

Prof. Dr. Gerhard Strube

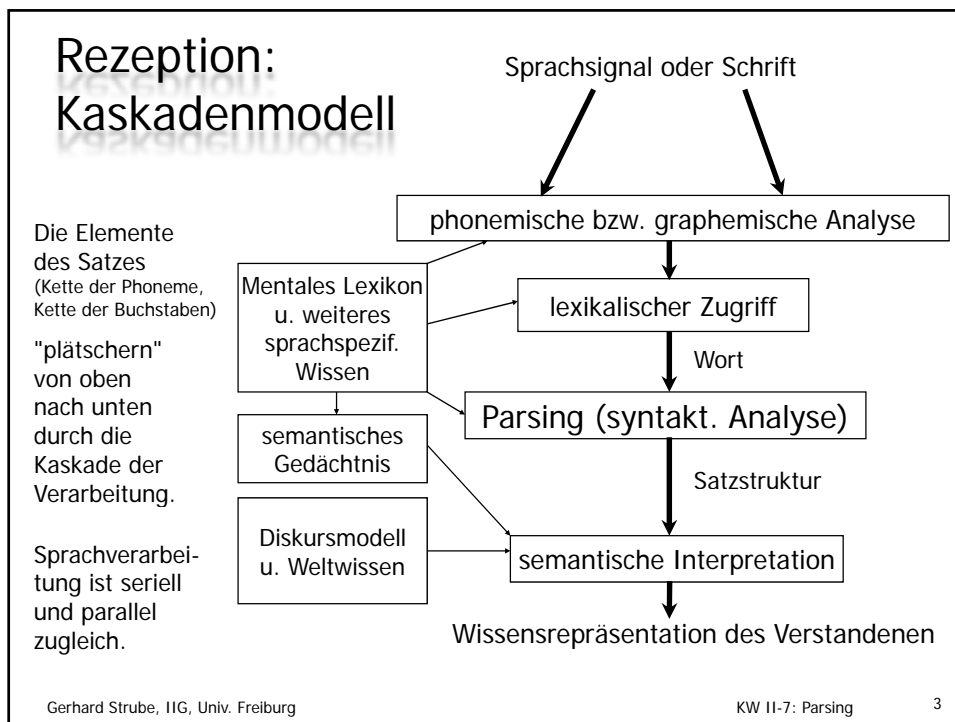
Allg. Psychologie I, Teilgebiet: Sprache

3: Syntaktische Analyse (Parsing) und Sprachverstehen

Grammatisches Wissen
Parsing: Reanalysemodelle und Anbindungspräferenzen
Empirische Belege: Blickbewegung und ERP
Argumentstruktur und thematische Rollen
Mentale Modelle beim Verstehen: Diskurs- und Situationsmodell

Wie steckt Grammatik in unseren Köpfen?

- Ein Beispiel für grammatisches Wissen:
 - Der Fährmann hat die Reisenden über**g**esetzt. ... hat ihr was vor**g**esetzt.
 - Der Redakteur hat den Aufsatz über**s**etzt. ... hat sie ver**s**etzt.
 - Gibt es hierbei Regularitäten?
 - Ja! Betonung (1. Silbe) + "ge" + Trennbarkeit
 - Der Fährmann setzt über, der Redakteur übersetzt / setzt ... vor, versetzt
- Wie ist grammatisches Wissen mental repräsentiert?
 - Grammatisches Wissen ist implizites Wissen:
Wir können in der Regel nicht darüber Auskunft geben.
 - NB: Was wir in der Schule über die Grammatik unserer Muttersprache gelernt haben, hat nur bedingt mit grammatischem Wissen zu tun.
 - Grammatisches Wissen besteht vorwiegend aus lexikalischem Wissen und Generalisierungen, die z.B. als Regeln modelliert werden können.
 - "Konstruktionen" im Sinne der sog. Construction Grammar
 - Art. "Konstruktionsgrammatik" (Wikipedia)

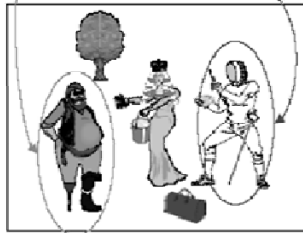


Unmittelbarkeit der Verarbeitung

- aus dem Kaskadenmodell folgt:
 - Jedes neue Element wird sofort verarbeitet
 - Bei normalem Sprechtempo also bis zu 4 Wörtern pro Sekunde
 - Bei schnellem Lesen ist der Prozess ressourcenbegrenzt
 - Bei langsamem Sprechen ist der Prozess datenbegrenzt
- Empirische Belege
 - Effekte experimentell induzierter Anomalien im EEG bereits 180-600 ms später erkennbar (davon 100-150 ms für lexikal. Zugriff)
 - Dasselbe gilt für Blickbewegungsstudien: erhöhte Lesezeiten und häufige Blickregressionen
 - Bei langsamer Vorgabe Wort für Wort wird jeweils die gesamte Verarbeitungskette durchlaufen, bevor das nächste Wort dargeboten wird (z.B. Visual World Paradigm)

Visual World Paradigm

"Die Prinzessin wäscht offensichtlich den Pirat."
"Die Prinzessin malt offensichtlich der Fechter."



Knoeferle et al. (2005)

- **Versuchsablauf**
 - Vpn sehen das linke Bild
 - nach 1 s wird der Satz langsam gesprochen (Dauer ca. 6 s)
 - Bild rechts: Kontrollbedingung
- **Ergebnisse und Interpretation**
 - Beim SVO-Satz (Die Prinzessin wäscht...) wird, während das Adverb gehört wird, der Blick mehrheitlich auf den Piraten (Objekt = Patient) gerichtet.
 - Beim OVS-Satz (Die Prinzessin [Akk.] malt ...) wird hingegen der Blick häufiger auf den Fechter (Subjekt = Agent) gerichtet.
 - Die Interpretation macht also vom Wissen um das Bild Gebrauch.
 - Dies gilt nur für die langsame Vorgabe, bei schneller Verarbeitung (Lesen) beginnt die syntaktische Analyse, bevor die Wortbedeutungen verfügbar sind!

Gerhard Strube, IIG, Univ. Freiburg

KW II-7: Parsing

5

Informationen für die syntaktische Analyse (Parsing)

- **Mentales Lexikon**
 - syntaktische Kategorie der Wörter: N, Adj, V, Det ...
 - Argumantstruktur für Verben, Präpositionen usw. (Spezifikation der notwendigen und optionalen Komplemente)
- **Morphologische Information**
 - Deklination und Konjugation
 - Komposita (semantisch transparent od. opak)
- **Wortstellung** (Reihenfolge der bisher verarbeiteten Wörter)
- **Prosodie des Satzes**
 - Akzentuierung (z.B. zur Fokus-Markierung)
 - Intonation zur Kennzeichnung (z.B. als Frage)
 - Pausenstruktur u. Intonation zur Gliederung
- **Semantik (Wortbedeutung):** nur bei interaktiven Parsern!
 - nicht bei modularen Parsern, weil sie erst später zur Verfügung steht

Gerhard Strube, IIG, Univ. Freiburg

KW II-7: Parsing

6

Wie werden Ambiguitäten bewältigt?

Das serielle Modell des Parsings

- In einer "ersten Analyse" wird ausschließlich lexikalisches und syntaktisches Wissen verwendet.
 - Parsing-Prinzipien erzeugen die präferierte syntaktische Struktur
 - Gewählt wird die häufigste Variante eines mehrdeutigen Wortes
 - Es wird die einfachste Struktur erzeugt ("minimal attachment")
- Falls die Analyse nicht "aufgeht", wird eine Reanalyse eingeleitet.
 - Reanalyse auch später, wenn die Integration mit der Semantik nicht klappt
- Reanalyse kostet Zeit
 - nicht immer intuitiv verspürbar, aber messbar, ca. 50-100 ms mehr bei Blickbewegungsmessung
 - Auch die Reanalyse führt nicht immer zu einem korrekten Ergebnis.
- Holzwegeffekt (garden-path effect)
 - wenn auch die Reanalyse scheitert, wird der Konflikt bewusst

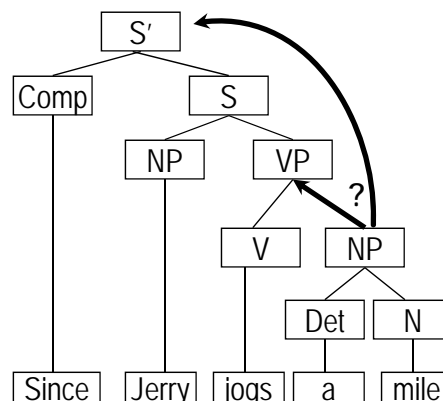
Gerhard Strube, IIG, Univ. Freiburg

KW II-7: Parsing

7

Ein Beispiel für Parsing-Prinzipien: "rechts anschließen" (late closure)

- Prinzip "Late Closure"
 - Das älteste Parsing-Prinzip (Kimball, 1973)
 - Binde eine neue Einheit (Wort, Konstituente) nach Möglichkeit an die zuletzt konstruierte Einheit.
 - Since Jerry jogs a mile seems a short distance to him.



Gerhard Strube, IIG, Univ. Freiburg

KW II-7: Parsing

8

Wann wird Reanalyse nötig?

- Wenn das Parsing nicht "aufgeht"
 - Beispiel: der Doktorand der Professorin
 - präferiert: der Professorin als Genitiv an die NP binden
 - Dass (der Dokt. d. Prof.) Urlaub abgetrotzt hat,...
 - Am Disambiguierungspunkt (abgetrotzt) wird die Reanalyse (der Dokt.) (der_{Dat} Prof.) erzwungen.
- Wenn die Integration mit der Semantik nicht klappt
 - Beispiel: Polizei erschießt 19jährigen mit Samuraischwert
 - präferiert: mit Samuraischwert als Instrument
 - scheint zunächst zu klappen, weil S. eine Waffe ist
 - klappt dann aber doch nicht, weil S. keine Schusswaffe ist
 - Also wird Reanalyse (19j. mit S.) erzwungen.

Wie ist Reanalyse messbar?

- Reanalyse wird von Lesern in der Regel nicht bemerkt.
- Vorgabe von Texten am Computer.
 - Auf Zeilenumbruch achten (Artefakte!)
 - Interessierende Regionen sollten nicht direkt am Satzende stehen (sentence wrap-up effect = erhöhte Lesezeit am Satzende)
- Vpn lesen, die Blickbewegungen werden registriert.
- Man kann im Labor beobachten:
 - Am Ende ambiger Regionen, d.h. am Disambiguierungspunkt, treten bei nicht präferierten Parsing-Varianten signifikant erhöhte Lesezeiten auf (Indiz für Reanalyse).
 - Während des Lesens ambiger Regionen gibt es dagegen keine erhöhten Lesezeiten.

Lokale Ambiguität und "garden path"

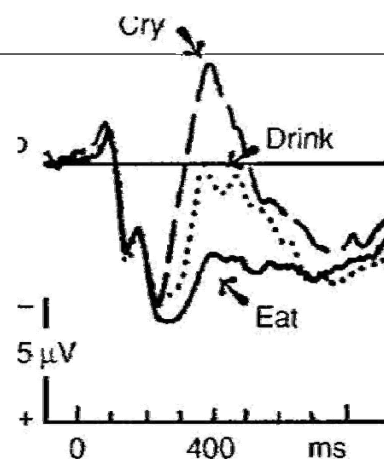
- Der garden path effect (Holzweg-Effekt)
 - tritt auf, wenn eine (meist durch eine lokale Ambiguität ausgelöste) Reanalyse scheitert.
 - Das Nicht-Verstanden-Haben wird bewusst.
- Beispiele:
 - The horse raced past the barn fell. (Bever, 1970)
 - Dass den Reportern die Prinzessin zu fragen erlaubt { hatte | worden war }, erstaunt mich. (Bader, 1991)
- Globale Ambiguität führt nicht zum Holzweg-Effekt
 - Eine der möglichen Lesarten wird "verstanden", Alternativen werden nicht analysiert: Vpn sind erstaunt, dass es noch andere Lesarten gibt, z.B. Müllers sahen die Störche auf ihrem Flug nach Afrika.

Gerhard Strube, IIG, Univ. Freiburg KW II-7: Parsing

11

Elektrophysiologische Studien zur Sprachverarbeitung

- Ereigniskorrelierte Potentiale (ERP)
 - Darbietung akustisch oder visuell-stationär (wortweise, Blickbewegungen unerwünscht), fester Takt 300-600 ms
 - Vorteil: Online-Technik, EEG während der Sprachverarbeitung
 - Vorteil: sehr gute zeitliche Auflösung (ca. 10-20 msec)
 - Nachteil: schlechte räumliche Auflösung (trotz 20-256 Elektroden)
 - Nachteil: 30-50 Durchgänge (mit demselben Satz!) nötig für verlässliche Mittelung des Signal



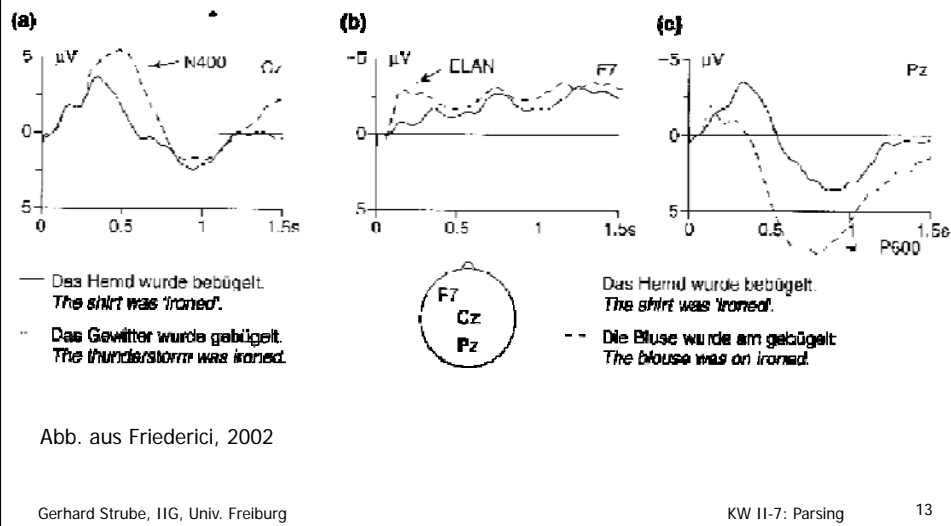
„The pizza was too hot to ...“
N 400, Abb. aus Kutas, Lindamood & Hillyard, 1984

Gerhard Strube, IIG, Univ. Freiburg

KW II-7: Parsing

12

Elektrophysiologische Befunde



Interpretation der ERP-Phänomene

- N-400
 - Dies ist ein Effekt, der immer auftritt, wenn kontextuelle Erwartungen verletzt werden
 - Im Beispiel ist es die Kombination von Gewitter und bügeln (vgl. früheres Beispiel: Pizza too hot to cry)
- ELAN (Early Left Anterior Negativity), ca. 180 ms
 - indiziert Probleme beim Aufbau der initialen Phrasenstruktur
- P-600 (Late positive shift)
 - indiziert Reanalyse bei syntaktisch-semantischer Integration

Fazit: zentrale Befunde zum Parsing

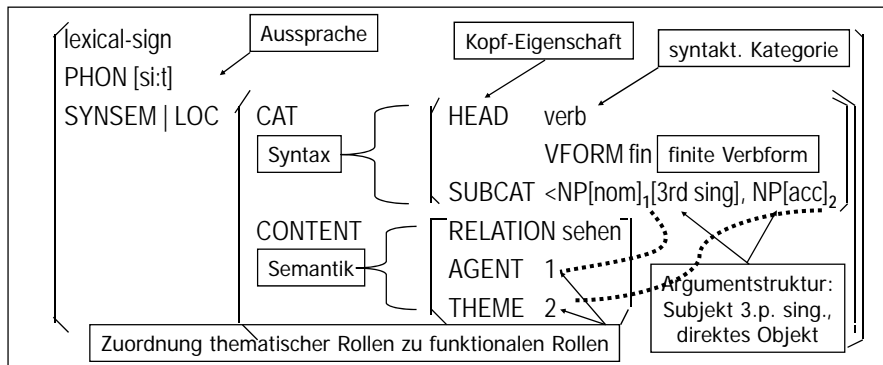
- Information wird zum frühest möglichen Zeitpunkt wirksam.
- Das Parsing verläuft inkrementell (Wort für Wort).
- Die (Re-)Konstruktion der Satzstruktur erfolgt gemäß Prinzipien, die nur von grammatischem Wissen abhängen.
- Bei Mehrdeutigkeiten (Ambiguitäten) wird oft nur eine Strukturvariante weiterverfolgt (Erste Analyse, first analysis). Scheitert diese aus syntaktischen oder semantischen Gründen, kommt es zur Reanalyse.
- Reanalyse funktioniert in den meisten Fällen gezielt, schnell und effizient.
- Scheitert die Reanalyse, so kommt es zum Holzweg-Effekt (garden path effect): Das Nicht-Verstehen wird bewusst
- Zuweilen wird eher "repariert" als reanalysiert, dann wird falsch verstanden.

Schnittstelle zur Semantik: die thematische Struktur:

- Tiefenkasus (Theta-Rollen, θ -Rollen)
 - spezifizieren die thematischen Rollen der Komplemente (bes. beim Verb),
 - z.B. Agent, Rezipient, Thema, Instrument, Lokation, Zeit usw.,
 - im Gegensatz zu deren sog. funktionalen Rollen, z.B. Subjekt.
- Lexikalischer Subkategorisierungs-rahmen von Verben, sog. Argumentstruktur
 - spezifiziert die θ -Rollen von Verb-Komplementen und deren syntaktische Funktionen (z.B. Agent = Subjekt, Patient = direktes Objekt, usw.)
- Beispiel: GEBEN
 - Agent = Subjekt (im Nominativ, Numerus muss mit dem Verb übereinstimmen, sog. agreement)
 - Thema (d.h., was gegeben wird) = direktes Objekt (im Akkusativ)
 - Rezipient = indirektes Objekt (im Dativ)
- Moderne Grammatiktheorien...
 - z.B. LFG (Bresnan 1982), HPSG (Pollard & Sag 1994)
 - verlagern einen großen Teil der grammatischen Regularitäten ins mentale Lexikon

Inhalt von Einträgen im Lexikon bei lexikalisierten Grammatiken

- Ein Beispiel aus der HPSG: "(er, sie, es) sieht"
(Head-Driven Phrase-Structure Grammar, Pollard & Sag, 1994):



Gerhard Strube, IIG, Univ. Freiburg

KW II-6: Syntax

17

Argumentstruktur als Grundlage der Interpretation von Äußerungen

- Die Bedeutung von Verbkomplementen
 - ist durch die thematischen Rollen markiert (Agent, Thema...)
- Syntax codiert thematische Rollen...
 - ... als funktionale Rollen (Subjekt, Objekte...)
 - entweder durch die Reihenfolge (SVO im Englischen)
 - oder durch die Kasusmorphologie

Beispiel:

- Der Junge sieht das Mädchen mit dem Fernglas
- "sehen" hat die Codierung:
 - Subjekt = Agent (wer sieht)
 - Thema = direktes Objekt (was gesehen wird)
- "Gedanke", dargestellt als ausgefülltes thematisches Schema:

- Ereignis: sehen
- Ort: ?
- Zeit: Gegenwart
- Agent: der Junge
- Thema: das Mädchen mit...

Gerhard Strube, IIG, Univ. Freiburg

Sprachpsychologie 6

18

Was muss repräsentiert werden?

- Diskursmodell
 - bei Gesprächen: mentales Modell der Gesprächssituation
 - einschließlich eines mentalen Modells des Gesprächspartners
 - immer: Modell des Diskurses/Textes
 - speziell:
Erwähnungsreihenfolge (zur Auflösung anaphorischer Referenzen)
- Situationsmodell
 - mentales Modell des Textinhalts
 - de-indexiertes Modell
 - z.B. Geschichte mit Hauptpersonen, Lokalitäten, etc.
 - bei Lehrtexten (expository texts):
mentale Modelle der im Text dargelegten Inhalte

Referenzobjekt (RefO)

- RefOs sind wesentlicher Bestandteil von Diskurs- und Situationsmodellen
- Ein RefO kann sein:
 - eine Person, ein Gegenstand, ein Ort, usw.
- Beispiel: Daniel denkt "Ach, würde Lara mich doch küssen"
 - RefO_1: Daniel, RefO_2: Lara
 - Propositionaler Gehalt: P1: (küssen, RefO_2, RefO_1)
 - Intentionale Einstellung dazu: P2: (wünschen, RefO_1, P1)
- Das Diskursmodell repräsentiert die Salienz der RefOs
 - Hauptpersonen vor Nebenrollen
 - auch die Erwähnungsreihenfolge ist wichtig
- Das Situationsmodell repräsentiert Relationen zwischen RefOs
 - auch bei Lehrtexten, z.B. (part-of, Motor, Zylinder)

Auflösung anaphorischer Referenzen: Unmittelbare Integration

- Marslen-Wilson & Tyler (1980; Tyler & Marslen-Wilson, 1982):
 - As Philip was walking back from the shop he saw an old woman trip and fall flat on her face in the street. She seemed unable to get up again.
 - {Philip ran | He ran | Running} towards {HIM | HER} ...
- Ergebnisse (Benennungslatenzen für HIM/HER):

Anapher	passend	unpassend	Diff.
Eigennamen	379 msec	429 msec	50 msec
Pronomen	385 msec	434 msec	49 msec
"Null"-A.	384 msec	420 msec	36 msec
- Unter allen Bedingungen wird das passende Pronomen (her) schneller benannt: Beleg für sofortige Integration in mentales Modell.
- Gilt aber nur für den aktuellen Hauptakteur! (Garrod & Sanford, 1985)

Gerhard Strube, IIG, Univ. Freiburg

KW 2: 9. Text & Diskurs

21

Inferenzen beim Textverstehen: Welche, wann, auf welcher Grundlage?

- Spontan werden nur solche Schlüsse gezogen, die für die Textkohäsion erforderlich sind (bridging inferences), sowie offenkundige Inferenzen
 - Online-Methoden: Priming, Satzlesezeiten, Antwortzeiten (auf Fragen)
 - Beispiel 1: Am Flughafen luden wir das Gepäck aus dem Taxi. Im Gedränge war es schwer, mit den Koffern durchzukommen. (Brückeninferenzen: Koffer sind Teil des Gepäcks; Gedränge herrscht im Flughafen.)
 - Beispiel 2: (prime): The director and the cameraman were ready to start shooting when suddenly the actress fell from the 14th floor. (target): dead (Inferenz als Ursache eines positiven Priming-Effekts: Ratcliff & McKoon, 1992)
- Basis für Inferenzen können auch analoge Repräsentationen sein (z.B. räumliche Distanzen in mentalen Modellen).
- Für das Lernen aus Texten sind weitere Inferenzen nötig (Elaboration)

Gerhard Strube, IIG, Univ. Freiburg

KW 2: 9. Text & Diskurs

22

Der Selbsterklärungseffekt

(Self-explanations: Chi, Bassok, Lewis, Reimann & Glaser, 1989)

- Was unterscheidet gute von schlechten Lernern?
 - Methode
 - Lehrtext lesen, "lautes Denken"
 - Mediansplit nach Lerngewinn
 - Post-hoc-Analyse der verbalen Protokolle
 - Ergebnis
 - Gute Lerner generieren Erklärungen des im Text Berichteten für sich selbst.
- Was sind Selbsterklärungen (SE)?
 - Konstruktion von Zusammenhängen, die der Text implizit enthält
 - unter Rekurs auf den Text (z.B. auf frühere Textabschnitte)
 - auch Rekurs auf eigenes Vorwissen
 - SE sind aufwendig: Haben nur gute Leser Ressourcen für SE frei?
- Nur **richtige** Selbsterklärungen helfen!
 - Wenn Lerner unzutreffende SE generieren, führt sie das in die Irre
 - "schlechte" Leser verlassen sich zu sehr auf ihr u.U. falsches Vorwissen (z.B. Wilkin, 1997)

Gerhard Strube, IIG, Univ. Freiburg

KW 2: 9. Text & Diskurs

23

Warum werden Lehrtexte so oft falsch interpretiert?

- Unterschiedliche Wissensstrukturen bei Autoren und Lesern
 - Autoren: Experten mit hoch differenzierter Struktur
 - "[Their knowledge structure is] very well developed: large, densely interconnected, and highly structured. But it is also highly proceduralized and highly automatized; these characteristics tend to make the content of the expert's mental structure opaque even to the expert" (Britton, 1994, 661)
 - Leser: ohne Vorwissen, oder mit unzutreffendem und kaum strukturiertem Vorwissen
- Die Illusion des Verstehens (Bromme & Jucks, 2004)
 - Experten-Laien-Kommunikation führt oft zum bloßen Eindruck, alles verstanden zu haben.
 - Beim Lesen kann mangelnde Aufmerksamkeit zur Illusion d. V. führen. Abhilfe: Elaborationen

Gerhard Strube, IIG, Univ. Freiburg

KW 2: 9. Text & Diskurs

24