

Verlaufscharakteristiken des Informationsabrufs beim Wiedererkennen und Verifizieren von Sätzen¹⁾

Franz Schmalhofer

Psychologisches Institut der Universität Freiburg

In einer empirischen und theoretischen Untersuchung wird gezeigt, daß ein Vergleich mittlerer Latenzzeiten richtiger Antworten selbst dann zu unzulässigen Schlußfolgerungen führen kann, wenn die Fehlerraten in den verschiedenen experimentellen Bedingungen identisch sind. Zur Vermeidung dieses Problems werden empirische und theoretische Verlaufscharakteristiken als Alternative zu Latenzzeitregistrierung und Wahlreaktionszeitmodellen vorgeschlagen. Eine Verlaufscharakteristik spezifiziert für die entsprechenden Bearbeitungszeitpunkte t eines bestimmten Zeitintervalls $0 < t < T$ die zu diesen Zeitpunkten vorliegenden Verarbeitungsergebnisse. Das vorgeschlagene Verfahren wurde zur Untersuchung des Wiedererkennens und Verifizierens von Sätzen angewendet. Die Ergebnisse zeigten, daß trotz der unterschiedlichen Aufgabenstellungen die gleichen Verarbeitungsstrategien angewendet wurden und daß bei beiden Aufgaben die Antworten der Personen stärker von im Gedächtnis abgespeicherten Situationsspezifikationen als von propositionalen oder wörtlichen Gedächtnisspuren des gelesenen Textes beeinflußt wurden.

Zur Erforschung kognitiver Prozesse und ihrer Dauer werden häufig Antwortlatenzzeiten unter verschiedenen experimentellen Bedingungen erhoben. Die vorliegende Untersuchung zeigt jedoch, daß ein Vergleich der mittleren Latenzzeiten richtiger Antworten selbst dann zu unzulässigen Schlußfolgerungen führen kann, wenn die Fehlerraten in den verschiedenen experimentellen Bedingungen identisch sind. Als Alternative zur Erhebung von einfachen Antwortlatenzzeiten wird daher vorgeschlagen Verlaufscha-

1) Diese Untersuchung wurde durch eine Reisebeihilfe der Deutschen Forschungsgemeinschaft unterstützt. Ich danke W. Kintsch, P. G. Polson, D. Albert und W. Hell für anregende und kritische Erörterungen über die berichteten Untersuchungen. W. Kintsch, P. G. Polson und dem Institute of Cognitive Science der University of Colorado habe ich für die freundliche Erlaubnis zur Nutzung des Computer Laboratory for Instruction in Psychological Research (CLIPR) und des „subject pools“ zu danken.

rakteristiken des Informationsabrufs zu erstellen und empirisch zu überprüfen.

Eine Verlaufscharakteristik spezifiziert für beliebige Bearbeitungspunkte t eines bestimmten Zeitintervalls $0 < t < T$, die zu diesen Zeitpunkten vorliegenden (möglicherweise vorläufigen) Verarbeitungsergebnisse. Die relative Häufigkeit richtiger Antworten, sowie andere Leistungsmaße oder auch subjektive Urteile können als solche Verarbeitungsergebnisse angesehen werden. Eine Verlaufscharakteristik ist somit eine Abbildung $v: t \rightarrow p(t)$, wobei p ein Maß für ein bestimmtes Verarbeitungsergebnis sei. Anders als bei Wahlreaktionszeitmodellen (Laming, 1968), bei denen die Latenzzeit und die Richtigkeit einer Antwort als abhängige Variablen betrachtet werden, wird bei Verlaufscharakteristiken das Verarbeitungsergebnis ähnlich wie bei speed accuracy trade-off Charakteristiken (Pachella, 1974) in Abhängigkeit der Bearbeitungszeit analysiert. In empirischen Untersuchungen können entsprechende experimentelle Daten erhoben werden. Es lassen sich daher theoretische und empirische Verlaufscharakteristiken voneinander unterscheiden. Theoretische Verlaufscharakteristiken werden ohne Einbeziehung empirischer Daten aus den Annahmen eines Modells hergeleitet. Eine in einem Experiment empirisch festgestellte Zuordnung oder Abbildung von Bearbeitungszeitpunkten zu den Ergebnissen kognitiver Verarbeitungsprozesse soll als empirische Verlaufscharakteristik bezeichnet werden.

Um ein Verarbeitungsergebnis für einen bestimmten Zeitpunkt empirisch zu bestimmen, werden häufig Antwort-Signale vorgegeben (Schmalhofer, 1978; 1982; Wickelgren, Corbett & Doshier, 1980). Dadurch können vorläufige Verarbeitungsergebnisse im Prinzip zu jedem beliebigen Zeitpunkt abgefragt werden. Dagegen bleibt bei einer einfachen Registrierung von Antwortlatenzzeiten der Zeitpunkt der Mitteilung eines Verarbeitungsergebnisses (Reaktion) der Versuchsperson selbst überlassen, so daß in diesem Falle sowohl Antwortrichtigkeit als auch Bearbeitungszeit abhängige Maße sind. Ein gravierendes Problem, das bei der Analyse der Kriteriumsvariablen Latenzzeit und der Kovariaten Antwortrichtigkeit auftreten kann, wird im folgenden am Beispiel der empirischen Überprüfung eines spezifischen Modells aufgezeigt. Es wird dann demonstriert, wie dieses Problem durch die Erstellung von Verlaufscharakteristiken umgangen werden kann.

Zur Überprüfung der Prozeßannahmen ihres Strategieauswahlmodells hat Reder (1982) einen Vergleich von Latenzzeiten zwischen verschiedenen experimentellen Bedingungen durchgeführt. Das Strategieauswahlmodell postuliert, daß trotz der unterschiedlichen Aufgabenanforderungen des Wiedererkennens und Verifizierens von Sätzen bei beiden Aufgaben die gleichen kognitiven Prozesse ausgeführt würden. Bei der Wiedererken-

nungsaufgabe soll bestimmt werden, ob ein vorgegebener Testsatz wörtlich im Text vorkam. Bei der Satzverifikation muß dagegen festgestellt werden, ob ein Testsatz bezüglich der im Text beschriebenen Sachverhalte richtig oder falsch ist.

Nach dem Strategieauswahlmodell von Reder werden Wiedererkennungs- und Verifikationsaufgaben mit denselben Strategien bearbeitet. Es wird angenommen, daß bei beiden Aufgabenstellungen zwei verschiedene Strategien angewandt werden können. *Die Strategie des direkten Vergleiches* (direct matches) bestimmt eine Antwort, indem der Testsatz mit den wörtlichen und propositionalen Gedächtnisspuren des zuvor gelesenen Textes verglichen wird. Wenn eine Übereinstimmung mit den geprüften Gedächtnisspuren vorliegt, wird ein Satz bei Anwendung dieser Strategie wiedererkannt oder als sachlich richtig eingestuft. Ein Testsatz kann darüber hinaus aufgrund von *Plausibilitätsurteilen* (plausibility judgements) beantwortet werden. Bei der Anwendung dieser Strategie überprüft die V_p , ob und wie gut der vorgegebene Testsatz zu den im Text beschriebenen Sachverhalten paßt und errechnet dann aufgrund der gefundenen Übereinstimmung eine Antwort. Bei einer hohen Übereinstimmung wird ein Satz wiedererkannt bzw. als sachlich richtig eingestuft. Reder's Modell postuliert, daß Plausibilitätsurteile zu schnelleren Antworten führen als das direkte Vergleichen eines Testsatzes mit wörtlichen und propositionalen Gedächtnisspuren. Es wird weiterhin angenommen, daß sowohl beim Wiedererkennen als auch beim Verifizieren beide Strategien zur Anwendung kommen, daß aber der Anteil von Plausibilitätsurteilen bei Verifikationsaufgaben deutlich höher liegt. Deshalb sollte ein Testsatz der wörtlich im Text vorkam, schneller verifiziert als wiedererkannt werden.

Für eine noch strengere Überprüfung des Strategieauswahlmodells und der Annahme, daß Personen Plausibilitätsurteile schneller ausführen können als direkte Vergleiche mit wörtlichen und propositionalen Gedächtnisspuren, stellte Reder folgende Hypothese auf:

„People are faster to judge a statement as plausible²⁾ even when it had not been presented in a text than to recognize that statement when it had been presented“ (S. 257).

Um im folgenden in kurzer Form die unterschiedlichen Testsätze ansprechen zu können, werden einige Abkürzungen eingeführt. Sätze, die wörtlich im Text vorkamen, werden als Original- oder O-Sätze bezeichnet. Falls die Bedeutung eines Testsatzes nicht expliziter Bestandteil der Textbasis (Kintsch, 1974; im Druck) des gelesenen Textes ist, so wird angenommen, daß die propositionale Darstellung des Testsatzes nicht zu der beim Lesen erstellten propositionalen Textenkodierung gehört. Dennoch kann ein sol-

2) d.h. Verifizieren von sachlich richtigen Sätzen.

cher Satz bezüglich der im Text angesprochenen Situationen korrekt sein. Solche sachlich richtigen Sätze, deren propositionale Darstellung nicht zur Textbasis gehört, werden als bedeutungsveränderte oder B-Sätze bezeichnet. Somit besagt Reders Hypothese, daß die Verifikation von B-Sätzen schneller erfolgt als das Wiedererkennen von O-Sätzen.

Die Latenzzeitexperimente von Reder

Um die obige Hypothese zu überprüfen, wurden von Reder (1982) für beide Aufgaben Antwortlatenzzeiten und Fehlerraten aufgezeichnet. Damit bei beiden Aufgaben nahezu identische Fehlerraten auftreten würden, wurden den Vpn on-line antwortkontingente Rückmeldungen gegeben (Reder, Experiment 2), welche je nach Bedarf die Vpn aufforderten entweder schneller oder mit weniger Fehlern zu antworten. Dadurch sollte verhindert werden, daß die durchschnittlichen Latenzzeiten wegen unterschiedlich hoher Fehlerraten nicht miteinander verglichen werden könnten (Problem des Geschwindigkeits-Genauigkeitsaustauschs; vgl. Pachella, 1974; Wender, Colonius & Schulze, 1980). Reder nimmt an, daß diese mittleren Latenzzeiten die Dauer der zum Wiedererkennen und Verifizieren verwendeten kognitiven Prozesse reflektieren. Aus dem Vergleich der empirisch erhobenen Antwortlatenzzeiten wurde von ihr der folgende Schluß gezogen:

„Finally, the prediction was confirmed that subjects should be faster to judge a statement as plausible when it was not presented than to recognize it when it was presented“ (S. 265).

Reder vertritt die Ansicht, daß die Ergebnisse ihres Latenzzeitexperimentes die Vorhersagen des Strategieauswahlmodells empirisch stützen. Falls es jedoch zutreffen sollte, — wie die folgende Analyse zeigen wird — daß nach Reders Modell für andere Fehlerraten eine umgekehrte Latenzzeitbeziehung vorhergesagt wird, können die von Reder berichteten Daten kaum als empirische Evidenz für ihr Modell interpretiert werden. Durch das Erstellen von Verlaufscharakteristiken für das Verifizieren von B- und das Wiedererkennen von O-Sätzen wollen wir im folgenden zeigen, daß aus Reders Modell sowohl die empirisch beobachtete Latenzzeitbeziehung als auch deren Umkehrung abgeleitet werden kann.

Theoretische Verlaufscharakteristiken

Es ist bekannt, daß die Analyse der Beziehung zwischen der Geschwindigkeit und Genauigkeit einer Antwort wichtige Erkenntnisse zur Identifi-

zierung separat ablaufender kognitiver Prozesse liefern kann (Link 1982). Bei der Spezifizierung der theoretischen Verlaufscharakteristiken des Strategieauswahlmodells ergibt sich für die Verifikation von B-Sätzen eine nicht monotone Funktion: Da Plausibilitätsurteile und die Strategie des direkten Vergleichens alternativ angewendet werden³⁾, entspricht die durchschnittliche Latenzzeit und die durchschnittliche Anzahl richtiger Antworten bei Vorliegen eines gegebenen Testsatzes den gewichteten durchschnittlichen Latenzzeiten bzw. richtigen Antworten, welche durch die beiden Strategien zustandekommen. Durch Plausibilitätsurteile wird bestimmt, ob ein Testsatz bezüglich der im Text dargestellten Sachverhalte korrekt ist. Somit erzeugen Plausibilitätsurteile bei Verifikationsaufgaben für B-Sätze korrekte Antworten, nämlich ja-Antworten. Die Strategie des direkten Vergleichens stellt demgegenüber fest, ob ein Satz wörtlich oder sinngemäß im Text vorkam. Durch das direkte Vergleichen würden somit beim Verifizieren von B-Sätzen falsche Antworten, nämlich nein-Antworten erzeugt. Für B-Sätze existieren nämlich weder wörtliche noch propositionale Gedächtnisspuren. Da Plausibilitätsurteile schneller erfolgen als direkte Vergleiche, sollte die Wahrscheinlichkeit einer richtigen Antwort mit steigender Bearbeitungszeit zunächst ansteigen bis ein Maximum erreicht wird. Danach wird mit zunehmender Bearbeitungszeit durch die erst später abgegebenen (und für B-Sätze falschen) Resultate des direkten Vergleichens die durchschnittliche Anzahl richtiger Antworten verringert.

Speziell nimmt das Modell von Reder an, daß Plausibilitätsurteile mit der Wahrscheinlichkeit x angewendet werden. Sowohl für O- als auch für B-Sätze führen diese Urteile zum Zeitpunkt t_1 zu einer ja-Antwort. Bei Antworten, die vor diesem Zeitpunkt erfolgen, müssen Personen raten, d. h. vor dem Zeitpunkt t_1 sei die Wahrscheinlichkeit einer ja-Antwort = 0.50 (fast guesses). Die theoretischen Verlaufscharakteristiken für Plausibilitätsurteile beim Wiedererkennen von O-Sätzen und beim Verifizieren von B-Sätzen sind in Abschnitt A der Abbildung 1 dargestellt. Nach dem Modell von Reder ergibt sich für die Plausibilitätsurteile bei beiden Testsätzen, die sich bzgl. ihrer spezifischen Plausibilität nicht unterscheiden sollten, die gleiche Charakteristik. Für die in Abschnitt A dargestellten Funktionen wurden zufällig variierende Urteilslatenzen mit einem Mittelwert von $t_1 = 2$ sec. angenommen.

3) Entsprechend den Annahmen von Reders Modell kann nach einer erfolglosen Anwendung der Strategie des direkten Vergleichens, bei der die gesuchten Gedächtnisspuren nicht aufgefunden wurden, zusätzlich ein Plausibilitätsurteil folgen. Da dies nach dem Modell von Reder insgesamt jedoch selten vorkommt, und die in diesem Artikel vorgebrachten Argumente dadurch nicht betroffen werden, können wir die konsekutive Anwendung von zwei Strategien (Plausibilitätsurteil nach erfolglosem direkten Vergleich) hier außer acht lassen.

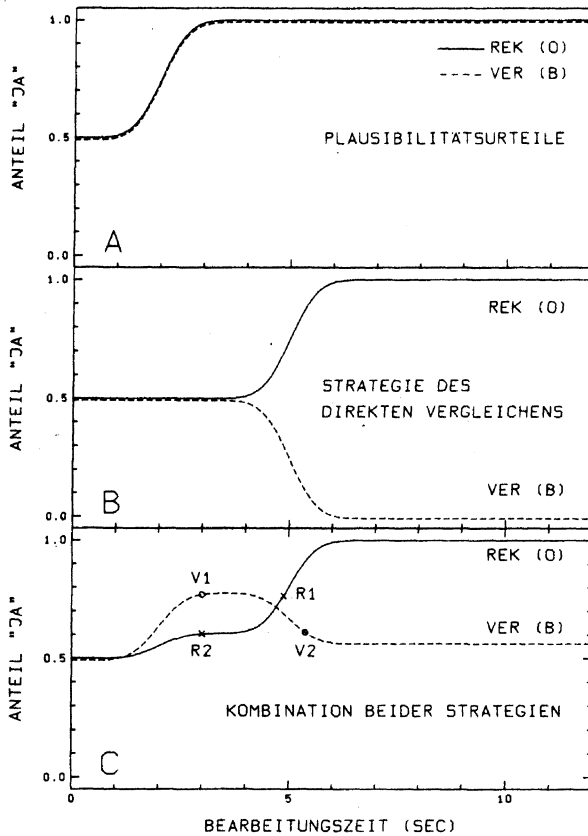


Abb. 1

Theoretische Verlaufscharakteristiken für A) Plausibilitätsurteile, B) die Strategie des direkten Vergleichens sowie C) eine Kombination beider Strategien beim Wiedererkennen von O-Sätzen und Verifizieren von B-Sätzen

Die Strategie des direkten Vergleichens, die mit einer Wahrscheinlichkeit von $(1-x)$ angewendet wird, liefert erst zu einem späteren Zeitpunkt t_2 ($t_2 > t_1$) eine Antwort. Im Gegensatz zu O-Sätzen liegen für B-Sätze keine wörtlichen oder propositionalen Gedächtnisspuren vor. Bei der Verifikation von B-Sätzen aufgrund des direkten Vergleichens sollten sich daher nein-Antworten ergeben. Abschnitt B der Abbildung 1 zeigt die theoretischen Verlaufscharakteristiken des direkten Vergleichens für das Wiedererkennen von O-Sätzen und das Verifizieren von B-Sätzen. Es wurde angenommen, daß die Dauer des direkten Vergleichens zufällig variiert und im Mittel $t_2=5$ Sekunden in Anspruch nimmt. Da im Modell nur die Wahrscheinlichkeit spezifiziert wird, mit der bei einer gegebenen Aufgabeninstruktion Plausibilitätsurteile bzw. die Strategie des direkten Vergleichens angewendet wird, können die in Abschnitt A und B angegebenen Verlaufscharakteristiken nicht unmittelbar empirisch evaluiert werden. Für eine empirische Überprüfung muß statt dessen zuerst eine durch gewichtete

Mittlung gebildete Kombination der in Abschnitt A und B angegebenen Verlaufscharakteristiken gebildet werden. Abschnitt C zeigt die Verlaufscharakteristik einer solchen Kombination beider Strategien. Dabei beträgt der Anteil von Plausibilitätsurteilen bei der Wiedererkennungsaufgabe $x=.21$. Bei der Verifikationsaufgabe beträgt der entsprechende Anteil $x'=.57$. Diese Angaben wurden aus der Tabelle 5 von Reder (1982, S. 267) entnommen.

Aus Abschnitt C der Abbildung ersieht man, daß die Verlaufscharakteristik für das Wiedererkennen von O-Sätzen monoton ansteigt. Beim Verifizieren von B-Sätzen ergibt sich dagegen eine nichtmonotone Verlaufscharakteristik. Man erkennt, daß sich die beiden Funktionen zum Zeitpunkt $t = 4.5$ sec. schneiden. Bei Bearbeitungszeiten, die unter 4.5 sec. liegen, ergeben sich daher beim Verifizieren von B-Sätzen mehr korrekte Antworten als beim Wiedererkennen von O-Sätzen. Umgekehrt ergeben sich bei Bearbeitungszeiten, die über 4.5 sec. liegen, für das Wiedererkennen von O-Sätzen mehr richtige Antworten als beim Verifizieren der B-Sätze.

Aus Reders Modell läßt sich leicht ableiten, daß solche Überschneidungen für alle Parameter x und x' entstehen, welche die grundlegende Modellannahme erfüllen, nämlich, daß $x < x'$. In Worten ausgedrückt bedeutet dies, daß Plausibilitätsurteile häufiger beim Verifizieren als beim Wiedererkennen angewendet werden. Daraus folgt, daß Reders Modell auch für niedrige Fehlerraten eine solche Überschneidung vorhersagt. Diese Vorhersage ist auch dann gültig, wenn die Wahrscheinlichkeit einer ja-Antwort durch Raten von 0.5 verschieden ist und auch wenn angenommen wird, daß direkte Vergleiche wie auch Plausibilitätsurteile in ihrem Resultat fehlerhaft sein können (z. B. wegen schwacher oder verfälschter Gedächtnisspuren). Direkte Vergleiche und Plausibilitätsurteile würden dann die vorhergesagte Antwort nur noch mit einer Wahrscheinlichkeit von $1-\epsilon$ erzeugen. Dabei soll sinnvollerweise gelten: $\epsilon < 0.5$. Nach einer kurzen Bearbeitungszeit, bei der die Plausibilitätsurteile schon abgeschlossen sind, die direkten Vergleiche jedoch noch andauern, ergeben sich für die beiden Aufgabeninstruktionen folgende Wahrscheinlichkeiten einer ja-Antwort:

$$W(\text{ja/Rek} \rightarrow \text{O}) = x \cdot (1-\epsilon) + (1-x) \cdot g \quad (1)$$

und

$$W(\text{ja/Ver} \rightarrow \text{B}) = x' \cdot (1-\epsilon) + (1-x') \cdot g \quad (2)$$

Da nach Voraussetzung gilt $x < x'$, ergibt sich für kurze Bearbeitungszeiten, daß $W(\text{ja/Rek} \rightarrow \text{O}) < W(\text{ja/Ver} \rightarrow \text{B})$. Nach längeren Bearbeitungszeiten, bei denen beide Strategien bereits ihre Antworten erzeugt haben, ergeben sich die in Gleichung 3 und 4 dargestellten Wahrscheinlichkeiten einer ja-Antwort:

$$W(\text{ja/Rek—O}) = x \cdot (1-\varepsilon) + (1-x) \cdot (1-\varepsilon) = 1-\varepsilon \quad (3)$$

$$W(\text{ja/Ver—B}) = x' \cdot (1-\varepsilon) + (1-x') \cdot \varepsilon \quad (4)$$

Da nach Annahme $\varepsilon < 0.5$ und $x' < 1$ ist, ergibt sich für längere Verarbeitungszeiten $W(\text{ja/Rek—O}) > W(\text{ja/Ver—B})$. Damit ist gezeigt, daß die Überschneidung der kontinuierlichen Verlaufscharakteristiken für das Verifizieren von B-Sätzen und das Wiedererkennen von O-Sätzen eine notwendige Bedingung des Strategieauswahlmodells ist.

Für eine konstante aber unspezifizierte Antwortrichtigkeit kann daher nicht entschieden werden, ob das Verifizieren von B-Sätzen oder das Wiedererkennen von O-Sätzen schneller vor sich geht. In Abhängigkeit der speziellen mittleren Richtigkeit der Antworten ist sowohl die Beobachtung, daß die Verifikation von B-Sätzen schneller erfolgt als das Wiedererkennen von O-Sätzen, mit dem Modell verträglich als auch das Gegenteil. Daraus wird ersichtlich, daß der von Reder durchgeführte Test nicht zur Überprüfung des Strategieauswahlmodells geeignet ist. So läßt sich aus Abschnitt C der Abbildung 1 erkennen, daß die Verifikationsaufgabe bei einer relativ niedrigen Fehlerrate schneller durchgeführt wird, als die Wiedererkennungsaufgabe (vgl. die Punkte V1 und R1). Bei einer höheren Fehlerrate wird dagegen die Wiedererkennungsaufgabe schneller ausgeführt als die Verifikationsaufgabe (siehe Punkte R2 und V2). Selbst wenn kein Geschwindigkeits-Genauigkeitsaustausch (speed accuracy trade-off) vorliegt, kann daher nicht entschieden werden, welche der beiden Aufgaben schneller durchgeführt wird.

Da Verlaufscharakteristiken jedoch im allgemeinen monoton ansteigen (Pachella, 1974), kann das Strategieauswahlmodell aufgrund der empirischen Vorhersage einer nicht-monotonen Verlaufscharakteristik kritisch überprüft werden. Im folgenden wird ein Experiment berichtet, bei dem für Wiedererkennungsaufgaben und Verifikationsaufgaben empirische Verlaufscharakteristiken erhoben wurden.

Erhebung empirischer Verlaufscharakteristiken

Durchführung: Zur Bestimmung empirischer Verlaufscharakteristiken wurde eine Untersuchung mit 64 englischsprachigen Studenten der University of Colorado durchgeführt. Die Vpn wurden instruiert, einen englischen Text (742 Worte) zu studieren, um dadurch Kenntnisse einer Programmiersprache zu erwerben. Dieser Text erläuterte die Regeln, nach denen Atome und S-Terme in einem bestimmten Dialekt der Programmiersprache LISP gebildet werden. Außerdem lieferte der Text einige Beispiele für korrekt sowie falsch gebildete Atome und S-Terme. Daran anschließend mußte jede Person

Sätze, die jeweils einzeln am Bildschirm dargeboten wurden, verifizieren bzw. wiedererkennen. Anhand von Antwortsignalen, die eine Sekunde vor der Darbietung eines Satzes sowie 1, 3, 5, 7, 9 und 11 sec. nach der Darbietung jedes Testsatzes vorgegeben wurden, mußte jede Vp zu den vorgegebenen Zeitpunkten, also für jeden dargebotenen Satz insgesamt siebenmal, eine Antwort abgeben. Eine Antwort wurde abgegeben, indem von der Vp die „ja“- oder „nein“-Taste gedrückt wurde. Diese spezielle Antwortmethode, die als speed accuracy-tapping procedure (Wickelgren et al. 1980) bezeichnet wird, wurde von Schmalhofer & Glavanov (im Druck) detailliert beschrieben.

Neben den bereits oben eingeführten Original- (O-) und bedeutungsveränderten (B-) Sätzen, wurden zwei weitere Arten von Testsätzen vorgegeben. Um bestimmen zu können, inwieweit die Strategie des direkten Vergleichens auf wörtliche oder propositionale Gedächtnisspuren zurückgreift, wurden auch paraphrasierte Testsätze (P-Sätze) vorgegeben. Paraphrasierte Testsätze stimmen bezüglich ihrer Bedeutung mit dem Originalsatz überein, unterschieden sich jedoch im Wortlaut. Zur zusätzlichen Kontrolle wurden darüber hinaus Testsätze gebildet, die bezüglich der Programmiersprache LISP sachlich falsch waren. Dadurch konnte zusätzlich der Einfluß der in einem Text ausgedrückten situativen Spezifikationen auf das Verifizieren und Wiedererkennen von Sätzen untersucht werden. Diese sachlich falschen Sätze werden im folgenden als F-Sätze bezeichnet. Von einem Originalsatz wurden z. B. folgende vier Testsätze hergeleitet bzw. übernommen:

BIO.-200 is not a legal atom, because it constains a dot.	(O-Satz)
BIO.-200 is not a legal atom, since it has a period.	(P-Satz)
BIO.-200 is not a legal atom, since it has a dash.	(B-Satz)
BIO200 is not a legal atom, because it concludes with a number.	(F-Satz)

Jeder Vp wurden 24 Testsätze zum Verifizieren sowie 24 Testsätze zum Wiedererkennen vorgegeben. Es wurden jeweils fünf O-, fünf P-, fünf B- und fünf F-Sätze vorgegeben. Zur Einübung wurden darüber hinaus 4 weitere Sätze dargeboten. Um die apriori-Wahrscheinlichkeiten des richtigen Erratens einer Antwort zwischen den Rekognitions- und Verifikationsaufgaben anzugleichen, wurden je nach Aufgabeninstruktion zusätzlich vier O-Sätze bzw. vier F-Sätze verwendet, die jedoch nicht in die Auswertung miteingingen. Bei der Wiedererkennungsaufgabe wurde die Vp instruiert, zu entscheiden, ob der vorgelegte Testsatz wörtlich im Text vorkam. Beim Verifizieren mußte die Vp dagegen beurteilen, ob ein Satz bezüglich der Programmiersprache LISP eine sachlich richtige Aussage macht. Die Testsätze für die beiden Aufgaben wurden gleichmäßig über die Vpn verteilt, d. h. in beiden Aufgaben wurden alle Testsätze gleich häufig vorgelegt und kein Testsatz wurde einer Vp zweimal dargeboten.

Ergebnisse und Diskussion: In Tabelle 1 ist die relative Häufigkeit der ja-Antworten als eine Funktion der Bearbeitungszeit für die 4 Satztypen bei Vorliegen der beiden Aufgabeninstruktionen wiedergegeben. Die entsprechenden relativen Häufigkeiten wurden für jede Vp separat arcsin-transformiert und einer $2 \times 4 \times 6$ -Varianzanalyse mit den Faktoren Aufgabeninstruktion, Art des Testsatzes und Bearbeitungszeit unterzogen. Dabei ergaben sich für die Aufgabeninstruktionen $F(1,504) = 36.74$, $MSE = 1.74$, $p < .0001$, die Satztypen $F(3,504) = 215.36$, $MSE = 1.74$, $p < .0001$ und die Bearbeitungszeiten $F(5,2520) = 33.26$, $MSE = 0.22$, $p < .0001$, statistisch signifikante Unterschiede. Darüber hinaus waren sämtliche Interaktionseffekte signifikant, $p < .05$. Um herauszufinden, ob O-Sätze unter Wiedererkennung- oder Verifikationsinstruktionen leichter durchgeführt werden, wurden die arcsin-transformierten relativen Häufigkeiten der O-Sätze einer 2×6 Varianzanalyse unterzogen. Die Analyse ergab, daß beim Verifizieren signifikant mehr korrekte Antworten gegeben werden als beim Wiedererkennen, $F(1,126) = 11.37$, $MSE = 1.37$, $p < .001$. Für die Bearbeitungszeit ergab sich ebenfalls ein signifikanter Effekt, $F(5,630) = 5.67$, $MSE = 1.13$, $p < .0001$. Der Interaktionseffekt war dagegen nicht signifikant, $F < 1$. Dies zeigt, daß das Verifizieren von Sätzen leichter fällt als das Wiedererkennen der gleichen Sätze.

Die Ergebnisse dieser Analyse stehen im Einklang mit Reders Befunden, die anhand einfacher Latenzzeitmessung gewonnen wurden und stimmen somit mit den Annahmen des Strategieauswahlmodells überein. Darüber hinaus zeigten die Daten jedoch auch, daß beim Verifizieren eine bessere asymptotische Leistung erreicht wird als beim Wiedererkennen der gleichen Sätze. Daraus könnte man folgern, daß Plausibilitätsurteile weniger fehleranfällig sind als die Strategie des direkten Vergleichens.

Tabelle 1

Anteil von „ja“-Antworten nach unterschiedlich langen Bearbeitungszeiten bei den 4 Testsatztypen und unter zwei verschiedenen Aufgabeninstruktionen

Anteil (ja)	Wiedererkennen von				Verifikation von			
	O	P	B	F	O	P	B	F
nach 1 sec	.71	.69	.68	.62	.76	.70	.76	.64
nach 3 sec	.74	.69	.66	.39	.82	.80	.78	.37
nach 5 sec	.76	.67	.60	.21	.85	.86	.74	.21
nach 7 sec	.77	.68	.59	.16	.86	.88	.74	.16
nach 9 sec	.79	.67	.55	.16	.87	.88	.74	.15
nach 11 sec	.79	.67	.53	.13	.87	.87	.75	.15

Um eine kritische Überprüfung des Strategieauswahlmodells durchzuführen, sollte jedoch die Leistung beim Wiedererkennen von O-Sätzen mit der entsprechenden Leistung beim Verifizieren von B-Sätzen verglichen werden. Im Gegensatz zum Wiedererkennen von O-Sätzen sollte beim Verifizieren von B-Sätzen eine nichtmonotone Verlaufscharakteristik beobachtet werden. Die relativen Häufigkeiten der ja-Antworten der beiden Aufgaben, die in Abbildung 2 dargestellt sind, wurden einer 2×6 -Varianzanalyse unterzogen. Es ergaben sich keine signifikanten Haupteffekte, $F < 1$. Jedoch zeigte eine signifikante Wechselwirkung $F(5,630) = 2.71$, $MSE = 0.20$, $p < .05$, daß die Überkreuzung der beiden Verlaufscharakteristiken (Abbildung 2) statistisch reliabel ist.

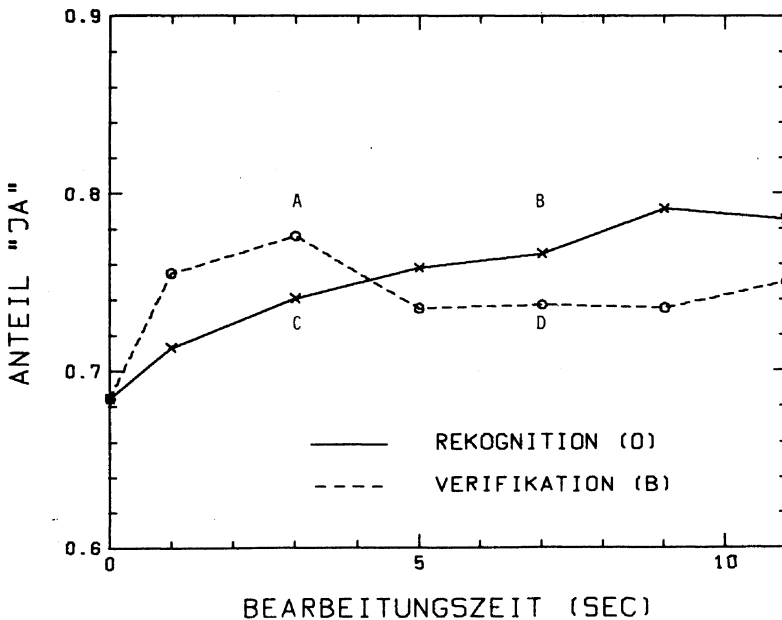


Abb. 2

Empirische Verlaufscharakteristiken beim Wiedererkennen von O-Sätzen und Verifizieren von B-Sätzen

Außerdem verdeutlichen die vorliegenden Ergebnisse, daß selbst in Abwesenheit einer Geschwindigkeits-Genauigkeitsaustausch-Konstellation der einfache Vergleich von Latenzzeiten in beliebiger Weise zu einem der beiden Ergebnisse führen kann. Dies ist darauf zurückzuführen, daß sich die beiden Verlaufscharakteristiken überkreuzen. Bei einer Antwortrichtigkeit von 0.77 wird das Verifizieren der B-Sätze schneller durchgeführt als das Wiedererkennen von O-Sätzen (Punkte A und B). Bei einer Antwort-

richtigkeit von 0.74 erfolgt dagegen das Wiedererkennen von O-Sätzen schneller als das Verifizieren von B-Sätzen (Punkte C und D). Da die Strategie des direkten Vergleichens sowohl auf wörtliche als auch auf propositionale Gedächtnisstrukturen zurückgreift, soll im folgenden der Beitrag wörtlicher und propositionaler Informationen beim Wiedererkennen und Verifizieren von Sätzen miteinander verglichen werden. Paraphrasierte (P-), bedeutungsveränderte (B-), und sachlich falsche (F-) Sätze kamen weder wörtlich noch sinngemäß im Text vor. Eine wörtliche Gedächtnisspur kann daher nur für Original (O-) Sätze bestehen. Da O- und P-Sätze sich in ihrer propositionalen Darstellung nicht unterscheiden und beide Sätze in bezug auf die im Text beschriebenen Situationen sachlich richtig sind, können die Unterschiede in der Beantwortung von O- und P-Sätzen zur Bestimmung der wörtlichen Gedächtnisspur verwendet werden.

In ähnlicher Weise unterscheiden sich P- und B-Sätze ausschließlich aufgrund ihrer propositionalen Darstellung. Beide Sätze sind sachlich richtig, wurden in dem Text jedoch nicht wörtlich dargeboten. Deshalb können die Unterschiede in den Antworten zu beiden Testsätzen zur Bestimmung der Stärke der propositionalen Gedächtnisspur verwendet werden. Schließlich unterscheiden sich B- und C-Sätze nur bezüglich ihrer sachlichen Richtigkeit. Dies ist die kritische Information, die bei Plausibilitätsurteilen verwendet wird. (B—F)-Satzpaare können somit dazu verwendet werden, den Beitrag situativer Information für die Beantwortung von Testsätzen zu beurteilen. Wie aus Tabelle 2 ersichtlich wird, kann man demzufolge den Beitrag wörtlicher, propositionaler und situativer Information mittels der Unterschiede zwischen den Antworten der Vpn auf O—P, P—B und B—F Satzpaare feststellen.

Tabelle 2
Art des Beitrags von wörtlichen, propositionalen und situativen Gedächtnisspuren für jeden der vier Testsätze

Art der Gedächtnisspur	Testsatz			
	O-Satz	P-Satz	B-Satz	F-Satz
wörtlich	+	-	-	-
propositional	+	+	-	-
situativ	+	+	+	-

Die „+“ und „-“ zeigen an, ob eine Gedächtnisspur positive oder negative Evidenz für das Wiedererkennen oder Verifizieren von Sätzen bereitstellt.

Der Beitrag wörtlicher, propositionaler und situativer Gedächtnisspuren wurde entsprechend der Theorie der Signalentdeckung (Baird & Noma,

1978) durch die Berechnung von d' -Werten geschätzt. Diese d' -Werte wurden von den relativen Häufigkeiten der ja-Antworten für die entsprechenden Satzpaare errechnet. Zunächst wurde für jeden erhobenen Bearbeitungszeitpunkt, für jede V_p , sowie für die beiden Aufgabeninstruktionen separat ein d' -Wert errechnet. In Abbildung 3, welche die mittleren d' -Werte sämtlicher V_p n zeigt, ist der Beitrag wörtlicher, propositionaler und situativer Information zum Wiedererkennen und Verifizieren von Sätzen dargestellt.

Für jede dieser drei Gedächtnisspuren wurde eine zweifaktorielle Varianzanalyse mit den Faktoren Aufgabeninstruktion (2 Stufen) und Bearbeitungszeit (6 Stufen) durchgeführt. Für wörtliche Information ergaben sich weder für die Aufgabeninstruktion, $F(1,126) = 3.88$, $p > 0.05$, noch für die Bearbeitungszeit signifikante Unterschiede, $F < 1$. Informationsverarbeitungsunterschiede wurden dagegen durch eine signifikante Interaktion aufgezeigt, $F(5,630) = 2.90$, $MSE = 0.57$, $p < 0.05$. Abbildung 3 deutet an, daß beim Wiedererkennen von Sätzen wörtlich abgespeicherte Informationen verwendet werden. Beim Verifizieren ist dies nicht der Fall. Bezüglich der propositionalen Information erwies sich nur die Bearbeitungszeit als signifikant, $F(5,630) = 14.74$, $MSE = 8.73$, $p < 0.0001$. Dies bedeutet, daß propositional abgespeicherte Informationen gleichermaßen zum Wiedererkennen und Verifizieren von Sätzen eingesetzt werden. Für Situationsinformation war sowohl die Bearbeitungszeit, $F(5,630) = 62.84$, $MSE = 0.70$, $p < 0.0001$, als auch die Aufgabeninstruktion, $F(1,126) = 8.67$, $MSE = 6.06$, $p < 0.005$ signifikant. Es lag keine signifikante Interaktion vor, $F < 1$. Daraus folgt, daß beim Verifizieren von Sätzen situative Informationen eine größere Rolle spielen als beim Wiedererkennen von Sätzen. Aber selbst bei Wiedererkennungsaufgaben, bei denen eigentlich wörtliche Gedächtnisspuren abgefragt werden, war der Anteil der situativen Information höher als der Anteil von wörtlichen und propositionalen Informationen zusammengekommen. Dies bestätigt Reders These, daß Plausibilitätsurteile, welche auf situativer Information basieren, auch beim Wiedererkennen von Sätzen eine sehr große Rolle spielen.

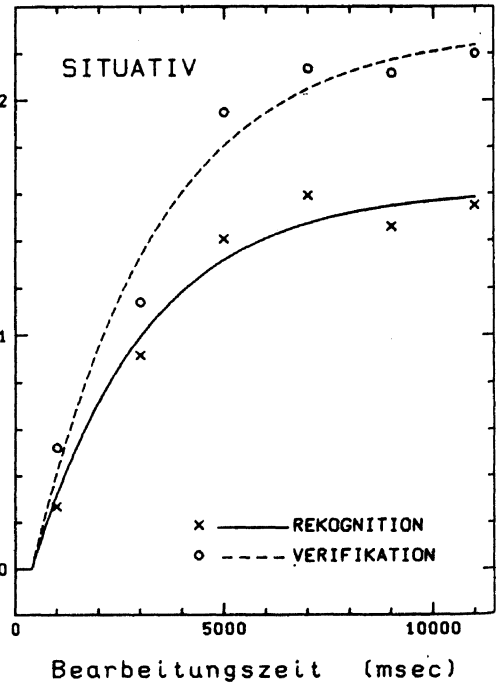
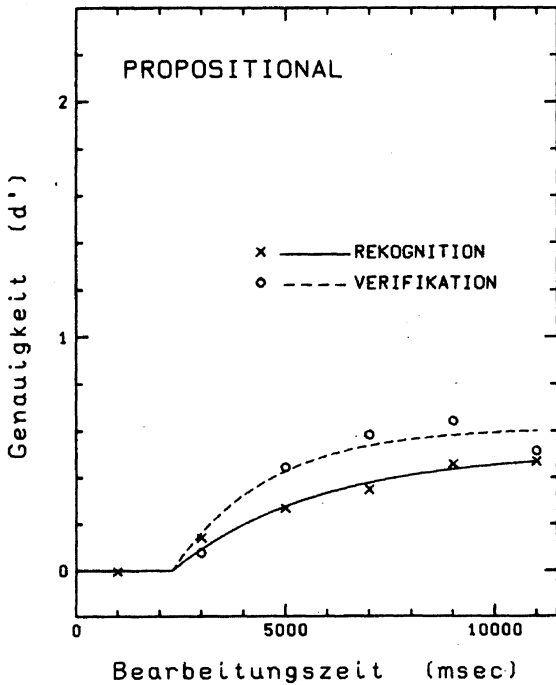
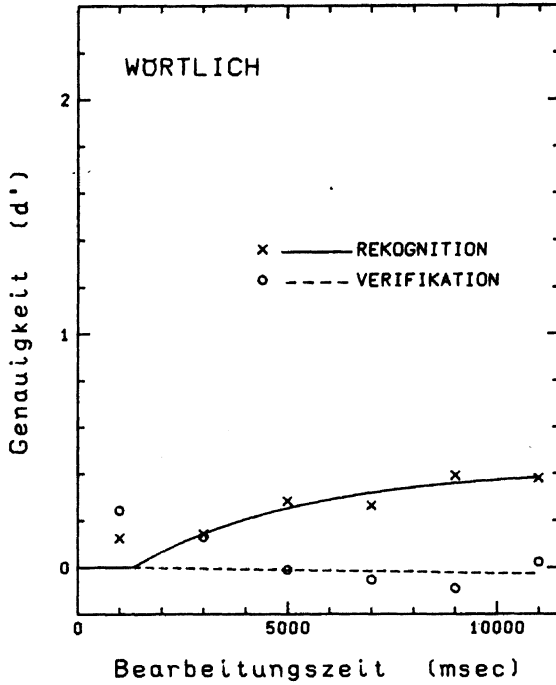
Um den zeitlichen Verlauf des Gedächtnisabrufs vereinfacht darzustellen, wurden die empirischen d' -Werte durch eine dreiparametrische Funktion approximiert (vgl. Wickelgren et al. 1980):

$$d(t) = \begin{cases} \lambda (1 - \exp - \{ -\beta [t - \delta] \}), & \text{falls } t > \delta \\ 0, & \text{falls } t \leq \delta. \end{cases} \quad (5)$$

In dieser Funktion stellt $d(t)$ die d' -Genauigkeit zum Zeitpunkt t dar. Der Parameter δ spezifiziert die minimale Bearbeitungszeit, die erforderlich ist, um eine d' -Genauigkeit zu erhalten, die sich von Null unterscheidet. Der

Abb. 3

Resultat des Gedächtnisabrufs von wörtlicher, propositionaler und situativer Information nach verschiedenen Bearbeitungszeiten bei Wiedererkennungs- und Verifikationsaufgaben



Parameter λ gibt die Asymptote an und β ist der Wachstumsparameter, der die Schnelligkeit beschreibt, mit der diese Asymptote erreicht wird. Die drei Funktionsparameter wurden aus den vorliegenden Daten mit Hilfe des STEPIT-Programms (Chandler, 1965) und einem Kleinstquadrat-Kriterium geschätzt. Für die beiden Aufgabeninstruktionen wurde bei jeder Darstellungsform jeweils ein identischer δ -Parameter postuliert. Die aus der Schätzung resultierenden kontinuierlichen Funktionen, die in Abbildung 3 dargestellt sind, weisen darauf hin, daß situative Information schneller abgerufen wird als propositionale Information. Dies gilt sowohl für die Wiedererkennung- als auch für die Verifikationsaufgabe. In der Terminologie von Reders Modell bedeutet dies, daß nichtmonotone Verlaufscharakteristiken beim Verifizieren von B-Sätzen dadurch zustandekommen, daß zwei Abrufkomponenten verwendet werden, die sich durch ihren Zeitbedarf und bezüglich ihres Ergebnisses unterscheiden. Während beim Verifizieren von B-Sätzen die wörtlichen und propositionalen Vergleiche zu langsamen und falschen Antworten führen, erzeugen die schnelleren Plausibilitätsurteile, welche auf situative Informationen zurückgreifen, schnelle und korrekte Antworten.

Abbildung 3 stützt darüber hinaus die Annahme, daß die unterschiedlichen Informationsarten in sehr ähnlicher Weise sowohl zum Wiedererkennen als auch zum Verifizieren von Sätzen verwendet werden. Obwohl für die asymptotischen Abrufresultate von wörtlicher, propositionaler und situativer Information eine signifikante Interaktion zwischen der Aufgabenstellung und den drei Informationsarten gefunden wurde, $F(2,252) = 4.68$, $MSE = 1.75$, $p < 0.05$, so sind die Unterschiede zwischen den Informationsarten, $F(2,252) = 58.50$, $MSE = 1.75$, $p < 0.0001$, doch sehr viel größer als der Unterschied zwischen den beiden Aufgabenanforderungen, $F(1,126) = 3.21$, $MSE = 1.20$, $p = 0.08$.

Schlußfolgerungen

Die Analyse des Strategieauswahlmodells hat gezeigt, daß die Richtigkeit einer Antwort beim Verifizieren von B-Sätzen, die nicht im Text vorkamen, zuerst mit zunehmender Bearbeitungszeit bis zu einem Maximum ansteigen sollte, um danach wieder abzufallen. Daraus ergibt sich, daß eine einfache Latenzzeitregistrierung für die Überprüfung des Strategieauswahlmodells nicht geeignet ist. Gleichzeitig ergibt sich dadurch jedoch auch eine neue Überprüfungsmöglichkeit von Reders Modell. Anhand des vorgeschlagenen Verfahrens kann die Hypothese, daß zum Wiedererkennen und Verifizieren von Sätzen mehr als ein Prozeß verwendet wird, evaluiert werden. Wenn die durchschnittliche Richtigkeit von Antworten mit zunehmender Bearbeitungszeit bis zu einem Maximum ansteigt und danach wieder ab-

fällt, so darf angenommen werden, daß (mindestens) zwei Komponenten, welche sich bezüglich der Bearbeitungszeit und bezüglich des Bearbeitungsergebnisses unterscheiden, an der Beantwortung der Aufgabe beteiligt sind. Die Brauchbarkeit dieser Methode wurde durch die Erstellung empirischer Verlaufscharakteristiken demonstriert.

Die Ergebnisse zeigen, daß Vpn wörtliche und propositionale Informationen zum Wiedererkennen und Verifizieren von Sätzen heranzogen, daß jedoch situative Informationen, die vermutlich durch Plausibilitäsurteile erstellt werden, den größten Beitrag zu den Antworten der Vpn lieferten. Reder postulierte, daß die Vpn zum Zeitpunkt des Verifizierens bzw. Wiedererkennens eines Satzes situative Information aus der Textenkodierung extrahieren, um darauf aufbauend ein Plausibilitäsurteil treffen zu können. Die Ergebnisse des vorliegenden Experiments weisen allerdings darauf hin, daß situative Informationen schneller abgerufen werden als wörtliche oder propositionale Informationen. Aus diesem Grund dürfte es unwahrscheinlich sein, daß die Vpn erst zum Testzeitpunkt situative Informationen aus der propositionalen Textenkodierung gewinnen. Es ist vielmehr anzunehmen, daß bereits beim Textverstehen eine Darstellung der realen oder hypothetischen Situationen, auf die sich ein Text bezieht, ausgebildet werden. Diese Annahme wurde von van Dijk & Kintsch (1983) aufgestellt und theoretisch begründet. Diese Autoren postulieren, daß die Konstruktion und das Erweitern von situativen Enkodierungen einen Teilprozeß des Textverstehens darstelle. Durch die Analyse von Lesezeiten konnten Schmalhofer und Glavanov auch die während der Textenkodierung stattfindende Konstruktion situativer Darstellungen nachweisen.

Somit soll angenommen werden, daß wörtliche, propositionale und situative Informationen bereits während des Textenkodierens gespeichert werden. Im Vergleich zu wörtlichen und propositionalen Gedächtnisspuren kann situative Information schneller und leichter abgerufen werden. Dadurch könnte auch erklärt werden, weshalb Personen zum Wiedererkennen von Sätzen situative Informationen heranziehen. Zusammenfassend läßt sich sagen, daß die Ergebnisse der berichteten Untersuchung zeigen, daß das Erstellen von Verlaufscharakteristiken einen größeren Einblick in den zeitlichen Verlauf kognitiver Prozesse erlaubt, als der Vergleich mittlerer Latenzzeiten.

Summary

An analysis of Reder's (1982) strategy selection model showed that for the verification of plausible sentences which did not occur in the text, response accuracy should first increase with processing time up to some maximum accuracy and decrease thereafter. It was shown that because of this non-monotonic speed accuracy operating characteristic, Reder's re-

sponse latency experiments cannot be considered an appropriate test of her model. Alternatively, speed accuracy operating or time-course characteristics were suggested for testing the strategy selection assumption. An application of this method indicated that the time-course characteristic for the verification of plausible sentences was indeed non-monotonic, supporting Reder's two-process assumption. Furthermore, the results showed that similar strategies were being used in both the recognition and verification of sentences. In both cases, the subjects' responses relied more heavily upon the general situational descriptions addressed by the text than upon propositional or verbatim text information.

Literatur

- Baird, C. B. & Noma, E.: Fundamentals of scaling and psychophysics. New York: John Wiley & Sons, 1978.
- Chandler, P. J.: Subroutine STEPIT: An algorithm that finds the values of the parameters which minimize a given continuous function. Bloomington: Indiana University, Quantum Chemistry Program Exchange, 1965.
- Dosher, B. A.: Effect of sentence size and network distance on retrieval speed. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 1982, 173—207.
- Kintsch, W.: The representation of meaning in memory. Hillsdale, NJ: Erlbaum, 1974.
- Kintsch, W.: Beiträge zu einer Theorie des Textverstehens, in Bosshart, H. G. (Hrsg.) *Sprachpsychologie: In memoriam Hans Hörmann (im Druck)*.
- Laming, D. R.: Information theory of choice-reaction times. London: Academic Press, 1968.
- Link, S.: Correcting response measures for guessing and partial information. *Psychological Bulletin*, 92, 1982, 469—486.
- Pachella, R. G.: The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Erlbaum, 1974.
- Reder, L. M.: Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, 89, 1982, 250—280.
- Schmalhofer, F.: Experimente zur gleichzeitigen Messung von sensorischer Beanspruchung und Aufmerksamkeit. Diplomarbeit, Universität Regensburg, 1978.
- Schmalhofer, F.: Comprehension of a technical text as a function of expertise. Doctoral Dissertation, University of Colorado, University Microfilms International, Pub.No.: 83-09-879, 1982.
- Schmalhofer, F. & Glavanov, D.: Three components of understanding a programmer's manual: Verbatim, propositional and situational representations. *Journal of Memory and Language (formerly, Journal of Verbal Learning and Verbal Behavior)* (in press).
- van Dijk, T. A. & Kintsch, W.: Strategies of discourse comprehension. New York: Academic Press, 1983.
- Wender, K., Colonius, H. & Schulze, H. H.: Modelle des menschlichen Gedächtnisses. Stuttgart: Kohlhammer, 1980.
- Wickelgren, W. A., Corbett, A. T. & Dosher, B. A.: Priming and retrieval from short-term memory: A speed accuracy trade-off analysis. *Journal of Verbal Learning and Verbal Behavior*, 19, 1980, 387—404.

Anschrift des Verfassers: Franz Schmalhofer, Psychologisches Institut der Universität Freiburg, Niemensstraße 10, 7800 Freiburg i. Br.