkade lag bei 155.9 ms.

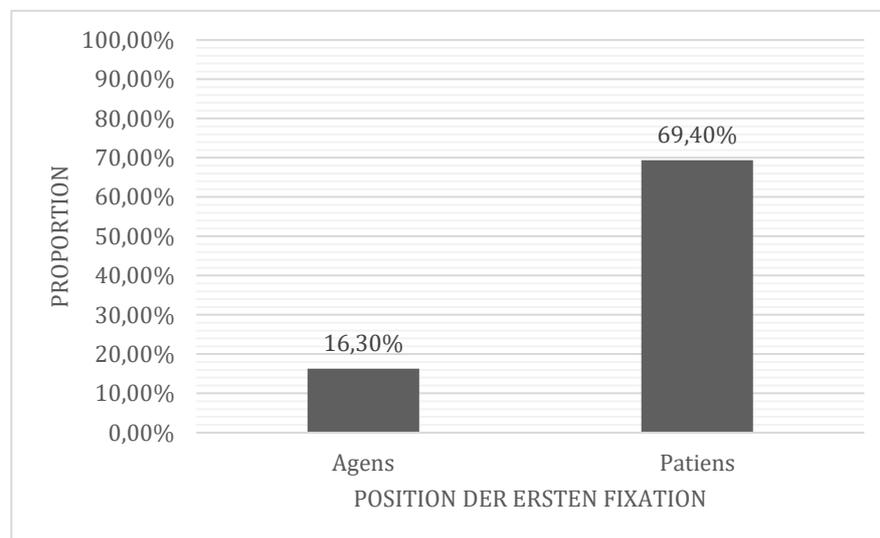


Abbildung 6.3: Proportion der zweiten Sakkaden, die auf einen anderen Referenten als den zuerst fixierten gerichtet wurden, im Verhältnis zur Position der ersten Fixationen

Effekt	Est.	Std. Error	z-Value
Intercept	-0.4373	0.2197	-1.990
Position der EF	1.2539	0.2190	5.726***

Tabelle 6.3: Logistische Regressionsanalyse auf die Abhängigkeit der Zielselektion der zweiten Sakkaden von der Position der ersten Fixationen

Visuelle/semantische Salienz und syntaktische Struktur

Wenn die erste Fixation auf dem Agens landete, waren 99% der produzierten Sätze aktiv. Ferner ist es wichtig anzumerken, dass keine Passivsätze gebildet wurden, wenn das Patiens zuerst fixiert wurde (vgl. Abbildung 6.4). Zudem wurden in den

beiden Bedingungen keine Scrambled-Aktivsätze mit OVS-Struktur produziert. Logistisch regressionsanalytisch ließ sich kein Zusammenhang zwischen der Position der ersten Fixation und der selektierten syntaktischen Strukturen feststellen (vgl. Tabelle 6.4). Die Ergebnisse sprechen daher dafür, dass, obwohl die frühe Aufmerksamkeit bzw. die erste Fixation von der visuellen Salienz bestimmt wurde, ihre Effekte nicht bis zur grammatischen Verarbeitungsphase aufrechterhalten wurden. Die erste Fixation auf eine Figur führte nicht zur häufigeren Produktion von Sätzen, in denen die an erster Stelle fixierte Figur am Satzanfang steht. Stattdessen wurde das Agens immer als der *starting point* der Verbalisierung selektiert und satzinitial angeordnet.

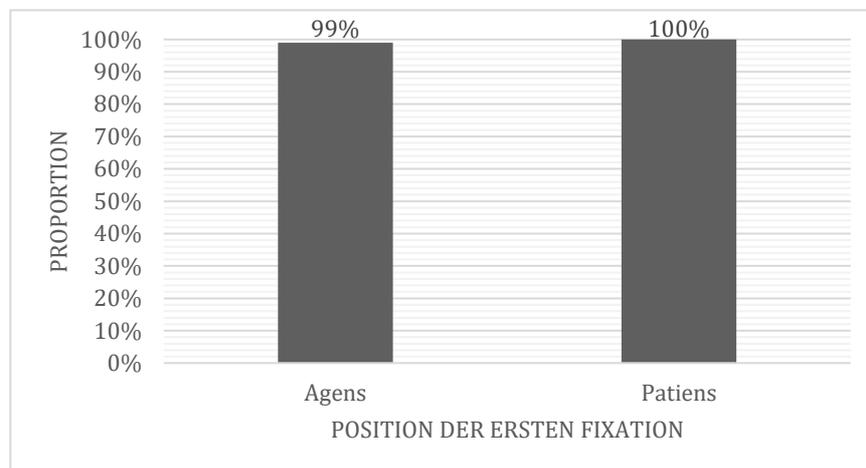


Abbildung 6.4: Proportion produzierter kanonischer Aktivsätze im Verhältnis zur Position der ersten Fixationen

Fixed-Effekt	Est.	Std. Error	z-value
Intercept	-58.56	330.10	-0.177
Position der EF	-22.25	330.09	-0.067

Tabelle 6.4: Logistische Regressionsanalyse auf die Abhängigkeit der syntaktischen Selektion von der Position der ersten Fixationen

6.4 Diskussion

In diesem Experiment wurde untersucht, ob die deutsche Satzproduktion in einer pragmatischen neutralen Situation als linear oder strukturell inkrementeller

Verarbeitungsprozess beschrieben werden kann. Zur Beantwortung dieser Frage wurde besonderes Gewicht auf die Größe der konzeptuellen Vorausplanung gelegt, die sich in zwei Indikatoren – die selektierten syntaktischen Strukturen und die sich parallel zur Entwicklung der Formulierung vollziehenden Augenbewegungen – zeigen kann.

Der linearen Inkrementalitätshypothese zufolge wird die Selektion von *starting points* überwiegend von der Eigenschaft individueller Referenten, wie zum Beispiel ihre Belebtheit oder visuelle Auffälligkeit, beeinflusst. Im Gegensatz hierzu ist die strukturelle Inkrementalitätshypothese der Auffassung, dass diejenigen Referenten, die eine höhere semantische Salienz implizieren, eher als *starting points* der Verbalisierung ausgewählt werden. Das bedeutet also, dass bei dieser Hypothese der Effekt der Ereignisstruktur im Vordergrund steht. Daher wurden aufgrund dieses Unterschiedes der beiden Annahmen in diesem Experiment die visuelle und die semantische Salienz einander gegenübergestellt. Es galt zu beobachten, welchem Faktor eine zentrale Bedeutung bei der Selektion von *starting points* bzw. bei der syntaktischen Enkodierung zukommt.

Um die semantische Salienz zu manipulieren, wurden transitive Ereignisse ausgewählt. Es traten jeweils immer ein belebtes Agens und ein belebtes Patiens auf, die gemeinsam eine Handlung ausführten, die sich mit einem transitiven Verb beschreiben ließ (z. B. *bespritzen, jagen, boxen, aufwachen, treten, erschießen usw.*). Dass das Agens semantisch salienter als das Patiens ist und daher bei der Informationsverarbeitung bevorzugt behandelt wird, konnten frühere Studien schon zeigen (z. B. Dobel et al., 2007).

Ein exogener Cue wurde kurz vor der Präsentation des Target-Ereignisses auf die Position eines Referenten gesetzt, um wiederum die visuelle Salienz dieses Referenten zu beeinflussen. Das Resultat vorangegangener Studien war, dass jene Entität, die durch den visuellen Cue markiert wurde, dadurch mehr anfängliche visuelle Aufmerksamkeit auf sich gezogen hatte als die andere und dadurch früher

enkodiert wurde (z. B. Gleitman et al., 2007; Hwang & Kaiser, 2015; Myachykov & Tomlin, 2008).

Wenn die Satzproduktion im Deutschen nun ein linear inkrementell verlaufender Prozess ist, sollten diejenigen Figuren, denen mehr visuelle Aufmerksamkeit gewidmet wurde, als angemessene *starting points* der Verbalisierung ausgewählt werden. Wenn dieser Prozess allerdings strukturell inkrementell verläuft, dann sollte die Selektion dieser *starting points* nicht auf den visuellen Eigenschaften einzelner Ereignisteilnehmer, sondern auf der semantisch strukturellen Repräsentation des Ereignisses beruhen.

Von diesen beiden Hypothesen ausgehend, sollen nun nachfolgend die aufgezeichneten visuellen und sprachlichen Daten diskutiert werden. Die sprachspezifischen Unterschiede zwischen den Befunden der deutschen und englischen Satzplanung sollen erklärt und begründet werden. Zusätzlich zur Diskussion über die Arbeitsweise des Satzplanungssystems soll auch mithilfe der aufgezeichneten sakkadischen Parameter auf den Zeitverlauf der semantischen Strukturierung eingegangen werden. Im Fokus steht dabei die Frage, wie lange es dauert, bevor deutsche Sprecher die semantische Struktur aus dem zu beschreibenden transitiven Ereignis entnommen haben. Abschließend wird dann die Stärke des visuellen und semantischen Salienzeinflusses miteinander verglichen, damit daraus eine Prioritätenliste der bei der Satzplanung mitwirkenden Faktoren entwickelt werden kann.

Lineare vs. strukturelle Inkrementalität

In Bezug auf den Effekt der visuellen Salienz wurde in erster Linie festgestellt, dass die erste Fixation wesentlich öfter auf die gezeichnete Entität gesetzt wurde. Dieses Ergebnis gleicht der von Gleitman et al. (2007) gefundenen Tendenz, dass die anfängliche visuelle Aufmerksamkeit bzw. der Wahrnehmungsprozess der Ereigniskodierung von der visuellen Eigenschaft individueller Referenten manipuliert werden kann.

Da die *low-level* perzeptive Enkodierung in die *high-level* konzeptuelle Verarbeitung münden muss, wurde weiterhin geprüft, ob die visuelle Salienz auch einen Einfluss auf die Selektion von *starting points* hat. Die Annahme in diesem Zusammenhang ist, dass die Satzproduktion inkrementell abläuft und die Entitäten, sobald sie als *starting points* der Verarbeitung ausgewählt wurden, umgehend an die grammatische Enkodierungsebene weitergereicht und dort enkodiert werden (vgl. Kempen & Hoenkamp, 1987). Dann erlauben es die konstruierten syntaktischen Strukturen, das Verhältnis zwischen der visuellen Aufmerksamkeit und der Auswahl von *starting points* zu identifizieren. Das wirft die Frage auf, ob eine Übereinstimmung zwischen dem zuerst fixierten und dem zuerst verbalisierten Referenten besteht.

Ein Ergebnis des ersten Experimentes ist, dass die initiale visuelle Aufmerksamkeit keinen Einfluss auf die Selektion von *starting points* hatte. Das Verhältnis der gebildeten Passivsätze erhöhte sich nicht, wenn zuerst das Patiens fixiert wurde. Fast alle produzierten Sätze - unabhängig davon, auf welche Figur die erste Fixation gelenkt wurde - waren kanonische Aktivsätze. Ferner wurden keine Scrambled-Aktivsätze produziert, in denen das Patiens in Form eines Objekts satzinitial in Erscheinung trat. Diese Ergebnisse über die selektierten syntaktischen Strukturen können daher als erster Hinweis darauf gedeutet werden, dass deutsche Sprecher tendenziell das Agens als *starting point* der Satzplanung auswählen. Obwohl die overte Aufmerksamkeit gleich nach der Bildpräsentation auf den durch den Cue markierten Referenten gerichtet wurde, wurde dieser Referent nicht weiter konzeptuell verarbeitet, außer wenn sich dieser nach der Extraktion der semantischen Struktur als das Agens herausgestellt hatte. Dieses Prinzip gilt allerdings nur für die Fälle, in denen keiner der beiden Referenten explizit als Diskurstopik markiert ist. Nachher in Experiment 2 werden wir sehen, wie diese sog. Agens-Präferenz von der kontextuellen Information modifiziert werden kann.

Die fehlende Korrespondenz zwischen dem zuerst fixierten und dem zuerst erwähnten Referenten und die durchgängige Anordnung des Agens in satzinitialer Position gleichen den Befunden von Myachykov et al. (2010) und Hwang und Kaiser

(2015), bei denen ebenfalls kein visueller Cueing-Effekt auf die syntaktische Konstruktion im Finnischen und Koreanischen gefunden wurden. Unsere Ergebnisse zur Satzproduktion im Deutschen liefern somit Evidenz dafür, dass die Rolle der visuellen Aufmerksamkeit, die in der Studie von Gleitman et al. (2007) als ausschlaggebender Faktor für die englische Satzplanung identifiziert wurde, wohl nicht universal gleich, sondern von Sprache zu Sprache verschieden ist.

Obwohl die oben beschriebenen Ergebnisse die Gültigkeit der linearen Inkrementalität zur Beschreibung des deutschen Satzproduktionsvorgangs in Frage stellen, erlauben sie aber auch keine klare Deutung für die strukturell inkrementelle Verarbeitungsweise. Um eine präzisere und kräftigere Aussage bezüglich der Mechanismen der Satzplanung im Deutschen treffen zu können, konzentrieren wir uns daher auf die Zielpositionen der zweiten Sakkaden.

Im Gegensatz zur ersten Sakkade ist die zweite Sakkade in der Tat die erste selbstbestimmte und freiwillig ausgeführte Augenbewegung, die nicht automatisch und reflexartig durch einen unerwarteten visuellen Cue verursacht wurde, sondern einer Top-Down-Verarbeitung von Sprechern zuzuschreiben ist. Aufgrund dieses Unterschiedes lässt sich ein Vergleich zwischen den Zielpositionen der beiden Sakkaden ziehen, um festzustellen, ob die visuelle Salienz an der Planung der zweiten Sakkade beteiligt ist, oder die semantische Salienz nun als Faktor wirkt.

Zudem zeigte sich, dass, wenn der Blick der Sprecher visuell bedingt zuerst auf das Patiens fiel, die zweite Sakkade dann größtenteils auf das Agens gerichtet wurde. Fiel die erste Sakkade jedoch auf das Agens, verharrten die Versuchspersonen mit ihrem Blick zum überwiegenden Teil auch weiterhin auf dem Agens. Dieser Befund lässt die Schlussfolgerung zu, dass, anstatt dass deutsche Sprecher ihren Blick auf die visuell salientere Entität richteten, sie tendenziell diejenige Figur anschauten, deren semantische Funktion am prominentesten war - in diesem Falle also das Agens.

Dieser Befund widerspricht der von der linearen Inkrementalitätshypothese ausgehenden Ansicht, nach der Sprecher ihren Blick von demjenigen Referenten, auf

den eine reizbezogene Sakkade gerichtet wurde, nicht abwenden. Anhand der zweiten Sakkade lässt sich feststellen, dass das Agens mit seiner höheren semantischen Salienz deutlich mehr Anziehungskraft ausübt als das Patiens. Da die Ausprägung des Agens auf die Augenbewegungen den Aufbau der semantischen Struktur zur Voraussetzung hat, kann man aus den Blickbewegungsdaten schlussfolgern, dass die semantische Struktur, oder mindestens das Agens, bereits in der konzeptuellen Vorausplanungsphase abgerufen wurde.

Aus dieser Analyse der visuellen und linguistischen Daten lässt sich folgendes Fazit ziehen: Obwohl die visuelle Salienz einen Einfluss auf die visuelle Wahrnehmung hat, dauert dieser nicht bis zur Selektion von *starting points* an. Stattdessen steht die Selektion im ursächlichen Zusammenhang mit der semantisch-konzeptuellen Struktur des Ereignisses. Daher lässt sich sagen, dass die deutsche Satzproduktion in einer pragmatisch neutralen Situation ein strukturell inkrementell verlaufender Prozess ist.

Vergleich mit vorangegangenen Studien

Unsere Befunde, dass der Selektion von *starting points* der Aufbau der strukturellen Repräsentation des Gesamtereignisses zugrunde liegt und die Satzplanung im Deutschen daher in einer strukturell inkrementellen Arbeitsweise abläuft, weichen von den Ergebnissen von Gleitman et al. (2007) ab. Die Autoren sehen einen Zusammenhang zwischen der visuellen Eigenschaft der Referenten und der Selektion von *starting points* der Verbalisierung und unterstützen somit die lineare Inkrementalitätshypothese. Aufgrund dieses einzelsprachbezogenen Unterschiedes zwischen den Satzproduktionsmechanismen im Deutschen und Englischen ist nun die Frage, worin der Grund dafür liegt.

Ein möglicher Grund für den Unterschied liegt in dem typologischen Unterschied bezüglich des Verhältnisses zwischen Satzpositionen und syntaktischen Funktionen im Deutschen und Englischen. Die englische Wortstellung ist relativ fest und erfordert prinzipiell die Abfolge Subjekt-Verb-Objekt (SVO). Die relativ

unveränderliche Wortstellung ist für die syntaktische Strukturierung von großer Bedeutung, da mit der Festlegung des zuerst zu äußernden Konzepts die syntaktische Struktur bzw. die Wortstellung dementsprechend festgelegt wird. Wenn beispielsweise bei der Satzplanung im Englischen eine Figur mit der Patiens-Rolle aufgrund ihrer höheren visuellen Salienz die initiale visuelle Aufmerksamkeit auf sich gezogen hat und danach als das erste zu verbalisierende lexikalische Konzept selektiert wird, wird die syntaktische Struktur bzw. die Wortstellung des zu bildenden Satzes automatisch generiert, d.h., ein Passiv- oder Ergativsatz wird produziert. Aufgrund dieser Eigenschaft wird die syntaktische Enkodierung beschleunigt, wenn der erste zu äußernde Referent aufgrund seiner höheren visuellen Salienz umgehend nach Stimulus-Onset als *starting points* ausgewählt und syntaktisch verarbeitet wird.

Im Gegensatz zum Englischen ist Deutsch eine sog. Kasus-Sprache, in der syntaktische Funktionen weniger mit den Satzpositionen zu tun haben, sondern viel mehr durch eine Kasusmarkierung sichtbar werden. Die Verfügbarkeit der Kasusmarker ermöglicht daher eine *flexible* Wortstellung innerhalb deutscher Sätze. So muss eine bestimmte Position (z. B. die satzinitiale Position) nicht unbedingt an eine bestimmte syntaktische Funktion (z. B. das Subjekt) gebunden sein. Das Patiens kann, neben der Subjekt-Funktion in einer Passivstruktur, auch als das Objekt in einer Aktivstruktur realisiert sein.

Diese freie Verbindung hat zur Folge, dass deutsche Sprecher nach der Selektion des ersten lexikalischen Konzeptes eine weitere Entscheidung über dessen syntaktische Funktionen treffen müssen. Aufgrund dieser grammatischen Eigenschaft der deutschen Sprache stellen wir uns nun vor, wie der Verarbeitungsprozess aussähe, wenn deutsche Sprecher bei der Satzplanung die lineare inkrementelle Verarbeitungsweise verfolgen würden, bei der die Selektion des ersten lexikalischen Konzeptes visuell bedingt ist und semantische Funktionen erst in einer späteren Phase zugewiesen werden. Wir nehmen die Selektion eines Referenten mit der Patiens-Rolle als *starting point* zum Beispiel.

Wenn ein Referent reizbedingt als *starting point* ausgewählt wird und dieser sich zu einem späteren Zeitpunkt nach dem Aufbau der semantischen Struktur als das Patiens des Ereignisses herausstellt, dann haben deutsche Sprecher zur syntaktischen Enkodierung dieses Referenten zwei Alternativen: Entweder wird er als Subjekt realisiert und satzinitial angeordnet, was zur Bildung eines Passivsatzes führen kann (z. B. *Ein Hund wird von einer Katze gejagt*), oder er wird der Objekt-Funktion zugeordnet, wodurch ein Scrambled-Aktivsatz produziert wird (z. B. *Einen Hund jagt eine Katze*). Obwohl Sprecher aufgrund der Zugänglichkeitshierarchie im Allgemeinen dazu tendieren, die Subjekt-Funktion zuerst zuzuteilen, was in dieser Situation eine Präferenz für die Auswahl der Passivstruktur verursachen kann, ist eine Zuweisung der Objekt-Funktion und somit die Aktivierung der Scrambled-Aktivform nicht unmöglich. So kann es sein, dass anstatt wie bei der englischen syntaktischen Enkodierung ein Beschleunigungseffekt mit der reizgetriebenen Selektion von *starting points* ausgelöst wird, eine neue Konkurrenz hinsichtlich der Zuteilung syntaktischer Funktionen an dem zuerst zu erwähnenden Wort hervorgerufen werden kann, was die grammatische Enkodierung erschweren und verlangsamen könnte.

Diese Beeinträchtigung der Satzplanung können deutsche Sprecher dadurch vermeiden, indem sie die zu verbalisierende Information in einer strukturell inkrementellen Art und Weise verarbeiten und das Agens als *starting point* der Satzplanung selektieren. Es sei daran erinnert, im Gegensatz zum Patiens ist die syntaktische Funktion des Agens an der ersten Stelle eines Satzes relativ fest. Wenn das Agens als das erste lexikalische Konzept ausgewählt wird, kann ihm im unmarkierten Fall nur die Subjekt-Funktion zugewiesen werden.⁶ Somit wird die syntaktische Struktur im Deutschen mit der Festlegung des Agens als *starting point* der Satzplanung automatisch konstruiert, d.h. es wird ein kanonischer Aktivsatz

⁶ Eine Voranstellung des Agens in Form eines Objekts ist im Deutschen zwar möglich (z. B. *Von der Katze wird der Hund gejagt*), sie ist allerdings eine Form der Topikalisierung. Denn die Satzproduktion in diesem Experiment findet in einer pragmatisch neutralen Situation statt, wird diese Konstruktion nicht berücksichtigt.

gebildet. Weitere Entscheidungen über die Zuteilung syntaktischer Funktionen an das erste Konzept sind in diesem Rahmen nicht mehr nötig. Aufgrund dieser grammatischen Beschränkung im Deutschen ist es daher sinnvoll, während der Satzproduktion mehr Aufmerksamkeit auf die semantische Struktur als auf die visuelle Information zu richten und das Agens als *starting point* auszuwählen. Dadurch wird die Satzproduktion nicht von Aktivierungen verschiedener syntaktischer Strukturen beeinträchtigt.

Eine andere Studie, die hier ebenfalls zum Vergleich aufgeführt wird, ist von Kuchinsky und Bock (2010). Kuchinsky und Bock kommen in ihrer Studie zu dem Ergebnis, dass die lineare und die strukturelle Inkrementalität nicht zwei sich gegenseitig ausschließende Arbeitsweisen des Satzproduktionssystems sind, sondern beide Sprechern zur Verfügung stehen. Welche Strategie letztendlich für die Satzproduktion ausgewählt wird, ist von der relativen Kodierbarkeit individueller Charaktere und der Ereignisstruktur abhängig. In Bezug auf die Kodierbarkeit sprechen die Autorinnen von der Zeit, die zum Verständnis einer Entität oder einer Struktur benötigt wird. Sie fanden heraus, dass, wenn die strukturelle Repräsentation eines Ereignisses schwer kodierbar war, sich englische Sprecher umgehend der Identifizierung individueller Referenten bzw. Konzepten widmeten, um einen flüssigen Sprechfluss garantieren und eine mögliche Beeinträchtigung abwenden bzw. vermeiden zu können. Diese Strategie, dass leichtere Informationen früher verarbeitet werden, damit sie sich nicht im Arbeitsgedächtnis akkumulieren und zu hohem kognitivem Aufwand führen, findet sich im sog. *minimal-load principle* für die Sprachproduktion von Levelt (1989): „When there are more alternatives, the principle says, do the simplest thing first“ (S. 144). Aber ist eine ähnliche Tendenz auch im vorliegenden Experiment zu beobachten?

Zur Beantwortung dieser Frage betrachten wir das Versuchsdesign unseres Experimentes genauer. Im Gegensatz zu vielen vorangehenden Studien, in denen die Ereignisse nur isoliert und ohne Kontext präsentiert wurden, wurden die deutschen Probanden in diesem Experiment vor der Beschreibung des Target-Ereignisses zuerst

einer sog. Preview-Phase ausgesetzt. In dieser Phase wurden die Ereignisteilnehmer sowohl visuell als auch lexikalisch eingeführt. Diese Preview-Phase führte dann zur Steigerung sowohl der perzeptuellen als auch der lexikalischen Verfügbarkeit der Ereignisteilnehmer. So sind sie im Vergleich zur Ereignisstruktur, die erst nach der Darbietung des Target-Bildes extrahiert werden kann, einfacher zu enkodieren. Gemäß dem *minimal-load principle* von Levelt und nach der Voraussage von Kuchinsky und Bock sollten die Versuchspersonen die Verarbeitung in diesem Zusammenhang eher mit den individuellen Charakteren als den strukturellen Repräsentationen des Target-Ereignisses beginnen. Das ist aber hier nicht der Fall. Abgesehen von der Voraktivierung der visuellen und lexikalischen Information der Ereignisteilnehmer, konstruierten deutsche Sprecher zunächst immer eine holistische Ereignisrepräsentation, bevor sie die *starting points* selektierten. Diese Ergebnisse stehen somit im Kontrast zu der Aussage von Levelt (1989) bzw. den Befunden von Kuchinsky und Bock (2010) und deuten darauf hin, dass der Aufbau einer holistischen Interpretation des Gesamtereignisses wohl obligatorisch für die Selektion von *starting points* bzw. für den Beginn der grammatischen Enkodierung im Deutschen ist.

Basierend auf der obigen Analyse kann geschlussfolgert werden, dass im Gegensatz zu englischen Sprechern deutsche Sprecher während der Satzproduktion wenig von der visuellen Salienz und lexikalischen Kodierbarkeit individueller Referenten beeinflusst werden. Stattdessen verfolgen sie eine strikte strukturelle inkrementelle Verarbeitungsweise, bei der die semantisch-konzeptuelle Struktur eine dominante Rolle spielt.

Zeitverlauf der semantischen Strukturierung

In Bezug auf den Konflikt bzw. den Gegensatz der linearen und der strukturellen Inkrementalität steht auch der Zeitverlauf der semantischen Extraktion im Fokus. Da die Planung der zweiten Sakkade in den Zeitraum der konzeptuellen Vorausplanung fällt, lässt sich aus dem Befund, dass die semantische Salienz bzw. das Agens einen

Einfluss auf die Positionierung der zweiten Fixation hat, die Erkenntnis ableiten, dass die Konstruktion der semantischen Struktur (oder präziser gesagt, die Zuweisung der Agens-Rolle) bereits in der Vorausplanungsphase abgeschlossen ist. Griffin und Bock (2000) zufolge erstreckt sich die konzeptuelle Vorausplanung auf einen Zeitraum von ca. 400 ms nach Stimulus-Onset, wodurch die semantische bzw. konzeptuelle Struktur wiederum innerhalb dieser Zeitspanne aufgebaut werden kann.

Ist es jedoch noch möglich, aufgrund des Verhältnisses zwischen der zweiten Sakkade und der semantischen Konstruktion das Zeitfenster der semantischen Strukturierung weiter einzugrenzen? Um diese Frage beantworten zu können, ist das intersakkadische Intervall zwischen der ersten und zweiten Sakkade von großer Bedeutung, da es die für die Auslösung der zweiten Sakkade in Anspruch genommene Zeit reflektieren kann. Infolge dessen, dass die zweite Sakkade semantisch gesteuert ist und deren Planung daher den Aufbau der semantischen Struktur voraussetzt, können wir durch die Festlegung des initialen Zeitpunkts dieser Sakkade auf das Zeitfenster der semantischen Strukturierung rückschließen.

Das durchschnittliche temporale Intervall zwischen der ersten und zweiten Sakkade lag in diesem Experiment bei 155.9 ms. Da frühe Studien gezeigt haben, dass zur Planung und Ausführung einer Sakkade in der Regel 200 - 300 ms benötigt werden (vgl. Fischer & Weber, 1993; Saslow, 1967), war die hier gemessene Planungszeit für die zweite Sakkade deutlich geringer als die normalerweise zur Generierung einer Sakkade erforderliche Zeit. So ergibt sich die Frage, worauf das verkürzte Zeitintervall zurückzuführen und welche Erkenntnis daraus abzuleiten ist.

Zur Erklärung der verkürzten intersakkadischen Latenz nehmen viele Psychologen an, dass sie ein Resultat der parallelen Planung einer vorigen und einer nachfolgenden Sakkade ist (z. B. McPeck, Skavenski & Nakayama, 2000; McPeck & Keller, 2002). Das bedeutet, dass Sprecher bereits vor dem Abschluss der ersten sakkadischen Bewegung angefangen haben, die zweite zu planen, was zu einer zeitlichen Überlappung und somit einer verkürzten intersakkadischen Latenz der

ersten und zweiten sakkadischen Planung führt. Aufgrund dieses Befunds, dass die zweite Sakkade semantisch angetrieben war, können wir aus der planerischen Zeitüberlappung der beiden Sakkaden schlussfolgern, dass die semantischen Funktionen, oder mindestens die Agens-Rolle, vor dem Abschluss der ersten Sakkade wahrscheinlich bereits zugewiesen wurden (vgl. Abbildung 6.4).

Von dieser Schlussfolgerung ausgehend, stellt sich nun eine weitere Frage, nämlich inwiefern die Planung der ersten und zweiten Sakkade überlappend war. Da die Sakkadengenerierung – wie in Abschnitt 5.3.2 eingeführt wurde – in drei Phasen (nämlich Verarbeitung visueller Informationen, *dead-time* und Ausführung) verläuft, kann die Frage wie folgt konkretisiert werden: In welcher Planungsphase der ersten Sakkade begann die Planung der zweiten Sakkade?

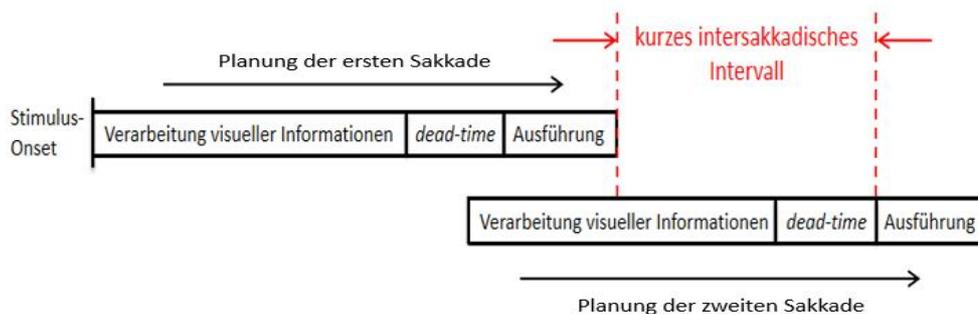


Abbildung 6.5: Simultane Programmierungen der ersten und zweiten Sakkade

Die Initiierung der zweiten sakkadischen Planung während der letzten Phase der ersten Sakkade, also der Ausführungsphase, wird zunächst ausgeschlossen. Mehrere Studien zur visuellen Verarbeitung haben gezeigt, dass die Verarbeitung räumlicher Informationen, welche ausschlaggebend für die Festlegung der sakkadischen Zielposition ist, während des sog. Blick-Überspringens blockiert wird. Für diese Schlussfolgerung liefern beispielsweise Iriwn und Brockmole (2004) eine Evidenz. In ihrer Studie wurden Versuchspersonen Bild-Stimuli mit Objekten (z. B. ein Vogel, ein Stuhl, ein Fahrrad usw.), die immer in einer aufrechten Orientierung erschienen, dargestellt und sie gebeten zu entscheiden, ob die Objekte nach links oder rechts ausgerichtet waren. Währenddessen mussten sie entweder eine lange oder eine kurze

Sakkade durchführen. Da die Entscheidung über die Richtung der Objekte die Enkodierung räumlicher Informationen in Anspruch nimmt bzw. benötigt, lässt sich aus den Reaktionszeiten auf die räumlichen Stimuli-Eigenschaften herleiten, ob die sakkadischen Übersprünge einen Einfluss auf die räumliche Verarbeitung haben. Es zeigte sich, dass die Versuchspersonen deutlich mehr Zeit für die Beurteilung benötigten, wenn sie zeitgleich eine lange Sakkade als eine kurze durchführten. Dieses Ergebnis lässt die Schlussfolgerung zu, dass die Verarbeitung räumlicher Informationen während der sakkadischen Bewegung blockiert wurde. Da die Selektion sakkadischer Zielpositionen ebenfalls eine Identifikation eines Raumpunktes benötigt, wird angenommen, dass die zweite sakkadische Planung bereits vor der Ausführungsphase der ersten Sakkade initiiert wurde. Da die Planung der zweiten Sakkade den Aufbau der semantischen Struktur voraussetzt, ist es mit dieser Analyse möglich, das Zeitfenster der semantischen Aktivierung auf die Zeitspanne vor der ersten sakkadischen Ausführung einzugrenzen, was wiederum zwei weitere Interpretationsmöglichkeiten zulässt.

Die erste Möglichkeit ist, dass die Planung der zweiten Sakkade während der *dead-time*-Phase initiiert wurde. Sobald die Planung der ersten Sakkade in die *dead-time*-Phase übergegangen ist, wird nämlich deren Ausführung „versiegelt“. Das bedeutet, dass die Auslösung der ersten Sakkade nicht von der Planung einer weiteren gestoppt oder zurückgezogen wird. Dies steht im Einklang mit unserer Beobachtung – nämlich, dass trotz der parallelen Planung die erste reizgesteuerte Sakkade aufrechterhalten und nicht von der zweiten Sakkade, die semantisch gesteuert war, beeinflusst wurde. Dies bedeutet auch, dass Sprecher die semantische Ereignisstruktur während des Zeitintervalls (zum Teil) aufgebaut haben.

Die zweite Möglichkeit ist, dass die Extraktion der semantischen Struktur bereits vor der *dead-time*-Phase der ersten Sakkade abgeschlossen wurde. Diese Hypothese halten wir allerdings für unmöglich. In Anbetracht dessen, dass vor der *dead-time*-Phase die Planung der ersten Sakkade noch nicht „immun“ gegen die Beeinflussung anderer Faktoren wie etwa der semantischen Salienz ist, sollten wir

eine Agens-Präferenz bei der Platzierung der ersten Fixation vermuten, sobald die semantisch strukturelle Repräsentation in dieser Phase aufgebaut worden ist. Dies ist – wie die aufgezeichneten Blickbewegungsdaten gezeigt haben – nicht der Fall. In diesem Experiment wurde kein Einfluss der Agens-Rolle auf die Positionierung der ersten Fixation beobachtet, was darauf hindeutet, dass die Zuweisung der Agens-Rolle in der ersten Planungsphase, also vor Beginn der *dead-time*-Phase, noch nicht konstruiert wurde.

Resümierend liefert das kurze intersakkadische Intervall zwischen der ersten und der zweiten Sakkade einen Hinweis auf eine zeitliche Überlappung zwischen den Planungen der beiden Sakkaden. Da die Planung der zweiten Sakkade vor allem von der semantischen Struktur angetrieben ist, führt dieses Ergebnis zu der Schlussfolgerung, dass der Aufbau der semantischen Struktur eines Ereignisses (oder präziser gesagt, die Identifizierung des Agens) vor der Ausführung der ersten Sakkade, oder genauer gesagt, in ihrer *dead-time*-Phase, vorgenommen wird.

Prioritätenliste der mitwirkenden Faktoren: visuelle vs. semantische Salienz

Um eine Prioritätenliste der mitwirkenden Faktoren zu entwickeln, werden zunächst die Einflussstärken der visuellen und der semantischen Salienz auf die visuelle Wahrnehmung miteinander verglichen. Dies gelingt vor allem durch die Beobachtung der Vorhersagbarkeit der beiden Faktoren auf den Platzierungsort der ersten Fixation.

Es konnte ein starker Manipulationseffekt des visuellen Cues auf die Positionierung der ersten Fixation festgestellt werden. Des Weiteren wurde dabei kein Agens-Einfluss gefunden. Da die erste Fixation Indikator für visuelle Wahrnehmung bzw. perzeptuelle Verarbeitung war, weist dieses Ergebnis darauf hin, dass die visuelle Wahrnehmung überwiegend von der visuellen Salienz gesteuert wurde.

Zudem ließ sich eine Bedeutsamkeit der beiden Faktoren für die Auswahl von *starting points* beobachten. Sowohl die Zielposition der zweiten Sakkade als auch die

produzierten Satzstrukturen implizierten eine Agens-Präferenz, was wiederum die Vermutung nahelegt, dass die semantische Salienz in diesem Prozess eine entscheidende Rolle spielt.

Aufgrund dieser Ergebnisse ist daher anzumerken, dass unterschiedliche Verarbeitungsprozesse von unterschiedlichen Faktoren beeinflusst werden. Während die visuelle Verarbeitung überwiegend von visuellen Eigenschaften der Stimuli bestimmt wird, wird die darauffolgende konzeptuelle und linguistische Verarbeitung vor allem von der semantischen Salienz kontrolliert. Insgesamt nimmt die semantische Salienz im Gegensatz zur visuellen Salienz eine vorrangige Position auf der Prioritätenliste ein.

Fazit

Zusammenfassend kann im Rahmen des vorliegenden Experiments festgehalten werden, dass die Satzproduktion im Deutschen in einer pragmatisch neutralen Situation ein strukturell inkrementell verlaufender Prozess ist. Dabei kommt der semantischen Struktur eines Ereignisses, sowohl für die Selektion von *starting points* als auch für die syntaktische Enkodierung, eine zentrale Bedeutung zu. Obwohl die anfängliche visuelle Aufmerksamkeit aufgrund des präsentierten Cues auf die durch diesen Cue markierte Entität gesetzt wurde, dauerte dieser visuelle Effekt nicht bis zur Selektion von *starting points* an. Stattdessen wurde dieser Effekt sofort unterbrochen und von einer sog. Agens-Präferenz abgelöst, sobald die semantisch-konzeptuelle Struktur der Target-Ereignisse aufgebaut wurde. Zudem war es anhand des intersakkadischen Intervalls möglich, eine Aussage über den Zeitverlauf der semantischen Strukturierung zu treffen. Bereits vor der Ausführung der ersten Sakkade wird die semantische Strukturierung vorgenommen. Diese Schlussfolgerung dient weiterhin als Beleg für die strukturell inkrementelle Verarbeitungsweise während der Satzproduktion im Deutschen. Unsere Befunde aus dem vorliegenden Experiment erlauben es auch, eine sog. Prioritätenliste der mitwirkenden Faktoren zu entwickeln. Die semantische Salienz spielt insgesamt eine

ausschlaggebende Rolle bei der Konzeptualisierung und nimmt im Gegensatz zur visuellen Salienz daher eine vorrangige Position in der Prioritätenliste ein.

Allerdings bleibt noch zu erwähnen, dass im vorliegenden Experiment nur auf die Auswirkungen stimulusbasierter Faktoren (sowohl die visuelle als auch die semantische Salienz) eingegangen wurde. Neben solchen stimulusbasierten Faktoren werden Sprecher zumeist spezifischen Kommunikationssituationen ausgesetzt, in denen situative Eigenschaften ebenfalls eine Rolle spielen. Während in Experiment 1 der Prozess der Satzproduktion und seine Ergebnisse im Wesentlichen an sich dargeboten werden, wird in Experiment 2 die Satzproduktion im Kontext betrachtet, um die Effekte eines anderen wichtigen Faktors, nämlich der kontextuellen Salienz, zu untersuchen.

7 Experiment 2

In Experiment 1 wurde festgestellt, dass sich die deutsche Satzproduktion in einer pragmatisch neutralen Situation als strukturell inkrementeller Verarbeitungsprozess beschreiben lässt, bei der die Selektion von *starting points* aufgrund der semantischen Struktur bzw. der semantischen Salienz top-down gesteuert ist. Jedoch ist die semantische Struktur nicht der einzige Faktor, der einen Top-Down-Verarbeitungsprozess verursacht. Ein weiterer Faktor, der ähnlich wie die semantische Struktur ebenfalls der Mikroplanung der Konzeptualisierung angehört und sich daher top-down auswirkt, ist die Informationsstruktur. Diese bezieht sich auf die Zuweisung der Diskursprominenz an Referenten (vgl. Cowles & Ferreira, 2012; Wagner, 2016). Lambrecht (1994) zufolge können syntaktische Selektionen die kontextuelle Salienz der Referenten, die in einem Satz gegeben sind, reflektieren. Des Weiteren haben einige psycholinguistische Studien gezeigt, dass Referenten mit höherer kontextueller Salienz zumeist in prominenteren Positionen einer Äußerung erscheinen (z. B. Bock & Irwin, 1980; Prat-Sala & Branigan, 2000; Cowles & Ferreira, 2012).

Trotz dieser - nach Ansicht des Autors - weitreichenden Erkenntnisfortschritte für das Verständnis des Zusammenhangs zwischen der kontextuellen Information und der Online-Satzproduktion existieren diesbezüglich heute noch immer grundsätzliche Fragen (Arnold et al. (2013): „How is information status represented?“ (S. 409); Wagner (2016): „How it affects word order?“ (S. 521)).

Diese in diesem Forschungsgebiet bestehende Herausforderung lässt erkennen, wie unzureichend das Verhältnis zwischen der kontextuellen Information und der Satzplanung bis heute verstanden ist. Von diesem Forschungsstand ausgehend, zielt das im Folgenden beschriebene Experiment zum einen darauf ab, die Befunde vorangegangener Studien hinsichtlich des Einflusses der kontextuellen Salienz auf die Satzproduktion zu replizieren und zum anderen, die Art der Verarbeitung der kontextuellen Information während des Satzproduktionsprozesses herauszuarbeiten.

7.1 Hypothesen

7.1.1 Hypothesen über die Rolle der kontextuellen Salienz bzw. der Topikalität bei der Satzproduktion

Blickbewegungsdaten

Um feststellen zu können, welche Rolle die kontextuelle Salienz bei der Satzplanung einnimmt, wird besonderes Augenmerk auf das Verhältnis der Topikalität zur Selektion von *starting points* gelegt. Wie bereits im ersten Experiment ausgeführt, kann der Selektionsprozess durch die Zielposition der zweiten Sakkade prognostiziert werden. Wenn die Topikalität einen Einfluss auf die Selektion von *starting points* hat, dann sollte die zweite Sakkade nicht nur von der semantischen Funktion einzelner Referenten, sondern ebenfalls von deren jeweiligem Diskursstatus beeinflusst werden. Demzufolge wird der Befund aus Experiment 1, dass die zweite Sakkade größtenteils auf das Agens gerichtet wird, herausgefordert, insbesondere unter der Bedingung, dass das Agens nicht das Satztopik der zu bildenden Äußerung ist.

Im Vergleich zur oben erwähnten Hypothese hinsichtlich des Verhältnisses zwischen der kontextuellen Salienz und der Selektion von *starting points* ist zudem auch eine stärkere Hypothese über den Einfluss der kontextuellen Salienz auf die Satzplanung möglich: Bereits in der visuellen Wahrnehmungsphase, die in der Regel der visuellen low-level-Verarbeitung unterliegt, kann sich die kontextuelle Salienz anfangen auszuwirken, was wiederum zu einer Korrespondenz zwischen der topikalischen und der zuerst fixierten Figur führen kann. Der Grund für diese Annahme ist der, dass sich die visuelle und kontextuelle Salienz möglicherweise durch einen geteilten Mechanismus, also Aufmerksamkeitslenkung, auswirken (vgl. Gleitman et al., 1996, 2007; Wagner, 2016). In vielen Studien zur visuellen Verarbeitung wurde gezeigt und bestätigt, dass die Aufmerksamkeit die Zuweisung von Figur- und Grund-Rolle beeinflusst. Beispielsweise haben Vecera et al. (2004) herausgefunden, dass die Darstellung eines exogenen Cues die Interpretation eines ambigen Bildes

beeinflussen kann: Die Regionen, an denen der exogene Cue vor der Bild-Darstellung aufgetreten ist, wurden anschließend eher als FIGUR angesehen als andere. Die sog. *figure-ground-perception* findet allerdings nicht ausschließlich im visuellen Kontext Verwendung. Dieser Begriff, der aus der Wahrnehmungspsychologie stammt, beschreibt in der Linguistik unterschiedliche Informationsgewichtungen bzw. –verhältnisse im Prozess der Sprachverarbeitung. Demnach richtet sich die Verteilung zwischen dem Topik und dem Kommentar nach Talmy (1978) nach denselben Mechanismen wie die zwischen der Figur und dem Grund im visuellen Kontext, wobei die Aufmerksamkeitsverteilung eine zentrale Rolle spielt.

Basierend auf dieser Argumentation kann daher vermutet werden, dass die kontextuelle Salienz wahrscheinlich auch einen Einfluss auf die visuelle Wahrnehmung dargebotener Stimuli hat. Da sich die visuelle Wahrnehmung – wie in Experiment 1 dargestellt wurde – vor allem in der Platzierung der ersten Fixation zeigt, wird hier beobachtet, ob die zuerst fixierte Figur mit der kontextuell salientesten Figur, nämlich der Figur mit dem Topikstatus, korrespondiert.

Produzierte Sätze

Die Rolle der kontextuellen Salienz bei der Satzplanung lässt sich ebenfalls anhand der produzierten Sätze überprüfen. Dabei ist die Überprüfung der Korrespondenz zwischen dem zuerst zum Ausdruck gebrachten Referenten und dem Referenten mit dem Topikstatus von großem Interesse. Spielt die kontextuelle Salienz in diesem Prozess eine Rolle, sollten wir eine Übereinstimmung zwischen dem ersten ausgesprochenen Referenten und dem Topik finden. Wenn das Topik allerdings bei der Satzplanung keinen besonderen Stellenwert hat, sollte keine Korrespondenz zwischen den beiden Referenten identifiziert werden können.

7.1.2 Hypothesen über den Abbildungsprozess des Topiks auf die Syntax

Mit den erhobenen linguistischen Daten lässt sich eine weitere Frage bezüglich des

Zeitverlaufs der Verarbeitung kontextueller Informationen beantworten. Im Zentrum dieser Arbeit steht die Frage, inwiefern Topiks, die nach vielen Psycholinguisten (z. B. Levelt, 1989; Klein & v. Stutterheim, 1992) der konzeptuellen Ebene zugeordnet werden, auf die syntaktische Oberflächenstruktur abgebildet werden. Dazu lassen sich vier Hypothesen aufstellen (vgl. Abbildung 7.1): (1) Das Topik wird unmittelbar auf die Linearisierungsebene abgebildet. (2) Das Topik wird zuerst auf die semantische Strukturierungsebene abgebildet, bevor es in die positionelle Enkodierungsebene gelangt. (3) Das Topik wird zuerst auf die syntaktische Funktionszuweisungsebene abgebildet, bevor es in die positionelle Enkodierungsebene gelangt. (4) Die semantische und syntaktische Enkodierung des Topiks gelten beide als Voraussetzung für dessen Anordnung in der Wortabfolge.

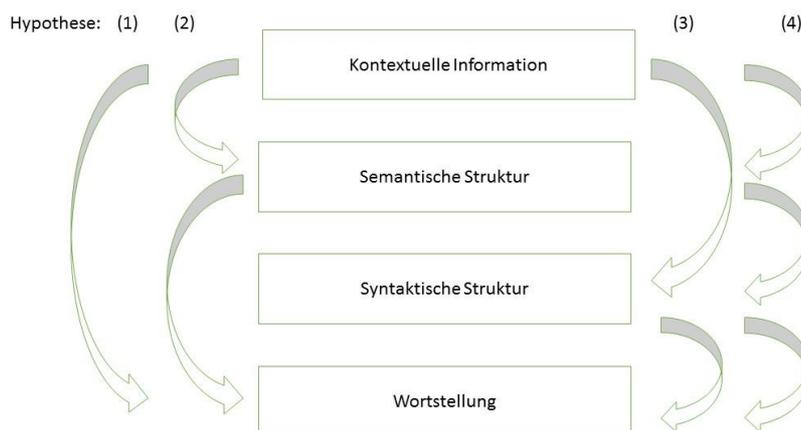


Abbildung 7.1: Vier Hypothesen über den Abbildungsprozess des Topiks auf die Wortstellung

Die erste Hypothese basiert auf der Beobachtung, dass weder zwischen dem Topik und der semantischen Funktion noch zwischen dem Topik und der syntaktischen Funktion zwingend ein Zusammenhang bestehen muss. Stattdessen korreliert das Topik mit seiner satzinitialen Belegungsposition mit der Wortstellung. Aufgrund dieses Verhältnisses ist es berechtigterweise anzunehmen, dass Topiks - sobald sie durch die Target-Frage in der Preview-Phase von Probanden aufgenommen wurden - sofort an die erste Position des zu produzierenden Satzes platziert und ausgesprochen

werden. Eine vorangehende Enkodierung auf der semantischen und der syntaktischen Funktionszuweisungsebene kann nicht nötig sein.

Wenn die simulierte Situation real ist, so ist zu erwarten, dass das Sprechen in diesem Experiment, im Vergleich zum ersten, beschleunigt wird. Denn in dem ersten Experiment waren keine Topiks involviert und das erste lexikalische Konzept konnte erst nach der Präsentation des Target-Ereignisses festgelegt werden.

In Bezug auf die zweite Hypothese wird angenommen, dass, obwohl Topiks in der Regel keine direkte Verbindung zum Gehalt des zu beschreibenden Ereignisses aufweisen, sie trotzdem in der semantischen Vorstellung des Ereignisses lokalisiert werden müssen. Diese Hypothese basiert auf dem Gedanken, dass Figuren mit Topikstatus aufgrund ihrer höheren kontextuellen Salienz als wahrscheinlichste Kandidaten für *starting points* der Satzplanung hervortreten. Aber diese Information ist für eine endgültige Entscheidung zur Selektion von *starting points* noch unzureichend. Sowohl Experiment 1 als auch weitere Studien (z. B. Dobel et al., 2007; Gleitman et al., 2007; van Nice & Dietrich, 2003) haben gezeigt, dass die Selektion von *starting points* in der Tat ein Resultat des Zusammenspiels verschiedener Faktoren ist. Beim Zusammenspiel dieser Faktoren spielen neben der kontextuellen Salienz auch die visuelle, semantische und konzeptuelle Salienz eine Rolle. In diesem Zusammenhang können Sprecher die Entscheidung über die Selektion von *starting points* bzw. den Beginn der linguistischen Verarbeitung verzögern, bis alle relevanten Referenten visuell und semantisch enkodiert sind.

Ausgehend von dieser Überlegung wird daher angenommen, dass eine Einordnung des Topiks in die semantische bzw. konzeptuelle Vorstellung des Target-Ereignisses für die nachfolgende linguistische Enkodierung notwendig ist. Im vorliegenden Experiment kann diese Hypothese jedoch aufgrund des experimentellen Designs nicht getestet werden. Der Grund dafür wird im Diskussionsteil dieses Kapitels näher besprochen.

Hinsichtlich der dritten Hypothese, die den Abbildungsprozess von Informationen

innerhalb des grammatischen Enkodierungssystems betrifft, dominieren derzeit zwei Annahmen: der sog. *grammatical function account* und der *word order account*. Der *grammatical function account* behauptet, dass die syntaktische Enkodierung in zwei streng chronologisch verlaufenden Phasen, also der syntaktischen Funktionszuweisungsphase (Erzeugung syntaktischer Strukturen) und der Konstituenten-Zusammenstellungsphase (Generierung von Wortstellungen), verläuft. Das Erstellen einer Wortstellung setzt die Zuweisung syntaktischer Funktionen voraus und wird daher erst nach der Festlegung syntaktischer Relationen aufgearbeitet. Im Gegensatz dazu geht der *word order account* davon aus, dass konzeptuelle Informationen unmittelbar auf die Wortstellung abgebildet werden, ohne die Rolle der syntaktischen Funktionen berücksichtigen zu müssen. Wie aber entscheidet ein Sprecher über die Abbildung des Topiks innerhalb des grammatischen Enkodierungssystems? Läuft diese streng chronologisch ab oder setzt sie sich über die syntaktische Funktionszuweisungsebene hinweg?

Um diesen Fragen nachzugehen, werfen wir nun zuerst einen Blick auf die Studie von Myachykov und Tomlin (1998), die sich mit demselben Thema in einer Sprache mit einer freien Wortstellung beschäftigten – und zwar dem Russischen. In ihrer Studie kamen die Autoren zu der Schlussfolgerung, dass konzeptuelle Informationen direkt auf die positionelle Enkodierungsebene abgebildet werden, ohne die syntaktische Funktionszuweisungsebene zu durchlaufen. Dieser Befund basiert auf der Beobachtung, dass derjenige Referent, der in einem Ereignis als Patiens auftrat und vorab von einem visuellen Cue angezeigt wurde, aufgrund seiner visuellen Salienz an den Satzanfang platziert wurde; und zwar vor allem durch Scrambling, als durch die Verteilung der Subjekt-Funktion an diesen Referenten.

Im Hinblick auf den Befund von Myachykov und Tomlin (1998) ist im vorliegenden Experiment eine vergleichbare Tendenz bei der Satzproduktion im Deutschen, welches ähnlich wie Russisch ebenfalls eine freie Wortstellung hat, zu entdecken, wenn die Erstellung der Wortfolge nicht unbedingt die syntaktische Funktionszuweisung voraussetzt. Das bedeutet, dass im Falle eines topikalischen

Patiens nicht nur Passivsätze, sondern auch OVS-Aktivsätze, in denen das Patiens als das Objekt erscheint, produziert werden. Wenn aber keine Scrambled-Aktivsätze in der beschriebenen Situation produziert werden, kann dann dieses Ergebnis zu der Schlussfolgerung führen, dass die Verarbeitung des Topiks die beiden Enkodierungsebenen durchlaufen muss? Diese Frage wird später in der Diskussion mithilfe der Befunde von Cowles und Ferreira (2012) geklärt.

Die Entscheidung über die vierte Hypothese ist im Wesentlichen von der Verifizierung der zweiten und dritten Hypothese abhängig.

7.1.3 Hypothesen über die Prioritätenliste der mitwirkenden Faktoren

Blickbewegungsdaten

Auf der Argumentation in dem letzten Abschnitt basierend, wird zunächst eine Korrespondenzprüfung zwischen der zuerst fixierten Figur und der Figur mit dem Topikstatus durchgeführt. Wenn durch diese Prüfung gezeigt werden kann, dass die kontextuelle Salienz ebenfalls eine Rolle bei der Planung der ersten Fixation spielt, werden deren Einflüsse mit denen der visuellen Salienz zum Vergleich gebracht. Es soll beobachtet werden, welcher Faktor stärker auf die Planung der ersten Sakkade und Fixation einwirkt. Der Vergleich wird unter der Bedingung durchgeführt, dass der visuelle Cue und der Topikstatus nicht auf dieselbe Figur fallen dürfen. Wirkt sich die kontextuelle Salienz stärker aus, so liegt demgemäß eine stärkere Korrespondenz zwischen der zuerst fixierten und der topikalischen Figur vor. Wenn aber die visuelle Salienz der dominante Faktor ist, sollte diese eine bessere Vorhersagbarkeit auf die erste Fixationsposition liefern können.

In Bezug auf die Selektion von *starting points* werden hierbei nur die Wirkungsstärken der beiden Top-Down-Faktoren, nämlich der semantischen und der kontextuellen Salienz, miteinander verglichen. Denn die visuelle Salienz - wie bereits in Experiment 1 gezeigt wurde - spielt in diesem Prozess keine Rolle. Als visueller Indikator dafür wird die Zielposition der zweiten Sakkade verwendet. Hierfür wird beobachtet, welcher Faktor eine bessere Vorhersagbarkeit auf die zweite

sakkadische Zielposition anbieten kann, wenn Sprecher einer Situation ausgesetzt sind, in der das Patiens das Topik der zu bildenden Sätze war. Es wird diesbezüglich angenommen, dass, wenn die semantische Salienz bei der zweiten sakkadischen Planung von größerer Bedeutung wäre, die zweite Sakkade unabhängig von der Zuweisung des Topikstatus durchgehend auf das Agens gerichtet sein sollte. Im Gegensatz dazu sollte das Patiens als angemessenes Ziel ausgewählt werden, wenn die kontextuelle Salienz dabei eine dominante Rolle spielt.

Produzierte Sätze

Der zweiten Sakkade entsprechend werden die produzierten Sätze auch verwendet, um die Wirkungsstärke der semantischen und der kontextuellen Salienz auf die Selektion von *starting points* zu überprüfen. Dazu wird die Übereinstimmung sowohl des Agens als auch des Topiks mit dem satzinitialen Wort miteinander verglichen. Derjenige Faktor, der zu einer höheren Vorhersagbarkeit führt, welcher der beiden Referenten zuerst erwähnt wird, soll eine vorrangige Position auf der Prioritätenliste einnehmen. Derjenige Faktor, der eine schwächere Auswirkung aufzeigt, wird dann an die hintere Position gestellt.

7.2 Methodisches Vorgehen

7.2.1 Versuchspersonen

Alle 32 beteiligten Versuchsteilnehmer (davon 27 Frauen, Median 25,2 Jahre, Bereich 19-32 Jahre) waren Studierende an der Universität Heidelberg und wuchsen mit Deutsch als Muttersprache auf. Ihre Sehstärke war normal oder mithilfe von Kontaktlinsen oder Brillen auf das Normalniveau korrigiert. Die Teilnahme erfolgte freiwillig und alle Probanden erhielten vom Experimentleiter eine Aufwandsentschädigung in Höhe von 7 EUR.

7.2.2 Technische Apparatur

Die technische Apparatur war dieselbe wie in Experiment 1.

7.2.3 Material

Im Prinzip waren alle Stimuli, die in Experiment 2 verwendet wurden, identisch mit denen in Experiment 1. Der einzige Unterschied betrifft das linguistische Material, also die in der Preview-Phase gestellte Target-Frage. Im Gegensatz zu Experiment 1 (*Sag mir bitte, was auf dem zweiten Bild passiert?*) wurden die Target-Fragen in dem vorliegenden Experiment nach dem Muster *Sag mir bitte, was mit [EINER FIGUR] passiert* gestaltet, durch den eine spezifische Perspektive ermittelt wurde, aus welcher die Antwort organisiert werden sollte.

7.2.4 Durchführung

Die Durchführung und der Ablauf eines Durchgangs waren identisch mit denen in Experiment 1. Zunächst wurden 3 Trials zur Übung durchgeführt (2 × kritische Bedingung, 1 × Filler-Bedingung). Danach wurden die experimentellen Trials abgespielt, wobei die Darstellungsreihenfolge pseudo-randomisiert wurde. Jeder Durchgang begann mit der Präsentation eines Einführungsbildes mit zwei nebeneinanderstehenden belebten Referenten, die dann als Ereignisteilnehmer im anschließenden Target-Bild nochmals präsentiert wurden. Während der Präsentation des Einführungsbildes bekamen die Versuchspersonen auditiven Input, zum einen in Form einer einführenden Aussage, durch die die Bezeichnungen der beiden Referenten eingeführt wurden, und zum anderen einer sog. Target-Frage, die sich auf das kommende Target-Bild bezog. In der kritischen Bedingung lautete so eine Frage bspw. *Auf dem nächsten Bild siehst du die Maus wieder; sag mir bitte, was mit der Schildkröte passiert?* und in der Filler-Bedingung zum Beispiel *Haben sich ihre Positionen auf dem zweiten Bild geändert?*. In einer Hälfte der kritischen Trials bezog sich die *mit*-Phrase auf das Agens, während in der anderen Hälfte das Patiens von der Frage näher spezifiziert wurde. Die Probanden wurden vor Experimentbeginn darüber informiert, dass der Interviewer ihre Antworten anschließend abhört und beurteilt, ob die gestellten Fragen damit auch beantwortet wurden. Aus diesem Grund sollten Probanden es vermeiden, die gestellten Fragen zu ignorieren oder die vorhandene Aufgabe einfach als eine simple

Beschreibungsaufgabe anzusehen. Stattdessen sollte sich die Planung der Antwort seitens der Probanden immer nach der konkreten Fragestellung richten. Nach den Fragen erschien ein Fixationskreuz in der Mitte des Bildschirms. Die Versuchspersonen wurden aufgefordert, das Fixationskreuz zu fixieren. Nach 500 ms wurde ein leerer Bildschirm gezeigt und nach weiteren 200 ms blitzte der exogene Cue auf. *Das Einzige, das sich vom ersten Experiment unterscheidet, ist das Auftreten des Cues nur in der Region des aufkommenden Patiens.* Nach dem Erscheinen des Cues erschien anschließend das Target-Bild. Die Versuchspersonen wurden dann aufgefordert, anhand dieses Bildes die zuvor gestellte Frage mit einem kompletten Satz so schnell wie möglich zu beantworten (vgl. Abbildung 7.2).

7.2.5 Datenkodierung

Die produzierten Sätze und Augenbewegungsmessungen wurden genau so aufbereitet wie in Experiment 1.

7.2.6 Datenbearbeitung und -analyse

In Bezug auf den Einfluss der kontextuellen Salienz auf die Satzplanung sollte zuerst die Frage geklärt werden, ob Topiks, die durch die Target-Fragen vermittelt wurden, die anfängliche visuelle Aufmerksamkeitsverteilung beeinflussen (oder nicht). Ein möglicher Effekt diesbezüglich sollte sich in dem ersten Fixationsort widerspiegeln bzw. zeigen. Dabei wird beobachtet, ob der Effekt des visuellen Cues, der die Positionierung der ersten Fixation reizbedingt manipulieren kann, vom kontextuellen Status der durch den Cue spezifizierten Referenten modifiziert werden kann. Eine logistische Regressionsanalyse mit gemischten Effekten wird durchgeführt, in der *der visuelle Cueingeffekt* die abhängige Variable ist und *der kontextuelle Status des durch den Cue spezifizierten Patiens (Topik vs. Nicht-Topik)* die feste unabhängige Variable ist. Dazu werden *Stimuli* und *Probanden* als zufällige unabhängige Variablen ausgewertet.

Neben der Relation zwischen der Topikalität und der visuellen Wahrnehmung steht die Rolle des Topiks bei der Konzeptualisierung nach der Wahrnehmungsphase auch

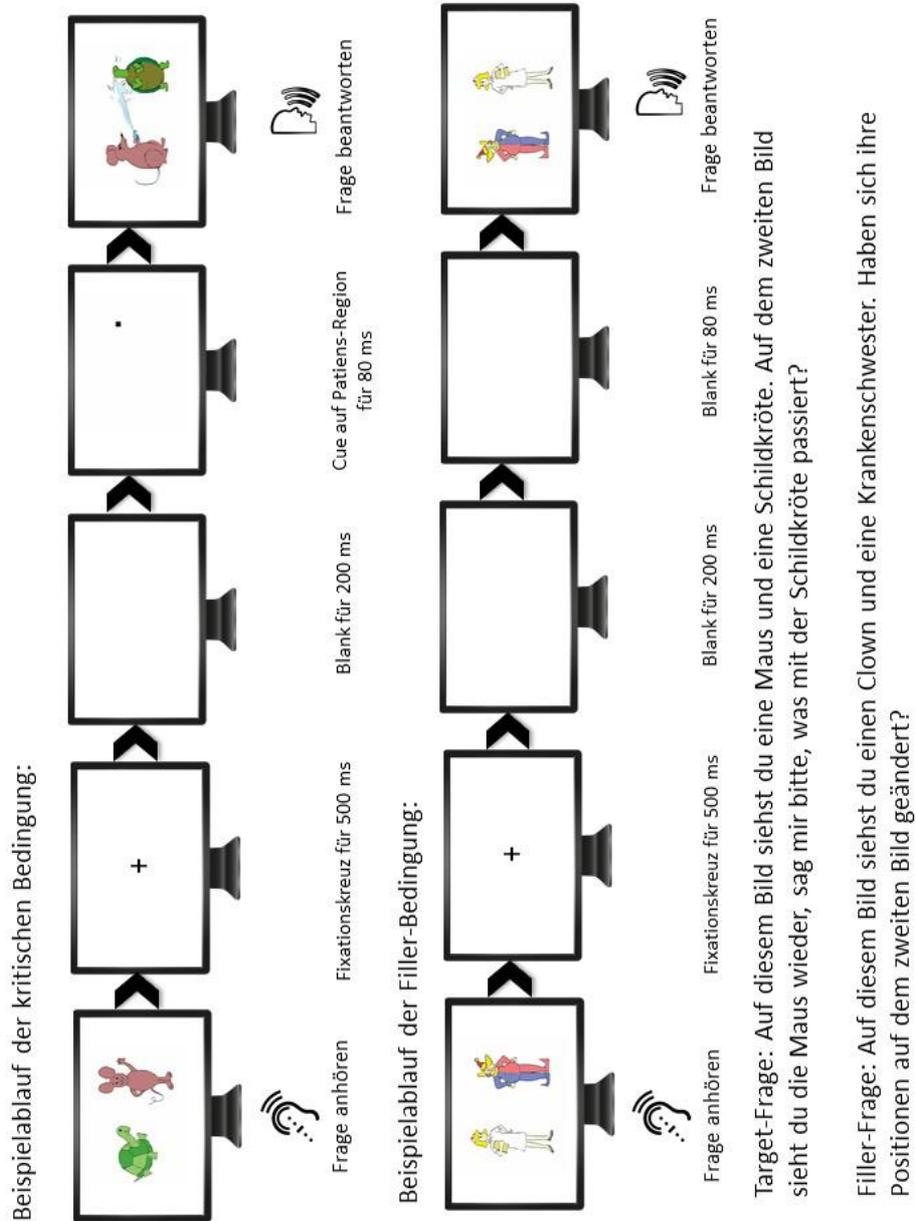


Abbildung 7.2: Beispielablauf der kritischen Bedingung und der Filler-Bedingung (Experiment 2)

im Mittelpunkt dieses Experimentes. Hierfür wird *der Effekt des Agens*, der sich in Experiment 1 als ausschlaggebend für die Selektion von *starting points* bzw. die syntaktische Strukturierung erwies, als abhängige Variable betrachtet. Es gilt, herauszufinden, ob der Effekt des Agens, der sich sowohl in der Selektion von syntaktischen Strukturen als auch in der Zielposition der zweiten Sakkade widerspiegelt, von seinem kontextuellen Status beeinflusst wird. Dafür wird noch einmal das logistische Regressionsmodell mit gemischten Effekten verwendet, bei dem *der kontextuelle Agens-Status (Topik vs. Nicht-Topik)* als feste und *Stimuli* und *Probanden* als zufällige unabhängige Variablen definiert werden.

Durch die Auswertung der Speech-Onset-Latenzen der produzierten Sätze ohne (Experiment 1) und mit der Beteiligung des Topiks (Experiment 2) ist es möglich zu testen, ob das Topik sofort in die lineare Wortabfolge integriert wird, sobald es durch die Target-Frage bekannt ist. Oder anders formuliert: Ist es möglich, dass die lineare Anordnung des Topiks ohne Aufbau der Ereignis- und der syntaktischen Struktur durchgeführt werden kann?

Da die Satzproduktionssituationen in Experiment 1 und 2 unterschiedlich waren, sollten nur die Sätze ausgewählt werden, die unter vergleichbaren Bedingungen gebildet wurden. In Experiment 1 wurden lediglich Aktivsätze produziert. Damit die Differenz zwischen der Speech-Onset-Latenz von Aktiv- und Passivsätzen nicht als störender Faktor für die Analyse auftritt, werden dementsprechend aus Experiment 2 nur Aktivsätze selektiert, in denen das Agens als Satztopik hervortritt. In Experiment 2 erschien der visuelle Cue nur auf dem Sichtfeld des Patiens. Demgemäß wurden aus Experiment 1 nur diejenigen Sätze ausgewählt, bei deren Produzieren das Patiens durch den exogenen visuellen Cue angezeigt wurde. Auf diese Weise waren die miteinander verglichenen selektierten Sätze unter einer gleichen Bedingung produziert, mit dem einzigen Unterschied, dass das Agens in Experiment 1 kontextuell neutral und in Experiment 2 das Topik ist. Zur statistischen Auswertung wird ein zweistichproben-*t*-Test aus der Sicht der *Stimuli* durchgeführt, um die Speech-Onset-Latenzen der beiden Satzgruppen miteinander zu vergleichen. Eine

t-Test-Analyse aus der Perspektive der Probanden ist hier nicht möglich, da verschiedene Probanden an Experiment 1 und 2 teilnahmen.

Ein weiteres Ziel dieses Experiments, neben der Messung des Einflusses der kontextuellen Salienz auf die Satzplanung, ist es, eine Prioritätenliste für die mitwirkenden bzw. involvierenden Faktoren zu bilden. Wie bereits oben erwähnt sind diese Faktoren die visuelle, semantische als auch die kontextuelle Salienz.

Die Anordnung dieser Faktoren in einer Liste ist von deren Einflussstärken abhängig. Dafür wurden dieselben Kriterien wie im ersten Experiment gewählt: die Vorhersagbarkeiten der Faktoren zum einen für den ersten Fixationsort, welcher als Indikator für die visuelle Wahrnehmung herangezogen wird, und zum anderen für den zweiten Fixationsort sowie den ersten artikulierten Referenten, welche die Selektion von *starting points* widerspiegeln können. Für diese Analyse sind insbesondere Ergebnisse aus der Situation relevant, wo die zum Vergleich gebrachten Faktoren im Widerspruch zueinanderstehen (z. B. beziehen sich der visuelle Cue und die *mit*-Phrase in der Target-Frage auf verschiedene Referenten).

Zur Identifizierung der Verarbeitungspriorität der Faktoren für die visuelle Wahrnehmung werden die Effekte der visuellen und kontextuellen Salienz auf die Positionierung der ersten Fixation miteinander verglichen. Da es im ersten Experiment keinen Effekt der semantischen Salienz auf die visuelle Wahrnehmung gab, wird dieser deswegen nicht in die Berechnung mit einbezogen. Ein zweistichproben-*t*-Test, sowohl aus der Sicht der *Probanden* als auch der *Stimuli*, wurde durchgeführt, um die Vorhersagbarkeit des exogenen Cues und der Topikalität auf die erste Fixation zu vergleichen. Dabei ist die topikalische Figur das Agens und stimmt daher mit dem immer durch den Cue angezeigten Patiens nicht überein.

Hinsichtlich der zweiten Fixation werden die Vorhersagbarkeit des Topiks und die des Agens nur unter der Bedingung miteinander verglichen, dass das Patiens topikalisiert wird, nachdem die erste Sakkade visuell bedingt auf das Patiens gerichtet wurde (Topik/Patiens vs. Agens). In Ähnlichkeit dazu wird die Vorhersagbarkeit der beiden

Top-Down-Faktoren auf das Anfangswort der gebildeten Sätze unter der Bedingung, dass das Patiens Topik und das Agens nicht Topik ist, mit einem *t*-Test überprüft.

7.3 Ergebnisse

Fehler

Insgesamt wurden 896 Sätze produziert. Davon werden ca. 4.6 % der erhobenen Sätze aufgrund von Fehlern von der statistischen Analyse ausgeschlossen. Es wurden dieselben Fehlerkategorien angelegt wie in Experiment 1. Tabelle 7.1 zeigt genau die Entstehung der Fehler.

Fehlerkategorie	Anzahl
Technische Messfehler	4 (0.4%)
Fehler Performanz der Sätze insgesamt	38 (4.2%)
Fehler Performanz der Sätze detailliert	38 (100%)
Sätze mit Vertauschung	0 (0%)
Sätze mit intransitiven Verben	31 (82%)
Sätze mit nur einem Ereignisteilnehmer	3 (8%)
Sätze mit koordinierten NPs	1 (2%)
Antworten aus mehreren Sätzen	3 (8%)

Tabelle 7.1: Fehler (Experiment 2)

Kontextuelle Salienz bzw. Topikalität und erste Fixation

Zuerst wurden die ersten Fixationspositionen analysiert. In allen Durchgängen im ersten Experiment wurden 69% der ersten Fixationen auf den durch den Cue markierten Charakter platziert. In diesem Experiment wurden die Effekte des visuellen Cues allerdings beträchtlich von dem Diskursstatus der markierten Figur, also des Patiens beeinflusst: Während 79.3% der ersten Fixationen auf das Patiens gerichtet wurden, wenn dem Patiens der Topikstatus zugewiesen wurde, sinkt dieses Verhältnis auf 37.1%, wenn hingegen das visuell nicht markierte Agens das Topik war

(vgl. Abbildung 7.3). Ein signifikanter Effekt der kontextuellen Salienz konnte durch eine logistische Regressionsanalyse entdeckt werden, was wiederum darauf hinweist, dass die kontextuelle Information bereits in der visuellen Wahrnehmungsphase eine Rolle gespielt hat (vgl. Tabelle 7.2).

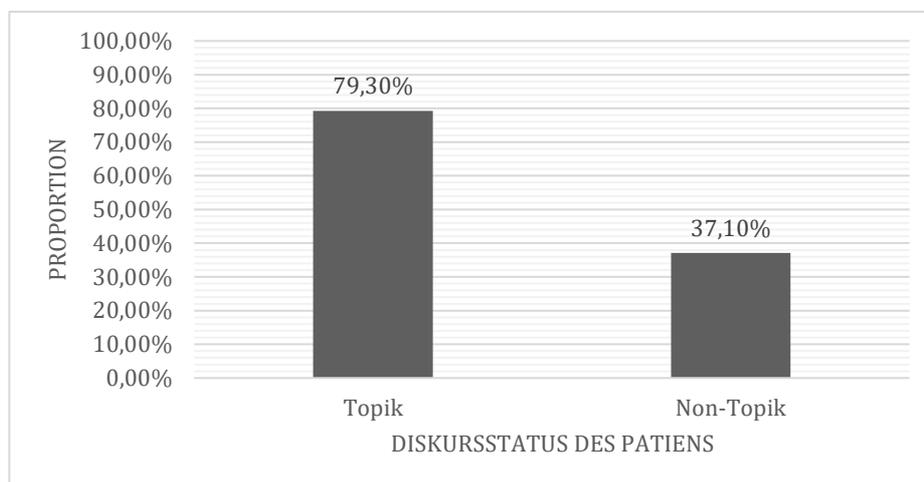


Abbildung 7.3: Proportion der durch den visuellen Cue getriebenen ersten Fixationen im Verhältnis zum Diskursstatus des Patiens (Das Patiens wurde immer durch den Cue markiert)

Effekt	Est.	Std. Error	z-value
Intercept	0.5232	0.1295	4.039
Fester Effekt	0.9659	0.1287	7.504***

Tabelle 7.2: Logistische Regressionsanalyse auf die Abhängigkeit des visuellen Cueingeffekts vom Diskursstatus des Patiens

Doch bedeutet das, dass sich der visuelle Cue, welcher einen Bottom-Up-Einfluss ausübt, nach der Einführung der kontextuellen Salienz als Einflussfaktor nicht mehr auf die Planung der ersten Sakkade auswirkt? Durch eine logistische Regressionsanalyse konnte ein signifikanter visueller Cueingeffekt auf den topikalischen Effekt festgestellt werden (vgl. Tabelle 7.3). Wenn das Topik das Agens war und daher visuell vorab nicht markiert wurde, lag die Vorhersagbarkeit des ersten Fixationsorts durch das Topik bei 54.7%. Dieses Verhältnis erhöht sich auf 79.3%, wenn das Topik sich auf das visuell markierte Patiens bezog (vgl. Abbildung 7.4). Aufgrund dieses Ergebnisses ergibt sich, dass die Stärke des Topik-Effekts davon abhängig ist, ob der topikalische Referent vorab durch einen Cue angezeigt wurde.

Trotz des starken Manipulationseffekts der Topikalität auf die erste Fixation wurde der Effekt aufgrund der visuellen Salienz weiterhin bewahrt.

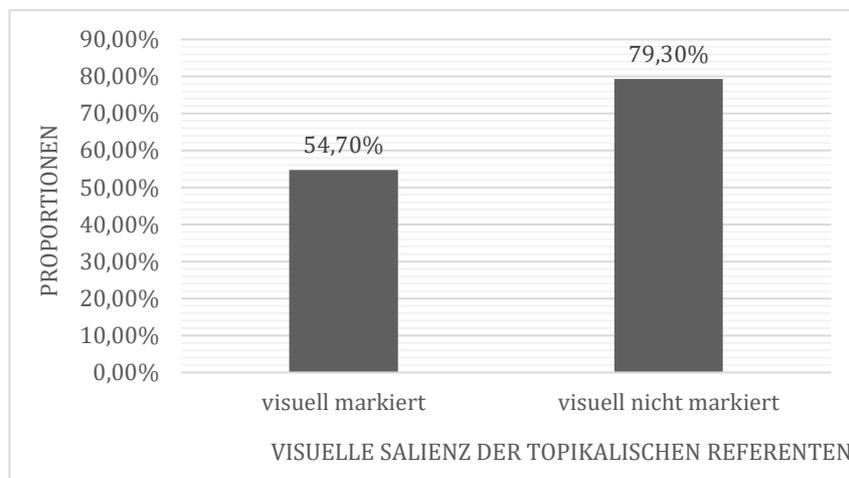


Abbildung 7.4 Proportion der durch das Topik getriebenen ersten Fixationen im Verhältnis zur visuellen Salienz der topikalischen Referenten

Effekt	Est.	Std. Error	z-value
Intercept	0.8607	0.1150	7.487
Fester Effekt	0.6292	0.1270	4.952***

Tabelle 7.3: Logistische Regressionsanalyse auf die Abhängigkeit des Topikalitäts-Effekts von der visuellen Salienz der topikalischen Referenten

Kontextuelle Salienz bzw. Topikalität und Zielposition der zweiten Sakkade

Die kontextuelle Salienz bzw. die Topikalität hat nicht nur einen Effekt auf die visuelle Wahrnehmung, sondern auch auf die Selektion von *starting points*. Dies spiegelt sich in erster Linie in der Zielselektion der zweiten Sakkade wider. In Experiment 1 wurde festgestellt, dass, wenn die erste Fixation aufgrund des visuellen Cues auf dem Patiens landete, die zweite Sakkade größtenteils auf das Agens gesteuert wurde, was eine sog. *Agens-Präferenz* andeutet. So eine Präferenz wurde allerdings in diesem Experiment nur beobachtet, wenn das Agens das Topik der nachfolgenden Äußerung ist. In diesem Fall richteten die Versuchspersonen 73.9% ihrer zweiten Sakkaden auf das Agens. Wurde hingegen dem Patiens im Gegensatz zum Agens der Topikstatus zugewiesen, verringerte sich das Verhältnis auf 22% (vgl.

Abbildung 7.5). Die statistische Analyse hat auch bestätigt, dass ein signifikanter Effekt der Topikalität auf die Planung der zweiten Sakkade gegeben ist (vgl. Tabelle 7.4).

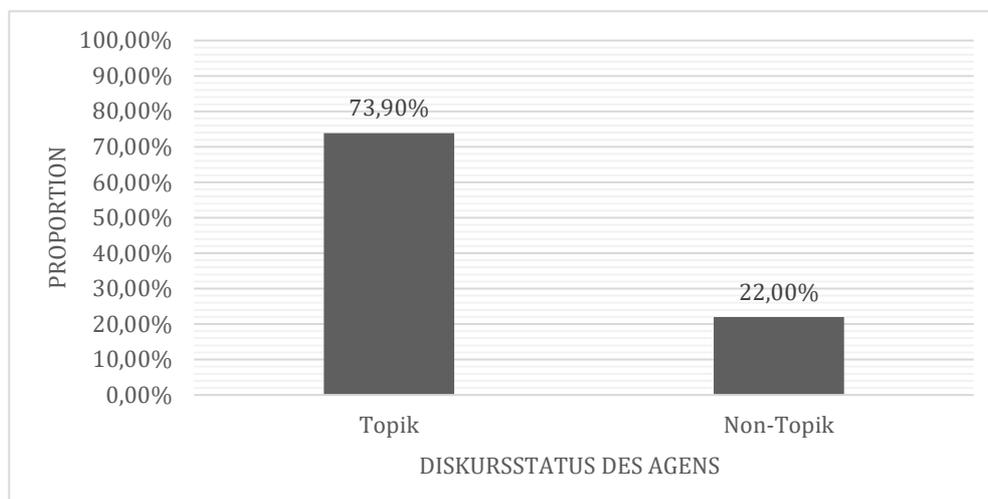


Abbildung 7.5 Proportion der zweiten Sakkaden, die auf das Agens gesteuert wurden, im Verhältnis zum Diskursstatus des Agens

	Est.	Std. Error	z-value
Intercept	0.08333	0.16325	0.51
Fester Effekt	1.32576	0.11770	11.26***

Tabelle 7.4: Logistische Regressionsanalyse auf die Abhängigkeit der Selektion der zweiten sakkadischen Zielposition vom Diskursstatus des Agens

kontextuelle Salienz bzw. Topikalität und syntaktische Struktur

In Bezug auf die syntaktische Struktur der produzierten Sätze wurde in Experiment 1 festgestellt, dass deutsche Testpersonen in pragmatisch neutralen Satzproduktionssituationen das Agens immer zuerst erwähnt haben. Dieses Ergebnis deutet daher auf eine ausschlaggebende Rolle der semantischen Salienz bei der Satzplanung im Deutschen hin. In diesem Experiment jedoch wurde außerdem noch ein Manipulationseffekt der kontextuellen Salienz auf die syntaktische Strukturierung festgestellt. Wenn das Agens das Topik war, waren 99.6% der produzierten Sätze Aktivsätze, wohingegen sich die Proportion der gebildeten Aktivsätze bei einer Nicht-Übereinstimmung zwischen dem Agens und dem Topik auf 3.2% verminderte

(vgl. Abbildung 7.6). Es wurden keine Scrambled-Aktivsätze produziert, in denen das Patiens als das Objekt in der satzinitialen Position aufgetreten ist. Durch eine logistische Regressionsanalyse lässt sich eine Abhängigkeit der Effekte des Agens von dessen kontextuellen Status feststellen (vgl. Tabelle 7.5). Das bedeutet, ob das Agens als der erste Referent versprachlicht wird oder nicht, ist folglich im großen Maß davon abhängig, ob es das Topik des zu bildenden Satzes ist.

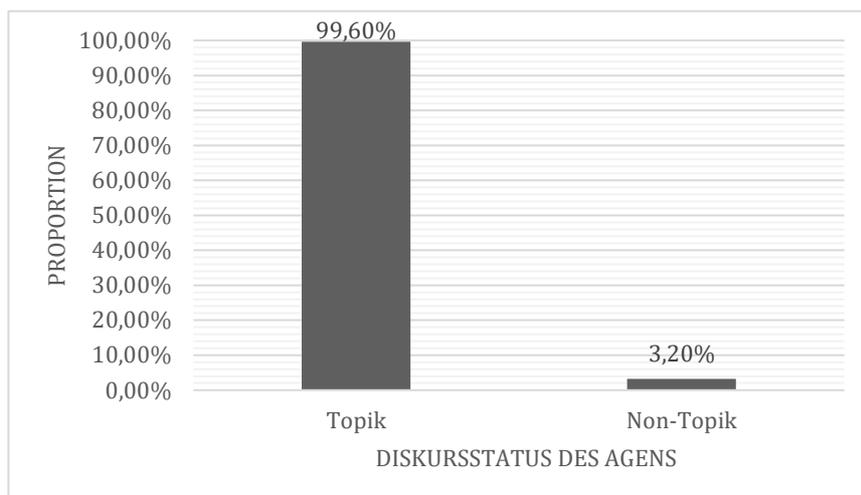


Abbildung 7.6: Proportion der gebildeten Aktivsätze im Verhältnis zum Diskursstatus des Agens

	Est.	Std. Error	z-value
Intercept	0.8215	1.1897	0.691
Fester Effekt	6.4122	1.2557	5.107***

Tabelle 7.5: Logistische Regressionsanalyse auf die Abhängigkeit der Selektion syntaktischer Strukturen vom Diskursstatus des Agens

Speech-Onset-Latenz der produzierten Sätze mit und ohne Teilnahme des Topiks

Hinsichtlich des Zeitverlaufs der Satzproduktion wurde ausgewertet, ob die Präsenz des Topiks die Speech-Onset-Latenz beeinflusst hat. Die erhobenen Sätze, die nach den in Abschnitt 7.2.6 erwähnten Kriterien jeweils aus Experiment 1 und 2 selektiert wurden, wurden miteinander verglichen. Die gemittelte Speech-Onset-Latenz der selektierten Sätze aus Experiment 1 (ohne Teilnahme des Topiks) betrug 1639 ms, die aus Experiment 2 (mit Teilnahme des Topiks) 1642 ms (vgl. Abbildung 7.7). Ein

t-Test ergab keinen signifikanten Unterschied, $t(27) = -0.10765$, $p > .05$.

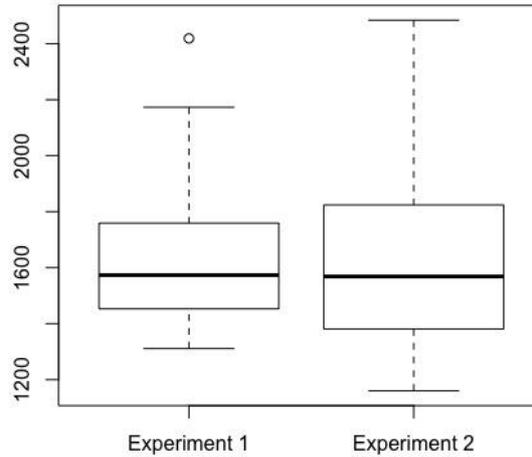


Abbildung 7.7 Speech-Onset-Latenz der selektierten Sätze aus Experiment 1 und 2

Einflussstärken der mitwirkenden Faktoren bezüglich der visuellen Wahrnehmung: kontextuelle vs. visuelle Salienz

In Experiment 1 wurde herausgefunden, dass die visuelle Salienz eine zentrale Rolle bei der visuellen Wahrnehmung spielt. Da in diesem Experiment ein topikalischer Einfluss auch auf die Planung der ersten Fixation beobachtet wurde, sollte in erster Linie festgestellt werden, welcher Faktor eine stärkere Rolle in diesem Prozess spielt. Hierzu wurde die Vorhersagbarkeit des ersten Fixationsorts durch den visuellen Cue und die durch das Topik verglichen, unter der Bedingung, dass der Cue und der Topikstatus nicht auf dieselbe Figur gefallen waren. Während die Vorhersagbarkeit durch den Cue 37.1% betrug, lag sie durch das Topik bei 54.7%. Es konnte ein signifikanter Vorhersagbarkeitsunterschied errechnet werden, wobei die kontextuelle Salienz bessere Ergebnisse erzielte, $t_1(27) = -3.6352$, $p_1 < 0.01$, $t_2(31) = -3.0142$, $p_2 < 0.01$.

Einflussstärken der mitwirkenden Faktoren bezüglich der Selektion von starting points: kontextuelle vs. semantische Salienz

In Bezug auf die Selektion von *starting points* wurde die Stärke des kontextuellen und des semantischen Salienzeinflusses miteinander verglichen. Hierfür wurden zwei Indikatoren – der zweite Fixationsort und der erst artikulierte Referent – verwendet.

Bei der Planung der zweiten Sakkade rief die kontextuelle Salienz einen stärkeren Effekt als die semantische Salienz hervor. In allen Durchgängen, bei denen die erste Sakkade aufgrund des visuellen Cues auf das Patiens gerichtet wurde, betrug die Wahrscheinlichkeit, dass der Blick der Sprecher in der Patiensregion blieb, ganze 77%, wenn das Patiens zeitgleich das Topik war. Unter der gleichen Bedingung wurden hingegen nur 22% der zweiten Sakkade auf das Agens gerichtet. Außerdem wurde ein signifikanter Unterschied in der Vorhersagbarkeit des Agens und des Topiks auf die zweite Fixation errechnet, $t_1(27) = 13.127$, $p_1 < 0.01$, $t_2(31) = 10.584$, $p_2 < 0.01$.

Eine analoge Tendenz zeigt sich ebenfalls in der Selektion des ersten Referenten der gebildeten Sätze. Wenn das Patiens das Topik war, lag die Proportion der gebildeten Passivsätze bei 96.7%, wohingegen im Gegensatz dazu der Wert der kanonischen Aktivsätze lediglich 3.2% betrug. Dieser Unterschied kann als signifikantes Ergebnis gewertet werden. Diese Ergebnisse verdeutlichen die Tatsache, dass die kontextuelle Salienz in Bezug auf die Selektion von *starting points* eine dominante Rolle spielt.

7.4 Diskussion

Eine Zielsetzung dieses Experiments war zum einen die Untersuchung der Rolle des Kontextes, welche in vorangehenden Studien zur Satzproduktion wenig behandelt wurde. Die zentrale Frage hierbei war, ob die kontextuelle Salienz bzw. die Topikalität die Konzeptualisierung und die grammatische Enkodierung der Satzplanung beeinflusst. Während des Experiments wurde die visuelle Präsentation gleich gehalten. Durch eine Preview-Phase vor der Präsentation des Target-Ereignisses unterschieden sich die Diskursprominenzen der Ereignisteilnehmer. Denn dadurch erhielt entweder das Agens oder das Patiens durch eine spezifizierte

Frage den Topik-Status. Wenn die kontextuelle Salienz bzw. der Topikstatus tatsächlich einen Einfluss auf die Satzplanung im Deutschen hat, sollte sich dieser in den erhobenen visuellen und linguistischen Daten finden lassen.

Neben der Relation zwischen der kontextuellen Salienz und der Satzproduktion wird auch besonderes Gewicht auf den Zeitverlauf der Topik-Verarbeitung während der Satzplanung gelegt. Der zentrale Punkt hierbei ist die Frage, in welcher Art und Weise Topiks von der konzeptuellen Repräsentation auf die Syntax abgebildet werden. Die diesbezügliche Diskussion gliedert sich in zwei Teile. Im ersten Teil sollte der Mapping-Prozess des Topiks innerhalb der grammatischen Enkodierungsphase identifiziert werden. Die zentrale Frage hierbei ist, in welcher Art und Weise die topikalische Information auf die lineare Position abgebildet wird (*Abbildungsprozess innerhalb der grammatischen Enkodierung*). Im zweiten Teil muss weiterhin diskutiert werden, ob die topikalische Information zuerst auf die semantische Repräsentation abgebildet werden muss, bevor sie an die grammatische Enkodierungsebene weitergereicht wird (*Abbildungsprozess innerhalb der Konzeptualisierung bzw. der Mikroplanung*).

Des Weiteren lag in unserem Interesse herauszufinden, inwiefern kontextuelle Informationen, hier vor allem das Topik, während des Satzproduktionsprozesses repräsentiert werden. Dafür gibt es zwei Möglichkeiten (vgl. Arnold et al., 2013): Die erste Möglichkeit ist, dass das Topik lediglich ein linguistisches Element ist und daher nur sprachlich repräsentiert wird. Sollte diese Hypothese stimmen, dann sollte ein Einfluss der Topikalität nur innerhalb des linguistischen Verarbeitungssystems erkennbar sein. Die zweite Möglichkeit ist, dass das Topik neben seiner sprachlichen auch eine konzeptuelle Repräsentation impliziert, die sich außerhalb des linguistischen Systems befindet. Dann sollte in diesem Fall ein Einfluss des Topiks auch auf die prälinguistische Enkodierung, also die Konzeptualisierung und sogar auf die visuelle Wahrnehmung, zu sehen sein.

Ähnlich wie beim ersten Experiment war das Ziel des vorliegenden Experiments

ebenfalls die Erstellung einer Prioritätenliste, auf der die bei der Satzproduktion mitwirkenden Faktoren nach deren Wirksamkeit auf die Satzplanung gereiht sind. Im ersten Experiment konnte bereits ein Teil solch einer Liste, die die relative Einflussstärke der visuellen und semantischen Salienz auf die Satzplanung beschreibt, festgestellt werden. Durch die anschließende Diskussion wird die kontextuelle Salienz in die Prioritätenliste eingeordnet. Von diesen Fragen ausgehend werden nun im Folgenden die gewonnenen Ergebnisse ausgewertet.

Kontextuelle Salienz und Satzplanung

Es wurde zuerst die Hypothese aufgestellt, dass sogar bei der visuellen Wahrnehmung die kontextuelle Salienz eine wichtige Rolle spielt. Zur Überprüfung dieser Hypothese wurde das Verhältnis zwischen der Topikalität und dem visuellen Cueingeffekt, der bereits in Experiment 1 identifiziert wurde, betrachtet. Dabei galt es herauszufinden, ob eine Korrespondenz zwischen dem reizbedingt als erstes fixierten Referenten und dem mit dem Topikstatus besteht. Das Vorhandensein einer derartigen Korrespondenz soll als Beleg für den Effekt der Topikalität auf die perzeptuelle Verarbeitung dienen.

Die durchgeführte statistische Analyse hat einen deutlichen Effekt der kontextuellen Salienz auf die Positionierung der ersten Fixation gezeigt: Ob die Augenbewegungen gleich nach Stimulus-Onset auf die durch den Cue angezeigte Figur gelenkt wurden oder nicht, war überwiegend davon abhängig, ob diese Figur zeitgleich das Topik des darauffolgenden Diskurses war. Wenn eine Figur gleichzeitig visuell markiert und das Topik war - im Vergleich zum ersten Experiment, indem die visuell markierte Figur pragmatisch neutral war - war eine leichte Steigerung in der Vorhersagbarkeit der ersten Fixation durch den exogenen Cue (von 68.9% auf 79.3%) zu erkennen. Im Gegensatz dazu wurde der Effekt der visuellen Salienz erkennbar beeinträchtigt, wenn eine andere Figur als die visuell markierte Entität das Topik der zu bildenden Äußerung war (von 68.9% auf 37.1%). Diese Ergebnisse bestätigen daher die Hypothese, dass die perzeptuelle Verarbeitung eines Ereignisses von der kontextuellen Salienz jeweiliger Ereignisteilnehmer beeinflusst wird.

Die Frage ist jedoch, wie lange der beobachtete topikalische Effekt anhält. Ist dieser Effekt, ähnlich wie der visuelle Cueingeffekt, nur vorübergehend und wird nach dem Aufbau der semantischen Struktur umgehend abgelöst? Oder kann er den Effekt der semantischen Salienz überwinden?

In diesem Experiment haben wir deutliche Hinweise darauf gefunden, dass die kontextuelle Salienz bzw. die Topikalität nicht nur bei der perzeptuellen Verarbeitung, sondern auch bei der Auswahl von *starting points* eine entscheidende Rolle spielt. Folgende Forschungsergebnisse hinsichtlich der produzierten Sätze und der ausgerichteten zweiten Sakkade sind Grundlage für diese Schlussfolgerung:

In Bezug auf die produzierten Sätze war anstelle der Agens-Präferenz eine starke Topik-Präferenz zu beobachten, bei der Figuren mit Topikstatus unabhängig von deren semantischen Funktionen sehr häufig der Subjekt-Funktion zugeordnet bzw. satzinitial angeordnet wurden. Das Agens, das in Experiment 1 durchgehend satzinitial aufgetreten ist, wurde meistens nur dann in das Vorfeld eines Satzes platziert, wenn es gleichzeitig das Topik war. Wenn dem Patiens allerdings der Topik-Status zugeschrieben wurde, dann wurden die produzierten Sätze in den meisten Fällen im Passiv formuliert. Die starke Korrespondenz zwischen dem Topik und dem ersten Referenten bzw. dem Satz-Subjekt legt nahe, dass der Effekt der kontextuellen Salienz, im Gegensatz zu dem Effekt der visuellen Salienz, welcher nach dem Aufbau der semantischen Struktur umgehend abgelöst wird, bis zur grammatischen Verarbeitungsstufe anhält.

In den Daten der zweiten Sakkade lassen sich weitere Evidenzen für den Effekt der kontextuellen Salienz auf die Selektion von *starting points* finden. Im Unterschied zu Experiment 1, in dem die zweite Sakkade größtenteils auf das Agens gerichtet wurde, was auf ein konsistentes Auswählen des Agens als *starting points* hindeutet, wurden die zweiten Sakkaden in diesem Experiment größtenteils vom Topik angezogen, ohne dabei dessen semantische Funktion zu berücksichtigen. Die hohe Antizipation der kontextuellen Salienz für den zweiten Fixationsort dient daher als Indikator dafür,

dass die kontextuelle Salienz bzw. die Topikalität einen Einfluss auf die Selektion von *starting points* der Satzplanung hat.

Die oben erwähnten visuellen und linguistischen Befunde lassen den Rückschluss zu, dass die Selektion von *starting points* bzw. die syntaktische Enkodierung von der kontextuellen Salienz bzw. der Topikalität beeinflusst wird. Damit können die in der Literatur beschriebenen Resultate bestätigt werden (u.a. Cowles & Ferreira, 2012). Des Weiteren wurde unser Erkenntnisstand über das Verhältnis zwischen der kontextuellen Salienz und der Satzplanung dadurch erweitert, dass wir einen Manipulationseffekt der Topikalität auf die erste Fixationsposition festgestellt haben. Dieses Ergebnis weist darauf hin, dass auch die visuelle Wahrnehmungsphase unter dem Einfluss kontextueller Informationen steht.

Abbildung des Topiks auf die Syntax

Des Weiteren geht das Experiment der Frage nach, wie die Topik-Information, die der konzeptuellen Enkodierungsebene angehört und durch die Mikroplanung erstellt wird, linear positioniert wird. Bock und Levelt (1994) zufolge durchläuft ein Konzept nach dessen Aktivierung zuerst die semantische bzw. konzeptuelle Ebene und dann die Funktionszuweisungs- und Linearisierungsebene. Dementsprechend lassen sich hinsichtlich des Mappings der Topik-Information von der konzeptuellen Repräsentation auf die Wortstellung vier Hypothesen aufstellen (vgl. Abbildung 7.1 in Abschnitt 7.1.2): (1) Das Topik überspringt alle dazwischenliegenden Ebenen und trifft direkt auf die Linearisierungsebene. (2) Das Topik muss zuerst auf die semantische Enkodierungsebene abgebildet werden, bevor es unmittelbar in die Linearisierungsebene gelangt. (3) Das Topik muss zuerst auf die syntaktische Enkodierungsebene abgebildet werden, bevor es in die Linearisierungsebene gelangt. (4) Die Mappings auf die semantische und syntaktische Struktur gelten beide als Voraussetzung für die Anordnung des Topiks in einer linearen Wortsequenz.

Um die erste Hypothese, also die unmittelbare Abbildung der Topik-Information von der konzeptuellen Repräsentation auf die Wortstellung, zu überprüfen, wurden die

Speech-Onset-Latenzen der jeweils aus Experiment 1 und 2 selektierten Sätze miteinander verglichen. Wie bereits in Abschnitt 7.2.6 erläutert, erfüllten diese selektierten Sätze folgende Kriterien: (1) Alle Sätze waren Aktivsätze. (2) Aus Experiment 1 wurden nur diejenigen Aktivsätze ausgewählt, bei deren Produzieren der visuelle Cue auf der Patiens-Region auftauchte. (3) Aus Experiment 2 wurden nur diejenigen Aktivsätze ausgewählt, bei deren Produktion das Agens als Topik hervortrat.

Die Ergebnisse stützen die Annahme, dass das Topik nach dessen Aktivierung nicht direkt in die Linearisierungsebene gelangt. In den Sprechdaten fanden sich keine signifikanten Unterschiede der Speech-Onset-Latenzen zwischen den selektierten Aktivsätzen aus Experiment 1 und 2. Der Sprechvorgang in Experiment 2 wurde aufgrund der Topik-Einführung in der Preview-Phase nicht beschleunigt. Somit lässt sich annehmen, dass die Linearisierung in Experiment 2 aufgrund des in der Preview-Phase eingeführten Topiks nicht vorgezogen wurde und ein Topik nicht unmittelbar nach dessen Angabe satzinitial angeordnet wird.

Es gibt eine alternative Erklärung, der zufolge das Fehlen von Unterschieden der Speech-Onset-Latenzen aufgrund eines syntaktischen Buffersystems resultiert, welches sich zwischen der positionellen und der phonologisch-phonetischen Enkodierungsebene befindet (vgl. z. B. Levelt 1989). Dieses System hält die Topik-Information, die nach Bekanntwerden bzw. Erscheinen in der Preview-Phase direkt auf diese Ebene abgebildet wird, während der Verarbeitung in einem sog. "Arbeitsspeicher". Diese Erklärung halten wir allerdings für nicht überzeugend. Darauf gehen wir in Experiment 3 näher ein.

Ausgehend von dem Befund, dass Topiks wahrscheinlich nicht direkt auf die positionelle Verarbeitungsebene abgebildet werden, eröffnen sich drei weitere Möglichkeiten:

(1) Das Topik wird zunächst nur auf die semantische Ebene abgebildet, bevor es danach direkt in die Wortstellung gelangt.

(2) Das Topik wird zunächst nur auf die syntaktische Funktionszuweisungsebene abgebildet, bevor es in die Wortstellung gelangt.

(3) Die semantische und syntaktische Enkodierung des Topiks gelten beide als Voraussetzung für dessen Anordnung in einer Wortabfolge.

Im Folgenden beginnen wir zuerst mit der Überprüfung der zweiten Möglichkeit, die sich mit dem Abbildungsprozess innerhalb des grammatischen Enkodierungssystems befasst. Die Aufstellung dieser Hypothese geht von einer Auseinandersetzung zwischen dem *grammatical function account* und dem *word order account* aus (vgl. Abschnitt 2.4). In Bezug auf den *grammatical function account* werden präverbale Botschaften einschließlich der kontextuellen Information zuerst syntaktisch verarbeitet und danach in einer linearen Sequenz angeordnet. Dabei wird die Rolle syntaktischer Funktionen für die Konstituenten-Zusammenstellung hervorgehoben. Im Gegensatz dazu landen konzeptuelle Informationen dem *word order account* nach unmittelbar auf der Wortstellung, ohne syntaktisch funktional enkodiert werden zu müssen. Um diese beiden Annahmen im Rahmen der deutschen Satzproduktion voneinander unterscheiden zu können, werden die erhobenen Daten nun nachfolgend mithilfe von der Studie von Cowles und Ferreira (2012), die bereits im theoretischen Teil vorgestellt wurde, analysiert.

Ein wichtiges Ergebnis in diesem Zusammenhang ist, dass das Topik immer satzinitial angeordnet wurde. Obwohl das Subjekt als auch das Objekt am Satzanfang auftreten können, wurde das satzinitiale Topik immer als das Subjekt realisiert. Diese Korrespondenz weist darauf hin, dass deutsche Sprecher bewusst dazu tendieren, den Referenten mit Topikstatus einer bestimmten syntaktischen Funktion (hier dem Subjekt) zuzuordnen. Dieses Vorgehen spricht grundsätzlich für die Annahme des *grammatical function account* (nähere Diskussion siehe Myachykov & Tomlin (2008) und Branigan & Feleki (1999)).

Eine stärkere Evidenz diesbezüglich liefert die Studie von Cowles und Ferreira (2012), die in Abschnitt 4.4.2 vorgestellt wurde. Ihre Studie zielte darauf ab, den

Topikalitäts-Effekt auf die grammatische Enkodierung bzw. die syntaktische Selektion zu identifizieren. Einer der wichtigsten Befunde, der festgestellt wurde, ist die Voraussetzung der syntaktischen Funktionszuweisung des topikalischen Referenten für das Auftreten des Topik-Effekts. Die Probanden wurden nämlich dahingehend instruiert, anstatt einen Satz mit vorgegebenen Wörtern, unter denen ein bestimmtes Wort vorab durch einen Einführungssatz topikalisch markiert wurde, zu bilden, diese einfach in Form von koordinierten NPs wiederzugeben. Dabei war keine Funktionszuweisung nötig und es wurde kein Topik-Einfluss auf die Reihenfolge der wiedergegebenen Wörter entdeckt. Im Vergleich zu dem bei der Satzbildung erhaltenen Effekt, geht Wagner (2016) in seiner Diskussion über diese Studie davon aus, dass die Ausprägung des Topik-Effekts wahrscheinlich durch die Zuweisung syntaktischer Funktionen beeinflusst wird. Wenn in einem Verbalisierungsprozess keine syntaktische Funktionszuweisung in Anspruch genommen wird, bleibt wahrscheinlich ein Einfluss der Topikalität aus.

Aufgrund unseres eigenen Befundes und der Ergebnisse der Studie von Cowles und Ferreira (2012) nehmen wir daher an, dass Topiks sehr wahrscheinlich zuerst auf die Funktionszuweisungsebene abgebildet werden, bevor sie im Anschluss in die Linearisierungsebene überführt werden. Diese Verarbeitungsweise steht in seinen Kernpunkten im Einklang mit dem zweistufigen Modell bzw. dem *grammatical function account*. Darüber hinaus kann eine parallele Abbildung des Topiks auf die Funktionszuweisungs- und Linearisierungsebene nicht möglich sein. Denn wäre dies der Fall, sollten, ähnlich wie die Ergebnisse aus dem Japanischen (Tanaka et al., 2011, vgl. Abschnitt 2.4), zumindest ein paar Scrambled-Sätze produziert worden sein.

Zur Untersuchung der ersten Hypothese, nach der eine Lokalisation des Topiks in der semantisch strukturellen Repräsentation des Target-Ereignisses erforderlich ist, wird der Unterschied zwischen der Speech-Onset-Latenz von Aktiv- und Passivsatz herangezogen. In einer gleichen Satzproduktionssituation wurden Aktivsätze in der Regel schneller als Passivsätze produziert. Dieser Unterschied wird an dieser Stelle dazu genutzt, um zu überprüfen, ob die semantischen Funktionen des Topiks vor

Beginn der syntaktischen Enkodierung identifiziert werden.

Der Hypothese nach, dass das Topik vor der Weiterleitung in die syntaktische Ebene zuerst auf die semantisch strukturelle Repräsentation abgebildet wird, sollte zwischen den Speech-Onset-Latenzen der produzierten Aktiv- und Passivsätze, in denen das Agens und das Patiens jeweils als das Satztopik hervortreten, ein signifikanter Unterschied bestehen. Wenn aber Topiks, sobald sie in der Preview-Phase bekannt geworden sind, direkt an die linguistische Verarbeitungsebene weitergereicht und verbalisiert werden, ohne dass eine Abbildung auf die semantische Strukturierungsebene vorgenommen werden muss, sollte in dieser Hinsicht kein signifikanter Unterschied identifizierbar sein. Die Begründung liegt darin, dass topikalische Referenten mit der Agens-Rolle in den topik-initialen Aktivsätzen und mit der Patiens-Rolle in den topik-initialen Passivsätzen, ohne semantische Enkodierung eine gleichwertige semantische Salienz implizieren. Das bedeutet, dass deutsche Sprecher zum Zeitpunkt der Anordnung und Artikulation des Topiks wohl nicht wissen, was für eine semantische Rolle das Topik einnimmt und was für eine Satzstruktur sie produzieren. Demnach ist kein Unterschied der Sprechanfängszeitpunkte zu finden.

Allerdings konnte die oben erwähnte Analyse aufgrund des vorhandenen Versuchsdesigns nicht durchgeführt werden. Der Grund dafür liegt darin, dass das Produzieren der Aktiv- und Passivsätze in diesem Experiment nicht unter einer gleichen visuellen Bedingung erfolgte. Diese Diskrepanz wurde von der Position des visuellen Cues verursacht. Der visuelle Cue, der die erste Fixationsverteilung beeinflussen kann, tauchte in dem vorliegenden Experiment immer auf der Patiens-Position auf. So wird aus der Perspektive der visuellen Salienz eine vorteilhafte Situation für das Produzieren von Passivsätzen und eine nachteilige Situation für das von Aktivsätzen generiert. Bei der Bildung eines Aktivsatzes können Sprecher zuerst das Patiens fixieren, bevor sie ihren Blick auf den tatsächlichen *starting point*, also das Agens, richten. Ein derartiger Sakkadenwechsel ist allerdings beim Produzieren der Passivsätze in den meisten Fällen nicht nötig. Da die Planung

des sakkadischen Wechsels bei der Bildung eines Aktivsatzes zusätzliche Planungszeit benötigt, die ebenfalls in die Sprechanfängszeit eingeplant wird, lässt sich nicht genau ableiten, wie viel Zeit deutschsprachige Probanden zur Bildung von Aktivsätzen gebraucht haben. In Anbetracht der im Allgemeinen unterschiedlichen visuellen Reaktionsmuster während der Produktion der Aktiv- und Passivsätze wurde eine Abbildung der kontextuellen Information auf die semantische Ebene in diesem Experiment nicht überprüft. Dieser Hypothese wird in Experiment 3 nachgegangen.

Repräsentation und Wirkungsweise des Topiks innerhalb des Satzproduktionssystems

An dieser Stelle kann nun festgehalten werden, dass die kontextuelle Salienz im Satzproduktionsprozess einen hohen Stellenwert einnimmt. Allerdings ist bislang noch nicht klar, inwiefern kontextuelle Informationen, hier vor allem das Topik, mental repräsentiert sind und in welcher Art und Weise sie bei der Satzplanung wirken. Im Folgenden werden die beiden Fragen thematisiert.

Unsere Befunde untermauern in Verbindung mit denen von Cowels und Ferreira (2010), dass ein unmittelbares Mapping des Topiks von der Konzeptualisierungs- auf die Linearisierungsebene, welches sich ohne Rücksicht auf die syntaktischen Funktionen der betroffenen Entität vollzieht, unmöglich zu sein scheint. Stattdessen werden topikalische Referenten höchst wahrscheinlich zuerst in die syntaktische Funktionszuweisungsebene überführt und dort syntaktisch funktional enkodiert, bevor sie dann später auf die positionelle Verarbeitungsebene treffen und satzinitial angeordnet werden. Angesichts des engen Zusammenspiels zwischen dem Diskursstatus einer Entität und deren syntaktischer Funktion (z. B. die hohe beobachtete Korrespondenz zwischen dem Topik und dem Subjekt) wird angenommen, dass einzelne syntaktische Elemente (z. B. das Subjekt) gewissermaßen mit informationsstrukturellem Status (z. B. Topik-Status) „beflaggt“ werden können (vgl. Rizzi, 1997). Damit die positionellen Bedarfe solcher kontextuellen Eigenschaften, die an syntaktische Funktionen gekoppelt sind, „zufriedengestellt“ werden können, müssen die syntaktischen Rollen ihren Weg zur

richtigen Position innerhalb des zu produzierenden Satzes finden (vgl. Arnold et al., 2013). Mit anderen Worten, die Anordnung einer Figur in einer linearen Wortsequenz wird nicht nur von ihrer syntaktischen Funktion beeinflusst (z. B. das Subjekt im Deutschen kann sowohl am Satzanfang als auch im Mittelfeld erscheinen), sondern sie muss auch der Eigenschaft der kontextuellen Information, die an die syntaktische Funktion gebunden ist, Rechnung tragen.

Somit ist unser Befund, dass die Subjekte der gebildeten Sätze immer zum Satzanfang arrangiert und keine Scrambled-Sätze mit OVS-Struktur produziert werden, nicht zufällig. Im Gegenteil, das verläuft planmäßig: Obwohl das Subjekt im Deutschen sowohl an den Satzanfang als auch ins Mittelfeld gestellt werden kann, wird diese Selektionsflexibilität aufgrund der Verbindung zwischen dem Subjekt und dem Topik begrenzt. Aufgrund der Tatsache, dass Topiks typischerweise an das Vorfeld eines Satzes gebunden sind, ist es für Sprecher daher nötig, das Subjekt an diese Position zu positionieren, um dem Bedürfnis des Topiks, satzinitial zu stehen, zu entsprechen.

Neben der linguistischen Repräsentation ist die Topik-Erstellung Levelt (1989) zufolge ein wichtiger Bestandteil der Mikroplanung, die innerhalb des Konzeptualisierungssystems stattfindet. Aus diesem Grund ist die Annahme berechtigt, dass das Topik wahrscheinlich auch eine prälinguistische konzeptuelle Repräsentation hat. Da sich der Konzeptualisierungsprozess in diesem Experiment vor allem in der ersten und zweiten Sakkade/Fixation ausdrückt, betrachten wir nun zuerst die Beziehung zwischen der Topikalität und den beiden Sakkaden. Das Ziel dabei ist, die Art und Weise, mit der das Topik die Konzeptualisierung manipuliert, zu identifizieren.

Das zentrale Ergebnis diesbezüglich ist, dass sowohl die erste Sakkade, welche normalerweise reizbezogen ist und den visuellen Wahrnehmungsprozess widerspiegelt, als auch die zweite Sakkade, durch die die Selektion von *starting points* reflektiert werden kann, größtenteils auf das Topik gerichtet wurden. Aus kognitionspsychologischer Sicht läuft die Selektion einer sakkadischen Zielposition

nicht wahllos ab, sondern ihr liegt ein sog. Filtervorgang zugrunde, durch den nur die relevanten Informationen aufgenommen und zuverlässig verarbeitet werden. Diesen Vorgang gelingt durch die sog. *selektive Aufmerksamkeit* (Niemann & Gauggel 2010, S. 147).

Mithilfe der Verknüpfung zwischen der Sakkade und der Aufmerksamkeit kann aus der Tatsache, dass Topiks sowohl die Planung der ersten als auch der zweiten Sakkade beeinflussen können, abgeleitet werden, dass sie grundsätzlich in einem engen Zusammenhang mit dem Aufmerksamkeitsstatus einer Entität stehen. Denjenigen Entitäten, die als Topik einer Äußerung hervortreten und daher kontextuell prominenter sind, wird entsprechend mehr Aufmerksamkeit gewidmet als denjenigen, die eine kontextuell niedrigere Prominenz aufweisen (vgl. Lambrecht, 1994; Levelt, 1989; Reinhart, 1982). Daher ist anzunehmen, dass Topiks neben ihrer linguistischen Repräsentation auch konzeptuell einen hohen Aufmerksamkeitsstatus haben können.

Doch in welcher Weise wird die Konzeptualisierung durch den Aufmerksamkeitsstatus des Topiks beeinflusst? Um diese Frage zu beantworten gehen wir nun zuerst auf ein berühmtes Modell ein, welches ursprünglich vom amerikanischen Psychologen Michael Posner (1980) zur Beschreibung der grundlegenden Selektionsmechanismen der Aufmerksamkeit vorgeschlagen wurde: das sog. *Spotlight-Modell*.

Dieses Modell nimmt grundlegend an, dass Menschen ihre Aufmerksamkeit sozusagen wie einen „Scheinwerfer“ „umherbewegen“, um verschiedene Punkte des Sichtfeldes zu fokussieren. Diese Annahme basiert auf der Beobachtung, dass obwohl jede Sekunde unmessbare visuelle Reize aus der Umwelt auf das menschliche Gehirn einwirken, die Menschen immer in der Lage sind, relevante Reize aus dem gesamten Feld herauszufiltern und ihre Aufmerksamkeit schnell auf den ausgewählten Reiz zu richten (Broadbent, 1966). Doch wie gelingt dies?

Laut Posner (1980) kann die selektive Aufmerksamkeit als sog. attentionaler Strahl eines mentalen Scheinwerfers bezeichnet werden, welcher konstant nach Ausschnitten

des Reizangebotes sucht, die einen hohen Aufmerksamkeitsstatus aufweisen. Wenn sich jedoch eine Entität aufgrund ihres hohen Aufmerksamkeitsstatus von den anderen abhebt, wird diese vom attentionalen Scheinwerfer “beleuchtet” und in den Fokus der Verarbeitung gerückt, was wiederum zu einer früheren Enkodierung führen kann. Die anderen Figuren, die sich außerhalb des Lichtkegels befinden, werden dann wenig bis kaum verarbeitet.

Die Vorstellung, die dem Spotlight-Modell zugrunde liegt, stammt ursprünglich von der Studie von Posner (1980). In seiner Studie hat er mithilfe des *spatial-cueing*-Paradigmas die covert Aufmerksamkeit ausrichtung untersucht. Das Experiment lief wie folgt ab: Die Versuchspersonen sollten einen zentralen Fixationspunkt in der Mitte eines Bildschirms fixieren. Per Tastendruck mussten sie dann so schnell wie möglich das Auftreten von Zielreizen, die entweder auf der linken oder der rechten Seite des Punktes erschienen, melden. Ein symbolischer, endogener Cue (Pfeil) wurde kurz vor Einsetzen des Zielreizes präsentiert. Die Probanden wurden darüber informiert, dass ein Zielreiz mit einer hohen prozeduralen Wahrscheinlichkeit von 80% auf der durch den Cue markierten Seite erscheinen wird (valide Bedingung). In 20% der Fälle war der Cue falsch und ein Zielreiz tauchte am nicht indizierten Sichtfeld auf (invalide Bedingung). Neben diesen zwei Bedingungen gab es noch eine Kontrollbedingung, in der ein Cue (zentrales Kreuz) mit einer Validität von 50% keine ortsbezogene Information lieferte.

Die Ergebnismuster der Reaktionslatenz unterschieden sich unter der validen und invaliden Bedingung. Es gab einen sog. Beschleunigungseffekt (sog. Nutzen-Effekt), wenn die richtige Zielposition vorab durch den endogenen Cue angezeigt wurde. Die Reaktionszeit in Bezug auf den Zielreiz fiel im Gegensatz zur Kontrollbedingung deutlich kürzer aus. Dagegen tauchte ein Interferenzeffekt auf, wenn der endogene Cue die Position des Zielreizes nicht angeben konnte (sog. Kosten-Effekt). Doch wie lässt sich dieser Unterscheid der Reaktionsgeschwindigkeit erklären?

Posner interpretierte den beobachteten Nutzen- und Kosten-Effekt als experimentellen

Hinweis auf Effekte der versteckten Aufmerksamkeitsverlagerung. Er nimmt an, dass durch den vorab dargebotenen Cue überproportional viele Aufmerksamkeitsressourcen an die durch den Cue angezeigte Region weitergeleitet wurden. Die Konzentration der Aufmerksamkeit auf ein bestimmtes Sichtfeld kann metaphorisch als eine Art Beleuchtung eines attentionalen Scheinwerfers angesehen werden. Wenn der Zielreiz in demselben Sichtfeld erscheint, welches vom kognitiven Scheinwerfer beleuchtet wird, wird er aufgrund der intensiven Aufmerksamkeit früher und intensiver verarbeitet, was dann zu einer schnelleren Reaktion führt. Angesichts der Annahme, dass das Aufmerksamkeitssystem limitierte Kapazitäten besitzt, bedeutet die gehäufte Aufmerksamkeit auf eine Position aber auch, dass anderen Positionen weniger Aufmerksamkeit gewidmet wird. Dies kann wiederum eine schlechtere Verarbeitung zur Folge haben.

Ursprünglich wurde die Spotlight-Metapher von Posner zur Beschreibung der Selektionsmechanismen der visuellen Aufmerksamkeit bei der räumlichen Verarbeitung vorgeschlagen und steht in keinem direkten Zusammenhang mit der hier untersuchten Satzplanung. Trotzdem kann das Modell aber auch Anregungen für die Interpretation des in diesem Experiment identifizierten topikalischen Effekts liefern, da die Satzproduktion ähnlich wie visuelle Selektion auch selektive Aufmerksamkeit erfordert.

Zur Interpretation der Verarbeitungsmechanismen der Topikalität wird hier zuerst ein Vergleich zwischen dem experimentellen Design des vorliegenden Experiments und dem von Posner (1980) durchgeführt.

In der Preview-Phase des vorliegenden Experiments wurden den Sprechern anhand eines Einführungsbilds und eines Einführungssatzes zwei Figuren, sowohl visuell als auch lexikalisch, präsentiert. Da den Probanden vor Experimentbeginn mitgeteilt wurde, dass die beiden Entitäten in einem später gezeigten Target-Bild ein transitives Ereignis bilden werden, wurden beide bis zum Zeitpunkt der Einführung des Topiks als gleichwertige Kandidaten für den *starting point* der Satzplanung angesehen. Dies

gleich der Phase des Experiments von Posner (1980), in der der endogene Cue noch nicht präsentiert wurde und sowohl die linke als auch die rechte Hälfte des Bildschirms mögliche Darstellungsorte des Zielreizes waren.

In einem zweiten Schritt des vorliegenden Experiments wurde einem bestimmten Ereignisteilnehmer durch die Stellung einer Target-Frage der Topikstatus der kommenden Äußerung zugewiesen, der dann wiederum mit einer höheren Wahrscheinlichkeit als *starting point* selektiert wird. Diese Target-Frage ist im Wesentlichen mit dem visuellen Hinweisreiz in Posners Experiment vergleichbar, der die Position des nachfolgenden Zielreizes mit einer hohen Vorhersagbarkeit anzeigt.

In unserem Experiment kam heraus, dass die erste Sakkade anstatt reizbedingt auf die durch den Cue angezeigten Figuren, größtenteils auf diejenigen Figuren gerichtet wurde, die den Topikstatus innehatten und im späteren Satzproduktionsprozess zuerst verbalisiert wurden. Dieses Blickmuster kann als Hinweis darauf gesehen werden, dass mit der Suche nach *starting points* in dem vorliegenden Experiment dementsprechend früher begonnen wurde. Da der Suchvorgang massive Aufmerksamkeit beansprucht, lässt sich aus der Präferenz der ersten Fixation für topikalische Referenten ableiten, dass die Aufmerksamkeit direkt nach der Zuweisung des Topikstatus vom kognitiven System auf die topikalische Figur gelenkt wird. Das bedeutet, dass der Referent, welcher durch einen Kontext zum Topik gemacht werden, vom attentionalen Scheinwerfer beleuchtet, nach Stimulus-Onset sofort fixiert und dann konzentriert verarbeitet wird. Diejenige Figur, die eine verringerte Diskursprominenz aufweist und sich deswegen außerhalb des Spotlights der konzeptuellen Verarbeitung befindet, wird ausgeblendet und möglicherweise erst zu einem späteren Zeitpunkt vom mentalen Scheinwerfer beleuchtet.

Prioritätenliste der mitwirkenden Faktoren

Eine andere, ebenfalls für die vorliegende Studie zentrale Frage ist, mit welcher Gewichtung die bei der Satzproduktion mitwirkenden Faktoren in den Konzeptualisierungsprozess eingehen. Mit dem Beitrag der Topikalität zur

Satzproduktion wird in diesem Teil die Wirksamkeit der visuellen, semantischen und kontextuellen Salienz miteinander verglichen.

In Bezug auf die visuelle Wahrnehmung, die sich in der Ausrichtung der ersten Sakkade/Fixation äußert, wird der Effekt der visuellen und kontextuellen Salienz prüfend gegenübergestellt. Die semantische Salienz wird nach dem Befund aus Experiment 1 aus dieser Auswertung ausgeschlossen. Dieser besagt, dass die Positionierung der ersten Sakkade nicht von den semantischen Funktionen der Referenten modifiziert wird.

Wie bereits im vorangegangenen Teil der Diskussion vorgestellt, haben die durchgeführten Analysen gezeigt, dass sowohl die Topikalität als auch die visuelle Salienz einen Einfluss auf die Bestimmung der ersten Sakkade haben. Das bedeutet, dass beide Faktoren im Wahrnehmungsprozess mitgewirkt haben. Allerdings sind ihre Einflüsse nicht gleich stark. Wenn der exogene Cue und der Topikstatus auf verschiedene Referenten fielen, war der Referent mit dem Topikstatus dem visuell markierten Referenten hinsichtlich der Vorhersagbarkeit auf den ersten Fixationsort überlegen. Dieses Ergebnis spricht dafür, dass bei der Festlegung der ersten Fixation bzw. während der visuellen Stimuli-Wahrnehmung die kontextuelle Salienz einen stärkeren Einfluss ausübt als die visuelle Salienz.

Hinsichtlich der Selektion von *starting points* wird die Effekte der kontextuellen und der semantischen Salienz miteinander verglichen. Da in Experiment 1 kein modifizierender Einfluss der visuellen Salienz zu finden war, wird dieser Faktor daher nicht in die Analyse einbezogen. Die Auswertung erfolgt in der Situation, in der das Patiens das Topik der zu bildenden Sätze war. Es ging nämlich darum, ob die zweite fixierte Figur eher das Agens oder das topikalische Patiens war. Während nur eine sehr kleine Anzahl der zweiten Sakkade nach der Fixation auf dem Patiens wieder auf das Agens gerichtet wurde, blieben die meisten weiterhin auf dem topikalischen Patiens. Der signifikante Unterschied in der Vorhersagbarkeit auf den zweiten Fixationsort weist darauf hin, dass der kontextuellen Salienz bei der Planung der

zweiten Sakkade eine dominante Rolle zukommt, wohingegen die semantische Salienz in ihrer Bedeutung nach der Topik-Einführung heftig beeinträchtigt wird.

Der stärkere Manipulationseffekt der kontextuellen Salienz auf die Selektion von *starting points* wurde ebenfalls durch die syntaktischen Strukturen der produzierten Sätze nachgewiesen. Das Agens, das in Experiment 1 durchgehend in der satzinitialen Position stand, wurde nur dann ins Vorfeld eines Satzes gerückt, wenn es gleichzeitig das Satztopik war. Wenn allerdings das Patiens das Topik war, dann wurden die meisten produzierten Sätze passivisch realisiert, in denen das Patiens dann als Subjekt satzinitial kodiert wurde. Die starke Korrespondenz zwischen dem Topik und dem ersten Referenten eines Satzes und der Abbau der sog. Agens-Präferenz lassen schlussfolgern, dass die Selektion von *starting points* überwiegend von der kontextuellen Salienz beeinflusst wird. Es lässt sich also festhalten, dass die kontextuelle Salienz eine ausschlaggebende Rolle bei der Konzeptualisierung spielt.

Fazit

Experiment 2 hat insgesamt gezeigt, dass der kontextuellen Salienz eine dominante Rolle in dem Konzeptualisierungsvorgang zukommt. Sowohl die Stimuli-Wahrnehmung als auch die Selektion von *starting points* werden größtenteils von der kontextuellen Salienz der an einem Ereignis beteiligten Referenten beeinflusst.

Auf dem Status der kontextuellen Salienz basierend, stellt sich die Frage, wie das Topik mental repräsentiert ist und wie es bei der Satzproduktion wirkt. Als Fazit wird angenommen, dass das Topik bei der Satzplanung sowohl linguistisch als auch konzeptuell repräsentiert wird. Einerseits fungiert das Topik als informationsstrukturelle „Etikette“, die syntaktischen Funktionen wie etwa dem Subjekt eine positionelle Eigenschaft verleiht. Andererseits steht es im Mittelpunkt des Gesprächsinteresses und weist somit einen hohen Aufmerksamkeitsstatus auf. Aufgrund der linguistischen Repräsentation des Topiks wird angenommen, dass die Einordnung eines Referenten in eine Wortfolge nicht nur von seiner syntaktischen

Funktion beeinflusst wird, sondern auch die positionelle Anforderung des Topiks, das an diese syntaktische Funktion gebunden ist, gerecht werden muss. Von der konzeptuellen Repräsentation des Topiks ausgehend, lässt sich das Phänomen, dass Topiks sehr häufig als *starting points* der Satzplanung ausgewählt werden, anhand der Selektionsmechanismen der Aufmerksamkeit erklären. Für Posner (1980) ist dieser mit einem sich im menschlichen Gehirn bewegenden „Scheinwerfer“ vergleichbar. Da das Topik einen höheren Aufmerksamkeitsstatus aufweist, werden diejenigen Figuren, die den Topikstatus innehaben, dementsprechend früher vom kognitiven Scheinwerfer beleuchtet und dem Sprachproduktionssystem zur Verfügung gestellt.

In dieser Studie wurde zudem der Zeitverlauf der Topik-Verarbeitung thematisiert, über den verschiedene Hypothesen aufgestellt wurden. Die Hypothese, dass das Topik nach dessen Aktivierung unmittelbar auf die lineare Position des zu bildenden Satzes abgebildet wird, wurde nicht belegt. Hierzu wurde die Speech-Onset-Latenz der selektierten Sätze aus Experiment 1, an deren Produzieren die kontextuelle Salienz als Faktor nicht beteiligt war, und der aus Experiment 2, in denen der zuerst erwähnte Referent das Satztopik war, verglichen. Es wurde kein signifikanter temporaler Unterschied entdeckt. Die Satzproduktion der Sätze in Experiment 2 wurde trotz der Kenntnisnahme des Topiks vor Stimulus-Onset nicht beschleunigt. Ferner wurde überprüft, wie das Topik innerhalb des grammatischen Enkodierungssystems überführt wird. Aufgrund der gefundenen Korrespondenz zwischen dem Topik und dem Subjekt wurde der *grammatical function account* untermauert. Das bedeutet, dass sich ein Mapping der kontextuellen Information auf die syntaktische Funktionszuweisungsebene zuerst vollziehen sollte, bevor sie zum Schluss in einer linearen Sequenz angeordnet wird. Diese Schlussfolgerung wird auch durch die Resultate von Cowles und Ferreira (2012) gestützt.

In Bezug auf die Wirksamkeit der in Experiment 2 mitwirkenden Faktoren, also der semantischen, visuellen und kontextuellen Salienz, wurden die Einflüsse der kontextuellen Salienz bei der Konzeptualisierung und der linguistischen Verarbeitung hervorgehoben. Die Effekte der visuellen und semantischen Salienz wurden nach

Topik-Einführung einigermaßen abgeschwächt oder sogar abgebaut.

Bisher ist die Frage, ob es eine Verbindung zwischen der kontextuellen Information und der semantischen Funktion gibt, die sich darin zeigt, dass die kontextuelle Information zuerst auf die semantische Struktur abgebildet werden muss, bevor sie an die linguistische Verarbeitung weitergereicht wird, noch unzureichend beantwortet. Sie ermöglicht einen Einblick in den Mechanismen der Mikroplanung auf der Konzeptualisierungsebene und soll im nachfolgenden Kapitel beantwortet werden.

8 Experiment 3

In Experiment 3 sollte die Verarbeitungsmechanismen der Mikroplanung, die für die Generierung der semantischen und kontextuellen Information zuständig ist, untersucht werden. Die Frage, die beantwortet werden soll, lautet wie folgt: Wird die kontextuelle Information bzw. das Satztopik vor Beginn der grammatischen Enkodierung zuerst semantisch identifiziert, oder setzt sich die grammatische Enkodierung des Topiks über dessen Einordnung in eine semantische Rollenstruktur hinweg?

Um diese Frage beantworten zu können - wie bereits in Abschnitt 7.4 angemerkt - lässt sich der Unterschied zwischen der Speech-Onset-Latenz der produzierten topik-initialen Aktiv- und Passivsätze heranziehen. Allgemein gesagt werden Aktivsätze unter gleichen Bedingungen schneller produziert als Passivsätze. Diese Diskrepanz bezüglich der Speech-Onset-Latenz ist auf die unterschiedlichen semantischen Reihenfolgen der gebildeten Sätze zurückzuführen. Daher ist anzunehmen, dass die Speech-Onset-Latenzen der topik-initialen Aktiv- und Passivsätze einen signifikanten Unterschied aufweisen sollten, wenn semantische Funktionen bereits vor Beginn der grammatischen Enkodierung des Topiks zugewiesen würden. Im Gegensatz dazu wäre kein signifikanter Unterschied zu beobachten, wenn Topiks, ohne dass sie in die semantische Vorstellung des Target-Ereignisses eingeordnet werden, direkt an die linguistische Enkodierungsebene weitergereicht werden. Denn ohne die Zuweisung der semantischen Funktion zum Zeitpunkt der Anordnung und Artikulation des Topiks wissen Sprecher noch nicht, ob das Topik die Agens- oder Patiens-Rolle einnimmt bzw. ob sie einen Aktiv- oder Passivsatz bilden.

Diese Auswertung wurde in Experiment 2 allerdings nicht durchgeführt. Der Grund dafür liegt darin, dass das Produzieren der topik-initialen Aktiv- und Passivsätze im letzten Experiment mit dem visuellen Cue auf der Patiens-Region unter ungleichmäßigen visuellen Bedingungen erfolgte (vgl. die Diskussion in Abschnitt

7.4). Um eine gleichmäßige visuelle Bedingung für die Produktion der zu vergleichenden Sätze zu schaffen, wird der visuelle Cue in Experiment 3 lediglich auf die Agens-Position gerichtet. Die Speech-Onset-Latenzen der topik-initialen Passivsätze, die in Experiment 2 produziert wurden, und die der topik-initialen Aktivsätze, die in Experiment 3 produziert wurden, werden dann miteinander verglichen.

Ein weiteres Ziel dieses Experiments ist es, die zentralen Befunde aus Experiment 2, dass nämlich die kontextuelle Salienz sowohl bei der visuellen Wahrnehmung als auch bei der Selektion von *starting points* eine dominante Rolle spielt, zu replizieren. Die visuellen Augenbewegungen und die produzierten Sätze werden, ähnlich wie in Experiment 2, weiterhin als Indikatoren für genannte Prozesse angeführt.

8.1 Hypothesen

8.1.1 Hypothesen über den Abbildungsprozess des Topiks auf die Semantik

Um zu überprüfen, ob Topiks vor deren grammatischer Enkodierung zuerst auf die semantisch strukturelle Repräsentation eines Ereignisses abgebildet werden oder nicht, werden die Speech-Onset-Latenz der in Experiment 3 gebildeten topik-initialen Aktivsätze und die der in Experiment 2 gebildeten topik-initialen Passivsätze miteinander verglichen. Die zum Vergleich ausgewählten Sätze wurden sowohl visuell als auch kontextuell unter einer vergleichbaren Bedingung produziert. Der einzige Unterschied zwischen den beiden Satzgruppen besteht in der semantischen Rolle des Topiks bzw. der semantischen Reihenfolge der Sätze.

Wie oben angeführt sollte das Sprechen mit der Aktivkonstruktion früher begonnen werden als das mit der Passivkonstruktion, wenn das Topik bereits vor der grammatischen Enkodierung in die semantische Vorstellung eingeordnet wird. Im Gegensatz dazu sollte die Speech-Onset-Latenz der Aktivsätze nicht deutlich kleiner ausfallen als die der Passivsätze, wenn das Topik nicht auf die semantische

Strukturierungsebene abgebildet wird.

8.1.2 Hypothesen über den Effekt der kontextuellen Salienz bzw. der Topikalität auf die Satzplanung

Die Hypothesen über die Funktionalität der kontextuellen Salienz bzw. der Topikalität bei der Satzplanung waren dieselbe wie in Experiment 2. Sowohl die Blickbewegungsdaten als auch die produzierten Sätze werden zur Überprüfung ausgewertet. In Bezug auf das Verhältnis zwischen dem Topik und der visuellen Wahrnehmung wird analog zum zweiten Experiment auf die Auswahl des ersten Fixationsortes eingegangen. Es soll diesbezüglich geprüft werden, ob der visuelle Cueingeffekt von der kontextuellen Salienz der markierten Figur beeinflusst wird. Liegt der visuellen Wahrnehmung die kontextuelle Salienz der jeweiligen Referenten zugrunde, sollte eine Abhängigkeit des visuellen Cueingeffekts vom Diskursstatus des visuell markierten Referenten, also des Agens, festgestellt werden.

Sowohl die Korrespondenz zwischen Topik und dem ersten Referenten der gebildeten Sätze als auch die Korrespondenz zwischen Topik und der Zielposition der zweiten Sakkaden stehen im Fokus des Experimentes. Auf diese Weise ist es möglich zu bewerten, ob die kontextuelle Salienz einen Einfluss auf die Selektion von *starting points* hat. Wenn die kontextuelle Salienz dabei eine Rolle spielt, sollte, ähnlich wie in Experiment 2, ein Abbau der Agens-Präferenz beobachtet werden. Anstatt einer starken Korrespondenz zwischen dem ersten Referenten und dem Agens, sollte das erste Wort generell mit dem Topik übereinstimmen. Ferner sollte die zweite Sakkade größtenteils auf topikalische Referenten gerichtet werden, ohne deren semantische Rollen berücksichtigen zu müssen.

8.2 Methodisches Vorgehen

8.2.1 Versuchspersonen

Es wurden 32 Versuchspersonen (davon 26 Frauen, Median 22,6 Jahre, Bereich 19-31 Jahre) untersucht. Alle Teilnehmer waren Studierende an der Universität Heidelberg

und wuchsen mit dem Deutschen als Muttersprache auf. Die Sehkraft der Probanden war auf dem normalen Niveau oder wurde mithilfe von Kontaktlinsen oder Brillen dahingehend korrigiert. Die Teilnahme erfolgte freiwillig, und alle Probanden erhielten vom Experimentleiter eine Aufwandsentschädigung in Höhe von 7 Euro.

8.2.2 Technische Apparatur

Die technische Apparatur war dieselbe wie in Experiment 1 und 2.

8.2.3 Material

Im Prinzip waren alle Bild-Stimuli und linguistischen Stimuli, die in Experiment 3 verwendet wurden, identisch mit denen in Experiment 2.

8.2.4 Durchführung

Die Durchführung und der Ablauf eines Durchgangs waren genauso wie in Experiment 2, mit dem einzigen Unterschied, dass *der visuelle Cue vor der Präsentation der Target-Stimuli immer im Sichtfeld des Agens gezeigt wurde.*

8.2.5 Datenkodierung

Die produzierten Sätze und die Augenbewegungsmessungen wurden genau so aufbereitet wie in Experiment 1 und 2.

8.2.6 Datenbearbeitung und -analyse

Zur Überprüfung, ob die kontextuelle Information auf die semantisch strukturelle Repräsentation eines Ereignisses abgebildet wird, werden *die Speech-Onset-Latenzen der aufgezeichneten topik-initialen Aktiv- und Passivsätze* mittels eines Zweistichproben-*t*-Tests aus der Perspektive der *Stimuli* ausgewertet. Eine Analyse auf der Perspektive der *Probanden* ist hier nicht möglich, da Experiment 2 und 3 nicht von denselben Probanden erledigt wurden.

Um den Einfluss der Topikalität auf die visuelle Wahrnehmung und die Selektion von *starting points* auszuwerten, werden analog zu Experiment 1 die Zielpositionen der ersten und zweiten Sakkade betrachtet. Zuerst wird getestet, ob die erste Fixation trotz

der Präsenz des visuellen Cues auch von der Topikalität manipuliert wird. Zur Auswertung dieses Verhältnisses wird *der kontextuelle Status des Agens (Topik vs. Nicht-Topik)*, welches in diesem Experiment immer vom visuellen Cue indiziert wurde, als feste unabhängige Variable in eine logistische Regressionsanalyse mit gemischten Effekten eingeführt. Es wäre zu beobachten, ob der Effekt des visuellen Cues vom kontextuellen Status des Agens modifiziert wird.

Neben dem Zusammenhang zwischen der visuellen Wahrnehmung und der Topikalität wird auch die Relation zwischen der Selektion von *starting points* und der Topikalität anhand der produzierten syntaktischen Strukturen betrachtet.

Hierbei wird geprüft, ob die Agens-Präferenz, die sich vor allem in der Zielposition der zweiten Sakkade und der selektierten syntaktischen Struktur widerspiegelt, vom kontextuellen Status des Agens beeinflusst wird. Dazu wird analog zum zweiten Experiment noch einmal die logistische Regression durchgeführt, in der *der kontextuelle Status des Agens (Topik vs. Nicht-Topik)* als unabhängige Variable dienen. Die abhängigen Variablen sind jeweils *die Zielposition der zweiten Sakkade (Agens vs. Nicht-Agens)* und *die selektierten syntaktischen Strukturen (kanonische-Aktivsätze vs. Passivsätze)*.

8.3 Ergebnisse

Fehler

Insgesamt wurden 896 Sätze produziert. Davon werden ca. 5 % der erhobenen Sätze aufgrund von Fehlern von der statistischen Analyse ausgeschlossen. Es wurden dieselben Fehlerkategorien angelegt wie in Experiment 1 und 2. Tabelle 8.1 zeigt genau die Entstehung der Fehler.

Fehlerkategorie	Anzahl
Technische Messfehler	1 (0.1%)
Fehler Performanz der Sätze insgesamt	44 (4.9%)
Fehler Performanz der Sätze detailliert	44 (100%)
Sätze mit Vertauschung	2 (4.5%)
Sätze mit intransitiven Verben	39 (88.7%)
Sätze mit nur einem Ereignisteilnehmer	1 (2.3%)
Sätze mit koordinierten NPs	0 (0%)
Antworten aus mehreren Sätzen	2 (4.5%)

Tabelle 8.1: Fehler (Experiment 3)

Speech-Onset-Latenz der selektierten topik-initialen Aktiv- und Passivsätze

Die Frage, ob Topiks vor Beginn der grammatischen Enkodierung semantischen Funktionen zugeordnet werden, wurde durch den Vergleich zwischen den Speech-Onset-Latenzen der ausgewählten Aktiv- und Passivsätze beantwortet. Das Ergebnis war, dass Aktivsätze, in denen das Agens das Satztopik war, signifikant schneller produziert wurden als diejenigen, in denen das Patiens als Satztopik satzinitial aufgetreten ist: Während die durchschnittliche Sprechanfängszeit der zum Vergleich gebrachten Passivsätze bei 1641 ms lag, lag die der Aktivsätze bei 1466 ms, $t(27) = - 4.9127$, $p < .01$ (vgl. Abbildung 8.1). Dieser temporale Unterschied deutet daher darauf hin, dass die semantischen Funktionen von Topiks identifiziert werden, bevor deutsche Sprecher die anschließende grammatische Enkodierung beginnen.

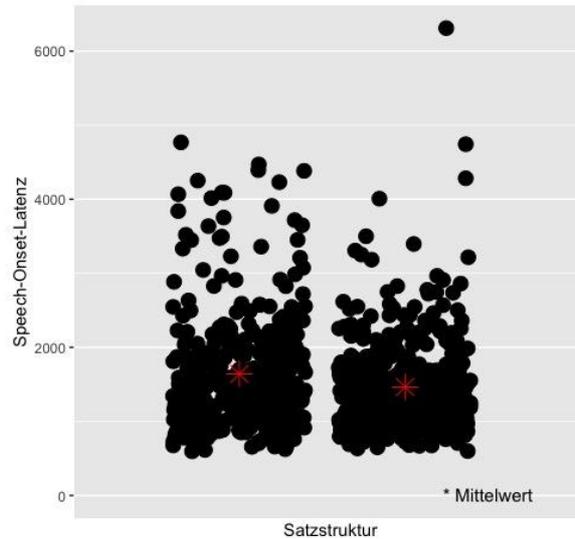


Abbildung 8.1 Punktdiagramm der Speech-Onset-Latenz der topik-initialen Passivsätze (links) und der topik-initialen Aktivsätze (rechts) (in Millisekunde)

Kontextuelle Salienz bzw. Topikalität und erste Fixation

Die *z*-Werte der logistischen Regressionsanalyse spiegeln die gleichen Verhältnisse wider, die wir bereits in Experiment 2 erkannt haben (vgl. Tabelle 8.2). Ein Haupteffekt der Topikalität wurde bei der Planung der ersten Sakkade gezeigt. Wenn die vom visuellen Cue indizierte Figur, also das Agens, zeitgleich auch das Topik der weiteren Rede war, wurden ca. 79.1% der ersten Sakkaden auf diese Figur gerichtet. Wenn stattdessen dem Patiens anstelle des Agens der Topikstatus zugeteilt wurde, wurde das Agens nur in ca. 36.8% aller Trials als erstes fixiert (vgl. Abbildung 8.2).

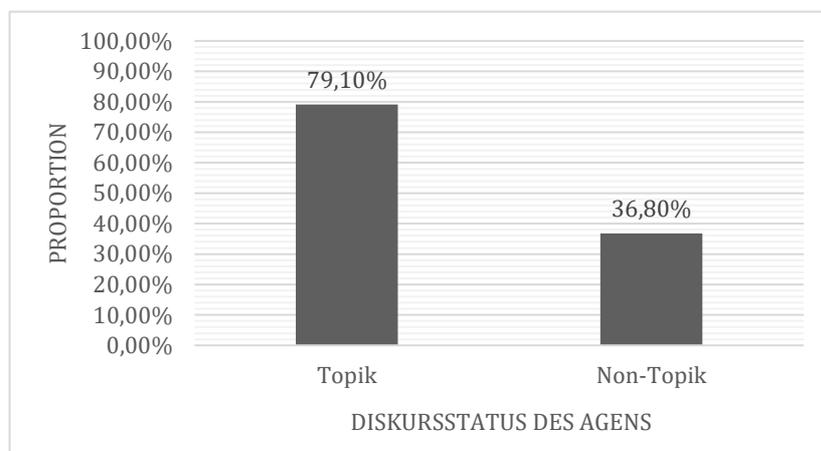


Abbildung 8.2 Proportion der durch den visuellen Cue getriebenen ersten Fixationen im Verhältnis zum Diskursstatus des Agens (Das Agens wurde immer durch den Cue markiert)

Effekt	Est.	Std. Error	z-value
Intercept	0.02975	0.13219	0.225
Fester Effekt	0.97133	0.10646	9.124***

Tabelle 8.2 Logistische Regressionsanalyse mit gemischten Effekten auf die Abhängigkeit der visuellen Salienz vom Diskursstatus des Agens

Kontextuelle Salienz bzw. Topikalität und Zielposition der zweiten Sakkade

Das Ergebnis bezüglich des Zusammenhangs zwischen der kontextuellen Salienz und Zielpositionen der zweiten Sakkade aus Experiment 2 wurde im vorliegenden Experiment bestätigt. Während 74.9% der zweiten Sakkaden auf das Agens gerichtet wurden, wenn das Agens das Topik der nachfolgenden Äußerung ist, sinkt dieses Verhältnis auf 19.9%, wenn hingegen nicht dem Agens, sondern dem Patiens der Topikstatus zugewiesen wurde (vgl. Abbildung 8.3). Die statistische Analyse hat bestätigt, dass ein signifikanter Effekt der Topikalität auf die Planung der zweiten Sakkade gegeben ist (vgl. Tabelle 8.3).

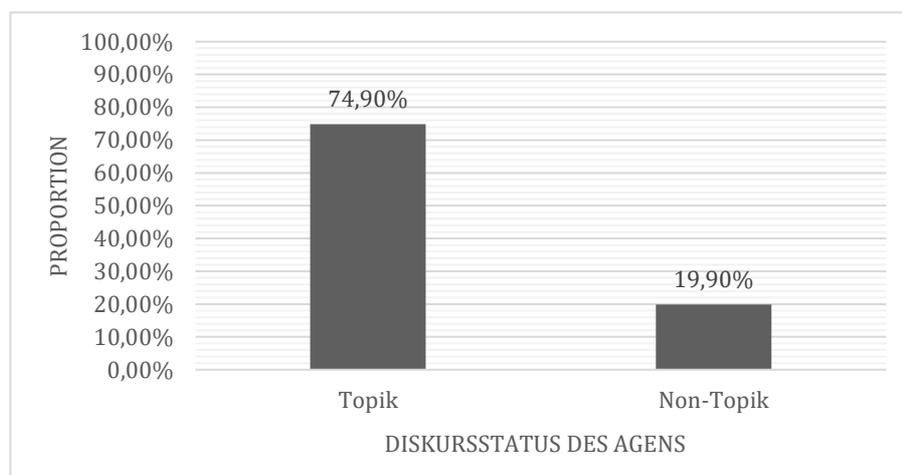


Abbildung 8.3 Proportion der zweiten Sakkaden, die auf das Agens gesteuert wurden, im Verhältnis zum Diskursstatus des Agens

	Est.	Std. Error	z-value
Intercept	0.1684	0.1666	1.011
Fester Effekt	1.3989	0.1040	13.450***

Tabelle 8.3 Logistische Regressionsanalyse mit gemischten Effekten auf die Abhängigkeit der Selektion der zweiten sakkadischen Zielposition vom Diskursstatus des Agens

Kontextuelle Salienz bzw. Topikalität und syntaktische Struktur

In diesem Experiment zeigt sich, ähnlich wie in Experiment 2, eine Abhängigkeit der selektierten Satzstrukturen vom Diskursstatus der Referenten (vgl. Tabelle 8.4). Wenn das Agens das Topik war, waren 99% der produzierten Sätze Aktivsätze. Wenn das Patiens das Topik der zu formulierenden Äußerung war, reduzierte sich der Wert auf 3% (vgl. Abbildung 8.4). Es wurden keine scrambled-Aktivsätze produziert, in denen das Patiens als das Objekt in der satzinitialen Position aufgetreten ist. Diese Ergebnisse lässt die Schlussfolgerung zu, dass die Selektion syntaktischer Strukturen von der kontextuellen Salienz der Ereignisteilnehmer manipuliert wurde. Ob das Agens als der erste Referent verbalisiert wird oder nicht, steht in einem engen Zusammenhang mit seinem Diskursstatus.

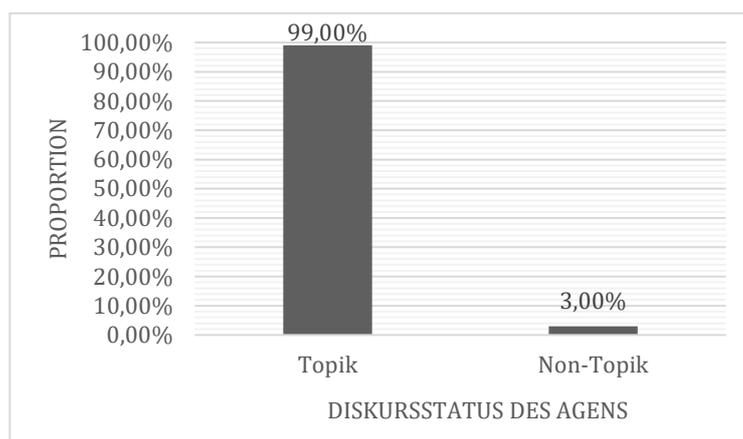


Abbildung 8.4 Proportion der gebildeten Aktivsätze im Verhältnis zum Diskursstatus des Agens

	Est.	Std. Error	z-value
Intercept	1.852	1.733	1.069
Fester Effekt	6.658	1.541	4.319***

Tabelle 8.4 Logistische Regressionsanalyse mit gemischten Effekten auf die Abhängigkeit der Selektion syntaktischer Strukturen vom Diskursstatus des Agens

8.4 Diskussion

Als ergänzende Studie, die auf den Ergebnissen von Experiment 2 beruht, hatte Experiment 3 zwei Ziele: Zum einen sollten die in Experiment 2 festgestellten Befunde bezüglich der Funktionalität der kontextuellen Salienz während des Sprechens repliziert werden. Zum anderen sollte mithilfe des Unterschiedes bezüglich der Speech-Onset-Latenz von Aktiv- und Passivsätzen eine mögliche Interaktion zwischen der Informationsstruktur und der semantischen Struktur auf der Mikroplanungsebene getestet werden.

Abbildung des Topiks auf die Semantik

Zur Untersuchung des Abbildungsprozesses innerhalb der Mikroplanung wurden die Speech-Onset-Latenzen der topik-initialen Aktiv- und Passivsätze miteinander verglichen. Die statistische Analyse zeigte, dass die Sprechanfängszeitpunkte der Passivsätze beträchtlich über denen der Aktivsätze lagen. Im Hinblick darauf, dass die ausgewählten Sätze unter vergleichbaren kontextuellen und visuellen Bedingungen produziert wurden, kann der identifizierte Unterschied bezüglich der Speech-Onset-Latenz lediglich auf die Zuweisung semantischer Rolle an den Referenten mit Topik-Status bzw. die thematischen Reihenfolgen der gebildeten Sätze zurückgeführt werden. Aufgrund der Tatsache, dass die Auswirkung der semantischen Reihenfolge den Aufbau der semantischen Struktur voraussetzt, kann aus dem gefundenen Unterschied nun geschlossen werden, dass die semantische Rolle des Topiks bereits vor dessen Weiterleitung in die syntaktische Verarbeitungsebene identifiziert wurde. Wenn Topiks direkt und unverzüglich ohne Einordnung in die semantische Vorstellung an die syntaktische Enkodierungsstufe weitergeleitet werden, sobald sie in der Preview-Phase von Sprechern zur Kenntnis genommen wurden, sollte kein Unterschied bezüglich der Speech-Onset-Latenz beobachtet werden. Denn in diesem Fall wissen Sprecher zu dem Zeitpunkt der Artikulation noch nicht, ob die satzinitialen Topiks das Agens oder das Patiens sind und ob die zu erstellenden Sätze in Form aktiver oder passiver Struktur realisiert werden. In diesem Zusammenhang

deuten unsere Befunde darauf hin, dass eine Interaktion zwischen der semantischen Struktur und der Informationsstruktur auf der Mikroplanungsebene stattfindet.

Von der oben aufgeführten Diskussion ausgehend gibt es noch zwei alternative Interpretationsmöglichkeiten, die nach Ansicht des Autors jedoch nicht als angemessene Erklärungen für den beobachteten temporalen Unterschied betrachtet werden können. Die erste Interpretation ist die, dass es eine sog. Default-Korrespondenz zwischen dem Topik und dem Agens im Gehirn des Sprechers gibt. Daher sei die beobachtete längere Speech-Onset-Latenz der Passivsätze, in denen der Topikstatus vom Patiens belegt ist, auf die Verletzung eben dieser Korrespondenz zurückzuführen.

Diese Erklärung geht von der Beobachtung aus, dass Sprecher im Allgemeinen vielmehr kanonische Aktivsätze als Passivsätze produzieren. Da in einem kanonischen Aktivsatz das Agens immer in der Satzanfangsposition auftritt, welche meistens vom Topik belegt wird, ist es Sprechern möglich, daraus eine implizite Default-Korrespondenz zwischen dem Topik und dem Agens zu entwickeln. Diese Korrespondenz kann wiederum dazu führen, dass das Topik des zu bildenden Satzes, nachdem es durch die in der Preview-Phase gestellte Frage eingeführt wurde, anhand der Default-Korrespondenz automatisch als Agens betrachtet wird. Zudem kann es dann als *starting point* auf der grammatischen Enkodierungsebene der Subjekt-Funktion zugeordnet werden. Und das wiederum kann eine Voraktivierung der aktiven Konstruktion hervorrufen. Wenn sich dieses Topik zu einem späteren Zeitpunkt, nach dem Aufbau der semantischen Struktur, als das Patiens des Ereignisses herausstellte, mussten Sprecher ihre vorherige Planung, die von der Default-Korrespondenz zwischen dem Topik und dem Agens ausging, unterbrechen und reparieren. Diese sog. Reparatur kann sich in der Speech-Onset-Latenz der Passivsätze manifestieren, das heißt, dass zusätzliche Planungszeit für den Passivsatzbau gebraucht wird.

Ein Schwachpunkt dieser Interpretation liegt darin, dass sie die kontextuelle Salienz

bzw. die Topikalität als den einzigen Faktor, der bei der Selektion von *starting points* wirksam ist, ansieht. Diese Annahme steht daher im Widerspruch zu unseren Befunden aus Experiment 1 und 2, in denen gezeigt wurde, dass die Entscheidung über *starting points* nicht beliebig getroffen wird, sondern im Wesentlichen ein Resultat des Zusammenspiels aller involvierten Faktoren ist.

Ein Beispiel für diese Aussage findet sich in unserem Befund aus Experiment 1 bezüglich der Einflüsse der visuellen und semantischen Salienz auf die Selektion. Es wurde festgestellt, dass, obwohl die visuell durch den Cue angezeigte Figur zuerst fixiert und dementsprechend mit größerer Wahrscheinlichkeit früher konzeptuell und syntaktisch enkodiert wurde, die Entscheidung über *starting points* nicht umgehend nach der visuellen Wahrnehmung, sondern erst nach dem Aufbau der semantischen Struktur getroffen wird. Wenn deutschsprachige Probanden der genannten alternativen Interpretation entsprechend *starting points* nur anhand eines Faktors, also der visuellen Salienz, auswählen, ohne die zu einem späteren Zeitpunkt festgelegte semantische Salienz jeweiliger Ereignisteilnehmer zu berücksichtigen, dann sollte in Experiment 1 kein Einfluss der semantischen Salienz beobachtet werden. Der beobachtete semantische Effekt bei der Selektion von *starting points* bestätigt deshalb die Annahme, dass sich Sprecher bei der Entscheidung von *starting points* wahrscheinlich nicht beeilen, sondern abwarten, bis ihnen alle mitwirkenden Faktoren, einschließlich der semantischen Salienz aller Beteiligten, vorliegen. Erst mithilfe einer umfassenden Auswertung aller betroffenen Faktoren können *starting points* endgültig ausgewählt werden.

Des Weiteren gibt es bisher – soweit dem Autor dieser Arbeit bekannt – keine Evidenz für die Existenz der von dieser Interpretation angenommenen Default-Korrespondenz zwischen dem Topik und dem Agens. Es wird gegenteilig angenommen, dass die Passivstruktur bzw. das Patiens in einem engeren Zusammenhang mit der Topikalität stehen kann (Halliday, 1967a; Crystal, 1997). Es wird zum Beispiel von Beedham (1982, S. 34) angenommen, dass die Passivierung im Wesentlichen ein Mittel der Topikalisierung ist (vgl. auch Halliday 1967b: 216; Quirk, Greenbaum, Leech &

Svartvik 1985: 943). Ähnlich nimmt Givon (1979) an, dass „the function of passive sentences in language is to code sentences in the context in which the non-agent is more topical“ (S. 57). Der Grund dafür, dass deutsche Sprecher die Passivstruktur trotz deren Markiertheit und der Komplexität in manchen Situationen wählen, kann daher in der informationsstrukturellen Spezifizierung der zu formulierenden Botschaft liegen. Mit der Passivierung wird das Patiens an die Satzanfangsposition gestellt und damit als Satztopik realisiert. Gleichwohl kann in diesem Zusammenhang eine Verbindung zwischen dem Patiens und dem Topik trotz geringer Anzahl von Passivsätzen sehr auffällig sein.

Basierend auf der oben aufgeführten Argumentation wird die Interpretation, dass der temporale Unterschied zwischen den Speech-Onset-Latenzen der topik-initialen Aktiv- und Passivsätze auf die Verletzung der sog. Default-Korrespondenz zwischen Topik und Agens zurückzuführen ist, als nicht überzeugend angesehen.

Die zweite alternative Interpretation zum identifizierten temporalen Unterschied ist die, dass Topiks bereits in der Preview-Phase als *starting points* ausgewählt und satzinitial angeordnet werden. Jedoch werden sie nach der Anordnung nicht sofort lautlich enkodiert und artikuliert, sondern erst verzögert, bis der Aufbau der semantischen Struktur abgeschlossen und das Topik in der semantischen Repräsentation lokalisiert ist, damit potenzielle Fehler möglichst ausgeschlossen werden können. Mit anderen Worten, diese Interpretation würde eine Art Buffer zwischen der grammatischen und phonologischen Enkodierungsebene implizieren, in dem das als *starting point* selektierte Topik nach dessen satzinitialer Positionierung temporär gespeichert wird. Die Verarbeitung wird erst fortgesetzt, nachdem die semantische Rolle dieses Topiks identifiziert wurde. Diese Vermutung trifft daher mit der Idee der *ideal delivery* zusammen, nach der Sprecher dazu tendieren, mit dem Sprechbeginn so lange zu warten, bis sie klar wissen, was sie sagen wollen (vgl. Clark & Clark, 1977; Ferreira & Swets, 2002). Diese Vermutung wird ebenfalls aus den folgenden zwei Gründen als problematisch betrachtet:

Erstens spricht diese Interpretation im Kern dafür, dass Sprecher nur mithilfe der kontextuellen Salienz und ohne nähere Betrachtung anderer mitwirkender Faktoren die *starting points* bereits in der Preview-Phase festlegen können. Diese Annahme ist jedoch nicht kompatibel mit unseren Ergebnissen, die eine Interaktion zwischen verschiedenen Faktoren während des Selektionsprozesses aufweisen. Obwohl die Verzögerung des Sprechbeginns den Sprechern nach der Reihung des ersten Referenten eine Chance bietet, bei einer irrtümlichen Selektion von *starting points* die Auswahl noch einmal vorzunehmen bzw. zu korrigieren, kann die Fehlerkorrektur aufwendig sein. Denn Sprecher müssen neben der neuen Auswahl den Abbildungsprozess von der konzeptuellen auf die grammatische Enkodierungsebene noch einmal durchführen. Eine ökonomischere Herangehensweise wäre, dass Sprecher auf der konzeptuellen Enkodierungsebene vorsichtig mit der Selektion von *starting points* sind und den Beginn der grammatischen Enkodierung verzögern, bis ihnen alle mitwirkenden Faktoren vorliegen und sie anhand derer die Entscheidung treffen. In diesem Sinne gleicht die von der vorliegenden Arbeit gehaltene Annahme ebenfalls dem Konzept der *ideal delivery*.

Zweitens kann diese alternative Interpretation nicht erklären, warum es keinen signifikanten Unterschied gibt zwischen den Speech-Onset-Latenzen der Aktivsätze aus Experiment 1, welche in einer pragmatisch neutralen Situation produziert wurden, und denen aus Experiment 2, in denen das Agens als Topik die satzinitiale Position besetzt (vgl. Abschnitt 7.4). In Experiment 1, in dem die Selektion von *starting points* ohne die Beteiligung des Topiks semantisch gestützt war, fand die positionelle Anordnung des ersten Referenten immer nach Stimulus-Onset statt. Das heißt, die Planungszeit für die positionelle Enkodierung des ersten Referenten ist in der Speech-Onset-Latenz mit einkalkuliert. Wenn diese Interpretation richtig wäre und die Reihung des Topiks bereits in der Preview-Phase, also vor der Stimulus-Präsentation vorgenommen wird, wäre beim Produzieren der topik-initialen Sätze im Vergleich zum ersten Experiment ein Beschleunigungseffekt zu erwarten. Denn in dieser Situation wird die für die positionelle Anordnung erforderliche Zeit nicht in der

Speechanfängszeit miteinberechnet. Dies ist – wie die aufgezeichneten Speech-Onset-Latenzen gezeigt haben – nicht der Fall. Es wurde kein signifikanter Unterschied bezüglich der Speech-Onset-Latenzen mit und ohne die Teilnahme des Topiks beobachtet, was darauf hindeutet, dass das Topik vor Stimulus-Onset wahrscheinlich nicht gereiht wurde.

Insgesamt deuten die oben erwähnten Ergebnisse und die Diskussion darauf hin, dass das Topik auf der Mikroplanungsebene auf die semantische strukturelle Repräsentation abgebildet wird und dieser kontextuell-semantische Interaktionsprozess eine Voraussetzung für die grammatische Enkodierung dieses Topiks bildet.

Topikalität und Satzplanung

Der Befund aus Experiment 2, dass die kontextuelle Salienz bzw. die Topikalität einen Einfluss, sowohl auf die visuelle Wahrnehmung als auch auf die Selektion von *starting points*, hat, wurde im vorliegenden Experiment repliziert. In Bezug auf den Zusammenhang zwischen der ersten Fixation und der Topikalität hat die durchgeführte Analyse einen Haupteffekt der Topikalität gezeigt: Die erste Sakkade wurde meistens dann auf die visuell markierte Figur, also das Agens, gerichtet, wenn es zeitgleich das Topik der darauffolgenden Rede war. Wenn das Patiens durch die eingeführte Target-Frage den Topikstatus erhielt, wurde die Anziehungskraft des durch den Cue angezeigten Agens für die erste Fixation erkennbar vermindert. Somit stützt dieses Ergebnis die Schlussfolgerung hinsichtlich des Effekts der kontextuellen Salienz auf die visuelle Wahrnehmung aus Experiment 2.

Dennoch beschränkt sich der modifizierende Effekt der Topikalität nicht nur auf die perzeptuelle Verarbeitung. Es wurden deutliche Evidenzen dafür gefunden, genau wie in Experiment 2, dass die Selektion von *starting points* von der kontextuellen Salienz der Ereignisteilnehmer abhängig ist. Bei den ausgewählten syntaktischen Strukturen manifestierte sich häufig eine Topik-Präferenz. Die Figuren, die Topik-Status besaßen, wurden unabhängig von ihren semantischen Funktionen stets als Subjekt realisiert

und ins Vorfeld eines Satzes gebracht. Des Weiteren wurden die zweiten Sakkaden, statt eine konsistente Präferenz für das Agens aufzuweisen, größtenteils vom Topik angezogen.

Zusammenfassend lässt sich feststellen, dass der gefundene Verarbeitungsvorteil bzw. die Rolle des Topiks bei der Satzplanung noch einmal mit den visuellen und linguistischen Daten aus Experiment 3 untermauert wurde.

Fazit

In diesem Experiment wurden zum einen erneut die Befunde bestätigt, die bereits in Experiment 2 herausgefunden wurden. Sowohl bei der visuellen Wahrnehmung als auch bei der Selektion von *starting points* kommt der kontextuellen Salienz eine dominante Rolle zu. Zum anderen wurde mithilfe des Unterschiedes bezüglich der Speech-Onset-Latenz der produzierten topik-initialen Aktiv- und Passivsätze der Abbildungsprozess des Topiks innerhalb des Mikroplanungssystems überprüft. Das Ergebnis war, dass, sobald Topiks als solche identifiziert wurden, diese nicht automatisch an die grammatische Enkodierung weitergereicht, sondern zuerst in eine semantisch strukturelle Repräsentation eingeordnet wurden. Eine Verbindung der kontextuell und semantisch strukturellen Repräsentation des Ereignisses auf der Mikroplanungsebene scheint für den Beginn der grammatischen Enkodierung notwendig zu sein. Diese Studie liefert daher auch Erkenntnisse über die Mechanismen der Mikroplanung während des Konzeptualisierungsvorgangs.

9 Allgemeine Diskussion

9.1 Zusammenfassung der Ergebnisse

Die vorliegende Arbeit befasst sich mit den Verarbeitungsprozessen der Satzproduktion im Deutschen. Folgende Fragestellungen sollten im Rahmen der Untersuchung beantwortet bzw. näher beleuchtet werden: Zum einen ist die Frage, ob die Satzproduktion im Deutschen, sowohl in pragmatisch neutralen Situationen als auch in Situationen, in denen die zu verbalisierenden Referenten einen unterschiedlichen Diskursstatus haben, linear oder strukturell inkrementell verläuft. Davon ausgehend soll weiterhin festgestellt werden, in welcher prozeduralen Einbindung sich die kontextuelle Salienz bzw. die Topikalität bei der Konzeptualisierung und der grammatischen Enkodierung auswirkt. Außerdem soll herausgefunden werden, ob während der Informationsverarbeitung eine Prioritätenliste der bei der Satzplanung mitwirkenden Faktoren – hier also der visuellen, semantischen und kontextuellen Salienz – besteht.

Im theoretischen Teil dieser Arbeit wurde zunächst das von vielen Forschern angenommene modulare System der Satzproduktion vorgestellt. Dieses besteht aus insgesamt vier verschiedenen Planungsstufen, nämlich der Konzeptualisierung, der grammatischen Enkodierung, der phonologisch-phonetischen Enkodierung sowie der Artikulation. Für die vorliegende Arbeit sind die Stufen der Konzeptualisierung und der grammatischen Enkodierung relevant. Auf der Konzeptualisierungsstufe wird vom Inhalt bis zur Art und Weise des Sprechens alles geplant und organisiert - so die Annahme. Der Output dieser ersten Stufe ist eine propositionale Struktur, welche dann wiederum als Input für die anschließende grammatische Enkodierung dient. Die grammatische Enkodierung wiederum verläuft in zwei Schritten: der funktionalen und der positionellen Verarbeitung. Während der funktionalen Verarbeitung werden passende Lemmata aus dem mentalen Lexikon ausgewählt und dann mit syntaktischen Funktionen belegt, und im darauffolgenden positionellen

Verarbeitungsschritt werden die Lemmata flektiert und in einer Wortfolge angeordnet.

Es ist zu beachten, dass, obwohl der Satzplanungsprozess schematisch durch aufeinander folgende Verarbeitungsebenen dargestellt wird, die Verarbeitung auf den einzelnen Ebenen jedoch nicht strikt nacheinander, sondern auch simultan ablaufen kann. Sobald ein Teil der präverbalen Botschaft (ein sog. Inkrement) auf einer Verarbeitungsebene erzeugt wurde, wird es unmittelbar an die darauffolgende Ebene weitergereicht. Da Sätze generell nicht in einem Arbeitsschritt als Ganzes, sondern stückweise geplant werden, müssen Sprecher vor Beginn der Satzplanung ein lexikalisches Konzept auswählen, welches die zu vollziehende Versprachlichung einleitet (das sog. *starting points problem*). Aber die Frage ist, wie dieses Konzept festgelegt wird. Um diese Frage zu beantworten, wurden in früheren Studien drei Hypothesen aufgestellt.

Laut der linearen Inkrementalitätshypothese sind inhärente Eigenschaften einzelner Ereignispartizipanten ausschlaggebend für die Selektion von *starting points*, einschließlich deren visueller oder konzeptueller Salienz. Beispielsweise haben einige frühere Studien gezeigt, dass diejenigen Figuren, die visuell salienter sind und daher früher wahrgenommen und identifiziert werden, mit höherer Wahrscheinlichkeit als *starting points* selektiert werden als diejenigen, die visuell weniger auffällig sind (z. B. Gleitman et al., 2007; Tomlin, 1995, 1997; Myachykov & Tomlin, 2008). Demnach begrenzt sich die Vorausplanung lediglich auf die Verarbeitung eines individuellen lexikalischen Konzepts. Im Gegensatz zur linearen Inkrementalitätshypothese wird von der strukturellen Inkrementalitätshypothese angenommen, dass der Selektion von *starting points* die semantisch strukturelle Repräsentation eines Gesamtereignisses zugrunde liegt. Diejenigen Figuren, deren semantische Rollen eine vordere Position in der thematischen Hierarchie belegen und daher eine höhere semantische Salienz innehaben (z. B. das Agens), werden wahrscheinlicher als *starting points* ausgewählt als diejenigen Figuren, die semantisch weniger salient sind (z. B. das Patiens). Dementsprechend kann der Umfang der Vorausplanung so groß wie die semantische bzw. konzeptuelle Struktur eines Gesamtereignisses sein. Eine dritte

Erklärungsmöglichkeit sieht beide Hypothesen gemeinsam als eine kontinuierliche Erklärung für die Arbeitsweise des Satzproduktionssystems. Die Auffassung ist, dass die Selektion von *starting points* eng mit der relativen Kodierbarkeit der semantischen Struktur eines Gesamtereignisses einerseits und den individuellen Ereignisteilnehmern andererseits korreliert. Zugänglichere Informationen sollen früher verarbeitet werden, damit sich Sprecher mehr Zeit, die sie für die Enkodierung komplexerer Informationen benötigen, verschaffen und somit mögliche Sprechverzögerungen vermeiden können (vgl. Kuchinsky & Bock, 2010; Konopka & Meyer, 2014).

In dieser Arbeit wurde die strukturelle Inkrementalitätshypothese experimentell sowohl für den Satzproduktionsprozess in pragmatisch neutralen Situationen (Experiment 1) als auch in Situationen, in denen die zu versprachlichenden Referenten einen unterschiedlichen Diskursstatus haben (Experiment 2 & 3), bestätigt.

Um die Mechanismen der Satzplanung in pragmatisch neutralen Situationen zu untersuchen, wurden die semantische und die visuelle Salienz, deren Effekte auf die Satzproduktion in früheren Studien bereits mehrfach bestätigt wurden, kontrastiert. Das wirft die Frage auf, welcher Faktor denn eine dominante Rolle während des Satzplanungsprozesses spielt. Zur Manipulation der semantischen Salienz wurde das Experiment im Kontext transitiver Ereignisse durchgeführt, in denen zwei belebte Ereignisteilnehmer, jeweils als das Agens und das Patiens, eine transitive Handlung vollzogen (z.B. *Eine Maus bespritzt eine Schildkröte*). Ein exogener visueller Cue wurde wiederum dazu verwendet, um die visuelle Salienz zu manipulieren. Dazu zeigte der Cue (ein kleines schwarzes Viereck) vor Stimulus-Onset entweder auf die Position des dann erscheinenden Agens- oder Patiens-Referenten, wodurch die frühe Aufmerksamkeit bzw. die erste Fixation auf den entsprechenden Referenten gelenkt werden konnte. Doch an dieser Stelle muss man sich fragen, welche Befunde uns nun die Schlussfolgerung erlauben, die Satzproduktion im Deutschen verlaufe strukturell inkrementell.

In einem ersten Schritt wurde der Manipulationseffekt des exogenen Cues auf die frühe visuelle Aufmerksamkeit überprüft. Es zeigte sich, dass im Vergleich zur Zufallswahrscheinlichkeit von 50% die erste Fixation durch den Cue deutlich öfter auf die dadurch angedeutete Entität fiel. Dennoch wurden die Referenten, die zuerst fixiert wurden, nicht als *starting points* ausgewählt und in die grammatische Enkodierungsebene überführt - außer wenn sie sich später als das Agens herausgestellt haben. Die Anzahl der gebildeten Passivsätze hat durch erste Fixationen auf dem Patiens nicht zugenommen. Dieses Ergebnis widerspricht der Hypothese der linearen Inkrementalität, nach der die erste Fixation den satzeinleitenden Referenten bestimmen kann, ohne dessen semantische Rolle berücksichtigen zu müssen.

In einem zweiten Schritt liefern Blickbewegungsdaten einen zweiten Beweis für die strukturelle Inkrementalitätshypothese und zeigen, dass die semantische Struktur bereits vor der Auslösung der zweiten Sakkade (mindestens zum Teil) aufgebaut ist. Dies zeigt sich vor allem in der zweiten Sakkade, welche in der Tat die erste eigenständige und willentlich ausgeführte Sakkade ist. Im Gegensatz zur ersten Sakkade bzw. Fixation zeigte die zweite Sakkade durchgehend eine Präferenz für das Agens. Da die Präferenz für eine bestimmte semantische Rolle den Aufbau der semantischen Struktur voraussetzt, lässt sich aus den Daten der zweiten Sakkade der Schluss ziehen, dass die semantischen Rollen (hier vor allem das Agens) bereits vor der zweiten sakkadischen Bewegung zugewiesen wurden. Die hohe Korrespondenz zwischen dem Agens und dem satzeinleitenden Referenten und die Präferenz der zweiten Sakkade für das Agens liefern somit gemeinsam eine Evidenz dafür, dass die Satzproduktion nach der visuellen Wahrnehmung strukturell inkrementell verläuft.

Unsere Befunde widersprechen grundsätzlich der Idee, dass die inkrementelle Verarbeitung eine hohe Variabilität aufweist. In Experiment 1 sind die Referenten, die das zu beschreibende transitive Ereignis konstituierten, sowohl visuell als auch lexikalisch bereits in der Preview-Phase eingeführt. Deswegen war die Enkodierung der individuellen Referenten während der Beschreibung im Gegensatz zum Aufbau der semantischen Struktur, die vorab nicht bekannt ist und erst nach der Präsentation

des Ereignisses benannt werden kann, weniger aufwändig. Gemäß dem *minimal-load principle* von Levelt (1989) und der Variabilitätsannahme von Kuchinsky und Bock (2010; vgl. auch Konopka & Meyer, 2014) sollte die Satzplanung in dieser Situation mit der Enkodierung individueller Ereignispartizipanten beginnen, was eine linear inkrementelle Verarbeitung auslöst. Dies ist jedoch nicht der Fall. Stattdessen hat sich erwiesen, dass die deutschsprachigen Probanden ihre Satzplanung immer mit der Konstruktion der semantischen Struktur begonnen haben, ohne dabei die konzeptuelle und lexikalische Kodierbarkeit einzelner Ereignisteilnehmer zu berücksichtigen.

Zusammenfassend lässt sich also sagen, dass die Satzproduktion im Deutschen in pragmatisch neutralen Situationen strukturell inkrementell verläuft und der semantischen Struktur dabei eine zentrale Rolle zukommt.

Nach Levelt (1989) und weiteren Linguisten (z. B. v. Stutterheim & Klein, 2002) ist die semantische Struktur nicht die einzige vorsprachliche Bedeutungsstruktur, die in der Mikroplanungsphase der Konzeptualisierung generiert wird. Darüber hinaus müssen Sprecher noch eine Perspektive, aus der Informationen sprachlich dargestellt werden, auswählen. Die Auswahl einer Perspektive führt zur Erstellung einer spezifischen Informationsstruktur. In Experiment 2 wurde die Theorie der strukturellen Inkrementalität, die in älterer Literatur nur die semantische Struktur betraf, auf die pragmatische Verarbeitungsebene erweitert und das Verhältnis zwischen der Satzproduktion und der kontextuellen Salienz untersucht.

Der Ablauf des Trials und die Manipulation der visuellen und semantischen Salienz verliefen im Allgemeinen wie in Experiment 1 – mit dem einzigen Unterschied, dass der visuelle Cue in Experiment 2 nur auf der Patiens-Region erschien. Durch eine in der Preview-Phase gestellte Frage, die eine *mit*-Phrase enthielt (z. B. *Sag mir bitte, was mit der Schildkröte passiert?*), wurde der kontextuelle Status der Ereignisteilnehmer festgelegt. Diejenigen Figuren, die durch die *mit*-Phrase näher bestimmt wurden, bekamen in den Antworten der Probanden den Topik-Status. Auf diese Art und Weise wurde eine bestimmte Perspektive der Informationsorganisation

vermittelt.

Da bei der syntaktischen Strukturierung anstelle der Agens- eine starke Topikpräferenz festgestellt wurde, konnte gezeigt werden, dass die kontextuelle Salienz einen Einfluss auf die Satzplanung hat. Den Figuren mit Topikstatus wurde ohne Rücksicht auf deren thematische Rollen immer die Subjekt-Funktion zugeordnet bzw. die Figuren wurden satzinitial angeordnet. Die hohe Korrespondenz zwischen dem Topik und dem zuerst genannten Referenten liefert einen Hinweis darauf, dass der Effekt der kontextuellen Salienz eine zentrale Rolle bei der syntaktischen Enkodierung spielt.

Ein weiteres Indiz für den Einfluss der kontextuellen Salienz bzw. der Topikalität ist die Zielposition der zweiten Sakkade. Im Gegensatz zu Experiment 1, in dem die zweite Sakkade, also auch die erste top-down getriebene Sakkade, zum großen Teil auf das Agens gerichtet wurde, wurde sie in Experiment 2 stets von dem Topik angezogen.

Ein weiterer in diesem Zusammenhang bedeutender Befund war, dass die kontextuelle Salienz auch einen Einfluss auf die visuelle Wahrnehmung hat. Dies spiegelt sich vor allem in den Blickbewegungsdaten der ersten Fixation wider. Es wurde nämlich herausgefunden, dass der visuelle Salienzeffekt stark beeinträchtigt wurde, wenn eine andere Figur als die durch den visuellen Cue markierte das Topik des zu bildenden Satzes war. Insgesamt wurde die erste Sakkade deutlich häufiger auf den Referenten mit Topikstatus als auf das Non-Topik gerichtet. Die Abhängigkeit des ersten Fixationsorts vom kontextuellen Status der Ereignisteilnehmer erlaubt eine Schlussfolgerung hinsichtlich des Zeitverlaufs der Produktion insofern, als die kontextuelle Salienz bereits in einer sehr frühen Phase der Satzplanung, also während der visuellen Wahrnehmung, hinzugetreten ist.

Die oben erwähnten Resultate hinsichtlich des Einflusses der kontextuellen Salienz auf die Satzplanung wurden in Experiment 3 mit einem ähnlichen Versuchsdesign repliziert. Das Ergebnis war, dass die Satzplanung sowohl auf der visuellen und

konzeptuellen als auch auf der linguistischen Verarbeitungsebene durch den kontextuellen Status des zu versprachlichenden Referenten modifiziert wurde.

Hinsichtlich des enormen Einflusses der kontextuellen Salienz bzw. der Topikalität sind nun die Fragen offen, inwiefern das Topik in dem Satzproduktionssystem repräsentiert ist und in welcher Art und Weise dessen Auswirkung abläuft.

Da das Topik sowohl einen Einfluss auf die Konzeptualisierung als auch auf die syntaktische Strukturierung hat, wurde dementsprechend dessen Repräsentation auf der konzeptuellen und der linguistischen Ebene diskutiert.

In Bezug auf die linguistische Repräsentation des Topiks wird angenommen, dass das Topik als kontextuelle „Etikette“ fungiert und syntaktischen Funktionen wie etwa der Subjekt-Funktion positionelle Merkmale verleiht. Diese Annahme basiert auf dem Ergebnis, dass das Subjekt der gebildeten Sätze, das im Deutschen sowohl satzinitial als auch im Mittelfeld auftreten kann, aufgrund des topikalischen Status an den Satzanfang positioniert wurde (vgl. auch Arnold et al., 2013).

In Bezug auf die konzeptuelle Repräsentation des Topiks nehmen wir an, dass Referenten mit Topik-Status einen hohen Aufmerksamkeitsstatus aufweisen. Dies spiegelt sich vor allem im Verhältnis zwischen der overten Aufmerksamkeit, also den gemessenen Blickbewegungen, und dem Diskursstatus der Referenten wider. Die erste Sakkade, die in pragmatisch neutralen Situationen grundsätzlich bottom-up gesteuert ist, kann aufgrund dieser Eigenschaft als Indikator für den visuellen Wahrnehmungsprozess angesehen werden. Im Gegensatz dazu reflektiert die zweite Sakkade den top-down-Verarbeitungsprozess während der Selektion von *starting points*. Beide Sakkaden wurden im Experiment vom kontextuellen Status der Ereignisteilnehmer beeinflusst.

Zur Erklärung der Effekte der Topikalität innerhalb der Konzeptualisierung wurde der kognitionspsychologische Begriff der *selektiven Aufmerksamkeit* benutzt. Durch diese Aufmerksamkeit selektierten Menschen Reize und konzentrieren sich nur auf die für

sie relevante Entität oder Position und blenden andere irrelevante oder unwichtige Reize in ihrem Umfeld aus. Sprecher müssen so schnell wie möglich, bevor sie anfangen zu sprechen, den relevantesten oder auffälligsten Reiz, also den *starting point*, aufmerksam auswählen und beachten. Daher lässt sich mithilfe der Selektionsmechanismen der Aufmerksamkeit, die Art und Weise, in der die Topikalität die Selektion von *starting points* beeinflusst, erklären.

An dieser Stelle soll nochmals die sog. Spotlight-Metapher von Posner (1980) aufgegriffen werden, um die Mechanismen der selektiven Aufmerksamkeit zu erläutern. Dieser Metapher nach bewegt sich die menschliche Aufmerksamkeit wie ein mentaler „Scheinwerferstrahl“ durch das Sichtfeld. Wenn sich ein bestimmter Teil des Sichtfeldes durch seinen hohen Aufmerksamkeitsstatus abhebt, wird dieser vom attentionalen Scheinwerfer beleuchtet. Das wiederum kann dann zu einer schnelleren und gründlicheren Reizverarbeitung führen. Referenten mit Topikstatus werden früher und intensiver vom Scheinwerfer beleuchtet und verarbeitet, da sie im Vergleich zu Non-Topiks einen hohen Aufmerksamkeitsstatus besitzen. Andere Figuren, die sich außerhalb des Spotlights der konzeptuellen Verarbeitung befinden, werden dann ausgeblendet und möglicherweise erst zu einem späteren Zeitpunkt beleuchtet und enkodiert.

Eine weitere Frage, die beantwortet werden sollte, betrifft den Abbildungsprozess der Topik-Information von der konzeptuellen auf die positionelle Verarbeitungsebene, auf der die Wortstellung generiert wird. Diese Frage ist im Hinblick auf das Verständnis des Verarbeitungszeitverlaufs des Topiks von großer Bedeutung. In dieser Studie wurde dieser Prozess in zwei Unterprozessen kategorisiert und analysiert. Der erste Unterprozess betrifft die Abbildung innerhalb des grammatischen Enkodierungssystems. Dabei steht die Auseinandersetzung zwischen dem *grammatical function account* und dem *word order account* im Mittelpunkt. Der zweite Unterprozess betrifft die Abbildung innerhalb des Mikroplanungssystems. Es sollte herausgefunden werden, ob Topik-Information zuerst in die semantisch strukturelle Repräsentation eingeordnet wird, bevor sie an die grammatische Enkodierungsebene

weitergereicht wird.

In Bezug auf den Abbildungsprozess innerhalb der grammatischen Enkodierung lässt sich in Verbindung mit den Ergebnissen von Cowles und Ferreira (2012) als relativ sicher annehmen, dass das Topik wahrscheinlich zuerst auf die funktionale Enkodierungsebene abgebildet und dort mit syntaktischen Funktionen belegt wird, bevor es dann zum Schluss in die positionelle Enkodierungsebene gelangt. Diese Schlussfolgerung basiert auf den folgenden zwei Beobachtungen. Im Gegensatz zum Russischen und Griechischen, die die Voranstellung der visuell oder konzeptuell salienteren Referenten überwiegend durch Scrambling vornehmen (Myachykov & Tomlin, 2008; Branigan & Feleki, 1999), erfolgt das im Deutschen immer durch die Zuweisung der Subjekt-Funktion an dem Topik. Außerdem haben Cowles und Ferreira (2012) herausgefunden, dass der topikalische Effekt nur in einer Art Produktion erschien, bei der eine syntaktische Funktionszuweisung beansprucht wurde (z. B. der Satzproduktion). Wenn die Versuchspersonen koordinierte NPs produzierten und es sich nur um einen Linearisierungsprozess handelte, waren keine Topikalitäts-Effekte zu sehen. Wagner (2016) zufolge lässt sich daraus ableiten, dass die kontextuelle Salienz nur dann wirkt, wenn Referenten während der Sprachproduktion syntaktisch funktional verarbeitet werden. Diese Schlussfolgerung passt in ihren Grundzügen mit der Annahme des *grammatical function account* zusammen.

In Bezug auf den Mappingprozess innerhalb der Mikroplanung wurde herausgefunden, dass es eine Verbindung zwischen der kontextuell und der semantisch strukturellen Repräsentation gibt. Der signifikante Zeitabstand hinsichtlich der Speech-Onset-Latenz beim Produzieren der topik-initialen Aktivsätze aus Experiment 3 und der topik-initialen Passivsätze aus Experiment 2 dient als Evidenz für diesen Befund. Die zugrundeliegende Hypothese lautet wie folgt: Es sollte kein Unterschied zwischen der Speech-Onset-Latenz der verglichenen topik-initialen Aktiv- und Passivsätze erkennbar sein, wenn das Topik unmittelbar an die grammatische Enkodierung weitergereicht und dann versprachlicht wird, sobald es in der

Preview-Phase bekannt ist und bevor die semantische Struktur aufgebaut wurde. Denn in dieser Situation können Sprecher zum Zeitpunkt der Artikulation noch nicht wissen, ob sie einen Aktiv- oder Passivsatz bilden. Diese Hypothese wurde nicht bestätigt. Denn die Tatsache, dass die Aktivsätze im Gegensatz zu den Passivsätzen schneller gebildet wurden, was auf eine differente semantische Reihenfolge der beiden Satzstrukturen zurückzuführen ist, liefert den Hinweis darauf, dass das Topik bereits vor der grammatischen Enkodierung in die semantische bzw. konzeptuelle Vorstellung des Target-Ereignisses eingeordnet wurde.

Basierend auf der oben aufgeführten Argumentation soll nachfolgend ein weiterentwickeltes Satzproduktionsmodell aufgezeigt werden, das zwar auf dem Modell von Bock & Levelt (1994) beruht, jedoch den Schwerpunkt auf den Verarbeitungsablauf der kontextuellen Information legt, welcher von Bock und Levelt eher unbeachtet blieb. Es handelt sich hierbei um den Abbildungsprozess des Topiks während der Satzplanung. Dieser Prozess ist in diesem Modell in orange gekennzeichnet (vgl. Abbildung 9.1).

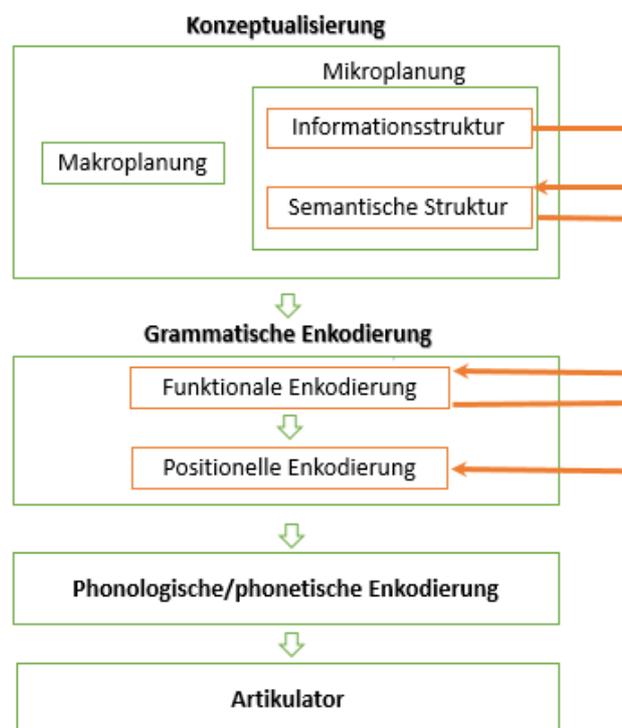


Abbildung 9.1 Der Abbildungsprozess des Topiks innerhalb der Satzplanung

Das letzte Ziel der vorliegenden Arbeit war, die Wirksamkeit der Einflussfaktoren auszuwerten und daraus anhand deren Bedeutsamkeit für die Satzplanung eine sog. Prioritätenliste zu entwickeln. So eine Liste kann als Leitfaden für die Selektion von *starting points* dienen, insbesondere dann, wenn mehrere Faktoren mit unterschiedlichen Referenten korrelieren.

Zusammengefasst kann man sagen, dass die semantische und die kontextuelle Salienz, die die strukturelle Repräsentation eines Ereignisses betreffen, höher gewertet werden als der Faktor, der die visuelle Eigenschaft individueller Referenten beschreibt. Dies äußert sich vor allem in den zweiten sakkadischen Zielpositionen und satzeinleitenden Wörtern, welche als zuverlässige Indikatoren für die ausgewählten *starting points* der Versprachlichung angesehen werden können.

Zuerst wurde durch Experiment 1 gezeigt, dass die oben erwähnten Indikatoren während der Beschreibung in pragmatisch neutralen Situationen eine deutliche Agens-Präferenz aufweisen. Die visuelle Salienz, die bei der Planung der ersten Fixation eine tragende Rolle spielt, wurde im Selektionsprozess von *starting points* kaum berücksichtigt. Aus diesem Ergebnis lässt sich ableiten, dass die semantische Salienz in der Prioritätenliste einen höheren Rang bzw. Status innehat.

In Experiment 2 und 3 wurde die kontextuelle Salienz dann als zusätzlicher Faktor eingeführt. Eine Auswertung der Kongruenz des zweiten fixierten Referenten mit dem Topik zeigte, dass die meisten zweiten Sakkaden auf den topikalischen Referenten gerichtet wurden, was eine Topik-Präferenz bedeutet. Die in Experiment 1 beobachtete Agens-Präferenz, die den Manipulationseffekt der semantischen Salienz widerspiegelt, wurde dadurch abgelöst. Die aufgezeichneten sprachlichen Daten bestätigen dieses Ergebnis. Anstatt derjenigen Referenten, die als Agens im Ereignis in Erscheinung traten, wurden diejenigen Referenten, den der Topik-Status zugeordnet war, immer zuerst enkodiert und verbalisiert. Aus diesen Ergebnissen kann man also zu dem Schluss kommen, dass die Selektion von *starting points* überwiegend von der kontextuellen Salienz beeinflusst wird.

9.2 Versuch einer eigenen Konzeption

Im nachfolgenden Abschnitt soll nun eine eigene Modellbeschreibung über die Selektionsmechanismen von *starting points* präsentiert werden (vgl. 9.2). Als Hilfe soll die Spotlight-Metapher von Posner (1980) als Erklärung dienen. Dieser Versuch soll als kreativer Anstoß oder als theoretische Exploration verstanden werden.

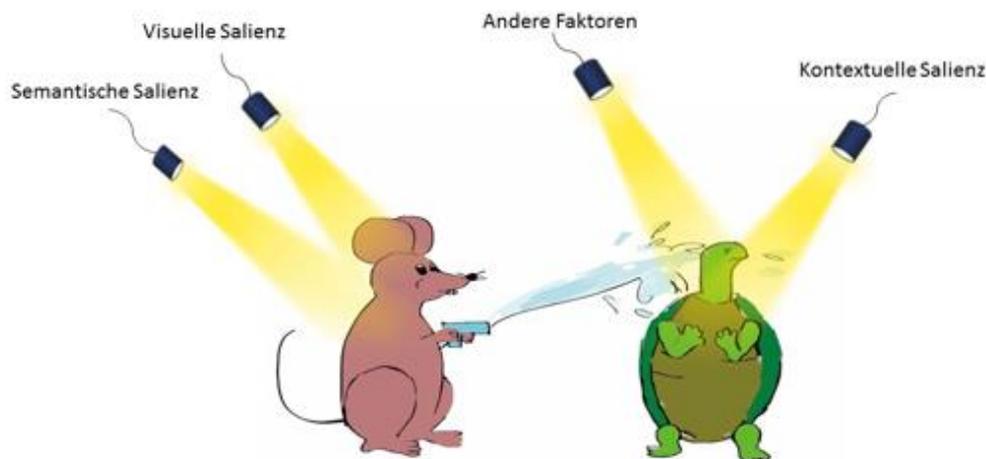


Abbildung 9.2 Spotlight-Modell zur Beschreibung der Selektionsweise von *starting points*

Dieses Modell geht im Wesentlichen von der von Posner (1980) aufgestellten *Spotlight*-Metapher aus, deren Grundannahme das Konzept der selektiven Aufmerksamkeit ist, die auch bei der Selektion von *starting points* während der Satzplanung eingesetzt wird. Die selektive Aufmerksamkeit kann mit einem attentionalen Scheinwerferstrahl verglichen werden, der sich konstant auf verschiedene Teile eines Reizangebots richtet und dabei den Zielreiz aus den gesamten Reizen selektiert. Diejenige Figur, die vom Strahl beleuchtet und daher in den Fokus der Verarbeitung gerückt wird, wird letztendlich als Zielreiz ausgewählt. Der Faktor, der diesen attentionalen „Lichtstrahl“ erzeugt, dient sozusagen als Scheinwerfer.

In der Studie von Posner (1980), die sich auf die räumliche visuelle Wahrnehmung konzentriert, war die visuelle Salienz der einzige Faktor, der die selektive

Aufmerksamkeit steuert. Deswegen wurde in seiner Studie nur ein einziger attentionaler Scheinwerfer angenommen, der für die visuelle Salienz steht. Der Selektionsprozess von *starting points* während der Satzplanung läuft jedoch viel komplizierter ab. Wie bereits in den drei für diese Studie durchgeführten Experimenten beschrieben, wurde die selektive Aufmerksamkeit in diesem Prozess nicht nur von der visuellen, sondern auch von der semantischen und kontextuellen Salienz beeinflusst. Daher nimmt diese Arbeit aufgrund dieses Zusammenhangs drei attentionale Scheinwerfer an, die den drei Salienzfaktoren entsprechen sollen und während des Satzproduktionsprozesses verschiedene Teile eines Ereignisses beleuchten können.

Eine wichtige Annahme dieses Modells ist, dass die sog. Stärke der „Strahlung“ oder Wirkung von den jeweiligen mentalen „Scheinwerfern“ unterschiedlich sind. Die Menge dieser Scheinwerfer spiegelt die durch die entsprechenden Faktoren ausgelöste Aufmerksamkeit wider. Diese Annahme fußt auf der in dieser Studie gefundenen ungleichen Wirksamkeit der Einflussfaktoren auf die Satzplanung (vgl. die Prioritätenliste der Einflussfaktoren). Die visuelle Salienz wirkte sich in dieser Studie beispielsweise nicht auf die Auswahl von *starting points* aus. Ein Einfluss der semantischen Salienz hingegen konnte nur in einer pragmatisch neutralen Situation beobachtet werden. Wenn Ereignisteilnehmern verschiedene kontextuelle Status zugewiesen waren, dann wurden meistens diejenigen als *starting points* der Versprachlichung ausgewählt, die eine höhere kontextuelle Salienz innehatten.

Des Weiteren nimmt das Modell an, dass die Strahlstärke eines Scheinwerfers, also die Wirksamkeit eines Faktors, auch je nach Satzproduktionssituation unterschiedlich sein kann. Diese sog. Flexibilität der Strahlstärke äußert sich zum einen in dem Verhältnis zwischen der Manipulationsweise eines Faktors und der daraus resultierenden Menge der Aufmerksamkeit. Wieviel Aufmerksamkeit von einem Faktor hervorgerufen und dann auf eine bestimmte Figur gerichtet wird, ist in gewisser Weise davon abhängig, wie diese Figur im Satzplanungsprozess deutlich gemacht wird. Damit lassen sich dann auch die unterschiedlichen Ergebnisse in den

Studien von Myachykov et al. (2012) und Gleitman et al. (2007) hinsichtlich des Manipulationseffekts des visuellen Cues auf die frühe visuelle Aufmerksamkeit im Rahmen dieses Modells erklären. In beiden Studien wurde zur Manipulation der visuellen Salienz des Ereignisteilnehmers das *visual-cueing*-Paradigma eingesetzt, in dem die Position der zu manipulierenden Referenten durch einen vorab präsentierten visuellen Cue angezeigt wurde, jedoch in unterschiedlichen Zeitlängen. Während der visuelle Cue in der Studie von Myachykov et al. (2012) in einer endogenen Weise für 700 ms präsentiert wurde, dauerte die Präsentation des visuellen Cues in der Studie von Gleitman et al. (2007) in einer exogenen Weise nur ca. 70 ms. Dieser Unterschied hinsichtlich der Darstellungsdauer führt zu verschiedenen Effektstärken des visuellen Cues: In Myachykov et al. Studie wurden 96% der ersten Fixationen auf die durch den Cue angezeigte Figur gerichtet, wohingegen das Verhältnis in der Studie von Gleitman et al. mit ca. 72% deutlich darunterliegt. Diese Diskrepanz deutet daher darauf hin, dass attentionale Scheinwerfer je nach experimentellen Designs unterschiedlich viel Strahlung aussenden.

Außerdem kann die Strahlstärke eines Scheinwerfers, also die Wirksamkeit eines Faktors, je nach Zielsprache des Sprechens verschieden sein. Diese Annahme basiert auf der Beobachtung, dass ein Faktor in unterschiedlichen Sprachen in unterschiedlichem Grad hervorgehoben wird und daher eine unterschiedliche Menge von Aufmerksamkeit hervorruft. So lässt sich auch zum Beispiel die Diskrepanz der Resultate hinsichtlich des Einflusses der visuellen Salienz auf die Selektion von *starting points* im Englischen (vgl. Gleitman et al, 2007) und Deutschen (vgl. Experiment 1) im Rahmen dieses Modells erklären. Obwohl die visuelle Salienz eines Referenten in diesen beiden Studien auf die gleiche Art und Weise manipuliert wurde, war ihre Wirkung auf die Selektion von *starting points* unterschiedlich. Die englischen Sprecher tendierten in pragmatisch neutralen Situationen dazu, die visuell auffälligeren Figuren als *starting points* der Versprachlichung auszuwählen. Eine solche Tendenz war im deutschen Satzproduktionsprozess nicht zu erkennen. Stattdessen spielte im Deutschen die semantische Struktur in diesem Zusammenhang

eine dominante Rolle. Der Unterschied bezüglich der Wirksamkeit eines Faktors in unterschiedlichen Sprachen steht wahrscheinlich in einem engen Zusammenhang mit der grammatischen Beschränkung in der jeweiligen Zielsprache (vgl. Abschnitt 6.4).

Des Weiteren impliziert der Zeitverlauf des Wirkens der jeweiligen Faktoren ebenfalls eine Flexibilität. Bei der Selektion von *starting points* werden diejenigen Scheinwerfer, die für die Einflussfaktoren stehen, nicht gleichzeitig “eingeschaltet”. Stattdessen steht die Reihenfolge der Beleuchtung in einem engen Zusammenhang mit (1) den kognitiven und psychophysiologischen Verarbeitungsmechanismen von Menschen und (2) der Reihenfolge des Auftretens der Einflussfaktoren.

Unter den kognitiven und psychophysiologischen Verarbeitungsmechanismen wird vor allem das verstanden, was die Enkodierung eines bildhaft dargebotenen Ereignisses betrifft (vgl. Abschnitt 4.2). Das Verständnis eines Ereignisses folgt immer einem Bottom-Up-Verarbeitungsprozess, in dem Menschen das Ereignis immer zuerst wahrnehmen müssen, bevor sie danach mit der Extraktion der Ereignisstruktur und der Identifizierung von Referenten beginnen können. Eine *high-level* konzeptuelle Enkodierung ohne eine davor vorgenommene *low-level* visuelle Wahrnehmung ist nicht möglich. Diesen Verarbeitungsmechanismen entsprechend tritt bei der Auswahl von *starting points* die visuelle Salienz immer vor der semantischen Salienz in die Rolle. Der attentionale Scheinwerfer, der von der semantischen Salienz gesteuert wird, wird sozusagen erst “eingeschaltet”, nachdem Sprecher das Ziel-Ereignis perzeptuell einigermaßen verarbeitet haben.

Darüber hinaus hat die Reihenfolge des Auftretens der Einflussfaktoren auch einen Einfluss auf den Zeitverlauf des Wirkens von Einflussfaktoren. Diejenigen Faktoren, die früher ins Spiel treten, wirken sich dementsprechend früher aus als diejenigen, die erst zu einem späteren Zeitpunkt eingeführt werden. In Experiment 2 und 3 wurde der Diskursstatus der Ereignisteilnehmer vor der Präsentation des Ziel-Ereignisses eingeführt und deshalb bereits auf der visuellen Wahrnehmungsstufe ein Topikalitäts-Effekt beobachtet. Wenn diese kontextuelle Information nach dem

Aufbau der semantischen Struktur vermittelt würde, sollte der visuelle Wahrnehmungsprozess nicht von ihr beeinflusst werden.

Hinsichtlich des Selektionskriteriums des Zielreizes besagt das Spotlight-Modell von Posner (1980), dass nur derjenige Referent oder Reiz, der im Mittelpunkt des Scheinwerferkegels steht, als Zielreiz ausgewählt wird. Diejenigen Figuren, die sich außerhalb des Lichtkegels befinden, werden dann wenig oder sogar kaum verarbeitet. Dieses Kriterium gilt allerdings nur für eine Situation, in der nur ein Scheinwerfer vorhanden ist und dementsprechend nur ein Referent oder ein Teilgebiet des Sichtfelds beleuchtet wird. Im Selektionsprozess von *starting points* können mehrere attentionale Scheinwerfer zeitgleich ihr Licht auf die verschiedenen Referenten eines Ereignisses werfen, was die Frage aufwirft, wie sich Sprecher in einer solchen Situation bezüglich *starting points* entscheiden.

Obwohl verschiedene Referenten zeitgleich von verschiedenen attentionalen Scheinwerfern beleuchtet werden, können sich deren Beleuchtungsstärken aufgrund der unterschiedlichen Einflussstärken der jeweiligen Faktoren voneinander unterscheiden. Hieraus ergibt sich, dass der gesamte Beleuchtungsgrad eines Referenten für die Selektion von *starting points* entscheidend ist, da er die gesamte Menge der Aufmerksamkeit, die auf die Verarbeitung dieses Referenten gerichtet ist, widerspiegelt. So werden die Referenten, die am „stärksten“ beleuchtet werden, als *starting points* der Satzplanung ausgewählt und diejenigen, die weniger beleuchtet werden, erst zu einem späteren Zeitpunkt verarbeitet. Das bedeutet aber nicht, dass diejenigen Figuren, die von den meisten Scheinwerfern beleuchtet werden, unbedingt als *starting points* selektiert werden müssen. Ob eine Figur selektiert wird oder nicht, hat nichts mit der Anzahl der Scheinwerfer, die sie beleuchten, sondern vor allem mit der Gesamtmenge der Aufmerksamkeit, die von den jeweiligen Scheinwerfern (den jeweiligen Faktoren) ausgestrahlt wird, zu tun. Beispielsweise wurde in einer Bedingung in Experiment 3 die Figur *Maus* bei der Beschreibung des Ereignisses *Eine Maus bespritzt eine Schildkröte* vom visuellen exogenen Cue markiert, wohingegen der Topikstatus der Figur *Schildkröte* zugeordnet wurde. Des Weiteren

war die Figur *Maus* auch das Agens dieses Ereignisses. Obwohl *Maus* von zwei Scheinwerfern beleuchtet wurde, wählten die Sprecher *Schildkröte* mit höherer Wahrscheinlichkeit als *starting points* ihrer Beschreibung aus. Dies kann darauf zurückzuführen sein, dass die Wirksamkeit der kontextuellen Salienz die Gesamtmenge der Wirksamkeit der semantischen und der visuellen Salienz übertrifft.

9.3 Ausblicke

In Experiment 2 und 3 zeigte sich ein starker Effekt des Einflusses des Diskursstatus auf die konzeptuelle und die syntaktische Enkodierung. Jedoch betrifft dieser Einfluss in der Tat nur das Topik. Der Fokus-Einfluss, der während der syntaktischen Strukturierung ebenfalls eine Rolle spielt, wurde nicht überprüft. Der Grund dafür ist, dass die in dieser Studie untersuchte Satzproduktion im Kontext transitiver Ereignisse stattfand. Mit der Festlegung des Topikstatus eines Referenten wurde der übrige Referent automatisch als Fokus identifiziert. An dieser Stelle soll nicht unerwähnt bleiben, inwiefern sich der Fokus auf die Satzplanung auswirkt.

Die mündliche Beschreibung ditransitiver Ereignisse mit drei Ereignisteilnehmern kann dazu dienen, den Fokus zu untersuchen. Im Gegensatz zur Beschreibung eines transitiven Ereignisses, bei der mit der Selektion eines Referenten als *starting point* der Versprachlichung die syntaktische Struktur bereits ungefähr festgelegt wird, müssen sich Sprecher bei der Beschreibung eines ditransitiven Ereignisses weiterhin über die Reihenfolge innerhalb des Prädikates entscheiden. An dieser Stelle soll folgendes Ereignis als Beispiel angeführt werden: *Der Professor übergibt das Buch an die Studentin*. Wenn das Agens (*der Professor*) aufgrund seines Topikstatus als *starting point* der Verbalisierung ausgewählt und satzinitial angeordnet wird, ist die syntaktische Struktur des Beschreibungssatzes dennoch immer noch nicht festgelegt. Sprecher müssen stattdessen entscheiden, in welcher Reihenfolge sie den Rezipienten (*die Studentin*) und das Thema (*das Buch*) erwähnen. In dieser Situation stehen deutschen Sprechern im Prinzip zwei Möglichkeiten zur Verfügung: Erstens können

sie den Rezipienten vor Thema anordnen, was eine DO-Konstruktion hervorruft (*Der Professor übergibt der Studentin das Buch*). In dieser Konstruktion ist der Rezipient das indirekte und das Thema das direkte Objekt. Zweitens kann der Rezipient satzfinal positioniert werden, was dann zur Bildung einer PO-Konstruktion führt, mit dem Rezipienten als das Präpositionalobjekt und dem Thema als das direkte Objekt (*Der Professor übergibt das Buch an die Studentin*). Welche von diesen beiden Konstruktionen letztendlich zur Verbalisierung des Kommentars ausgewählt wird, kann von der Zuweisung des Fokusstatus an den Referenten abhängig sein. Es wäre daher interessant zu sehen, wie sich der Fokus in einem derartigen Prozess auswirkt bzw. mit anderen Faktoren wie zum Beispiel der semantischen oder visuellen Salienz interagiert.

Diese Frage kann auch mit dem Versuchsdesign dieser vorliegenden Arbeit beantwortet werden, aber muss in zwei Punkten geändert werden. Zum einen betrifft die Änderung die Target-Frage, die in der Preview-Phase gestellt wird. Dieses Experiment handelt vom Prozess der Fokus-Einbindung und daher soll mit der Target-Frage der Fokus- und nicht der Topik-Status hinterfragt werden. Eine Methode dafür wären zum Beispiel W-Fragen. Wenn man beispielsweise den Rezipienten *die Studentin* als Fokus des zu bildenden Satzes erfragen möchte, dann stellt man folgende Frage: *Wer bekommt das Buch vom Professor?*. Mit der Frage *Was bekommt die Studentin vom Professor?* weist man hingegen dem Thema *das Buch* das Fokuspotential zu. Die zweite Änderung betrifft die Aufgabe der Probanden. Da wir uns nur für den Linearisierungsprozess des Rezipienten und des Themas innerhalb des Prädikates interessieren, steht die Positionierung des Agens als Subjekt in einem ditransitiven Satz nicht im Fokus. Aus diesem Grund wird der Satzanfang (z.B. *Der Professor übergibt ...*) immer angegeben, und die Aufgabe der Probanden besteht nur darin, den Satz mithilfe des Target-Bildes im Hinblick auf die Fragestellung zu vervollständigen.

Selbstverständlich hat dieses Design aber auch Schwachpunkte, auf die ich eingehen möchte. Ein Problem ist der mögliche Einfluss der konzeptuellen Zugänglichkeit bzw.

die Belebtheit des Rezipienten und des Themas. In einem ditransitiven Ereignis ist der Rezipient meistens belebt und das Thema in der Regel unbelebt. Ein zweites Problem betrifft die visuelle Salienz des Themas und des Rezipienten. Denn der Rezipient kann oft physikalisch größer als das Thema sein. Und deswegen ist es wichtig, bei der Vorbereitung von Versuchsmaterialien bzw. bei der Erstellung von Stimuli diese zwei Faktoren zu berücksichtigen.

Des Weiteren ist zu sagen, dass die Konzeptualisierung der Satzproduktion ein hochgradig komplexer kognitiver Prozess ist, auf den sich eine Reihe von Faktoren auswirken. In der vorliegenden Arbeit wurden nur drei Einflussfaktoren untersucht und anhand deren Wirksamkeit eine Verarbeitungsprioritätenliste entwickelt. Weitere mögliche Einflussfaktoren wie etwa die Belebtheit oder die Konkretheit sind in dieser Arbeit nicht berücksichtigt worden. Daher sind weitere Experimente diesbezüglich sinnvoll, um die relative Einflussstärke anderer Faktoren zu untersuchen und die Prioritätenliste zu vervollständigen.

Weiterhin wäre es interessant zu untersuchen, inwieweit und ob die in dieser Arbeit formulierte bzw. entwickelte Prioritätenliste von Einflussfaktoren für andere Sprachen gültig ist. Durch den Vergleich zwischen den Befunden aus der englischen und der deutschen Satzproduktion lässt sich schon absehen, dass Faktoren in unterschiedlichen Zielsprachen mit unterschiedlichen Gewichtungen in den Konzeptualisierungs- und syntaktischen Enkodierungsprozess einfließen. Daher ist es sinnvoll, Einflussfaktoren auf die Satzproduktion in unterschiedlichen Sprachen empirisch nachzugehen.

9.4 Allgemeine Schlussfolgerung

Die vorliegende Dissertation hatte zum Ziel, psycholinguistische Verarbeitungsprozesse in der Satzproduktion weiterführend zu untersuchen. In insgesamt drei Experimenten beschrieben deutschsprachige Probanden hierfür visuell dargebotene Stimuli, wobei die visuelle, semantische und kontextuelle Salienz im

Rahmen dieser Studie systematisch manipuliert wurden. Ziel der Studie war es, die Art der Informationsverarbeitung während des Sprechens sowohl in pragmatisch neutralen als auch in pragmatisch variierenden Kommunikationssituationen festzustellen.

Zur Beschreibung der Arbeitsweise des Satzproduktionssystems wird in der psycholinguistischen Forschung im Allgemeinen zwischen der linearen und der strukturellen Inkrementalität unterschieden. In Bezug auf die lineare Inkrementalität wird angenommen, dass der Selektion von *starting points* und der darauffolgenden grammatischen Enkodierung Eigenschaften einzelner Ereignispartizipanten, wie etwa die visuelle Salienz, zugrunde liegen. Bei der strukturellen Inkrementalität geht man davon aus, dass der Aufbau einer strukturellen Ereignisrepräsentation ausschlaggebend für genannte Prozesse ist. In früheren Studien (vgl. hierzu u.a. Griffin & Bock, 2000) bezieht sich die strukturelle Repräsentation überwiegend auf die semantische Struktur eines Ereignisses. In der vorliegenden Studie wurde hingegen zusätzlich eine andere Art struktureller Ereignisrepräsentation, nämlich die Informationsstruktur, als Faktor der Satzplanung mit einbezogen.

Durch die Analyse der Sprech- und Blickbewegungsdaten wurde in der hier durchgeführten Studie festgestellt, dass in pragmatisch neutralen Situationen der semantischen Struktur eines Ereignisses eine zentrale Bedeutung für die Selektion von *starting points* zukommt. Die visuelle Salienz hingegen wirkte sich lediglich auf die visuelle Wahrnehmung aus. Sobald die semantische Struktur aufgebaut wurde, wird der Effekt der visuellen Salienz umgehend abgelöst. Diese Ergebnisse stimmen mit der Annahme der strukturellen Inkrementalität überein und weichen daher von Studienergebnissen zur englischen Sprache (vgl. u.a. Gleitman et al., 2007) ab, welche die lineare Inkrementalitätshypothese untermauerten. Dieser sprachspezifische Unterschied zwischen dem Deutschen und dem Englischen deutet darauf hin, dass die Verarbeitungsmechanismen von Informationen während des Sprechens nicht universal gleich, sondern von Sprache zu Sprache verschieden sind.

Aktuell von besonderem Interesse für die psycholinguistische Forschung ist die Untersuchung der Satzproduktion im Rahmen eines Kontextes, bei dem verschiedene Ereignisteilnehmer einen unterschiedlichen Diskurstatus haben. In diesem Zusammenhang konnte in der vorliegenden Studie der Einfluss eines weiteren, aber bisher wenig betrachteten Faktors gezeigt werden: die kontextuelle Salienz. Dabei wurde deutlich, dass die Wirkung der kontextuellen Salienz die der semantischen Salienz übertreffen kann und als Grundlage der Selektion von *starting points* dient. Die kontextuelle Salienz kann die Informationsstruktur, welche als eine Art struktureller Ereignisrepräsentation auch auf der Mikroplanungsebene generiert ist, widerspiegeln. Aus diesem Grund kann der Befund, dass die kontextuelle Salienz eine ausschlaggebende Rolle bei der Selektion von *starting points* spielt, als Beweis für die strukturelle Inkrementalität angesehen werden. Durch die Studie wurde demnach bestätigt, dass die strukturelle Inkrementalitätshypothese auch für eine pragmatisch variierende Kommunikationssituation gelten kann.

Der Befund bezüglich des Einflusses der kontextuellen Salienz auf die Satzplanung ist insofern wichtig, als er die bisherigen Satzproduktionsmodelle (vgl. hierzu u.a. Bock & Levelt, 1994), welche der Rolle der kontextuellen Information nur eine vergleichsweise geringe Beachtung geschenkt haben, bereichern kann. Darüber hinaus lässt sich die Inkrementalitätstheorie auch durch diesen Befund erweitern. In älteren Studien erwies sich bisher nur der semantischen Ereignisstruktur als zuverlässiger Indikator für die strukturelle Inkrementalitätshypothese. Durch die Identifizierung des kontextuellen Einflusses im Rahmen der vorliegenden Dissertation wird ein neuer Maßstab zur Unterscheidung der linearen und der strukturellen Inkrementalität gesetzt, nämlich die Informationsstruktur. Wenn der Satzproduktion einer Sprache eine der beiden strukturellen Repräsentationen zugrunde liegt, kann dieser Vorgang dann als ein strukturell inkrementell verlaufender Prozess beschrieben werden. Des Weiteren liefert dieser Befund ebenfalls eine neue Perspektive zur Interpretation der Ergebnisse von Gleitman et al. (2009): Dass sich die strukturell inkrementelle Arbeitsweise in der Satzproduktion im Englischen bisher nicht feststellen ließ, kann daran liegen, dass

keine kontextuellen Informationen in dem Produktionsvorgang mit einbezogen wurden. In diesem Zusammenhang ist es interessant zu testen, welcher inkrementellen Arbeitsweise die Satzproduktion im Englischen folgt, wenn der Kontext mit einbezogen wird.

Vom Befund bezüglich des Verhältnisses zwischen dem Kontext und der Satzproduktion ausgehend, wurde in der Studie weiterhin festgestellt, in welcher prozeduralen Einbindung sich die kontextuelle Salienz bzw. die Topikalität bei der Satzplanung auswirken. Dies betraf vor allem die Repräsentation des Topiks und dessen Wirkungsweise. In Bezug auf die Repräsentation des Topiks haben meine Ergebnisse gezeigt, dass Topiks sowohl linguistisch als auch konzeptuell repräsentiert sind. Als linguistisches Element verleiht das Topik syntaktischen Funktionen wie etwa dem Subjekt eine positionelle Eigenschaft, was zu einer Belegung der Satzanfangsposition führen kann. Auf der konzeptuellen Ebene steht das Topik im Mittelfeld des Gesprächsinteresses und weist somit einen hohen Aufmerksamkeitsstatus auf. Dies kann eine frühere konzeptuelle Enkodierung des topikalischen Referenten hervorrufen.

Besonders neu und interessant ist der Verarbeitungszeitverlauf des Topiks innerhalb des Satzproduktionssystems. Es wurde in dieser Studie herausgefunden, dass das Topik auf die semantische Struktur abgebildet, bevor es an die syntaktische Funktionszuweisungsebene weitergereicht und schlussendlich satzinitial angeordnet wird. Der Befund, dass eine Verbindung zwischen der Informationsstruktur und der semantischen Struktur auf der Mikroplanungsebene stattfindet, steht im Wesentlichen im Einklang mit der Grundidee der strukturellen Inkrementalität. Sie deutet darauf hin, dass die beiden Arten struktureller Repräsentationen während der Satzplanung zusammenwirken. Meine Studie ist somit – soweit mir als Autor der Arbeit bekannt ist – die erste Studie, die den Zeitverlauf der Verarbeitung kontextueller Information bzw. des Topiks innerhalb des Satzproduktionssystems identifizierte und somit Erkenntnisse über die Mechanismen der Mikroplanung liefert.

Des Weiteren konnten die in dieser Studie durchgeführten Experimente die Wirksamkeit der genannten Einflussfaktoren, also der visuellen, semantischen und kontextuellen Salienz, bei der visuellen Wahrnehmung und der Selektion von *starting points* genauer darstellen als es in bisherigen Studien der Fall war. Im Gegensatz zu früheren Studien, in denen Einflussfaktoren meistens isoliert untersucht wurden, wurden sie in der vorliegenden Studie in einer und derselben Satzproduktionssituation zusammengeführt, was mir erlaubte, die Interaktion zwischen diesen Einflussfaktoren zu untersuchen und ihre Wirksamkeit direkt miteinander zu vergleichen. Es wurde herausgefunden, dass jeder dieser Einflussfaktor eine Auswirkung auf die Satzplanung hat, jedoch zu unterschiedlichen Zeiten und in unterschiedlichen Ausmaß.

Demgemäß wurde eine Prioritätenliste in der vorliegenden Arbeit entwickelt, die die Wirksamkeit der jeweiligen Faktoren während der Satzplanung beschreibt. Die kontextuelle Salienz nimmt eine vorrangige Position auf dieser Liste ein, was darauf hindeutet, dass die Verarbeitung kontextueller Information bei der Satzplanung wesentlichen Vorrang hat. Die semantische Salienz hingegen hat eine wichtige Rolle in pragmatisch neutralen Satzproduktionssituationen und im Gegensatz dazu, steht die visuelle Salienz von allen drei Faktoren an unterster Stelle. Diese Liste ist wichtig, da sie nicht nur einen Einblick im Verhältnis mitwirkender Faktoren während der Satzproduktion gewährt, sondern auch Sprechern einen Leitfaden zur Aufarbeitung komplexerer Kommunikationssituationen, in denen die unterschiedlichen Faktoren mit unterschiedlichen Referenten korrelieren, liefert.

Die durchgeführte Studie kann auch als ein Versuch für ein neues experimentelles Paradigma angesehen werden, welches es ermöglicht, Interaktionen zwischen verschiedenen Faktoren zu untersuchen. Es ist ein flexibles Paradigma und kann je nach untersuchter Faktoren in weiteren Studien angepasst und variiert werden. Zudem bilden die gefundenen Resultate und das hier entwickelte Versuchsparadigma eine anregende Grundlage für weitere Untersuchungen.

10 Literaturverzeichnis

- Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: an activation-based model. *Cognitive Science*, 26, 39-83.
- Arnold, J. E., Kaiser, E., Kahn, J. M., & Kim, L. K. (2013). Information structure: linguistic, cognitive, and processing approaches. *WIREs Cognitive Science*, 4, 403-413.
- Arnold, J. E., Wasow, T., Losongco, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76(1), 28-55.
- Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Becker, W. (1991). Saccades. In R. H. S. Carpenter (Hg.), *Eye Movements*. Boston: CRC Press, 95-137.
- Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double-step stimuli. *Vision Research*, 19, 967-983.
- Beedham, C. (1982). *The Passive Aspect in English, German and Russian*. Tübingen: Gunter Narr Verlag.
- Biederman, I. (1981). On the semantics of a glance at a scene. In: M. Kubovy, & J. R. Pomerantz (Hg.), *Perceptual Organization*. Hillsdale, NJ: Erlbaum, 213-253.
- Bierwisch, M. (1983). Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In R. Ruzicka, & W. Motsch (Hg.), *Untersuchung zur Semantik*. Berlin: Akademie Verlag, 61-99.
- Bierwisch, M., & Lang, E. (1987). Etwas länger – viel tiefer – immer weiter: Epilog zum Dimensionsadjektivprojekt. In M. Bierwisch, & E. Lang (Hg.), *Grammatische und konzeptuelle Aspekte von Dimensionsadjektiven (Studia grammatic 26/27)*. Berlin: Akademie-Verlag, 649-699.
- Bock, J. K. (1977). The effect of a pragmatic presupposition on syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior*, 16, 723-734.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355-387.
- Bock, J. K. (1987a). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, 26, 119-137.
- Bock, J. K. (1987b). Coordinating words and syntax in speech plans. In A. Ellis (Hg.), *Progress in the Psychology of Language, Vol. 3*. London: Erlbaum, 337-390.
- Bock, J. K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, 8(1), 57-99.

- Bock, J. K. & Ferreira, V. S. (2014). Syntactically speaking. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Hg.), *The Oxford Handbook of Language Production*. New York: Oxford University Press, 21-46.
- Bock, J. K. & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior*, 19, 467-484.
- Bock, J. K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, 48, 653-685.
- Bock, J. K., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. In F. Ferreira & J. Henderson (Hg.), *The Integration of Language, Vision, and Action: Eye Movements and the Visual World*. New York: Psychology Press, 249-278.
- Bock, J. K., & Levelt, W. J. M. (1994). Language production: grammatical encoding. In M. Gernsbacher (Hg.), *Handbook of Psycholinguistics*. San Diego: Academic Press, 945-984.
- Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: bridging the syntactic cleft. *Psychological Review*, 99, 150-171.
- Bock, K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47-67.
- Branigan, H. P., & Feleki, E. (1999). Conceptual accessibility and serial order in Greek language production. In M. Hahn, & S. C. Stoness (Hg.), *Proceedings of the 21th Conference of the Cognitive Science Society*. Mahwah: Lawrence Erlbaum, 96-101.
- Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118, 172-189.
- Bresnan, J., & Kanerva, J. M. (1992). The thematic hierarchy and locative inversion in UG. A reply to Paul Schachter's comments. In T. Stowell, & E. Wehrli (Hg.): *Syntax and Semantics 26: Syntax and the Lexicon*. New York: Academy Press. 111-125.
- Broadbent, D. E. (1966). A difficulty in assessing bimodality in certain distributions. *Mathematical and Statistical Psychology*, 19(1), 125-126.
- Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequenas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition*, 109, 274-280.
- Cai, Z. G., Pickering, M., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language*, 66, 833-849.
- Carpenter, R. H., & Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, 377(6544), 59-62.
- Carroll, J. B. (1958) Process and content in psycholinguistics. In R. Glaser (Hg.), *Current Trends in the Description and Analysis of Behavior*. Pittsburgh: Pittsburgh Press, 175-200

- Caspi, A., Beutter, B. R., & Eckstein, M. (2004). The time course of visual information accrual guiding eye movement decisions. *PNAS*, *101*(35), 13086-13090.
- Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. N. Li (Hg.), *Subject and Topic*. New York, Academic Press, 27-55.
- Chang, F., Bock, K., & Goldberg, A. (2003). Can thematic roles leave traces of their places? *Cognition*, *90*, 29-49.
- Choi, H. W. (1996). *Scrambling: Optimality-theoretic Interaction between Syntax and Discourse*. Dissertation, Stanford University.
- Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*, *98*, 105-135.
- Clark, H. H., & Clark, E. V. (1977). *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Clark, H., & Haviland, S. (1977). Comprehension and given-new contract. In: R. Freedle (Hg.), *Discourse Production and Comprehension*. Norwood, New Jersey: Ablex Publishing Corporation, 1-40.
- Clark, E. V. (1997). Conceptual perspective and lexical choice in language acquisition. *Cognition*, *64*, 309-343.
- Claus, B. (2011). Establishing salience during narrative text comprehension: A simulation view account. In C. Chiarcos, B. Claus, & M. Grabski (Hg.), *Salience: Multidisciplinary Perspectives on its Function in Discourse*. Berlin: Walter de Gruyter, 251-277.
- Cowles, H. W., & Ferreira, V. (2012). The influence of topic status on written and spoken sentence production. *Discourse Processes*, *49*(1), 1-28.
- Crystal, D. (1997). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *3*, 283-321.
- Dell, G. S., Chang, F. & Griffin, Z. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517-542.
- De Smedt, K. (1990). IPF: an incremental parallel formulator. In R. Dale, C. Mellish, & M. Zock (Hg.), *Current Research in Natural Language Generation*. London: Academic Press, 167-192.
- Dietrich, R. (1999). On the production of word order and the origin of incrementality. In: R. Klabunde & C. v. Stutterheim (Hg.), *Representations and Processes in Language Production*. Wiesbaden: Deutscher Universitäts-Verlag, 57-87.
- Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*(2), 129-143.

- Dryer, M. (2013). Order of subject, object and verb. In M. S. Dryer, & M. Haspelmath (Hg.), *The World Atlas of Language Structure Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/81> [zuletzt geprüft:: 11.12.2018]
- Dowty, D. (1995). Toward a minimalist theory of syntactic structure. In H. Bunt, & A. van Horck (Hg.), *Discontinuous Constituency*. Berlin und New York: Mouton de Gruyter, 11-62.
- Dürscheid, C. (2012). *Syntax: Grundlagen und Theorien*. Göttingen: Vandenhoeck & Ruprecht.
- Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, 42(5), 621-636.
- Everling, S., & Fischer, B. (1998). The antisaccade: a review of basic research and clinical studies. *Neuropsychologia*, 36, 885-899.
- Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46, 57-84.
- Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. *Journal of Memory and Language*, 48, 379-398.
- Ferreira, V. S. & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40(4), 296-340.
- Ferreira, F., & Henderson, J. M. (1998). Linearization strategies during language production. *Memory and Cognition*, 26, 88-96.
- Ferreira, V. S. & Slevc, R. L. (2009). Grammatical encoding. In M. G. Gaskell (Hg.), *The Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press, 453-469.
- Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, 32(6), 669-692.
- Fillmore, C. J. (1968). The case for case, in E. Bach, & R. T. Harms (Hg.), *Universals in Linguistic Theory*. New York: Holt Rinehart & Winston, 1-88.
- Fillmore, C. J. (1977). Topics in lexical semantics, in R. W. Cole (Hg.), *Report of the 22nd Annual Roundtable Meeting on Linguistics and Language Studies*. Washington, DC: Georgetown University Press, 35-56.
- Fischer, B., & Weber, H. (1993). Express saccades and visual attention. *Behavioral and Brain Sciences*, 16, 553-610.
- Fischer, K. (2013). *Satzstrukturen im Deutschen und Englischen. Typologie und Textgeneralisierung*. Berlin: Akademie-Verlag.

- Fleischer, Z., Pickering, M. J., & Mclean, J. F. (2012). Shared information structure: Evidence from cross-linguistic Priming. *Bilingualism: Language and Cognition*, *15*, 568-579.
- Frey, W. (2004). A medial topic position for German. *Linguistische Berichte*, *198*, 153-190.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Hg.), *The Psychology of Learning and Motivation (Vol. 9)*. New York: Academic Press, 133-177.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Hg.), *Language Production, Vol 1: Speech and Talk*. London: Academic Press, 177-220.
- Gerwien, J., & Flecken, M. (2015). There is no prime for time: The missing link between form and concept of progressive aspect in L2 production. *International Journal of Bilingual Education and Bilingualism*, *18(5)*, 561-587.
- Gilchrist, I. D. (2011). Saccades. In: S. P. Liversedge, I. D. Gilchrist, & S. Everling (Hg.), *The Oxford Handbook of Eye Movements*. Oxford: Oxford University Press, 85-94.
- Givon, T. (1979). *On Understanding Grammar*. New York: Academic Press.
- Glaser, W. R., & Döngelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception Performance*, *10(5)*, 640-654.
- Gleitman, L. R., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. *Cognition*, *58*, 321-376.
- Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, *57*, 544-569.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Hg.), *Universals of Human Language*. Cambridge: MIT Press, 73-113.
- Griffin, Z. M. (2001). Gaze duration 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.
- Griffin, Z. M., & Ferreira, V. S., (2006). Properties of spoken language production. In M. J. Traxler, & M. A. Gernsbacher (Hg.), *Handbook of Psycholinguistics*. Amsterdam, Boston: Elsevier/Academic Press, 21-59
- Gundel, J. K. (1988). Universals of topic-comment structure. In M. Hammond, E. Moravcsik, & J. Wirth (Hg.), *Studies in Syntactic Typology*. Amsterdam: John Benjamins, 209-239.

- Härtl, H. (1999). „fürchten“ vs. „ängstigen“. Thematische Rollen und Ereignisstrukturen psychischer Verben in einem Modell der Sprachproduktion. In I. Waschmuth, & B. Jung (Hg.), *KogWis 99: Proceedings der 4. Fachtagung der Gesellschaft für Kog. –wissenschaft*. St. Augustin: Infix, 189-194.
- Hallet, P. E. (1998). Primary and secondary saccades to goals defined by instruction. *Vision Research*, 20, 329-339.
- Hallet, P. E., & Adams, B. D., (1980). The predictability of saccadic latency in a novel voluntary oculomotor task. *Vision Research*, 20(4), 329-339.
- Hallet, P. E. & Zacks, R. T. (1998). The predictability of saccadic latency in a novel voluntary oculomotor task. *Vision Research*, 20, 329-339.
- Halliday, M. A. K. (1967a). *Grammar, Society and the Noun*. London: H.K Lewis for University College London.
- Halliday, M. A. K. (1967b). Notes on transitivity and theme in English, Part 2. *Journal of Linguistics* 3, 199-244.
- Hare, M. L., & Goldberg, A. E. (1999). Structural priming: Purely syntactic? In M. Hahn, & S. C. Stones (Hg.), *Proceedings of the 21th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum, 208-211.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, 50, 243-271.
- Henderson, J. M., & Macquistan, A. D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics*, 53(2), 221-230.
- Herweg M., & Maienborn, C. (1992). Konzept, Kontext, Bedeutung – Zur Rolle der Zwei-Ebenen-Semantik in einem Modell der Sprachproduktion. *Hamburger Arbeitspapiere zur Sprachproduktion*, 9, 7-31.
- Hockett, Ch. F. (1958). *A Course in Modern Linguistics*. New York: McMillan.
- Hoffman, D. D., & Singh, M. (1997). Saliency and visual parts. *Cognition*, 63, 29-78.
- Hooge, I. T., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, 58(7), 969-976.
- Humphreys, K. R., & Bock., J. K. (2005). Notional number agreement in English. *Psychonomic Bulletin & Review*, 12(4), 689-695.
- Hwang, H. & Kaiser, E. (2014). Having a syntactic choice is not always better: the effects of syntactic flexibility on Korean production. *Language, Cognition and Neuroscience*, 29(9), 1115-1131.
- Hwang, H. & Kaiser, E. (2015). Accessibility effects on production vary cross-linguistically: Evidence from English and Korean. *Journal of Memory and Language*, 84, 190-204.

- Ibbotson, P., Lieven, E., & Tomasello, M. (2013). The attention-grammar interface: Eye-gaze cues structural choice in children and adults. *Cognitive Linguistics*, 24(3), 457-481.
- Itti, L., & Koch, C. (2009). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10-12), 1489-1506.
- Irwin, D. E., & Brockmole, J. R. (2004). Suppressing where but not what: The effect of saccades on dorsal- and ventral-stream visual processing. *Psychological Science*, 15(7), 467-473.
- Jackendoff, R.S. (1983). *Semantics and Cognition*. Cambridge: MIT Press.
- Jackendoff, R.S. (1990). *Semantic Structures*. Cambridge: MIT Press.
- Jespersen, O. (1924). *The Philosophy of Grammar*. London: Allen & Unwin.
- Johnson-Laird, P. N. (1968). The choice of the passive voice in a communicative task. *British Journal of Psychology*, 59, 7-15.
- Keenan, E. & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8, 63-99.
- Keil, F.C. (1979). *Semantic and Conceptual Development: An Ontological Perspective*. Cambridge, MA: Harvard University Press.
- Kempen, G., & Harbusch, K. (2004). Word order scrambling as a consequence of incremental sentence production. In H. Härtl, & H. Tappe (Hg.), *Mediating between Concepts and Grammar*. Berlin: Mouton De Gruyter, 141-164.
- Kempen, G., & Hoenkamp, E. (1982). Incremental sentence generation: implications for the structure of a syntactic processor. In J. Horecký (Hg.), *Proceedings of the 9th International Conference on Computational Linguistics*. Amsterdam: North-Holland, 151-156.
- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201-258.
- King, J. (1996). Structured propositions and sentence structure. *Journal of Philosophical Logic*, 25, 495-521.
- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *The American Journal of Psychology*, 77(4), 576-588.
- Klein, W., & Stutterheim, C. v. (1987). Quaestio and referentielle Bewegung in Erzählungen. *Linguistische Berichte*, 109, 163-185.
- Klein, W., & Stutterheim, C. v. (1992). Textstruktur und referentielle Bewegung. *Zeitschrift für Literatur und Linguistik*, 86, 67-92.
- Klein, W., & Stutterheim, C. v. (1998). Referential movement in descriptive and narrative discourse. In R. Dientrich, & C. F. Graumann (Hg.), *Language Processing in Social Context*. North Holland: Elsevier Science Publishers B.V., 39-76.

- Konokpa A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, 73, 1-40.
- Krifka, M. (2007). Basic notions of information structure. In: C. Féry, G. Fanselow & M. Krifka (Hg.): *Interdisciplinary Studies on Information Structure 6*. Potsdam: Universitätsverlag, 13-55.
- Kuchinsky, S. E., & Bock, K. (2010). *From seeing to saying: Perceiving, planning, production*. Aufsatz präsentiert auf 23rd Meeting of the CUNY Human Sentence Processing Conference, New York, NY.
- La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll, & A. M. de Groot (Hg.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press, 289-307.
- La Heij, W., Boelens, H., & Kuipers, J. R. (2010). Object interference in children's colour and position naming: Lexical interference or task-set competition? *Language and Cognitive Processes*, 25(4), 568-588.
- Lambrecht, K. (1994). *Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press
- Lee, E. K., Brown-Schmidt, S., & Watson, D. W. (2013). Ways of looking ahead: incrementality in language production. *Cognition*, 129, 544-562.
- Lenz, A. (2010). Zum Salienzbeffriff und zum Nachweis salienter Merkmale. In C. Anders, M. Hundt, & A. Lasch (Hg.), *Perceptual Dialectology: Neue Wege der Dialektologie*. New York/Berlin: de Gruyter, 89-110.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1992). Accessing words in speech production: stages, processes and representations. *Cognition*, 42, 1-22.
- Levelt, W. J. M. (1994). The skill of speaking. In: P. Bertelson, P. Eelen, & G. d'Ydewalle (Hg.), *International Perspectives on Psychological Science: Vol. I: Leading Themes*. Hillsdale: Lawrence Erlbaum, 89-104.
- Levelt, W. J. M. (1999a). A blueprint of the speaker. In C. Brown, & P. Hagoort (Hg.), *The Neurocognition of Language*. Oxford: Oxford University Press, 83-122.
- Levelt, W. J. M. (1999b). Models of word production. *Trends in Cognitive Science*, 3 (6), 223-232.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavior Brain Science*, 11, 1-38.
- Ludwig, C. J. H., Mildinhall J. W., & Gilchrist, L. D. (2007). A population coding account for systematic variation in saccade dead time. *Journal of Neurophysiology*, 97, 795-805
- MacWhinney, B. (1977). Starting points. *Language*, 53(1), 152-168.

- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- McCarley, J. S., & Kramer, A. F. (2007). Eye movements as a window on perception and cognition. In R. Parasuraman & M. Rizzo (Hg.), *Neuroergonomics: The Brain at Work*. Oxford: Oxford University Press, 95-112.
- McDonald, J. L., Bock, J. K., & Kelly, M. H. (1991). Word and word order: semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, 25, 188-230.
- McPeck, R. M., & Keller, E. L. (2002). Saccade target selection in the superior colliculus during a visual search task. *Journal of Neurophysiology*, 88(4), 2019-2034.
- McPeck, R. M., Skavenski, A. A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, 40(18), 2499-2516.
- Messenger, K., Branigan, H. P., & McLean, J. F. (2011). Evidence for (shared) abstract structure underlying children's short and full passives. *Cognition*, 7, 268-274.
- Myachykov, A., & Tomlin, R. S. (2008). Perceptual priming and structural choice in Russian sentence production. *Journal of Cognitive Science*, 6, 31-48.
- Myachykov, A., Garrod, S., & Scheepers, C. (2009). Attention and syntax in sentence production: A critical review. *Discours*, 4. <https://journals.openedition.org/discours/7594#quotation> [zuletzt geprüft: 11.12.2018]
- Myachykov, A., Garrod, S., & Scheepers, C. (2010). Perceptual priming of structural choice during English and Finnish sentence production. In R. K. Mishra, & N. Srinivasan (Hg.), *Language and Cognition: State of the Art*. Munich: Lincom Europa, 54-72.
- Myachykov, A., Thompson, D., Scheepers, C., Garrod, S. (2011). Referential and visual cues to structural choice in visually situated sentence production. *Frontiers in Psychology*, 2, 396.
- Meyer, A. (2003). Die phonologische Realisierung. In G. Rickheit, T. Herrman, & W. Deutsch (Hg.), *Psycholinguistik*, 346-354
- Mizzo, M., & Caramazza, A. (1997). The retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1-14.
- Momma, S., Slevc, L. R., & Phillips, C. (2015). The timing of verb planning in Japanese sentence production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 42, 813-824.
- Musan, R. (2010). *Informationsstruktur*. Heidelberg: Universitätsverlag Heidelberg.
- Niemann, H., & Gauggel, S. (2010). Störungen der Aufmerksamkeit. In: P. Frommelt, & H. Lösslein (Hg.), *NeuroRehabilitation*. Berlin/Heidelberg: Springer, 145-170.

- Neumann, O. (1980). *Informationsselektion und Handlungssteuerung*. Dissertation. Universität Bochum.
- Norcliffe, E., Konopka, A. E., & Levinson, S. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, *30*, 1187-1208.
- Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: a computational model of fixation durations in scene viewing. *Psychological Review*, *117*(2), 382-405.
- Pechmann, T. (1994). *Sprachproduktion zur Generierung komplexer Nominalphrasen*. Opladen: Westdeutscher Verlag.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633-651.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3-25.
- Potter, M. C. (1975). Meaning in visual search. *Science*, *187*(4180), 965-966.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 509-522.
- Prat-Sala, M. (1997). *The Production of Different Word Orders: A Psycholinguistics and Developmental Approach*. Dissertation. University of Edinburgh.
- Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, *42*, 168-182.
- Prat-Sala, M., Shillcock, R., & Sorace, A. (2000). Animacy effects on the production of object-dislocated descriptions by Catalan-speaking children. *Journal of Child Language*, *27*, 97-117.
- Prentice, J. (1966). Response strength of single words as an influence in sentence behavior. *Journal of Verbal Learning and Verbal Behavior*, *5*, 429-433.
- Primus, B. (1996). Dependenz und Serialisierung: das Deutsche im Sprachvergleich. In E. Lang, & G. Zifonun (Hg.), *Deutsch – Typologisch: Jahrbuch des Instituts für Deutsche Sprache*. Berlin: de Gruyter, 57-91.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A Comprehensive Grammar of the English Language*. London: Longman.
- Pappert, S., & Pechmann, T. (2014). Priming word order by thematic roles: No evidence for an additional involvement of phrase structure. *The Quarterly Journal of Experimental Psychology*, *67* (11), 2260-2278.
- Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*, 460-499.

- Reinhart, T. (1982). Pragmatics and linguistics: an analysis of sentence topic. *Philosophica*, 27, 53-94.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Hg.), *Elements of Grammar: Handbook of Generative Syntax*. Dordrecht: Kluwer, 281-337.
- Roelof, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.
- Roelof, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 904-921.
- Roelof, A. (2011). Modeling the attentional control of vocal utterances: From Wernicke to WEAVER++. In J. Guendouzi, F. Loncke, & M. Williams (Hg.), *The Handbook of Psycholinguistic and Cognitive Processes: Perspectives in Communication disorders*. New York/Hove: Psychology Press, 189-208.
- Rosengren, I. (1993). Wahrfreiheit mit Konsequenzen – Scrambling, Topikalisierung und FHG im Dienste der Informationsstrukturierung. In: M. Reis (Hg.), *Wortstellung und Informationsstruktur*. Tübingen: Miemeyer, 251-312.
- Rosenholtz, R. (1999). A Simple saliency model predicts a number of motion popout phenomena. *Vision Research*, 39, 3157-3163.
- Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition*, 10, 257-275.
- Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, 57(8), 1030-1033.
- Sauppe S., Norcliffe E. J., Konopka A. E., Van Valin R. D., & Levinson S. C. (2013). Dependencies first: Eye tracking evidence from sentence production in Tagalog. In M. Knauff, M. Pauen, N, Sebanz, & I, Wachsmuth (Hg.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 1265–1270.
- Segalowitz, N. S. (1982). The perception of semantics relations in pictures. *Memory & Cognition*, 10, 381-388.
- Shin, J., & Christianson, K. (2009). Syntactic processing in Korean-English bilingual production: evidence from cross-linguistic structural priming. *Cognition*, 112, 175-180.
- Stutterheim, C. v., & Klein, W. (2002). Quaestio and L-Perspectivation. In C. F. Graumann, & K. Werner (Hg.), *Perspective and Perspectivation in Discourse*. Amsterdam: John Benjamins, 59-88.

- Talmy, L. (1975). *Figure and Ground in Complex Sentences. BLS 1*. Berkeley, California: Linguistics Society, 419-430.
- Talmy, L. (1978). Figure and ground in complex sentences. In J. H. Greenberg (Hg.), *Universals of Human Language. Volume 4: Syntax*. Stanford: Stanford University Press, 625-649.
- Talmy, L. (2000). *Toward a Cognitive Semantics, Vol I: Concept Structuring Systems*. Cambridge: The MIT Press.
- Tanaka, M. N. (2006). Animacy and syntactic structure in sentence formulation. In H. Shoboo (Hg.), *Proceedings of the Seventh Tokyo Conference on Psycholinguistics*. Tokyo, 273-298.
- Tanaka, M. N., Branigan, H. p., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language, 65*, 318-330.
- Tannenbaum, P., & Williams, F. (1968). Generation of active and passive sentence as a function of subject or object focus. *Journal of Verbal Learning and Verbal Behavior, 7*, 246-250.
- Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: capture of the eyes by new objects. *Psychological Science, 9*, 379-385.
- Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on eye movement control. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1595-1608.
- Thiering, M. (2011). Figure-ground reversals in language, *Gestalt Theory, 33*, 245-276.
- Tomlin, R. (1986). *Basic Word Order: Functional Principles*. London: Croom Helm.
- Tomlin, R. (1995). Focal attention, voice, and word order: An experimental, cross-linguistic study. In M. Noonan, & P. Downing (Hg.), *Word Order in Discourse*. Amsterdam: John Benjamins, 521-558.
- Tomlin, R. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In J. Nuyts, & E. Pederson (Hg.), *Language and Conceptualization*. Cambridge: Cambridge University Press, 162-189.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12(1)*, 97-136.
- Ullmann, S. (1996). *High-level Vision: Object Recognition and Object Cognition*. Cambridge, MA: MIT Press
- van Nice, K. Y. & Dietrich, R. (2003). Task-sensitivity of animacy effects: evidence from German picture descriptions. *Linguistics, 5*, 825-849.
- Vecera, S. P., Flevaris, A. V., & Filapek, J. C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science, 15(1)*, 20-26.

- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8, 314-319.
- Walker, R., & McSorley, E. (2006). The parallel programming of voluntary and reflexive saccades. *Vision Research*, 46(13), 2082-2093.
- Wagner, M. (2016). Information structure and production planning. In: C. Féry & S. Ishihara (Hg.), *The Oxford Handbook of Information Structure*. Oxford: Oxford University Press, 541–561.
- Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 423-440.
- Webb, A., Knott, A., & MacAskill, M. R. (2010). Eye movements during transitive action observation have sequential structure. *Acta Psychologica*, 133(1), 51-56.
- Weidner, R., & Fink, G. R. (2007). Wahrnehmung und Aufmerksamkeit. In F. Schneider & G. Fink (Hg.), *Funktionelle MRT in Psychiatrie und Neurologie*. Berlin/Heidelberg: Springer-Verlag, 301-318.
- Wheeldon, L. R., Meyer, A. S., & Smith, M. C. (2003). Language production, incremental. In L. Nadal (Hg.), *Encyclopedia of Cognitive Science*. London: Nature Publishing Group, 760-764.
- Wheeldon, L. (2013). Producing spoken sentences: The scope of incremental planning. In S. Fuchs, M. Weirich, D. Pape, & P. Perrier (Hg.), *Speech Production and Perception: Speech Planning and Dynamics*. Frankfurt: Peter Lang International Academic Publishers, 97-118.
- Wingfield, A. (1967). Perceptual and response hierarchies in object identification. *Acta Psychologica*, 26, 216-226.
- Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *American Journal of Psychology*, 81, 226-234.
- Wolfe, J. M. (1992). The parallel guidance of visual attention. *Current Directions in Psychological Science*, 1, 125-128
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Hg.), *Attention*. London: University College London Press, 13-74.
- Wundt, W. (1900). *Die Sprache*. Leipzig: Kroner.
- Zarcone, A., van Schijndel, M., Vogels, J., & Demberg, V. (2016). Saliency and attention in supralingual accounts of language processing. *Frontiers in Psychology*, 7, 844.
- Zimmer, R. (2005). *Handbuch der Sinneswahrnehmung: Grundlagen einer ganzheitlichen Bildung und Erziehung*. Freiburg im Breisgau: Verlag Herder.
- Zimmermann, R. (1997). Dimensionen des mentalen Lexikons aus der Perspektive des L2-Gebrauchs. In W. Börner, & K. Vogel (Hg.), *Kognitive Linguistik und*

Fremdsprachenerwerb: Das mentale Lexikon. Tübingen: Gunter Narr Verlag, 107-128.

Zionun, G., Hoffmann, L., Strecker, B. & Ballweg, J. (1997). *Grammatik der deutschen Sprache, Band 1.* Berlin, New York: Walter de Gruyter.