

**Subjekt-Verb-Kongruenz bei  
komplexen, numerusambigen Nominalphrasen**

Schriftliche Arbeit  
zur Erlangung des Akademischen Grades  
„Master of Arts“ (MA)  
an der Neuphilologischen Fakultät,  
Ruprecht-Karls-Universität Heidelberg

vorgelegt bei  
Erstgutachter: Dr. Johannes Gerwien  
Zweitgutachterin: PD Dr. Sandra Pappert  
Institut für Deutsch als Fremdsprachenphilologie

vorgelegt von  
Philip Kehl  
Fahrbachweg 1  
69126 Heidelberg  
Matrikelnr.: 3228720  
E-Mail: [philip.kehl@posteo.de](mailto:philip.kehl@posteo.de)

Heidelberg, 12. September 2019

# Inhaltsverzeichnis

<b>Einleitung .....</b>	<b>1</b>
<b>1 Kongruenz und syntaktische Struktur .....</b>	<b>5</b>
1.1 Kongruenz als linguistischer Begriff .....	5
1.2 Subjekt-Verb-Kongruenz im Deutschen .....	6
1.3 Kongruenz in komplexen syntaktischen Strukturen .....	7
1.3.1 Merkmale von Appositionen .....	9
1.3.2 Appositionen und Numerus-Kongruenz .....	10
1.4 Zwischenfazit .....	11
<b>2 Sprachverstehen .....</b>	<b>13</b>
2.1 Grundlagen .....	13
2.1.1 Das ‚Wissen‘ des menschlichen Satzanalysesystems .....	13
2.1.2 Modulare Erklärungsansätze .....	16
2.1.3 Nicht-modulare Erklärungsansätze .....	19
2.2 Die Verarbeitung von Kongruenz .....	20
2.2.1 Die Rolle des Arbeitsgedächtnisses .....	20
2.2.2 Die Rolle von Erwartungen .....	23
2.2.3 Agreement Attraction .....	25
2.2.4 Good-enough-Parsing .....	28
2.3 Die Verarbeitung von ‚Beides‘ + KoordP + Verb .....	30
<b>3 Empirische Untersuchung .....</b>	<b>36</b>
3.1 Die Methode des Eyetracking .....	36
3.1.1 Das Visual World Paradigm .....	36
3.1.2 Grundlagen für Visual-World-Studien zum Sprachverstehen .....	37
3.1.3 Die Hypothese des kognitiven Aufwands .....	39
3.2 Design der empirischen Untersuchung .....	40
3.2.1 Zusammensetzung der Items .....	41
3.2.2 Präsentation der Items .....	45
3.2.3 Abhängige Variablen und Erwartungen .....	47

3.3 Datenerhebung .....	49
3.3.1 Vorbereitung .....	49
3.3.2 Durchführung .....	49
3.4 Auswertung der Daten .....	50
3.4.1 Tastendrucke .....	50
3.4.2 Eyetracking-Daten .....	53
<b>4 Diskussion .....</b>	<b>56</b>
<b>Literatur .....</b>	<b>61</b>
<b>Anhang .....</b>	<b>66</b>

## **Tabellenverzeichnis**

Tabelle 1: Grammatische und semantische Kongruenz relevanter Elemente .....	8
Tabelle 2: Sätze und Hypothesen zu ihrer Verarbeitung .....	31
Tabelle 3: Sätze und Matches mit Numerus-Cues .....	34
Tabelle 4: Zusammensetzung der Critical- und Control-Items .....	41
Tabelle 5: Absolute Zahlen zur Akzeptabilität der Sätze .....	51
Tabelle 6: Tastendrucke: Statist. Vergleich zw. kritischen u. Kontrollbedingung .....	53
Tabelle 7: Tastendrucke: Statist. Vergleich zw. den kritischen Bedingungen .....	53
Tabelle 8: Ausprägungen der Satzstruktur und Kongruenzverhältnisse mit V .....	57

## **Abbildungsverzeichnis**

Abbildung 1: Statistische Auswertung der Tastendrucke .....	52
Abbildung 2: Fixationsproportionen der Areas of Interest .....	54

# Abkürzungsverzeichnis

AOI – Area of Interest

ConPlu – Conceptual Plural

Distr – Distraktor

KoordP – Koordinationsphrase

LongDep – Long Dependency

N – Nomen

NP – Nominalphrase

Plu – Plural

Prox – Proximity

Ref – Referent

S – Subjekt

Sing – Singular

SO – Stimulus-Onset

SOV – Subjekt Objekt Verb (lineare  
Anordnung in einem Satz)

V – Verb

Vers – Version (eines Satzes)

VP – Verbalphrase

## Einleitung

Kongruenz ist ein über Sprachen hinweg beobachtbares, in verschiedenen Ausprägungen auftretendes Phänomen. So ist es in vielen Sprachen der Fall, dass bestimmte Elemente hinsichtlich eines Merkmals, z. B. Numerus, Genus oder Belebtheit, mit anderen übereinstimmen müssen, um einen grammatisch wohlgeformten Satz zu bilden. Üblicherweise fordert ein Element bei einem später auftretenden entsprechende Kongruenzmerkmale. Für das Sprachverarbeitungssystem stellen Kongruenzphänomene eine Herausforderung dar, weil sich zwischen zwei kongruierenden Elementen oft andere befinden, die wiederum untereinander kongruieren. Bei der Sprachproduktion muss das kongruierende Element in der passenden Form gebildet werden; während des Sprachverstehens ist es die Aufgabe des Verarbeitungssystems, im Falle von zwischen den kongruierenden Elementen stehenden Wörtern, zu klären, welche Elemente miteinander kongruieren und welche nicht (vgl. Nicol et al., 1997, 569).

Im Rahmen meiner Arbeit habe ich das Verstehen von Sätzen wie in (1) untersucht, in denen das der Fall ist: Zwischen mehreren nominalen Elementen, in diesem Fall drei, und einem finiten Verb können potentiell Kongruenzbeziehungen aufgebaut werden.

- (1)        a. *Beides, Plastiktüte und Sprühdose, schadet dem Klima.*  
              b. *Beides, Plastiktüte und Sprühdose, schaden dem Klima.*  
              c. *Beides, Plastiktüte und Sprühdosen, schaden dem Klima.*

Aus formal-syntaktischer Perspektive scheint der Fall klar zu sein: Auf das erste nominale Element, das substantivierte Indefinitpronomen ‚*beides*‘, folgt ein Einschub in Form einer koordinierte Nominalphrase, die das Pronomen semantisch spezifiziert, bevor sich die Verbalphrase anschließt. Das darin stehende finite Verb steht im Singular, weil ‚*beides*‘ als formales Subjekt des Satzes dies fordert. ‚*Beides*‘ kann unabhängig vom Einschub als Subjekt auftreten und fordert ebenso ein Verb im Singular (2a.)

- (2)        a. *Beides schadet/\*schaden dem Klima.* (aber unklar, auf was referiert wird)  
              b. *Plastiktüte und Sprühdose schaden dem Klima.*

Die eingeschobene koordinierte Nominalphrase könnte, ließe man ‚*beides*‘ weg, auch eigenständig die Funktion des Subjekts übernehmen, ohne das sich die Semantik des Satzes verändert. Sie stellt einen semantisch-konzeptuellen und durch die Konjunktion ‚*und*‘ auch einen formalen Plural dar und würde ein Verb im Plural fordern (2b.).

Die drei Sätze in (1) sind nach meiner Intuition alle grammatisch akzeptabel und sollen im Rahmen der vorliegenden Arbeit untersucht werden. Die folgenden Hypothesen zur Auflösung von Abhängigkeiten zwischen nominalen Elementen und Verb sollen überprüft werden:

- Satz (1) a. – Long-Dependency-Hypothese

Das formale Subjekt ‚*Beides*‘ kongruiert mit dem Verb im Singular; die koordinierte NP [N1+N2] und das Verb kongruieren nicht miteinander:

┌──────────────────┐  
*Beides*, N1(Sing) + N2(Sing), V(Sing) ...

- Satz (1) b. – Conceptual-Plural-Hypothese

Die koordinierte NP wird als konzeptueller Plural [N1(Sing)+N2(Sing)] verarbeitet und kongruiert mit dem Verb im Plural; ‚*beides*‘ und Verb kongruieren nicht miteinander:

┌──────────┐  
*Beides*, [N1(Sing) + N2(Sing)], V(Plu) ...

- Satz (1) c. – Proximity<sup>1</sup>-Hypothese

Das unmittelbar präverbale N2(Plu) kongruiert mit V(Plu), weil es näher als die anderen potentiell kongruierenden Elemente am Verb steht:

┌───┐  
*Beides*, N1(Sing) + N2(Plu), V(Plu) ...

Wegerer (2012) erhob in ihrer Dissertation zur Numerus-Kongruenz von Subjekt und finitem Verb im Deutschen anhand eines Fragebogens die grammatische Akzeptabilität sogenannter sprachlicher Zweifelsfälle, u. a. auch der komplexen Satzstruktur ‚*Beides*‘ + *KoordP* + *Verb*. Sprachliche Zweifelsfälle dieser Art zeichnen sich dadurch aus, dass eine einschlägige Anzahl kompetenter Sprecher\*innen des Deutschen nicht ohne weiteres entscheiden kann, ob es ein Singular- oder ein Plural-Verb braucht, um die Struktur zu einem grammatikalisch korrekten Satz zu ergänzen. Die Ergebnisse zeigen, dass eine große Mehrheit der Befragten trotz des formalen Subjekts im Singular ein Plural-Verb wählte. Der Autorin

---

<sup>1</sup> *proximity* (engl.) – Nähe, Nachbarschaft, nähere Umgebung

zufolge ist in Sätzen mit einem bedeutungsschwachen Subjekt die Proximitätskongruenz ausschlaggebend.

Um ebenfalls zu prüfen, ob es sich bei den Konstruktionen um sprachliche Zweifelsfälle handelt und ob Proximitätskongruenz wirkt, wurden im Rahmen meiner Studie muttersprachlichen Sprecherinnen und -Sprechern des Deutschen innerhalb eines experimentellen Settings Sätze mit dieser und ähnlicher Struktur vorgespielt, zu denen sie ihre Grammatikalitätsurteile abgaben (mit der Auswahlmöglichkeit korrekt/inkorrekt). Neben dieser Art von Daten wurden in meinem Experiment anhand der Methode des Eyetracking Blickbewegungsdaten erhoben, um eine psycholinguistische Perspektive auf solche Satzstrukturen zu bekommen. Diese Methode liefert Online-Daten, die Aussagen über bestimmte Zeitpunkte oder Phasen innerhalb des Sprachverarbeitungsprozesses zulassen. Die Online-Daten, gepaart mit Grammatikalitätsurteilen, sind als Grundlage für die Analyse sprachlicher Verarbeitungsprozesse aufschlussreicher als die alleinige Auswertung einer Befragung bzw. der Abfrage von einer Person bewusst getroffener Entscheidungen, welche nur das Resultat eines Verarbeitungsprozesses widerspiegeln.

In der vorliegenden Studie wird Eyetracking vor dem Hintergrund des Visual World Paradigm eingesetzt, das systematische Wechselbeziehungen zwischen sprachlichen Verarbeitungsprozessen und visueller Information annimmt. Objekte, die in einem auditiven linguistischen Input genannt werden, werden auf entsprechenden Abbildungen präferiert angeschaut. Für meine Untersuchung gehe ich von der Hypothese des kognitiven Aufwandes aus, die annimmt, dass Blickbewegungen Aufmerksamkeit und den kognitiven Aufwand bei der Sprachverarbeitung widerspiegeln (Griffin, 2004), insbesondere dann, wenn Versuchspersonen anspruchsvolle Aufgaben zu lösen haben. In der vorliegenden Studie besteht die Aufgabe in der Beurteilung von grammatischen Zweifelsfällen.

In den drei zu untersuchenden Satzstrukturen wird das menschliche Satzanalyse-system mit verschiedenen möglichen Kongruenz-Beziehungen zwischen den präverbalen nominalen Elementen untereinander sowie dem finiten Verb konfrontiert. Vor diesem Hintergrund lässt sich vermuten, dass sich die Blickbewegungen, die während dem Hören der drei Sätze ausgeführt werden, unterscheiden, und zwar insbesondere in einer kritischen Phase der Verarbeitung: in den (Milli-)Sekunden, nachdem die Versuchspersonen das finite Verb gehört haben. In dieser Phase müssen die Merkmale der kongruenzfähigen Elemente miteinander abgeglichen werden. Konkret sollte sich dies in Blicken zu den entsprechenden Referenten widerspiegeln.

Die Untersuchung war explorativ gestaltet, in dem Sinne, dass die Messgröße, auf die hin die Eyetracking-Daten der kritischen Phase ausgewertet werden sollten, bis zum Abschluss der Erhebung offengelassen wurden (z. B. Anzahl der Sakkaden, zu welchem Objekt geht die erste Sakkade?, Fixationsproportionen).

Die Auswertung der Grammatikalitätsurteile ergab eine signifikant niedrigere Akzeptabilität für Satz (1a.) als für (1b.) und (1c.). Bei der Auswertung der Fixationsproportionen für das kritische Zeitfenster zeigten sich wie erwartet Unterschiede zwischen Kontroll- und kritischen Bedingungen; in der kritischen Bedingung zu Satz (1c.) ergab die Auswertung eine anfänglich starke Präferenz für das Bild, in dem N2(Plu) abgebildet war. Es lässt sich vermuten, dass bei der Verarbeitung der Sätze die formal-syntaktische Kongruenzregel entweder eine geringe Rolle spielt, oder gar keine, weil das formale Subjekt ‚beides‘ nicht mitanalysiert wird.

Grundlage für die vorliegende Untersuchung ist zum einen das Phänomen der Kongruenz. In Kapitel 1 wird der Begriff eingeführt und es wird thematisiert, wie Subjekt-Verb-Kongruenz in der syntaktischen Umgebungen auf verschiedenen Ebenen hergestellt werden kann. Anschließend wird versucht, die zu untersuchenden Sätze bzw. ihre Struktur ‚*beides*‘ + *KoordP* + *Verb* mit Hilfe des Begriffs der Apposition syntaktisch zu fassen, bevor auf die grammatische Akzeptabilität ebendieser Struktur in der Studie von Wegerer (2012) eingegangen wird.

Um eine allgemeine Vorstellung über Sprachverstehensprozesse zu bekommen, wird im Kapitel 2 dargelegt, wie das menschlichen Sprachanalysesystem funktioniert, wie es verschiedenen Herausforderungen bei der Verarbeitung begegnet und welche grob zu unterscheidenden Ansätze es zur Erklärung von Sprachverstehensprozessen gibt. Daraufgehend geht es darum, wie Arbeitsgedächtnis und wie Erwartungen die Verarbeitung von Kongruenzbeziehungen beeinflussen. Anschließend wird das Phänomen Agreement Attraction erörtert und der Erklärungsansatz des Good-enough-Parsing vorgestellt. Zum Abschluss von Kapitel 2 werden unter Berücksichtigung des bis dahin Ausgeführten Überlegungen zur Verarbeitung der Struktur ‚*beides*‘ + *KoordP* + *VP* angestellt.

In Kapitel 3 wird sich zunächst mit der Methode des Eyetracking und dem Visual World Paradigm auseinandergesetzt; darauf folgen eine ausführliche Beschreibung des Designs der Untersuchung, die Erhebung der Daten sowie die Datenauswertung, bevor abschließend die Ergebnisse diskutiert werden.



# 1 Kongruenz und syntaktische Struktur

## 1.1 Kongruenz als linguistischer Begriff

Der deutsche Grammatik-Duden (2016, 952) bezeichnet Kongruenz als die Übereinstimmung zweier unterschiedlicher Elemente in einem Satz bezüglich bestimmter morphologischer bzw. grammatischer Merkmale:

„Wenn zwei Bestandteile des Satzes oder eines Textes miteinander in einer engen Beziehung stehen, übernimmt der eine Bestandteil häufig vom anderen bestimmte grammatische Merkmale. Die beiden Bestandteile haben dann diese Merkmale gemeinsam. Man spricht hier von Übereinstimmung oder **Kongruenz**. Die Kongruenz verdeutlicht den Zusammenhang zwischen den jeweiligen Bestandteilen. Kongruenz kann die folgenden grammatischen Merkmalklassen betreffen:

- Person: 1., 2., 3. Person
- Numerus (grammatische Zahl): Singular und Plural
- Genus (grammatisches Geschlecht): Maskulinum, Femininum, Neutrum
- Kasus (Fall): Nominativ, Akkusativ, Dativ, Genitiv.“

Elemente, bei denen diese Merkmale morphologisch ausgedrückt und verändert werden, können den Merkmalklassen entsprechend flektiert werden. Kongruenz besteht in den Flexionskategorien, die die kongruierenden Elemente gemeinsam haben. Im Deutschen werden nominale Elemente in allen vier Flexionskategorien flektiert, verbale Elemente in Person und Numerus.

Ein Überblick über Elemente, zwischen denen im Deutschen Kongruenz besteht, lässt sich in *grammis*, dem Grammatischen Informationssystem des IDS Mannheim (2019), finden:

„1. In der Nominalphrase besteht Kongruenz in Numerus und Kasus zwischen Nomen, Artikel und attributivem Adjektiv (das Genus von Artikel und Adjektiv wird vom Nomen regiert!): *mit diesen rostigen Werkzeugen; ohne dieses rostige Werkzeug*

2. Im Satz besteht zwischen einer Nominalphrase in Subjektsfunktion und dem finiten Verb Kongruenz im Numerus: *Der Nagel ist zu kurz. Die Nägel sind zu kurz.*

3. Kasuskongruenz besteht zwischen einem Prädikativkomplement und seinem Bezugsnomen: *Ich bin ein gelehrter Mann. Man nennt mich einen gelehrten Mann.*

4. Kasuskongruenz besteht zwischen Adjunktphrase und Bezugsnomen: *Ich als blutiger Laie verstehe nichts davon. Mir als blutigem Laien kannst du alles erzählen.*

5. Kasuskongruenz besteht zwischen Apposition und Bezugsnomen: *Zu Hanne, dieser Langweilerin, brauchst du mich nicht mehr mitzunehmen. Auf Hanne, diese Langweilerin, kann ich gern verzichten.*“

Dem Grammatik-Duden und dem Überblick bei *grammis* zufolge handelt es sich bei Kongruenz also um ein nur die Form betreffendes Phänomen, wobei die Kategorien lediglich syntaktische Funktionen übernehmen. Der Begriff ist auf die Betrachtung von Beziehungen zwischen linguistischen Elementen in dem linguistischen System der Syntax beschränkt. Diese auf die formal-syntaktische Struktur gerichtete Sichtweise gilt es zu ergänzen um eine inhaltsbezogene. Morpheme, in denen die Kategorie Person kodiert ist, referieren nicht nur innerhalb des Sprachsystems aufeinander sondern auch auf außersprachliche Entitäten (vgl. Wegerer, 2012, 35).

## 1.2 Subjekt-Verb-Kongruenz im Deutschen

Es scheint eine einfache Regel zu geben, nach der im Deutschen Subjekt und finites Verb in einem Satz bezüglich der Kategorie Numerus miteinander zu kongruieren haben: Ein singularisches Subjekt fordert ein Verb im Singular, ein pluralisches Subjekt fordert ein Verb im Plural.

Die Subjekt-Verb-Kongruenz im Deutschen hängt im Grunde genommen von der Anzahl der Referenten ab, auf die sich das Subjekt bezieht. Sie wird aber nicht nur durch grammatische Bedingungen bestimmt. Im Vergleich zur Kategorie Genus, die als lexikalisches Merkmal eines Nomens ihm inhärent ist, und der Kategorie Kasus, die dem Nomen abhängig von seiner syntaktischen Funktion zugewiesen wird, ist die Kategorie Numerus variabel, d. h. ein Nomen wird entweder als Singular- oder als Pluralform realisiert, abhängig vom Inhalt der Aussage. Die Wahl des Numerus folgt eher semantischen Prinzipien. Genus und Kasus können nicht gewählt werden, weil sie syntagmatisch gebunden sind (vgl. Jaeger, 1992, 18). Dieser Analyse folgend können drei Arten von Kongruenz beschrieben werden.

*Kategorien der Numerus-Kongruenz* (vgl. Wegerer, 2012, 69, nach Jaeger, 1992, 56):

- *Grammatische Kongruenz*: Die grammatische Kongruenz bezeichnet die Übereinstimmung von Subjekt und Prädikat hinsichtlich grammatisch-formaler Eigenschaften und wird morphologisch realisiert. Bsp.: *Der Nagel ist zu kurz. Die Nägel sind zu kurz.*

- *Semantische Kongruenz*: Fallen Form und Inhalt eines Nomens nicht zusammen, kann der Inhalt ausschlaggebend für die morphologische Markierung des Verbs sein. In diesem Fall ist laut Grammatik-Duden (2016, 953) von einer „constructio ad sensum“ oder einer Konstruktion „kata synesin“ – „nach dem Sinn“ die Rede.

„Diese Sinnkongruenz ist semantisch begründet und richtet sich nach der begrifflichen Bedeutung, nicht nach einer grammatischen Kategorie. Sie fällt vor allem dort auf, wo grammatische und semantische Merkmale nicht aufeinander bezogen sind“ (Wegerer, 2012, 69f.).

Im Vergleich zum Deutschen gibt es Sprachen, in denen eine „constructio ad sensum“ häufiger vorkommt. So ist es in den folgenden Konstruktionen nicht entscheidend, dass das Subjekt im Singular steht, sondern dass mit dem Subjekt mehrere Personen bezeichnet werden: Im Englischen gilt der Satz *The police were chasing the crook* als korrekt; im Deutschen dagegen ist *\*Die Polizei jagten den Betrüger* nicht akzeptabel.

Der Unterschied zwischen formaler und semantischer Ebene wird in Ausdrücken wie ‚eine Menge Gläser steht/stehen im Schrank‘ oder ‚eine Vielzahl von Personen kam/kamen zur Demo‘ deutlich. Die Kopf-Nomina in der Subjektphrase haben morphologisch eine Singular-Form, referieren aber auf mehrere Entitäten. In diesen Fällen sind sowohl Singular- als auch Plural-Verb akzeptabel.

- *Proximitätskongruenz*: Diese Kategorie der Kongruenz beruht auf der linearen Position von Konstituenten im Satz, die Eigenschaften haben, nach denen sie auch als Subjekt fungieren könnten. Die Proximitätskongruenz bezeichnet die formale Übereinstimmung des finiten Verbs mit dem unmittelbar präverbal auftretenden Element. Sie wirkt anders als die beiden anderen Kategorien mittelbar, indem sie Elemente, die potentiell grammatisch und/oder semantisch mit anderen kongruieren, darin fördert, Bezug auf das Verb zu nehmen.

### 1.3 Kongruenz in komplexen syntaktischen Strukturen

Bisher wurde in Kapitel 1 darauf eingegangen, welche Elemente im Deutschen grammatisch-morphologisch miteinander kongruieren und welche Kategorien der Flexion dabei betroffen sein können. Im Fall von Subjekt und Verb bezieht sich Kongruenz vor allem auf Übereinstimmungen auf grammatischer Ebene, es können sich jedoch die Semantik oder



Bisher wurde die Koordinationsphrase als Einschub bezeichnet. Um einen Überblick über die Merkmale eines solchen Einschubs zu bekommen, wird im folgenden Abschnitt der syntaktische Status des Einschubs in der Konstruktion ‚*Beides*‘ + *KoordP* + *Verb* anhand des Begriffs der Apposition geklärt.

### 1.3.1 Merkmale von Appositionen

Wenn mit dem Begriff der Apposition gearbeitet wird, sollte vorweg erwähnt werden,

„[...] daß die nominalen Attribute, die als „Appositionen“ bezeichnet werden, bzgl. des syntaktischen Status, i. e. hinsichtlich der strukturellen Position [...] und der kategorialen Zuordnung, keine einheitliche, sondern eine äußerst heterogene Gruppe darstellen“ (Lawrenz, 1993, 3).

Manch eine Konstruktion, die von Autor\*innen als Apposition eingeordnet wird, besitzt bestimmte, von anderen als wichtig erachtete Merkmale nicht oder erfüllt gewisse Kriterien nicht. Außerdem werden im Zusammenhang mit diesem Begriff häufig Konstruktionen thematisiert, die nicht gefestigt sind und sich im Umbruch befinden (Eisenberg & Thierhoff, 2013, 256-259). Vor diesem Hintergrund soll im Folgenden versucht werden, die Konstruktion ‚*Beides*‘ + *KoordP* + *Verb* formal-syntaktisch grob einzuordnen.

Appositionen werden als eingeschobene Elemente wahrgenommen, was sich in vielen Fällen zum einen darin widerspiegelt, dass sie, betrachtet man die orthografische Ebene, von Kommata eingegrenzt sind, und zum anderen, dass sie auf phonologischer Ebene intonatorisch gekennzeichnet sind.

„Appositionen sind Nominalphrasen, die postnominal als Supplemente einer Nominalphrase eingesetzt werden können, mit der sie meist in Numerus und Kasus übereinstimmen. Appositionen – daher ihre Bezeichnung – wirken stets appositiv, d.h. nicht-restriktiv [und charakterisierend, P. K.]. Sie unterscheiden sich darin von Namensbestandteilen wie *Karl der Große* [...] die restriktiv zu interpretieren sind.“ (grammis, 2019)

Im Gegensatz dazu weist Lawrenz (1993, 61f.) darauf hin, dass Appositionen in vielen Fällen nicht weggelassen werden können, ohne dass sich die Semantik des Satzes verändert. Sie wirken sich in diesen Fällen restriktiv spezifizierend auf das Bezugsnomen aus; siehe die folgenden Beispielsätze:

(5) *Man interessierte sich nur für Müller, den Fußballer, nicht für Müller, den Menschen.*

(6) *Beides, Plastiktüte und Sprühdose, schadet dem Klima.*

(7) *Beides schadet dem Klima.*

Der Satz ohne Koordinationsphrase ist zwar grammatikalisch korrekt, ohne Kontext bleibt aber unklar, auf was referiert wird.

Für eine Charakterisierung von ‚*Beides*‘ + *KoordP* + *Verb* als lose Apposition spricht, dass die Koordinationsphrase bewegt werden, also z. B. im Vor- oder Nachfeld extraponiert erscheinen kann:

(8) *Plastiktüte und Sprühdose, beides schadet dem Klima.*

(9) *Beides schadet dem Klima, Plastiktüte und Sprühdose.*

Die drei zu untersuchenden Satzstrukturen erfüllen demnach einige typische Kriterien von Appositionen.

Wie ‚*beides*‘ und die *KoordP* hierarchisch genau zu einander stehen, soll an dieser Stelle nicht weiter erörtert werden. ‚*Beides*‘ ist das formale Subjekt des Satzes und die *KoordP* füllt dieses Subjekt semantisch. Die *KoordP* ist nicht ohne weiteres als Konstituente in der Nominalphrase des nominalen Bezugselements ‚*beides*‘ zu verorten, zumindest nicht so eindeutig wie ‚*der Kinder*‘ in

(10) *Beides der Kinder trank eine Limonade.*

### 1.3.2 Appositionen und Numerus-Kongruenz

Wegerer (2012) beschäftigte sich in ihrer Dissertation mit den Problembereichen der Numerus-Kongruenz von Subjekt und Prädikat im Deutschen. In diesem Rahmen führte sie eine empirische Studie mittels Online-Fragebögen durch, in der sie muttersprachliche Deutsch-Sprecher\*innen bat, Lückensätze, die in anerkannten Grammatiken als Zweifelsfälle bezüglich der Numerus-Kongruenz eingestuft werden, mit einem Verb<sup>2</sup> im Singular oder im Plural zu vervollständigen. Vor dem Hintergrund der Fachliteratur und der Ergebnisse aus den Fragebögen erörterte sie, bei welchen Konstruktionen welche Entscheidungsprozeduren von Bedeutung sind.

---

<sup>2</sup> Die Infinitiv-Form eines Verbs wurde vorgegeben.

Für meine Studie von Interesse sind Wegerers Ausführungen zu Satzstrukturen vom Typ „Singularisches Subjekt mit Apposition im Plural“, zunächst jene zur Theorie (2012, 91), und anschließend jene zu den Ergebnisse der Befragung (2012, 168). In Fällen wie

(11) *Dieser Film, vor allem die Verfolgungsszenen, ist/sind sehr gut gelungen.*

(12) *Beides, Rasen und Weg, ist/sind mit Laub bedeckt.*

kongruiert das Verb entweder mit dem formalen Subjekt im Singular oder es stimmt mit der Apposition im Plural überein. Bei letzterem wird als Erklärung auf das Proximitätsprinzip verwiesen. Beide Möglichkeiten sind laut Grammatik-Duden (2016, 1027) und Zweifelsfälle-Duden (2011, 559) akzeptabel.

Die Auswertung der Ergebnisse für Satz (11) ergab eine leichte Tendenz zum Singular des finiten Verbs: In 58 % der Antworten wurde ‚ist‘ gewählt, in 42 % ‚sind‘. Der Autorin zufolge könnte ein Grund dafür sein, dass „das grammatische Subjekt durch ein Substantiv, auf das mit dem Demonstrativpronomen ‚dieser‘ in adverbialer Verwendung speziell hingewiesen wird, repräsentiert“ wird (Wegerer, 2012, 168f.). In Satz (12) hingegen „tritt das grammatische Subjekt eindeutig in den Hintergrund“ (Wegerer, 2012, 169): Hier wählte eine Mehrheit der Befragten ein Finitum im Plural (82 %) und deutlich weniger eines im Singular (18 %). Die Autorin kommt zu dem Schluss,

„dass in Sätzen mit einem bedeutungsschwachen Subjekt die Proximitätskongruenz [...] eine große Rolle spielt. Durch die Kongruenz mit dem nächststehenden Ausdruck heben die Sprecher\*innen die semantische Vielheit hervor, was die Singular-Plural-Verteilung bei diesem Beispielsatz bestätigt.“ (Wegerer, 2012, 169)

#### 1.4 Zwischenfazit

Abgesehen davon, wie genau das Indefinitpronomen ‚beides‘ und die Koordinationsphrase hierarchisch zueinander stehen, kann man festhalten, dass es bei der Analyse verschiedene Ebenen der Kongruenz zu berücksichtigen gilt. Das Subjekt ‚beides‘ ist eine Singular-Form, es verweist jedoch auf zwei Entitäten. Der Einschub ist meines Erachtens durch die Konjunktion ‚und‘ eine Plural-Form und referiert, unabhängig von den Formen der Einzel-elemente, auf mehrere Entitäten. Das gesamte präverbale Phrasenmaterial besteht aus mehreren nominalen Elementen, die sich hinsichtlich ihrer Numerus-Merkmale unterscheiden. Es ist unklar, welcher Numerus zum höchsten Knoten dieses Materials projiziert wird. Dadurch hat es im Hinblick auf seine Kongruenzbeziehung zum Verb eine ambige Wirkung.

Syntaktische Konstruktionen mit komplexen, numerusambigen Nominalphrasen sind in formal-syntaktischer Hinsicht die Kongruenz-Relationen betreffend nicht eindeutig einzuordnen. Dies zeigt sich darin, dass bei der Beurteilung solcher Konstruktionen Zweifel bei muttersprachlichen Sprecher\*innen aufkommen, ob ein Singular- oder ein Plural-Verb grammatikalisch korrekt ist.

Die Akzeptabilität der vorliegenden Konstruktionen ist anhand grammatisch-formaler Kategorien nur unzureichend zu beschreiben. Deshalb erscheint es sinnvoll, die ausgeführten Beschreibungen um psycholinguistische Aspekte zu ergänzen.

Die psycholinguistische Perspektive soll Erkenntnisse darüber liefern, wie Sprache als Teil kognitiver Prozesse im menschlichen Gehirn funktioniert. Dazu gehört die Untersuchung von Verarbeitungsprozessen mit Hilfe von Methoden, die Echtzeit-Daten liefern. Das Verstehen der thematisierten Satzstrukturen stellt für das Sprachverarbeitungssystem eine Herausforderung dar, weil mehrere nominale Elemente Subjektmerkmale aufweisen, wegen der sie mit dem finiten Verb eine Kongruenzbeziehung eingehen können.

Im folgenden Kapitel werden zunächst grundlegende Eigenschaften des Sprachanalyse-systems und dessen Arbeitsweise, wenn es mit verschiedene Konstruktionen konfrontiert wird, dargelegt. Anschließend sollen spezifischere Ansätze vorgestellt werden, die sich der Verarbeitung von Kongruenzphänomenen in komplexen Sätzen widmen. Abschließend werden diese Ansätze in Zusammenhang mit den genannten Hypothesen zur Satzstruktur ‚*Beides*‘ + *KoordP* + *Verb* gebracht, bevor im dritten Kapitel auf die empirische Untersuchung eingegangen wird.



## 2 Sprachverstehen

Neben der Sprachproduktion ist die Sprachrezeption die andere wichtige Verarbeitungsrichtung des Sprachverhaltens. Werden bei der Sprachproduktion Informationen im Allgemeinen von der pragmatischen über die semantische und syntaktische zur motorischen Ebene weitergegeben, laufen während der Sprachrezeption Verarbeitungsprozesse in die entgegengesetzte Richtung ab: Sprachliches Material wird zunächst sensorisch wahrgenommen (visuell beim Lesen, auditiv beim Hören), anschließend syntaktisch und semantisch analysiert, bevor es auf der pragmatischen Ebene in den Kontext der Kommunikationssituation eingeordnet wird (vgl. Rickheit et al., 2002, 93).

Der Begriff der Sprachrezeption beinhaltet also explizit auch die sensorische Wahrnehmung von Sprache. Mit dem Begriff des Sprachverstehens dagegen werden eher Prozesse auf syntaktischer, semantischer und pragmatischer Ebene in den Blick genommen. Da diese Ebenen in der vorliegenden Arbeit relevanter als sensorische Verarbeitungsmechanismen sind, wird der Begriff *Sprachverstehen* verwendet.

### 2.1 Grundlagen

#### 2.1.1 Das ‚Wissen‘ des menschlichen Satzanalysesystems

Das Satzanalyseystem des Menschen bezeichnet man als menschlichen *Parser*. Die zwei großen Herausforderungen an den Parser sind zum einen die Komplexität der syntaktischen Beziehungen und zum anderen der Umgang mit syntaktischen Mehrdeutigkeiten. Er muss anhand der syntaktischen Information, die durch die Worterkennung aktiviert ist, Teilen der prosodischen Information und der Wortreihenfolge eine syntaktische Struktur aufbauen (vgl. Dietrich & Gerwien, 2017, 183).

Die syntaktische Struktur besteht aus den Beziehungen zwischen den Wortformen in einer Phrase, der sogenannten Konstituentenstruktur, und den Funktionen der Wortformen in einer Phrase und im gesamten Satz, der funktionalen Gliederung. Für die Analyse der lexikalischen Information greift der Parser auf die im Mentalen Lexikon vorhandenen Wortformen zu, die durch die Worterkennung abrufbar sind. Bei Wortformen, die mehrdeutig sind bzw. mehrere Funktionen übernehmen können, werden alle Alternativen zur Verfügung gestellt (vgl. Dietrich & Gerwien, 2017, 186).

Trotz der Komplexität der Beziehungen und der potentiellen Mehrdeutigkeiten schafft es der Parser erstaunlicherweise, die syntaktische Analyse relativ schnell erfolgreich und oft ohne Irritationen durchzuführen. Worauf kann der Parser dabei zurückgreifen?

Spezifischer formuliert sind die folgenden Fragen Grundlage für die Modellierung eines Parsers, der Satzstrukturen ausschließlich syntaktischen Gesichtspunkten nach analysiert und sich keinen andersartigen Informationen bedient:

- Auf welcher Wissensbasis beruht die Arbeitsweise des Parsers bei der Rekonstruktion der Satzstruktur? Wie sind Konstituentenstrukturen und die funktionale Gliederung gestaltet?
- Wird der Parser bei seinen Analysen von bestimmten Prinzipien geleitet?

### *Konstituentenstruktur*

Die Wortformen in einem Satz stehen in verschiedenen Arten von Beziehungen zu einander. Eine Art sind die sogenannten Präzedenzbeziehungen. Das System dieser Beziehungen regelt, welche lineare Anordnungen von Elementen möglich sind. Das andere wichtige System ist das der Dominanzbeziehungen. Es grenzt ein, nach welchen Kriterien Phrasen gebildet werden können. Diese zwei Systeme aufeinander beziehend gilt das sogenannte Lokalisierungsprinzip, welches besagt, dass Elemente, die zur selben Phrase gehören, nebeneinander stehen.

Sätze lassen sich intuitiv in Einheiten beieinander stehender Wörter zergliedern, die als eher zusammengehörig eingestuft werden als andere. Diese Einheiten entsprechen im allgemeinen den grammatischen Phrasen. Gleichzeitig ist die Konstituentenstruktur eine Anordnung hierarchisch zueinander stehender Elemente. Manche Elemente dominieren andere, welche wiederum andere auf einer niedrigeren Ebene dominieren usw.. Vice versa sind Elemente niedriger Ebenen Konstituenten von direkt über ihnen stehenden Elementen.

Dass der Parser die Regeln bezüglich der Konstituentenstruktur während der Analyse berücksichtigt, konnte experimentell bestätigt werden (s. Dietrich & Gerwien, 2017, 189, für entsprechende Studien).

### *Syntaktische Funktionen*

Innerhalb der Konstituentenstruktur übernehmen Elemente sogenannte syntaktische Funktionen. Das lässt sich am folgenden Satz illustrieren:

(13) *Der Spion, der den Auftrag bekommen hatte, öffnete unbeobachtet die Haustür.*

Syntaktisch beschrieben fungiert ‚*der Spion*‘ als Subjekt und ‚*die Haustür*‘ als direktes Objekt des Hauptsatzes. Nach dem Modell der Konstituentenstruktur kann die Subjektfunktion als die Nominalphrase definiert werden, die sich im Satz am hierarchisch höchsten Knoten befindet. An welcher Stelle in der Abfolge des Satzes die Phrase positioniert ist, spielt bei der Übernahme einer syntaktischen Funktion keine Rolle. In gleicher Weise fungiert diejenige Nominalphrase als direktes Objekt, welche zum verbalen Element hierarchisch die geringste Distanz hat (vgl. Dietrich & Gerwien, 2017, 189).

### *Kopfelemente und Argumente*

Für einen grammatisch wohlgeformten Satzbau bestehen noch weitere Bedingungen/Beschränkungen. Viele Wörter fordern Ergänzungen, Verben bspw. häufig Nominalphrasen; Nomina häufig Präpositionalphrasen etc.. Diese Ergänzungen werden in der Theorie *Argumente* genannt, die Gesamtheit aller Argumente zu einem fordernden Wort *Argumentrahmen*. Das fordernde Wort wird in diesem Kontext als Kopfelement oder *Kopf* bezeichnet. Der Kopf einer Phrase bestimmt auch deren Kategorie (d. h. eine Verbalphrase hat ein Verb als Kopf, eine Nominalphrase ein Nomen etc.) (vgl. Dietrich & Gerwien, 2017, 190).

### *Prinzipien des Phrasenaufbaus*

Das Sprachanalysesystem verlässt sich bei seiner Arbeit auf eine vergleichsweise geringe Anzahl grammatischer Prinzipien, nach denen die vielen möglichen syntaktischen Strukturen der Sprache aufgebaut sind. Für die verschiedenen Phrasenstrukturen gelten folgende drei essentielle Prinzipien:

- Eine Phrase hat nur einen Kopf (*Kopf-Prinzip*).
- Der Kopf trägt das morphologische Merkmal der Phrase. (*Kopf-Vererbungsprinzip*)
- Die hierarchischen Beziehungen bestehen dergestalt, dass die Phrase eines Typs eine komplexe Einheit dieses Typs darstellt, aber nicht anders herum. (*Ebenen-Prinzip*)

(vgl. Dietrich & Gerwien, 2017, 191, insbesondere für eine formale Beschreibung des Ebenen-Prinzips)

Des Weiteren von Bedeutung ist das sogenannte *Theta-Raster*, das besagt, dass in einer Phrase jeder Argumentposition, die ein Verb fordert, eine semantische Rolle entspricht und umgekehrt. Dieses Raster stellt für den Parser eine Analysemaxime dar, der er bei jedem seiner Schritte versucht zu folgen. Ausgehend davon postuliert Pritchett (1992, 138) ein allgemeineres Prinzip, das der Parser während des Verarbeitungsprozesses befolgt: Versuche jedem Prinzip der Syntax zu jeder Zeit so weit wie möglich gerecht zu werden (vgl. Dietrich & Gerwien, 2017, 192-194). Dem folgend werden lokale Analysen durchgeführt, die oftmals aufwendig revidiert werden müssen, und dennoch gelingt es dem Parser im Allgemeinen, schnell und erfolgreich zu arbeiten. Das ist u. a. dem nicht-syntaktischen Wissen zu verdanken. Auf dieses Wissen wird, falls relevant, im weiteren Verlauf der Arbeit an entsprechenden Stellen eingegangen (→ Prosodie, Häufigkeiten des Auftretens von syntaktischen Konstruktionen, Prinzip der Kooperation, Weltwissen).

### 2.1.2 *Modulare Erklärungsansätze*

Modulare Ansätze des Sprachverstehens nehmen an, dass im menschlichen Gehirn verschiedene kognitive Module vorhanden sind, wobei jedes für sich als hochspezialisierte Einheit für einen abgrenzbaren Aufgabenbereich der Sprachverarbeitung zuständig ist. Dieser Vorstellung entsprechend geht man von einzelnen funktionalen Ebenen aus, zwischen denen, allerdings nur monodirektional, Informationseinheiten als jeweils abgeschlossenes Output-„Paket“ weitergegeben werden. Durch die sogenannte Informationsabkapselung ist es nur direkt miteinander verbundenen Modulen möglich, Information voneinander zu verarbeiten; „weiter entfernte“ können keinen Einfluss mehr auf Repräsentationen auf niedrigeren Ebenen ausüben. Das bedeutet, dass die syntaktische Analyse beispielsweise von der prosodischen Materials unberührt bleibt.

Um der Frage nachzugehen, bei welchen syntaktischen Strukturen und ab wann die syntaktische Analyse möglicherweise doch durch die semantische ergänzt wird, werden Sätze mit strukturellen Mehrdeutigkeiten, sogenannten *Ambiguitäten*, untersucht. Denn hier ist der Parser gefragt, Strategien zur Auflösung dieser Mehrdeutigkeiten zu kreieren. Wie tut er das? Wird hier womöglich auf die semantische Analyse zugegriffen?

Es wird zwischen zwei Arten von Ambiguitäten unterschieden (Drenhaus, 2012, 97; Dietrich & Gerwien, 2017, 185): *Globale* oder *endgültige* Ambiguitäten sind solche, die bis über das Satzende hinaus bestehen bleiben; *Lokale* oder *temporäre* Ambiguitäten werden

im weiteren Verlauf des Satzes durch eine Information aufgelöst. An den zwei folgenden Beispielsätzen lassen sich die zwei Arten von Ambiguitäten nachvollziehen:

(14) *Der Junge berührte den Hut mit der Feder.*

Für Satz (14) kommen über sein Ende hinaus zwei verschiedene syntaktische Analysen in Frage. Es sind zwei Lesarten möglich; der Satz ist somit global ambig. 1. Analysevariante: Die Präpositionalphrase (*,mit der Feder‘*) wird mit dem Objekt (*,Hut‘*) verbunden; in der Vorstellung des Rezipienten entsteht in diesem Fall das Bild eines Hutes, an dem eine Feder befestigt ist; dieser Hut wird von einem Jungen berührt. 2. Analysevariante: Die Präpositionalphrase wird als Teil der Verbalphrase analysiert; *,Mit der Feder‘* fungiert instrumental; es entsteht also ein Bild eines Jungen, der eine Feder in seiner Hand hält, und damit in Kontakt mit einem Hut kommt. Der Satz ist für den Parser ohne andere, nicht-syntaktische Information nicht zu verarbeiten. Beim folgenden Satz verhält es sich dagegen anders:

(15) *Das Mädchen betrachtete den Mann mit dem Regenschirm.*

In Satz (15) entsteht eine Ambiguität, die mit dem Wort *,Regenschirm‘* aufgelöst wird. Die Semantik von *,Regenschirm‘* und die des Verbs *,betrachtete‘* lassen sich nicht zusammenbringen (weil man i. d. R. nicht mit Hilfe eines Schirms etwas betrachtet), weshalb der Parser die Präpositionalphrase *,mit dem Regenschirm‘* nur als Attribut des Objekts *,Mann‘* analysieren kann; es bleibt also nur eine Lesart des Satzes.

Ein sehr einflussreicher modularer Ansatz, der sich damit befasst, wie syntaktische Ambiguitäten aufgelöst werden, stellt das *Garden-Path-Modell* von Lyn Frazier (1987) dar. Innerhalb dieses Modells wird angenommen, dass der Parser zunächst die Analyse nach syntaktischen Regeln durchführt und jedes Wort in die aufgebaute Struktur integriert. Ist diese abgeschlossen, wird die Struktur semantischen Gesichtspunkten nach interpretiert. Das Prinzip, zunächst eine Struktur nur aus syntaktischer Information zu generieren, wird als *First-Pass-Parsing* bezeichnet (vgl. Rickheit et al., 2002, 96f.; Drenhaus, 2012, 98). Dieses Analysemodell arbeitet *deterministisch* (Schritt für Schritt absolute Entscheidungen gegen/für eine Analyse treffend), *seriell* (verschiedene Informationstypen werden pro Schritt einzeln und nacheinander berücksichtigt) und schwach semantisch eingeschränkt (vgl. Dietrich & Gerwien, 2017, 200).

Stellt sich jedoch im Laufe des Satzes aufgrund einer neuen Information heraus, dass die bisherige Analyse falsch war, muss eine Reanalyse gestartet werden. Bildlich ausgedrückt – der Parser ist mit seiner Analyse in eine Sackgasse oder auf einen Holzweg (engl. *garden*

*path*) geraten. Der *Garden-Path-Effekt*, der durch den Aufwand der Reanalyse entsteht, kann z. B. in Selbstlese-Experimenten erzeugt werden: Er tritt in Form von längeren Lesezeiten an Stellen (den sogenannten *Disambiguierungspunkten*) auf, an denen Mehrdeutigkeiten aufgelöst werden.

*Prinzipien des Garden-Path-Modells* (vgl. Drenhaus, 2012, 99f.)

#### *Minimal Attachment*

Nach dem Minimal-Attachment-Prinzip muss der Parser das kommende Wort in die vorhandene Struktur integrieren, so dass die einfachste Struktur entsteht.

Im folgenden Beispielsatz besteht beim Verb ‚*read*‘ eine lokale Ambiguität (a.). Es kann entweder ein Partizip Perfekt eines reduzierten Relativsatzes oder ein Hauptverb im Simple Past sein.

- (16)      a. *The author read ...*  
            b. *The author read by the student was ...*  
            c. *The author read a newspaper*

Nach dem Minimal-Attachment-Prinzip muss der Parser annehmen, dass der Satz wie c. fortgesetzt wird, weil die Struktur in c. einfacher ist als jene in b. (d. h. in c. müssen weniger Verzweigungen erstellt werden als in b., um neue Konstituenten einzufügen (Dietrich & Gerwien, 2017, 200)). Wird der Satz wie in b. fortgesetzt, wird der Parser von dieser Struktur „überrascht“, und am Disambiguierungspunkt ‚*by the student*‘ kommt es zu einer Verlangsamung der Lese-Geschwindigkeit (Frazier & Rayner, 1982).

#### *Late Closure*

Wie geht der Parser vor, wenn in einem Satz eine lokale Ambiguität zwei gleichermaßen komplexe Strukturen möglich macht? In so einem Fall wird das Late-Closure-Prinzip angewendet, welches besagt, dass eintreffendes Material in die Phrase integriert wird, die gerade verarbeitet wird, und nicht in eine bereits abgeschlossene. Das bedeutet, dass eine aktuell bearbeitete Phrase so lange wie möglich offen gelassen wird, um neue Elemente integrieren zu können. Einmal abgeschlossene Phrasen können während des weiteren Verlaufs des First-Pass-Parsings nicht mehr überarbeitet werden. Im folgenden Satz (17) wird nach dem Late-Closure-Prinzip der Relativsatz eher an der hierarchisch niedrigeren (*das Stilleben*) als an der höheren NP (*das Gemälde*) platziert.

(17) *Das Gemälde mit dem Stillleben, das sich auf dem Tisch befand,  
wurde gestohlen.*

Man spricht in solchen Fällen entsprechend von der Präferenz des *Low Attachments* gegenüber des *High Attachments*.

Im Garden-Path-Modell wird eine zweischrittiges Analysesystem (*two-step-analysis*) angenommen. Zunächst werden Wörter zu Phrasen verbunden; die Phrasen werden jedoch erstmal nicht aneinander gekoppelt. Anschließend werden unter Einfluss von extra-syntaktischem Wissen die Phrasen so arrangiert, dass ein Satz entsteht (vgl. Dietrich & Gerwien, 2017, 201).

### 2.1.3 Nicht-modulare Erklärungsansätze

Nicht-modulare Ansätze des Sprachverstehens postulieren, dass *parallel* zur syntaktischen, eventuell sogar vorher, auch eine semantische Analyse durchgeführt wird. Der Parser integriert also bei jedem neuen Wort syntaktisches als auch semantisches Wissen. Die an der Verarbeitung beteiligten Arbeitseinheiten können untereinander Information austauschen, bezüglich der Weitergaberichtung herrschen nicht so strenge Regeln wie zwischen Modulen, eine Verkapselung der Information findet nicht statt. Dies lässt zu, dass jedes neue Wort inkrementell verarbeitet wird<sup>3</sup> und lokale Ambiguitäten oft aufgelöst werden können. Vor dem Hintergrund der Annahme einer relativ großen Durchlässigkeit bezüglich extra-syntaktischer, einschränkender Faktoren werden diese Modelle auch als *interaktiv* oder *constraint-based* bezeichnet (Rickheit et al., 2002, 97). Ein Parser, der im Vergleich zu einer modularen Verarbeitungsweise insgesamt ‚offener‘ funktioniert, arbeitet *probabilistisch* (Entscheidung für eine Analysealternative nach Wahrscheinlichkeit ihres Auftretens), durch Informationen lexikalischer, kontextueller etc. Art *beschränkt, interaktiv* und *parallel*.

Die syntaktische Analyse beeinflussende Einschränkungen (*Constraints*) hängen bspw. mit Folgendem zusammen:

- Semantische Rollen: Es ist wahrscheinlicher, dass als erstes Nomen im Satz ein Agens auftritt als z. B. ein Patiens; Verben weisen Nomina eher die eine als die andere semantische Rolle zu (*Thematic Bias*, Tanenhaus et al. (2000)).

---

<sup>3</sup> *Inkrementell* bedeutet, dass die Verarbeitung schrittweise erfolgt, Wort für Wort.

- das Wahrscheinlichkeitsverhältnis zwischen Hauptsatz-Alternative und Nicht-Hauptsatz-Alternative
- im Englischen: Ist die Verbform ‚*carried*‘ als Partizip oder als Past-Form häufiger?

Alle Constraints werden bei jedem Schritt der Analyse neu berücksichtigt und tragen ggf. jedes Mal auf andere Weise zu den Entscheidungen des Parsers bei. Je schwieriger die Entscheidung zu fällen ist (also je ähnlicher die Wahrscheinlichkeiten der Alternativen), desto länger dauert der Verarbeitungsprozess an (vgl. Dietrich & Gerwien, 2017, 201f.).

Nachdem die Grundlagen des Sprachverstehens dargelegt wurden, wird im folgenden Teil darauf eingegangen, wie das Arbeitsgedächtnis und wie Erwartungen die Verarbeitung von Kongruenzbeziehungen beeinflussen. Anschließend wird das Phänomen Agreement Attraction erörtert und der Erklärungsansatz des Good-enough-Parsing vorgestellt. Zum Abschluss von Kapitel 2 werden unter Berücksichtigung des bis dahin Ausgeführten Überlegungen zur Verarbeitung der Struktur ‚*beides*‘ + *KoordP* + *VP* angestellt.

## 2.2 Die Verarbeitung von Kongruenz

### 2.2.1 Die Rolle des Arbeitsgedächtnisses

Neben modularen und interaktiven / constraint-based Erklärungsansätzen zum Sprachverstehen gibt es solche, die einen starken Fokus auf die Rolle des Speichersystems während der Sprachverarbeitung legen. Essentielle Fragen dabei lauten (Drenhaus, 2012, 106):

1. Welche Rolle spielt die Kapazität des Arbeitsgedächtnisses bei der Verarbeitung von sprachlichem Input?
2. Besteht das Arbeitsgedächtnis aus einem Gebilde, das als domänenübergreifende Komponente arbeitet? Oder existieren in diesem Gebilde ausdifferenzierte Subkomponenten?

Einen wichtigen Beitrag zur Charakterisierung von Sprachverstehensprozessen lieferten Just und Carpenter (1992) mit ihrer These, dass alle linguistischen Prozesse die gleichen Speicherressourcen beanspruchen. Sie gehen von dem folgenden Zusammenhang aus: Eine Überschreitung der Speicherkapazität führt zu einer Verlangsamung der Verarbeitung bzw. einem Zerfall der Repräsentationen im Speicher; dies spiegelt sich wiederum in Problemen bzw. Fehlern bei der Verarbeitung wieder. Basierend auf den Ergebnissen ihrer Lesezeitstudie kamen Just und Carpenter zu dem Schluss, dass das Verstehen von Sätzen mit der Ka-



pazität des Arbeitsgedächtnisses zusammenhängt. Sie konnten zeigen, dass Versuchspersonen, denen zuvor eine höhere Kapazität als anderen zugeschrieben wurde, während des First-Pass-Parsings nicht nur syntaktisches Wissen nutzten, sondern auch auf die semantische Kategorie Belebtheit zugreifen konnten (für eine genauere Darstellung der Studie s. Drenhaus, 2012, 106f.).

Ein Faktor, der sich negativ auf die Leistung des Arbeitsgedächtnisses auswirkt, stellt die zunehmende Distanz zwischen einem Argument und dessen Kopf dar (Nicenboim et al., 2015, 2). Belege dafür finden sich in Lese-Studien, die berichten, dass größere Dependenz-Längen sogenannte *Locality-Effekte* hervorrufen, d. h. die Lesegeschwindigkeit an der Stelle der Dependenz-Resolution abnimmt bzw. die Anzahl regressiver Sakkaden steigt (z. B. Self-paced-reading- (Gibson, 2000) bzw. Eyetracking-Studien (Demberg & Keller, 2008)).

Als Ansätze, die Verarbeitungsschwierigkeiten bei zunehmender Argument-Kopf-Distanz in Zusammenhang mit der Leistung des Arbeitsgedächtnisses sehen, sollen im Folgenden zunächst die *Dependency Locality Theory* und anschließend das *Activation-Based Model* erläutert werden.

#### *Dependency Locality Theory (Gibson, 2000):*

Eine Theorie des Sprachverstehens, so Gibson (2000, 95), muss berücksichtigen, wie der eintreffende, aus verschiedenen Informationsquellen stammende Input in Echtzeit zu einer Satzstruktur zusammengesetzt wird (Parsing). In diesem Prozess wirkt beschränkend, in welchem Maße über verarbeitende Ressourcen verfügt werden kann. Zwei wichtige Komponenten des Parsings beanspruchen solche Ressourcen:

1. *Die strukturelle Integration:* Neue Wörter müssen in die bereits aufgebaute Struktur integriert werden. Der kognitive Aufwand des Integrierens beruht auf der Distanz zwischen Argument und Kopf. Die Distanz hängt von der Anzahl der sich dazwischen befindlichen Diskursreferenten ab.

2. *Das Im-Gedächtnis-Behalten der Struktur (Storage),* was beinhaltet, offene Abhängigkeiten weiterzuverfolgen. Der kognitive Aufwand hierfür hängt davon ab, wie viele Phrasenköpfe noch benötigt werden, um aus dem aktuell vorhandenen sprachlichen Material bzw. der bereits aufgebauten Struktur einen grammatikalisch korrekten Satz zu generieren.

Der kognitive Aufwand scheint unberührt davon zu bleiben, wie lange (bezogen auf die Zeit) eine offene Dependenz im Gedächtnis behalten werden muss (Gibson et al., 2005). Beides, *Integration* und *Storage*, greift auf den gleichen Ressourcen-Pool zu (Just & Carpenter, 1992). Die Kapazität dieses Pools ist unveränderlich und begrenzt. Jeder neu dazu kommende Phrasenkopf beansprucht eine feste Menge an Ressourcen. Je mehr Ressourcen für das Arbeitsgedächtnis benötigt werden, desto langsamer kann das neue Material integriert werden (Gibson, 2000, 115).

*Activation-Based Model (Lewis & Vasishth, 2005):*

Während die Dependency Locality Theory die Verarbeitungsprozesse beim Verstehen eines Satzes mit linguistischen Mitteln beschreiben zu versucht, nimmt das *Activation-Based Model* eine allgemeinere kognitive Perspektive ein (Nicenboim et al., 2015, 2). Diesem Modell (Lewis & Vasishth, 2005) zufolge sind linguistische Elemente im Arbeitsgedächtnis als Merkmalbündel repräsentiert, deren Stabilität über den Zeitverlauf negativ beeinflusst wird, zum einen dadurch, dass sie in ihrer Repräsentation wieder zerfallen, und zum anderen dadurch, dass die Eigenschaften anderer Elemente störend auf sie wirken. Den Locality-Effekt erklärt dieses Modell damit, dass der Wiederabruf der nicht-lokalen Argumente nicht funktioniert, wobei der Wiederabruf den Reizen (*Cues*) folgt, die zum Zeitpunkt der Dependenz-Resolution einsetzen. Um auf ein Argument zugreifen zu können, müssen die Eigenschaften der Wiederabruf-Reize mit den potentiellen Elementen im Arbeitsgedächtnis abgeglichen werden. Dieser Zugriff wird erschwert,

- je mehr Zeit seit der Encodierung des Arguments vergeht (*Zerfall der Repräsentation*) und,
- wenn Elemente mit ähnlichen Eigenschaften auftreten und als Distraktoren wirken (*Similarity-Based Interference*).

Im Activation-Based Model wird der Faktor *Storage* nicht explizit berücksichtigt, trotzdem verursachen im Arbeitsgedächtnis befindliche Elemente Interferenz-Effekte.

Nach beiden Ansätzen, sowohl der Dependency Locality Theory als auch dem Activation-Based Models, wird in den unten stehenden Sätzen (18 a-c) der Wiederabruf des Arguments mit zunehmender Distanz zwischen dem nicht-lokalen Element (*,who'*) und dem Verb (*,supervised'*) schwieriger (Nicenboim et al., 2015, 2). Dies wird von Grodner und

Gibson (2005) bestätigt, die in einer Studie von Locality-Effekten in (Objekt-externen) Relativsätzen folgender Art berichten:

- (18)      a. *The administrator who the nurse supervised ...*  
            b. *The administrator who the nurse from the clinic supervised ...*  
            c. *The administrator who the nurse who was from the clinic supervised ...*

Im Gegensatz dazu gibt es jedoch auch immer mehr Belege für den *Anti-Locality-Effekt*: Bei SOV-Strukturen (z. B. im Deutschen, Levy & Keller, 2013) konnte gezeigt werden, dass sich mit zunehmender Distanz an der Stelle, an der die Dependenz vervollständigt wird, die Lesegeschwindigkeit erhöht. In vielen Fällen kann dies anhand des Activation-Based Models erklärt werden, da die Elemente förderlich auf die Stabilität des kommenden Phrasenkopfes wirken, in dem sie ihn durch Modifikation aktivieren (Vasishth & Lewis, 2006). Nach Theorien dieser Art hat das zur Folge, dass die Verarbeitung des Phrasenkopfes erleichtert wird, weil er bereits generiert worden ist; anders ausgedrückt – die Verbalphrase ist *prä-aktiviert*. Das ist v. a. für SOV-Sprachen von Relevanz, in denen die Argumente der Verbalphrase vorausgehen. Sie wird also erst modifiziert, bevor anschließend ihr Kopf verarbeitet wird. Für solche Fälle würde das Activation-Based Model voraussagen: Je länger die Distanz, desto schneller die Verarbeitung (wobei der Zerfall der Repräsentation den Nutzen der Aktivierung nicht ausgleichen darf) (vgl. Nicenboim et al., 2015, 2).

### 2.2.2 Die Rolle von Erwartungen

An der Verarbeitung beim Sprachverstehen sind nicht nur Bottom-up-Prozesse, also der Aufbau der syntaktischen Struktur aus dem eintreffenden Input, sondern auch Top-down-Prozesse maßgeblich beteiligt. Ein solcher Top-down-Prozess besteht in Erwartungen bezüglich des kommenden Inputs. Nicenboim et al. (2015, 3) nennen zwei Gründe, durch die in einem syntaktisch eingegrenzten Kontext die Vorhersage über das Auftreten eines Wortes erleichtert werden kann:

- durch lokale syntaktische Beschränkungen, die mit den Eigenschaften des Verbs zusammenhängen, oder
- durch die Möglichkeit bzw. die Fähigkeit des Parsers, unter der Berücksichtigung grammatischer und wahrscheinlichkeitsbasierter Information top-down eine Struktur aufzubauen.

Levy (2008) folgt in seiner Erwartungsbezogenen Theorie (*Expectation-Based Theory*) diesen Annahmen und benennt als Hauptgrund für Schwierigkeiten bei der Verarbeitung den Informationsgehalt oder Überraschungswert eines Wortes (*Surprisal*, mathematisch ausgedrückt der negative Logarithmus der bedingten Wahrscheinlichkeit), den ein Wort hat, bevor es in einen bestehenden Kontext integriert werden muss. Dahinter verbirgt sich die Vorstellung, dass über lexikalisches Material, das eher überraschend auftritt, auch schwerer Vorhersagen zu treffen sind (vgl. Nicenboim et al., 2015, 3).

Im Falle einer Long-Dependency-Resolution geht man davon aus, dass bei Leser\*innen bzw. Hörer\*innen eine Erwartung eines Verbs (samt seines Subkategorisierungsrahmens) entsteht, das irgendwann kommen *muss* (die Frage ist wann). Jede Konstituente eines Typs, die auf eine des selben Typs folgt, verringert zum einen die Erwartung, dass nochmal eine solche auftritt; zum anderen erhöht sich gleichzeitig die Erwartung, dass eine Konstituente anderen Typs (z. B. ein finites Verb) auftritt. Basierend auf der erhöhten Erwartungshaltung, dass als nächstes Element ein Verb kommt, wird dessen Verarbeitung dann erleichtert. Somit kann angenommen werden, dass in einer endlichen Satzstruktur sich mit jeder Konstituente, die kein finites Verb ist, die Wahrscheinlichkeit erhöht, dass das nächste eintreffende Element ein Verb sein wird (vgl. Nicenboim et al., 2015, 3).

Im folgenden eine Übersicht über die Auswirkung von Elementen, die zwischen dem Subjekt und dem Verb stehen (Vasishth, 2017, 50):

Der Effekt, der durch diese Zwischenelemente entsteht, scheint davon abzuhängen,

- welche Wiederabrufprozesse nach den Zwischenelementen angestoßen werden;
- welche sprachlichen Wiederabruf-Reize während des Wiederabrufs eingesetzt werden (z. B. Reize morphosyntaktischer Art);
- welche individuellen Unterschiede bezüglich der Kapazität des Arbeitsgedächtnisses und der kognitiven Kontrolle bestehen;
- ob eine systematische Unterspezifizierung vorliegt, möglicherweise mit der Folge, dass das Subjekt Abhängigkeiten nicht vollständig aufbaut;
- was vorausgesagt wird (kann um die individuelle Fähigkeit variieren, Vorhersagen aufrechtzuerhalten);
- welche Art von Information in den Zwischenelementen enthalten ist (kann Vorhersagen präziser oder gröber machen).

### 2.2.3 Agreement Attraction

Kongruenz-Phänomene, die Beziehungen zwischen finitem Hauptverb und anderen Elementen als dem formalen Subjekt betreffen, sind sowohl in der formal-syntaktischen Linguistik als auch in der Psycholinguistik ein oft untersuchter Forschungsgegenstand. Eine einflussreiche Veröffentlichung im Bereich der Psycholinguistik stellt jene von Kathryn Bock und Carol A. Miller dar (*Broken Agreement*, 1991). Der unkomplizierten Regel ‚Ein singularisches Subjekt fordert ein Verb im Singular, ein pluralisches Subjekt fordert ein Verb im Plural‘ wird offensichtlich nicht immer Rechnung getragen. In ihrer experimentellen Sprachproduktionsstudie berichten die Autorinnen von sogenannten *Attraction*-Fehlern, bei denen das Hauptverb fälschlicherweise mit einem zwischen dem Kopfnomen der Subjektphrase und dem Verb liegenden Element im Numerus kongruiert. Wenn das Kopfnomen im Singular steht und das lokale, unmittelbar präverbale Nomen im Plural, zeigte sich bei den Versuchspersonen eine robuste Tendenz, ein Verb mit Pluralmarkierung zu bilden, obwohl die Subjektphrase eindeutig im Singular steht und so ein Singular-Verb fordert (Bock & Miller, 1991; Franck et al., 2002).

(19) *The time for fun and games \*are over. (time fordert is)*

(20) *The readiness of our conventional forces \*are at an all-time low.  
(readiness fordert is)*

(21) *The key to the cabinets \*were on the table. (key fordert was)*

(aus dem gesprochenen Englisch; Bock & Miller, 1991, 46)

Sätze wie diese wurden in der Linguistik häufig thematisiert und sind sowohl in alltäglichen Gesprächen als auch in redaktionell überarbeiteten Texten zu finden. Sprachproduktionsstudien liefern eine Menge gut belegter Faktoren, die die statistische Wahrscheinlichkeit beeinflussen, dass formal-syntaktische Kongruenzregeln sich nicht durchsetzen und stattdessen ein lokales, dem Verb näher stehendes Element als sogenannter *Attractor* fungiert. Faktoren dabei können folgende sein (vgl. Wagers et al., 2009, 2017, auch für Studien dahingehend):

- die Numerus-Eigenschaften der möglichen Attractor-Elemente,
- die strukturelle Tiefe der Attractor-Elemente im Bezug zum Regens, und
- die Wortreihenfolge.

Beschreibungen von *Agreement Attraction* bei der Sprachproduktion teilen die Auffassung, dass Nomina bzw. Nominalphrasen in einer komplexen Subjektphrase ihre eigenen spezifischen Numerus-Eigenschaften besitzen, die nicht miteinander verbunden sind (z. B. Eberhard, 1997; Franck et al., 2002). Die Beschreibungen unterscheiden sich jedoch hinsichtlich der Ursache dafür, dass diese Eigenschaften Einfluss auf den Numerus des Verbs nehmen, und so ein grammatikalisch inkorrekt gebildet wird. Wagers et al. (2009, 207) nennen zwei wichtige Erklärungsansätze:

1. Aufgrund hierarchisch unklar strukturierter Repräsentationen breiten sich Merkmale fälschlicherweise in der Nominalphrase aus (*Feature Percolation*).
2. Während die Verbform geplant wird geraten die Merkmale gleichzeitig aktivierter Konstituenten durcheinander.

Der Hauptunterschied zwischen den beiden Erklärungsansätzen betrifft die Repräsentation des Subjekts selbst. Der erste Ansatz führt fehlerhafte Numerus-Markierungen am Verb darauf zurück, dass die Repräsentation des Subjekts in sich selbst inkonsistent oder fehlerhaft ist, was an den Eigenschaften der kombinatorischen Schritte liegt, die die Repräsentation bilden und aufrechterhalten (*Repräsentationaler Ansatz*). Dem zweiten Ansatz zufolge gibt es keinerlei Probleme beim Aufrechterhalten der Repräsentation oder in ihrer Struktur selbst. Unstimmigkeiten entstehen alleine dadurch, dass die Repräsentationen der einzelnen Konstituenten während des Verarbeitungsprozesses verändert oder abgerufen werden (*Cue-Based-Retrieval-Ansatz*, orientiert am *Activation-Based Model*, Lewis & Vasishth, 2005).

Ergänzt man die oben erwähnten Phänomene aus Studien zur Sprachproduktion um die Ergebnisse aus solchen zum Sprachverstehen, zeigt sich ein konvergentes Bild: Werden in bestimmten syntaktischen Kontexten bei Elizitationsaufgaben häufiger Kongruenz-Fehler produziert, haben Versuchspersonen beim Verstehen entsprechender Konstruktionen weniger Schwierigkeiten bei der Verarbeitung (z. B. Pearlmutter et al., 1999).

Repräsentationale Ansätze können *Agreement Attraction* bei der Sprachproduktion verlässlich vorhersagen: Eine Satzstruktur mit einem lokalen Attractor-Nomen wie ‚*The key to the cabinets ...*‘ sollten demnach Versuchspersonen relativ häufig fälschlicherweise mit einem Verb im Plural fortsetzen – das bestätigte sich. Für Verstehens-Kontexte würden solche Ansätze voraussagen: Grammatisch korrekte Sätze mit lokalem Attractor-Nomen wie ‚*The key to the cabinets is ...*‘ sollten relativ häufig als inkorrekt wahrgenommen werden – dies scheint jedoch nicht der Fall zu sein (vgl. Wagers et al., 2009; Tanner et al., 2014). Diese

grammatische Asymmetrie lässt darauf schließen, dass die Mechanismen, die sich hinter dem Phänomen Agreement Attraction verstecken, sich für Produktion und Verstehen zumindest teilweise unterscheiden (vgl. Schlueter et al., 2018, 75).

Agreement Attraction in Sprachverstehenskontexten liegt also vor, wenn in einem Satz zwischen Subjekt und Verb bezüglich Numerus keine Kongruenz besteht, aber zwischen dem Verb und einem anderen Element, und dieser Satz als grammatikalisch korrekt wahrgenommen wird. In anderen Worten – wird ein Satz wie

(21)       ,The key to the cabinets are rusty‘,

in dem der Numerus ‚Plural‘ des Nomens ‚cabinets‘ strukturell betrachtet keinen Einfluss auf die gesamte komplexe Subjekt-NP hat, verarbeitet, ist die Wahrscheinlichkeit geringer, dass er als grammatikalisch inkorrekt eingestuft wird als in einem Satz ohne ein solches Pluralnomen. Dass dieses Nomen als Attractor wirkt, kann mit dem *Cue-Based-Retrieval-Ansatz* erklärt werden (Wagers et al., 2009; Tanner et al., 2014, Schlueter et al., 2018).

Um der Abhängigkeitsregel ‚Singularisches Subjekt fordert ein Verb im Singular, pluralisches Subjekt fordert ein Verb im Plural‘ gerecht zu werden, muss das Subjekt wieder aus dem Arbeitsgedächtnis abgerufen werden. Nach dem Cue-Based-Retrieval-Ansatz stellt in Satz (21) das Verb einen Numerus-Cue [Plural] und einen strukturellen Cue [Subjekt] bereit. Wenn eines der (nominalen) Elemente im Arbeitsgedächtnis Eigenschaften besitzt, die mit beiden Cues übereinstimmen, ist wahrscheinlich, dass es abgerufen wird. Natürlich ist es auch möglich, dass ein Element, das nicht das Subjekt ist, vorhanden ist und durch ein Numerus-Cue-Match eine verstärkte Aktivierung erfährt. In einem grammatisch wohlgeformten Satz stellen die Eigenschaften des Subjekts ein optimales Ziel für die beiden Wiederabrufer-Cues des Verbs dar, weshalb das Subjekt auch ohne Schwierigkeiten abgerufen wird. In grammatisch inkorrekten Sätzen hingegen, in denen Subjekt und Verb nicht kongruieren (wie in Satz (21)), kann fälschlicherweise ein Attractor-Nomen abgerufen werden. Weil es aufgrund der Ähnlichkeit der Merkmale zu einer Störung kommt, ist in diesem Zusammenhang von *Similarity-Based Interference* die Rede. Das Subjekt erfährt keine verstärkte Aktivierung durch den Numerus-Cue, sondern ausschließlich durch den strukturellen Cue. Das Attractor-Nomen hingegen wird durch den Numerus-Cue verstärkt aktiviert. Dies ermöglicht in manchen Fällen einen fälschlichen Abruf des Attractors anstatt des Subjekts. Dieser Mechanismus stellt eine Strategie dar, wie das Sprachverarbeitungssystem den Schwierigkeiten begegnet, die auftreten, wenn Kongruenzregeln missachtet werden (vgl. Schlueter et al., 2018, 75f.).

#### 2.2.4 *Good-enough-Parsing*

Viele theoretische Ansätze, die sich mit dem Sprachverstehen in Diskursen beschäftigen, gehen davon aus, dass in einem Diskurs Verstehensprozesse auf lexikalischer, syntaktischer, semantischer und der Diskursebene vollständig abgeschlossen werden. Das bedeutet, dass am Ende der Prozesse keinerlei ambige Interpretationen von Sätzen mehr bestehen. In vielen Situationen – beim Lesen, Hören oder in Gesprächen – werden Sprachverarbeitungsprozesse jedoch nicht abgeschlossen; die Satzanalyse bleibt unvollständig (vgl. Sanford & Graesser, 2006, 99f.).

Eine nicht unerhebliche Anzahl sprachlicher Fehlinterpretationen und missverständlicher Kommunikationsprozesse sind nicht zufälligen Fehlern des menschlichen Sprachverarbeitungssystems geschuldet, sondern sind darauf zurückzuführen, dass der Parser bestimmte Verarbeitungsweisen gegenüber anderen bevorzugt. Er präferiert schnelle und einfache, weniger aufwendige Prozesse, deren Output den Anforderungen einer funktionierenden Kommunikation in den meisten Fällen genügt, und nimmt dafür Fehler in Kauf. Verarbeitungsstrategien, die dieses Prinzip verfolgen, werden in der psycholinguistischen Forschungsliteratur mit den Bezeichnungen *Good-enough*, *Underspecified* oder *Shallow<sup>4</sup> Parsing* beschrieben.

Es ist bemerkenswert, dass der menschliche Parser bei der ständig auf ihn einprasselnden Menge sprachlichen Inputs diesen größtenteils so erfolgreich verarbeitet wie er es tut. Man könnte also annehmen, dass hinter Fehlinterpretationen keine Systematik steckt. Obwohl Fehlinterpretationen gerade bei herausfordernden Aufgabenstellungen häufig zu beobachten sind, wurden bis vor nicht allzu langer Zeit solche Phänomene psycholinguistisch eher selten untersucht. Es ist davon auszugehen, dass sie bisher eher auf Unaufmerksamkeit oder Störfaktoren aus der Umwelt (wie z. B. Geräusche) zurückgeführt wurden und nicht auf Sprachverarbeitungsprozesse, die auf systematisch angelegten Mechanismen beruhen (vgl. Christianson, 2016, 817).

Traditionell wurde sich in der Parsing-Forschung häufig auf temporäre Mehrdeutigkeiten in Garden-Path-Sätzen bezogen. Für Sätze dieser Art wurde und wird häufig immer noch angenommen, dass ihre ambigen syntaktischen Strukturen vom Parser vollständig aufgelöst werden müssen, um überhaupt als ‚verstanden‘ bezeichnet werden zu können. Chris-

---

<sup>4</sup> *shallow* (engl.) – oberflächlich



tianson (2016, 818) spricht in diesem Zusammenhang von einem Alles-oder-nichts-Blick auf Sprachverstehensprozesse. Dieser Blick spiegelt sich in der Methodenwahl bei Sprachverstehensstudien wieder, in denen oft Grammatikalitätsurteile anhand von zwei Auswahloptionen (grammatisch/ungrammatisch) erhoben und auf dieser Grundlage Rückschlüsse auf den Verarbeitungsprozess gezogen werden. Als Beispiel dafür geht er auf eine Studie von Ferreira und Henderson (1991) ein, in der sie über den sogenannten ‚*Head position effect*‘ in Garden-Path-Sätzen berichten. Der Effekt nimmt eine einfachere Verarbeitung für (23a.) als für (23b.) an:

- (23)        a. *After the Martians invaded the town was evacuated.*  
              b. *After the Martians invaded the town the city bordered was evacuated.*

Als Grund dafür wird angegeben, dass ‚*the town*‘ die semantische Rolle ‚*Patiens*‘ in (b.) länger zugeschrieben bleibt als in (a.), weil sich in (b.) mehr Elemente zwischen ‚*the town*‘ und dem disambiguierenden Verb ‚*was*‘ befinden als in (a.). Weiterhin gehen sie davon aus, dass das zwischen den kritischen Elementen befindliche Material dafür sorgt, dass es nach der falschen Analyse dem Parser nicht mehr möglich ist, eine weitere durchzuführen. Wie oben bereits erwähnt wurden stufenlos erhobene Grammatikalitätsurteile als Anzeichen dafür interpretiert, ob dem Parser in den temporär ambigen Garden-Path-Sätzen eine vollständige Korrektur (*Recovery*) der falschen syntaktischen Analyse gelingt (in den als korrekt beurteilten Sätzen) oder nicht (in den als inkorrekt beurteilten).

In dieser wie auch in weiteren Studien (z. B. Frazier & Rayner (1982), Eyetracking-Lesestudie, Satzstrukturen wie in (23 a.&b.), regressive Blicke als Beleg für *Recovery* interpretiert), die die Vorstellung einer kompletten Korrektur stützen, wurden die ersten, falschen Analysen des Parsers bzw. daraus entstehende Interpretationen größtenteils vernachlässigt. *Recovery* bedeutet nicht zwangsläufig, dass die Struktur aus der ersten Analyse vollständig ‚gelöscht‘ wird und gar keinen Einfluss auf den Reanalyseprozess bzw. die schlussendliche Interpretation hat (vgl. Christianson, 2016, 818).

Vor dem Hintergrund dieser Kritik führten Christianson et al. (2001) Experimente durch, die zum Ziel hatten, die Beziehung zwischen ursprünglicher Analyse und Reanalyse zu untersuchen. Versuchspersonen hatten die Aufgabe, Sätze wie die folgenden zu lesen und darauf abgestimmte Fragen zu beantworten:

- (24)        *While the man hunted the deer that was brown and graceful ran into the woods.*  
              - *Did the man hunt the deer?*

- (25) *While Anna dressed the baby that was cute and cuddly played in the crib.*  
- *Did Anna dress the baby?*

Es zeigte sich, dass die Versuchspersonen die Fragen sehr oft inkorrektweise mit ‚ja‘ beantworteten, wobei sie die Korrektheit ihrer Antworten als hoch einschätzten. Die Fragen gleichen Typs zu Sätzen ohne Garden-Path-Effekt (d. h. in denen ‚the deer‘ und ‚the baby‘ nicht als Objekt-NP interpretiert werden konnten) wurden weitaus seltener mit ‚ja‘ beantwortet. Außerdem antworteten die Versuchspersonen auf Fragen wie ‚Did the deer run into the woods?‘ und ‚Did the baby play in the crib?‘ mit einer Korrektheit von fast 100 Prozent mit ‚ja‘. Nach Christianson (2016, 819) lassen diese Ergebnisse darauf schließen, dass es, nachdem klar wird, dass die initiale Analyse falsch war, nicht zu einem kompletten Abbruch der Verarbeitung kommt (was zu einem Urteil ‚ungrammatisch‘ führen würde), sondern dass eine irgendwie geartete Umstrukturierung stattfinden muss. Des Weiteren wurden Sätze eher selten missverstanden, in denen die ambige Stelle relativ kurz war (26a.) und in denen die Endinterpretation unplausibel erschien (26b.):

- (26) a. *While the man hunted the deer paced into the woods.*  
b. *While the man hunted the deer paced in the zoo.*

Diese und noch weitere Ergebnisse lassen sich als Beleg dafür deuten, dass Versuchspersonen Garden-Path-Sätze signifikant öfter falsch verstehen als bis dato angenommen und sie dabei offenbar nur zu einem geringen Teil bewusst Störungen wahrnehmen (vgl. Christianson, 2016, 819). Mit dem oben erwähnten Alles-oder-nichts-Blick auf Sprachverstehensprozesse lassen sich diese Befunde kaum in Einklang bringen.

### 2.3 Die Verarbeitung von ‚Beides‘ + KoordP + Verb

Im Folgenden sollen nun die drei Sätze und die entsprechenden, in der Einleitung bereits formulierten Hypothesen vor dem Hintergrund der besprochenen Prinzipien und Strategien beim Sprachverstehen bzw. der Subjekt-Verb-Kongruenz erörtert werden. Zur Wiederholung an dieser Stelle noch einmal die Hypothesen:

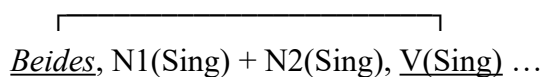
Wie werden die folgenden Sätze verarbeitet?

	Sätze	Hypothese zur Verarbeitung
Vers1	<i>Beides</i> , N1(Sing) + N2(Sing), V(Sing) ...	Long-Dependency-Hypothese
Vers2	<i>Beides</i> , N1(Sing) + N2(Sing), V(Plu) ...	Conceptual-Plural-Hypothese
Vers3	<i>Beides</i> , N1(Sing) + N2(Plu), V(Plu) ...	Proximity-Hypothese

Tabelle 2: Sätze und Hypothesen zu ihrer Verarbeitung

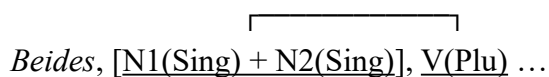
- Vers1 – Long-Dependency-Hypothese

Das formale Subjekt ‚*Beides*‘ kongruiert mit dem Verb im Singular; die koordinierte NP [N1+N2] und das Verb kongruieren nicht miteinander:


  
*Beides*, N1(Sing) + N2(Sing), V(Sing) ...

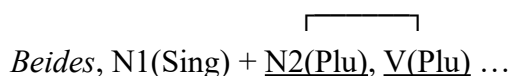
- Vers2 – Conceptual-Plural-Hypothese

Die koordinierte NP wird als konzeptueller Plural [N1(Sing)+N2(Sing)] verarbeitet und kongruiert mit dem Verb im Plural; ‚*beides*‘ und Verb kongruieren nicht miteinander:


  
*Beides*, [N1(Sing) + N2(Sing)], V(Plu) ...

- Vers3 – Proximity-Hypothese

Das unmittelbar präverbale N2(Plu) kongruiert mit V(Plu), weil es näher als die anderen potentiell kongruierenden Elemente am Verb steht:


  
*Beides*, N1(Sing) + N2(Plu), V(Plu) ...

Während der Verarbeitung dieser Sätze können also, wie oben grafisch gekennzeichnet, Kongruenz-Beziehungen zwischen verschiedenen Elementen hergestellt werden. Formalgrammatisch muss in jedem der Sätze ein Singular-Verb stehen, weil das formale Subjekt ‚*beides*‘ dies fordert. Dem Grammatik-Duden (2016, 1027) und dem Zweifelsfälle-Duden (2011, 559) zufolge sind beide Numerus-Formen akzeptabel.

Die Fragebogen-Untersuchung von Wegerer (2012), in der die Lücke im Satz

(27) *Beides, Rasen und Weg, \_\_\_\_\_ mit Laub bedeckt*

mit einer Singular- oder Plural-Form des Verbs ‚sein‘<sup>5</sup> gefüllt werden sollte, zeigt jedoch, dass muttersprachliche Sprecher\*innen des Deutschen für eine Konstruktion dieses Typs die Plural-Form bevorzugen.

Vor dem Hintergrund der Studie von Wegerer (2012), die über eine deutliche Präferenz des grammatisch inkorrekten Plural-Verbs berichtet, stellen sich bezüglich der Verarbeitung folgende Fragen:

- Wie reagiert der Parser auf Vers1, wo er mit einer Singular-Form konfrontiert wird?
- Stellt er trotz der Präferenz eines Plural-Verbs die Long-Dependency-Kongruenz her?
- Spiegeln sich diese Schwierigkeiten in den Grammatikalitätsurteilen wider?

Das grammatikalisch inkorrekte Plural-Verb dürfte also keine Schwierigkeiten verursachen, sondern den Verarbeitungsprozess eher vereinfachen. Für Vers2 und Vers3 lässt sich vermuten, dass das Phänomen *Agreement Attraction* auftreten wird (Strukturen, die Kongruenz mit anderen Elementen außer dem Subjekt enthalten, werden als korrekt beurteilt). Da in Vers3 neben der KoordP als Gesamtes auch noch das unmittelbar präverbale Nomen mit dem Verb kongruiert, könnte das dazu führen, dass diese Konstruktion im Vergleich zu Vers2 noch besser verarbeitet werden kann.

Für ein Auftreten von *Agreement Attraction* könnte außerdem folgendes sprechen:

„[...] it is easy to imagine that subject-verb agreement would derail most often when the head subject noun has few of the prototypical features of subjects, and a preverbal noun has many of them.“ (Bock & Miller, 1991, 48)

In der Konstruktion ‚*beides*‘ + *KoordP* besitzt ‚*beides*‘ weniger prototypische Subjekt-Merkmale als die zwei nominalen Elemente in der *KoordP*: ‚*Beides*‘ referiert als kataphorisches Pronomen auf etwas Abstraktes; es wird erst durch die Nomina der *KoordP* semantisch gefüllt. Die *KoordP* dagegen enthält zwei Nomina, die auf Konkrete verweisen.

---

<sup>5</sup> Die Infinitiv-Form des Verbs wurde vorgegeben.

Die in der vorliegenden Arbeit zu untersuchenden Konstruktionen unterscheiden sich offensichtlich von Garden-Path-Sätzen, die in Übersichtsdarstellungen oder zur Modellierung von Sprachverstehensprozessen angeführt werden. In der Satzstruktur ‚*Beides*‘ + *KoordP* + *VP* gibt es keine alternative syntaktischen Analysen in dem Sinne, dass sich die Semantik verändert und so unterschiedliche Lesarten entstehen könnten. Nach dem Garden-Path-Modell werden während des Verarbeitungsprozesses zwei unabhängige syntaktische Analysen durchgeführt. Es steht zur Debatte, ob der Parser bei einer Satzstruktur wie ‚*beides*‘ + *KoordP* + *VP* das tut. Bei der Verarbeitung der vorliegenden Konstruktionen mit komplexen, numerusambigen Nominalphrasen kann davon ausgegangen werden, dass kein vergleichbarer Reanalyseprozess, wie er in modularen Ansätzen angenommen wird, stattfindet.

Wenn man das *Late-Closure-Prinzip* auf die Verarbeitung der vorliegenden Konstruktion anwendet, ergeben sich folgende Überlegungen: Die ‚*beides*‘-NP könnte beim Eintreffen des Verbs bereits abgeschlossen sein. Dann könnte das Verb an die gesamte *KoordP* angeschlossen werden und sollte somit im Plural stehen (→ Vers2 – ConPlu-Hypothese); es läge Low Attachment vor. Egal, wie eine *KoordP* im Einzelnen gegliedert ist – in Vers3 (→ Prox-Hypothese) müsste das Verb ebenso im Plural stehen. Würde man davon ausgehen, dass das zweite nominale Element in einer *KoordP* dem ersten untergeordnet ist, wäre in Vers3 ebenfalls Low Attachment vorhanden, hier aber nicht an die gesamte *KoordP*, sondern nur an N2. Nimmt man an, dass kein klares hierarchisches Verhältnis zwischen der ‚*beides*‘-NP und der *KoordP* besteht, wäre die Kongruenz zwischen einem Singular-Verb und ‚*beides*‘ (→ Vers1 – LongDep-Hypothese) kein High Attachment.

Es gibt zwar Belege dafür, dass die hierarchischen Beziehungen bei der Verarbeitung von Kongruenz den höchsten Stellenwert haben, die lineare Nähe eines Elements zum Verb jedoch Einfluss hat, wenn die Hierarchie ausgeglichen ist (Haskell & MacDonald, 2005). Ist die *KoordP* also auf der gleichen syntaktischen Ebene wie das formale Subjekt anzuordnen, wäre es nicht überraschend, dass das Verb mit den näher stehenden Elementen kongruiert, insbesondere mit dem unmittelbar präverbalen (vgl. Keung & Staub, 2018, 2).

Den Studien von Wagers et al. (2009), Tanner et al. (2014) und Schlueter et al. (2018) folgend, die über Attraction-Phänomene bei Subjekt-Verb-Kongruenz berichten, soll im Folgenden das *Cue-Based-Retrieval-Modell* herangezogen werden.

Die Dependenz-Resolution beginnt, wenn der Parser das Verb samt seiner gesamten Merkmale – in den vorliegenden Fällen relevant: die Numerus-Kennzeichnung und die Subjekt-

Dependenz – verarbeitet. Ab dann werden Wiederabrufprozesse anhand des Numerus-Cues und des strukturellen Cues in Gang gesetzt. Welches sind potentielle Ziel-Elemente? In allen drei Satzstrukturen passt der strukturelle Cue [Subjekt] zu ‚*beides*‘. Die Matches des Numerus-Cue [Sing; Plu] mit den nominalen Elementen sind in der folgenden Tabelle aufgeführt:

	Sätze	Matches mit Numerus-Cue
Vers1	<i>Beides</i> , N1(Sing)+N2(Sing), V(Sing) ...	Cue[Sing] – <i>Beides</i> ; N1; N2
Vers2	<i>Beides</i> , N1(Sing)+N2(Sing), V(Plu) ...	Cue[Plu] – KoordP
Vers3	<i>Beides</i> , N1(Sing)+N2(Plu), V(Plu) ...	Cue[Plu] – KoordP; N2

Tabelle 3: Sätze und Matches mit Numerus-Cues

Der Annahme der *Expectation-Based-Theory* folgend, dass das Auftreten von Konstituenten gleichen Typs die Erwartungen und damit die Verarbeitung beeinflusst, führt zu der Frage, wie ähnlich die Konstituenten im präverbalen Material (‚*Beides*‘; N1; N2; KoordP) sind. Es ist davon auszugehen, dass die Wahrscheinlichkeiten für das präverbale Auftreten eines kataphorisch wirkenden Pronomens und die eines Konkretums unterschiedlich sind. Dennoch sind es drei nominale Elemente, die aufeinander folgen, sodass mit der Verarbeitung von N2 entsprechend (hohe) Erwartungen bezüglich einer kommenden Verbalphrase entstehen, was sich förderlich auf die Verarbeitung des schließlich eintreffenden Verbs auswirkt.

Die Beziehungen zwischen den drei nominalen Elementen sind unübersichtlich, was dem Parser Schwierigkeiten bei der Herstellung von Kongruenz-Beziehungen bereitet. An dieser Stelle lässt sich auf den *Good-enough-Parsing-Ansatz* zurückgreifen, der annimmt, dass der Parser unter Umständen weniger aufwendige (gleichzeitig falsche) Analysen gegenüber komplizierten präferiert und während des weiteren Verarbeitungsverlauf ggf. gar nicht mehr korrigiert. Möglicherweise ist also eine vollständige syntaktische Analyse gar nicht notwendig, um einen Satz so zu interpretieren, dass man Anforderungen in kommunikativen Situationen gerecht wird.

Hierzu ein Beispiel von Tabor et al. (2004), die in einer Self-paced-reading-Studie über Interferenz-Effekte aufgrund lokaler Kohärenz berichten: Versuchspersonen analysierten im

folgenden Satz (28) den Teil ‚*the player tossed a frisbee*‘ verhältnismäßig oft als Hauptsatz.

(28) a. *The coach smiled at the player who was tossed a frisbee*

b. *The coach smiled at the player ~~who was~~ tossed a frisbee*

In einer kommunikativen Situation ist es mitunter nicht so wichtig, dass man unterscheidet, ob dem\*der Spieler\*in die Frisbee zugeworfen wurde (a. – Relativsatz; b. – reduzierter Relativsatz) oder die Frisbee selbst warf (falsche Hauptsatzanalyse in b.). Die Hauptsatzanalyse wird gewählt, weil sie weniger aufwendig ist als die Analyse der komplexen Konstruktion (Matrix-/Relativsatz).

Ähnliches, und zwar dass die gesamte Satzstruktur (Matrixsatz, eingebetteter Relativsatz) aufgebrochen wird, könnte bei der Verarbeitung der Konstruktion ‚*Beides*‘ + *KoordP* + *Verbalphrase* passieren: ‚*Beides*‘ wird völlig von *KoordP* + *VP* abgelöst und nicht mitanalysiert, weil ‚*beides*‘ für das Verstehen (gerade in einer kontextfreien Situation) quasi nicht relevant ist. Die Dependenz zwischen Verb und *KoordP* aufzulösen gestaltet sich nun weniger aufwendig, weil ein potentielles Subjekt-Element wegfällt.

Am Ende der Arbeit wird vor dem Hintergrund der bisherigen Ausführungen und den Daten der empirischen Untersuchung nochmal ausführlich auf die Verarbeitung der Struktur eingegangen.

Nachdem thematisiert wurde, wie sich die Kongruenzverhältnisse in der Struktur ‚*beides*‘ + *KoordP* + *VP* syntaktisch gestalten, welche Faktoren bei der Herstellung von Kongruenz im Sprachverarbeitungsprozess eine Rolle spielen und wie die zu untersuchende Struktur verarbeitet werden könnte, wird im folgenden Teil der Arbeit die empirische Untersuchung dargestellt.

### 3 Empirische Untersuchung

Um die Kongruenz zwischen finitem Verb und der komplexen, numerus-ambigen Struktur ‚beides‘ + *KoordP* empirisch zu untersuchen, sollte eine Studie durchgeführt werden, die auf einer Basis fußt, die sowohl aus Offline- (Grammatikalitätsurteile) als auch aus Online-Daten (Blickbewegungen) besteht. Im folgenden Kapitel 3.1 wird die Methode des Eye-tracking und die damit verbundenen Prämissen über linguistischen und visuellen Input dargestellt. Darauf folgen Ausführungen über das Design der Untersuchung (Kap. 3.2), die Erhebung der Daten (Kap. 3.3) sowie die Datenauswertung (Kap. 3.4), bevor abschließend die Ergebnisse diskutiert werden.

#### 3.1 Die Methode des Eyetracking

Mit der Nachverfolgung von Augenbewegungen ist es möglich zu untersuchen, worauf Menschen ihre Aufmerksamkeit lenken und für wie lange sie das tun. Weiterhin kann man Rückschlüsse darauf ziehen, woran sie möglicherweise gerade denken und wie viel kognitiven Aufwand sie das kostet. Dazu ist es für gewöhnlich noch nicht einmal nötig, sie eine Aufgabe bearbeiten zu lassen. Es reicht, wenn Versuchspersonen alltägliche Dinge tun, wie z. B. einen Text lesen oder Bilder bzw. Videos auf einem Computerbildschirm betrachten. Eyetracking liefert eine große Menge an Echtzeit-Daten, wodurch es eine oft eingesetzte Methode in der Psychologie, der Psycholinguistik und in Bereichen der angewandten Linguistik wie der Übersetzungswissenschaft ist. Man sollte jedoch beachten, dass die Methode nicht ein „Allheilmittel“ darstellt, welches linguistische Studien von vornherein weniger künstlich macht. Sie bleibt ein Werkzeug, das es abhängig von Forschungsfragen etc. wohlbedacht einzusetzen gilt (Conklin et al., 2018, xiii).

##### 3.1.1 *Das Visual World Paradigm*

Die Grundlage für das sogenannte *Visual World Paradigm* wurde 1974 von R. M. Cooper gelegt, als er im Rahmen einer Studie Versuchspersonen kurze Erzählungen vorspielte und sie gleichzeitig Bilder mit alltäglichen Objekten, zum Teil mit, zum Teil ohne Beziehung zum Text, betrachten ließ. In der Instruktion des Experiments hieß es, dass die Pupillenweite der Versuchspersonen gemessen werde und sie hinschauen dürften, wohin sie wollten. Es zeigte sich, dass die Proband\*innen zu solchen Objekten blickten, die entweder im Text erwähnt oder in irgendeiner anderen Weise in Beziehung zu ihm standen. Außerdem



standen die Blickbewegungen in einem sehr engen zeitlichen Verhältnis zum Text: Über 90 % der Fixationen eines kritischen Objekts wurden gemessen, während das entsprechende Wort gesprochen wurde oder innerhalb von 200 msec nach Offset des Wortes.

Möglicherweise weil sich der Einsatz von Eyetracking-Geräten zu dieser Zeit noch recht mühsam gestaltete und weil mit J. A. Fodors These von der Modularität des Geistes (1983) ein einflussreiches Konzept zunächst noch nicht vorhanden war, blieb Coopers Ansatz von 1974 bis in die Mitte der 1990er-Jahre in der Psycholinguistik größtenteils unbeachtet. Erst mit dem Artikel von Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy (1995) gewann die Annahme systematischer Beziehungen zwischen Blickbewegungen und Sprachverarbeitung, dann namentlich unter dem Begriff Visual World Paradigm (Allopenna, Magnuson & Tanenhaus, 1998), in der Psycholinguistik massiv an Bedeutung (vgl. Huettig et al., 2011, 152).

Die Grundannahmen des Visual World Paradigm bestehen darin, dass

- zwischen linguistischer und visueller Informationsverarbeitung ein Abgleich bzw. Wechselwirkungen stattfinden;
- sich Augenbewegungen auf die im linguistisch-auditiven Input genannten Objekte richten, während und bis ca. 200 msec nach deren Nennung.

### *3.1.2 Grundlagen für Visual-World-Studien zum Sprachverstehen*

Visual-World-Experimente zum Sprachverstehen folgen für gewöhnlich einem relativ unkomplizierten Grundaufbau: Pro Trial wird der Versuchsperson auf einem Computerbildschirm ein visueller Stimulus präsentiert, parallel dazu hört sie einen oder mehrere gesprochene Sätze. Währenddessen werden anhand eines Eyetrackers die Blickbewegungen aufgenommen, um sie anschließend zu analysieren. Oftmals werden als visueller Input Strichzeichnungen oder Photos eingesetzt, die einzeln für sich stehen oder mehr oder weniger realistische Szenen darstellen; der auditive Input besteht aus Sätzen, die das Abgebildete unmittelbar aufgreifen oder kommentieren oder in irgendeinem Zusammenhang dazu stehen. Normalerweise werden in den Sätzen erwähnte Objekte und nicht-erwähnte, sogenannte Distraktor-Objekte abgebildet. Anhand szenischer Darstellungen lässt sich untersuchen, wie die Wahrnehmung der Szene und/oder das Weltwissen über Szenen und Ereignisse das Verstehen der mündlichen Äußerungen beeinflusst. Werden die Objekte in von einander erkennbar abgetrennten Bildern dargestellt, d. h. sie sind nicht in einen gemeinsamen visuellen Kontext eingebettet, spielt das Wissen über Szenen beim Verstehen keine Rolle.

Stattdessen eignet sich diese Art von Stimulus-Arrangement dazu, herauszufinden, wie die Aktivierung konzeptuellen und lexikalischen Wissens mit einzelnen Wörtern zusammenhängt (vgl. Huettig et al., 2011, 152).

Üblicherweise sind jene Blickbewegungen von besonderem Interesse, die zeitlich an die Nennung eines bestimmten Wortes innerhalb der Äußerung gekoppelt sind bzw. während oder kurz nach dessen Nennung ausgeführt werden. Der visuelle Stimulus wird meistens gleichzeitig mit oder etwas (z. B. 1 sec) vor dem Onset der Äußerung eingeblendet, was nach sich zieht, dass die Versuchsperson sich visuell orientiert und Informationen aus dem mentalen Lexikon abgerufen werden (vgl. Huettig et al., 2011, 153).

Aufgaben für die Versuchspersonen können entweder sein, selbst bewusst eine Handlung auszuführen (‘direct action tasks‘, hervorgerufen durch eine entsprechende Instruktion in der Äußerung), oder der Äußerung passiv zuzuhören und währenddessen auf den Bildschirm zu schauen (egal wohin) (‘look & listen tasks‘, Äußerungen sind Beschreibungen des Bildmaterials bzw. beziehen sich in irgendeiner Weise darauf) (Huettig et al., 2011, 153).

Es ist nicht schwer nachzuvollziehen, dass Versuchspersonen ein Objekt anschauen, wenn sie dazu aufgefordert sind, es zu berühren oder zu bewegen. Aber auch wenn sie keine solche Aufgabe zu erfüllen haben, blicken sie zu Objekten, die erwähnt werden. Sie versuchen, das Geäußerte mit dem visuellen Input in Zusammenhang zu bringen (z. B. Altmann & Kamide, 1999). Vermutlich hat diese Verhaltensweise etwas mit Gewohnheit zu tun, außerdem wird sie erwartet und hilft im Alltag. Ganz oft sind auditive und visuelle Reize komplementär zu einander, wie z. B. bei wissenschaftlichen Textformen (Textabschnitte und Diagramme nehmen aufeinander Bezug), in Nachrichtenvideos im Internet, bei der Wettervorhersage (Sprecheräußerungen und Landkarte), in Tutorialvideos etc., und es ist nützlich, die verschiedenen Informationstypen bei ihrer Verarbeitung miteinander zu verknüpfen (Huettig et al., 2011, 153).

### *Datenanalyse*

Was gilt es bei der Datenanalyse in Visual-World-Studien zum Sprachverstehen zu beachten? Eine Grafik, in der anhand einer Kurve Fixationsproportionen über einen zeitlichen Verlauf dargestellt werden, vermittelt den Eindruck, dass es sich dabei um eine Messung einer kontinuierlichen Größe handelt (Tanenhaus & Trueswell, 2006, 869). Man sollte sich jedoch bewusst darüber sein, dass Sakkaden und Fixationen separate, einzeln für sich ste-

hende Ereignisse sind. Daten aus einzelnen Trials können keinen Aufschluss geben über die Verarbeitung linguistischen Materials, die prozesshaft verläuft und sich so über einen gewissen Zeitraum erstreckt. Ermittelt man jedoch einen Durchschnittswert über Trials und Versuchspersonen hinweg, kann man berechnen, wie wahrscheinlich es ist, dass eine Person zu einem bestimmten Zeitpunkt in eine bestimmte Area of Interest (AOI) blickt. Vor diesem Datenhintergrund lassen sich Rückschlüsse auf den zeitlichen Verlauf von kognitiven Abläufen ziehen (Huettig et al., 2011, 153).

### *Die Messgröße Fixationsproportionen*

Fixationsproportionen bezeichnen die anteiligen Blicke auf AOIs zu einem Zeitpunkt  $x$ ; über einen Zeitraum hinweg lassen sich Veränderungen in den Fixationsproportionen nachvollziehen. Wenn man die Werte der Fixationsproportionen aller AOIs zu einem Zeitpunkt  $x$  addiert, ist das Ergebnis quasi nie exakt 1,0 (bzw. 100 %), weil Blicke schließlich auch abseits der AOIs landen können. Bei der Festlegung von AOIs sollten Bereiche definiert werden, die etwas mehr als nur das Objekt erfassen, da auch Fixationen, die nicht direkt, aber in unmittelbarer Nähe stattfinden, dazu führen, dass ein Objekt wahrgenommen, identifiziert und im (Arbeits-)Gedächtnis eine Repräsentation von ihm aufgebaut wird.

Weil eine Fixation eine gewisse Zeit andauert (~150–250 msec), sind Proportionen in verschiedenen Zeitfenstern nicht unabhängig voneinander. Um eine größere Unabhängigkeit der Daten zu schaffen, könnte man die Anzahl neuer Sakkaden zu einer AOI in einem Zeitfenster auswerten (Tanenhaus & Trueswell, 2006, 869). Wie bereits erwähnt konnte gezeigt werden, dass etwa 200 msec nach dem phonetischen Input, der die Wahrscheinlichkeit steigen lässt, dass ein kritisches Bild bevorzugt angeschaut wird, dies auch dementsprechend geschieht (Tanenhaus & Trueswell, 2006, 870).

### *3.1.3 Die Hypothese des kognitiven Aufwands*

Nach Griffin (2004) ist im Kontext der Lösung kognitiver Aufgaben davon auszugehen, dass Blicke Aufmerksamkeit und kognitiven Aufwand widerspiegeln. Versuchspersonen tendieren dazu, Objekte anzuschauen, an die sie gerade denken, bzw. sie richten ihre Blicke auf Dinge, die in irgendeiner Weise mit dem in Zusammenhang stehen, über was sie gerade nachdenken. Außerdem steigt die Wahrscheinlichkeit, dass Blickbewegungen kognitive Prozesse abbilden, mit dem kognitiven Aufwand, der die Bearbeitung einer Aufgabe erfor-

dert. Daraus schließt die Autorin, dass die bei der Sprachproduktion schwierigsten Teilaufgaben sich in den Blickbewegungen niederschlagen sollten.

Bezogen auf meine Studie – die Versuchspersonen haben eine nicht ganz einfache Aufgabe zu lösen: Sie sollen die Grammatikalität der Sätze beurteilen. Somit ist es nicht unwahrscheinlich, dass die Interpretation von Blickbewegungsdaten Aufschluss über die Verarbeitung der Satzstrukturen geben kann.

An dieser Stelle sei erwähnt, dass die Aufgabe, die den Versuchspersonen gestellt werden, generell einen großen Einfluss auf ihr Verhalten hat. Blickbewegungen unterscheiden sich von Aufgabe zu Aufgabe. Zudem verhält sich jede Person anders. Deshalb sollte man darauf achten, eine große und repräsentative Datenstichprobe zu erheben sowie eine Analyse-methode zu wählen, die eine randomisierte Verteilung der Versuchspersonen und Items einbezieht (vgl. Conklin et al., 2018, 207).

### 3.2 Design der empirischen Untersuchung

Aus den dargelegten Annahmen des Visual World Paradigm und der Hypothese des kognitiven Aufwandes stellen sich für die empirische Untersuchung folgende Fragen:

- Wie lässt sich Verarbeitung der unklaren Kongruenz-Beziehungen fassen?

Mit Offset des Verbs entsteht für den Parser die Herausforderung, Übereinstimmungen bezüglich der Kategorie Numerus zwischen Verb und infrage kommenden nominalen Elementen herzustellen. Die Elemente (Verb und Nomina) stimmen untereinander nur eingeschränkt überein. Bei der Dependenz-Resolution entsteht also ein besonderer kognitiver Aufwand.

- Wie zeigt sich der kognitive Aufwand in Blickbewegungen?

Versuchspersonen schauen Objekte, die in einem auditiven Stimulus genannt werden (Nomina in der KoordP), in entsprechenden Bildern an, weil linguistische und visuelle Information miteinander abgeglichen werden. Sie tun das außerdem, weil sie zum einen den Satz grammatikalisch beurteilen sollen, und zum anderen, weil sie wegen der erschwerten Dependenz-Resolution ihnen besondere Aufmerksamkeit widmen. Die Blicke zu den Objekten wirken, so meine Annahme, unterstützend auf die kognitive Verarbeitung der Nomina.

Im Verlauf dieses Kapitels wird nun das Design der Studie vorgestellt. Es wird v. a. darauf eingegangen, wie die verschiedenen Itemtypen zusammengesetzt sind, wie ein Trial aufge-

baut und ein Gesamtdurchlauf für eine Versuchsperson gestaltet ist, bevor die Erwartungen bezüglich der Daten dargelegt werden.

### 3.2.1 Zusammensetzung der Items

Ein Item setzt sich aus einem visuellen und einem auditiven Stimulusteil zusammen – drei Bildern und einem gesprochenen Satz. Die Versionen eines Satzes werden mit Bildern zu einem Item kombiniert. Critical- und Control-Items setzen sich so zusammen, dass der Numerus der zwei Nomina in der Apposition im Satz mit der Anzahl der Referenten in den jeweiligen Bildern übereinstimmt. Durch die Übereinstimmung der sprachlichen und visuellen Elemente bezüglich des Merkmals Numerus bzw. Anzahl sollte der Verarbeitungsaufwand, der für den Abgleich der beiden Informationstypen anfällt, möglichst klein gehalten werden. Im dritten Bild ist eine Entität zu sehen, die nicht im Zusammenhang mit dem Satz steht (Distraktor). Die Critical-Items enthalten die zu untersuchenden Sätze, die Control-Items Sätze, die den zu untersuchenden ähneln, jedoch hinsichtlich der Numerus-Merkmale der präverbalen nominalen Elemente nicht ambig sind.

Um eine Präferenz von Bildern unabhängig von den Satzstrukturen nicht zu begünstigen, wurde bei der Auswahl darauf geachtet, dass sie mehr oder weniger ähnlich salient<sup>6</sup> sind.

<i>Sätze</i>	<i>Bildkombinationen</i>	<i>Itemtypen</i>
Vers1: N1(Sing), N2(Sing)	Ref1(Sing), Ref2(Sing), Distr	Critical1
Vers2: N1(Sing), N2(Sing)	Ref1(Sing), Ref2(Sing), Distr	Critical2
Vers3: N1(Sing), N2(Plu)	Ref1(Sing), Ref2(Plu), Distr	Critical3
Vers4: N1(Plu), N2(Plu)	Ref1(Plu), Ref2(Plu), Distr	Control

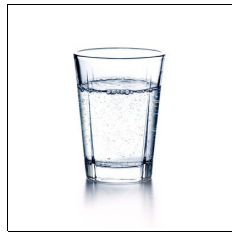
*Tabelle 4: Zusammensetzung der Critical- und Control-Items*

---

<sup>6</sup> *Salienz* – das Hervortreten, Herausstechen eines Reizes

*Beispiele für die Itemtypen:*

Critical1:



„Beides, Glas und Flasche, ist in einem Restaurant zu finden.“

Critical2:



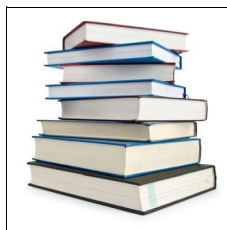
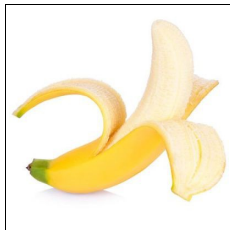
„Beides, Glühbirne und Kerze, strahlen Licht aus.“

Critical3:



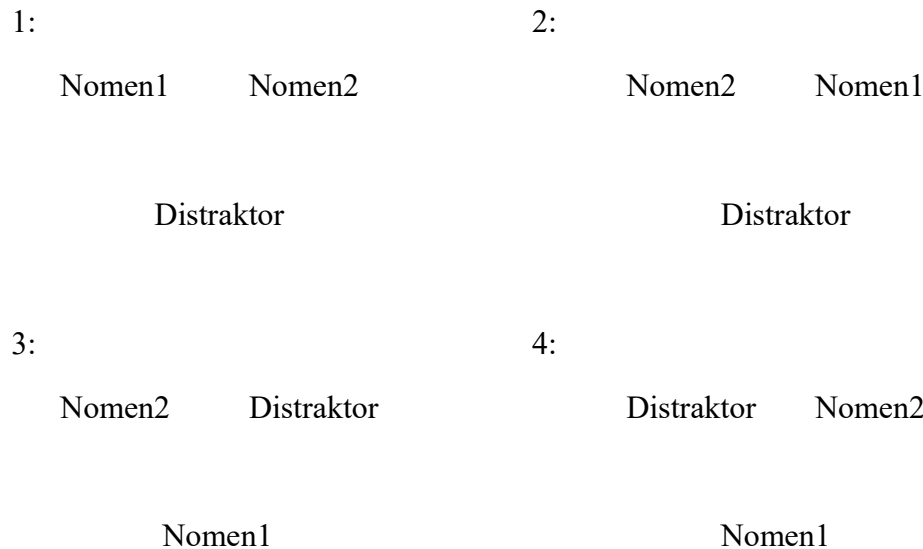
„Beides, Plastiktüte und Sprühdosen, schaden dem Klima.“

Control:



„Alle Bücher und Regale besitzen rechtwinklige Ecken.“

Die Positionen der Bilder wurde in allen Itemtypen, sowohl in den Critical- und Controls auch in den Filleritems systematisch variiert, um die Position als mögliche Störvariable auszuschließen. Hier die vier verwendeten Variationen der Bildpositionen:



(An den obigen Beispielen für die vier Itemtypen lassen sich die vier Positionsvariationen beispielhaft nachvollziehen.)

*Itemtypen:*

- Critical1 (entspricht Vers1); 4 Stück pro Versuchsperson
- Critical2 (entspricht Vers2); 4 Stück pro Versuchsperson
- Critical3 (entspricht Vers3); 4 Stück pro Versuchsperson
- Control (entspricht Vers4); 4 Stück pro Versuchsperson
- Filler: 64 Stück pro Versuchsperson; eine Hälfte grammatikalisch korrekt, die andere inkorrekt. Der visuelle Stimulus besteht (wie bei Criticals und Controls) aus drei Bildern. Die drei vor dem Verb genannten Nomina im Satz entsprechen den Referenten in den Bildern<sup>7</sup>. Im Folgenden die verwendeten Konstruktionstypen (jeweils mit Beispielen und der Anzahl pro Versuchsperson):

---

<sup>7</sup> Bei der Zusammensetzung der Filler-Items wurde nicht explizit darauf geachtet, dass sprachliche und visuelle Elemente bezüglich des Merkmals Numerus bzw. Anzahl übereinstimmen.



- Reihung mit drei Nomina: Koffer, Reisetasche und Rucksack fahren auf dem Gepäckband. (*korrekt*); 11 Stück. Saxophon, Fagott und Klarinette gehören vor das Holzblasinstrument. (*inkorrekt*); 11 Stück.
- 3er-Reihung, *alle...* : Maus, Meerschweinchen und Ratte, alle gehören zu den Nagetieren. (*korrekt*); 11 Stück. Herz, Lunge und Darm, alle sind ohne menschliche Organe. (*inkorrekt*); 11 Stück.
- *Sowohl N als auch N und N...*: Sowohl Fernglas als auch Lupe und Mikroskop lassen Dinge größer erscheinen. (*korrekt*); 5 Stück. Sowohl Milch als auch Joghurt und Käse stehen im Kühlregal drunter. (*inkorrekt*); 5 Stück.
- *Sowohl N und N als auch N...*: Sowohl Enten und Schwäne als auch Biber leben im Wasser. (*korrekt*); 5 Stück. Sowohl Sprungbrett und Rutsche als auch Becken sind im Schwimmbad zum Schwimmen. (*inkorrekt*); 5 Stück.
- Intro-Items (vom Typ wie Filler); 4 Stück

Gesamtanzahl der Items pro Versuchsperson: 84 Stück

Die 4 Intro-Items sowie die 64 Filler-Items waren für alle Versuchspersonen die selben.

### *Verteilung der Criticals und Controls*

Es existieren 16 semantisch verschiedene Sätze wie z. B.

*Beides, Nagel und Schraube, ist aus Eisen.*

*Beides, Kamera und Mikrofon, dient der Überwachung. etc.*

Jeder dieser Sätze ist in vier Ausprägungen (3 Criticals, 1 Control) vorhanden. Die daraus erstellten 64 Items wurden so in vier Gruppen aufgeteilt, dass jede Versuchsperson vier Items jedes Typs präsentiert bekam, wobei keiner Person Sätze mit gleichen Referenten gezeigt wurden. Zudem wurde darauf geachtet, dass die Positionsvariationen innerhalb der Itemtypen gleichmäßig verteilt sind.

### 3.2.2 *Präsentation der Items*

Die Präsentation eines Items läuft wie folgt ab:

- In der Mitte des weißen Bildschirms erscheint ein Kreuz. Sobald die Versuchsperson das Kreuz 500 msec lang fixiert hat, beginnt der Trial (Start des Trials durch die Funktion des sog. AOI-Triggers).

- Drei Bilder werden eingeblendet. Bei Critical-Stimulis entsprechen zwei davon den Nomina in der KoordP, das dritte Bild steht in keinem semantischen Zusammenhang zum Satz. Ab dem Stimulus-Onset beginnt die Versuchsperson, sich visuell zu orientieren. Möglichen Salienzeffekten wird damit vorgebeugt.
- Zwei Sekunden nach dem visuellen SO hört die Versuchsperson einen gesprochenen Satz. (Dauer des Satzes ca. 4 sec)
- Die Bilder bleiben nach Ende des Satzes noch zwei Sekunden eingeblendet.
- Anschließend erscheint die Frage ‚Korrekt?‘ auf dem Bildschirm, wodurch die Versuchsperson aufgefordert wird, eine Entscheidung zur grammatikalischen Korrektheit des Satzes zu treffen. Die Frage ist drei Sekunden lang sichtbar. Währenddessen soll die Versuchsperson eine entsprechende Taste drücken, entweder ‚A‘ für ‚ja‘ oder ‚+‘ (im Ziffernblock der Tastatur) für ‚nein‘.

Aus dem Design eines Items ergibt sich eine Dauer von ca. 11,5 Sekunden pro Item. Aus der Anzahl von 84 Items folgt somit eine Gesamtdurchlaufdauer von ca. 17 Minuten.

Der Aufbau eines Durchlaufs hat folgende Struktur:

- 4 Intro-Items (innerhalb dieser 4er-Gruppe randomisiert)
  - 4 Filler (innerhalb dieser 4er-Gruppe randomisiert)
  - 1 Critical oder Control (innerhalb aller Criticals/Controls randomisiert)
  - 4 Filler
  - 1 Critical oder Control
  - 4 Filler
- usw.

### 3.2.3 Abhängige Variablen und Erwartungen

*Unabhängige Variable* – Satzstruktur (4 Ausprägungen; Vers1–4)

Vers1 – *Beides*, N1(Sing) und N2(Sing) + V(Sing)

Vers2 – *Beides*, N1(Sing) und N2(Sing) + V(Plu)

Vers3 – *Beides*, N1(Sing) und N2(Plu) + V(Plu)

Vers4 – *Alle* N1(Plu) und N2(Plu) + V(Plu)

*Abhängige Variable 1* – Grammatikalitätsurteile

Grammatikalitätsurteile werden operationalisiert in Form von Tastendrücker, die am Ende jedes einzelnen Trials von den Versuchspersonen getätigt und von der Software registriert und gespeichert werden. Die Anzahl der Tastendrücker ‚korrekt‘/‚inkorrekt‘ wird dann interpretiert, zu je einer Ausprägung der Satzstruktur, über alle Versuchspersonen hinweg.

*Erwartungen bezüglich der Grammatikalitätsurteile*

Wie in Kapitel 2.3 (*Die Verarbeitung von ‚Beides‘ + KoordP + Verb, → ab S. 30*) bereits ausgeführt, kann mit *Agreement Attraction* (Strukturen, die Kongruenz mit anderen Elementen außer dem Subjekt enthalten, werden als korrekt beurteilt) bei Vers2 und Vers3 gerechnet werden, weil Wegerer (2012) für die Konstruktion ‚*beides*‘ + *KoordP* + *VP* eine deutliche Präferenz des Plural-Verbs feststellte. D. h. bei Vers2 & 3 sollte es eine deutlich höhere Anzahl von Tastendrücker für ‚korrekt‘ geben als für ‚inkorrekt‘.

Bei Vers1, folgt man Wagers et al. (2009) und Tanner et al. (2014), ist eher *nicht* von einem umgekehrten *Attraction*-Effekt auszugehen (d. h. die *KoordP* wirkt *nicht* als *Attractor* auf ein Plural-Verb; das Verb im Singular führt *nicht* zu einem Urteil ‚grammatikalisch inkorrekt‘) (→ S. 26).

Nimmt man eine Abtrennung des formalen Subjekts vom Rest an (→ S. 35), wäre in Vers1 für den Parser Struktur (29a.) zu analysieren, die durchaus akzeptabel ist; ähnlich akzeptabel sind die ‚Reststrukturen‘, die sich in Vers2 (29b.) und Vers3 (29c.) ergäben:

- (29)
- a. *Buch und Regal besitzt rechtwinklige Ecken.*
  - b. *Buch und Regal besitzen rechtwinklige Ecken.*
  - c. *Buch und Regale besitzen rechtwinklige Ecken.*

Weil nicht mit einem umgekehrten Attraction-Effekt zu rechnen ist und bei einem möglichen Aufbruch der Gesamtstruktur eine grammatikalisch akzeptable Reststruktur entstünde, ist auch für Vers1 eine hohe Anzahl von Tastendrücker für ‚korrekt‘ zu erwarten.

Für die Kontrollbedingung (Vers4) ist von einer deutlich höheren Akzeptabilität als für die anderen Ausprägungen des Satzes auszugehen, weil hier das gesamte präverbale Material hierarchisch klarer geordnet ist als in Vers1–3 und Kongruenz zwischen allen nominalen Elementen untereinander sowie dem Verb herrscht. Entsprechend können bei Vers4 deutlich mehr Tastendrücker für ‚korrekt‘ als für ‚inkorrekt‘ erwartet werden.

#### *Abhängige Variable 2 – Eyetracking-Daten*

Wie genau kann ‚Aufmerksamkeit‘ oder ‚Präferenz einer AOI‘ operationalisiert werden? Im Vorfeld wurde sich nicht für eine bestimmte Art der Operationalisierung festgelegt. Dementsprechend wurde als Arbeitsschritt zunächst die Erhebung der Rohdaten geplant und durchgeführt. Die Fixationsproportionen erwiesen sich dann als interessant. Je mehr Fixationen anteilig in einer AOI registriert werden, desto mehr Aufmerksamkeit wird dem dort abgebildeten Objekt gewidmet. Bei der Auswertung gilt es, mögliche Präferenzen von AOIs über den Zeitverlauf zu interpretieren.

#### *Erwartungen bezüglich der Eyetracking-Daten*

Zu Beginn von Kapitel 3.2 (*Design der empirischen Untersuchung*, → S. 40) wurde bereits darauf eingegangen, dass es zum Zeitpunkt der Dependenz-Resolution, also bei der Verarbeitung des Verbs, zu einem erhöhten kognitiven Aufwand kommt und anzunehmen ist, dass Blicke auf Bilder der Referenten die Verarbeitung der Nomina erleichtern. Somit ist die Wahrscheinlichkeit hoch, dass die mit den beiden Nomina der KoordP übereinstimmenden Bilder mehr Blicke auf sich ziehen als das Distraktor-Bild, und zwar ab einem Zeitpunkt von 200 msec nach Offset des Verbs, für eine Dauer von ca. zwei bis drei Sekunden. Spiegeln die Blickbewegungen kognitiven Aufwand wider, ist zu erwarten, dass sich in dem genannten Zeitfenster die Daten der Kontrollbedingung sichtlich von denen der kritischen Bedingungen unterscheiden.

### 3.3 Datenerhebung

#### 3.3.1 Vorbereitung

Über eine Bildersuchmaschine fand ich im Internet die benötigten Dateien (bildliche Darstellungen der Nomina aus den Sätzen und Distraktor-Objekte), die anschließend, falls noch nicht so vorhanden, mit der Bildbetrachtungs-/ bearbeitungssoftware IrfanView einheitlich auf die Maße 400 x 400 Pixel gebracht und ins jpg-Format konvertiert wurden. Durch die Einheitlichkeit der Bildmaße konnten die AOIs einfach definiert werden.

Die Sätze wurden von mir eingesprochen und über ein im Laptop integriertes Mikrofon mit der Software Audacity aufgenommen. Der einzelne Satz wurde anschließend, mit einer künstlich generierten, jeweils 2 Sekunden andauernden Stille vor und nach der Aufnahme, als Sounddatei im wav-Format gespeichert. Aus technischen Gründen, die in der Software SMI Experiment Center liegen, mussten die Sounddateien dann, kombiniert mit einer weißen jpg-Bilddatei, in mp4-Filmdateien konvertiert werden.

Anhand einer systematisch konzipierten Excel-Datei wurden die Bild- und Filmdateien sowie die Abläufe, in welchen Kombinationen und welcher Reihenfolge welche Itemtypen erscheinen, in SMI Experiment Center importiert.

#### 3.3.2 Durchführung

Die Datenerhebung fand in einem Zeitraum von ca. sechs Wochen (im Dezember 2018 und Januar 2019) im Eyetracking-Labor des Instituts für Deutsch als Fremdsprachenphilologie der Universität Heidelberg statt.

Zunächst wurde den Versuchspersonen der grobe Ablauf des Experiments geschildert: Es werde etwa 20 Minuten dauern, sie würden auf dem Computerbildschirm Bilder zu sehen bekommen und gleichzeitig würden ihnen über Lautsprecher Sätze vorgespielt. Sie müssten dann, jeweils nach Ende eines Satzes, eine Entscheidung bezüglich der Grammatikalität dieses Satzes treffen und eine entsprechende Taste für ‚korrekt‘ oder ‚inkorrekt‘ drücken. Außerdem wurde den Versuchspersonen eine Instruktion (inkl. Beispielitems) in Papierform vorgelegt (s. beigelegter Datenträger). Darin wurden die sie darauf hingewiesen, verschiedene Dinge zu beachten, u. a.:

- Bequem hinsetzen, da möglichst wenig Bewegung erwünscht (→ Einstellung des Eyetrackers)

- Es geht nicht darum, ob der gesprochene Satz mit den Bildern inhaltlich übereinstimmt oder die inhaltliche Aussage des Satzes an sich stimmt. (→ nur Grammatikalität des Satzes ist von Interesse)
- Die Reaktionszeit spielt keine Rolle.
- Finger auf den relevanten Tasten halten (,A‘ und ,+‘ im Ziffernblock)
- Nicht mit dem Versuchsleiter interagieren während des Experiments

Um ablenkend wirkende Reize gering zu halten, platzierte ich unmittelbar vor Beginn des Durchlaufs eine verschiebbare Trennwand zwischen der Versuchsperson und mir. Die Bilder wurden auf einem Computerbildschirm präsentiert, der Satz wurde über Lautsprecherboxen vorgespielt.

Aus den Aufnahmen der 28 Probanden und Probandinnen, die ins Labor kamen, konnte in 25 Fällen durch SMI Experiment Center eine Datei generiert werden, die weiter zu verarbeiten war. Bei drei der 28 Versuchspersonen funktionierte die Kalibrierung des Eye-trackers nicht so wie erhofft<sup>8</sup>, sodass ich mich dazu entschied, mit diesen Personen das Experiment nicht durchzuführen. Fünf der 25 weiter zu verarbeitenden Datensätze stammten aus Probedurchläufen vor Beginn der eigentlichen Erhebungsphase. Diese Durchläufe wurden teils nicht vollständig bzw. mit Deutsch-als-Fremdsprache-Sprecherinnen durchgeführt, so dass die daraus generierten Daten für die Endauswertung nicht infrage kamen. 20 Datensätze konnten somit schlussendlich ausgewertet werden.

## 3.4 Auswertung der Daten

### 3.4.1 *Tastendrucke*

Für jede der vier Ausprägungen wurden 80 Items über alle Versuchspersonen hinweg gezeigt. Es kam vor, dass zu einem Stimulus mehrfach gedrückt wurde. In diesen Fällen wurden Mehrfachdrücke nicht gezählt. Bei fünf Items wurden sowohl Tastendrucke für ‚korrekt‘ als auch für ‚inkorrekt‘ registriert (in diesen Fällen revidierten Versuchspersonen offensichtlich ihr Urteil), für drei Items wurde gar kein Tastendruck erfasst (s. genaue Auflistung auf dem beigelegten Datenträger). Aus diesen Gründen ergibt die Summe pro Ausprägung nicht immer genau 80.

---

<sup>8</sup> Bei der Kalibrierung wird der Einfallswinkel zwischen der Eyetracking-Kamera und den Augen der Person gemessen. Um valide/reliable Daten erheben zu können, sollte die Abweichung der Grad-Werte einen bestimmten Rahmen nicht überschreiten.

	Anzahl der Tastendrucke	als ,korrekt‘ bewertet	als ,nicht korrekt‘ bewertet	gesamt
	Sätze			
Vers1	Beides, Buch und Regal + SingV	37	46	83
Vers2	Beides, Buch und Regal + PluV	58	20	78
Vers3	Beides, Buch und Regale + PluV	64	16	80
Vers4	Alle Bücher und Regale + PluV	75	6	81

*Tabelle 5: Absolute Zahlen zur Akzeptabilität der Sätze*

Wie erwartet zeigt sich für die Kontrollbedingung (Vers4) eine deutlich höhere Anzahl an Drücken für ‚korrekt‘ (75) als für ‚inkorrekt‘ (6). Bei keiner anderen Ausprägung ist das in diesem Maße der Fall.

Für die kritischen Bedingungen Vers2 und Vers3 waren ebenfalls mehr Drücke für ‚korrekt‘ als für ‚inkorrekt‘ zu erwarten. Dies bestätigte sich, jedoch ist das Verhältnis bei weitem nicht so deutlich wie bei der Kontrollbedingung: Bei Vers2 entfielen etwa 3/4 der Drücke auf ‚korrekt‘ und etwa 1/4 auf ‚inkorrekt‘; bei Vers3 wurden 4/5 für ‚korrekt‘ und 1/5 für ‚inkorrekt‘ verzeichnet.

Bezüglich Vers1 herrscht ein relativ ausgeglichenes Verhältnis, mit ganz leichter Tendenz zur Bewertung ‚inkorrekt‘ (37 – ‚korrekt‘, 46 – ‚inkorrekt‘). Die Erwartung zu Vers1 (eine ebenfalls deutlich höhere Anzahl von Tastendrucke für ‚korrekt‘) bestätigte sich demnach nicht.

## Statistische Auswertung der Tastendrucke

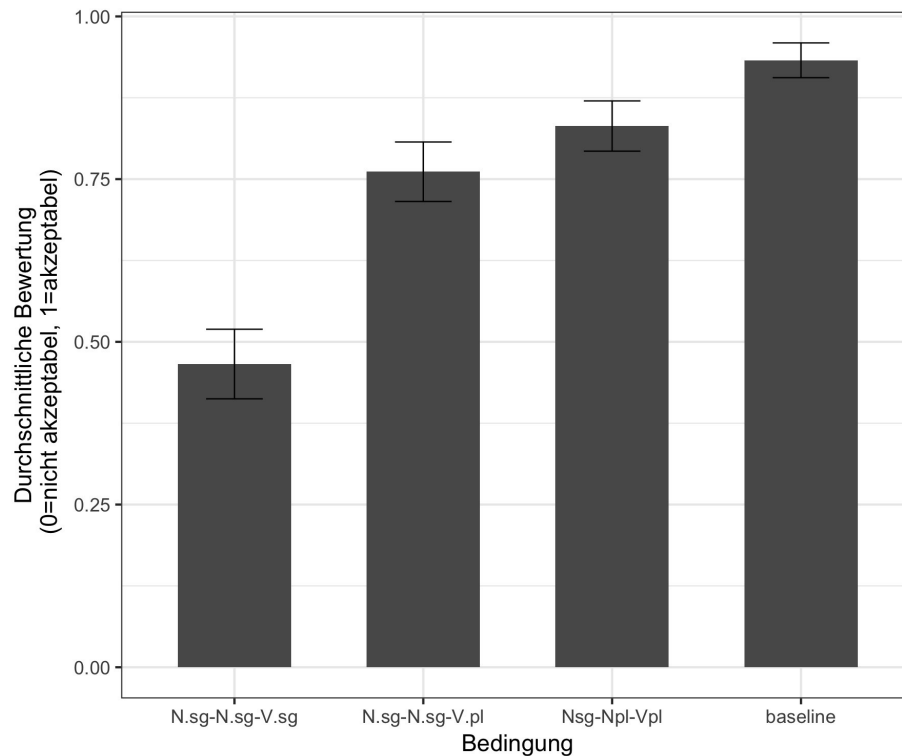


Abbildung 1: Statistische Auswertung der Tastendrucke

Die Bedingungen entsprechen von links nach rechts Vers1-4.

Die Analyse der binomialen Daten (akzeptabel = 1, nicht akzeptabel = 0) erfolgte mit einem generalisierten mixed-effects Modell.

Als abhängige Variable wurde die Entscheidung der Versuchspersonen eingesetzt, als fester Faktor die experimentelle Bedingung („Dummy Coding“ mit Baseline als Vergleichsbedingung).

Bei den Zufallsvariablen wurde ein random intercept für die Variable Versuchsperson und ein random intercept und random slope für die Variable Stimulus spezifiziert (die Spezifikation von random intercepts und random slopes für beide Zufallsvariablen ließ das Modell nicht konvergieren).

Tabelle 6 zeigt die Ergebnisse; alle kritischen Bedingungen unterscheiden sich signifikant von der Baseline.



	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	3.766	0.899	4.189	0.000	***
baseline vs Nsg-Nsg-Vsg (Vers1)	-4.039	0.865	-4.672	0.000	***
baseline vs Nsg-Nsg-Vpl (Vers2)	-2.327	0.864	-2.692	0.007	**
baseline vs Nsg-Npl-Vpl (Vers3)	-1.927	0.868	-2.220	0.026	*

Tabelle 6: Tastendrucke: Statist. Vergleich zw. kritischen u. Kontrollbedingung

In einem zweiten Schritt sollte ermittelt werden, ob sich die kritischen Bedingungen untereinander in Bezug auf die Akzeptabilitätsurteile unterscheiden.

Hierfür wurde das selbe statistische Modell wie oben verwendet mit dem einzigen Unterschied, dass als Vergleichsbedingung nun „Nsg-Nsg-Vpl“ (Vers2) ausgewählt wurde (ebenefalls „Dummy Coding“).

Tabelle 7 zeigt die Resultate; die Nsg-Nsg-Vpl-Bedingung (Vers2) unterscheidet sich signifikant von der Nsg-Nsg-Vsg-Bedingung (Vers1), aber nicht von der Nsg-Npl-Vpl-Bedingung (Vers3).

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	1.439	0.434	3.313	0.001	***
Nsg-Nsg-Vpl vs Nsg-Nsg-Vsg (Vers2 vs Vers1)	-1.711	0.404	-4.232	0.000	***
Nsg-Nsg-Vpl vs Nsg-Npl-Vpl (Vers2 vs Vers3)	0.400	0.432	0.928	0.354	

Tabelle 7: Tastendrucke: Statist. Vergleich zw. den kritischen Bedingungen

### 3.4.2 Eyetracking-Daten

Wie oben bereits angesprochen wurde die Messgröße, hinsichtlich welcher die Eyetracking-Daten ausgewertet werden sollten, vor der Erhebung nicht festgelegt. Die Auswertung bezüglich der *Fixationsproportionen* lieferte dann Werte, die in einer Weise aussagekräftig waren, dass sie als Grundlage für Interpretationen im Rahmen der vorliegenden Arbeit genommen werden konnten. In der folgenden Abbildung sind die Fixationsproportionen der AOIs grafisch dargestellt:

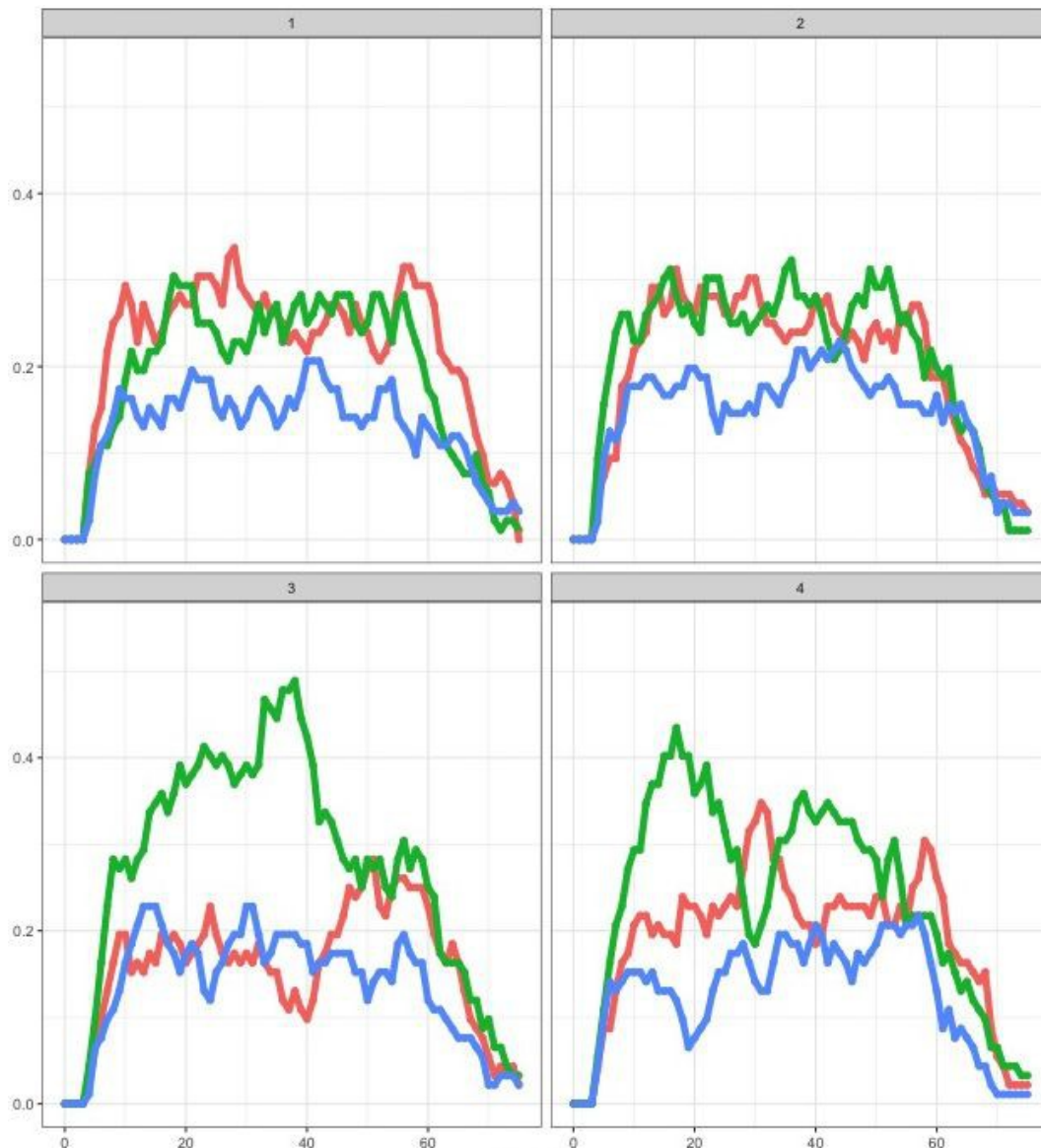


Abbildung 2: Fixationsproportionen der Areas of Interest

Größe der AOIs entsprechen den Bilddateien 400 x 400 Pixel; ab 200 msec nach Verb-Offset, zu den Ausprägungen der Satzstruktur (Vers1–4), über alle 20 Versuchspersonen.

Kurven: rot – Ref1; grün – Ref2; blau – Distraktor.

x-Achse: Zeiteinheiten. 20 entsprechen 1 sec.

y-Achse: Anteile der Fixationen in einer AOI. 0,2 entspricht 20 % etc..

Zunächst lässt sich konstatieren, dass sich die Erwartung bestätigt hat, dass sich die Blickbewegungen zwischen der Kontrollbedingung einerseits und den kritischen Bedingungen andererseits merklich unterscheiden.

In den Grafiken zu Vers1 und 2 sind keine deutlichen Präferenzen für einen Referenten erkennbar. Es ist lediglich eine tendenzielle Bevorzugung der beiden Referenten gegenüber dem Distraktor über den Großteil des Zeitfensters hinweg auszumachen.

Die Grafik zu Vers3 belegt eine frühe, bis  $\sim 2,4$  sec anhaltende, deutliche Präferenz für Ref2.

Die Grafik zur Kontrollbedingung (Vers4) zeigt eine frühe, bis  $\sim 1,3$  sec anhaltende, deutliche Präferenz für Ref2. Im weiteren Zeitverlauf lässt sich ein Wechsel der Präferenz in der Abfolge Ref1–Ref2–Ref1 erkennen, und zwar bis  $\sim 3,3$  sec, wobei die Differenzen zwischen den Proportionen nicht so deutlich sind wie vor dieser Abfolge.

## 4 Diskussion

*Wie lassen sich die Anzahlen der Tastendrucke interpretieren?*

Die verhältnismäßig hohe Anzahl von Tastendrucke für ‚korrekt‘ in der Kontrollbedingung ist meiner Ansicht nach ein Anzeichen dafür, dass die Versuchspersonen die Satzstruktur dieser Ausprägung relativ reibungslos verarbeiten konnten. Die Verhältnisse der Anzahl der Tastendrucke in den drei kritischen Bedingungen deuten darauf hin, dass es sich bei den entsprechenden Sätzen um sprachliche Zweifelsfälle handelt. Es lässt sich vermuten, dass Schwierigkeiten bei der Verarbeitung auftraten, jedoch sollten an diesem Punkt die Eyetracking-Daten (die Einblicke in den Verarbeitungsprozess ermöglichen) hinzu gezogen werden.

*Wie kann die anfänglich deutliche Präferenz für Ref2 in Vers3 und 4 interpretiert werden?*

Die Präferenz für Ref2 kann *nicht alleinig* auf die unmittelbar präverbale Stellung von N2 zurückgeführt werden, da auch in Vers1 Kongruenz zwischen N2 und Verb vorliegt, aber keine Präferenz für Ref2. Ebenso wenig kann als alleinige Ursache dafür gelten, dass in beiden Sätzen die KoordP mit dem Verb kongruiert, da selbiges auf den Satz in Vers2 zutrifft, sich dort jedoch keinerlei Präferenz für Ref2 zeigt.

Betrachtet man die Grammatikalitätsurteile zu Vers3 und 4, wurden die Sätze in beiden Bedingungen relativ oft als ‚korrekt‘ beurteilt. Es bestehen also Ähnlichkeiten zwischen Vers3 und 4 bezüglich beider abhängigen Variablen. Die Strukturen der beiden Sätze hingegen unterscheiden sich deutlich – in Vers4 ist kein erhöhter kognitiver Aufwand für die Verarbeitung zu erwarten, weil die Kongruenzbeziehungen leichter herzustellen sind als in Vers3. Dadurch wird ersichtlich, dass die Präferenz für Ref2 nicht per se als kognitiver Aufwand interpretiert werden kann.

Wenn man davon ausgeht, dass die Grafik zu Vers4 einen relativ reibungslosen Verarbeitungsprozess widerspiegelt, und die Grafik zu Vers3 einen im Verhältnis dazu aufwendigeren, könnte sich dieser höhere Aufwand für Vers3 in der länger andauernden Präferenz für Ref2 zeigen (bis ~2,4 sec). Gleicht man diese Annahme mit den Grammatikalitätsurteilen zu Vers3 ab, lässt das auf eine immer noch relativ erfolgreiche Verarbeitung schließen.

Kann die Präferenz für Ref2 in Vers3 nun so interpretiert werden, dass entsprechende Blicke förderlich auf die Verarbeitung des Nomens wirken und/oder dessen erfolgreiche Verarbeitung widerspiegeln? Unabhängig davon kann für beide Bedingungen vermutet wer-

den, dass mit dem Ende der Präferenz für Ref2 die Kongruenz zwischen Verb und N2 hergestellt wurde.

Nimmt man an, dass die Grammatikalitätsurteile etwas über die Verarbeitung aussagen, spiegelt auch die Grafik zu Vers1 einen aufwendigen Prozess wider. In ähnlichem Maße gilt dies auch für die Grafik zu Vers2. Die ausbleibenden Präferenzen könnten in diesen beiden Bedingungen auf Verarbeitungsprozesse hinweisen, in denen Schwierigkeiten bei der Dependenz-Resolution auftreten. Zieht man an dieser Stelle die geringe Akzeptabilität der Sätze in Vers1 hinzu, kann man für Vers1 von tendenziell aufwendigen und relativ erfolglosen Verarbeitungsprozessen ausgehen.

Um die weiteren Schritte der Interpretation besser nachvollziehen zu können sind in der folgenden Tabelle die vier Ausprägungen der Satzstruktur und die Übereinstimmungen zwischen den nominalen Elementen und dem Verb dargestellt (*Häkchen*: Match; *Querstrich*: kein Match):

	Kongruenzverhältnisse mit V	Long Dependency Match (,Beides'-V)	Conceptual Plural Match (KoordP-V)	Proximity Match (N2-V)
	Ausprägungen der Satzstruktur			
Vers1	Beides, Buch und Regal + SingV	✓	–	✓
Vers2	Beides, Buch und Regal + PluV	–	✓	–
Vers3	Beides, Buch und Regale + PluV	–	✓	✓
Vers4	Alle Bücher und Regale + PluV	✓	✓	✓

*Tabelle 8: Ausprägungen der Satzstruktur und Kongruenzverhältnisse mit V*

In allen Sätzen wird das formale Subjekt ‚beides‘ syntaktisch als Singular und/oder semantisch-konzeptuell als Plural interpretiert und im Arbeitsgedächtnis zwischengespeichert. Die KoordP wird immer als Conceptual Plural zwischengespeichert. Die zunehmende Distanz zwischen formalem Subjekt und Verb sorgt dafür, dass die Leistung des Arbeitsgedächtnis negativ beeinflusst wird. Die Erwartung, dass ein finites Verb als nächstes Element eintrifft, steigt mit den aufeinander folgenden nominalen Elementen. Es ist nicht ausgeschlossen, dass durch die mehreren präverbale nominalen Elemente eine höhere Erwartung bezüglich eines Plural- als eines Singular-Verbs entsteht.

Trifft das Verb ein, können Kongruenzbeziehungen hergestellt werden. Elemente, mit denen das Verb kongruieren kann, sind ‚beides‘, N1, N2 und die KoordP, die alle als Merkmalsbündel im Arbeitsgedächtnis repräsentiert sind. Um auf ein Argument zugreifen zu können, müssen die Eigenschaften der Wiederabruf-Cues, die das Verb ‚aussendet‘, mit den potentiellen Elementen im Arbeitsgedächtnis abgeglichen werden. Im Falle von Vers2 und 3 erfährt das Subjekt keine verstärkte Aktivierung durch den Numerus-Cue [Plural], sondern ausschließlich durch den strukturellen Cue [Subjekt]. In Vers2 wird die KoordP und in Vers3 die KoordP *und* das Einzelelement N2 durch den Numerus-Cue [Plural] verstärkt aktiviert. In diesen beiden Fällen könnte es zu einem fälschlichen Abruf dieser Attractor-Elemente kommen. Der Zugriff auf das nicht-lokale Argument ‚beides‘ funktioniert nicht, weil die anderen Elemente mit ihren prototypische Subjekt-Merkmalen Interferenzen verursachen und es so nicht zu einem Abruf des formalen Subjekts kommt (→ Similarity-Based Interference).

Zieht man nun die empirischen Daten heran, kann festgehalten werden, dass weder die Grammatikalitätsurteile noch die Fixationsproportionen nahelegen, dass bei der Verarbeitung der Sätze 1–4 die Long-Dependency-Kongruenz diejenige ist, die hergestellt wird.

In Vers1 ist ein LongDep-Match vorhanden (‚beides‘–SingV), die Akzeptabilität ist jedoch, relativ zu den anderen Versionen, signifikant niedriger. Die Fixationsproportionen zu Vers1 zeigen keinerlei Präferenz für einen Referenten. Beide Datentypen zusammen lassen auf Schwierigkeiten bei der Verarbeitung schließen.

Vers3 wird dagegen viel häufiger akzeptiert – trotz eines LongDep-*Mismatches*. Zudem zeigen die Eyetracking-Daten eine deutliche Präferenz von Ref2. Diese beiden Umstände zusammen lassen sich als eine relativ reibungslose Verarbeitung des pluralischen N2 interpretieren. Bei Vers1 und 3 ist ein Prox-Match vorhanden, jedoch bei Vers3 zusätzlich ein ConPlu-Match. Die Kombination aus Prox- und ConPlu-Match (Vers3) wird also einfacher verarbeitet als die Kombi aus Prox- und LongDep-Match (Vers1). Es ist zu vermuten, dass der ConPlu-*Mismatch* in Vers1 für Verarbeitungsschwierigkeiten sorgt. Die immer noch relativ hohe Akzeptabilität für Vers1 (37mal akzeptiert, 46mal nicht) ist auf den Prox-Match zurückzuführen, jedoch nicht auf den LongDep-Match.

Vers2 wird relativ oft akzeptiert, trotz erschwelter Verarbeitung laut Eyetracking-Daten und trotz des Prox-*Mismatches*. Die relativ hohe Akzeptabilität lässt sich somit vermutlich auf den ConPlu-Match zurückführen. Der LongDep-*Mismatch* hat hierbei anscheinend einen geringen Einfluss, obwohl die grammatische Kongruenzregel missachtet wird.

In Vers3 kommen ConPlu- und Prox-Match zusammen – zum Numerus-Cue [Plural] des Verbs gibt es zwei passende Elemente, die wohl auch abgerufen werden. Eine hohe Akzeptabilität sowie eine starke anfängliche Präferenz von N2 laut den Eyetracking-Daten sprechen für die Annahme einer relativ störungsfreien Verarbeitung.

Im Einklang mit den Ergebnissen wäre der *Good-enough-Parsing-Ansatz*, nach dem man davon ausgehen könnte, dass im Falle der drei kritischen Bedingungen die Gesamtstruktur des Satzes aufgebrochen wurde. ‚Beides‘ wurde abgetrennt, nur die KoordP und die VP wurden analysiert. Eine Long-Dependency-Resolution stand als Analysealternative somit gar nicht zur Verfügung. Zusammenhängend analysiert worden wären dann Sätze wie die folgenden:

Vers1 – *Buch und Regal besitzt rechtwinklige Ecken.*

Vers2 – *Buch und Regal besitzen rechtwinklige Ecken.*

Vers3 – *Buch und Regale besitzen rechtwinklige Ecken.*

Diese Sätze weisen meiner Meinung nach eine ähnliche Akzeptabilität wie ihre Ursprungsversionen auf.

Die Auswertung der Grammatikalitätsurteile und sowie der Eyetracking-Daten lässt vermuten, dass das Analysesystem bei der Herstellung der potentiellen, oben beschriebenen Kongruenzbeziehungen zwischen mehreren Elementen Schwierigkeiten hat. Die formal-syntaktischen Regeln treten bei der Verarbeitung in den Hintergrund. Zwischen dem formalen Subjekt und dem finiten Verb stehende Elemente, die selbst prototypische Eigenschaften eines Subjekts aufweisen, wirken als Konkurrenten. Nach dem Cue-Based-Retrieval-Modell ist es gut möglich, dass diese Elemente einfacher wieder abgerufen als das formale Subjekt.

Es gibt zwar Belege dafür, dass die hierarchischen Beziehungen bei der Verarbeitung von Kongruenz den höchsten Stellenwert haben, die lineare Nähe eines Elements zum Verb jedoch Einfluss hat, wenn die Hierarchie ausgeglichen ist (Haskell & MacDonald, 2005). Ist die KoordP also auf der gleichen syntaktischen Ebene wie das formale Subjekt anzuordnen, wäre es nicht überraschend, dass das Verb mit den näher stehenden Elementen kongruiert, insbesondere mit dem unmittelbar präverbale (vgl. Keung & Staub, 2018, 2).

In den vorliegenden Fällen zeigt sich das Phänomen Agreement Attraction: Im Vergleich zur kritischen Bedingung mit Singular-Verb werden in den anderen beiden kritischen Bedingungen mit Plural-Verb eine signifikant höhere Anzahl an Tastendrücker für ‚korrekt‘ verzeichnet. Die Long-Dependency-Hypothese für Vers1 wurde widerlegt, insofern, dass die LongDep-Resolution selten stattfindet. Die Conceptual-Plural-Hypothese scheint sich für Vers2 und 3 zu bestätigen, insofern, dass die Sätze in diesen Bedingungen relativ zu Vers1 eine signifikant höhere Akzeptabilität erfahren. Proximitätskongruenz macht den Unterschied zwischen Sätzen in Vers2 und 3 aus und verursacht eindeutig andere Blickbewegungen, und zwar so, dass in Vers3 Ref2 besondere Aufmerksamkeit gewidmet wird. Ob diese nun tatsächlich förderlich auf die Verarbeitung des pluralischen N2 wirken und/oder dessen erfolgreiche Verarbeitung widerspiegeln, bleibt offen, da sich die Akzeptabilitäten für Vers2 und 3 nicht signifikant unterscheiden.

Es bleibt abschließend zu bemerken, dass Kongruenz-Phänomene ähnlicher Art nach meinem Kenntnisstand zwar in Lesestudien mit 2AFC<sup>9</sup> untersucht wurden (Keung & Staub, 2018), jedoch nicht unter dem Einsatz des Visual World Paradigm. Deshalb ist es nur eingeschränkt möglich, die Ergebnisse angemessen einzuordnen. Im Falle der vorliegenden empirischen Untersuchung wäre es sicher wünschenswert, noch mehr Daten von weiteren Versuchspersonen zu erheben, um der Interpretation eine breitere Basis geben zu können. Hätte man diese breitere Basis, wäre es außerdem interessant, die Eyetracking-Daten abhängig von den Grammatikalitätsurteilen ‚korrekt‘/‚inkorrekt‘ zu analysieren (Action-contingent analysis, Tanenhaus & Trueswell, 2006, 871).

---

<sup>9</sup> *two alternative forced choice* – Versuchspersonen werden zwei Alternativen als Antwort, Lösung etc. vorgegeben, wobei sie gezwungen sind, sich für eine der beiden zu entscheiden.



## Literatur

- Allopenna, Paul D.; Magnuson, James S.; Tanenhaus, Michael K. (1998): Tracking the Time Course of Spoken Word Recognition Using Eye Movements. Evidence for Continuous Mapping Models. In: *Journal of Memory and Language* 38 (4), S. 419–439. <https://doi.org/10.1006/jmla.1997.2558>.
- Altmann, Gerry T. M.; Kamide, Yuki (1999): Incremental interpretation at verbs. Restricting the domain of subsequent reference. In: *Cognition* 73 (3), S. 247–264. Online verfügbar unter <https://www.sciencedirect.com/science/article/pii/S0010027799000591>, zuletzt geprüft am 10.09.2019.
- Bock, Kathryn; Miller, Carol A. (1991): Broken agreement. In: *Cognitive Psychology* 23 (1), S. 45–93. Online verfügbar unter <https://www.sciencedirect.com/science/article/pii/0010028591900037>, zuletzt geprüft am 10.09.2019.
- Christianson, K.; Hollingworth, A.; Halliwell, J. F.; Ferreira, F. (2001): Thematic roles assigned along the garden path linger. In: *Cognitive Psychology* 42 (4), S. 368–407. <https://doi.org/10.1006/cogp.2001.0752>.
- Christianson, Kiel (2016): When language comprehension goes wrong for the right reasons. Good-enough, underspecified, or shallow language processing. In: *Quarterly journal of experimental psychology* 69 (5), S. 817–828. <https://doi.org/10.1080/17470218.2015.1134603>.
- Conklin, Kathy; Pellicer-Sánchez, Ana; Carroll, Gareth (2018): *Eye-tracking. A guide for applied linguistics research*. Cambridge, New York, NY: Cambridge University Press.
- Cooper, Roger M. (1974): The control of eye fixation by the meaning of spoken language. A new methodology for the real-time investigation of speech perception, memory, and language processing. In: *Cognitive Psychology* 6 (1), S. 84–107. Online verfügbar unter <https://www.sciencedirect.com/science/article/pii/001002857490005X>, zuletzt geprüft am 10.09.2019.
- Crocker, Matthew W.; Pickering, Martin; Clifton, Charles (Hg.) (2000 // 2006): *Architectures and mechanisms for language processing*. Digitally print. 1st paperback version. Cambridge: Cambridge University Press.
- Demberg, Vera; Keller, Frank (2008): Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. In: *Cognition* 109 (2), S. 193–210. <https://doi.org/10.1016/j.cognition.2008.07.008>.
- Dietrich, Rainer; Gerwien, Johannes (2017): *Psycholinguistik. Eine Einführung*. Kapitel 5: Sprachverstehen. S. 163–222. 3., aktualisierte und erweiterte Auflage. Stuttgart: J.B. Metzler. <https://doi.org/10.1007/978-3-476-05494-4>.
- Drenhaus, Heiner (2012): Kapitel 7: Sprachverarbeitung. In: Barbara Höhle (Hg.): *Psycholinguistik*. 2., unveränd. Aufl. Berlin: Akad.-Verl. (Akademie-Studienbücher), S. 95–109.
- Dudenredaktion; Eisenberg, Peter; Schneider, Jan Georg (Hg.) (2011): *Duden. Richtiges und gutes Deutsch. Das Wörterbuch der sprachlichen Zweifelsfälle*. 7., vollständig überarbeitete Auflage. Berlin: Dudenverlag (Der Duden, in zwölf Bänden: das Standardwerk zur deutschen Sprache; Band 9).

- Dudenredaktion; Wöllstein, Angelika (Hg.) (2016): Duden. Die Grammatik. Unentbehrlich für richtiges Deutsch. 9., vollständig überarbeitete und aktualisierte Auflage. Berlin: Dudenverlag (Der Duden, in zwölf Bänden; das Standardwerk zur deutschen Sprache; Band 4).
- Eberhard, Kathleen M. (1997): The Marked Effect of Number on Subject–Verb Agreement. In: *Journal of Memory and Language* 36 (2), S. 147–164. <https://doi.org/10.1006/jmla.1996.2484>.
- Eberhard, Kathleen M.; Spivey-Knowlton, Michael J.; Sedivy, Julie C.; Tanenhaus, Michael K. (1995): Eye movements as a window into real-time spoken language comprehension in natural contexts. In: *Journal of Psycholinguistic Research* 24 (6), S. 409–436. <https://doi.org/10.1007/BF02143160>.
- Eisenberg, Peter; Thieroff, Rolf (2013): Grundriss der deutschen Grammatik. Band 2: Der Satz. 4., aktualisierte und überarbeitete Auflage. Stuttgart, Weimar: Verlag J.B. Metzler. <https://doi.org/10.1007/978-3-476-00743-8>.
- Ferreira, Fernanda; Henderson, John M. (1991): Recovery from misanalyses of garden-path sentences. In: *Journal of Memory and Language* 30 (6), S. 725–745. Online verfügbar unter <https://www.sciencedirect.com/science/article/pii/0749596X9190034H>, zuletzt geprüft am 10.09.2019.
- Fodor, Jerry A. (1983 // 2008): *The modularity of mind. An essay on faculty psychology*. 15. printing. Cambridge, Mass.: MIT Press (A Bradford Book).
- Franck, Julie; Vigliocco, Gabriella; Nicol, Janet (2002): Subject-verb agreement errors in French and English. The role of syntactic hierarchy. In: *Language and Cognitive Processes* 17 (4), S. 371–404. <https://doi.org/10.1080/01690960143000254>.
- Frazier, Lyn (1987): Syntactic Processing: Evidence from Dutch. In: *Natural Language and Linguistic Theory* (5), S. 519–559. <https://doi.org/10.1007/BF00138988>.
- Frazier, Lyn; Rayner, Keith (1982): Making and correcting errors during sentence comprehension. Eye movements in the analysis of structurally ambiguous sentences. In: *Cognitive Psychology* 14 (2), S. 178–210. Online verfügbar unter <https://www.sciencedirect.com/science/article/pii/0010028582900081>, zuletzt geprüft am 10.09.2019.
- Gibson, Edward (2000): The Dependency Locality Theory: A Distance-Based Theory of Linguistic Complexity. In: Alec Marantz, Yasushi Miyashita & Wayne O'Neil (Hg.): *Image, language, brain. Papers from the First Mind Articulation Project Symposium*. Cambridge, Mass.: MIT Press, S. 95–126.
- Gibson, Edward; Desmet, Timothy; Grodner, Daniel; Watson, Duane; Ko, Kara (2005): Reading relative clauses in English. In: *Cognitive Linguistics* 16 (2), S. 313–353. <https://doi.org/10.1515/cogl.2005.16.2.313>.
- grammis - Grammatisches Informationssystem (2019): Appositionen. Hg. v. Leibniz-Institut für Deutsche Sprache. Mannheim. Online verfügbar unter <https://grammis.ids-mannheim.de/systematische-grammatik/1676>, zuletzt geprüft am 25.04.2019.
- grammis - Grammatisches Informationssystem (2019): Kongruenz. Hg. v. Leibniz-Institut für Deutsche Sprache. Mannheim. Online verfügbar unter <https://grammis.ids-mannheim.de/terminologie/577>, zuletzt geprüft am 25.04.2019.

- Griffin, Zenzi M. (2004): Why Look? Reasons for Eye Movements Related to Language Production. In: John M. Henderson & Fernanda Ferreira (Hg.): The interface of language, vision, and action. Eye movements and the visual world. [workshop at Michigan State University, May 2002]. New York, NY: Psychology Press, S. 213–247.
- Grodner, Daniel; Gibson, Edward (2005): Consequences of the serial nature of linguistic input for sentential complexity. In: *Cognitive science* 29 (2), S. 261–290. [https://doi.org/10.1207/s15516709cog0000\\_7](https://doi.org/10.1207/s15516709cog0000_7).
- Haskell, Todd R.; MacDonald, Maryellen C. (2005): Constituent structure and linear order in language production. Evidence from subject-verb agreement. In: *Journal of experimental psychology. Learning, memory, and cognition* 31 (5), S. 891–904. <https://doi.org/10.1037/0278-7393.31.5.891>.
- Henderson, John M.; Ferreira, Fernanda (Hg.) (2004): The interface of language, vision, and action. Eye movements and the visual world. [workshop at Michigan State University, May 2002]. New York, NY: Psychology Press.
- Höhle, Barbara (Hg.) (2012): *Psycholinguistik*. 2., unveränd. Aufl. Berlin: Akad.-Verl. (Akademie-Studienbücher).
- Huettig, Falk; Rommers, Joost; Meyer, Antje S. (2011): Using the visual world paradigm to study language processing. A review and critical evaluation. In: *Acta psychologica* 137 (2), S. 151–171. <https://doi.org/10.1016/j.actpsy.2010.11.003>.
- Jaeger, Christoph (1992): Probleme der syntaktischen Kongruenz. Theorie und Normvergleich im Deutschen. Berlin: De Gruyter (Reihe Germanistische Linguistik, 132). <https://doi.org/10.1515/9783111676623>.
- Just, Marcel A.; Carpenter, Patricia A. (1992): A capacity theory of comprehension. Individual differences in working memory. In: *Psychological Review* 99 (1), S. 122–149. <https://doi.org/10.1037/0033-295X.99.1.122>.
- Keung, Lap-Ching; Staub, Adrian (2018): Variable agreement with coordinate subjects is not a form of agreement attraction. In: *Journal of Memory and Language* 103, S. 1–18. <https://doi.org/10.1016/j.jml.2018.07.001>.
- Lawrenz, Birgit (1993): Apposition. Begriffsbestimmung und syntaktischer Status. Zugl.: Köln, Univ., Diss., 1991 u. d. T.: Lawrenz, Birgit: Zum syntaktischen Status der Apposition in der DP/NP. Tübingen: Narr (Studien zur deutschen Grammatik, 44).
- Levy, Roger (2008): Expectation-based syntactic comprehension. In: *Cognition* 106 (3), S. 1126–1177. <https://doi.org/10.1016/j.cognition.2007.05.006>.
- Levy, Roger P.; Keller, Frank (2013): Expectation and Locality Effects in German Verb-final Structures. In: *Journal of Memory and Language* 68 (2), S. 199–222. <https://doi.org/10.1016/j.jml.2012.02.005>.
- Lewis, Richard L.; Vasishth, Shravan (2005): An activation-based model of sentence processing as skilled memory retrieval. In: *Cognitive science* 29 (3), S. 375–419. [https://doi.org/10.1207/s15516709cog0000\\_25](https://doi.org/10.1207/s15516709cog0000_25).
- Marantz, Alec; Miyashita, Yasushi; O'Neil, Wayne (Hg.) (2000): Image, language, brain. Papers from the First Mind Articulation Project Symposium. Cambridge, Mass.: MIT Press.
- Nicenboim, Bruno; Vasishth, Shravan; Gattei, Carolina; Sigman, Mariano; Kliegl, Reinhold (2015): Working memory differences in long-distance dependency resolution. In: *Frontiers in psychology* 6: Article 312. <https://doi.org/10.3389/fpsyg.2015.00312>.

- Nicol, Janet; Forster, Kenneth; Verres, C. (1997): Subject-verb agreement processes in comprehension. In: *Journal of Memory and Language* 36 (4), S. 569–587.  
<https://doi.org/10.1006/jmla.1996.2497>.
- Pearlmutter, Neal J.; Garnsey, Susan M.; Bock, Kathryn (1999): Agreement Processes in Sentence Comprehension. In: *Journal of Memory and Language* 41 (3), S. 427–456.  
<https://doi.org/10.1006/jmla.1999.2653>.
- Pritchett, Bradley L. (1992): *Grammatical competence and parsing performance*. Chicago, Ill.: Univ. of Chicago Press.
- Rickheit, Gert; Sichelschmidt, Lorenz; Strohner, Hans (2002): Psycholinguistik. Die Wissenschaft vom sprachlichen Verhalten und Erleben. Kapitel 7. Mitteilungen verstehen: Sprachrezeption. S. 93-112. Tübingen: Stauffenburg-Verl. (Stauffenburg-Einführungen, 7).
- Sanford, Anthony J.; Graesser, Arthur C. (2006): Shallow Processing and Underspecification. In: *Discourse Processes* 42 (2), S. 99–108.  
[https://doi.org/10.1207/s15326950dp4202\\_1](https://doi.org/10.1207/s15326950dp4202_1).
- Schlueter, Zoe; Williams, Alexander; Lau, Ellen (2018): Exploring the abstractness of number retrieval cues in the computation of subject-verb agreement in comprehension. In: *Journal of Memory and Language* 99, S. 74–89.  
<https://doi.org/10.1016/j.jml.2017.10.002>.
- Tabor, Whitney; Galantucci, Bruno; Richardson, Daniel (2004): Effects of merely local syntactic coherence on sentence processing. In: *Journal of Memory and Language* 50 (4), S. 355–370. <https://doi.org/10.1016/j.jml.2004.01.001>.
- Tanenhaus, Michael K.; Spivey-Knowlton, Michael J.; Eberhard, Kathleen M.; Sedivy, Julie C. (1995): Integration of visual and linguistic information in spoken language comprehension. In: *Science* (268), S. 1632–1634. Online verfügbar unter <https://science.sciencemag.org/content/268/5217/1632>, zuletzt geprüft am 10.09.2019.
- Tanenhaus, Michael K.; Spivey-Knowlton, Michael J.; Hanna, Joy E. (2000 // 2006): Modeling thematic and discourse context effects with a multiple constraints approach: Implications for the architecture of the language comprehension system. In: Matthew W. Crocker, Martin Pickering & Charles Clifton (Hg.): *Architectures and mechanisms for language processing*. Digitally print. 1st paperback version. Cambridge: Cambridge University Press, S. 90–118.
- Tanenhaus, Michael K.; Trueswell, John C. (2006): Eye Movements and Spoken Language Comprehension. In: Matthew J. Traxler & Morton A. Gernsbacher (Hg.): *Handbook of psycholinguistics*. 2nd Edition. Amsterdam: Elsevier, S. 863–900.
- Tanner, Darren; Nicol, Janet; Brehm, Laurel (2014): The time-course of feature interference in agreement comprehension. Multiple mechanisms and asymmetrical attraction. In: *Journal of Memory and Language* 76, S. 195–215.  
<https://doi.org/10.1016/j.jml.2014.07.003>.
- Traxler, Matthew J.; Gernsbacher, Morton A. (Hg.) (2006 // 2011): *Handbook of psycholinguistics*. 2nd Edition. Amsterdam: Elsevier.
- Vasishth, Shravan (2017): *Introduction to Sentence Comprehension*. Online verfügbar unter <http://www.ling.uni-potsdam.de/~vasishth/courses/pdfs/LitomyslVasishth.pdf>, zuletzt geprüft am 16.06.2019.

- Vasishth, Shravan; Lewis, Richard L. (2006): Argument-Head Distance and Processing Complexity. Explaining both Locality and Antilocality Effects. In: *Language* 82 (4), S. 767–794. <https://doi.org/10.1353/lan.2006.0236>.
- Wagers, Matthew W.; Lau, Ellen F.; Phillips, Colin (2009): Agreement attraction in comprehension. Representations and processes. In: *Journal of Memory and Language* 61 (2), S. 206–237. <https://doi.org/10.1016/j.jml.2009.04.002>.
- Wegerer, Martina (2012): Die Numeruskongruenz von Subjekt und finitem Verb im Deutschen. Untersuchung der grammatischen Entscheidungsprozeduren bei zweifelhaften Kongruenzrelationen. Dissertation. Universität Wien. Online verfügbar unter [http://othes.univie.ac.at/18319/1/2012-02-01\\_9800525.pdf](http://othes.univie.ac.at/18319/1/2012-02-01_9800525.pdf), zuletzt geprüft am 10.09.2019.

# Anhang

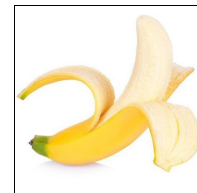
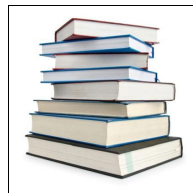
## Sätze und Bilder der Criticals/Controls

1. Beides, Glas und Flasche, ist in einem Restaurant zu finden. (Vers1)  
Beides, Glas und Flasche, sind in einem Restaurant zu finden. (Vers2)  
Beides, Glas und Flaschen, sind in einem Restaurant zu finden. (Vers3)  
Viele Gläser und Flaschen sind in einem Restaurant zu finden. (Vers4)



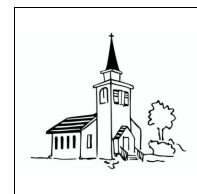
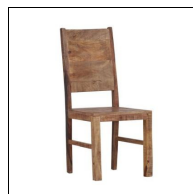
Distraktor

2. Beides, Buch und Regal, besitzt rechteckige Ecken. (Vers1)  
Beides, Buch und Regal, besitzen rechteckige Ecken. (Vers2)  
Beides, Buch und Regale, besitzen rechteckige Ecken. (Vers3)  
Alle Bücher und Regale besitzen rechteckige Ecken. (Vers4)



Distraktor

3. Beides, Tisch und Stuhl, steht auf vier Beinen. (Vers1)  
Beides, Tisch und Stuhl, stehen auf vier Beinen. (Vers2)  
Beides, Tisch und Stühle, stehen auf vier Beinen. (Vers3)  
Alle Tische und Stühle stehen auf vier Beinen. (Vers4)



Distraktor

4. Beides, Auto und Fahrrad, bevölkert deutsche Innenstädte. (Vers1)  
 Beides, Auto und Fahrrad, bevölkern deutsche Innenstädte. (Vers2)  
 Beides, Auto und Fahrräder, bevölkern deutsche Innenstädte. (Vers3)  
 Viele Autos und Fahrräder bevölkern deutsche Innenstädte. (Vers4)



Distraktor

5. Beides, Schiff und Baumstamm, schwimmt auf Wasser. (Vers1)  
 Beides, Schiff und Baumstamm, schwimmen auf Wasser. (Vers2)  
 Beides, Schiff und Baumstämme, schwimmen auf Wasser. (Vers3)  
 Alle Schiffe und Baumstämme schwimmen auf Wasser. (Vers4)



Distraktor

6. Beides, Nagel und Schraube, ist aus Eisen. (Vers1)  
 Beides, Nagel und Schraube, sind aus Eisen. (Vers2)  
 Beides, Nagel und Schrauben, sind aus Eisen. (Vers3)  
 Alle Nägel und Schrauben sind aus Eisen. (Vers4)



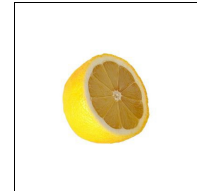
Distraktor

7. Beides, Glühbirne und Kerze, strahlt Licht aus. (Vers1)  
 Beides, Glühbirne und Kerze, strahlen Licht aus. (Vers2)  
 Beides, Glühbirne und Kerzen, strahlen Licht aus. (Vers3)  
 Alle Glühbirnen und Kerzen strahlen Licht aus. (Vers4)



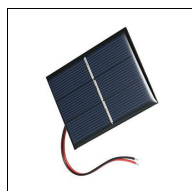
Distraktor

8. Beides, Mütze und Handschuh, schützt vor Kälte. (Vers1)  
 Beides, Mütze und Handschuh, schützen vor Kälte. (Vers2)  
 Beides, Mütze und Handschuhe, schützen vor Kälte. (Vers3)  
 Alle Mützen und Handschuhe schützen vor Kälte. (Vers4)



Distraktor

9. Beides, Windrad und Solarzelle, erzeugt Strom. (Vers1)  
 Beides, Windrad und Solarzelle, erzeugen Strom. (Vers2)  
 Beides, Windrad und Solarzellen, erzeugen Strom. (Vers3)  
 Alle Windräder und Solarzellen erzeugen Strom. (Vers4)



Distraktor

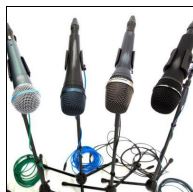
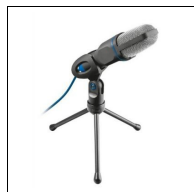
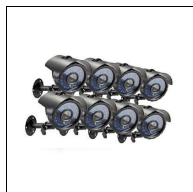


10. Beides, Kamera und Mikrofon, dient der Überwachung. (Vers1)

Beides, Kamera und Mikrofon, dienen der Überwachung. (Vers2)

Beides, Kamera und Mikrofone, dienen der Überwachung. (Vers3)

Viele Kameras und Mikrofone dienen der Überwachung. (Vers4)



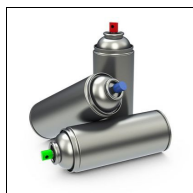
Distraktor

11. Beides, Plastiktüte und Sprühdose, schadet dem Klima. (Vers1)

Beides, Plastiktüte und Sprühdose, schaden dem Klima. (Vers2)

Beides, Plastiktüte und Sprühdosen, schaden dem Klima. (Vers3)

Viele Plastiktüten und Sprühdosen schaden dem Klima. (Vers4)



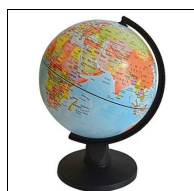
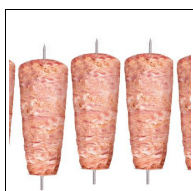
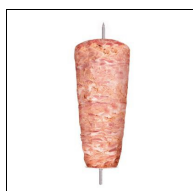
Distraktor

12. Beides, Döner-Spieß und Globus, dreht sich um die eigene Achse. (Vers1)

Beides, Döner-Spieß und Globus, drehen sich um die eigene Achse. (Vers2)

Beides, Döner-Spieß und Globusse, drehen sich um die eigene Achse. (Vers3)

Alle Döner-Spieße und Globusse drehen sich um die eigene Achse. (Vers4)



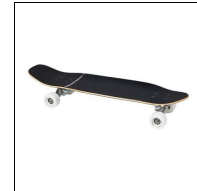
Distraktor

13. Beides, Hemd und Bluse, hängt an der Kleiderstange. (Vers1)  
 Beides, Hemd und Bluse, hängen an der Kleiderstange. (Vers2)  
 Beides, Hemd und Blusen, hängen an der Kleiderstange. (Vers3)  
 Alle Hemden und Blusen hängen an der Kleiderstange. (Vers4)



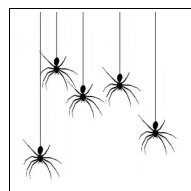
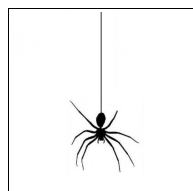
Distraktor

14. Beides, Kette und Ring, lockt Einbrecher an. (Vers1)  
 Beides, Kette und Ring, locken Einbrecher an. (Vers2)  
 Beides, Kette und Ringe, locken Einbrecher an. (Vers3)  
 Viele Ketten und Ringe locken Einbrecher an. (Vers4)



Distraktor

15. Beides, Lampe und Spinne, hängt von der Decke herab. (Vers1)  
 Beides, Lampe und Spinne, hängen von der Decke herab. (Vers2)  
 Beides, Lampe und Spinnen, hängen von der Decke herab. (Vers3)  
 Viele Lampen und Spinnen hängen von der Decke herab. (Vers4)



Distraktor

16. Beides, Radiergummi und Stift, befindet sich in einem Mäppchen. (Vers1)

Beides, Radiergummi und Stift, befinden sich in einem Mäppchen. (Vers2)

Beides, Radiergummi und Stifte, befinden sich in einem Mäppchen. (Vers3)

Viele Radiergummis und Stifte befinden sich in einem Mäppchen. (Vers4)



Distraktor

---

Auf dem beigelegten Datenträger ist folgendes zu finden:

- Bilder: Criticals, Filler, Intro-Items
- Alle Sätze als Audio-Dateien
- Liste der Tastendrucke
- Liste der Filler-Sätze mit Positionsvariationen
- Instruktionen für die Versuchspersonen