

Diskrete dynamische Systeme: Kompositionen höherer Ordnung bei Bediensequenzen

Burkhard Müller und Joachim Funke

Psychologisches Institut der Universität Bonn

Schlüsselwörter: Diskrete dynamische Systeme, Bediensequenzen, Komposition, finite Grammatik, competitive chunking

Zusammenfassung: Die Theorie zum „competitive chunking“ von Servan-Schreiber und Anderson (1990) wurde auf den Bereich diskreter dynamischer Systeme angewandt, um zu prüfen, inwieweit auch bei Bediensequenzen die Bildung von Chunks höherer Ordnung („Phrasen“) nachweisbar ist. In einer Memorierphase mußten zunächst 48 aufgrund einer finiten Grammatik generierte Bediensequenzen in 24 Sets á zwei Sequenzen zweimal richtig reproduziert werden. Neben einer Kontrollbedingung, in der von vornherein die vollständigen Sequenzen zu reproduzieren waren, wurden als weitere Bedingungen realisiert, daß (a) zunächst nur die ersten beiden Eingabepaare („Worte“), (b) zunächst nur die letzten beiden Eingabepaare der beiden Sequenzen reproduziert werden mußten. In einer anschließenden für die Vpn unerwarteten Diskriminationsphase mußte für bis dahin unbekannte Bediensequenzen entschieden werden, ob sie einer eventuell wahrgenommenen Systematik entsprachen. Die Ergebnisse belegen die Bildung von Chunks höherer Ordnung. Die multinomiale Modellierung des Urteilsprozesses verdeutlicht jedoch auch, daß im Unterschied zum Originalexperiment die Tendenz bestand, die *letzten* beiden „Worte“ zu einer „Phrase“ zusammenzufassen. Als Erklärung für diesen Effekt wird im Rahmen der Theorie des „competitive chunking“ der Einfluß der zeitlich sukzessiven Darbietung der Sequenzen diskutiert. Die Interpretierbarkeit der Befunde wird als Beleg dafür angesehen, daß der Ansatz finiter Automaten erfolgreich dazu genutzt werden kann, die häufig gegenläufigen Erfordernisse objektiverer Erkenntnisse und ökologischer Validität zu integrieren.

Discrete Dynamic Systems: Higher Order Composition of Operating Sequences

Key words: Discrete dynamic systems, operating sequences, composition, finite grammar, competitive chunking

Summary: The theory of competitive chunking (Servan-Schreiber & Anderson, 1990) was utilized in the realm of discrete dynamic systems to investigate whether the creation of higher order chunks can be shown for sequences of operations. In a memorization task 48

sequences of a finite grammar in 24 sets with 2 sequences each had to be reproduced correctly for two times. In the control condition only whole sequences had to be reproduced. In the experimental conditions whole sequences had to be reproduced after (a) the first and second pair of operations of a sequence or (b) the second and third pair of operations of a sequence were reproduced correctly for two times. In an unexpected discrimination task subjects had to judge for new sequences whether the sequence was compatible or incompatible with a perhaps perceived systematicity of the previous learned sequences. Results confirm the creation of higher order chunks. A multinomial model of the grammatical decisions, however, shows as clear difference to the original experiment that there was a tendency to chunk the *last* two "words" of sequences to "phrases". As explanation within the theory of competitive chunking the impact of temporal succession of presentation is discussed. We interpret the findings as indicating, that the framework of finite automata can successfully be used to integrate the often divergent necessities of generalizability of knowledge and ecological validity.

Umgang mit diskreten dynamischen Systemen

Täglich hantieren wir mit Videorecordern, Waschmaschinen, Fahrkartenautomaten, Software oder ähnlichen Systemen. Der Umgang mit diesen Systemen verlangt, daß ausgehend von einem Anfangszustand durch entsprechende Maßnahmen Folgezustände erreicht werden, die angestrebte Handlungsziele realisieren. Die Veränderung von Zuständen dieser Systeme und darauf bezogene Interventionsmöglichkeiten lassen sich im Rahmen der Theorie endlicher Automaten allgemein beschreiben (vgl. Funke & Buchner, 1992). Diese Theorie bietet den Vorteil, die Struktur diskreter dynamischer Systeme in Form von Übergangsmatrizen darstellen zu können, die aus der Sicht von Benutzern Anfangszustände, Interventionen und Folgezustände in Beziehung setzen. Als Basiselemente des Bedienwissens können Tripel, bestehend aus Anfangszustand, Intervention und Folgezustand, angesehen werden, die für spezifische Bedienziele sequentiell kombiniert werden müssen. Aus dieser Annahme ergibt sich die Möglichkeit, Wissen über das System und seine Bedienung dadurch zu diagnostizieren, daß (1) bei vorgebenem Anfangs- und Folgezustand nach der nötigen Intervention gefragt wird (Interpolationsfrage), (2) bei vorgebenem Anfangszustand und einer Intervention nach dem Folgezustand gefragt wird (Prognosefrage) oder (3) bei vorgegebener Intervention und Folgezustand nach dem Anfangszustand gefragt wird (Retrognosefrage). Ein weiterer Vorteil dieser Konzeptualisierung diskreter dynamischer Systeme besteht darin, daß für einzelne Bearbeitungsziele aus den Übergangsmatrizen algorithmisch optimale Sequenzen von Interventionen ableitbar sind. Damit kann in Untersuchungen zu Erwerb und Anwendung von Wissen die Performanz von Versuchspersonen objektiv bestimmt werden. Die Analyse der Übergangsmatrizen kann auch zur Bewertung und Verbesserung von Instruktionen genutzt werden. So konnten Funke und Gerdes (1993) am Beispiel von Videorecordern zeigen, daß Funktionsbeschreibungen in zugehörigen Manualen an wichtigen Stellen unvollständig und sogar fehlerhaft waren. Der Vergleich von Gruppen, die entweder das Gerät unter Verwendung des Originalmanuals oder einer auf der Grundlage dieser Analyse verbesserten Version zu bearbeiten hatten, erbrachte klare Vorteile für die Gruppe mit der verbesserten Variante.

In Erweiterung der genannten Annahmen über die Repräsentation von Wissen über die Bedienung diskreter dynamischer Systeme wird davon ausgegangen, daß mit zunehmender Erfahrung im Umgang mit einem spezifischen System Bedienwissen als serielle Ordnungsinformation über mögliche „Zustand-Intervention-Folgezustand“-Übergänge repräsentiert wird. Hinweise auf derartige Repräsentationen von Systemwissen haben sich z.B. in Verifikationsaufgaben ergeben, bei denen beurteilt werden mußte, ob vorgegebene Zustandsübergänge des Systems zulässig sind (Buchner & Funke, 1993). Die Urteilszeiten waren dann signifikant geringer, wenn die Präsentationsreihenfolge der einzelnen zu beurteilenden Übergänge der „natürlichen“ Sequentialität des Systems entsprach. In weiteren Experimenten hat sich gezeigt, daß für Wissen über Bediensequenzen analog zu klassischen Befunden sequentiellen Lernens Kompositionseffekte nachweisbar sind, die die Strukturierung einer Bedienoberfläche mit Hilfe visueller und akustischer Merkmale reflektieren (Müller, Funke & Buchner, 1994): Das Erlernen aufeinanderfolgender Bedienschritte führte insbesondere dann zu Transfervorteilen bei einer neu zu erlernenden Bediensequenz, wenn die übertragbaren Schritte aus Operationen *innerhalb* eines einheitlich markierten Bereichs der Bedienoberfläche bestanden. Dieser Effekt belegt, daß nicht lediglich einfache Assoziationen zwischen zeitlich aufeinanderfolgenden Bedienschritten gebildet werden, sondern Wissen über Bediensequenzen in sogenannten Chunks (Miller, 1956) organisiert wird.

Chunk-Bildung und Komposition von Wissensstrukturen

Das Konzept des Chunks umfaßt den Sachverhalt der subjektiven Organisation einzelner Wissensbestandteile in neuen, übergeordneten Einheiten und wurde von Miller (1956) zur Erklärung des inhaltsvarianten Informationsgehaltes bei angenommener Kapazitätsbeschränkung kurzzeitiger Speicherung eingeführt. In modernen Theorien zum Erwerb von Fertigkeiten bezeichnet Chunk-Bildung oder Komposition Prozesse der Wissensorganisation, die dem ursprünglichen Konzept entsprechend als Zusammenballungen von Elementen angenommen werden, um bei begrenzter Verarbeitungskapazität die Performanzsteigerung mit zunehmender Übung zu erklären (z.B. Anderson, 1983; Lewis, 1987; MacKay, 1982; Neves & Anderson, 1981; Newell, 1990). Gegenüber einer einfachen assoziativen Verknüpfung aufeinanderfolgender Informationen und Aktionen wird dabei angenommen, daß Basiselemente des Wissens zu neuen Einheiten zusammengefaßt werden, aus denen auf einer höheren Ebene wiederum neue Einheiten gebildet werden.

In einer neueren Arbeit stellen Servan-Schreiber und Anderson (1990) ihre Theorie des „competitive chunking“ vor, mit der sie Phänomene des nicht-intentionalen Lernens strukturierten Lernmaterials (z.B. regelhafte Buchstabenreihen) modellieren konnten. Die Autoren beschreiben die für derartige Experimente typischen Lernprozesse als Etablierung hierarchischer Netzwerke von Chunks, deren unterste Schicht aus einzelnen, bekannten Ereignissen (z.B. den Buchstaben der Sequenz S-V-P) besteht. Diese Hierarchie entsteht dadurch, daß zunächst einzelne Elemente einer Sequenz (z.B. Buchstaben) aufgrund ihrer Kontingenz zu Sub-Chunks zu-

sammengefaßt werden. Aus den Sub-Chunks werden wiederum Chunks höherer Ordnung gebildet. Die Anzahl der Chunks im resultierenden Perzept einer Sequenz ist ein Maß für die Kompaktheit der Stimulusrepräsentation und damit auch zugleich für dessen Vertrautheit. Auf der Basis der Vertrautheit sollen dann die Grammatikalitätsurteile erfolgen. Mit Hilfe dieser Annahmen waren die Autoren in der Lage, Lerneffekte klassischer Experimente in diesem Bereich (Miller, 1958; Reber, 1967) zu modellieren. Im Unterschied zu Miller (1956), der Chunk-Bildung als bewußte Rekodierung von Information auffaßte, und der Auffassung von Reber (1967, 1976), wonach das unbewußte Erfassen von abstrakten Regularitäten des Reizmaterials sogar noch leichter fallen soll, gehen Servan-Schreiber und Anderson davon aus, daß die Lernprozesse bei der Bildung von Chunks sequentieller Ereignisse immer unbewußt ablaufen und lediglich deren Ergebnisse mehr oder weniger bewußt sein können.

Da wir uns in der vorliegenden Arbeit an Theorie und experimentellem Paradigma von Servan-Schreiber und Anderson (1990) orientieren, sei deren Vorgehen etwas ausführlicher dargestellt. Das Experiment zum Nachweis von Kompositionen höherer Ordnung (Experiment 2) gliederte sich in zwei Phasen. In einer Memorierphase mußten Versuchspersonen (Vpn) unter drei verschiedenen Bedingungen wiederholt jeweils drei vorgegebene Sätze richtig reproduzieren. Die einzelnen Sätze bestanden jeweils aus drei Paaren von Konsonanten (z.B. HK LF JZ), die als Worte dienten. Zur Induktion von Chunks auf der Wortebene wurden jeweils zwei Buchstabenpaare („Worte“) durch einen Zwischenraum getrennt. In einer unerwarteten Diskriminationsphase mußte dann für unterschiedliche Typen von Sätzen entschieden werden, ob sie einer eventuell wahrgenommenen Systematik entsprachen oder nicht. Die verschiedenen Sätze wurden mit Hilfe einer finiten Grammatik erstellt (Abbildung 1).

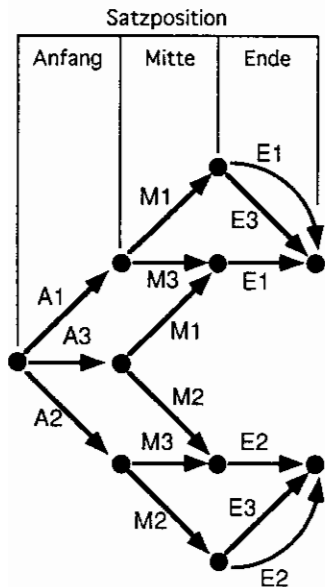


Abb. 1 Grammatik der Buchstabenfolgen von Servan-Schreiber und Anderson (1990, Experiment 2). Jedes Symbol der Satzpositionen (A1-3, M1-3, E1-3) konnte durch zwei verschiedene Buchstabenpaare („Worte“) instantiiert werden, so daß durch die Grammatik 64 verschiedene Buchstabenfolgen generierbar sind.

Für jede der drei Positionen eines Satzes (Anfang, Mitte, Ende) wurden drei verschiedene Wortklassen gebildet, die aus jeweils zwei Worten bestanden (= 18 verschiedene Worte). Als Grundtypen dienten die Sätze mit den Worten aus Wortklasse 1 der jeweiligen Position (z.B.: HK LF JZ) und die mit den Worten aus Wortklasse 2 (z.B.: NG VS TQ). Die Sätze der Memorierphase wurden dadurch gebildet, daß bei den Grundtypen (z.B.: HK LF JZ) das erste, zweite oder dritte Wort durch ein Wort aus Wortklasse 3 ersetzt wurde (z.B.: PX LF JZ, HK BR JZ oder HK LF DW). Daraus ergeben sich 48 verschiedene Sätze, die in 16 Sets mit jeweils 3 Sätzen vorgegeben wurden. Eine Besonderheit bestand darin, daß bei jeweils zwei aufeinanderfolgenden Sets (= 6 Sätze) die einzelnen Sätze aus anderen Worten bestanden. Um die Bildung von Phrasen zu prüfen, wurde zwischen Versuchsgruppen variiert, ob bei einem Set die einzelnen Sätze von vornherein vollständig oder zunächst nur erstes und zweites bzw. zweites und drittes Wort zu reproduzieren waren.

In der sich anschließenden Diskriminationsphase wurden entweder solche Sätze vorgegeben, die den beiden Grundtypen entsprachen (z.B.: HK LF JZ oder NG VS TQ), oder solche, bei denen das erste, zweite oder dritte Wort durch ein Wort des jeweils anderen Grundtyps ersetzt worden war (z.B.: NG LF JZ, HK VS JZ oder HK LF TQ). Die verschiedenen Satztypen unterschieden sich also nicht in der Vertrautheit der einzelnen Worte, sondern in der Überlappung von Wortpaaren („Phrasen“) mit Sätzen der Memorierphase. Es zeigte sich erwartungsgemäß für die Gruppen, die in der Memorierphase zunächst nur Teilsätze reproduzieren mußten, daß die Ablehnungsquoten in der Diskriminationsphase gegenüber denen für die Grundtypen insbesondere bei den Sätzen erhöht war, bei denen die vermutete Phrasenstruktur verletzt war. Für die Kontrollgruppe zeigte sich, daß tendenziell die Zusammenfassung von erster und zweiter Position zu einer Phrase wahrscheinlicher ist als die von zweiter und dritter Position.

Im Rahmen des vorliegenden Experimentes wurde der Frage nachgegangen, inwieweit die Vorstellungen von Servan-Schreiber und Anderson (1990) zur Bildung von Chunks höherer Ordnung auch für die Bedienung diskreter dynamischer Systeme gelten. Um im direkten Vergleich prüfen zu können, inwieweit die Theorie des „competitive chunking“ auf eine Situation generalisierbar ist, in der nicht Buchstaben, sondern Bedienschritte miteinander zu assoziieren bzw. zu komponieren sind, wurde versucht, möglichst weitgehend die experimentellen Bedingungen beizubehalten, die die Autoren gewählt haben.

Methode

Das Experiment bestand aus einer Memorier- und einer anschließenden Diskriminationsphase. In der Memorierphase bestand die Aufgabe für die Versuchspersonen (Vpn) darin, zwei dargebotene Tastenabfolgen durch Eingaben mit Hilfe eines Mauszeigers zu reproduzieren. Die Sequenzen umfaßten sechs Tasten, die jeweils paarweise aufeinanderfolgend in drei von links nach rechts nebeneinander angeordneten Bereichen plaziert waren. Die Bereiche waren durch unterschiedliche visuelle Hintergrundmuster der entsprechenden Tasten gekennzeichnet. In der anschließenden

den und für die Vpn unerwarteten Diskriminationsphase waren einzeln vorgeführte Sequenzen danach zu klassifizieren, ob sie einer eventuell wahrgenommenen Systematik entsprachen. Die Sequenzen der beiden Phasen waren nach der gleichen Grammatik wie bei Servan-Schreiber und Anderson (1990) strukturiert (vgl. Abbildung 1) und wurden in der Diskriminationsphase durch solche ergänzt, die die Struktur der finiten Grammatik auf höherer Ebene verletzten.

Versuchsmaterial

Die „Bedienoberfläche“ wurde auf dem Bildschirm eines Macintosh SE Computers als horizontale Anordnung von ovalen Tasten in drei visuell gut unterscheidbaren Bereichen dargestellt (Abbildung 2).

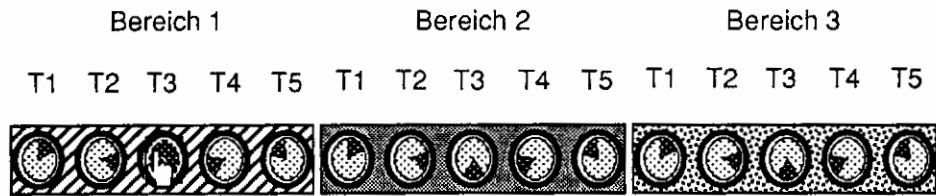


Abb. 2 Bildschirmdarstellung der „Bedienoberfläche“ mit jeweils drei Bereichen zu je fünf Tasten. Die erste „anzuklickende“ Taste wird angezeigt

Die Bezeichnungen in Abbildung 2 sind zur Erläuterung der Sequenzstruktur angegeben und waren im Experiment nicht zu sehen. Jeder Bereich umfaßte fünf Tasten, die zur besseren Diskriminierbarkeit innerhalb der Bereiche durch spezifische Symbole gekennzeichnet waren und bei entsprechender Eingabe spezifische Töne auslösten. Tasten unterschiedlicher Bereiche mit der gleichen Position innerhalb des Bereiches wiesen das gleiche Symbol auf und eine Eingabe löste den gleichen Ton aus.

Zur Generierung grammatisch strukturierter Sequenzen wurden für jeden der drei Bereiche drei verschiedene Klassen von paarweisen Tastenabfolgen gebildet (Tabelle 1).

Tab. 1 Tastenpaare nach Bereichen und Klassen

Klasse	Bereich		
	1 (Anfang)	2 (Mitte)	3 (Ende)
I	T1 - T4	T3 - T4	T1 - T3
	T4 - T1	T4 - T3	T3 - T1
II	T2 - T5	T1 - T2	T3 - T5
	T5 - T2	T2 - T1	T5 - T3
III	T1 - T5	T2 - T3	T2 - T4
	T5 - T1	T3 - T2	T4 - T2

Jede Klasse eines Bereiches bestand aus zwei Tastenpaaren, die sich lediglich durch die zeitliche Abfolge der Tasten unterschieden (z.B. T1 - T4 vs. T4 - T1 für Klasse I und Bereich 1). Grundtypen bestanden aus drei Tastenpaaren der für den jeweiligen Bereich ersten oder zweiten Klasse. Die räumlichen Abstände zwischen den Tasten eines Bereiches war für diese beiden Klassen identisch (T1 - T4 vs. T2 - T5 für Bereich 1), um eine einfache Diskriminationsmöglichkeit auszuschließen.

Die Sequenzen der Memorierphase wurden dadurch erstellt, daß bei den Grundtypen entweder das erste, zweite oder dritte Tastenpaar durch ein Paar aus der dritten Klasse ersetzt wurde (z.B.: T1 - T5, T3 - T4, T1 - T3; T1 - T4, T2 - T3, T1 - T3 oder T1 - T4, T3 - T4, T2 - T4). Daraus ergeben sich 48 Sequenzen, die paarweise in 24 Sets vorgegeben wurden. Bei der Zusammenstellung der Sets wurde analog zum Originalexperiment berücksichtigt, daß die Sequenzen eines Sets aus verschiedenen Tastenpaaren bestanden und ein und dasselbe Tastenpaar nur in jedem dritten Set enthalten war. Diese kombinatorischen Beschränkungen wurden dadurch realisiert, daß bei jeweils drei aufeinanderfolgenden Sequenzen neben den Tastenpaaren der Klasse 3 an der ersten, zweiten oder dritten Position an den jeweils verbleibenden Positionen entweder nur Tastenpaare aus der Klasse 1 der jeweiligen Position oder nur Tastenpaare der Klasse 2 der jeweiligen Position vorkamen. Im Unterschied zum ersten und letzten von drei aufeinanderfolgenden Sets á zwei Sequenzen waren bei den Sequenzen des mittleren Sets Tastenpaare aller drei Klassen beteiligt. Innerhalb der Versuchsbedingungen wurde je zur Hälfte variiert, ob bei drei aufeinanderfolgenden Sets von Sequenzpaaren generell das erste aus Tastenpaaren der Klasse 1 und 3 und das letzte aus Tastenpaaren der Klasse 2 und 3 bestanden oder umgekehrt. Die Abfolge der acht verschiedenen Sechsergruppen (= 3 Sets) wurde zufällig variiert.

Als Sequenzen der Diskriminationsphase wurden zum einen die Grundtypen (NG) eingesetzt, zum anderen wurden solche vorgegeben, bei denen das erste (EA), das zweite (EM) oder das dritte Paar (EE) eines Grundtyps durch ein entsprechendes Paar des anderen Grundtyps ersetzt wurde (z.B.: T2 - T5, T3 - T4, T1 - T3; T1 - T4, T1 - T2, T1 - T3 oder T1 - T4, T3 - T4, T3 - T5). Ebenfalls analog zum Originalexperiment wurden in der Diskriminationsphase die neuen grammatischen Sequenzen zweimal, die anderen Sequenzen nur einmal dargeboten, so daß 32 neue grammatische und jeweils 16 der drei Mischtypen zur Beurteilung vorgegeben wurden. Die Abfolge der Sequenzen war in zwei Blöcke unterteilt, die jeweils die gleichen Anteile an neuen grammatischen und den verschiedenen Mischtypen aufwiesen. Innerhalb der Blöcke war die Reihenfolge der Sequenzen unsystematisch variiert und für alle Vpn gleich.

Versuchsplan

Als unabhängige Variable wurde für die Memorierphase unterschieden, ob zunächst nur die erste und die zweite paarweise Abfolge von Tasten (Anfang-Mitte), die zweite und dritte (Mitte-Ende) und erst danach die gesamte Sequenz oder von vornherein die gesamte Sequenz (Kontrolle 1) reproduziert werden mußte. Eine weitere Gruppe (Kontrolle 2) mußte lediglich die Diskriminationsphase bearbeiten.

Als *abhängige Variablen* wurden in der Memorierphase die Zahl der Wiederholungen bis zum Erreichen des Lernkriteriums erfaßt. In der Diskriminationsphase wurden die Ablehnungsquoten (Ja/Nein-Paradigma) für die verschiedenen Sequenztypen sowie die Latenzzeiten für Ablehnung und Zustimmung erfaßt. Neben dem direkten Vergleich werden die Ablehnungsquoten im Rahmen eines multinomialen Modells (Riefer & Batchelder, 1988) analysiert. Diese Modellierung bietet die Vorteile, daß zum einen die Wahrscheinlichkeiten der Phrasenbildung sowie die Antwortneigung getrennt ermittelt werden können und zum anderen die Effekte mit *einem* Test der Modellanpassung inferenzstatistisch abgesichert werden können (vgl. Vaterrodt-Plünnecke, 1994).

Hypothesen

Entsprechend den Ergebnissen von Servan-Schreiber und Anderson (1990) wird erwartet, daß sich die Gruppen nicht in der Zahl der Wiederholungen unterscheiden, die in der Memorierphase pro Set nötig sind, um das Lernkriterium zu erreichen. Dagegen sollten sich zwischen den Gruppen Unterschiede in der Diskrimination der verschiedenen Sequenztypen zeigen. Diejenigen, die zunächst Anfang und Mitte eines Sequenzpaares reproduzieren mußten, sollten gegenüber neuen grammatischen Sequenzen solche besser diskriminieren können, bei denen das erste oder zweite Tastenpaar ausgetauscht wurde. Diejenigen, die zunächst Mitte und Ende eines Sequenzpaares reproduzieren mußten, sollten dagegen gegenüber den Grundtypen solche Sequenzen besser diskriminieren können, bei denen das zweite oder dritte Paar ausgetauscht wurde. Bei der Kontrollgruppe 1 sollte sich ein Muster zeigen, das mit der ersten Gruppe vergleichbar ist, die Diskrimination sollte jedoch geringer ausgeprägt sein.

Die Rekonstruktion der Ablehnungsquoten im Rahmen des multinomialen Modells sollte für die drei eben genannten Gruppen erhöhte Neigungen zu „Ja“-Antworten ausweisen, da die einzelnen Sequenzelemente sehr vertraut sind. Bei der Kontrollgruppe 2, die lediglich die Sequenzen der Diskriminationsphase bearbeitet hat, wird demgegenüber erwartet, daß aufgrund der höheren Häufigkeit von neuen grammatischen Abfolgen ein zu Kontrollgruppe 1 vergleichbares Muster in der Diskrimination der verschiedenen Sequenztypen festzustellen ist, jedoch keine erhöhte Neigung zu „Ja“-Antworten eintritt.

Versuchspersonen

In Anlehnung an Servan-Schreiber und Anderson (1990) wurden 12 Vpn für jede der vier Bedingungen des Versuchsplans vorgesehen. Damit kann nach Cohen (1977) für einen einseitigen *t*-Test ein Effekt großer Stärke ($d = 0.80$) unter fairen Testbedingungen ($\alpha = \beta$) nur mit $\alpha = .17$ aufgedeckt werden; bei konventionellem $\alpha = .05$ resultiert ein $\beta = .40$. Für die Testung der Modellpassung ($N = 12 \cdot 80 = 960$) kann für das χ^2 -verteilte Divergenzmaß von Read und Cressie (1988) bei einem aufzudeckenden schwachen Effekt ($\omega = .1$) bei $\alpha = .05$ ein $\beta = .13$ realisiert werden.¹

¹ Die Berechnungen beruhen auf dem Programm G-Power (Version 1.2; Buchner, Faul & Erdfelder, 1992).

Als Vpn konnten durch Aushänge 15 männliche und 33 weibliche Studierende der Universität Bonn angeworben werden, die zwischen 18 und 36 Jahren alt waren ($M = 23.52$; $SD = 4.27$). Die Teilnahme am Experiment wurde mit 20,- DM oder einer für das Studium erforderlichen Teilnahmebescheinigung honoriert.

Versuchsablauf

Das Experiment wurde als Einzelversuch durchgeführt. Nach einer einleitenden Instruktion hatten die Vpn der experimentellen Gruppen in einer Übungsphase Gelegenheit, sich anhand einer Sequenz mit sechs Eingaben an die Versuchsumgebung sowie den Mauszeiger als Eingabeinstrument zu gewöhnen.

In der *Memorierphase* wurden je nach Bedingung paarweise vollständige Sequenzen oder Teilsequenzen vorgeführt, die dann von der Vpn reproduziert werden mußten. Bei der Vorführung erschien das Symbol des Mauszeigers auf der entsprechenden Taste auf dem Bildschirm, worauf die Taste für ca. 333 ms invertiert dargestellt wurde. Weitere 250 ms später erschien der Mauszeiger auf der folgenden Taste, die danach ebenfalls invertiert dargestellt wurde usw. Beim Wechsel zur zweiten Sequenz erschien der Mauszeiger nicht nach 250, sondern nach 500 ms. Der Beginn des Reproduktionsversuchs wurde durch einen Signalton angezeigt. Bei fehlerhaften Eingaben wurde immer die richtige Taste invertiert. Am Ende eines Reproduktionsversuchs wurde die Anzahl richtiger Eingaben zurückgemeldet. Im Anschluß daran wurde die entsprechende Sequenz von Tasten noch einmal vorgeführt und erneut Gelegenheit zur Reproduktion gegeben. Die Vorgaben mußten insgesamt zweimal richtig reproduziert werden. Dieses gegenüber den Untersuchungen von Servan-Schreiber und Anderson erhöhte Lernkriterium erschien sinnvoll, um die im Gegensatz zu den 21 Konsonanten des Alphabets bei fünf Tasten eines Bereiches erhöhte Wahrscheinlichkeit einer zufällig richtigen Reproduktion zu kompensieren. Im Anschluß an die zweimalige richtige Reproduktion wurden in den Bedingungen Anfang-Mitte (A-M) und Mitte-Ende (M-E) die vollständigen Sequenzen vorgegeben, in der Bedingung Kontrolle 1 wurde das nächste Set präsentiert. Nach jedem dritten Set wurden die Vpn zu einer längeren Pause ermuntert, um Auswirkungen von Ermüdung zu reduzieren.

In der *Diskriminationsphase* wurden die Vpn darauf hingewiesen, daß die Sequenzen der Memorierphase systematisch strukturiert waren. Sie wurden aufgefordert, im folgenden für einzeln vorgegebene Sequenzen zu entscheiden, ob sie dieser Systematik entsprächen oder nicht. Die Vpn wurden ermuntert, ihre Entscheidung „intuitiv“ zu fällen, wenn sie keine Idee hätten, worin die Systematik bestand. Die Sequenzen wurden wie in der Memorierphase auf dem Bildschirm präsentiert, auf dem in dieser Phase jedoch nur eine Zeile mit drei nebeneinander angeordneten Bereichen zu sehen war. Nach der Präsentation erschien ein Dialog, der mit der Frage „Entsprach die Folge der Systematik?“ überschrieben war und als Antwortalternativen Tasten mit den Bezeichnungen „JA“ und „NEIN“ zeigte. Der Mauszeiger wurde zwangsweise so positioniert, daß er zu beiden Tasten den gleichen Abstand aufwies. Die Antworten und Latenzzeiten wurden aufgezeichnet.

Ergebnisse

Zunächst werden die Ergebnisse aus der Memorierphase berichtet, für die Diskriminationsphase werden dann die Ablehnungsquoten und Latenzzeiten analysiert. Im Anschluß daran wird auf die multinomiale Modellierung des Entscheidungsprozesses und die entsprechenden Parameterschätzungen eingegangen.

Memorierphase

In Abbildung 3 ist für die einzelnen Gruppen getrennt dargestellt, wie häufig während der Memorierphase die Sequenzen eines Sets fehlerhaft reproduziert wurden.

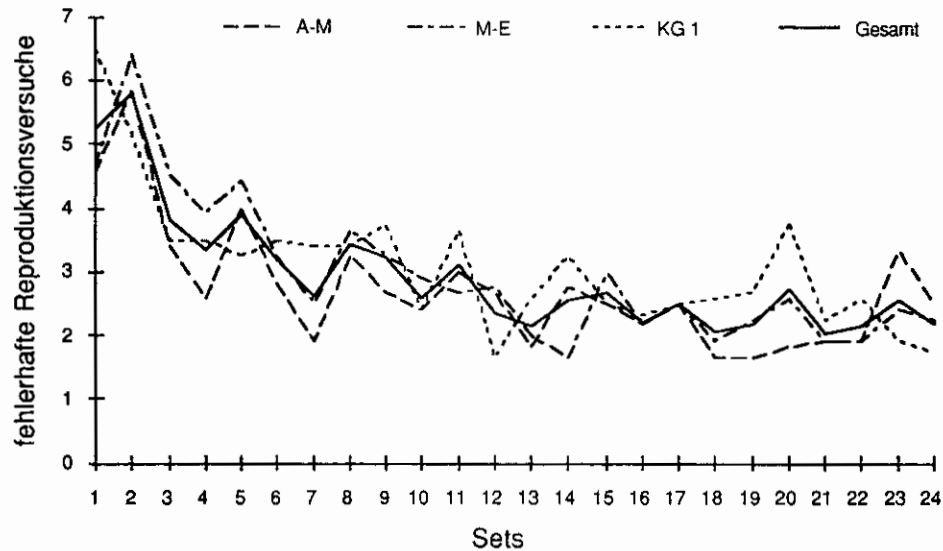


Abb. 3 Anzahl der fehlerhaften Reproduktionsversuche nach experimentellen Bedingungen und Sets getrennt

Es zeigt sich, daß diese Häufigkeiten für die einzelnen Sets sehr unterschiedlich ausfallen. Unabhängig von den experimentellen Bedingungen ergibt sich eine Reduktion der über das Lernkriterium hinaus gehenden Reproduktionsversuche von durchschnittlich 5.25 Wiederholungen für das erste Set zu 2.17 Wiederholungen für das letzte Set. Zwischen den Gruppen zeigen sich insgesamt keinerlei Unterschiede, die Interaktion zwischen Sets und experimentellen Bedingungen ist ebenfalls nicht signifikant. Interessant ist, daß bei jeweils drei aufeinanderfolgenden Sets im Durchschnitt die Anzahl zusätzlicher Präsentationen für das mittlere Set erhöht ist. Hierbei handelte es sich jeweils um Paare von Sequenzen, die zwar wie alle anderen

Sequenzpaare aus jeweils verschiedenen „Worten“ bestanden, sich aber durch eine Eigenschaft von diesen unterschieden, die aus den bei der Konstruktion der Sets zu berücksichtigenden kombinatorischen Beschränkungen resultiert: Während bei jedem der beiden anderen Sets zwei der Tastenpaare eines Bereiches sich lediglich durch die Abfolge der Tasten unterschieden, bestanden bei diesen Sets sämtliche „Worte“ aus anderen Elementen. Inferenzstatistisch ist dieser Effekt signifikant ($F_{(1,759)} = 22.22$).

Diskriminationsphase

In Tabelle 2 sind die Ablehnungsquoten nach experimentellen Bedingungen und Sequenztypen getrennt angegeben. Es zeigen sich für alle Gruppen im Durchschnitt geringere Ablehnungsquoten für solche neue grammatische Sequenzen (NG), die die häufiger gesehene Kombination von erstem und zweitem sowie zweitem und drittem Tastenpaar repräsentieren. Für die Kontrollgruppe 2, die nur die Sequenzen der Diskriminationsphase bearbeitet hat, zeigt sich, daß die Ablehnungsquote für diesen Sequenztyp ebenfalls gegenüber den anderen Sequenztypen reduziert ist.

Tab. 2 Ablehnungsquoten (%) in der Diskriminationsphase nach experimentellen Bedingungen und Sequenztyp getrennt

Experimentelle Bedingung	Sequenztyp			
	NG	EA	EM	EE
A-M	33.3	45.8	51.0	44.3
M-E	38.3	41.1	46.9	45.3
Kontrolle 1	36.5	39.1	53.1	50.5
Kontrolle 2	47.4	53.6	59.9	54.7

Erwartungsgemäß zeigen sich für die Gruppe A-M höhere Ablehnungsquoten für die Sequenztypen EA und EM, allerdings ist der Unterschied zwischen den Ablehnungsquoten für die Sequenztypen EA und EE minimal. Bei insgesamt geringen Ablehnungsquoten zeigt sich für die Gruppe M-E ebenfalls das aufgrund der Ausbildung von Phrasenstrukturen zu erwartende Muster, d.h. Sequenzen vom Typ EM und EE werden in höherem Maße abgelehnt als Sequenzen vom Typ EA, bei denen die hypothetische Phrasenstruktur intakt ist. Die Ablehnungsquoten der Gruppe Kontrolle 1 weisen ein Muster auf, daß entgegen den Erwartungen nicht dem der Gruppe A-M, sondern dem der Gruppe M-E ähnlich ist, d.h. bei dieser Gruppe finden sich höhere Ablehnungsquoten für Sequenzen vom Typ EM und EE gegenüber EA. Dieser Effekt ist im Vergleich mit der Gruppe M-E sogar ausgeprägter. Für die Gruppe Kontrolle 2 finden sich vergleichbar hohe Ablehnungsquoten für die Sequenzen vom Typ EA und EE sowie erhöhte Ablehnungsquoten für Sequenzen vom Typ EM. In dieser Bedingung scheint keine besondere Präferenz für eine der beiden

Phrasenstrukturen zu bestehen, so daß Sequenzen vom Typ EM, die beide Formen verletzen, in besonderem Maße abgelehnt wurden. Inferenzstatistisch sind für die Gruppe A-M die Unterschiede zwischen den Ablehnungsquoten für Sequenzen vom Typ NG und EA ($t_{(22)} = 1.90$) sowie zwischen NG und EM ($t_{(22)} = 2.69$) signifikant, für die Kontrollgruppe 1 sind die Unterschiede zwischen den Sequenzen vom Typ NG und EM ($t_{(22)} = 2.52$) sowie zwischen NG und EE ($t_{(22)} = 2.13$) signifikant. Für die Kontrollgruppe 2 ist lediglich der Unterschied zwischen den Ablehnungsquoten für Sequenzen vom Typ NG und EM ($t_{(22)} = 1.90$) bedeutsam.

Tab. 3 Latenzzeiten (s) in der Diskriminationsphase nach experimentellen Bedingungen und Antwortkategorie getrennt

Experimentelle Bedingung	Antwortkategorie		Gesamt
	Ja	Nein	
A-M	2.49	2.38	2.44
M-E	2.42	2.28	2.35
Kontrolle 1	2.50	2.23	2.36
Kontrolle 2	2.59	2.33	2.46
Gesamt	2.50	2.31	2.40

In Tabelle 3 sind die Latenzzeiten für Ja- und Nein-Antworten für die einzelnen Gruppen getrennt angegeben. Es zeigt sich, daß zwischen den Gruppen kaum Unterschiede hinsichtlich der Latenzzeiten für die beiden Antwortkategorien bestehen. Entgegen den Erwartungen zeigt sich jedoch für alle Gruppen, daß die Latenzzeiten für Ja-Antworten größer sind als für Nein-Antworten. Dieser Effekt ist inferenzstatistisch signifikant ($F_{(1,44)} = 6.66$).

Modellierung der Entscheidungsprozesse mit Hilfe eines multinomialen Modells

Für die Entscheidungen in der Diskriminationsphase wurde ein multinomiales Modell (Riefer & Batchelder, 1988) spezifiziert, in dem als Zustände unterschieden werden, ob (a) eine Satzstruktur, (b) eine der beiden möglichen Phrasenstrukturen oder (c) keine dieser Strukturen ausgebildet wurden. Das nicht-restringierte Modell enthält vier freie Parameter: Parameter *s* bezeichnet die Wahrscheinlichkeit, daß eine Satzstruktur ausgebildet wurde. Parameter *p1* steht für die Wahrscheinlichkeit, daß die Eingabepaare im ersten und zweiten Bereich zu einer Phrase zusammengefaßt wurden, Parameter *p2* für die Wahrscheinlichkeit, daß die Eingabepaare im zweiten und dritten Bereich zu einer Phrase zusammengefaßt wurden. Parameter *r* steht für die Wahrscheinlichkeit, daß mit „Ja“ geantwortet wird, obwohl keine der möglichen Strukturen ausgebildet wurde. Es wird angenommen, daß die Ausbildung

einer Phrasenstruktur darüber entscheidet, in welchem Maße die verschiedenen Sequenztypen abgelehnt werden. In Abbildung 4 ist das Modell als Entscheidungsbaum für die neuen grammatischen Sequenzen dargestellt.

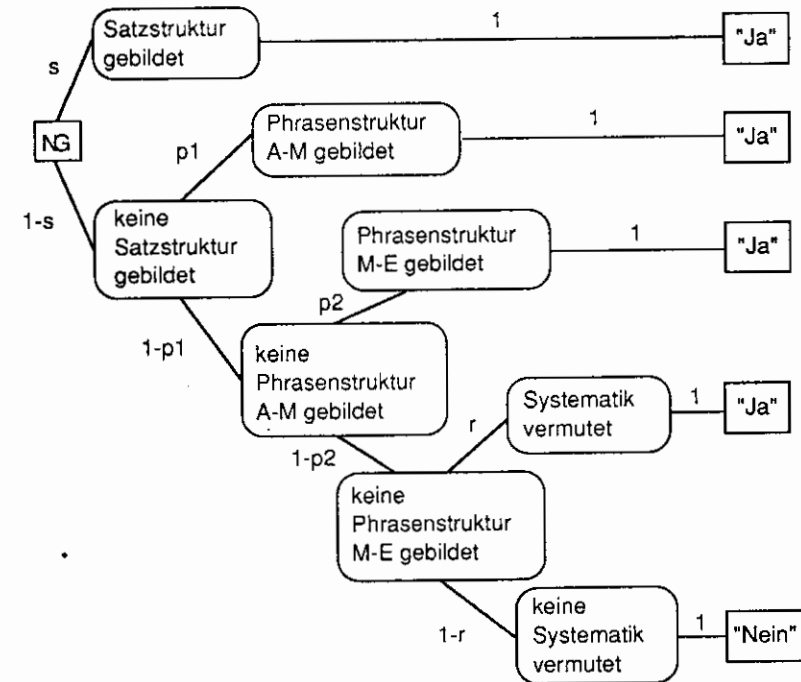


Abb. 4 Entscheidungsbaum für die Zustimmung („Ja“) bzw. Ablehnung („Nein“) neuer grammatischer Sequenzen in der Diskriminationsphase

Es wird erwartet, daß mit „Ja“ geantwortet wird, wenn die Satzstruktur gebildet wurde oder entweder die ersten beiden Tastenpaare oder die letzten beiden Tastenpaare zu einer „Phrase“ zusammengefaßt wurden oder geraten wurde. Für die anderen Sequenztypen ergeben sich jeweils andere Erwartungen: Bei den Sequenzen vom Typ EA wird angenommen, daß die Ausbildung einer Satzstruktur für neue grammatische sowie die Zusammenfassung der ersten beiden Tastenpaare zu einer „Phrase“ dazu führt, daß mit „Nein“ geantwortet wird. Bei Sequenzen vom Typ EM wird angenommen, daß lediglich dann mit „Ja“ geantwortet (geraten) wird, wenn keine der Strukturen ausgebildet wurde. Bei Sequenzen vom Typ EE wird angenommen, daß die Zusammenfassung der ersten beiden Tastenpaare zu einer „Phrase“ dazu führt, daß mit „Ja“ geantwortet wird, während die Ausbildung anderer Phrasen zur Ablehnung der Sequenzen führt. Zur Schätzung der Parameter ergeben sich damit die acht Gleichungen, die in Abbildung 5 gezeigt sind.

NG $p(\text{„Ja“}) = s + (1-s)*p1 + (1-s)*(1-p1)*p2 + (1-s)*(1-p1)*(1-p2)*r;$ (1)
 $p(\text{„Nein“}) = (1-s)*(1-p1)*(1-p2)*(1-r)$ (2)

EA $p(\text{„Ja“}) = (1-s)*(1-p1)*p2 + (1-s)*(1-p1)*(1-p2)*r;$ (3)
 $p(\text{„Nein“}) = s + (1-s)*p1 + (1-s)*(1-p1)*(1-p2)*(1-r)$ (4)

EM $p(\text{„Ja“}) = (1-s)*(1-p1)*(1-p2)*r;$ (5)
 $p(\text{„Nein“}) = s + (1-s)*p1 + (1-s)*(1-p1)*p2 + (1-s)*(1-p1)*(1-p2)*(1-r)$ (6)

EE $p(\text{„Ja“}) = (1-s)*p1 + (1-s)*(1-p1)*(1-p2)*r;$ (7)
 $p(\text{„Nein“}) = s + (1-s)*(1-p1)*p2 + (1-s)*(1-p1)*(1-p2)*(1-r)$ (8)

Abb. 5 Gleichungssystem mit Gleichungen (1) bis (8) für die Parameterschätzung des Entscheidungsbaums aus Abbildung 4

Das in Abbildung 5 dargestellte Modell ist prinzipiell saturiert, so daß die Parameter s , $p1$, $p2$ und r durch Gleichungsumformungen berechnet werden können. Da jedoch angenommen wird, daß die Bildung einer Satzstruktur nicht stattgefunden hat, also Parameter $s = 0$ sein sollte, ergibt sich für das Modell ein Freiheitsgrad. Die Gleichungen und Daten für die verschiedenen experimentellen Bedingungen wurden einer Parameterschätzung durch ein Programm von Hu (1991) unterzogen. Aufgrund der günstigen Eigenschaften der entsprechenden Parameterschätzungen (Hu, 1994; Riefer & Batchelder, 1988) wurde „maximum likelihood“ als Schätzme-

Tab. 4 Parameterschätzungen, Konfidenzintervalle und Modellpassung für das multinomiale Modell nach experimentellen Bedingungen getrennt

Bedingung	Parameter				Modellpassung
	s	p1	p2	r	
A-M	0	0.101 (0.047–0.155)	0.095 (0.036–0.154)	0.581 (0.549–0.613)	0.738
M-E	0	0.023 (–0.031–0.077)	0.066 (0.011–0.121)	0.578 (0.550–0.607)	0.038
Kontrolle 1	0	0.026 (–0.028–0.080)	0.144 (0.085–0.199)	0.563 (0.531–0.594)	0.000
Kontrolle 2	0	0.058 (0.003–0.113)	0.073 (0.015–0.130)	0.456 (0.424–0.488)	0.024

thode gewählt. Tabelle 4 zeigt für die verschiedenen Bedingungen die Parameterwerte sowie die nach Read und Cressie (1988) ermittelten Divergenzmaße des Modells, die χ^2 -verteilt sind.

Insgesamt zeigt sich für alle experimentellen Bedingungen eine hohe Entsprechung zwischen Modell und Daten ($\chi^2_{(1)}\text{-krit}_{(\alpha=.05)} = 3.84$). Die Annahme, daß keine Satzstruktur ausgebildet wurde ($s = 0$), erscheint damit bestätigt. Die Parameterrelationen bestätigen darüber hinaus die Interpretationen der Ablehnungsquoten. Für die Bedingung A-M zeigt sich, daß die Wahrscheinlichkeit ($p1$), die Eingabepaare des ersten und zweiten Bereichs zu einer Phrase zusammenzufassen, signifikant von 0 verschieden ist, allerdings gegenüber der Wahrscheinlichkeit, die Eingabepaare des zweiten und dritten Bereichs zusammenzufassen, nur leicht erhöht ist. Im Unterschied zur Inferenzstatistik auf der Basis der direkten Vergleiche der Ablehnungsquoten zeigt sich für die Gruppe M-E, daß die Wahrscheinlichkeit ($p2$), die Eingabepaare des zweiten und dritten Bereichs zusammenzufassen, bedeutsam von 0 verschieden ist. Für die Kontrollgruppe 1 bestätigt sich, daß diese Wahrscheinlichkeit ($p2$) noch höher ausgeprägt ist. Für die Kontrollgruppe 2 zeigt sich für beide Parameter bezüglich der Phrasenstruktur ($p1$ und $p2$) vergleichbar hohe von 0 verschiedene Ausprägungen. Für die Gruppen mit Memorierphase (alle außer Kontrolle 2) zeigt der Parameter r an, daß eine signifikant erhöhte Tendenz bestand, mit „Ja“ zu antworten, wenn die Sequenzen keiner hypothetischen Struktur widersprachen. Dieser Wert ist für die drei Gruppen vergleichbar groß. Für die Kontrollgruppe 2 hingegen zeigt sich eine erhöhte Tendenz, bei Unsicherheit über die Systematik der Sequenz mit „Nein“ zu antworten. Dieser Unterschied reflektiert die durch die Memorierphase bedingte höhere Vertrautheit der einzelnen Eingaben und Eingabepaare.

Diskussion

In der Memorierphase hat sich gezeigt, daß die Anzahl von Reproduktionsversuchen bis zum Lernkriterium mit steigender Anzahl der Lernsets abnahm. Entsprechend den Befunden von Servan-Schreiber und Anderson (1990) unterschieden sich die experimentellen Gruppen nicht in der durchschnittlichen Anzahl von falschen Reproduktionen. Eine Besonderheit bestand darin, daß vom zweiten Set an für jedes dritte darauf folgende Set eine Erhöhung der Anzahl von Reproduktionsversuchen zu beobachten war. Dieser Effekt korrespondiert damit, daß gegenüber den anderen Sets sämtliche Tastenpaare aus anderen Klassen der jeweiligen Position stammten (s.o.), also diese Sets weniger redundant waren.

In der Diskriminationsphase hat der direkte Vergleich der Ablehnungsquoten für die verschiedenen Sequenztypen ähnlich den Befunden von Servan-Schreiber und Anderson (1990) angedeutet, daß abhängig von der experimentellen Manipulation in der Memorierphase Vpn in unterschiedlichem Maße gegenüber Verletzungen der „Phrasenstruktur“ der Sequenzen sensitiv waren. Allerdings waren die deskriptiven Unterschiede für die Gruppe M-E inferenzstatistisch nicht absicherbar und die Effekte in der Kontrollbedingung wiesen auf eine Präferenz für die Zusammenfassung der letzten beiden „Worte“ zu einer „Phrase“ hin. Interessanterweise waren die Latenzzeiten für „Ja“-Antworten signifikant höher als für „Nein“-Antworten. Die An-

nahme einer höheren Flüssigkeit der Wahrnehmung als Erklärung impliziter Effekte ist mit diesem Unterschied nicht ohne weiteres vereinbar. Dieser Befund entspricht auch Ergebnissen von Buchner (1994), der zwar für Rekognitionsurteile, nicht jedoch für grammatische Urteile einen positiven Zusammenhang mit Identifikationsleistungen gefunden hat. Allerdings ist darauf hinzuweisen, daß in dem vorliegenden Experiment die Antwortklasse mit der Richtung der Mausbewegung konfundiert war („Ja“-Taste links, „Nein“-Taste rechts).

Multinomiale Modellierung des Entscheidungsprozesses

Die Rekonstruktion der Entscheidungen im Rahmen eines multinomialen Modells hat verdeutlicht, daß die Parameterschätzungen für die Wahrscheinlichkeiten der Ausbildung von Phrasenstrukturen von Null verschieden sind und für die verschiedenen Memorierbedingungen zwar nicht erwartungsgemäß, jedoch unterschiedlich ausfallen. Als Erklärung für diese Befundlage deutet sich an, daß bei Bediensequenzen im Unterschied zu Buchstabenfolgen eher eine Tendenz zur Zusammenfassung der letzten beiden Eingabenpaare zu bestehen scheint. Die Interpretation der Parameter in diesem Sinne wird durch die Reanalyse der Daten des Experimentes von Servan-Schreiber und Anderson (1990) nach diesem Modell (Gleichungen (1) bis (8) in Abbildung 5) gestützt. Die Ergebnisse dieser Reanalyse (Tabelle 5) bestätigen im wesentlichen die Interpretationen der Autoren. Als Besonderheit zeigt sich aber auch, daß die Modellpassung des restringierten Modells ($s = 0$) für die Bedingung A-M über dem kritischen χ^2 -Wert (3.84) liegt. Für diese Bedingung muß angenommen werden, daß über die Phrasenstrukturen hinaus auch korrekte Satzstrukturen für die neuen grammatischen Sätze gebildet wurden. Damit wird auch die für diese Bedingung fehlende Erhöhung der Tendenz zu „Ja“-Antworten erklärbar, da mit den Satzstrukturen gegenüber den Gruppen in den anderen Bedingungen ein valides Konzept für neue grammatische Sätze zur Verfügung stand.

Generell erweist sich damit die multinomiale Modellierung als äußerst hilfreiches Verfahren, bei komplexen Datenmustern Annahmen über theoretisch zugrundeliegende Zusammenhänge zu prüfen. Neben der Möglichkeit, inferenzstatistische Tests

Tab. 5 Parameterschätzungen, Konfidenzintervalle und Modellpassung für die Reanalyse von Experiment 2 von Servan-Schreiber und Anderson (1990) nach dem multinomialen Modell

Bedingung s	Parameter				Modellpassung
	p1	p2	r		
A-M	0.132 (0.025–0.240)	0.180 (0.102–0.259)	0.000 (–0.109–0.111)	0.469 (0.419–0.518)	0.201
A-M	0	0.233 (0.180–0.286)	0.083 (0.014–0.151)	0.431 (0.393–0.470)	5.576
M-E	0	0.012 (–0.042–0.065)	0.157 (0.011–0.121)	0.559 (0.527–0.590)	0.156
Kontrolle	0	0.108 (0.053–0.162)	0.017 (–0.043–0.078)	0.570 (0.540–0.600)	0.004

durchzuführen, zeigt die Reanalyse der Daten von Servan-Schreiber und Anderson (1990), daß dadurch auch zusätzlich die Kompositionsbildung bestätigende Information resultieren kann.

Um die insbesondere für die Kontrollgruppen der beiden Experimente feststellbaren Unterschiede in der Präferenz zur Zusammenfassung der ersten oder letzten Elemente der Buchstaben- bzw. Tastenfolgen inferenzstatistisch zu überprüfen, wurden die Ablehnungsquoten für die Kontrollgruppen der beiden Experimente in einem gemeinsamen multinomialen Modell (vgl. Hu, 1994) analysiert. Hierbei zeigte sich, daß sich die Parameterausprägungen signifikant unterscheiden. Die Differenz der Modellpassung bei sechs freien Parametern und der bei drei freien Parametern ($p1$, $p2$ und r wurden für beide Gruppen gleichgesetzt) ergibt einen Divergenzwert > 16 , der bei drei Freiheitsgraden über dem kritischen χ^2 -Wert (7.81) liegt.

Dieser Effekt kann nicht einfach als Rezens-Effekt in der Diskriminationsphase angesehen werden, da bei Kontrollgruppe 2 erwartungsgemäß von Null verschiedene, jedoch vergleichbar große Parameter für die Phrasenbildung geschätzt wurden. Im folgenden wird diskutiert, wie dieser Effekt im Rahmen der Theorie des „competitive chunking“ dadurch erklärt werden kann, daß die zeitliche Sukzession der Darbietung der Bediensequenzen in der Memorierphase berücksichtigt wird.

Einfluß zeitlicher Sukzession auf das Chunking

Nach der Theorie des „competitive chunking“ ergibt sich die Wahrscheinlichkeit, daß ein Chunk abgerufen bzw. gebildet wird, gemäß Formel (1) als Funktion der Stärke der Konstituenten und dem Competition-Parameter c .

$$p(\text{chunk}) = \frac{1 - e^{-c \cdot \text{support}}}{1 + e^{-c \cdot \text{support}}}, \text{ wobei } \text{support} = \sum_{i=1}^n \frac{S_i}{n}, \text{ und } S = \sum_{i=1}^n t_i^{-d} \quad (1)$$

Für den Einfluß zeitlicher Sukzession ist dabei relevant, daß die Stärkekfunktion mit zunehmendem Zeitintervall generell negativ beschleunigt abfällt. Das führt dazu, daß bei kleinen Zeitintervallen gleiche absolute Zeitabstände einen größeren Effekt auf die Stärke haben als bei großen. Dieser Effekt wirkt sich auf die Wahrscheinlichkeitsfunktion umso stärker aus, je geringer das bis dahin akkumulierte Stärkeniveau der Konstituenten ist. Abbildung 6 zeigt den Zusammenhang zwischen Zeitintervall und Wahrscheinlichkeit des Abrufs bzw. der Bildung von Chunks für unterschiedlich starke Konstituenten, wenn Decay-Parameter d und Competition-Parameter c auf .5 gesetzt wurden.

Der Vergleich der beiden Kurvenverläufe verdeutlicht, daß zeitliche Sukzession für den Abruf bzw. die Bildung von Chunks umso relevanter ist, je geringer die Stärke der Konstituenten dieser Chunks ist. Im Unterschied zum Experiment von Servan-Schreiber und Anderson (1990) wurden in diesem Experiment wenig vertraute Sequenzelemente eingesetzt, so daß die zeitliche Sukzession während der Darbietung der Sequenzelemente in der Memorier- wie in der Diskriminationsphase umso mehr als potentielle Erklärung der Effekte angesehen werden kann.

In der Memorierphase kann dieser Zusammenhang dazu geführt haben, daß generell mit höherer Wahrscheinlichkeit die letzten Tastendrucke zu Paaren sowie die

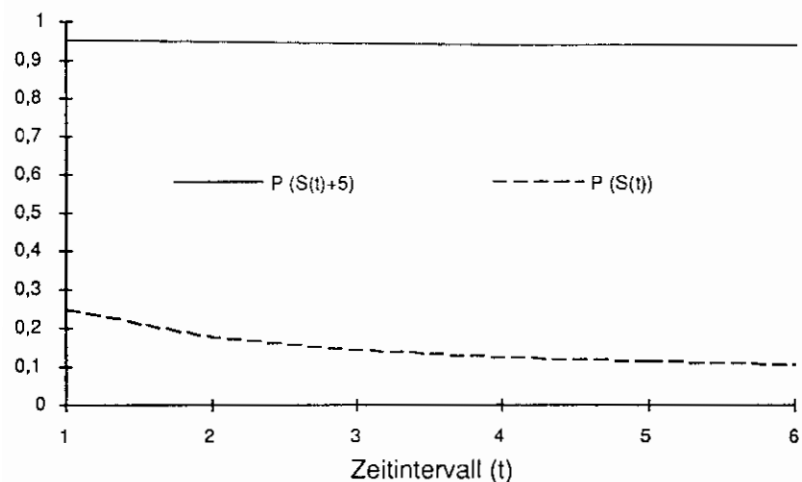


Abb. 6 Der Zusammenhang von zeitlicher Sukzession der Darbietung (t) von Chunkkonstituenten, deren vorher erreichter Stärke ($S(t)$ vs. $S(t) + 5$) und der Wahrscheinlichkeit, den Chunk abzurufen bzw. zu bilden ($P(S)$) nach der Theorie des „competitive chunking“. Der Einfluß zeitlicher Sukzession auf die Wahrscheinlichkeit des Abrufs bzw. der Bildung von Chunks ist umso größer, je weniger Stärke die Konstituenten akkumuliert haben. Weitere Erläuterungen finden sich im Text

letzten beiden Paare zu einer „Phrase“ zusammengefaßt wurden. Damit konnten diese Einheiten auch in höherem Maße Stärke akkumulieren. Allerdings muß hierbei zusätzlich angenommen werden, daß die Kompositionsbildung generell erst nach der Vorführung oder erfolgreichen Reproduktion der Sequenzelemente stattfindet. In der Diskriminationsphase würde dieser Effekt dadurch zum Tragen kommen, daß der Abruf der Phrase, die die letzten beiden Eingabepaare als Konstituenten hat, wahrscheinlicher wird. Inwiefern dieser Zusammenhang und damit die Erklärung der für die Paradigmen unterschiedlichen Effekte in nachfolgenden Experimenten bzw. Simulationen bestätigt werden kann, bleibt abzuwarten.

Der Ansatz finiter Automaten zur integrativen Rekonstruktion von Forschungsparadigmen

Im Interesse eines Brückenschlags zwischen Grundlagen- und Anwendungsfor-schung wurde exemplarisch die Übertragbarkeit von theoretischen Konzeptionen zur Ausbildung komplexer Wissensstrukturen auf den Bereich diskreter dynamischer Systeme geprüft. Dazu wurde auf der formalen Basis finiter Automaten ein experimentelles Paradigma gewählt, daß als abstrakte Variante eines Automaten angesehen werden kann und gleichzeitig maximale Ähnlichkeit zu dem empirischen Paradigma aufweist, innerhalb dessen die theoretischen Vorstellungen zur Ausbildung

von Chunks höherer Ordnung belegt werden konnten. Generell hat sich dabei gezeigt, daß die Annahmen der Theorie des „competitive chunking“ auf den Bereich von Bediensequenzen übertragbar sind. Wenngleich das eingesetzte Paradigma in einer Reihe von Eigenschaften von tatsächlich existierenden Mensch-Maschine-Systemen abweicht, bestätigt aus unserer Sicht die Interpretierbarkeit der Befunde, daß der Ansatz finiter Automaten geeignet ist, Konzepte und Theorien grundlagenorientierter Forschung zu integrieren und damit die Basis schaffen kann, anwendungsorientierte Fragen und Probleme der Handlungssteuerung zu beantworten. Damit wollen wir unseren Beitrag auch als Versuch verstanden wissen, im Spannungsfeld zwischen der häufig geforderten ökologischen Validität und den Anforderungen einer objektivierbaren Methodologie praktikable Lösungen aufzuzeigen.

Literatur

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Buchner, A. (1994). Indirect effects of synthetic grammar learning in an identification task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 550-566.
- Buchner, A., Faul, F. & Erdfelder, E. (1992). G-Power: A priori, post-hoc and compromise power analyses for the Macintosh [Computer-Programm]. Bonn: Psychologisches Institut der Universität Bonn.
- Buchner, A., & Funke, J. (1993). Finite state automata: Dynamic task environments in problem solving research. *Quarterly Journal of Experimental Psychology*, 46A, 83-118.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Revised edition. Hillsdale, NJ: Lawrence Erlbaum.
- Funke, J., & Buchner, A. (1992). Finite Automaten als Instrumente für die Analyse von wissensgeleiteten Problemlöseprozessen: Vorstellung eines neuen Untersuchungsparadigmas. *Sprache & Kognition*, 11, 27-37.
- Funke, J., & Gerdes, H. (1993). *Manuale für Videorekorder: Auswahl von Textinhalten unter Verwendung der Theorie endlicher Automaten*. *Zeitschrift für Arbeitswissenschaft*, 47, 44-49.
- Hu, X. (1991). *General program for processing tree models* [Computer-Programm]. Irvine, CA: University of California at Irvine.
- Hu, X. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika*, 59, 21-48.
- Lewis, C. (1987). Composition of productions. In D. Klahr, P. Langley, & R. Neches (Eds.), *Production system models of learning and development* (pp. 329-358). Cambridge, MA: MIT Press.
- MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy in skilled behavior. *Psychological Review*, 89, 483-506.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Miller, G. A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology*, 56, 485-491.
- Müller, B., Funke, J. & Buchner, A. (1994). Diskrete dynamische Systeme: Der Einfluß perzeptueller Strukturierung auf Komposition und Transfer von Wissen über Bediensequenzen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 41, 443-472.

- Neves, D.M., & Anderson, J.R. (1981). Knowledge compilation: mechanisms for the automatization of cognitive skills. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition*. (pp. 57-84). Hillsdale, NJ: Erlbaum.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Read, T. R. C. & Cressie, N. A. C. (1988). *Goodness-of-fit statistics for discrete multivariate data*. New York: Springer-Verlag.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 5, 855-863.
- Reber, A.S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 88-94.
- Riefer, D. M. & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95, 318-339.
- Servan-Schreiber, E. & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 592-608.
- Vaterrödt-Plünnecke, B. (1994). Multinomiale Modellierung impliziter Gedächtnisprozesse: Ein alternativer Ansatz. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 41, 295-314.

Autorenhinweis

Die berichteten Befunde stammen aus dem Projekt „Analyse von Lernprozessen beim Umgang mit finiten Automaten (ALFA)“, gefördert von der Deutschen Forschungsgemeinschaft (Az. Fu 173/3).

Für die Unterstützung bei der Programmierung und Datenerhebung danken wir René Jerusalem sowie Melanie Steffens. Für hilfreiche Kommentare und Anregungen zu früheren Fassungen des Manuskripts danken wir Dr. Bianca Vaterrödt-Plünnecke, Dr. Michael Ziebler sowie zwei anonymen Gutachtern.

Eingang des Manuskripts 21.3.1994

Eingang des überarbeiteten Manuskripts 27.7.1994

Anschrift der Verfasser: Dr. Burkhard Müller, PD Dr. Joachim Funke: Psychologisches Institut der Universität Bonn, Römerstr. 164, D-53117 Bonn, Bundesrepublik Deutschland.