

Zwölfte Sitzung: Klausurrelevante Wiederholung

Berry Claus

Computerlinguistik
Universität des Saarlandes

[Klausur]

Termin: nächste Woche, 18. Juli, 9:15 Uhr

Raum: wird noch bekannt gegeben → **Web-Seite und e-mail!**

Dauer: 90 Minuten

Format: offene Fragen

Zwei Teile

Teil 1 (40% der Punkte): 8 kurze Fragen (~ 5 Minuten pro Frage)

Teil 2 (60% der Punkte): 3 längere Fragen (~ 15 Minuten pro Frage)

Auswahlmöglichkeit: 4 Fragen → 3 auswählen

Erlaubtes Hilfsmittel: DIN A4 Papier (doppelseitig) mit **selbst geschriebenen** Notizen

[Grundlagen]

Sprache vs. tierische Kommunikationssysteme

Merkmale von Sprache (Semantizität, Arbitrarität, Raum-/Zeit-Unabhängigkeit, Kreativität)

tierische Kommunikationssystem ≠ Sprache

„Sprechende“ Schimpansen

Schimpansen sind in der Lage, Assoziationen zwischen Bezeichnungen und Objekten/Aktivitäten zu lernen

aber: (bisher) keine Evidenz für den Erwerb von Sprache

Biologische Voraussetzungen für Spracherwerb

angeborene Ausstattung die Spracherwerb ermöglicht? (Chomskys Language Acquisition Device)

bei den meisten Menschen: linke Hemisphäre = dominant

Broca-Aphasie (Beeinträchtigung der Sprachproduktion: zögerliche Sprechweise, Wortfindungsstörungen, eingeschränkte Grammatik)

Wernicke-Aphasie (Starke Beeinträchtigung des Sprachverstehens, flüssiges Sprechen, wegen Wortverwechslungen, -verdrehungen und -neuschöpfungen häufig unverständlich)

[Grundlagen]

Kognitive Voraussetzungen

Kognitionshypothese: Kognitive Fähigkeiten sind Voraussetzung für Spracherwerb (Objektpermanenz und Benennungsexplosion)

aber: keine Korrelation zw. Objektpermanenz und sprachlicher Entwicklung;

Kinder mit Williams-Beuren-Syndrom (durchschnittlicher IQ = 50)

können flüssig und korrekt sprechen

Sapir-Whorf-Hypothese

Sprache bestimmt Denken

Linguistischer Determinismus (Form und Charakteristika der Sprache bestimmen Gedanken, Erinnerung und Wahrnehmung)

Linguistischer Relativismus (unterschiedliche Sprachen erzeugen unterschiedliche kognitive Strukturen)

Evidenz nur für schwache Version der Hypothese (Unterschiede zwischen Sprachen beeinflussen bestimmte Prozesse, bei denen sprachliche Kodierung bedeutsam ist)

[Grundlagen]

Sprachverarbeitende Computerprogramme

ELIZA (festgelegte Antworten auf bestimmte Schlüsselwörter, z.B. Vater; Satzerkennung über Mustererkennung)

SHRDLU (Blockswelt; „Verständnis“ nur für solche Äußerungen zu denen SHRDLU Handlungen ausführen konnten; Syntaktische Analyse nur für Blocksweltsätze)

Augmented Transition Networks (ATN) (top-down-Parsing → Start = Netzwerk für Satz; korrekte Analyse der syntaktischen Struktur von jedem korrekten Satz)

ATN als Modell für menschliches Parsing?

Sätze können mit unterschiedlichen syntaktischen Kategorien beginnen

→ ATN: häufiges backtracking

Menschliches Parsing: bottom-up-Komponente

[Grundlagen]

Konnektionismus

Modellierung kognitiver Prozesse mit künstlichen neuronalen Netzen
parallel verteilte Verarbeitung: miteinander vernetzte Einheiten
keine festgelegte Regeln

Aktivierung

zentrales Konzept für viele Modelle (insbesondere konnektionistische)
sich ständig verändernde Größe (Zu- und Abnahme von Aktivierung)
Aktivierungsausbreitung von hochaktivierten zu benachbarten Einheiten

Modularität in der Sprachverarbeitung

Modul = in sich geschlossene Menge von Prozessen (Prozesse innerhalb eines Moduls arbeiten unabhängig von Prozessen anderer Module)

Verschiedene Ebenen der Sprachverarbeitung (Worterkennung, syntaktische Analyse, Bedeutungsbestimmung)

→ Welche Beziehung besteht zwischen den Ebenen?

Interaktive Modelle → Ebenen beeinflussen sich

Autonome Modelle → Sprachverarbeitung ist modular

[Psycholinguistische Methoden]

Zeitmessung

Lesezeiten (satzweise / wortweise Präsentation)

Reaktionszeiten, z.B. Lexikalische Entscheidungsaufgabe, Benennung

Priming

Beeinflusst die Präsentation eines *primes* (Wort, Satz, Bild) die Verarbeitung eines nachfolgend präsentierten *targets*?

Blickbewegungsmessung

beim Lesen und beim Betrachten von Bildern + Hören von Äußerungen

Messung von Aktivierung im Gehirn

Ereigniskorrelierte Potentiale (EKP)

Bildgebende Verfahren

[Experimentelle Methoden]

Experiment

systematische Manipulation von experimentellen Bedingungen
→ Auswirkung auf Verhalten

Unabhängige Variable (= Faktor)

experimentell manipulierte Variable

Ausprägungen der Variable = (Faktor)Stufen

Faktorstufenkombinationen = Bedingungen (bei nur einem Faktor:

Bedingungen = Faktorstufen)

Abhängige Variable

erhobene/gemessene Größe (z.B. Reaktionszeit bei Benennung)

Versuchsplan/Design

→ Anzahl Faktoren

→ Anzahl Stufen

→ *between-subjects* oder *within-subject* manipulierte Faktoren

[Experimentelle Methoden]

between-subjects

n Bedingungen → n Probandengruppen
jede Probandengruppe wird einer der Bedingungen zugeordnet
→ jeder Pb wird nur unter einer der Bedingungen getestet
Nachteil: Bedingungen unterscheiden sich nicht nur in experimenteller Manipulation → Effekte von Unterschieden zwischen Pbn

within-subject

jeder Pb wird unter jeder experimentellen Bedingung getestet
Vorteil: Bedingungen unterscheiden sich nicht in Bezug auf systematische Unterschiede zwischen Pbn
potentielle Nachteile: Übungs-, Ermüdungs-, Reihenfolgeeffekte (können verhindert/kontrolliert werden)

mixed design

Kombination von between-subjects und within-subject manipulierten Faktoren

[Experimentelle Methoden]

Untersuchungsmaterial

Experimentelle Items

Manipulation von Bedingungen über experimentelle Items; Optimum: bis auf Bedingungsmanipulation keine Unterschiede zwischen Versionen; ausgeglichen in Bezug auf Störvariablen

Filler-Items

dienen zur Verschleierung der untersuchten Fragestellung

Zuordnung von Bedingungen zu Items und Pbn

bei within-subject Faktoren

jeder Pb erhält jedes experimentelle item nur in einer der Versionen
Zuteilung?

randomisiert (selten)

nach Lateinischem Quadrat (häufig)

n Bedingungen → n Pbn- und n Item-Gruppen

[Inferenzstatistik]

Statistische Tests

Ist das für eine untersuchte Stichprobe gefundene Ergebnis (Unterschiede zwischen Bedingungen) generalisierbar?
statistische Tests testen gegen die Nullhypothese (keine Unterschiede in Population) und ermitteln Wahrscheinlichkeit (p)
 p = Wahrscheinlichkeit dafür, Unterschiede zwischen den experimentellen Bedingungen zu finden, wenn es in der Population keine Unterschiede gibt (= wenn Nullhypothese gilt)
wenn $p < .05$ / $p < .01$, nennt man den Effekt der experimentellen Manipulation signifikant

[Inferenzstatistik]

t-Test

zur Überprüfung einer Unterschiedshypothese bei einem ein-faktoriellen Design mit zwei Bedingungen und einer intervallskalierten abhängigen Variable

t-Test für unabhängige Stichproben

bei between-subjects manipulierten Faktor
Prüfgröße: Verhältnis zwischen Mittelwertunterschied und Streuung der beiden Stichproben

t-Test für abhängige Stichproben

bei within-subject manipulierten Faktor
Prüfgröße: Verhältnis zwischen Mittelwert der Differenzen und Streuung der Differenzen

[Inferenzstatistik]

Varianzanalyse (VA)

zur Überprüfung einer Unterschiedshypothese bei einem oder mehr Faktoren mit zwei oder mehr Stufen und einer intervallskalierten abhängigen Variable

Grundprinzip: Varianzanalyse zerlegt die vorhandene Varianz der Messwerte in Komponenten, allgemein: treatment- und Fehleranteil
Prüfgröße: verallgemeinert: Verhältnis zwischen treatment-Varianz und Fehlervarianz

VA mit Messwiederholung: within-subject-Manipulation von Faktoren

VA (ohne Messwdhg): between-subjects-Manipulation von Faktoren

VA und Interaktionen

Mehrfaktorielle VA → Testen von Interaktion

→ Wechselwirkung von 2 oder mehr Faktoren auf abhängige Variable

z.B. signifikanter Effekt der Interaktion von 2 Faktoren, A und B:

Unterschiedlicher Effekt von Faktor A für die Stufen des Faktors B

[Worterkennung]

Mentales Lexikon

mentale Repräsentation des Wissens über Worte

Auditive (AWE) vs. visuelle Worterkennung (VWE)

zeitliche Erstreckung: +/-; Flüchtigkeit: +/-; klare Wortgrenzen: -/+

Effekte

Frequenz

semantisches Priming

Wort/Nicht-Wort

Uniqueness point (nur bei auditiv)

[Worterkennung]

Modelle der Worterkennung

Forsters Modell der autonomen seriellen Suche (VWE + AWE)

Logogenmodell (VWE + AWE)

Interactive Activation and Competition Modell (nur VWE)

Kohortenmodell (nur AWE)

TRACE (nur AWE)

Suchmodelle vs. Modelle des direkten Zugriffs?

Konnektionistisch?

Erklärung der Worterkennungseffekte?

Schwierigkeiten?

Lexikalische Ambiguität

Aktivierung von Lesarten

→ Experiment von Swinney (*bugs*)

[Bedeutung und Textverstehen]

Flache semantische Verarbeitung

Moses-Illusion

→ semantischer Effekt (Bemerken hängt von semantischer Relation ab)

Erklärung über Merkmalstheorien

Beeinflussung der Verarbeitungstiefe

Nicht-wörtliche Bedeutung

z.B. indirekte Sprechakte

Verarbeitung: traditioneller Ansatz → von wörtlicher zur nicht-wörtlicher Bedeutung

Negation

längere Lesezeiten → traditionelle Erklärung

pragmatische Angemessenheit

[Sprachproduktion]

Versprecher

Versprecherklassifikation: Antizipation, Vertauschung, Substitution

Analyse von Versprechern: Aufschluss über Sprachproduktionssystem

Komponenten der Sprachproduktion

Konzeptualisierungskomponente

Grammatikalische Komponente

Funktionale Verarbeitung (lexikalische Auswahl + Funktionszuweisung)

Positionale Verarbeitung (lexikalischer Abruf + Konstituentenverknüpfung)

Phonologische Komponente

Abgrenzung von grammatikalischer und phonologischer Komponente

Gründe für Abgrenzung?

Versprechertypen

[Situiertes Sprachverstehen]

Visual-World-Paradigma

Sprache + Bild

Blickbewegungsmessung (→ Aufschluss über Prozesse beim Sprachverstehen und Einfluss von Abbildungen auf Sprachverstehen)

Experiment von Tanenhaus et al. (1995) zur syntaktischen Ambiguität

Put the apple on the towel in the box

→ Effekt des visuellen Kontextes auf syntaktische Verarbeitung

→ syntaktische Verarbeitung ≠ vollständig autonom

Antizipatorische Blickbewegungen

Blicke zu einem abgebildeten Gegenstand bevor dieser genannt wurde

[Situiertes Sprachverstehen]

Experiment von Knoeferle et al. (2005)

Verarbeitung strukturell ambiger Sätze (SVO vs. OVS)

Die Prinzessin wäscht offensichtlich den Piraten / malt offensichtlich der Fechter.

Abgebildete Ereignisse (den-Piraten-waschende-Prinzessin / die-Prinzessin-malender-Fechter) beeinflussen Verarbeitung

Inkrementelle Nutzung von sprachlicher und visueller Information

Aspekte der sprachliche Äußerung → Blickverhalten

Aspekte der Abbildung → Blickverhalten

[Sprache und Embodiment]

Embodiment → Theorien

Gleiches Format von Repräsentationen beim Sprachverstehen und bei nicht-linguistischer Kognition

Zwaans Immersed Experienter Framework

Sprache = Instruktion für die mentale Simulation des Erlebens einer beschriebenen Situation

mentale Repräsentationen beruhen auf Erfahrungen

Referentielle Repräsentationen (Gedächtnisspuren für Konzepte)

Linguistische Repräsentationen (Gedächtnisspuren für visuelle/lautl. Muster)

Verbindungen zwischen Repräsentationen ← entstanden über gemeinsames Auftreten