

Manuela Friedrich

Der Erwerb hierarchisch ordenbarer Kategorien in einem neuronalen Netz der Adaptiven Resonanztheorie

Zusammenfassung

Gegenstand dieser Arbeit ist das Identifizieren von Bedingungen für die Bildung unterschiedlich allgemeiner Begriffe in einem eigenständig kategorisierenden neuronalen Netz. Innerhalb eines Simulationskomplexes wurde der Einfluß verschiedener Parameter des neuronalen Netzes ART1 auf den Abstraktionsgrad herausgebildeter Kategorien untersucht. Dafür wurde ein Datensatz konstruiert, der Strukturierungsmöglichkeiten auf mehreren Abstraktionsniveaus zuläßt. Über diesem Datensatz können Oberbegriffe, Basisbegriffe, Unterbegriffe und Exemplarbegriffe gebildet werden.

Die Simulationen zeigen, daß neben dem Parameter Vigilanz auch der Schwellwert in der Merkmalsschicht und die Resonanzdauer einen Einfluß auf den Abstraktionsgrad der gebildeten Begriffe haben. Weiterhin ändern die Begriffe häufig über die Lernzyklen hinweg ihren Abstraktionsgrad. Dabei konnte eine Entwicklung von der Bildung sehr allgemeiner Begriffe hin zur Bildung spezifischer Begriffe beobachtet werden. Es stellte sich insbesondere heraus, daß in bestimmten Lernzuständen eine Bildung spezifischer Begriffe selbst mit maximaler Vigilanz nicht möglich ist. Damit lassen sich experimentelle Befunde erklären, die auf eine Abhängigkeit der kindlichen Fähigkeit zur Teilkategorienbildung vom speziellen Lernzustand des allgemeinen Begriffs weisen.

Es hat sich weiterhin herausgestellt, daß die Variation des Parameters Vigilanz eine entscheidende Voraussetzung für das Nebeneinanderbestehen unterschiedlich abstrakter Begriffe im Netz ist. Unter konstanten Parameterkombinationen bestehen verschieden abstrakte Begriffe ausschließlich in einer zeitlich begrenzten Differenzierungsphase nebeneinander. Längerfristig entstehen im Netz bei konstanten Parametern nur Kategorien eines einzigen Abstraktionsniveaus. Wird jedoch die Vigilanz über die Lernzyklen hinweg variiert, so können die zeitlich nacheinander vom Netz erworbenen Begriffe mehrerer Abstraktionsgrade auch nebeneinander und damit gleichzeitig bestehen bleiben. Eine flexible Regulation der Vigilanz ist also für den Aufbau eines verschiedene Abstraktionsgrade umfassenden Begriffssystems im ART1 notwendig.

Problemstellung

Aus empirischen Untersuchungen mit Kindern ist bekannt, daß ihre Begriffsentwicklung von der Bildung allgemeiner Begriffe hin zur Bildung spezifischerer verläuft. Zuerst werden die allgemeinsten Begriffe erworben, deren Merkmale direkt sensorisch wahrnehmbar sind, erst später werden zusätzlich zu diesen spezifische sensorische Begriffe gebildet oder Begriffe, deren Merkmale nicht direkt wahrnehmbar sind (z.B. Rosch und Mitarbeiter, 1976; Mandler und Mitarbeiter, 1991). Es hat sich weiterhin gezeigt, daß spezifische Begriffe nicht zu jedem Zeitpunkt von bereits erworbenen allgemeinen Begriffen differenziert werden können. Neben allgemeinen kognitiven Voraussetzungen scheint der spezifische Lernzustand eines bestehenden allgemeinen Begriffes einen wesentlichen Einfluß auf die Fähigkeit zur weiteren begrifflichen Differenzierung seiner Exemplare zu besitzen. Diese Vermutung wird von Untersuchungen nahegelegt, in denen Kinder in bestimmten Bereichen (z.B. bei Autos) bereits spezifische Begriffe erwerben konnten, in anderen Bereichen (z.B. bei Blumen) dazu jedoch nicht in der Lage waren (Goede, 1989; Goede und Friedrich, 1995).

Anhand von Simulationen mit einem neuronalen Netz der Adaptiven Resonanztheorie (z.B. Grossberg, 1987 a, b, 1982, 1980) soll gezeigt werden, wie Lernzustände die Fähigkeit des Netzes zur feineren Differenzierung von Begriffen beeinflussen. Das nichtüberwacht lernende Netz ART1 (Carpenter & Grossberg, 1987) strukturiert sich Datensätze selbst, es werden keine Strukturierungen explizit vorgegeben. Existieren mehrere Strukturierungsmöglichkeiten einer Datenmenge, so kann von vornherein nicht sicher bestimmt werden, welche vom Netz gewählt wird. Insbesondere bei Vorliegen hierarchischer Beziehungen zwischen möglichen Strukturierungen ist der Abstraktionsgrad der vom Netz erworbenen Strukturen unklar. Aus diesem Grund wird speziell untersucht, welchen Allgemeinheitsgrad die eigenständig vom Netz erworbenen Begriffe unter verschiedenen Parameterbedingungen herausbilden und welche Lernzustände damit verbunden sind. Weiterhin wird die Entwicklung der Begriffe beobachtet, insbesondere, wie sich ihr Allgemeinheitsgrad mit der weiteren Darbietung von Exemplaren verändert. Schließlich interessiert in diesem Zusammenhang, unter welchen Bedingungen das Netz mehrere Begriffe mit unterschiedlichem Allgemeinheitsgrad über der gleichen Datenmenge erwerben kann.

Die Repräsentation hierarchischer Kategorien im ART1

Das ART1 besitzt neben der Inputschicht eine Merkmalschicht und eine Kategorienschicht. Das bedeutet, daß innerhalb des Netzes alle Kategorien „nebeneinander“ gespeichert sind. Dies gilt insbesondere auch für Kategorien, die sich entsprechend ihrer Merkmalsstruktur oder entsprechend der ihnen zugeordneten Exemplare in einer hierarchischen Ordnung befinden. Diese Annahme einer nichthierarchischen Repräsentation hierarchisch ordenbarer Kategorien hat sich bereits als sinnvoll bei der Erklärung empirischer Phänomene erwiesen. Die experimentell erzeugbaren zeitweiligen Untergeneralisierungen bei Kindern (Goede und Friedrich, 1995) und Erwachsenen (Goede und Friedrich, 1997) konnten auf dieser Grundlage ebenso simuliert werden wie die Abhängigkeit der Reaktionszeit von der Typikalität bei der Verifikation kategorialer Zugehörigkeit (Friedrich, 1994; Friedrich & Goede, 1997).

Die Struktur des Datensatzes und des Kategoriensystems

Um die Herausbildung und Entwicklung hierarchisch ordenbarer Kategorien im Netz genau zu verfolgen, wurde ein spezieller Datensatz konstruiert, dessen Exemplare jeweils zu verschiedenen allgemeinen Kategorien zusammengefaßt werden können. Ein einzelnes Exemplar kann gleichzeitig mehreren Kategorien verschiedener Abstraktionsgrade angehören. Der Datensatz (Tab.1) wurde aus 16 Items mit jeweils 28 Merkmalen so gebildet, daß je 8 Items einer allgemeinen begrifflichen Kategorie (Oberbegriff) angehören. Von diesen 8 Items formen jeweils 4 Items eine Kategorie mittleren Abstraktionsgrades (Basisbegriff), so daß für jeden Oberbegriff 2 solcher Basisbegriffe und insgesamt 4 Basisbegriffe möglich sind. Jeder Basisbegriff umfaßt wiederum 2 der insgesamt 8 spezifischen Kategorien (Unterbegriffe), die von jeweils 2 Items gebildet werden. Damit können in dem Datensatz vier hierarchisch ordenbare Kategorientypen verschiedener Abstraktionsgrade identifiziert werden: Oberbegriffe, Basisbegriffe, Unterbegriffe und Exemplarbegriffe. Jedes Item gehört jedem dieser Kategorientypen genau einmal an, es gehört also zu einem Oberbegriff, zu einem Basisbegriff, zu einem Unterbegriff und formt einen eigenen Exemplarbegriff.

Jeweils 2 Merkmale sind als definierende Merkmale für die Oberbegriffszuordnung relevant. Als Basisbegriffsmerkmale werden diese und jeweils zwei weitere Merkmale genutzt, so daß über die Basisbegriffszuordnung insgesamt 4 Merkmale entscheiden, wobei sich je zwei Basisbegriffe höchstens in ihren beiden Oberbegriffsmerkmalen überlagern. Je zwei Merkmale definieren in Verbindung mit den Ober- und Basisbegriffsmerkmalen einen Unterbegriff. Die Unterbegriffe überlagern sich z.T. in Oberbegriffs- und Basisbegriffsmerkmalen sowie in den

Unterbegriffsmerkmalen mit den Items des jeweils anderen Oberbegriffs. Jeweils zwei von 8 weiteren Merkmalen geben dem Item zusätzlich eine Exemplarspezifika. Diese überlagern sich innerhalb und zwischen den Unterbegriffen. Die über das Skalarprodukt definierte Ähnlichkeit zwischen den einzelnen Exemplaren ist in Tab.2 dargestellt. Daraus ergibt sich als mittlere Ähnlichkeit der Exemplare innerhalb der Unterbegriffe 7.5, zwischen den Unterbegriffen eines Basisbegriffes 4.25, innerhalb der Basisbegriffe 5.875, zwischen den Basisbegriffen eines Oberbegriffes 2.125, innerhalb der Oberbegriffe 4 und zwischen den Oberbegriffen 1.

Tab.1: Datensatz mit mehreren Strukturierungsmöglichkeiten

Item Nr.	Exemplar- Merkmale																														
	Unterbegriffs- Merkmale																														
	Basisbegriffs- Merkmale																														
	Oberbegriffs- Merkmale																														
1	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
2	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
3	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0
4	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0
5	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0
6	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0
7	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1
8	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	1
9	0	0	1	1	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
10	0	0	1	1	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
11	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1
12	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1
13	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0
14	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0
15	0	0	1	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0
16	0	0	1	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0

Tab.2: Ähnlichkeitsmatrix für die Items des Datensatzes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	8	7	4	4	2	2	2	3	1	1	0	1	2	1	1	1
2	7	8	5	4	2	2	2	2	1	1	0	0	1	2	2	1
3	4	5	8	7	2	2	2	2	1	1	1	1	0	1	2	1
4	4	4	7	8	3	2	2	2	2	1	1	1	0	0	1	2
5	2	2	2	3	8	7	4	4	2	1	1	1	1	1	0	1
6	2	2	2	2	7	8	5	4	1	2	2	1	1	1	0	0
7	2	2	2	2	4	5	8	7	0	1	2	1	1	1	1	1
8	3	2	2	2	4	4	7	8	0	0	1	2	2	1	1	1
9	1	1	1	2	2	1	0	0	8	7	4	4	2	2	2	3
10	1	1	1	1	1	2	1	0	7	8	5	4	2	2	2	2
11	0	0	1	1	1	2	2	1	4	5	8	7	2	2	2	2
12	1	0	1	1	1	1	1	2	4	4	7	8	3	2	2	2
13	2	1	0	0	1	1	1	2	2	2	2	3	8	7	4	4
14	1	2	1	0	1	1	1	1	2	2	2	2	7	8	5	4
15	1	2	2	1	0	0	1	1	2	2	2	2	4	5	8	7
16	1	1	1	2	1	0	1	1	3	2	2	2	4	4	7	8

Mit diesem Datensatz wurden mehrere Simulationsreihen durchgeführt. Um das Zuordnungsverhalten des Netzes zu klassifizieren, wurden die gebildeten Kategorien in verschiedene Kategorientypen eingeteilt, die ihrem jeweiligen Abstraktionsgrad entsprechen. Oberbegriffe wurden als Kategorientyp 1 definiert, Basisbegriffe als Kategorientyp 3, Unterbegriffe als Kategorientyp 5 und Exemplarbegriffe als Kategorientyp 7. Zusätzlich wurden die Kategorientypen 2, 4 und 6 als Übergangsformen von Oberbegriffen zu Basisbegriffen, von Basisbegriffen zu Unterbegriffen und von Unterbegriffen zu Exemplarbegriffen festgelegt. Die Zugehörigkeit der Kategorien zu dem jeweiligen Kategorientyp wurde parallel über das Zuordnungsverhalten und die Anzahl der gebildeten Kategorien bestimmt, zusätzlich wurde die interne Merkmalsstruktur der Kategorien kontrolliert.

Oberbegriffe werden demnach dadurch charakterisiert, daß ihnen die jeweils 8 Items mit gleichen Oberbegriffsmerkmalen zugeordnet werden. Das Netz besitzt den *Kategorientyp 1*, wenn es genau zwei solcher Kategorien und keine weitere erworben hat. Je nach Schärfegrad der Kategorien können von ihnen entweder nur die beiden definierenden Oberbegriffsmerkmale

oder zusätzlich noch weitere Merkmale entsprechend ihrer Häufigkeit kodiert werden. Basisbegriffe wurden ausgebildet, wenn ihnen die jeweils 4 Items mit gleichen Basisbegriffsmerkmalen zugeordnet werden. Der *Kategorientyp 3* bestimmt sich bei dem verwendeten Datensatz aus genau vier solcher Kategorien. Diese können intern wieder schärfer oder unschärfer ausgebildet sein, müssen jedoch in jedem Fall die vier Basisbegriffsmerkmale besitzen. Unterbegriffe wurden differenziert, wenn ihnen jeweils die beiden, einen Unterbegriff formenden Items und keine weiteren zugeordnet werden. Damit werden beim *Kategorientyp 5* genau 8 Kategorien erworben. Die Unterbegriffe enthalten alle sechs Unterbegriffsmerkmale. Den Exemplarbegriffen wird jeweils nur ein Item zugeordnet. Beim *Kategorientyp 7* werden genau 16 Exemplarbegriffe gebildet. Sie kodieren den vollständigen Merkmalsatz eines einzigen Exemplares. Der *Kategorientyp 2* ist dadurch charakterisiert, daß genau drei Kategorien in Erscheinung treten. Davon stellt eine einen Oberbegriff dar, eine weitere war bisher Oberbegriff, jedoch werden einzelne zu ihr gehörende Items bereits zu einer neu etablierten Kategorie zusammengefaßt. Die Kategorienstrukturen können je nach Entwicklungsstand variieren. Bei dem *Kategorientyp 4* verhält es sich ähnlich. Hier liegen fünf bis sieben Kategorien vor. Dabei wurden von den bisher ausgebildeten Basisbegriffen einige Items abgespalten, so daß bereits unterbegriffsähnliche Kategorien entstanden sind. Die Basisbegriffe selbst besitzen noch keinen Unterbegriffscharakter, entwickeln sich jedoch dahingehend. Die Kategorienstruktur ist hier ebenfalls uneindeutig. Der *Kategorientyp 6* ist durch das Vorhandensein von 11 bis 15 Kategorien charakterisiert, von denen einige Unterbegriffe darstellen, andere bereits Exemplarbegriffe sind.

Simulationsmethodik

In den Simulationen wurden die Vigilanz, der Schwellwert der Merkmalschicht und die Dauer der Resonanz als Parameter des Netzes variiert.¹ Die Vigilanz legt die Genauigkeit fest, mit der das Netz die Items kategorisiert, und entscheidet damit wesentlich über die Differenzierungsfähigkeit des Netzes (Carpenter & Grossberg, 1987). Der Schwellwert der Merkmalschicht bestimmt in der sogenannten Zwei-Drittel-Regel (Carpenter & Grossberg, 1987), ab welcher (top-down-) Gewichtstärke ein Merkmal von dem entsprechenden Begriff erwartet wird. Der Schwellwert beeinflusst über diesen Mechanismus den Schärfegrad der gebildeten Kategorien, d.h. die Anzahl und das jeweilige Gewicht der von der Kategorie kodierten

¹ Eine detaillierte Beschreibung der Dynamik des ART1 sowie eine Darstellung der Wirkung von Parametern finden sich in Friedrich (1994) oder in Friedrich und Goede (1997).

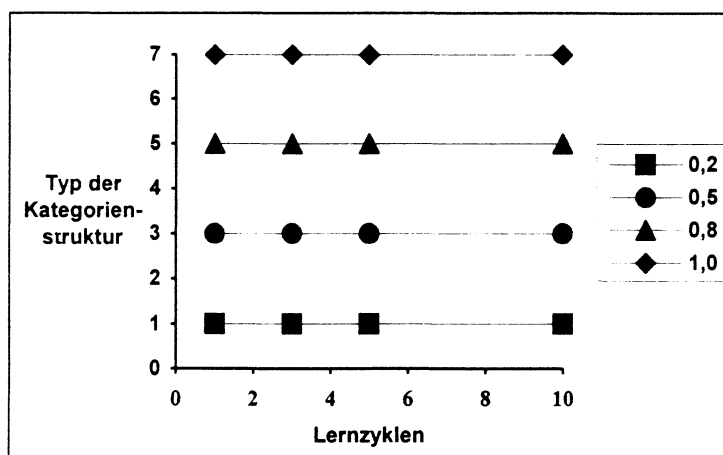
Merkmale (Friedrich, 1994). Die Dauer der Resonanz bestimmt, wie lange jede Kategorisierung eines Items gelernt wird, d.h. wieviel Information bei einer einzelnen Itemzuordnung längerfristig gespeichert wird. Die Dauer der Resonanz beeinflusst ebenfalls den Schärfegrad der entstehenden Kategorien und damit ihren Lernzustand (Friedrich, 1994; Friedrich und Goede, 1997).

Für die Exemplare wurde einmalig eine Zufallsreihenfolge erzeugt, die aus Gründen der Vergleichbarkeit konstant gehalten wurde. Über mehrere Lernzyklen (Durchläufe des Datensatzes) wurde beobachtet, wieviele Kategorien welchen Abstraktionsgrades vom Netz gebildet wurden. Die Simulationen teilen sich in zwei wesentliche Komplexe. Im ersten wurden alle gewählten Parameter über die Lernzyklen hinweg konstant gehalten, im zweiten wurde die Vigilanz zwischen den Lernzyklen variiert.

Simulationen mit konstanten Parameterkombinationen für alle Lernzyklen

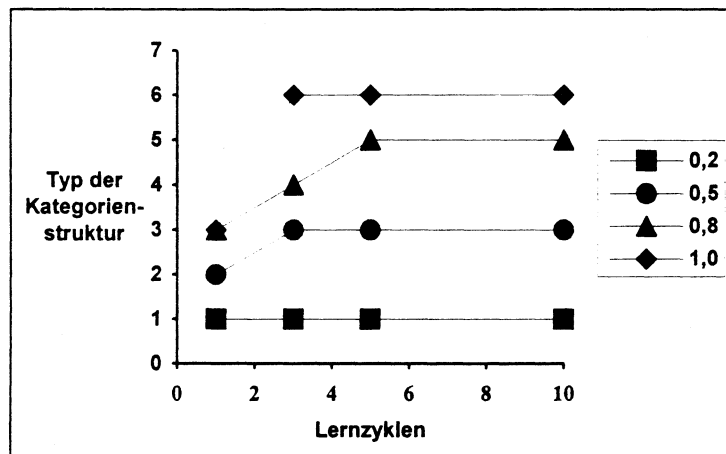
Die Abbildungen 1 bis 3 zeigen den Einfluß der Vigilanz auf den Abstraktionsgrad der gebildeten Kategorien für verschiedene Lernstadien bei einem Schwellwert von $zw=1.6$. In Abb.1 kann man sehen, daß sich im Falle einer langen Resonanzphase ($kk=10$) die Kategorientypen über die Lernzyklen hinweg nicht verändern. Bei niedriger Vigilanz (0.2) werden nur zwei Oberbegriffe ausgebildet (Kategorientyp 1), bei mittlerer Vigilanz (0.5) entstehen 4 Basisbegriffe (Kategorientyp 3), bei hoher Vigilanz (0.8) werden 8 Unterbegriffe (Kategorientyp 5) erzeugt und bei maximaler Vigilanz (1.0) erscheinen Exemplarbegriffe (Kategorientyp 7). Zu keinem Zeitpunkt existieren Kategorien verschiedener Abstraktionsgrade nebeneinander.

Abb. 1: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Vigilanz und der Anzahl der Lernzyklen für die Parameter $zw=1.6$ und $kk=10$



Verringert man die Resonanzdauer pro Item ($kk=3$), dann zeigt sich für mittlere und höhere Vigilanzwerte deutlich eine Veränderung der gebildeten Kategorientypen mit dem Lernzustand (Abb.2). Nach einem Lernzyklus wurden mit einer mittleren Vigilanz von $p=0.5$ noch nicht die vier möglichen Basisbegriffe gebildet. Während beide Oberbegriffe noch vorhanden sind, wird zusätzlich schon ein Begriff abgespalten, der sich im weiteren zu einem Basisbegriff entwickelt. Der zugehörige Oberbegriff entwickelt sich dann zu dem komplementären Basisbegriff. Diese zeitlich begrenzte Übergangsphase, in der kurzzeitig Kategorien verschiedener Abstraktionsgrade nebeneinander bestehen, wurde als Kategorientyp 2 klassifiziert. Nach drei Lernzyklen hat sich das Kategoriensystem hier stabilisiert, es sind jetzt ausschließlich Basisbegriffe und keine Oberbegriffe mehr vorhanden (Kategorientyp 3). Bei hoher Vigilanz (0.8) wirkt sich die verringerte Resonanzdauer noch stärker aus. Es werden nun in einem Lernzyklus zuerst Basisbegriffe (Typ 3) gebildet, die insgesamt 5 Lernzyklen benötigen, um zu Unterbegriffen differenziert zu werden. Nach drei Lernzyklen tritt eine Zwischenphase in Erscheinung, in der einige Unterbegriffe ausdifferenziert werden und somit Basis- und Unterbegriffe nebeneinander existieren (Kategorientyp 4). Jedoch entwickeln sich alle Basisbegriffe zu Unterbegriffen, so daß ab dem 5. Lernzyklus nur noch Unterbegriffe (Typ 5) existieren. Bei maximaler Vigilanz (1.0) werden nach einem Lernzyklus Basisbegriffe (Typ 3) gebildet, in zwei weiteren Lernzyklen entwickeln sich Unterbegriffe und aus diesen Exemplarbegriffe.

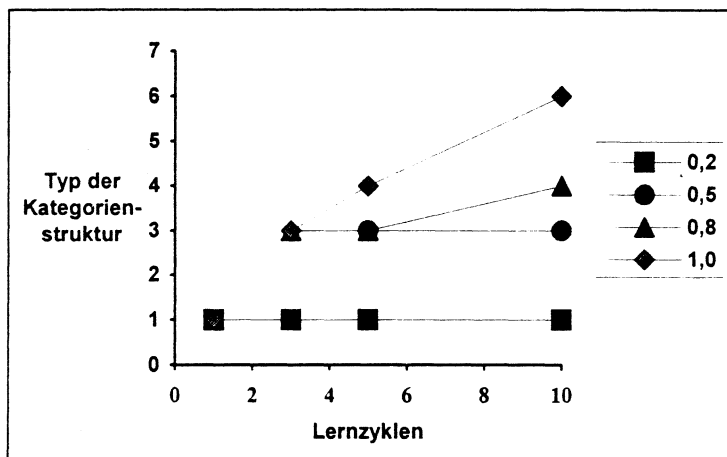
Abb. 2: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Vigilanz und der Anzahl der Lernzyklen für die Parameter $zw=1.6$, $kk=3$



Bei noch kürzerer Resonanzdauer ($kk=1$) werden mit einer Vigilanz von 0.5 selbst im 3. Lernzyklus noch keine Basisbegriffe gebildet (Abb.3). Diese entstehen hier erst im 5. Lernzyklus. Weiterhin werden mit einer hohen Vigilanz von $p=0.8$ sogar nach 5 Lernzyklen nur Basisbegriffe gebildet. Nach 10 Lernzyklen befindet sich das Netz noch in einem Übergangsstadium

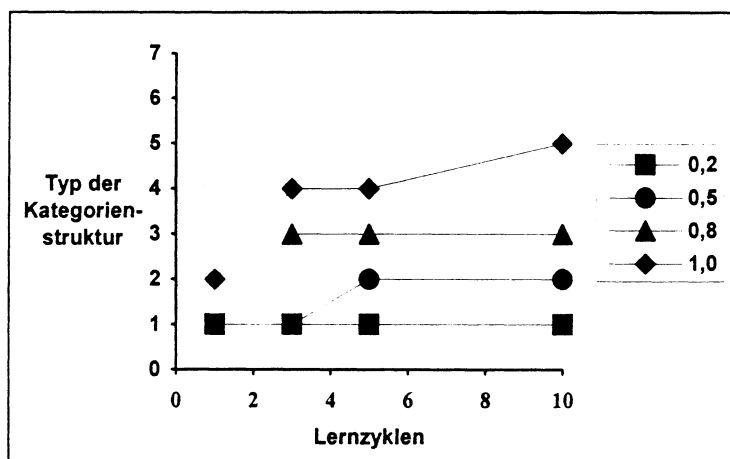
von Basisbegriffen zu Unterbegriffen. Letztere entstehen erst im 13. Lernzyklus. Selbst mit maximaler Vigilanz kann das Netz in einem Lernzyklus nur Oberbegriffe und in zwei weiteren Lernzyklen Basisbegriffe ausbilden. Nach 10 Lernzyklen mit $p=1.0$ erreicht das Netz einen Zustand, in dem Unterbegriffe und Exemplarbegriffe nebeneinander existieren.

Abb. 3: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Vigilanz und der Anzahl der Lernzyklen für die Parameter $zw=1.6$, $kk=1$



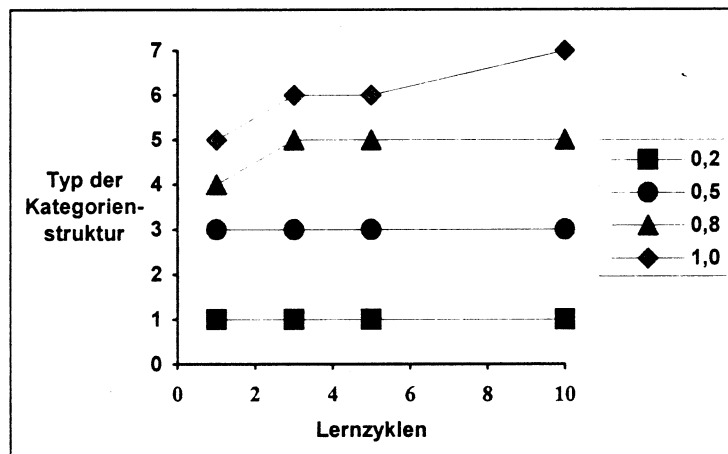
Variiert man den Schwellwert der Merkmalschicht, so zeigt sich ein noch differenzierteres Bild. Ein um 0.2 verringerter Schwellwert von 1.4 führt bei relativ geringer Resonanzdauer ($kk=3$) auch mit maximaler Vigilanz nach 10 Lernzyklen höchstens zur Bildung von Unterbegriffen. Bei der noch relativ hohen Vigilanz von 0.8 werden maximal Basisbegriffe gebildet, bei mittlerer Vigilanz (0.5) werden diese nicht einmal vollständig erreicht (Abb.4, vgl. mit Abb.2).

Abb. 4: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Vigilanz und der Anzahl der Lernzyklen für die Parameter $zw=1.4$, $kk=3$



Ein um 0.2 erhöhter Schwellwert (1.8) dagegen forciert die Ausdifferenzierung spezifischerer Kategorien. Er führt bei mittlerer Vigilanz (0.5) schon nach einen Lernzyklus zur Bildung von Basisbegriffen, bei hoher Vigilanz (0.8) werden in nur 3 Lernzyklen bereits Unterbegriffe gebildet. Mit maximaler Vigilanz entstehen nach 10 Lernzyklen Exemplarbegriffe (Abb.5. vgl. mit Abb.2 und 4).

Abb. 5: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Vigilanz und der Anzahl der Lernzyklen für die Parameter $zw=1.8$, $kk=3$

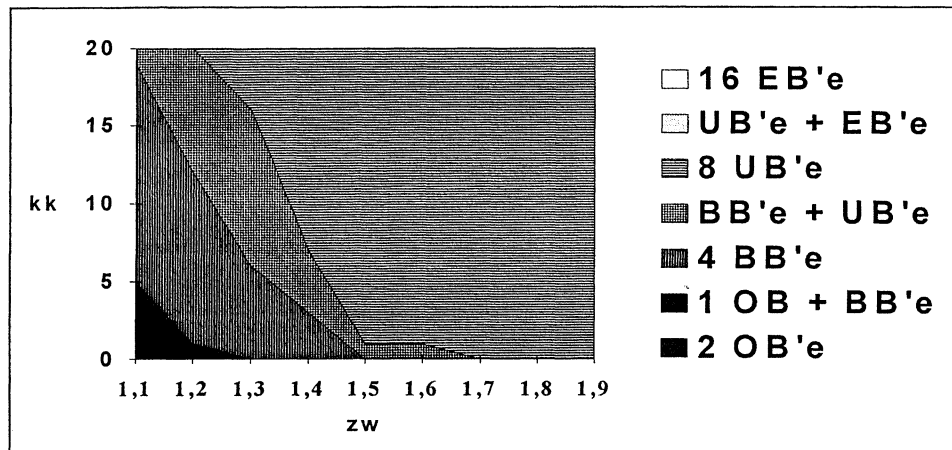


Es zeigt sich, daß bei einer sehr niedrigen Vigilanz (0.2) unter allen Parameterbedingungen nur Oberbegriffe gebildet werden, d.h. es kann überhaupt keine feinere begriffliche Differenzierung vom Netz vorgenommen werden. Mittlere und hohe Vigilanzwerte ermöglichen dem Netz zwar ein feineres Kategorisieren, jedoch ist die Güte der Differenzierung hier zusätzlich vom Schwellwert der Merkmalschicht, der Resonanzdauer und der Häufigkeit bisherigen Lernens abhängig.

In Abb. 6 ist für den gesamten Parameterraum der beiden Dimensionen Schwellwert (zw) und Resonanzdauer (kk) im Bereich von $kk=1$ bis $kk=20$ und $zw=1.1$ bis $zw=1.9$ dargestellt, welche Kategorientypen in 10 Lernzyklen mit einer hohen Vigilanz von 0.8 gebildet werden. Es zeigt sich, daß für geringe Schwellwerte (zw) verbunden mit einer geringen Resonanzdauer (kk) nur Oberbegriffe entstehen. Mit zunehmender Erhöhung des Schwellwertes oder mit zunehmender Resonanzdauer werden spezifischere Begriffe, d.h. Basisbegriffe und im weiteren Unterbegriffe, ausgebildet. Exemplarbegriffe entstehen hier nicht.

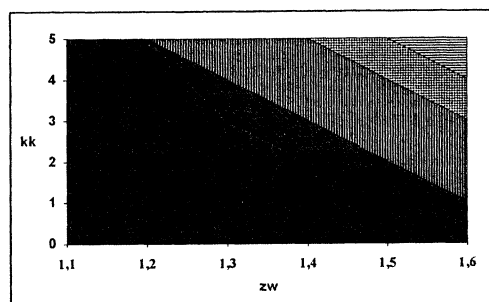
Abb. 6: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Resonanzdauer (kk) und dem Schwellwert der Merkmalschicht (zw) bei einer Vigilanz von 0.8 nach 10 Lernzyklen

OB'e: Oberbegriffe, BB'e: Basisbegriffe,
 UB'e: Unterbegriffe, EB'e: Exemplarbegriffe

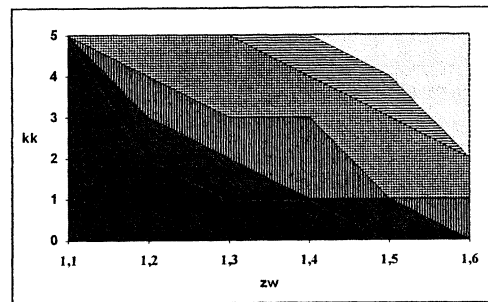


Die Abb.en 7a-d zeigen, welche Kategorientypen sich bei maximaler Vigilanz (1.0) im linken unteren Bereich des Parameterraumes mit Schwellwerten von 1.1 bis 1.6 und einer Resonanzdauer im Bereich von $kk=1$ bis $kk=5$ herausbilden. Nach einem Lernzyklus (Abb.7a) entstehen unter den meisten Bedingungen Oberbegriffe, d.h. sofern nicht gleichzeitig Schwellwert und Resonanzdauer relativ groß sind. Spezifischere Begriffe benötigen für ihre Konstitution höhere Schwellwerte und längere Resonanz. Selbst eine maximale Vigilanz muß also nicht notwendigerweise und unmittelbar zu einer feineren begrifflichen Differenzierung führen. Die durch die hohe Vigilanz bereitgestellte Differenzierungsfähigkeit entwickelt sich außerdem mit zunehmenden Lernzyklen. Während nach einem Lernzyklus höchstens Unterbegriffe gebildet werden können, erscheinen nach drei Lernzyklen (Abb.7b) erste Exemplarbegriffe. Jedoch dominieren auch hier in großen Teilen allgemeinere Begriffe. Mit zunehmender Zyklenzahl werden an allen Orten des Parameterraumes allgemeinere Begriffe durch spezifischere ersetzt (Abb.7c und d). Nach 10 Lernzyklen nehmen Unterbegriffe und Exemplarbegriffe den größten Raum ein. Nur bei sehr kleinen Schwellwerten, verbunden mit relativ kurzer Resonanz bleibt eine grobe Strukturierung erhalten. Unter diesen Bedingungen werden trotz maximaler Vigilanz weiterhin nur Oberbegriffe oder Basisbegriffe gebildet. Nach 10 Lernzyklen findet man also in diesem Ausschnitt des Parameterraumes Begriffe aller 4 möglicher Abstraktionsgrade.

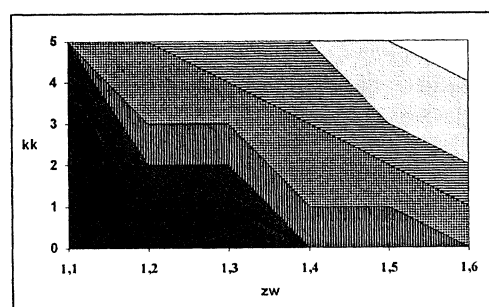
Abb. 7a-d: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Resonanzdauer (kk) und dem Schwellwert der Merkmalschicht (zw) bei einer Vigilanz von 1.0 Achsen und Legende wie in Abb. 6



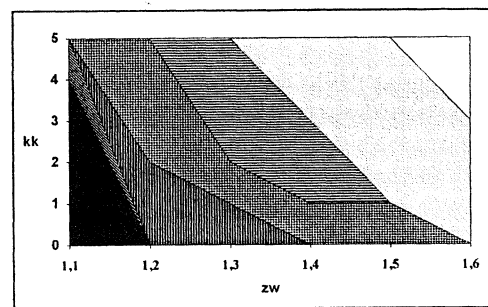
a) 1 Lernzyklus



b) 3 Lernzyklen



c) 5 Lernzyklen



d) 10 Lernzyklen

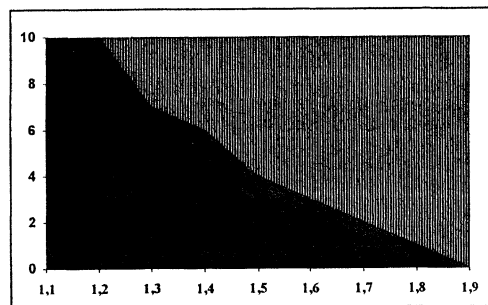
Veränderungen in der Differenzierungsfähigkeit des System über die Lernzyklen hinweg können auch für weitere Vigilanzstufen nachgewiesen werden. In Abb. 8 ist für den Parameterraum im Bereich von $kk=1$ bis $kk=10$ und $zw=1.1$ bis $zw=1.9$ für zwei Vigilanzstufen und verschiedene Lernstadien dargestellt, welchen Abstraktionsgrad die gebildeten Kategorien besitzen.

Bei einer Vigilanz von 0.5 (Abb. 8a, c und e) können mit niedrigem Schwellwert und kurzer Resonanz Oberbegriffe oder mit höherem Schwellwert und verlängerter Resonanz Basisbegriffe gebildet werden, wobei sich mit zunehmender Lerndauer Basisbegriffe anstelle der Oberbegriffe entwickeln. Insbesondere im diagonalen Bereich des Parameterraumes von kleinem Schwellwert mit langer Resonanz hin zu großem Schwellwert mit kurzer Resonanz treten solche Verschiebungen auf.

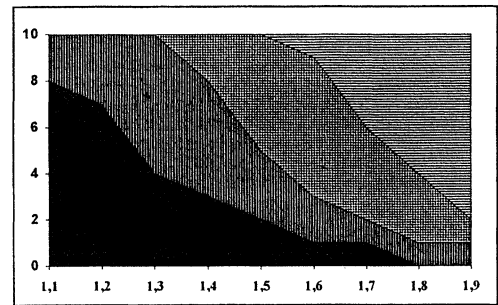
Bei einer höheren Vigilanz von 0.8 (Abb. 8b, d und f) kann man bis auf Exemplarbegriffe alle Abstraktionsgrade im Parameterraum finden. Ein niedriger Schwellwert und kurze Resonanz führen hier wieder zur Bildung von Oberbegriffen, ein höherer Schwellwert und lange Resonanz dagegen zu Unterbegriffen. Dazwischen treten als relativ breites Band Basisbegriffe und etwas schwächer Übergangsformen in Erscheinung. Auch hier im Falle hoher Vigilanz ersetzen Basisbegriffe bei weiterem Lernen zunehmend die Oberbegriffe. Weiterhin werden in

großen Bereichen des Parameterraumes Basisbegriffe von den Unterbegriffen verdrängt. Nur in den Eckbereichen von kleinem Schwellwert mit kurzer Resonanz und großem Schwellwert mit langer Resonanz ist eine Stabilität im Abstraktionsgrad der gebildeten Kategorien zu beobachten.

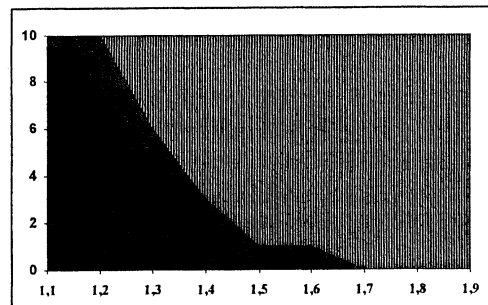
Abb. 8a-h: Abstraktionsgrad der gebildeten Kategorien in Abhängigkeit von der Resonanzdauer (kk) und dem Schwellwert der Merkmalschicht (zw)
 Achsen und Legende wie in Abb. 6



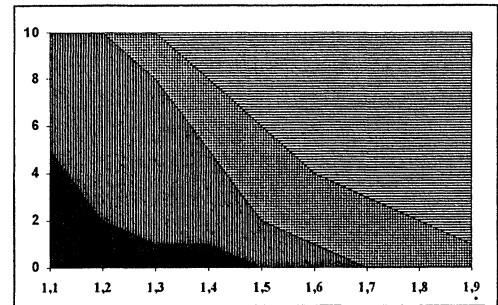
a) $p = 0.5$ 1 Lernzyklus



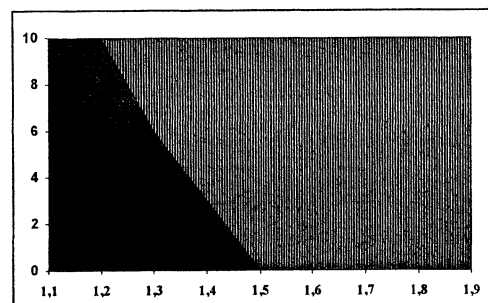
b) $p = 0.8$ 1 Lernzyklus



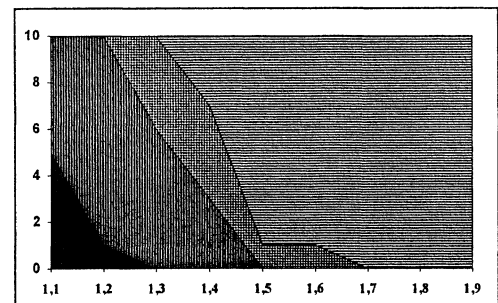
c) $p = 0.5$ 3 Lernzyklen



d) $p = 0.8$ 3 Lernzyklen



e) $p = 0.5$ 10 Lernzyklen



f) $p = 0.8$ 10 Lernzyklen

Insgesamt wird aus den Simulationen deutlich, daß in Abhängigkeit von internen Parametern des Netzes unterschiedlich abstrakte Begriffe gebildet werden. Niedrige Vigilanz führt zur Bildung allgemeiner Begriffe, nur eine höhere Vigilanz ermöglicht die Entwicklung spezifi-

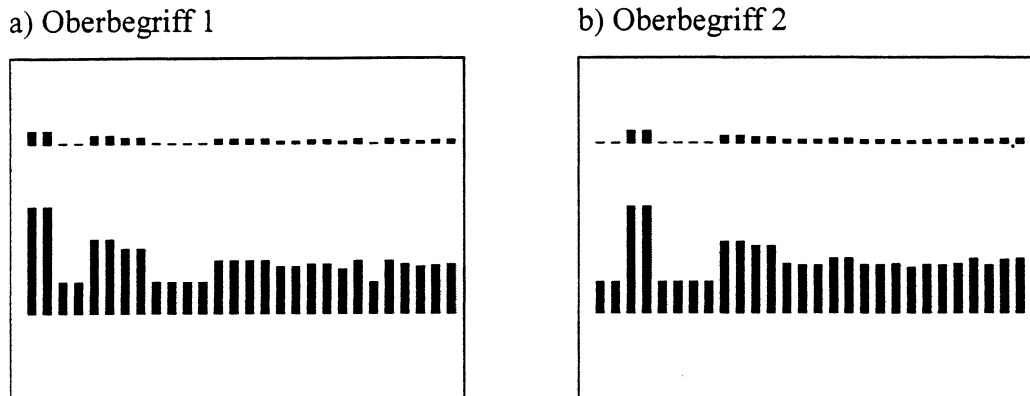
scherer Begriffe. Neben der Vigilanz haben auch der Schwellwert der Merkmals-schicht und die Resonanzdauer einen Einfluß auf den Abstraktionsgrad der zu einem Zeitpunkt gebildeten Begriffe. Höhere Schwellwerte und längere Resonanzphasen begünstigen die Ausbildung spezifischer Begriffe, sehr niedrige Schwellwerte und kurze Resonanzphasen verhindern diese Entwicklung. Insbesondere bei mittleren Werten entwickeln sich spezifische Begriffe erst langsam. So durchläuft das Begriffssystem Phasen, in denen nur Oberbegriffe oder nur Basisbegriffe existieren, um nach weiteren Lernzyklen Unterbegriffe auszubilden. Dabei muß hervorgehoben werden, daß unter konstanten Parameterbedingungen Kategorien unterschiedlicher Abstraktionsgrade ausschließlich in einer zeitlich begrenzten Differenzierungsphase nebeneinander bestehen. Längerfristig entstehen im Netz bei konstanten Parametern nur Kategorien eines Abstraktionsniveaus. Das bedeutet, daß z.B. nach dem Herausbilden von Basisbegriffen keine Oberbegriffe mehr im Netz existieren. Insbesondere kann für ein Item zu einem festen Zeitpunkt nur eine einzige Zuordnung getroffen werden, weil jeweils nur eine für dieses Item relevante Kategorie im Netz existiert. Über verschiedene Lernphasen hinweg kann sich die Zuordnung eines Items jedoch ändern, dadurch daß sich spezifischere Kategorien aus allgemeinen entwickeln und sich weitere spezifische etablieren. So wird ein Item z.B. anfänglich einem Oberbegriff zugeordnet werden, in späteren Lernzyklen kann es dann Basisbegriffen und schließlich auch Unterbegriffen angehören.

Erklärung der Simulationensergebnisse und Interpretation von empirischen Befunden: Der Lernzustand allgemeiner Begriffe

In den Simulationen hat sich gezeigt, daß unter speziellen Parameterbedingungen oder zu bestimmten Zeitpunkten die Bildung spezifischerer Kategorien auch mit maximaler Vigilanz nicht möglich ist. Dies ist auf den Lernzustand vorhandener allgemeinerer Begriffe zurückzuführen, was im folgenden näher erläutert wird.

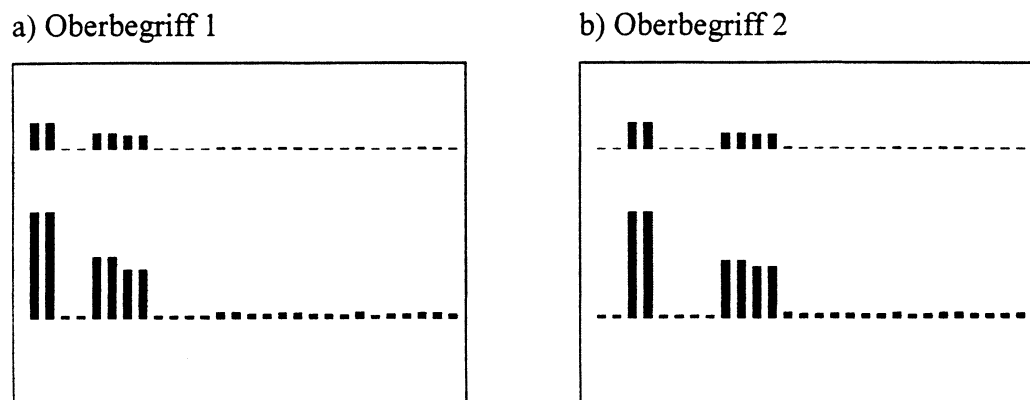
In Abb.8 ist der unter den Parameterbedingungen $kk=3$ und $zw=1.4$ erhaltene Lernzustand der beiden Oberbegriffe nach einem Lernzyklus dargestellt. Die in den Gewichten kodierte Merkmalsstruktur ist relativ unscharf. Es sind noch viele Merkmale repräsentiert, so daß eine diffuse Erwartung für die Merkmale der Exemplare gebildet wird. Diese stimmt auch mit vielen spezifischen Merkmalen eines einzelnen Items überein. Dadurch werden die Items hinreichend gut auf den allgemeinen Begriff abgebildet.

Abb. 8 a,b: Merkmalsstruktur der mit $kk=3$ und $zw=1.4$ gebildeten allgemeinen Begriffe nach einem Lernzyklus
 bottom-up Gewichte: oben, top-down Gewichte: unten



Demgegenüber sind in Abb.9 die gleichen Begriffe nach drei Lernzyklen abgebildet. Ihre interne Merkmalsstruktur hat sich deutlich verändert. Es werden jetzt die jeweiligen Oberbegriffsmerkmale und die Merkmale der zu den Oberbegriffen gehörenden Basisbegriffe kodiert. Weitere Merkmale sind nicht mehr vorhanden. Die Begriffe sind somit bereits schärfer geworden. Es werden gezielt die allgemeinen Merkmale für die Exemplare des Begriffs erwartet. Da die spezifischen Merkmale der Items keine Berücksichtigung mehr finden, werden bei einer hohen Vigilanz (0.8) bereits spezifischere Begriffe gebildet.

Abb. 9 a,b: Merkmalsstruktur der mit $kk=3$ und $zw=1.4$ gebildeten allgemeinen Begriffe nach drei Lernzyklen
 bottom-up Gewichte: oben, top-down Gewichte: unten



Bei hinreichender Schärfe der allgemeinen Kategorie, können Exemplare auch mit mittlerer Vigilanz eine neue Kategorie konstituieren. Nach Absinken weiterer Merkmalsbewichtungen (Abb.10) werden bei dem betrachteten Oberbegriff nur noch die Oberbegriffsmerkmale und die Merkmale eines Basisbegriffes (BB 4) stark genug kodiert und dementsprechend erwartet. Für Items des anderen Basisbegriffes (BB 3) werden daher ausschließlich die beiden Oberbegriffsmerkmale bei der Zuordnung zur bestehenden allgemeinen Kategorie erfaßt, was von einer niedrigen Vigilanz (0.2) toleriert wird, von einer mittleren jedoch bereits nicht mehr. Somit wird bei diesem Lernzustand des allgemeinen Begriffs im vierten Lernzyklus erstmals ein neuer Begriff bei einer Vigilanz von 0.5 gebildet.

Abb. 10: Merkmalsstruktur eines mit $kk=3$ und $zw=1.4$ gebildeten allgemeinen Begriffs im vierten Lernzyklus, Zeitpunkt des Konstituierens eines spezifischen Begriffs mit einer Vigilanz von 0.5
bottom-up Gewichte: oben, top-down Gewichte: unten

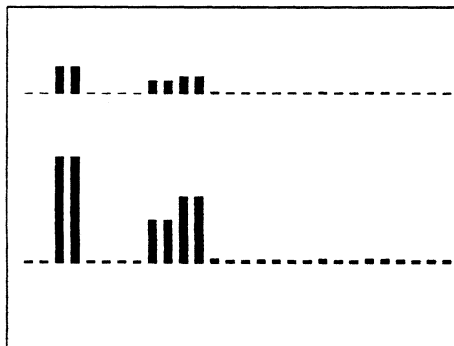
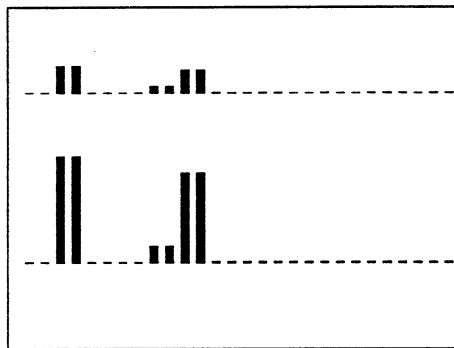


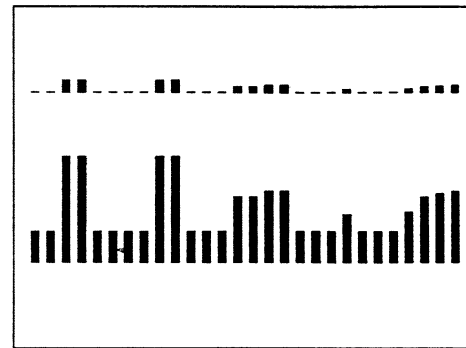
Abb.11 zeigt, wie sich der Oberbegriff 2 nach dem Konstituieren des Basisbegriffs 3 im 5. Lernzyklus zu dem komplementären Basisbegriff 4 entwickelt. Während die Merkmale des Basisbegriffs 3 weiter an Gewicht verlieren, werden die Merkmale des Basisbegriffs 4 zunehmend gestärkt. Der sich aus dem ehemals allgemeinen Begriff entwickelnde Basisbegriff 4 ist bereits sehr scharf, während der neu erworbene Basisbegriff 3 noch sehr unscharf ausgebildet ist. Die weitere begriffliche Differenzierung, d.h. die Herausbildung von Unterbegriffen wird demzufolge eher am Basisbegriff 4 erfolgen als am Basisbegriff 3.

Abb. 11 a,b: Merkmalsstruktur der mit $kk=3$ und $zw=1.4$ gebildeten Basisbegriffe nach fünf Lernzyklen
 bottom-up Gewichte: oben, top-down Gewichte: unten

a) Basisbegriff 4,
 früherer Oberbegriff 2



b) Basisbegriff 3,
 neu gebildet



Der Einfluß des Lernzustandes einer allgemeinen Kategorie auf die Fähigkeit des gesamten Systems zur Bildung entsprechender spezifischerer Kategorien erklärt unseren experimentellen Befund (Goede und Friedrich, 1995), nach dem Zweijährige nicht fähig waren, eine Teilkategorie für die Blumen zu bilden, diese Aufgabe jedoch für Autos bewältigten. Die allgemeine Kategorie der Autos ist aufgrund eines häufigeren Lernens besser gefestigt als die Kategorie der Blumen. Insbesondere ist sie bereits schärfer repräsentiert, d.h. es werden nicht mehr diffus viele, sondern nur noch die allgemeinen Merkmale bisheriger Exemplare kodiert. Damit werden von der begrifflichen Kategorie nur wenige, immer wieder auftretende Merkmale fokussiert. Bei geeigneter Vigilanz werden daher für die Autos bereits früh spezifischere Begriffe gebildet. Im Gegensatz dazu ist die begriffliche Kategorie der Blumen noch sehr unscharf. Es werden über eine unspezifische Erwartung Merkmale vieler unterschiedlicher Exemplare erfaßt, so daß diese noch hinreichend gut auf die allgemeine Kategorie abgebildet werden. Damit ist für die Blumen der Erwerb einer spezifischen Kategorie selbst mit maximaler Vigilanz nicht möglich, während eine solche Differenzierung bei Autos bereits erfolgen kann.

Weiterhin wird auch der empirische Befund von Goede (1989) verständlich, wonach die Fähigkeit zur Teilkategorienbildung zunimmt, sobald andere spezifische Begriffe für einen allgemeinen Begriff verfügbar sind. Das Etablieren dieser anderen Begriffe setzt nämlich genau den Lernzustand voraus, der auch zur Bildung spezifischer Begriffe unter experimentellen Bedingungen notwendig ist.

Simulationen mit variabler Vigilanz in den Lernzyklen

Bisher wurden nur Begriffe jeweils eines Abstraktionsniveaus vom Netz gebildet, wobei mit zunehmenden Lernzyklen allgemeine Begriffe durch spezifischere ersetzt wurden. Diese Entwicklung von allgemeinen zu spezifischen Begriffen wird auch bei Kindern beobachtet. Kinder haben weiterhin häufig Schwierigkeiten, einzelne Objekte gleichzeitig oder nacheinander auf verschiedenen Abstraktionsniveaus zu betrachten, was sowohl durch ein Nichtvorhandensein verschieden allgemeiner Gedächtnisstrukturen als auch durch die Unfähigkeit zur gleichzeitigen Nutzung bestehender Strukturen verursacht sein kann. Das Vorhandensein nur einer begrifflichen Abstraktion für ein Objekt ist im Netz in der Konstanz der Parameter begründet. Daß im neuronalen Netz ART1 über eine Variation der Vigilanz Begriffe mehrerer Abstraktionsgrade gebildet werden können und welche Probleme es bei der Nutzung dieser Begriffe gibt, wird in diesem Abschnitt gezeigt.

Tab.3 faßt für den Schwellwert 1.6 und eine Resonanzdauer von $kk=3$ noch einmal zusammen, welche Kategorientypen mit fester Vigilanz in maximal 20 Lernzyklen gebildet werden: Oberbegriffe werden in beliebig vielen Lernzyklen mit Vigilanz 0.2 oder in maximal einem Lernzyklus mit Vigilanz 0.5 gebildet. Basisbegriffe entstehen nach mindestens zwei Zyklen mit mittlerer Vigilanz (0.5), höchstens zwei Zyklen mit hoher Vigilanz (0.8) oder nach einem Zyklus mit maximaler Vigilanz (1.0). Unterbegriffe entstehen nach mindestens fünf Zyklen mit hoher Vigilanz oder drei Zyklen mit maximaler Vigilanz. Exemplarbegriffe benötigen mindestens 11 Zyklen mit maximaler Vigilanz zu ihrer Herausbildung.

Tab. 3: Anzahl der Lernzyklen mit fester Vigilanz, die für die Bildung und Aktivierung von Begriffen eines bestimmten Abstraktionsgrades bei einem Schwellwert von $zw=1.6$ und einer Resonanzdauer von $kk=3$ notwendig sind

Vigilanz	Anzahl der Lernzyklen			
p=0.2	1 - 20			
p=0.5	1	2 - 20		
p=0.8		1 - 2	5 - 20	
p=1.0		1	3	11 - 20
ausgebildete und aktivierte Kategorien	2 OB'e	4 BB'e	8 UB'e	16 EB'e

Tab.4 zeigt dagegen eine Situation, die bei gleichem F1-Schwellwert und gleicher Resonanzdauer eintritt, wenn die Vigilanz über die Lernzyklen hinweg gesteigert wird. Vier Lernzyklen mit einer niedrigen Vigilanz von 0.2 und zwei weitere Lernzyklen mit einer mittleren Vigilanz von 0.5 führen dazu, daß vom Netz sowohl 2 Oberbegriffe als auch 4 Basisbegriffe ausgebildet werden (Spalte 1). Die bottom-up Gewichte der einzelnen Merkmale sind bei den Basisbegriffen, noch nicht hoch genug, um starke Signale zu diesen Kategorien weiterzuleiten. Da im Ergebnis der Konkurrenz zwischen geringen Signalen alle bis auf das maximale ausgelöscht werden (Grossberg und Levine, 1975), kann ein Exemplar nur jeweils einem Begriff zugeordnet werden. Eine spontane Zuordnung entsprechend dem größten Signal erfolgt hier zu den Oberbegriffen. Bei einer niedrigen Vigilanz von 0.2 wird diese Kategorisierung der Exemplare auch akzeptiert, bei einer höheren Vigilanz (0.5) jedoch nicht, so daß nun die Basisbegriffe genutzt werden. Der Zugriff auf die Basisbegriffe erfolgt bei diesem Lernzustand ausschließlich über eine erhöhte Vigilanz, auf die Oberbegriffe dagegen nur bei niedriger Vigilanz.

Tab. 4: Bildung und Aktivierung von unterschiedlich abstrakten Begriffen
 aufgrund unterschiedlicher Vigilanz in den Lernzyklen
 bei einem Schwellwert von $zw=1.6$ und einer Resonanzdauer von $kk=3$

Vigilanz	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8
Anzahl der Lernzyklen	4	2	0	2	4	0	6	6	0
ausgebildete Kategorien	2 OB'e, 4 BB'e			2 OB'e, 4 BB'e			2 OB'e, 4 BB'e 3 UB'e		
Zuordnung der Items bei p=0.2	OB'e			BB'e			OB'e und BB'e gemeinsam		
Zuordnung der Items bei p=0.5	BB'e			BB'e			UB'e, BB'e		

Werden jedoch nach zwei Lernzyklen mit niedriger Vigilanz vier Lernzyklen mit mittlerer Vigilanz durchgeführt, so sind die Basisbegriffe besser gelernt als die Oberbegriffe, werden nun von ihren zugehörigen Items spontan aktiviert und vom Netz sowohl bei niedriger als

Tab.4 zeigt dagegen eine Situation, die bei gleichem F1-Schwellwert und gleicher Resonanzdauer eintritt, wenn die Vigilanz über die Lernzyklen hinweg gesteigert wird. Vier Lernzyklen mit einer niedrigen Vigilanz von 0.2 und zwei weitere Lernzyklen mit einer mittleren Vigilanz von 0.5 führen dazu, daß vom Netz sowohl 2 Oberbegriffe als auch 4 Basisbegriffe ausgebildet werden (Spalte 1). Die bottom-up Gewichte der einzelnen Merkmale sind bei den Basisbegriffen, noch nicht hoch genug, um starke Signale zu diesen Kategorien weiterzuleiten. Da im Ergebnis der Konkurrenz zwischen geringen Signalen alle bis auf das maximale ausgelöscht werden (Grossberg und Levine, 1975), kann ein Exemplar nur jeweils einem Begriff zugeordnet werden. Eine spontane Zuordnung entsprechend dem größten Signal erfolgt hier zu den Oberbegriffen. Bei einer niedrigen Vigilanz von 0.2 wird diese Kategorisierung der Exemplare auch akzeptiert, bei einer höheren Vigilanz (0.5) jedoch nicht, so daß nun die Basisbegriffe genutzt werden. Der Zugriff auf die Basisbegriffe erfolgt bei diesem Lernzustand ausschließlich über eine erhöhte Vigilanz, auf die Oberbegriffe dagegen nur bei niedriger Vigilanz.

Tab. 4: Bildung und Aktivierung von unterschiedlich abstrakten Begriffen
 aufgrund unterschiedlicher Vigilanz in den Lernzyklen
 bei einem Schwellwert von $zw=1.6$ und einer Resonanzdauer von $kk=3$

Vigilanz	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8
Anzahl der Lernzyklen	4	2	0	2	4	0	6	6	0
ausgebildete Kategorien	2 OB'e, 4 BB'e			2 OB'e, 4 BB'e			2 OB'e, 4 BB'e 3 UB'e		
Zuordnung der Items bei p=0.2	OB'e			BB'e			OB'e und BB'e gemeinsam		
Zuordnung der Items bei p=0.5	BB'e			BB'e			UB'e, BB'e		

Werden jedoch nach zwei Lernzyklen mit niedriger Vigilanz vier Lernzyklen mit mittlerer Vigilanz durchgeführt, so sind die Basisbegriffe besser gelernt als die Oberbegriffe, werden nun von ihren zugehörigen Items spontan aktiviert und vom Netz sowohl bei niedriger als

auch bei mittlerer Vigilanz akzeptiert (Spalte 2). Nicht mehr genutzt werden in diesem Fall die allgemeineren Begriffe. Ihr Zugriff ist möglicherweise dauerhaft verhindert.

Lernt das Netz dagegen in 6 Lernzyklen mit niedriger Vigilanz die Oberbegriffe hinreichend gut und werden in weiteren 6 Lernzyklen die Basisbegriffe erworben und gefestigt, wobei zusätzlich bereits Unterbegriffe entstehen, so kann bei niedriger Vigilanz jedes Item gleichzeitig seinen Basisbegriff und den dazugehörigen Oberbegriff aktivieren (Spalte 3). Bei mittlerer Vigilanz wird diese gemeinsame Zuordnung nicht mehr akzeptiert. Die bereits gebildeten Unterbegriffe werden jetzt genutzt und entwickeln sich z.T. zu redundanten Basisbegriffen, die jedoch nicht gemeinsam mit einem Oberbegriff aktiviert werden.

Wenn man dagegen nur Lernzyklen mit niedriger (0.2) und mit hoher (0.8) Vigilanz durchführt, so fehlen die bei mittlerer Vigilanz entstehenden Basisbegriffe völlig (Tab.5). In diesen Fällen werden Unterbegriffe zusätzlich zu den Oberbegriffen ausgebildet. Nach jeweils 4 Lernzyklen mit niedriger und hoher Vigilanz, werden spontan wieder die Oberbegriffe aktiviert (Spalte 1). Während diese Zuordnung bei einer Vigilanz von 0.2 auch bestehen bleibt, werden die Exemplare bei einer erhöhten Vigilanz den Unterbegriffen zugeordnet. Interessant ist hier, daß schon eine mittlere Vigilanz von 0.5 den Zugriff auf die Unterbegriffe ermöglicht, wenn diese bereits über eine höhere Vigilanz konstituiert wurden.

Nach zwei Lernzyklen mit niedriger und acht mit hoher Vigilanz erfolgt der Zugriff unter allen betrachteten Vigilanzbedingungen nur auf die besser gelernten Unterbegriffe (Spalte 2). Werden dagegen anfänglich sechs Zyklen mit niedriger Vigilanz und daran anschließend acht Zyklen mit hoher Vigilanz durchgeführt, so sind spontan die Oberbegriffe gemeinsam mit den Unterbegriffen aktiv (Spalte 3). Die gemeinsame Zuordnung zu jeweils einem Ober- und einem Unterbegriff erfolgt bei niedriger Vigilanz. Dadurch, daß die Aktivierungen der Unterbegriffe im Ergebnis der Konkurrenz mit den Oberbegriffen geschwächt wurden, ist die gemeinsame Erwartung von Ober- und Unterbegriff schlechter als die vom Unterbegriff allein ausgelöste. Aus diesem Grund wird die gemeinsame Zuordnung bei erhöhter Vigilanz verhindert. Es werden nun Exemplarbegriffe gebildet, die sich ihrerseits in Abhängigkeit von der Vigilanz zu neuen Basis- oder Unterbegriffen entwickeln können.

Tab. 5: Bildung und Aktivierung von unterschiedlich abstrakten Begriffen
 aufgrund unterschiedlicher Vigilanz in den Lernzyklen
 bei einem Schwellwert von $zw=1.6$ und einer Resonanzdauer von $kk=3$

Vigilanz	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8
Anzahl der Lernzyklen	4	0	4	2	0	8	6	0	8
ausgebildete Kategorien	2 OB'e, 8 UB'e		2 OB'e, 8 UB'e		2 OB'e, 8 UB'e (und 4 weitere)				
Zuordnung der Items bei p=0.2	OB'e		UB'e		OB'e und UB'e gemeinsam				
Zuordnung der Items bei p=0.5	UB'e		UB'e		EB'e, UB'e, BB'e				
Zuordnung der Items bei p=0.8	UB'e		UB'e		EB'e, UB'e				

Wie Ober- und Basisbegriffe sowie Ober- und Unterbegriffe können bei geeigneter Vigilanz auch Basis- und Unterbegriffe vom Netz ausgebildet werden und nebeneinander existieren. Analog zu den bisher betrachteten Fällen ist der Zugriff wieder in Abhängigkeit vom jeweiligen Lernzustand und der Vigilanz entweder nur auf Basisbegriffe, nur auf Unterbegriffe oder auf beide gleichzeitig möglich.

Es können jedoch Kategorien nicht nur auf zwei sondern auch auf mehreren Abtraktionsgraden gebildet werden und nebeneinander bestehen. Werden im Anschluß an die Lernzyklen mit niedriger und mittlerer Vigilanz zusätzlich Lernzyklen mit hoher Vigilanz durchgeführt, so entstehen im Netz neben Oberbegriffen und Basisbegriffen auch noch Unterbegriffe (Tab.6). Nach vier Lernzyklen mit Vigilanz 0.2, drei Lernzyklen mit Vigilanz 0.5 und weiteren 4 Lernzyklen mit Vigilanz 0.8 werden z.B. die beiden Oberbegriffe, die vier möglichen Basisbegriffe und alle acht Unterbegriffe gebildet (Spalte 1). Nach diesen Lernzyklen werden spontan von den Items nur Oberbegriffe aktiviert. Ein Zugriff auf Basis- und Unterbegriffe kann jedoch über entsprechend erhöhte Vigilanz erfolgen, was zu deren weiterer Ausdifferenzierung führt. Nach drei Lernzyklen mit Vigilanz 0.2, vier Lernzyklen mit Vigilanz 0.5 und weiteren 4 Lernzyklen mit Vigilanz 0.8 werden alle Items spontan ihren Basisbegriffen zugeordnet, obwohl

wiederum auch alle anderen Kategorientypen vom Netz ausgebildet wurden. Im Falle der Aufeinanderfolge von jeweils sechs Lernzyklen mit Vigilanz 0.2 und 0.5 sowie weiteren 3 Lernzyklen mit Vigilanz 0.8 wurden auch alle Kategorientypen erworben, der spontane Zugriff erfolgt nun auf Ober- und Basisbegriffe gemeinsam. Die Zugehörigkeit zu Unterbegriffen kann über eine erhöhte Vigilanz verifiziert werden. Der Fall, daß Items sowohl zu Ober- als auch zu Basis- und zu Unterbegriff zugeordnet werden, tritt ein, wenn zusätzlich die Zahl der Lernzyklen mit hoher Vigilanz bis auf 10 erhöht wird, so daß alle Kategorientypen gut gelernt wurden.

Tab. 6: Bildung und Aktivierung von unterschiedlich abstrakten Begriffen aufgrund von Vigilanzerhöhungen zwischen den Lernzyklen bei einem Schwellwert von $zw=1.6$ und einer Resonanzdauer von $kk=3$

Vigilanz	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8	p=0.2	p=0.5	p=0.8
Anzahl der Lernzyklen	4	3	4	3	4	4	6	6	3
ausgebildete Kategorien	2 OB'e, 4 BB'e 8 UB'e			2 OB'e, 4 BB'e, 8 UB'e			2 OB'e, 4 BB'e 8 UB'e		
Zuordnung der Items bei p=0.2	OB'e			BB'e			OB'e und BB'e gemeinsam		

Es zeigt sich also insgesamt, daß der Erwerb verschiedener Strukturierungen einer Datenmenge, insbesondere die Ausbildung unterschiedlich abstrakter Kategorien im neuronalen Netz ART1 durch eine Variation der Vigilanz erreicht werden kann. Nach dem Konstituieren mehrerer begrifflicher Kategorien, die sich in einer hierarchischen Relation befinden, ist ihre weitere Nutzung entscheidend von ihrem Lernzustand und von der aktuellen Vigilanz abhängig. Bei ungenügend gefestigten Kategorien kann in Abhängigkeit von der Vigilanz nur entweder ein allgemeiner oder ein spezifischer Begriff aktiviert werden. Dieser Zusammenhang konnte bereits die experimentell erzeugbaren Phänomene der Exklusion, Reduktion und Negation erklären, die spezielle Formen zeitweiliger Untergeneralisierungen sind. Demnach kann von bestimmten oder von allen Exemplaren bei zu hoher Vigilanz nur eine spezifischere und nicht mehr die allgemeine Kategorie aktiviert werden, obwohl letztere bei niedriger Vigilanz sicher verfügbar ist. Sind dagegen die begrifflichen Kategorien gut gelernt, was sich in einer starken

Bewichtung zeigt, so wird bei niedriger Vigilanz die gleichzeitige Zuordnung eines Items zu einem allgemeinen und einem spezifischen Begriff möglich. Damit erweist sich neben der Fähigkeit des Netzes zur Bildung spezifischer Begriffe auch seine Fähigkeit zur gleichzeitigen Nutzung unterschiedlich allgemeiner Begriffe als vom Lernzustand vorhandener begrifflicher Strukturen abhängig.

Zusammenfassende Bemerkungen

Im neuronalen Netz ART1 entscheidet die Vigilanz in Verbindung mit der Resonanzdauer und dem Schwellwert der Merkmalschicht über den Abstraktionsgrad der gebildeten Kategorien. Werden die Parameter konstant gehalten, so entstehen nur Kategorien eines einzigen Abstraktionsniveaus. Dabei müssen sich spezifischere Begriffe oft erst mit zunehmendem Lernen aus allgemeineren entwickeln. So finden sich Phasen der begrifflichen Entwicklung im Netz, in denen trotz hoher Vigilanz noch keine spezifischen Kategorien erworben werden können. Dies stimmt mit empirischen Beobachtungen bei Kindern überein. In Folge dieser zeitweiligen Unfähigkeit zur begrifflichen Differenzierung können analog zu empirischen Befunden auch im Netz Übergeneralisierungen auftreten, die in späteren Entwicklungsphasen wieder überwunden werden.

Gegenüber konstanten Parametern ermöglicht eine Variation der Vigilanz die Ausbildung unterschiedlich abstrakter Kategorien. Eine beliebige Nutzung mehrerer Abstraktionsgrade tritt aber erst mit einem bestimmten Lernzustand ein. Bis dahin ist der Zugriff auf die verschieden abstrakten Kategorien von der Vigilanz abhängig. An dieser Stelle wird die Bedeutung der *Vigilanzregulation* für den Aufbau und die Nutzung eines strukturierten Begriffssystems sichtbar. Eine solche Vigilanzregulation ist im neuronalen System ARTMAP (Carpenter und Mitarbeiter, 1991) als Teil einer intermodularen Anpassung entwickelt. In diesem System wird die Vigilanz eines Moduls als abhängige Variable so reguliert, daß die korrekte Zuordnung entsprechend dem jeweils angezieltem Fokus bzw. Abstraktionsgrad erfolgen kann. Welche Phänomene in diesem Zusammenhang auftreten, ist in Friedrich (1997) ausführlich dargestellt.

Literatur

- Carpenter, G.A. & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, 37, 54-115.
- Carpenter, G.A., Grossberg, S. & Reynolds, J. H. (1991). ARTMAP: supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, 4, 565-588.
- Friedrich, M. (1997). Die Entstehung und Überwindung von Übergeneralisierungen im neuronalen Netz ARTMAP. in diesem Band.
- Friedrich, M. (1994). Modellierung und Simulation kategorialer Strukturbildung - Eine Anwendung der Adaptiven Resonanztheorie auf die Begriffsbildung. Dissertation.
- Friedrich, M. & Goede, K. (1997). Eine Anwendung der Adaptiven Resonanztheorie auf den Erwerb und die Repräsentation begrifflichen Wissens. *Zeitschrift für Psychologie*, 205, 1-35.
- Goede, K. (1989). Möglichkeiten und Grenzen jüngerer Kinder bei der Herstellung einer Inklusionsrelation zwischen Begriffen. *Zeitschrift für Psychologie*, 197, 31-48.
- Goede, K. & Friedrich, M. (im Druck). Gefäße und Behälter: Das Phänomen der Exklusion bei Erwachsenen. *Zeitschrift für Psychologie*.
- Goede, K. & Friedrich, M. (1995). Wenn Rosen keine Blumen sind - Phänomene der Begriffsentwicklung. *Zeitschrift für Psychologie*, 203, 197-220.
- Grossberg, S. (Ed.). (1987a). The adaptive brain, 1: Cognition, learning, reinforcement, and rhythm. Amsterdam: North-Holland.
- Grossberg, S. (Ed.). (1987b). The adaptive brain, 2: Vision, speech, language, and motor control. Amsterdam: North Holland.
- Grossberg, S. (Ed.). (1982). Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control. Boston: Reidel Press.
- Grossberg, S. (1980). How does the brain build a cognitive code? *Psychological Review*, 87(1), 1-51.
- Grossberg, S. & Levine, D. (1975). Some developmental and attentional biases in the contrast enhancement and short term memory of recurrent neural networks. *Journal of Theoretical Biology*, 53, 341-380.
- Mandler, J. M.; Bauer, P. J. & Mc Donough, L. (1991). Separating the sheep from the goats: Differentiating global categories. *Cognitive Psychology*, 23, 263 - 298.
- Rosch, E.; Mervis, C.B.; Gray, W.D.; Johnson, D.M. & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.